Università di Roma Tor Vergata
Dipartimento di Matematica

Université Paris-Est Marne-La-Vallée
École Doctorale Mathématiques et STIC. Discipline: Mathématiques

Ph.D. Thesis

Option prices in stochastic volatility models

Giulia Terenzi

Advisors
Prof. Lucia Caramellino
Prof. Damien Lamberton

Academic year 2017/2018

Alla mia famiglia

Abstract

We study option pricing problems in stochastic volatility models. In the first part of this thesis we focus on American options in the Heston model. We first give an analytical characterization of the value function of an American option as the unique solution of the associated (degenerate) parabolic obstacle problem. Our approach is based on variational inequalities in suitable weighted Sobolev spaces and extends recent results of Daskalopoulos and Feehan (2011, 2016) and Feehan and Pop (2015). We also investigate the properties of the American value function. In particular, we prove that, under suitable assumptions on the payoff, the value function is nondecreasing with respect to the volatility variable. Then, we focus on an American put option and we extend some results which are well known in the Black and Scholes world. In particular, we prove the strict convexity of the value function in the continuation region, some properties of the free boundary function, the Early Exercise Price formula and a weak form of the smooth fit principle. This is done mostly by using probabilistic techniques.

In the second part we deal with the numerical computation of European and American option prices in jump-diffusion stochastic volatility models. We first focus on the Bates-Hull-White model, i.e. the Bates model with a stochastic interest rate. We consider a backward hybrid algorithm which uses a Markov chain approximation (in particular, a "multiple jumps" tree) in the direction of the volatility and the interest rate and a (deterministic) finite-difference approach in order to handle the underlying asset price process. Moreover, we provide a simulation scheme to be used for Monte Carlo evaluations. Numerical results show the reliability and the efficiency of the proposed methods.

Finally, we analyse the rate of convergence of the hybrid algorithm applied to general jump-diffusion models. We study first order weak convergence of Markov chains to diffusions under quite general assumptions. Then, we prove the convergence of the algorithm, by studying the stability and the consistency of the hybrid scheme, in a sense that allows us to exploit the probabilistic features of the Markov chain approximation.

Keywords: stochastic volatility; European options; American options; degenerate parabolic problems; optimal stopping; tree methods; finite-difference.

Résumé

L'objet de cette thèse est l'étude de problèmes d'évaluation d'options dans les modèles à volatilité stochastique. La première partie est centrée sur les options américaines dans le modéle de Heston. Nous donnons d'abord une caractérisation analytique de la fonction de valeur d'une option américaine comme l'unique solution du problème d'obstacle parabolique dégénéré associé. Notre approche est basée sur des inéquations variationelles dans des espaces de Sobolev avec poids étendant les résultats récents de Daskalopoulos et Feehan (2011, 2016) et Feehan et Pop (2015). On étudie aussi les propriétés de la fonction de valeur d'une option américaine. En particulier, nous prouvons que, sous des hypothèses convenables sur le payoff, la fonction de valeur est décroissante par rapport à la volatilité. Ensuite nous nous concentrons sur le put américaine et nous étendons quelques résultats qui sont bien connus dans le monde Black-Scholes. En particulier nous prouvons la convexité stricte de la fonction de valeur dans la région de continuation, quelques propriétés de la frontière libre, la formule de Prime d'Exercice Anticipée et une forme faible de la propriété du smooth fit. Les techniques utilisées sont de type probabiliste.

Dans la deuxième partie nous abordons le problème du calcul numérique du prix des options europénnes et américaines dans des modèles à volatilité stochastiques et avec sauts. Nous étudions d'abord le modèle de Bates-Hull-White, c'est-à-dire le modèle de Bates avec un taux d'intérêt stochastique. On considére un algorithme hybride rétrograde qui utilise une approximation par chaîne de Markov (notamment un arbre "avec sauts multiples") dans la direction de la volatilité et du taux d'intérêt et une approche (déterministe) par différence finie pour traiter le processus de prix d'actif. De plus, nous fournissons une procédure de simulation pour des évaluations Monte Carlo. Les résultats numériques montrent la fiabilité et l'efficacité de ces méthodes. Finalement, nous analysons le taux de convergence de l'algorithme hybride appliqué à des modèles généraux de diffusion avec sauts. Nous étudions d'abord la convergence faible au premier ordre de chaînes de Markov vers la diffusion sous des hypothéses assez générales. Ensuite nous prouvons la convergence de l'algorithme: nous étudions la stabilité et la consistance de la méthode hybride par une technique qui exploite les caractéristiques probabilistes de l'approximation par chaîne de Markov.

Mots clés : volatilité stochastique ; options américaines ; options européennes ; problèmes paraboliques dégénérés ; arrêt optimal ; approximation par arbres ; différences finies.

Contents

Introduction

The seminal work by Black and Scholes ([21], 1973) was the starting point of equity dynamics modelling and it is still widely used as a useful approximation. It owns its great success to its high intuition, simplicity and parsimonious description of the market derivative prices. Nevertheless, it is a well known fact that it disagrees with reality in a number of significant ways. Even F. Black, 15 years after the publication of the original paper, wrote about the flaws of the model [20]. Indeed, empirical studies show that in the real market the log-return process is not normally distributed and its distribution is often affected by heavy tail, jumps and high peaks. Moreover, the assumption of a constant volatility turns out to be too rigid to model the real world financial market. It is enough to analyse the so-called implied volatility (that is the value of the volatility parameter that, replaced in the Black and Scholes formula, gives the real market price) in a set of traded call options to recognize the well known *smile/skew effect*. In fact, if we plot the implied volatility against the strike price, we can observe that the resulting shape is not a horizontal line, as it should derive from assuming a constant volatility, but it is usually convex and can present higher values for high and low values of the strike price (a *smile*) or asymmetries (from which the term *skew*). Furthermore, the assumption of a constant volatility does not allow to properly price and hedge options which strongly depend on the volatility itself, such as the options on the realized variance or the *cliquet options*.

These results have called for more sophisticated models which can better reflect the reality. Various approaches to model volatility have been introduced over time, paving the way for a huge body of literature devoted to this subject. Let us briefly recall some of the most famous ones.

Roughly speaking, we can recognize two different classes of models. The first class is given by models in which the volatility is assumed to depend on the same noise source as the underlying asset. Here, we can find the so-called *local volatility models*, where the volatility is assumed to be a function of time and of the current underlying asset price. Therefore, the asset price S is modeled by a diffusion process of the type

$$dS_t = \mu(t, S_t)S_t dt + \sigma(t, S_t)S_t dB_t.$$

Under classical assumptions these models preserve the completeness of the market and all the Black-Sholes pricing and hedging theory can be adapted (see, for example, [22, Chapter 2]). The choice of a suitable local volatility function $\sigma = \sigma(t, S)$, is a delicate problem. Bruno Dupire proved in [46] that it is possible to find a function $\sigma = \sigma(t, S)$ which gives theoretical prices matching a given configuration of vanilla options' prices. Typically, the local volatility function is calibrated at $t = 0$ on the market smile and kept frozen afterwards. Therefore, it does not take into account the daily changes in the volatility smile observed in the market. For this reason, local volatility models seem to be an analytically tractable simplification of the reality rather than a representation of how volatility really evolves. Other different models presented in the literature belong to this first class, for instance *path dependent volatility models*, in which volatility depends on the whole past trajectory of the asset price (see [51, 60]).

The second class of models consists of the so-called *stochastic volatility models*. Here, the volatility is modelled by an autonomous stochastic process Y driven by some additional random noise. Typically, a stochastic volatility model is a Markovian model of the form

$$\begin{aligned} dS_t &= \mu_S(t, S_t)S_t dt + \sigma_S(Y_t)S_t dB_t, \\ dY_t &= \mu_Y(t, Y_t)dt + \sigma_Y(t, Y_t)dW_t, \end{aligned}$$

where B and W are possibly correlated Brownian motions. Moreover, often jumps are added to the dynamics of the assets prices and/or their volatilities. The literature on stochastic volatility models is huge. The most successful model is the one introduced by S. Heston [58], which will be extensively studied later on in this thesis. Among the others we cite, for example, the models by Hull and White [61], Bates [17] and Stein and Stein [90]. Moreover, there are also examples of *local-stochastic volatility models* (such as the famous SABR model [57]) in which the volatility coefficient $\sigma_S(Y_t)$ of the underlying asset price is more general and has the form $\sigma_S(S_t, Y_t)$, that is it depends also on the current asset price.

These models are, in general, not complete: the derivative securities are usually not replicable by trading in the underlying. However, this does not affect the practice since the market can be completed with well known procedures of market completion (for example by trading a finite number of vanilla options).

We point out that the research is still fervent in this area. For example, empirical studies have questioned the smoothness of the volatility dynamics. As a consequence, new models called *rough volatility models* have recently been introduced. They are non-Markovian models in which the volatility is driven by a Fractional Brownian motion, see the reference paper [54] and the comprehensive website [86], which gathers all the developments on this subject.

In this thesis we consider Markovian stochastic volatility models and we collect some results on the problem of pricing European and American options. It is divided into two strongly correlated parts. In the first one we study some theoretical properties of the American option prices in Heston-type models. In the second part, we deal with the problem of the numerical computation of the prices, describing and theoretically studying hybrid schemes for pricing European and American options in jump-diffusion stochastic volatility models. More precisely, the thesis is organized as follows:

- Part I: American option prices in Heston-type models
  - Chapter 1. Variational formulation of American option prices in Heston-type models;
  - Chapter 2. American option price properties in Heston-type models.
- Part II: Hybrid schemes for pricing options in jump-diffusion stochastic volatility models
  - Chapter 3. Hybrid Monte Carlo and tree-finite differences algorithm for pricing options in the Bates-Hull-White model;
  - Chapter 4. Weak convergence of Markov chains and numerical schemes for jump diffusion processes.

The above chapters are extracted, sometimes verbatim, from the papers [73, 74, 26, 27] respectively. We now give a brief outline of the main results collected in this thesis.

Part I: American option prices in Heston-type models

The model introduced by S. Heston in 1993 [58] is one of the most widely used stochastic volatility models in the financial world and it was the starting point for several generalizations. In this 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

$$\begin{cases} dS_t = (r-\delta)S_t dt + \sqrt{Y_t} S_t dB_t, & S_0 = s > 0, \\ dY_t = \kappa(\theta - Y_t)dt + \sigma\sqrt{Y_t} dW_t, & Y_0 = y \geq 0, \end{cases} \tag{0.0.1}$$

where B and W denote two correlated Brownian motions with

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

Here $r \geq 0$ and $\delta \geq 0$ are the risk free rate of interest and the continuous dividend rate respectively. The dynamics of the volatility follows a square-root diffusion process, which was originally introduced by E. Feller in 1951 [50] and then rediscovered by Cox, Ingersoll and Ross as an interest rate model in [38]. For this reason this process is known in the financial literature as the CIR process. The parameters $\kappa \geq 0$ and $\theta > 0$ are known respectively as the mean-reversion rate and the long run state, while the parameter $\sigma > 0$ is called the vol-vol (volatility of the volatility). One can observe that the volatility $(Y_t)_t$ tends to fluctuate around the value θ and that κ indicates the velocity of this fluctuation and determines its frequency. This is the mean reversion feature of the CIR process and justifies the names of the constants κ and θ.

It is well known (see, for example, [5, Section 1.2.4]) that under the so called Feller condition $2\kappa\theta \geq \sigma^2$, the process Y with starting condition $Y_0 = y > 0$ remains always positive. On the other hand, if the Feller condition is not satisfied, as happens in many cases of practical importance (see e.g. the calibration results in [30, 44]), Y reaches zero with probability one for any $Y_0 = y \geq 0$.

The great success of the Heston model is due to the fact that the dynamics of the underlying asset price 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 vanilla European options using Fourier transform.

In this framework, the price at time $t \in [0,T]$ of an American option with payoff function φ and maturity T is given by $P(t, S_t, Y_t)$, where

$$P(t,s,y) = \sup_{\tau \in \mathcal{T}_{t,T}} \mathbb{E}\left[e^{-r(\tau-t)}\varphi(S_\tau^{t,s,y})\right],$$

$\mathcal{T}_{t,T}$ being the set of all the stopping times with values in $[t,T]$ and $S^{t,s,y}$ denoting the solution to (0.0.1) with starting condition $S_t = s$, $Y_t = y$.

If we consider, as usual, the log-price process $X_t = \log S_t$, the 2-dimensional diffusion (X, Y) has infinitesimal generator given by

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

and defined on the set $\mathcal{O} = \mathbb{R} \times (0, \infty)$. Note that the differential operator $\mathcal{L}$ has unbounded coefficients and it is not uniformly elliptic: it degenerates on the boundary of $\mathcal{O}$, that is, when the volatility vanishes. This degenerate property gives rise to some technical difficulties when dealing with the theoretical properties of the model, in particular when the problem of pricing American options is considered. In the first part of this thesis we address some of these issues.

Chapter 1: Variational formulation of American option prices in Heston type models

Chapter 1 is devoted to the identification of the American option value function as the unique solution of the associated obstacle problem. Indeed, despite the great success of the Heston model, as far as we know, an exhaustive analysis of the analytic characterization of the value function for American options in Heston-type models is missing in the literature, at least for a large class of payoff functions which include the standard call and put options.

Our approach is based on variational inequalities and extends recent results of Daskalopoulos and Feehan [42, 43] and Feehan and Pop [48] (see also [32]). More precisely, we first study the existence and uniqueness of a weak solution of the associated degenerate parabolic obstacle problem in suitable weighted Sobolev spaces introduced in [42] (Section 1.3). Moreover, we also get a comparison principle. The proof essentially relies on the classical penalization technique (see [19]), with some technical devices due to the degenerate nature of the problem.

Once we have the existence and uniqueness of an analytical weak solution, in Section 1.4 we identify it with the solution to the optimal stopping problem, that is the American option value function. In order to do this, we use suitable estimates on the joint distribution of the log-price process and the volatility process. Moreover, we rely on semi-group techniques and on the affine property of the model.

Chapter 2: American option price properties in Heston type models

In Chapter 2 we study some qualitative properties of an American option value function in the Heston model. We first prove in Section 2.3 that, if the payoff function is convex and satisfies some regularity assumptions, then the option value function is increasing with respect to the volatility variable. Then, in Section 2.4, we focus on the standard put option, that is we fix the payoff function $\varphi(s) = (K - s)_+$, and we extend to the Heston model some results which are well known in the Black and Scholes world, mostly by using probabilistic techniques. In particular, in Section 2.4.1 we introduce the so called *exercise boundary* or *critical price*, that is the map

$$b(t, y) = \inf\{s > 0 \mid P(t, s, y) > (K - s)_+\}, \qquad (t, y) \in [0, T) \times [0, \infty),$$

and we study some features of this function such as continuity properties. Then, in Section 4.3.1 we prove that the American put value function is strictly convex with respect to the stock price in the continuation region, and we do it by using purely probabilistic arguments. In Section 2.4.3 we extend to the stochastic volatility Heston model the *early exercise premium formula*, that is, we prove that

$$P(0, S_0, Y_0) = P_e(0, S_0, Y_0) - \int_0^T e^{-rs}\mathbb{E}[(\delta S_s - rK)\mathbf{1}_{\{S_s \le b(s, Y_s)\}}]ds,$$

where $P_e(0, S_0, Y_0)$ is the price at time 0 of a European put with the same maturity T and strike price K of the original American put with price P. Finally, in Section 2.4.4 we prove a weak form of the *smooth fit principle*, a well known concept in optimal stopping theory.

Part II: Hybrid schemes for pricing options in jump-diffusion stochastic volatility models

In the second part of this thesis we face up with the problem of the numerical computation of European and American options prices in jump-diffusion stochastic volatility models. In particular, we consider the Heston model and some generalizations of it which have other random sources such as jumps and a stochastic interest rate (see [17, 61]).

From a computational point of view, the most delicate point is the treatment of the CIR dynamics for the volatility process in the full parameter regime - it is well known that the standard techniques fail when the square root process is considered. Moreover, one has to be careful in choosing the approximation method according to the European or American option case. In fact, when dealing with European options, i.e. solutions to Partial (Integro) Differential Equation (hereafter P(I)DE) problems, numerical approaches involve tree methods [2, 80], Monte Carlo procedures [3, 4, 6, 8, 98], finite-difference numerical schemes [34, 64, 92] or quantization algorithms [82]. When American options are considered, that is, solutions to specific optimal stopping problems or P(I)DEs with obstacle, it is very useful to consider numerical methods which are able to easily handle dynamic programming principles, for example trees or finite-difference.

In this thesis we consider a backward "hybrid" algorithm which combines:

- finite difference schemes to handle the jump-diffusion price process;
- Markov chains (in particular, multiple jumps trees) to approximate the other random sources, such as the stochastic volatility and the stochastic interest rate.

Chapter 3: Hybrid Monte Carlo and tree-finite differences algorithm for pricing options in the Bates-Hull-White model

In Chapter 3 we focus on the Bates-Hull-White model, where the volatility Y is a CIR process and the underlying asset price process S contains a further noise from a jump as introduced by Merton [77]. Moreover, the interest rate r is stochastic and evolves according to a generalized Ornstein-Uhlenbeck (hereafter OU) process. More precisely, under the pricing measure, we consider the following jump-diffusion model:

$$\begin{aligned} \frac{dS_t}{S_{t^-}} &= (r_t - \delta)dt + \sqrt{Y_t}\, dZ_t^S + dH_t, \\ dY_t &= \kappa_Y(\theta_Y - Y_t)dt + \sigma_Y \sqrt{Y_t}\, dZ_t^Y, \\ dr_t &= \kappa_r(\theta_r(t) - r_t)dt + \sigma_r dZ_t^r, \end{aligned}$$

where, as usual, δ denotes the continuous dividend rate, $S_0, r_0 > 0$, $Y_0 \geq 0$, Z^S, Z^Y and Z^r are correlated Brownian motions and H is a compound Poisson process with intensity λ and i.i.d. jumps $\{J_k\}_k$, that is,

$$H_t = \sum_{k=1}^{K_t} J_k,$$

K denoting a Poisson process with intensity λ. We assume that the random sources , given by the Poisson process K, the jump amplitudes $\{J_k\}_k$ and the 3-dimensional correlated Brownian motion (Z^S, Z^Y, Z^r), are independent.

We refer to the introduction of Chapter 3 for an overview on the existing numerical schemes for pricing options in this model.

Our pricing procedures work as follows. We first approximate both the stochastic volatility and the interest rate processes with a binomial "multiple jumps" tree approach which is based on the techniques originally introduced in [79]. Such a multiple jumps tree approximation for the CIR process was first introduced and analysed in [10], where it is shown to be reliable and accurate without imposing restrictions on the coefficients.

Then, we develop two different pricing procedures. In Section 3.3.3 we propose a (forward) Monte Carlo method, based on simulations for the model following the binomial tree in the direction of both the volatility and the interest rate, and a space-continuous approximation for the underlying asset price process coming from a Euler-Maruyama type scheme.

In Section 3.4, we describe a hybrid backward procedure which works following the tree method in the direction of the volatility and the interest rate and a finite-difference approach in order to handle the underlying asset price process. We also give a first theoretical result on this algorithm, studying some stability properties of the procedure.

Finally, Section 3.5.2 is entirely devoted to numerical results. Several experiments are provided, both for European and American options, with different values of the parameters of the model. In particular, we also consider cases in which the Feller condition for the volatility process is not satisfied. All numerical results show the reliability, the accuracy and the efficiency of both the Monte Carlo and the hybrid algorithm.

Chapter 4: Weak convergence rate of Markov chains and hybrid numerical schemes for jump-diffusion processes

We devote Chapter 4 to the study of the theoretical convergence of a generalization of the hybrid numerical procedure described in Chapter 3. Here we just briefly describe our main results, referring to Section 4.1 for an overview on the existing literature on the rate of convergence of numerical methods for pricing options in Heston-type models.

Recall that the hybrid algorithm uses tree approximations and that, in their turn, tree methods rely on Markov chains. So, we first consider in Section 4.3 a d-dimensional diffusion process $(Y_t)_{t\in[0,T]}$ which evolves according to the SDE

$$dY_t = \mu_Y(Y_t)dt + \sigma_Y(Y_t)dW_t.$$

Fix a natural number $N \geq 1$, $h = T/N$ and assume that $(Y_{nh})_{n=0,\dots,N}$ is approximated by a Markov chain $(Y^h_n)_{n=0,\dots,N}$. It is well known that the weak convergence of Markov chains to diffusions relies on assumptions on the local moments of the approximating process up to order 3 or 4. We prove that, stressing these assumptions, we can study the rate of the weak convergence. This analysis is independent of the financial framework but, as an example, we apply our results to the multiple jumps tree approximation of the CIR process introduced in [10] and used in [24, 25, 27]. Let us mention that our general convergence result (Theorem 4.3.1) may in principle be applied to more general trees constructed through the multiple jumps approach by Nelson and Ramaswamy [79], on which the tree in [10] is based – to our knowledge, a theoretical study of the rate of convergence for such trees is missing in the literature. And it could also be used in other cases, e.g. the recent tree method for the Heston model developed in [2].

Then, in Section 4.4 we combine the Markov chain approach with other numerical techniques in order to handle the different components in jump-diffusion coupled models. In particular, we link $(Y_t)_{t\in[0,T]}$ with a jump-diffusion process $(X_t)_{t\in[0,T]}$ which evolves according to a stochastic differential whose coefficients only depend on the process. In mathematical terms, we consider the stochastic differential equation system

$$\begin{cases} dX_t = \mu_X(Y_t)dt + \sigma_X(Y_t)dB_t + \gamma_X(Y_t)dH_t, \\ dY_t = \mu_Y(Y_t)dt + \sigma_Y(Y_t)dW_t, \end{cases}$$

where H is a compound Poisson process independent of the 2-dimensional Brownian motion (W, B). We generalize the hybrid procedure developed in [24, 25, 27] which works backwardly by approximating the process Y with a Markov chain and by using a different numerical scheme for solving a (local) PIDE allowing us to work in the direction of the process X. We study the speed of convergence of this hybrid approach. The main difficulty comes from the fact that, in general, the hybrid procedure cannot be directly written on a Markov chain, so we cannot apply the convergence results obtained in Section 4.3. Therefore, the idea is to follow the hybrid nature of the procedure: we use classical numerical techniques, that is an analysis of the stability and of the consistency of the method, but in a sense that allows us to exploit the probabilistic properties of the Markov chain approximating the process Y. Again, we provide examples from the financial framework, applying our convergence results to the tree-finite difference algorithm in the Heston or Bates model.

Part I

American option prices in Heston-type models

Chapter 1

Variational formulation of American option prices

1.1 Introduction

The Heston model is the most celebrated stochastic volatility model in the financial world. As a consequence, there is 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 [7]). 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.

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 [66] for the case of the Black and Scholes models).

Concerning the existing literature, E. Ekstrom and J. Tysk in [47] 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 in [42, 43] 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 also [32]). In another paper ([48]) 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 [49]. They also announce results for the parabolic case in [48].

The aim of this chapter is to give a precise analytical characterization of the American option price function in the Heston model 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 [42].

The chapter 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 [42] in the elliptic case. The proof relies, as in [42], on the classical penalization technique introduced by Bensoussan and Lions [19] 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.

1.2 Notations and main results

1.2.1 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 $\left(Y_t, \int_0^t Y_s ds\right)$ and the resulting degenerate model cannot be treated with the techniques we develop in this chapter. Moreover, this particular situation is not very interesting from a financial point of view.

Moreover, we recall that $r \geq 0$ 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 $\kappa > 0$, long run state $\theta > 0$ and volatility of the volatility $\theta > 0$. We stress that we do not require the Feller condition $2\kappa\theta \geq \sigma^2$: the volatility process Y can hit 0 (see, for example, [5, 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

$$X_t = \log S_t - \bar{c}t, \qquad \text{with } \bar{c} = r - \delta - \frac{\rho\kappa\theta}{\sigma}, \tag{1.2.1}$$

which satisfies

$$\begin{cases} dX_t = \left(\frac{\rho\kappa\theta}{\sigma} - \frac{Y_t}{2}\right)dt + \sqrt{Y_t}dB_t, \\ dY_t = \kappa(\theta - Y_t)dt + \sigma\sqrt{Y_t}dW_t. \end{cases} \tag{1.2.2}$$

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 $(\psi(t,X_t,Y_t))_{0\le t\le T}$ 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 $(X_s^{t,x,y},Y_s^{t,y})_{t\le s\le T}$ denotes the solution to (1.2.2) with the starting condition $(X_t,Y_t)=(x,y)$.

Our aim is to give an analytical characterization of the price function P. In this chapter we denote by $\mathcal{L}$ the infinitesimal generator of the two dimensional diffusion (X,Y), 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 $\partial\mathcal{O}=\mathbb{R}\times\{0\}$.

1.2.2 American options and variational inequalities

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

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

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:

(i) Study the obstacle problem

$$\begin{cases} \frac{\partial u}{\partial t} + \mathcal{L}u \leq 0, & u \geq \psi, \quad in\ [0,T] \times \mathcal{O}, \\ \left(\frac{\partial u}{\partial t} + \mathcal{L}u\right)(\psi - u) = 0, & in\ [0,T] \times \mathcal{O}, \\ u(T,x,y) = \psi(T,x,y). \end{cases} \tag{1.2.3}$$

(ii) Show that the discounted price function $\tilde{P}$ is equal to the solution of (1.2.3) 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 (1.2.3).

Weighted Sobolev spaces and bilinear form associated with the Heston operator

We consider the measure first introduced in [42]:

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

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

It is worth noting that in [42] 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 chapter, 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 [42]. The choice of these particular Sobolev spaces will allow us to formulate the obstacle problem (1.2.3) in a variational framework with respect to the measure $\mathfrak{m}_{\gamma,\mu}$.

Definition 1.2.1. *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 $H^0(\mathcal{O}, \mathfrak{m}_{\gamma,\mu}) := L^2(\mathcal{O}, \mathfrak{m}_{\gamma,\mu})$.

(i) 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

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

(ii) 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 1.2.2. *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

$$j_{\gamma,\mu}(x) = \frac{1}{2}\left(1 - \gamma sgn(x) - \mu\rho\sigma\right), \qquad k_{\gamma,\mu}(x) = \kappa - \frac{\gamma\rho\sigma}{2} sgn(x) - \frac{\mu\sigma^2}{2}. \tag{1.2.4}$$

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 chapter 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.

1.2.3 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 $(1+y)^{\frac{3}{2}}\Phi \in L^2([0,T]; H)$, $\frac{\partial \Phi}{\partial t} + \mathcal{L}\Phi \le 0$ 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 [42].

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 1.2.3. *Assume that ψ satisfies Assumption $\mathcal{H}^1$ together with $0 \le \psi \le \Phi$, 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*

$$\begin{cases} -\left(\frac{\partial u}{\partial t}, v-u\right)_H + a(u, v-u) \ge 0, \quad a.e. \ in \ [0,T] \quad v \in L^2([0,T]; V), \ v \ge \psi, \\ u \ge \psi \ a.e. \ in \ [0,T] \times \mathbb{R} \times (0,\infty), \\ u(T) = \psi(T), \\ 0 \le u \le \Phi. \end{cases} \tag{1.2.5}$$

The proof is presented in Section 3 and essentially relies on the penalization technique introduced by Bensoussan and Lions (see also [52]) with some technical devices due to the degenerate nature of the problem. We extend in the parabolic framework the results obtained in [42] for the elliptic case.

The second step is to identify the unique solution of the variational inequality (1.2.5) 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 $(t,x,y) \in [0,T] \times \mathbb{R} \times [0,\infty)$,

$$0 \le \psi(t,x,y) \le C(e^x + e^{Ly}), \tag{1.2.6}$$

and

$$\left|\frac{\partial \psi}{\partial t}(t,x,y)\right| + \left|\frac{\partial \psi}{\partial x}(t,x,y)\right| + \left|\frac{\partial \psi}{\partial y}(t,x,y)\right| \le C(e^{a|x|+by}), \tag{1.2.7}$$

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 (1.2.2)) such that ψ satisfies the assumptions of Theorem 1.2.3. Then, for such γ and μ, we get the following identification result.

Theorem 1.2.4. *Assume that ψ satisfies Assumption $\mathcal{H}^*$. Then, the solution u of the variational inequality* (1.2.5) *associated with ψ is continuous and coincides with the function u^* defined by*

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

1.3 Existence and uniqueness of solutions to the variational inequality

1.3.1 Integration by parts and energy estimates

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

Proposition 1.3.1. *If $u \in H^2(\mathcal{O}, \mathfrak{m})$ and $v \in H^1(\mathcal{O}, \mathfrak{m})$, we have*

$$(\mathcal{L}u, v)_H = -a(u,v). \tag{1.3.8}$$

This result is proved with the same arguments of [42, Lemma 2.23] or [43, 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 $\log S_t$.

Before proving Proposition 1.3.1, we show some preliminary results. The first one is about the standard regularization of a function by convolution.

Lemma 1.3.2. *Let $\varphi : \mathbb{R} \times \mathbb{R} \to \mathbb{R}^+$ be a C^∞ function with compact support in $[-1,+1] \times [-1,0]$ and such that $\int\int \varphi(x,y)dxdy = 1$. For $j \in \mathbb{N}$ we set $\varphi_j(x,y) = j^2\varphi(jx, jy)$. Then, for every function u locally square-integrable on $\mathbb{R} \times (0,\infty)$ and for every compact set K, we have*

$$\lim_{j\to\infty} \iint_K (\varphi_j * u - u)^2(x,y)dxdy = 0.$$

Proof. We first observe that, by using Jensen's inequality with respect to the measure $\varphi_j(\xi,\zeta)d\xi d\zeta$, we get

$$\begin{aligned}\iint_K (\varphi_j * u)^2(x,y)dxdy &\leq \iint_K dxdy \iint \varphi_j(\xi,\zeta)u^2(x-\xi, y-\zeta)d\xi d\zeta \\ &= \iint \varphi_j(\xi,\zeta)d\xi d\zeta \iint \mathbb{1}_K(x+\xi, y+\zeta)u^2(x,y)dxdy.\end{aligned}$$

We deduce, for j large enough,

$$\iint_K (\varphi_j * u)^2(x,y)dxdy \le \iint_{\bar{K}} u^2(x,y)dxdy,$$

where $\bar{K} = \{(x,y) \in \mathcal{O} | d_\infty\big((x,y),K\big) \le \frac{1}{j}\}$. Let ϵ be a positive constant and v be a continuous function such that $\iint_{\bar{K}}(u(x,y)-v(x,y))^2dxdy \le \epsilon$. By using the well known inequality $(x_1+\cdots+x_l)^2 \le l(x_1^2+\cdots+x_l^2)$, we have

$$\begin{aligned}
&\iint_K (\varphi_j * u - u)^2(x,y)dxdy \\
&\le 3\int\int_K (\varphi_j * u - \varphi_j * v)^2(x,y)dxdy + 3\iint_K (\varphi_j * v - v)^2(x,y)dxdy + 3\iint_K (v-u)^2(x,y)dxdy \\
&\le 3\iint_{\bar{K}} (v-u)^2(x,y)dxdy + 3\iint_K (\varphi_j * v - v)^2(x,y)dxdy + 3\iint_{\bar{K}} (v-u)^2(x,y)dxdy \\
&\le 6\epsilon + 3\iint_K (\varphi_j * v - v)^2(x,y)dxdy.
\end{aligned}$$

Since v is continuous, we have $|\varphi_j * v| \le \sup_{x,y\in\bar{K}} |v(x,y)|$ and $\lim_{j\to\infty} \varphi_j * v(x,y) = v(x,y)$ on K. Therefore, by Lebesgue Theorem, we can pass to the limit in the above inequality and we get

$$\limsup_{j\to\infty} \iint_K (\varphi_j * u - u)^2(x,y)dxdy \le 6\epsilon,$$

which completes the proof. □

Then, the following two propositions justify the integration by parts formulas with respect to the measure $\mathfrak{m}$.

Proposition 1.3.3. *Let us consider $u, v : \mathcal{O} \to \mathbb{R}$ locally square-integrable on $\mathcal{O}$, with derivatives u_x and v_x locally square-integrable on $\mathcal{O}$ as well. Moreover, assume that*

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

Then, we have

$$\int_{\mathcal{O}} u_x(x,y)v(x,y)d\mathfrak{m} = -\int_{\mathcal{O}} u(x,y)\left(v_x(x,y) - \gamma sgn(x)v\right)d\mathfrak{m}. \tag{1.3.9}$$

Proof. First we assume that v has compact support in $\mathbb{R}\times(0,\infty)$. For any $j\in\mathbb{N}$ we consider the C^∞ functions $u_j = \varphi_j * u$ and $v_j = \varphi_j * v$, with φ_j as in Lemma 1.3.2. Note that supp $v_j \subset$

supp v + supp φ_j and so, for j large enough, supp $v_j \subset \mathbb{R} \times (0, \infty)$. For any $\epsilon > 0$, integrating by parts, we have

$$\int_{-\infty}^{\infty} (u_j)_x(x,y) v_j(x,y) e^{-\gamma\sqrt{x^2+\epsilon}} dx = -\int_{-\infty}^{\infty} u_j \left((v_j)_x(x,y) - \gamma \frac{x}{\sqrt{x^2+\epsilon}} v_j(x,y) \right) e^{-\gamma\sqrt{x^2+\epsilon}} dx,$$

and, letting $\epsilon \to 0$,

$$\int_{-\infty}^{\infty} (u_j)_x(x,y) v_j(x,y) e^{-\gamma|x|} dx = -\int_{-\infty}^{\infty} u_j \big((v_j)_x(x,y) - \gamma sgn(x) v_j(x,y) \big) e^{-\gamma|x|} dx.$$

Multiplying by $y^{\beta-1}e^{-\mu y}$ and integrating in y we obtain

$$\int_{\mathcal{O}} (u_j)_x(x,y) v_j(x,y) d\mathfrak{m} = -\int_{\mathcal{O}} u_j(x,y) \big((v_j)_x(x,y) - \gamma sgn(x) v_j(x,y) \big) d\mathfrak{m}.$$

Recall that, for j large enough, v_j has compact support in $\mathbb{R} \times (0, \infty)$ and $\mathfrak{m}$ is bounded on this compact. By using Lemma 1.3.2, letting $j \to \infty$ we get

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

Now let us consider the general case of a function v without compact support. We introduce a $C^\infty-$function α with values in $[0,1]$, $\alpha(x,y) = 0$ for all $(x,y) \notin [-2,+2] \times [-2,+2]$, $\alpha(x,y) = 1$ for all $(x,y) \in [-1,+1] \times [-1,+1]$ and a $C^\infty-$function χ with values in $[0,1]$, $\chi(y) = 0$ for all $y \in [0, \frac{1}{2}]$, $\chi(y) = 1$ for all $y \in [+1, \infty)$. We set

$$A_j(x,y) = \alpha\left(\frac{x}{j}, \frac{y}{j}\right) \chi(jy), \qquad j \in \mathbb{N}.$$

For every $j \in \mathbb{N}$, A_j has compact support in $\mathcal{O}$ and we have

$$\int_{\mathcal{O}} u_x(x,y) A_j(x,y) v(x,y) d\mathfrak{m}$$

$$= -\int_{\mathcal{O}} u(x,y) \big(v_x(x,y) - \gamma sgn(x) v(x,y) \big) A_j(x,y) d\mathfrak{m} - \int_{\mathcal{O}} u(x,y) v(x,y) (A_j)_x(x,y) d\mathfrak{m}.$$

The function A_j is bounded by $\|\alpha\|_\infty \|\chi\|_\infty$ and $\lim_{j \to +\infty} A_j(x,y) = 1$ for every $(x,y) \in \mathcal{O}$. Moreover $(A_j)_x(x,y) = \frac{1}{j}\alpha_x\left(\frac{x}{j}, \frac{y}{j}\right)\chi(jy)$, so that

$$\left| \int_{\mathcal{O}} u(x,y) v(x,y) (A_j)_x(x,y) d\mathfrak{m} \right| \leq \frac{C}{j} \int_{\mathcal{O}} \mathbb{1}_{\{|x| \geq j\}} |u(x,y) v(x,y)| d\mathfrak{m},$$

where $C = \|\alpha_x\|_\infty \|\chi\|_\infty$. Therefore, we obtain (1.3.9) letting $j \to \infty$. □

Proposition 1.3.4. *Let us consider $u, v : \mathcal{O} \to \mathbb{R}$ locally square-integrable on $\mathcal{O}$, with derivatives u_y and v_y locally square-integrable on $\mathcal{O}$ as well. Moreover, assume that*

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

Then, we have

$$\int_{\mathcal{O}} yu_y(x,y)v(x,y)d\mathfrak{m} = -\int_{\mathcal{O}} yu(x,y)v_y(x,y)d\mathfrak{m} - \int_{\mathcal{O}} (\beta - \mu y)u(x,y)v(x,y)d\mathfrak{m}. \tag{1.3.10}$$

Proof. If v has compact support in $\mathcal{O}$, we obtain (1.3.10) as in the proof of Proposition 1.3.3. On the other hand, if v does not have compact support,

$$\begin{aligned}\int_{\mathcal{O}} yu_y(x,y)v(x,y)A_j(x,y)d\mathfrak{m} = &-\int_{\mathcal{O}} yu(x,y)v_y(x,y)A_j(x,y)d\mathfrak{m}\\ &-\int_{\mathcal{O}} (\beta-\mu y)u(x,y)v(x,y)A_j(x,y)d\mathfrak{m} - \int_{\mathcal{O}} yu(x,y)v(x,y)(A_j)_y(x,y)d\mathfrak{m},\end{aligned}$$

where $A_j(x,y) = \alpha(\frac{x}{j}, \frac{y}{j})\chi(jy)$, as in the proof of Proposition 1.3.3 but choosing χ such that, moreover, $\|y\chi'(y)\|_\infty < \infty$. We have $(A_j)_y(x,y) = \frac{1}{j}\alpha_y(\frac{x}{j}, \frac{y}{j})\chi(jy) + j\alpha(\frac{x}{j}, \frac{y}{j})\chi'(jy)$. Note that

$$\left|\int_{\mathcal{O}} yu(x,y)v(x,y)j\alpha\left(\frac{x}{j}, \frac{y}{j}\right)\chi'(jy)d\mathfrak{m}\right| \le \int_{\mathcal{O}} \mathbb{1}_{\left\{y \le \frac{1}{j}\right\}} |u(x,y)v(x,y)| \|\alpha\|_\infty \sup_{\zeta>0} |\zeta\chi'(\zeta)|d\mathfrak{m}.$$

The last expression goes to 0 as $j \to \infty$ since $\int_{\mathcal{O}} |u(x,y)v(x,y)|d\mathfrak{m} < \infty$. The assertion follows by passing to the limit $j \to \infty$. □

We can now prove Proposition 1.3.1.

Proof of Proposition 1.3.1. By using Lemma 1.3.3 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

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

□

Remark 1.3.5. *By a closer look at the proof of Proposition 1.3.1 it is clear that the choice of* $\bar{c}$ *in* (1.2.1) *allows to avoid terms of the type* $\int(u_x + u_y)vd\mathfrak{m}$ *in the associated bilinear form* a*. This trick will be crucial in order to obtain suitable energy estimates.*

Recall the well-known inequality

$$bc = (\sqrt{\zeta}b)\left(\frac{c}{\sqrt{\zeta}}\right) \le \frac{\zeta}{2}b^2 + \frac{1}{2\zeta}c^2, \qquad b, c \in \mathbb{R},\ \zeta > 0. \tag{1.3.11}$$

Hereafter we will often apply (1.3.11) in the proofs even if it is not explicitly recalled each time.

We have the following energy estimates.

Proposition 1.3.6. *For every* $u, v \in V$*, the bilinear form* $a(\cdot, \cdot)$ *satisfies*

$$|a(u, v)| \le C_1\|u\|_V\|v\|_V, \tag{1.3.12}$$

$$a(u, u) \ge C_2\|u\|_V^2 - C_3\|(1 + y)^{\frac{1}{2}}u\|_H^2, \tag{1.3.13}$$

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

$$\delta_0 = \sup_{s_1^2+t_1^2>0,\ s_2^2+t_2^2>0} \frac{|s_1s_2 + \rho\sigma s_1t_2 + \rho\sigma s_2t_1 + \sigma^2t_1t_2|}{2\sqrt{(s_1^2 + t_1^2)(s_2^2 + t_2^2)}}, \tag{1.3.14}$$

$$\delta_1 = \inf_{s^2+t^2>0} \frac{s^2 + 2\rho\sigma st + \sigma^2t^2}{2(s^2 + t^2)}, \tag{1.3.15}$$

and

$$K_1 = \sup_{x\in\mathbb{R}} \sqrt{j^2_{\gamma,\mu}(x) + k^2_{\gamma,\mu}(x)}. \tag{1.3.16}$$

It is easy to see that the constants δ_0, δ_1 and K_1 defined in (1.3.14) and (1.3.16) 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 (1.2.4) are bounded).

These energy estimates were already proved in [42, 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 1.3.6. In order to prove (1.3.13), 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 $\zeta = \delta_1/K_1$. (1.3.12) can be proved in a similar way. □

1.3.2 Proof of Theorem 1.2.3

Among the standard assumptions required in [19] 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 (1.3.6) we have, for every $u, v \in V$,

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

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

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 1.3.6. 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

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

The operator a_λ was introduced in [42] 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 1.3.7. *Let δ_0, δ_1, K_1 be defined as in* (1.3.14), (1.3.15) *and* (1.3.16) *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*

$$|a_\lambda(u,v)| \leq C\|u\|_V \|v\|_V, \qquad u, v \in V, \tag{1.3.17}$$

$$a_\lambda(u,u) \geq \frac{\delta_1}{2} \|u\|_V^2, \qquad u \in V, \tag{1.3.18}$$

and

$$|a_\lambda^{(M)}(u,v)| \leq C\|u\|_V \|v\|_V, \qquad u, v \in V, \tag{1.3.19}$$

$$a_\lambda^{(M)}(u,u) \geq \frac{\delta_1}{2} \|u\|_V^2, \qquad u \in V. \tag{1.3.20}$$

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

Proof. The proof for the bilinear form a_λ follows as in [42, 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 $u, v \in V$,

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

so that by straightforward computations we get

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

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

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

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

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

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

From now on in the rest of this chapter we assume $\lambda \geq \frac{\delta_1}{2} + \frac{K_1^2}{2\delta_1}$ as in Lemma 1.3.7. 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 1.3.8. *We stress that Lemma 1.3.7 gives us*

$$\sup_{M>0} \|a_\lambda^{(M)}\| \leq C, \tag{1.3.21}$$

where $C = \delta_0 + K_1 + \lambda$. This will be crucial in the penalization technique we are going to describe in Section 1.3.2. Roughly speaking, in order to prove the existence of a solution of the penalized coercive problem we will introduce in Theorem 1.3.10, 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 1.3.11). Then, thanks to (1.3.21)*, we can deduce estimates on the solution which are uniform in M (see Lemma 1.3.12) 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

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

Penalized problem

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

$$\zeta_\varepsilon(t,u) = -\frac{1}{\varepsilon}(\psi(t)-u)_+ = \frac{1}{\varepsilon}\zeta(t,u), \qquad t \in [0,T], u \in V. \tag{1.3.22}$$

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 [19]).

Lemma 1.3.9. *For any fixed $t \in [0,T]$ the penalizing operator* (1.3.22) *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 1.3.10. *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 $v \in L^2([0,T];V)$,*

$$\begin{cases} -\left(\frac{\partial u_{\varepsilon,\lambda}}{\partial t}(t), v(t)\right)_H + a_\lambda(u_{\varepsilon,\lambda}(t),v(t)) + (\zeta_\varepsilon(t,u_{\varepsilon,\lambda}(t)),v(t))_H = (g(t),v(t))_H, & a.e.\ in\ [0,T], \\ u_{\varepsilon,\lambda}(T) = \psi(T). \end{cases} \tag{1.3.23}$$

Moreover, the following estimates hold:

$$\|u_{\varepsilon,\lambda}\|_{L^\infty([0,T],V)} \leq K, \tag{1.3.24}$$

$$\left\|\frac{\partial u_{\varepsilon,\lambda}}{\partial t}\right\|_{L^2([0,T];H)} \leq K, \tag{1.3.25}$$

$$\frac{1}{\sqrt{\varepsilon}}\left\|(\psi-u_{\varepsilon,\lambda})^+\right\|_{L^\infty([0,T],H)} \leq K, \tag{1.3.26}$$

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 [19], so we omit the proof.

The proof of existence in Theorem 1.3.10 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 1.3.11. *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 $v \in L^2([0,T];V)$*

$$\begin{cases} -\left(\frac{\partial u_{\varepsilon,\lambda,M}}{\partial t}(t), v(t)\right)_H + a_\lambda^{(M)}(u_{\varepsilon,\lambda,M}(t), v(t)) + (\zeta_\varepsilon(t, u_{\varepsilon,\lambda,M}(t)), v(t))_H = (g(t), v(t))_H, & a.e.\ in\ [0,T), \\ u_{\varepsilon,\lambda,M}(T) = \psi(T). \end{cases} \tag{1.3.27}$$

Proof. (i) **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

$$\begin{cases} -\left(\frac{\partial u_j}{\partial t}(t), v\right)_H + a_\lambda^{(M)}(u_j(t), v) - \frac{1}{\varepsilon}((\psi_j(t) - u_j(t))_+, v)_H = (g(t), v)_H, & v \in V_j, \\ u_j(T) = \psi(T). \end{cases} \tag{1.3.28}$$

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. Note that the function $u \to Q_j((\psi_j(t) - u)_+)$ is Lipschitz continuous, since

$$\begin{aligned} \|Q_j((\psi_j(t) - u)_+) - Q_j((\psi_j(t) - v)_+)\|_{V_j} &\le C_j \|Q_j((\psi_j(t) - u)_+) - Q_j((\psi_j(t) - v)_+)\|_H \\ &\le C_j \|u - v\|_H. \end{aligned}$$

On the other hand, the function $(t,u) \to Q_j((\psi_j(t) - u(t)_+)$ is continuous with values in V_j. In fact, we can easily prove that it is weakly continuous, that is, for $v \in V_j$, the application $(t,u) \to (Q_j((\psi_j(t) - u)_+), v)$ is continuous. In fact

$$\begin{aligned} \left|\left(Q_j((\psi_j(t) - u)_+) - Q_j((\psi_j(s) - w)_+), v\right)\right| &\le \left|\left(Q_j((\psi_j(t) - u)_+) - Q_j((\psi_j(s) - u)_+), v\right)\right| \\ &+ \left|\left(Q_j((\psi_j(s) - u)_+) - Q_j((\psi_j(s) - w)_+), v\right)\right|. \end{aligned} \tag{1.3.29}$$

The second term in the right hand side of (1.3.29) goes to 0 by using the Lipschitz continuity proved above. On the other hand, it is easy to prove that for any $u \in V, v \in H^2(\mathcal{O}, \mathfrak{m})$, one has $|(u,v)_V| \leq C\|u\|_H\|v\|_{H^2(\mathcal{O}(\mathfrak{m}))}$. Since $v \in V_j$ we can assume without loss of generality that $v \in H^2(\mathcal{O}, \mathfrak{m})$, so that for the first term in the right hand side of (1.3.29), we easily get

$$\left|\left(Q_j((\psi_j(t)-u)_+) - Q_j((\psi_j(s)-u)_+), v\right)\right| \leq \|\psi_j(t) - \psi_j(s)\|_H \|v\|_{H^2(\mathcal{O},\mathfrak{m})},$$

which goes to 0. Finally, it is easy to see that the term $Q_j g$ belongs to $L^2([0,T]; V_j)$.

Therefore, we can use the Cauchy-Lipschitz Theorem and we deduce the existence and the uniqueness of a solution u_j of (1.3.28), continuous from $[0,T]$ into V_j, a.e. differentiable and with integrable derivative.

(ii) **Estimates on the finite dimensional problem** First, we take $v = u_j(t) - \psi_j(t)$ in (1.3.28). 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

$$\begin{aligned} &\tfrac{1}{4}\|u_j(t)\|_H^2 + \tfrac{\delta_1}{8}\int_t^T \|u_j(s)\|_V^2 ds + \tfrac{1}{\varepsilon}\int_t^T \|(\psi_j(s) - u_j(s))_+\|_H^2 ds \\ &\leq C\left(\left\|\tfrac{\partial \psi}{\partial t}\right\|_{L^2([t,T];V)}^2 + \|g\|_{L^2([t,T];H)}^2 + \|\psi\|_{L^2([t,T];V)}^2 + \|\psi(T)\|_H^2\right). \end{aligned} \tag{1.3.30}$$

We now go back to (1.3.28) 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

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

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

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

Recall that $\bar{a}_\lambda(u_j(t), u_j(t)) \geq \frac{\delta_1}{2}\|u_j(t)\|_V^2$, $|\tilde{a}^{(M)}(u, v)| \leq K_1 \int_{\mathcal{O}} y \wedge M|\nabla u||v| d\mathfrak{m}$ and

$\bar{a}_\lambda(\psi_j(T), \psi_j(T)) = \bar{a}_\lambda(\psi(T), \psi(T)) \leq \|\bar{a}_\lambda\| \|\psi(T)\|_V^2$, so that, for every $\zeta > 0$,

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

From (1.3.30), we already know that

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

then we can finally deduce

$$\begin{aligned}
&\int_t^T \left\| \frac{\partial u_j}{\partial t}(s) \right\|_H^2 ds + \|u_j(t)\|_V^2 + \frac{1}{2\varepsilon} \|(\psi_j(t) - u_j(t))_+\|_H^2 \\
&\qquad \leq C_{\varepsilon,M} \left(\left\| \frac{\partial \psi}{\partial t} \right\|_{L^2([t,T];V)}^2 + \|g\|_{L^2([t,T];H)}^2 + \|\psi\|_{L^2([t,T];V)}^2 + \|\psi(T)\|_V^2 \right),
\end{aligned} \tag{1.3.31}$$

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 (1.3.28) with respect to t for a fixed v independent of t, we obtain that $\bar{u}_j$ satisfies

$$-\left(\frac{\partial \bar{u}_j}{\partial t}(t), v \right)_H + a_\lambda^{(M)}(\bar{u}_j(t), v) - \frac{1}{\varepsilon} \left(\left(\frac{\partial \psi_j}{\partial t}(t) - \bar{u}_j(t) \right) \mathbb{1}_{\{\psi_j(t) \geq u_j(t)\}}, v \right)_H = \left(\frac{\partial g}{\partial t}(t), v \right)_H, \tag{1.3.32}$$

for any $v \in V_j$. As regards the initial condition, from (1.3.28) computed in $t = T$, for every $v \in V_j$ we have

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

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

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

that is, $\left\|\frac{\partial u_j(T)}{\partial t}\right\|_H \leq C\left(\|\psi(T)\|_{H^2(\mathcal{O},\mathfrak{m})} + \|(1+y)\psi(T)\|_H + \|g(T)\|_H\right)$.
We can take $v = \bar{u}_j(t)$ in (1.3.32) and we obtain

$$\begin{aligned}&-\left(\frac{\partial \bar{u}_j}{\partial t}(t), \bar{u}_j(t)\right)_H + a_\lambda^{(M)}(\bar{u}_j(t), \bar{u}_j(t)) - \frac{1}{\varepsilon}\left(\left(\frac{\partial \psi_j}{\partial t}(t) - \bar{u}_j(t)\right)\mathbb{1}_{\{\psi_j(t)\geq u_j(t)\}}, \bar{u}_j(t)\right)_H\\ &\quad= \left(\frac{\partial g}{\partial t}(t), \bar{u}_j(t)\right)_H,\end{aligned}$$

so that

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

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

$$\begin{aligned}&\|\bar{u}_j(t)\|_H^2 + \frac{\delta_1}{2}\int_t^T \|\bar{u}_j(s)\|_V^2 ds\\ &\quad\leq C_\varepsilon\left(\|\psi(T)\|_{H^2(\mathcal{O},\mathfrak{m})}^2 + \|(1+y)\psi(T)\|_H^2 + \|g(T)\|_H^2 + \left\|\frac{\partial \psi}{\partial t}\right\|_{L^2([t,T];H)}^2 + \left\|\frac{\partial g}{\partial t}\right\|_{L^2([t,T];H)}^2\right),\end{aligned} \tag{1.3.33}$$

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

(iii) **Passage to the limit**

Let ε and M be fixed. By passing to a subsequence, from (1.3.31) 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 (1.3.33) we deduce that $\frac{\partial u_{\varepsilon,\lambda,M}}{\partial t} \in L^2([0,T];V)$. Moreover, from (1.3.31), 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 $\|v - v_j\|_V \to 0$. 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 1.3.12. *Assume that, in addition to the assumptions of Proposition 1.3.11,* $\sqrt{1+y}\psi \in L^2([0,T];V)$, $\left|\frac{\partial \psi}{\partial t}\right| \leq \Psi$ *with* $\Psi \in L^2([0,T];V)$ *and* g *satisfies Assumption* $\mathcal{H}^0$*. Let* $u_{\varepsilon,\lambda,M}$ *be the solution of* (1.3.27)*. Then,*

$$\begin{aligned}&\int_t^T \left\|\frac{\partial u_{\varepsilon,\lambda,M}}{\partial s}(s)\right\|_H^2 ds + \|u_{\varepsilon,\lambda,M}(t)\|_V^2 + \frac{1}{\varepsilon}\|(\psi(t) - u_{\varepsilon,\lambda,M}(t))_+\|_H^2 \\ &\leq C\left(\|\Psi\|_{L^2([0,T];V)} + \|\sqrt{1+y}g\|_{L^2([0,T];H)} + \|\sqrt{1+y}\psi\|^2_{L^2([0,T];V)} + \|\psi(T)\|_V^2\right),\end{aligned} \tag{1.3.34}$$

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 (1.3.27), with simple passages we get

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

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

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

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 $n > 0$,

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

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

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

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

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

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

$$\begin{aligned} &\frac{1}{2}\int_{\mathcal{O}} w(t)^2\varphi_n d\mathfrak{m} + \frac{\delta_1}{2}\int_t^T ds\|w(s)\|_{V,n}^2 + \frac{1}{\varepsilon}\int_t^T ds\int_{\mathcal{O}}(-w(s))_+^2\varphi_n d\mathfrak{m} \\ &\leq \int_t^T \left(g(s) + \frac{\partial\psi}{\partial t}(s), w(s)\varphi_n \right)_H ds + \left| \int_t^T a_\lambda^{(M)}(\psi(s), w(s)\varphi_n)ds \right| + K_2\int_t^T \|w(s)\|_V^2 ds, \end{aligned} \tag{1.3.35}$$

and, in the case $n = 0$,

$$\begin{aligned} &\frac{1}{2}\|w(t)\|_H^2 + \frac{\delta_1}{2}\int_t^T \|w(s)\|_V^2 ds + \frac{1}{\varepsilon}\int_t^T ds\int_{\mathcal{O}}(-w(s))_+^2 d\mathfrak{m} \\ &\quad\leq \int_t^T \left(g(s) + \frac{\partial\psi}{\partial t}(s), w(s) \right)_H \|ds + \int_t^T \left| a_\lambda^{(M)}(\psi(s), w(s)) \right| ds. \end{aligned} \tag{1.3.36}$$

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

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

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

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

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

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

Now, if we chose $\zeta_1 = K_3\zeta_2 = \delta_1/4$ and we go back to (1.3.35) and (1.3.36), using $\left|\frac{\partial\psi}{\partial t}\right| \leq \Psi$ we get

$$\begin{aligned}
&\frac{1}{2}\int_{\mathcal{O}} w^2(t)\varphi_n d\mathfrak{m} + \frac{\delta_1}{4}\int_t^T \|w(s)\|^2_{V,n}ds + \frac{1}{\varepsilon}\int_t^T ds\int_{\mathcal{O}}(-w(s))^2_+\varphi_n d\mathfrak{m}\\
&\leq \frac{4}{\delta_1}\left(\|\sqrt{1+y}g\|^2_{L^2([t,T];H)} + \|\sqrt{1+y}\Psi\|^2_{L^2([t,T];H)}\right) + \frac{2K_3^2}{\delta_1}\int_t^T \|\psi(s)\|^2_{V,n}ds + K_2\|w\|^2_{L^2([t,T];H)},\\
&\leq \frac{4}{\delta_1}\left(\|\sqrt{1+y}g\|^2_{L^2([t,T];H)} + \|\sqrt{1+y}\Psi\|^2_{L^2([t,T];H)}\right) + \frac{4K_3^2}{\delta_1}\left\|\sqrt{1+y}\psi\right\|^2_{L^2([t,T];V)}\\
&\quad + K_2\|w\|^2_{L^2([t,T];H)},
\end{aligned} \tag{1.3.37}$$

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

$$\begin{aligned}
&\frac{1}{2}\|w(t)\|^2_H + \frac{\delta_1}{4}\int_t^T\|w(s)\|^2_V ds + \frac{1}{\varepsilon}\int_t^T ds\int_{\mathcal{O}}(-w(s))^2_+ d\mathfrak{m}\\
&\quad\leq \frac{4}{\delta_1}\left(\|g\|^2_{L^2([t,T];H)} + \|\Psi\|^2_{L^2([t,T];H)}\right) + \frac{2K_3^2}{\delta_1}\|\psi\|^2_{L^2([t,T];V)}.
\end{aligned} \tag{1.3.38}$$

From (1.3.38) recalling that $w = u - \psi$ we deduce

$$\begin{aligned}
\int_t^T\|u(s)\|^2_V ds &\leq \int_t^T 2(\|w(s)\|^2_V + \|\psi(s)\|^2_V)ds\\
&\leq \frac{32}{\delta_1^2}\left(\|g\|^2_{L^2([t,T];H)} + \|\Psi\|^2_{L^2([t,T];H)}\right) + \left(\frac{16K_3^2}{\delta_1^2} + 2\right)\|\psi\|^2_{L^2([t,T];V)}.
\end{aligned} \tag{1.3.39}$$

Moreover, combining (1.3.37) and (1.3.38), we have

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

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

$$\int_t^T \|y|\nabla u(s)|\|_H^2 ds \le K_4 \left(\|\sqrt{1+y}g\|^2_{L^2([t,T];H)} + \|\sqrt{1+y}\Psi\|^2_{L^2([t,T];H)} + \|\sqrt{1+y}\psi\|^2_{L^2([t,T];V)}\right), \tag{1.3.40}$$

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 (1.3.34). Taking $v = \frac{\partial u}{\partial t}$ in (1.3.27), we have

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

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

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

Moreover, if we take $v = \Psi(t)$ in (1.3.27), we get

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

so that

$$\frac{1}{\varepsilon}\left((\psi(t) - u(t))_+, \Psi(t)\right)_H \le \left\|\frac{\partial u}{\partial t}(t)\right\|_H \|\Psi(t)\|_H + \|a_\lambda^{(M)}\|\|u(t)\|_V\|\Psi(t)\|_V + \|g(t)\|_H\|\Psi(t)\|_H. \tag{1.3.41}$$

Therefore,

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

hence

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

Integrating between t and T, we get,

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

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

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

where the last inequality follows from (1.3.39) and (1.3.40). Rearranging the terms, we deduce that there exists a constant $C>0$ independent of M and ε such that

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

which concludes the proof. □

Proof of Theorem 1.3.10: 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 (1.3.34) we can repeat the same arguments as in the proof of Proposition 1.3.11 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 (1.3.24), (1.3.25) and (1.3.26) directly follow from (1.3.34) 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, $\left|\frac{\partial \psi}{\partial t}\right| \leq \Psi$ 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 [42, 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 $n \in \mathbb{N}$, $n \in \mathbb{N}$, $|\frac{\partial \psi_n}{\partial t}| \leq \Psi_n$ and all the regularity assumptions required in the first part of the proof are satisfied. Moreover, it is easy to see that $\|\sqrt{1+y}g_n - \sqrt{1+y}g\|_{L^2([0,T];H)} \to 0$, $\|\sqrt{1+y}\psi_n - \sqrt{1+y}\psi\|_{L^2([0,T];V)} \to 0$, $\|\Psi_n - \Psi\|_{L^2([0,T];V)} \to 0$, $\|\psi_n(T) - \psi(T)\|_V \to 0$ as $n \to \infty$. Therefore, the solution $u^n_{\varepsilon,\lambda,M}$ of the equation (1.3.23) with source function g_n and obstacle function ψ_n satisfies

$$\begin{aligned}&\int_t^T \left\|\frac{\partial u^n_{\varepsilon,\lambda,M}}{\partial s}(s)\right\|_H^2 ds + \|u^n_{\varepsilon,\lambda,M}(t)\|_V^2 + \tfrac{1}{\varepsilon}\|(\psi_n(t) - u^n_{\varepsilon,\lambda,M}(t))_+\|_H^2 \\ &\leq C\left(\|\sqrt{1+y}g_n\|_{L^2([0,T];H)} + \|\sqrt{1+y}\psi_n\|^2_{L^2([0,T];V)} + \|\Psi_n\|^2_{L^2([0,T];V)} + \|\psi_n(T)\|_V^2\right).\end{aligned} \tag{1.3.42}$$

Then, we can take the limit for $n \to \infty$ in (1.3.42) 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 1.3.13. *(i) 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* (1.3.23) *with obstacle function ψ_i and source function g. If $\psi_1 \leq \psi_2$, then $u^1_{\varepsilon,\lambda} \leq u^2_{\varepsilon,\lambda}$.*

(ii) 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 (1.3.23) *with obstacle function ψ and source function g_i. If $g_1 \leq g_2$, then $u^1_{\varepsilon,\lambda} \leq u^2_{\varepsilon,\lambda}$.*

(iii) 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 (1.3.23) *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 $\|u^1_{\varepsilon,\lambda} - u^2_{\varepsilon,\lambda}\|_\infty \leq \|\psi_1 - \psi_2\|_\infty$.*

Proposition 1.3.13 can be proved with standard techniques introduced in [19, Chapter 3] so we omit the proof.

Coercive variational inequality

Proposition 1.3.14. *Assume that ψ satisfies Assumption $\mathcal{H}^1$ and g satisfies Assumption $\mathcal{H}^0$. Moreover, assume that $0 \leq \psi \leq \Phi$ with $\Phi \in L^2([0,T]; H^2(\mathcal{O}, \mathfrak{m}))$ such that $\frac{\partial \Phi}{\partial t} + \mathcal{L}\Phi \leq 0$ and $0 \leq g \leq -\frac{\partial \Phi}{\partial t} - \mathcal{L}^\lambda \Phi$. 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*

$$\begin{cases} -\left(\frac{\partial u_\lambda}{\partial t}, v - u_\lambda\right)_H + a_\lambda(u_\lambda, v - u_\lambda) \geq (g, v - u_\lambda)_H, \quad a.e.\ in\ [0,T] \quad v \in L^2([0,T]; V),\ v \geq \psi, \\ u_\lambda(T) = \psi(T), \\ u_\lambda \geq \psi\ a.e.\ in\ [0,T] \times \mathbb{R} \times (0,\infty). \end{cases} \tag{1.3.43}$$

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

Proof. The uniqueness of the solution of (1.3.43) follows by a standard monotonicity argument introduced in [19, Chapter 3] (see [93]). As regards the existence of a solution, we follow the lines of the proof of [19, 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 (1.3.24) and (1.3.25), 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 (1.3.23) when $\psi = g = 0$, while $u = \Phi$ is the unique solution of (1.3.23) 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 1.3.13 implies that $0 \leq u_{\varepsilon,\lambda} \leq \Phi$. 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 $\mathcal{O}_\delta := (-\frac{1}{\delta}, \frac{1}{\delta}) \times (\delta, \frac{1}{\delta})$,

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

and it is enough to let δ goes to 0.

From (1.3.26) 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 $u_\lambda(t) \geq \psi(t)$.

Then we consider the penalized coercive equation in (1.3.23) replacing v by $v - u_{\varepsilon,\lambda}(t)$, with $v \geq \psi(t)$. Since $\zeta_\varepsilon(t, v) = 0$ and $(\zeta_\varepsilon(t, v) - \zeta_\varepsilon(t, u_{\varepsilon,\lambda}(t)), v - u_{\varepsilon,\lambda}(t))_H \geq 0$ we easily deduce that

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

so that, letting ε goes to 0, we have

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

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

The following Comparison Principle is a direct consequence of Proposition 1.3.13,.

Proposition 1.3.15. *(i) For $i = 1, 2$, assume that ψ_i satisfies Assumption $\mathcal{H}^1$, g satisfies Assumption $\mathcal{H}^0$ and $0 \leq \psi_i \leq \Phi$ with $\Phi \in L^2([0, T]; H^2(\mathcal{O}, \mathfrak{m}))$ such that $\frac{\partial\Phi}{\partial t} + \mathcal{L}\Phi \leq 0$ and $0 \leq g \leq -\frac{\partial\Phi}{\partial t} - \mathcal{L}^\lambda\Phi$. Let u^i_λ be the unique solution of* (1.3.43) *with obstacle function ψ_i and source function g. If $\psi_1 \leq \psi_2$, then $u^1_\lambda \leq u^2_\lambda$.*

(ii) For $i = 1, 2$, assume that ψ satisfies Assumption $\mathcal{H}^1$, g_i satisfy Assumption $\mathcal{H}^0$ and $0 \leq \psi \leq \Phi$ with $\Phi \in L^2([0, T]; H^2(\mathcal{O}, \mathfrak{m}))$ such that $\frac{\partial\Phi}{\partial t} + \mathcal{L}\Phi \leq 0$ and $0 \leq g_i \leq -\frac{\partial\Phi}{\partial t} - \mathcal{L}^\lambda\Phi$. Let u^i_λ be the unique solution of (1.3.43) *with obstacle function ψ and source function g_i. If $g_1 \leq g_2$, then $u^1_\lambda \leq u^2_\lambda$.*

(iii) For $i = 1, 2$, assume that ψ_i satisfies Assumption $\mathcal{H}^1$, g satisfies Assumption $\mathcal{H}^0$ and $0 \leq \psi_i \leq \Phi$ with $\Phi \in L^2([0, T]; H^2(\mathcal{O}, \mathfrak{m}))$ such that $\frac{\partial\Phi}{\partial t} + \mathcal{L}\Phi \leq 0$ and $0 \leq g \leq -\frac{\partial\Phi}{\partial t} - \mathcal{L}^\lambda\Phi$. Let u^i_λ be the unique solution of (1.3.43) *with obstacle function ψ_i and source function g. If $\psi_1 - \psi_2 \in L^\infty$, then $u^1_\lambda - u^2_\lambda \in L^\infty$ and $\|u^1_\lambda - u^2_\lambda\|_\infty \leq \|\psi_1 - \psi_2\|_\infty$.*

Non-coercive variational inequality

We can finally prove Theorem 1.2.3. Again, we first study the uniqueness of the solution and then we deal with the existence.

Proof of uniqueness in Theorem 1.2.3. Suppose that there are two functions u_1 and u_2 which satisfy (1.2.5). 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 $[0,T]$,

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

From the energy estimate (1.3.13), 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

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

where $\mu' < \mu$ and $\lambda > 0$. Since $C_2 = \int_{\mathcal{O}} dx dy (1+y) \Phi^2(s) y^{\beta-1} e^{-\gamma|x|} e^{-\mu' y} < \infty$, 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 1.2.3. Given $u_0 = \Phi$, we can construct a sequence $(u_n)_n \subset V$ such that

$$u_n \geq \psi \text{ a.e. in } [0,T]\times\mathcal{O}, \qquad n \geq 1, \tag{1.3.44}$$

$$-\left(\frac{\partial u_n}{\partial t}, v-u_n\right)_H + a(u_n, v-u_n) + \lambda((1+y)u_n, v-u_n)_H \geq \lambda((1+y)u_{n-1}, v-u_n)_H, \\ v\in V, \quad v\geq\psi, \quad \text{a.e. on } [0,T]\times\mathcal{O}, \qquad n\geq 1, \tag{1.3.45}$$

$$u_n(T) = \psi(T), \qquad \text{in } \mathcal{O}, \tag{1.3.46}$$

$$\Phi \geq u_1 \geq u_2 \geq \cdots \geq u_{n-1} \geq u_n \geq \cdots \geq 0, \qquad \text{a.e. on } [0,T]\times\mathcal{O}. \tag{1.3.47}$$

In fact, if we have $0 \leq u_{n-1} \leq \Phi$ for all $n \in \mathbb{N}$, then the assumptions of Proposition 1.3.14 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 (1.3.45) such that $0 \leq u_n \leq \Phi$. (1.3.47) is a simple consequence of Proposition 1.3.15. 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 1.3.16. *Keeping in mind our purpose of identifying the solution of the variational inequality* (1.2.5) *with the American option price we have considered the case without source term* ($g = 0$) *in the variational inequality* (1.2.5). *However, under the same assumptions of Theorem 1.2.3, we can prove in the same way the existence and the uniqueness of a solution of*

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

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 1.3.15 and the proof of Proposition 1.2.3.

Proposition 1.3.17. *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* (1.3.43) *with obstacle function ψ_i. Then:*

(i) If $\psi_1 \leq \psi_2$, then $u_\lambda^1 \leq u_\lambda^2$.

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

1.4 Connection with the optimal stopping problem

Once we have the existence and the uniqueness of a solution u of the variational inequality (1.2.3), 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

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

$\mathcal{T}_{t,T}$ 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.

1.4.1 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

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

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

First of all, we state the following result:

Proposition 1.4.1. *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*

$$\left(\frac{\partial u}{\partial t},v\right)_H+a_\lambda(u,v)=(f,v)_H,\quad v\in V. \tag{1.4.48}$$

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

$$\|u(t)\|_H^2+\frac{\delta_1}{2}\int_0^t\|u(s)\|_V^2ds\leq\|\psi\|_H^2+\frac{2}{\delta_1}\int_0^t\|f(s)\|_H^2ds \tag{1.4.49}$$

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 can be found in the appendix of this chapter. Moreover, we can prove a Comparison Principle for the equation (1.4.48) as we have done for the variational inequality.

We denote $u(t)=\bar{P}_t^\lambda\psi$ the solution of (1.4.48) corresponding to $u(0)=\psi$ and $f=0$. From (1.4.49) we deduce that the operator $\bar{P}_t^\lambda$ is a linear contraction on H and, from uniqueness, we have the semigroup property.

Proposition 1.4.2. *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 $u(t) = \int_0^t \bar{P}^\lambda_s f(t-s)ds = \int_0^t \bar{P}^\lambda_{t-s} f(s)ds$.

Proof. Note that V is dense in H and recall the estimate (1.4.49), so it is enough to prove the assertion for $f = \mathbb{1}_{(t_1,t_2]}\psi$, with $0 \le t_1 < t_2$ and $\psi \in V$. If we set $u(t) = \int_0^t \bar{P}^\lambda_{t-s} f(s)ds$, we have

$$u(t) = \mathbb{1}_{\{t\ge t_1\}} \int_{t_1}^{t\wedge t_2} \bar{P}^\lambda_{t-s}\psi ds = \begin{cases} \int_{t_1}^{t_2} \bar{P}^\lambda_{t-s}\psi ds = \int_{t-t_2}^{t-t_1} \bar{P}^\lambda_s \psi ds & \text{if } t \ge t_2 \\ \displaystyle\int_{t_1}^{t} \bar{P}^\lambda_{t-s}\psi ds = \int_0^{t-t_1} \bar{P}^\lambda_s \psi ds & \text{if } t \in [t_1, t_2) \end{cases}.$$

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

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

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

Remark 1.4.3. *It is not difficult to prove that* $\bar{P}^\lambda_t : L^p(\mathcal{O}, \mathfrak{m}) \to L^p(\mathcal{O}, \mathfrak{m})$ *is a contraction for every* $p \ge 2$, *and it is an analytic semigroup. This is not useful to our purposes so we omit the proof.*

1.4.2 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

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

The operator P^λ_t 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, [5, Section 4.2.1]). We include a proof for convenience.

Proposition 1.4.4. *Let z and w be two complex numbers with nonpositive real parts. The equation*

$$\psi'(t) = \frac{\sigma^2}{2}\psi^2(t) - \kappa\psi(t) + w \tag{1.4.50}$$

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 (1.4.50). 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

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

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

$$M_t=e^{w\int_0^t Y_s ds}F_{z,w}(T-t,Y_t) \tag{1.4.51}$$

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 1.4.5. *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

$$\psi'(t) = \frac{\sigma^2}{2}\psi^2(t) - \kappa\psi(t) + \lambda_2 \tag{1.4.52}$$

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 (1.4.52) 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 1.4.4: just note that the process $(M_t)_t$ defined as in (1.4.51) 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 1.4.6. *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 1.4.5, by using the optional sampling theorem we have

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

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 1.4.4 to prove suitable estimates on the joint law of the process (X, Y).

Lemma 1.4.7. *For every* $q > 0$ *there exists* $C > 0$ *such that for all* $y_0 \geq 0$,

$$\mathbb{E}_{y_0}\left(\int_0^t Y_v dv\right)^{-q} \leq \frac{C}{t^{2q}}. \tag{1.4.53}$$

Proof. If we take $\lambda_1 = 0$ and $\lambda_2 = -s$ with $s > 0$ in Proposition 1.4.5, 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 1.4.5 that $\psi'_{0,-s}(t) = -se^{\int_0^t(\sigma^2\psi(u)-\kappa)du}$. Therefore, since $\psi_{0,-s} = 0$, we have

$$\psi_{0,-s}(t) = -s\int_0^t e^{\int_0^u(\sigma^2\psi(v)-\kappa)dv}du. \tag{1.4.54}$$

Again from the proof of Proposition 1.4.5,

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

so, by using (1.4.54), 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

$$\begin{cases} d\tilde{X}_t = \left(\frac{\rho\kappa}{\sigma} - \frac{1}{2}\right) Y_t dt + \sqrt{1-\rho^2}\sqrt{Y_t} d\tilde{B}_t, \\ dY_t = \kappa(\theta - Y_t)dt + \sigma\sqrt{Y_t} dW_t. \end{cases} \tag{1.4.55}$$

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 1.4.8. *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 1.4.4.

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

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

Then the assertion follows by using Proposition 1.4.4. □

1.4.3 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 1.2.1 for $p \geq 1$. In order to prove the identification of the semigroups, we need the following property of the transition semigroup.

Theorem 1.4.9. *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 1.4.9 will also play a crucial role in order to prove Theorem 1.2.4. 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 1.4.10. *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 1.4.9 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 1.4.8,

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

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

1.4.4 Estimates on the joint law

In this section we prove Theorem 1.4.9. 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, [72, Section 6.2.2]) that the transition density of the process Y is given by

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

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

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

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

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

Therefore

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

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

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

So

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

and the assertion follows. □

We are now ready to prove Theorem 1.4.9, which we have used in order to prove the identification of the semigroups in Proposition 1.4.10 and which we will use again later on in this chapter.

Proof of Theorem 1.4.9. Note that

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

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 (1.4.55) 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 $e^{-\gamma|z| - \bar{\mu} Y_t} dz d\mathbb{P}_{y_0}$, where $\gamma > 0$ and $\bar{\mu} > 0$ will be chosen later on, gives, for every $p > 1$

$$P_t^\lambda f(x_0, y_0) \leq \left[\mathbb{E}_{y_0}\left(\int e^{-\gamma|z| - \bar{\mu} Y_t} \tilde{f}^p\left(\tilde{x}_0 + z, Y_t\right) dz \right) \right]^{1/p} J_q, \tag{1.4.56}$$

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 1.4.11 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

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

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

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

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

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

Note that

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

so that

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

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 $q_1=p_1/(p_1-1)$,

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

where the last inequality follows from Lemma 1.4.7.

We now apply Proposition 1.4.5 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

$$\lambda > \frac{\sigma^2}{2p(p-1)}\bar{\mu}^2 - \kappa\frac{\bar{\mu}}{p} + |\bar{\kappa}|\frac{\gamma}{p} + \frac{1}{2p^2}\gamma^2\bar{\rho}^2. \tag{1.4.57}$$

Going back to (1.4.56) under the condition (1.4.57), 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

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

□

1.4.5 Proof of Theorem 1.2.4

We are finally ready to prove the identification Theorem 1.2.4. We first prove the result under further regularity assumptions on the payoff function ψ, then we deduce the general statement by an approximation technique.

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 1.4.12. *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 (1.2.5) satisfies*

$$-\left(\frac{\partial u}{\partial t}, v\right)_H + a_\lambda(u, v) = (F, v)_H, \qquad a.e.\ in\ [0,T], \quad v \in V. \tag{1.4.58}$$

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 1.3.14. 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

$$\begin{aligned} &-\frac{1}{p}\|w(0)\|^p_{L^p(\mathcal{O},\mathfrak{m})} - \int_0^T a_\lambda((\psi - u_{\varepsilon,\lambda})(t), w^{p-1}(t))dt - \frac{1}{\varepsilon}\int_0^T \|w(t)\|^p_{L^p(\mathcal{O},\mathfrak{m})}dt \\ &= \int_0^T \left(g(t), w^{p-1}(t)\right)_H dt - \int_0^T \left(\frac{\partial \psi}{\partial t}(t), w_+^{p-1}(t)\right)_H dt + \int_0^T a_\lambda(\psi(t), w^{p-1}(t))dt. \end{aligned} \tag{1.4.59}$$

Now, with the usual integration by parts,

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

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 (1.4.59) and using Hölder's inequality,

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

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

$$\left\|\frac{1}{\varepsilon}(\psi - u_{\varepsilon,\lambda})_+\right\|_{L^p([0,T];L^p(\mathcal{O},\mathfrak{m}))} \leq C, \tag{1.4.60}$$

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 1.4.13. *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 1.4.13 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 [63] (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 1.4.14. *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 1.4.8, 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 1.4.5, if

$$\frac{\sigma^2}{2}\left(a\frac{\rho}{\sigma} + b\right)^2 - \kappa\left(a\frac{\rho}{\sigma} + b\right) + \left(a(\frac{\rho\kappa}{\sigma} - \frac{1}{2}) + \frac{a^2}{2}(1 - \rho^2) - \lambda\right) < 0, \tag{1.4.61}$$

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 (1.4.61) 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,

$$\sup_{(t,x,y)\in[0,T]\times K} \mathbb{E}_{x,y}\left(e^{a(1+\epsilon)|X_t|+b(1+\epsilon)Y_t-\lambda(1+\epsilon)\int_0^t Y_s ds}\right) < \infty. \tag{1.4.62}$$

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 1.4.13 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 (1.4.62) 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 1.4.15. *Fix $p > \beta + \frac{5}{2}$ and λ as in Theorem 1.4.9. 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 1.4.14, we have

(i) For every $t \in [0, T]$, $u(t) = P_t^\lambda\psi + \int_0^t P_s^\lambda f(t - s)ds$.

(ii) The function $(t, x, y) \mapsto u(t, x, y)$ is continuous on $[0, T] \times \mathbb{R} \times [0, +\infty)$.

(iii) If $\Lambda_t = \lambda\int_0^t(1 + Y_s)ds$, the process $(M_t)_{0\leq t\leq T}$, 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 1.4.2.

The continuity of $(t,x,y) \mapsto P_t^\lambda \psi(x,y)$ is given by Proposition 1.4.14. 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 1.4.9, we can write for $t \in [0,T]$ and $(x,y) \in K$

$$\begin{aligned}
\int_0^t P_s^\lambda |f_n - f|(t-s,\cdot,\cdot)(x,y)ds &\le \int_0^t \frac{C_{p,K,T}}{s^{\frac{2\beta+3}{2p}}} ds ||(f_n-f)(t-s,\cdot,\cdot)||_{L^p(\mathcal{O},\mathfrak{m})} \\
&\le C_{p,K,T}\left(\int_0^t ||(f_n-f)(t-s,\cdot,\cdot)||^p_{L^p(\mathcal{O},\mathfrak{m})}ds\right)^{1/p}\left(\int_0^t \frac{ds}{s^{\frac{2\beta+3}{2(p-1)}}}\right)^{1-\frac{1}{p}} \qquad (1.4.63) \\
&\le C_{p,K,T}\left(\int_0^T ||(f_n-f)(s,\cdot,\cdot)||^p_{L^p(\mathcal{O},\mathfrak{m})}ds\right)^{1/p}\left(\int_0^T \frac{ds}{s^{\frac{2\beta+3}{2(p-1)}}}\right)^{1-\frac{1}{p}}.
\end{aligned}$$

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,

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

□

We are now ready to prove the following proposition.

Proposition 1.4.16. *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* (1.2.5) *satisfies*

$$u(t,x,y) = u^*(t,x,y), \qquad \text{on } [0,T]\times\bar{\mathcal{O}}, \qquad (1.4.64)$$

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 1.4.12. Note that, thanks to the growth condition (1.2.6), 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 (1.2.7). Then, it is easy to see that we can choose γ and μ in the definition of the measure $\mathfrak{m}$ (see (1.2.2)) 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 1.4.12 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 1.4.15 are satisfied, so the process $(M_t)_{0\le t\le T}$ defined by

$$M_t = e^{-\Lambda_t}u(t,X_t,Y_t) + \int_0^t e^{-\Lambda_s}F(s,X_s,Y_s)ds, \tag{1.4.65}$$

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 (1.4.65) 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

$$\mathbb{E}_{x,y}[u(\tau \wedge \tau_n, X_{\tau\wedge\tau_n}, Y_{\tau\wedge\tau_n})] = u(0,x,y) - \mathbb{E}_{x,y}\left[\int_0^{\tau\wedge\tau_n} (F(s,X_s,Y_s) - \lambda(1+Y_s)u(s,X_s,Y_s))ds\right]. \tag{1.4.66}$$

Since $F - \lambda(1+y)u \geq 0$ we can pass to the limit in the right hand side of (1.4.66) 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 [67] 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 1.4.6. 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 (1.4.66) 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. □

Weaker assumptions on ψ

The last step is to establish the equality $u = u^*$ under weaker assumptions on ψ, so proving Theorem 1.2.4.

Proof of Theorem 1.2.4. 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 1.4.16. For every $n \in \mathbb{N}$, we set $u_n = u_n(t, x, y)$ the unique solution of the variational inequality (1.2.3) 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 1.4.16, 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 1.4.16. This can be done following the very same arguments of [66, Lemma 3.3] so we omit the technical details (see [93]). □

1.5 Appendix: Proof of Proposition 1.4.1

The proof of Proposition 1.4.1 can be carried out following the very same lines of the proof of Proposition 1.3.14. For this reason, we retrace here only the main steps of the proof. So, the first step is to solve the following truncated coercive problem.

Proposition 1.5.1. *Assume $\lambda \geq \frac{\delta_1}{2} + \frac{K_1^2}{2\delta_1}$. For every $\psi \in V$, $f \in L^2_{loc}(\mathbb{R}^+, H)$ and $M > 0$, there exists a unique function $u^{(M)} \in L^2_{loc}(\mathbb{R}^+, V)$, such that $u_t^{(M)} \in L^2_{loc}(\mathbb{R}^+, H)$, $u^{(M)}(0) = \psi$ and*

$$(u_t^{(M)}, v)_H + a_\lambda^{(M)}(u^{(M)}, v) = (f, v)_H, \qquad v \in V.$$

Moreover, for every $t \geq 0$,

$$\|u^{(M)}(t)\|_H^2 + \frac{\delta_1}{2} \int_0^t \|u^{(M)}(s)\|_V^2 ds \leq \|\psi\|_H^2 + \frac{2}{\delta_1} \int_0^t \|f(s)\|_H^2 ds \tag{1.5.67}$$

and

$$\begin{aligned} &\frac{1}{2} \int_0^t \|u_t^{(M)}(s)\|_H^2 ds + \frac{\delta_1}{4} \|u^{(M)}(t)\|_V^2 \\ &\leq \frac{1}{2} \bar{a}_\lambda(\psi, \psi) + \frac{1}{2} \int_0^t \|f(s)\|_H^2 ds + K_1 \int_0^t ds \iint y \wedge M |\nabla u^{(M)}(s)| |u_t^{(M)}(s)| d\mathfrak{m}. \end{aligned} \tag{1.5.68}$$

Proof. Fix $\psi \in V$ and $f \in L^2_{loc}(\mathbb{R}^+, H)$. Let $(V_j)_j$ be an increasing sequence of subspaces of V with finite dimension such that $\bigcup_j V_j$ is dense in V and $\psi \in V_0$. For every j, denote by u_j the unique solution of the differential equation

$$\left(\frac{\partial u_j}{\partial t}, v \right)_H + a_\lambda^{(M)}(u_j, v) = (f, v)_M, \qquad v \in V_j,$$

with $u_j(0) = \psi$.

Taking $v = u_j$ and using the inequality $a_\lambda^{(M)}(u, u) \geq \frac{\delta_1}{2} \|u\|_V$, we get

$$\begin{aligned} \left(\frac{\partial u_j}{\partial t}, u_j \right)_H + a_\lambda^{(M)}(u_j, u_j) &= (f, u_j)_H \\ \frac{1}{2} \frac{d}{dt} \|u_j(t)\|_H^2 + a_\lambda^{(M)}(u_j(t), u_j(t)) &= (f(t), u_j(t))_H \\ \frac{1}{2} \frac{d}{dt} \|u_j(t)\|_H^2 + \frac{\delta_1}{2} \|u_j(t)\|_V^2 &\leq (f(t), u_j(t))_H. \end{aligned}$$

Integrating between 0 and t, we get

$$\frac{1}{2} \|u_j(t)\|_H^2 + \frac{\delta_1}{2} \int_0^t \|u_j(s)\|_V^2 ds \leq \frac{1}{2} \|\psi\|_H^2 + \int_0^t \|f(s)\|_H \|u_j(s)\|_H ds.$$

So, if $f = 0$,

$$\|u_j(t)\|_H^2 + \delta_1 \int_0^t \|u_j(s)\|_V^2 ds \leq \|\psi\|_H^2,$$

and, for $f \neq 0$,

$$\frac{1}{2}\|u_j(t)\|_H^2 + \frac{\delta_1}{2}\int_0^t \|u_j(s)\|_V^2 ds \quad \leq \quad \frac{1}{2}\|\psi\|_H^2 + \frac{\delta_1}{4}\int_0^t \|u_j(s)\|_H^2 ds + \frac{1}{\delta_1}\int_0^t \|f(s)\|_H^2 ds.$$

Therefore,

$$\frac{1}{2}\|u_j(t)\|_H^2 + \frac{\delta_1}{4}\int_0^t \|u_j(s)\|_V^2 ds \quad \leq \quad \frac{1}{2}\|\psi\|_H^2 + \frac{1}{\delta_1}\int_0^t \|f(s)\|_H^2 ds.$$

By taking $v = \partial u_j/\partial t$, we get, using the symmetry of $\bar{a}_\lambda$,

$$\begin{aligned} \left\|\frac{\partial u_j}{\partial t}\right\|_H^2 + a_\lambda^{(M)}\left(u_j, \frac{\partial u_j}{\partial t}\right) &= \left(f, \frac{\partial u_j}{\partial t}\right)_H \\ \left\|\frac{\partial u_j}{\partial t}\right\|_H^2 + \bar{a}_\lambda\left(u_j, \frac{\partial u_j}{\partial t}\right) + \tilde{a}^{(M)}\left(u_j, \frac{\partial u_j}{\partial t}\right) &= \left(f, \frac{\partial u_j}{\partial t}\right)_H \\ \left\|\frac{\partial u_j}{\partial t}\right\|_H^2 + \frac{1}{2}\frac{d}{dt}\bar{a}_\lambda\left(u_j, u_j\right) + \tilde{a}^{(M)}\left(u_j, \frac{\partial u_j}{\partial t}\right) &= \left(f, \frac{\partial u_j}{\partial t}\right)_H, \end{aligned}$$

and, integreting from 0 to t,

$$\begin{aligned} \int_0^t \left\|\frac{\partial u_j}{\partial t}(s)\right\|_H^2 ds + \frac{1}{2}\bar{a}_\lambda\left(u_j(t), u_j(t)\right) = \frac{1}{2}\bar{a}_\lambda\left(\psi, \psi\right) + \int_0^t \left(f(s), \frac{\partial u_j}{\partial t}(s)\right)_H ds \\ - \int_0^t \tilde{a}^{(M)}\left(u_j(s), \frac{\partial u_j}{\partial t}(s)\right)_H ds. \end{aligned}$$

Therefore,

$$\begin{aligned} &\int_0^t \left\|\frac{\partial u_j}{\partial t}(s)\right\|_H^2 ds + \frac{\delta_1}{4}\left\|u_j(t)\right\|_V^2 \\ &\leq \frac{1}{2}\bar{a}_\lambda\left(\psi, \psi\right) + \int_0^t \left(f(s), \frac{\partial u_j}{\partial t}(s)\right)_H ds + K_1\int_0^t ds \int_{\mathcal{O}} y \wedge M|\nabla u_j(s,.)| \left|\frac{\partial u_j}{\partial t}(s,.)\right| d\mathfrak{m} \\ &\leq \frac{1}{2}\bar{a}_\lambda\left(\psi, \psi\right) + \int_0^t \|f(s)\|_H \left\|\frac{\partial u_j}{\partial t}(s)\right\|_H ds + \int_0^t ds \int_{\mathcal{O}} \left(\frac{K_1 y}{2\zeta}|\nabla u_j(s,.)|^2 + \frac{K_1 M\zeta}{2}\left|\frac{\partial u_j}{\partial t}(s,.)\right|^2\right) d\mathfrak{m} \\ &\leq \frac{1}{2}\bar{a}_\lambda\left(\psi, \psi\right) + \int_0^t \|f(s)\|_H \left\|\frac{\partial u_j}{\partial t}(s)\right\|_H ds + \frac{K_1}{2\zeta}\int_0^t \|u_j(s)\|_V^2 ds + \frac{K_1 M}{2}\zeta\int_0^t \left\|\frac{\partial u_j}{\partial t}(s)\right\|_H^2 ds. \end{aligned}$$

Then the assertion follows by passing to the limit as j tends to infinity and by using the estimates above. □

Then, we have the following Lemma.

Lemma 1.5.2. *If, in addiction to the assumptions of Proposition 1.5.1 we also assume* $\sqrt{1+y}f \in L^2_{loc}(\mathbb{R}^+, H)$*, we have*

$$\begin{aligned}\frac{1}{4}\int_0^t \|u_t^{(M)}(s)\|_H^2 ds + \frac{\delta_1}{4}\|u^{(M)}(t)\|_V^2 \quad &\leq \quad \frac{1}{2}\bar{a}_\lambda(\psi,\psi) + \frac{1}{2}\int_0^t \|f(s)\|_H^2 ds \\ &\quad + \frac{4K_1^2K_3}{\delta_1}\left(\|\sqrt{1+y}\psi\|_H^2 + \int_0^t ds\|\sqrt{1+y}f(s)\|_H^2\right).\end{aligned}$$

Proof. Let us denote $\phi_M(x,y) = y \wedge M$. Since ϕ_M and its derivatives are bounded, if $u^{(M)} \in V$, $u^{(M)}\phi_M \in V$. Then, taking $v = u^{(M)}\phi_M$, we get

$$\left(\frac{\partial u^{(M)}}{\partial t}, u^{(M)}\phi_M\right)_H + a_\lambda^{(M)}(u^{(M)}, u^{(M)}\phi_M) \quad = \quad \left(f, u^{(M)}\phi_M\right)_H,$$

which, setting $\phi'_M = \partial\phi_M/\partial y$, can be rewritten as

$$\begin{aligned}&\int_\mathcal{O} \frac{\partial u^{(M)}}{\partial t}u^{(M)}\phi_M d\mathfrak{m} + \int_\mathcal{O}\frac{y}{2}\left(\frac{\partial u^{(M)}}{\partial x}\frac{\partial u^{(M)}}{\partial x} + \sigma^2\frac{\partial u^{(M)}}{\partial y}\frac{\partial u^{(M)}}{\partial y} + 2\rho\sigma\frac{\partial u^{(M)}}{\partial x}\frac{\partial u^{(M)}}{\partial y}\right)\phi_M d\mathfrak{m} \\ &+ \int_\mathcal{O}\frac{y}{2}\left(\rho\sigma\frac{\partial u^{(M)}}{\partial x} + \sigma^2\frac{\partial u^{(M)}}{\partial y}\right)u^{(M)}\phi'_M d\mathfrak{m} + \int_\mathcal{O} y\left(\frac{\partial u^{(M)}}{\partial x}j_{\gamma,\mu} + \frac{\partial u^{(M)}}{\partial y}k_{\gamma,\mu}\right)u^{(M)}\phi_M d\mathfrak{m} \\ &+ \lambda\int_\mathcal{O}(1+y)(u^{(M)})^2\phi_M d\mathfrak{m} = (f, u^{(M)}\phi_M)_H.\end{aligned}$$

Then, by using $0 \leq \phi'_M \leq \mathbb{1}_{\{y\leq M\}}$,

$$\begin{aligned}&\frac{1}{2}\frac{d}{dt}\int_\mathcal{O}(u^{(M)})^2\phi_M d\mathfrak{m} + \delta_1\int_\mathcal{O} y\left|\nabla u^{(M)}\right|^2\phi_M d\mathfrak{m} + \lambda\int_\mathcal{O}(1+y)(u^{(M)})^2\phi_M d\mathfrak{m} \\ &\leq (f, u^{(M)}\phi_M)_H + K_1\int_\mathcal{O} y\left|\nabla u^{(M)}\right|\left|u^{(M)}\right|\phi_M d\mathfrak{m} + \int_\mathcal{O}\frac{y}{2}\left|\rho\sigma\frac{\partial u^{(M)}}{\partial x} + \sigma^2\frac{\partial u^{(M)}}{\partial y}\right|\left|u^{(M)}\right|\phi'_M d\mathfrak{m} \\ &\leq (f, u^{(M)}\phi_M)_H + K_1\int_\mathcal{O} y\left|\nabla u^{(M)}\right|\left|u^{(M)}\right|\phi_M d\mathfrak{m} + \frac{\sqrt{\rho^2\sigma^2+\sigma^4}}{2}\int_\mathcal{O} y\wedge M\left|\nabla u^{(M)}\right||u^{(M)}|d\mathfrak{m} \\ &\leq (f, u^{(M)}\phi_M)_H + \frac{K_1\zeta}{2}\int_\mathcal{O} y\left|\nabla u^{(M)}\right|^2\phi_M d\mathfrak{m} + \frac{K_1}{2\zeta}\int_\mathcal{O} y\left|u^{(M)}\right|^2\phi_M d\mathfrak{m} \\ &\quad + \frac{\sqrt{\rho^2\sigma^2+\sigma^4}}{2}\int_\mathcal{O} y\wedge M\left|\nabla u^{(M)}\right||u^{(M)}|d\mathfrak{m}.\end{aligned}$$

By taking $\zeta = \delta_1/K_1$ and noting that $\int_\mathcal{O} y\wedge M\left|\nabla u^{(M)}\right||u^{(M)}|d\mathfrak{m} \leq \|u^{(M)}\|_V^2$, we get

$$\begin{aligned}&\frac{1}{2}\frac{d}{dt}\int_\mathcal{O}(u^{(M)})^2\phi_M d\mathfrak{m} + \frac{\delta_1}{2}\int_\mathcal{O} y\left|\nabla u^{(M)}\right|^2\phi_M d\mathfrak{m} + \left(\lambda - \frac{K_1^2}{2\delta_1}\right)\int_\mathcal{O}(1+y)(u^{(M)})^2\phi_M d\mathfrak{m} \\ &\quad \leq (f, u^{(M)}\phi_M)_H + K_2\|u^{(M)}\|_V^2\end{aligned}$$

with $K_2 = \frac{\sqrt{\rho^2\sigma^2+\sigma^4}}{2}$ and, by using $\lambda \geq \frac{\delta_1}{2} + \frac{K_1^2}{2\delta_1}$ and integrating from 0 to t,

$$\begin{aligned}
&\frac{1}{2}\int_{\mathcal{O}} (u^{(M)})^2(t,.)\phi_M d\mathfrak{m} + \frac{\delta_1}{2}\int_0^t ds \int_{\mathcal{O}} \left(y\left|\nabla u^{(M)}(s)\right|^2 + (1+y)(u^{(M)})^2(s)\right)\phi_M d\mathfrak{m}\\
&\quad\leq \int_0^t (f(s), u^{(M)}(s)\phi_M)_H ds + \frac{1}{2}\int_{\mathcal{O}} \psi^2\phi_M d\mathfrak{m} + K_2\int_0^t ds \|u^{(M)}(s)\|_V^2 d\mathfrak{m}.
\end{aligned}$$

We have, for every $\zeta > 0$,

$$\int_0^t (f(s), u^{(M)}(s)\phi_M)_H ds \leq \frac{\zeta}{2}\int_0^t ds \int_{\mathcal{O}} \phi_M \left|u^{(M)}(s)\right|^2 d\mathfrak{m} + \frac{1}{2\zeta}\int_0^t ds \int_{\mathcal{O}} \phi_M \left|f(s)\right|^2 d\mathfrak{m}$$

and, taking $\zeta = \delta_1/2$,

$$\begin{aligned}
&\frac{1}{2}\int_{\mathcal{O}} (u^{(M)})^2(t,.)\phi_M d\mathfrak{m} + \frac{\delta_1}{4}\int_0^t ds \int_{\mathcal{O}} \left(y\left|\nabla u^{(M)}(s)\right|^2 + (1+y)(u^{(M)})^2(s)\right)\phi_M d\mathfrak{m}\\
&\quad\leq \frac{1}{\delta_1}\int_0^t ds \int_{\mathcal{O}} \phi_M \left|f(s)\right|^2 d\mathfrak{m} + \frac{1}{2}\int_{\mathcal{O}} \psi^2\phi_M d\mathfrak{m} + K_2\int_0^t \|u^{(M)}(s)\|_V^2 ds.
\end{aligned}$$

Then, by using (1.5.67),

$$\begin{aligned}
&\frac{1}{2}\int_{\mathcal{O}} (u^{(M)})^2(t,.)\phi_M d\mathfrak{m} + \frac{\delta_1}{4}\int_0^t ds \int_{\mathcal{O}} \left(y\left|\nabla u^{(M)}(s)\right|^2 + (1+y)(u^{(M)})^2(s)\right)\phi_M d\mathfrak{m}\\
&\quad\leq \frac{1}{\delta_1}\int_0^t ds \int_{\mathcal{O}} \phi_M \left|f(s)\right|^2 d\mathfrak{m} + \frac{1}{2}\int_{\mathcal{O}} \psi^2\phi_M d\mathfrak{m} + \frac{2K_2}{\delta_1}\|\psi\|_H^2 + \frac{4K_2}{\delta_1^2}\int_0^t \|f(s)\|_H^2 ds\\
&\quad\leq K_3\left(\|\sqrt{1+y}\psi\|_H^2 + \int_0^t ds\|\sqrt{1+y}f(s)\|_H^2\right),
\end{aligned}$$

where $K_3 = \max\left(\frac{1}{\delta_1}, \frac{1}{2}, \frac{2K_2}{\delta_1}, \frac{4K_2}{\delta_1^2}\right)$. Note that K_3 does not depend on M. We deduce from the last inequality that

$$\int_0^t ds \int_{\mathcal{O}} \left|\nabla u^{(M)}(s)\right|^2 \phi_M^2 d\mathfrak{m} \leq \frac{4K_3}{\delta_1}\left(\|\sqrt{1+y}\psi\|_H^2 + \int_0^t ds\|\sqrt{1+y}f(s)\|_H^2\right)$$

and, by using (1.5.68),

$$\begin{aligned}
&\frac{1}{2}\int_0^t \|u_t^{(M)}(s)\|_H^2 ds + \frac{\delta_1}{4}\|u^{(M)}(t)\|_V^2\\
&\quad\leq \frac{1}{2}\bar{a}_\lambda(\psi,\psi) + \frac{1}{2}\int_0^t \|f(s)\|_H^2 ds + K_1\int_0^t ds\int_{\mathcal{O}} y\wedge M|\nabla u^{(M)}(s)\||u_t^{(M)}(s)|d\mathfrak{m}\\
&\quad\leq \frac{1}{2}\bar{a}_\lambda(\psi,\psi) + \frac{1}{2}\int_0^t \|f(s)\|_H^2 ds + \frac{K_1\zeta}{2}\int_0^t ds\int_{\mathcal{O}} |u_t^{(M)}(s)|^2 d\mathfrak{m} + \frac{K_1}{2\zeta}\int_0^t ds\int_{\mathcal{O}} \phi_M^2|\nabla u^{(M)}(s)|^2 d\mathfrak{m}
\end{aligned}$$

By taking $\zeta = 1/(2K_1)$, we get

$$\frac{1}{4}\int_0^t \|u_t^{(M)}(s)\|_H^2 ds + \frac{\delta_1}{4}\|u^{(M)}(t)\|_V^2$$
$$\leq \frac{1}{2}\bar{a}_\lambda(\psi,\psi) + \frac{1}{2}\int_0^t \|f(s)\|_H^2 ds + \frac{4K_1^2K_3}{\delta_1}\left(\|\sqrt{1+y}\psi\|_H^2 + \int_0^t ds\|\sqrt{1+y}f(s)\|_H^2\right).$$

□

Now, in order to prove Proposition 1.4.1, it is enough to let M go to infinity.

Chapter 2

American option price properties in Heston type models

2.1 Introduction

One of the strengths of the Black and Scholes type models relies in their analytical tractability. A large number of papers have been devoted to the pricing of European and American options and to the study of the regularity properties of the price in this framework.

Things become more complicated in the case of stochastic volatility models. Some properties of European options were studied, for example, in [81] but if we consider American options, as far as we know, the existing literature is rather poor. One of the main reference is a paper by Touzi [93], in which the author studies some properties of a standard American put option in a class of stochastic volatility models under classical assumptions, such as the uniform ellipticity of the model.

However, the assumptions in [93] are not satisfied by the well known Heston model because of its degenerate nature and some of the analytical techniques used in [93] cannot be directly applied.

This chapter, which is extracted from [74], is devoted to the study of some properties of the American option price in the Heston model. Our main aim is to extend some well known results in the Black and Scholes world to the Heston type stochastic volatility models. We do it mostly by using probabilistic techniques.

In more details, the chapter is organized as follows. In Section 2.2 we set up our new notation. In Section 2.3, we prove that, if the payoff function is convex and satisfies some regularity assumptions, the American option value function is increasing with respect to the volatility variable. This topic was already addressed in [11] with an elegant probabilistic approach, under the assumption that the coefficients of the model satisfy the well known Feller condition. Here, we prove it without imposing conditions on the coefficients.

Then, in Section 2.4 we focus on the standard American put option. We first generalise to the Heston model the well known notion of critical price or exercise boundary and we study some properties of this function. Then we prove that the American option price is strictly convex in the continuation region with respect to the stock price. This result was already proved in [93] for uniformly elliptic stochastic volatility by using PDE techniques. Here, we extend the result to the degenerate Heston model by using a probabilistic approach. We also give an explicit formulation of the early exercise premium, that is the difference in price between an American option and an otherwise identical European option, and we do it by using results first introduced in [65]. Finally, we provide a weak formulation of the so called smooth fit property. The chapter ends with an appendix, which is devoted to the proofs of some technical results.

2.2 Notation

Recall that in the Heston model we have

$$\begin{cases} \frac{dS_t}{S_t} = (r-\delta)dt + \sqrt{Y_t}dB_t, & S_0 = s > 0, \\ dY_t = \kappa(\theta - Y_t)dt + \sigma\sqrt{Y_t}dW_t, & Y_0 = y \geq 0, \end{cases} \tag{2.2.1}$$

where B and W denote two correlated Brownian motions with correlation coefficient $\rho \in (-1,1)$. Through this chapter we denote by $\mathcal{L}$ the infinitesimal generator of the pair (S,Y), that is the differential operator given by

$$\mathcal{L} = \frac{y}{2}\left(s^2\frac{\partial^2}{\partial s^2} + 2s\rho\sigma\frac{\partial^2}{\partial s\partial y} + \sigma^2\frac{\partial^2}{\partial y^2}\right) + (r-\delta)\,s\frac{\partial}{\partial s} + \kappa(\theta-y)\frac{\partial}{\partial y}. \tag{2.2.2}$$

Let $(S_u^{t,s,y}, Y_u^{t,y})_{u\in[t,T]}$ be the solution of (2.2.1) which starts at time t from the position (s,y). When the initial time is $t=0$ and there is no ambiguity, we will often write $(S_u^{s,y}, Y_u^{y})$ or directly (S_u, Y_u) instead of $(S_u^{0,s,y}, Y_u^{0,y})$. We recall that the price of an American option with a nice enough payoff $(\varphi(S_t))_{t\in[0,T]}$ and maturity T is given by $P_t = P(t, S_t, Y_t)$, where

$$P(t,s,y) = \sup_{\tau\in\mathcal{T}_{t,T}} \mathbb{E}[e^{-r(\tau-t)}\varphi(S_\tau^{t,s,y})],$$

$\mathcal{T}_{t,T}$ being the set of the stopping times with values in $[t, T]$.

It will be useful in this chapter to consider the log-price process, so we set $X_t = \log S_t$. In this case, recall that the pair (X, Y) evolves according to

$$\begin{cases} dX_t = \left(r - \delta - \frac{1}{2}Y_t\right) dt + \sqrt{Y_t} dB_t, & X_0 = x = \log s \in \mathbb{R}, \\ dY_t = \kappa(\theta - Y_t)dt + \sigma\sqrt{Y_t}dW_t, & Y_0 = y \geq 0, \end{cases} \tag{2.2.3}$$

and has infinitesimal generator given by

$$\tilde{\mathcal{L}} = \frac{y}{2}\left(\frac{\partial^2}{\partial x^2} + 2\rho\sigma\frac{\partial^2}{\partial x \partial y} + \sigma^2\frac{\partial^2}{\partial y^2}\right) + \left(r - \delta - \frac{y}{2}\right)\frac{\partial}{\partial x} + \kappa(\theta - y)\frac{\partial}{\partial y}. \tag{2.2.4}$$

With this change of variables, the American option price function is given by $u(t, x, y) = P(t, e^x, y)$, which can be rewritten as

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

where $\psi(x) = \varphi(e^x)$.

2.3 Monotonicity with respect to the volatility

In this section we prove the increasing feature of the option price with respect to the volatility variable under the assumption that the payoff function φ is convex and satisfies some regularity properties. The same topic was addressed by Touzi in [93] for uniformly elliptic stochastic volatility models and by Assing *et al.* [11] for a class of models which includes the Heston model when the Feller condition is satisfied.

For convenience we pass to the logarithm in the $s-$variable and we study the monotonicity of the function u. Note that the convexity assumption on the payoff function $\varphi \in C^2(\mathbb{R})$ corresponds to the condition $\psi'' - \psi' \geq 0$ for the function $\psi(x) = \varphi(e^x)$.

Let us recall some standard notation. For $\gamma > 0$ we introduce the following weighted Sobolev spaces

$$L^2(\mathbb{R}, e^{-\gamma|x|}) = \left\{u : \mathbb{R} \to \mathbb{R} : \|u\|_2^2 = \int u^2(x)e^{-\gamma|x|}dx < \infty\right\},$$

$$W^{1,2}(\mathbb{R}, e^{-\gamma|x|}) = \left\{u \in L^2(\mathbb{R}, e^{-\gamma|x|}) : \frac{\partial u}{\partial x} \in L^2(\mathbb{R}, e^{-\gamma|x|})\right\},$$

$$W^{2,2}(\mathbb{R}, e^{-\gamma|x|}) = \left\{u \in L^2(\mathbb{R}, e^{-\gamma|x|}) : \frac{\partial u}{\partial x}, \frac{\partial^2 u}{\partial x^2} \in L^2(\mathbb{R}, e^{-\gamma|x|})\right\}.$$

Theorem 2.3.1. *Let ψ be a bounded function such that $\psi \in W^{2,2}(\mathbb{R}, e^{-\gamma|x|}) \cap C^2(\mathbb{R})$ and $\psi'' - \psi' \geq 0$. Then the value function u is nondecreasing with respect to the volatility variable.*

In order to prove Theorem 2.3.1, let us consider a smooth approximation $f_n \in C^\infty(\mathbb{R})$ of the function $f(y) = \sqrt{y^+}$, such that f_n has bounded derivatives, $1/n \leq f_n \leq n$, $f_n(y)$ is increasing in y, f_n^2 is Lipschitz continuous uniformly in n and $f_n \to f$ locally uniformly as $n \to \infty$. Moreover, assume that there exists a constant $A > 0$ such that $f_n(x) \leq A(1 + |x|)$.

Then, we consider the sequence of SDEs

$$\begin{cases} dX_t^n = \left(r - \delta - \frac{f_n^2(Y_t^n)}{2}\right) dt + f_n(Y_t^n) dB_t, & X_0^n = x, \\ dY_t^n = \kappa\left(\theta - f_n^2(Y_t^n)\right) dt + \sigma f_n(Y_t^n) dW_t, & Y_0^n = y. \end{cases} \tag{2.3.5}$$

Note that, for every $n \in \mathbb{N}$, the diffusion matrix $a_n(y) = \frac{1}{2}\Sigma_n(y)\Sigma_n(y)^t$, where

$$\Sigma_n(y) = \begin{pmatrix} \sqrt{1-\rho^2} f_n(y) & \rho f_n(y) \\ 0 & \sigma f_n(y) \end{pmatrix},$$

is uniformly elliptic. For any fixed $n \in \mathbb{N}$ the infinitesimal generator of the diffusion (X^n, Y^n) is given by

$$\tilde{\mathcal{L}}^n = \frac{f_n^2(y)}{2}\left(\frac{\partial^2}{\partial x^2} + 2\rho\sigma\frac{\partial^2 u}{\partial x \partial y} + \sigma^2\frac{\partial^2}{\partial y^2}\right) + \left(r - \delta - \frac{f_n^2(y)}{2}\right)\frac{\partial}{\partial x} + \kappa\left(\theta - f_n^2(y)\right)\frac{\partial}{\partial y}$$

and it is uniformly elliptic with bounded coefficients.

We will need the following result.

Lemma 2.3.2. *For any $\lambda > 0$, we have*

$$\lim_{n\to\infty} \mathbb{P}\left(\sup_{t\in[0,T]} |X_t^n - X_t| \geq \lambda\right) = 0 \tag{2.3.6}$$

and

$$\lim_{n\to\infty} \mathbb{P}\left(\sup_{t\in[0,T]} |Y_t^n - Y_t| \geq \lambda\right) = 0. \tag{2.3.7}$$

The proof is inspired by the proof of uniqueness of the solution for the CIR process (see [63, Section IV.3]). We postpone it to the Appendix.

From now on, let us set $\mathbb{E}_{x,y}[\cdot] = \mathbb{E}[\cdot|(X_0, Y_0) = (x, y)]$. For every $n \in \mathbb{N}$, we consider the American value function with payoff ψ and underlying diffusion (X^n, Y^n), that is

$$u^n(t, x, y) = \sup_{\tau\in\mathcal{T}_{0,T-t}} \mathbb{E}_{x,y}\left[e^{-r\tau}\psi(X_\tau^n)\right], \qquad (t, x, y) \in [0, T] \times \mathbb{R} \times [0, \infty).$$

We prove that u^n is actually an approximation of the function u, at least for bounded continuous payoff functions.

Proposition 2.3.3. *Let ψ be a bounded continuous function. Then,*

$$\lim_{n\to\infty} |u^n(t,x,y) - u(t,x,y)| = 0, \qquad (t,x,y) \in [0,T] \times \mathbb{R} \times [0,\infty).$$

Proof. For any $\lambda > 0$,

$$\begin{aligned}
&\left| \sup_{\tau\in\mathcal{T}_{0,T-t}} \mathbb{E}_{x,y}\left[e^{-r\tau}\psi(X^n_\tau)\right] - \sup_{\tau\in\mathcal{T}_{0,T-t}} \mathbb{E}_{x,y}\left[e^{-r\tau}\psi(X_\tau)\right] \right| \\
&\qquad \le \sup_{\tau\in\mathcal{T}_{0,T-t}} \left| \mathbb{E}_{x,y}\left[e^{-r\tau}(\psi(X^n_\tau) - \psi(X_\tau))\right] \right| \\
&\qquad \le \mathbb{E}_{x,y}\left[\sup_{t\in[0,T]} |\psi(X^n_t) - \psi(X_t)|\right] \\
&\qquad \le \mathbb{E}_{x,y}\left[\sup_{t\in[0,T]} |\psi(X^n_t) - \psi(X_t)| \mathbf{1}_{\{|X^n_t - X_t| \le \lambda\}}\right] + 2\|\psi\|_\infty \mathbb{P}\left(\sup_{t\in[0,T]} |X^n_t - X_t| > \lambda\right).
\end{aligned}$$

Then the assertion easily follows using (2.3.6) and the arbitrariness of λ. □

We can now prove that, for every $n \in \mathbb{N}$, the approximated price function u^n is nondecreasing with respect to the volatility variable.

Proposition 2.3.4. *Assume that $\psi \in W^{2,2}(\mathbb{R}, e^{-\gamma|x|}dx) \cap C^2(\mathbb{R})$ and $\psi'' - \psi' \ge 0$. Then $\frac{\partial u^n}{\partial y} \ge 0$ for every $n \in \mathbb{N}$.*

Proof. Fix $n \in \mathbb{N}$. We know from the classical theory of variational inequalities that u^n is the unique solution of the associated variational inequality (see, for example, [66]). Moreover, u^n is the limit of the solutions of a sequence of penalized problems. In particular, consider a family of penalty functions $\zeta_\varepsilon : \mathbb{R} \to \mathbb{R}$ such that, for each $\varepsilon > 0$, ζ_ε is a C^2, nondecreasing and concave function with bounded derivatives, satisfying $\zeta_\varepsilon(u) = 0$, for $u \ge \varepsilon$ and $\zeta_\varepsilon(0) = b$, where b is such that $\tilde{\mathcal{A}}^n\psi \ge b$ (see the proof of Theorem 3 in [71]). Then, there exists a sequence $(u^n_\varepsilon)_{\varepsilon>0}$ such that $\lim_{\varepsilon\to 0} u^n_\varepsilon = u^n$ and, for every $\varepsilon > 0$,

$$\begin{cases} -\frac{\partial u^n_\varepsilon}{\partial t} - \mathcal{A}^n u^n_\varepsilon + \zeta_\varepsilon(u^n_\varepsilon - \psi) = 0, \\ u^n_\varepsilon(T) = \psi(T), \end{cases}$$

where $\tilde{\mathcal{A}}^n = \tilde{\mathcal{L}}^n - r$. In order to simplify the notation, hereafter in this proof we denote by u the function u^n_ε.

Recall that, from the classical theory of parabolic semilinear equations, since $\psi \in C^2(\mathbb{R})$ we have that $u \in C^{2,4}([0,T), \mathbb{R} \times (0,\infty))$ (here we refer, for example, to [70]). Set now $\bar{u} = \frac{\partial u}{\partial y}$. Differentiating the equation satisfied by u^n, we get that $\bar{u}$ satisfies

$$\begin{cases} -\frac{\partial \bar{u}}{\partial t} - \bar{\mathcal{A}}^n \bar{u} = f_n(y) f_n'(y) \left(\frac{\partial^2 u}{\partial x^2} - \frac{\partial u}{\partial x} \right), \\ \bar{u}(T) = 0, \end{cases}$$

where

$$\bar{\mathcal{A}}^n = \frac{f_n^2(y)}{2} \left(\frac{\partial^2}{\partial x^2} + 2\rho\sigma \frac{\partial^2 u}{\partial x \partial y} + \sigma^2 \frac{\partial^2}{\partial y^2} \right) + \left(r - \delta - \frac{f_n^2(y)}{2} + 2\rho\sigma f_n(y) f_n'(y) \right) \frac{\partial}{\partial x}$$
$$+ \left(\kappa \left(\theta - f_n^2(y) \right) + \sigma^2 f_n(y) f_n'(y) \right) \frac{\partial}{\partial y} - 2\kappa f_n(y) f_n'(y) + \zeta_\varepsilon'(u_\varepsilon^n - \psi) - (r - \delta).$$

By using the Comparison principle, we deduce that, if $f_n(y) f_n'(y) \left(\frac{\partial^2 u}{\partial x^2} - \frac{\partial u}{\partial x} \right) \geq 0$, then $\bar{u} \geq 0$ and the assertion follows letting ε tend to 0.

Since f_n is positive and nondecreasing, it is enough to prove that $\frac{\partial^2 u}{\partial x^2} - \frac{\partial u}{\partial x} \geq 0$. We write the equations satisfied by $u' = \frac{\partial u}{\partial x}$ and $u'' = \frac{\partial^2 u}{\partial x^2}$. We have

$$\begin{cases} -\frac{\partial u'}{\partial t} - \tilde{\mathcal{A}}^n u' + \zeta_\varepsilon'(u - \psi)(u' - \psi') = 0, \\ u(T) = \psi, \end{cases} \tag{2.3.8}$$

and

$$\begin{cases} -\frac{\partial u''}{\partial t} - \tilde{\mathcal{A}}^n u'' + \zeta_\varepsilon''(u - \psi)(u' - \psi')^2 + \zeta_\varepsilon'(u - \psi)(u'' - \psi'') = 0, \\ u''(T) = \psi''. \end{cases} \tag{2.3.9}$$

Using (2.3.8) and (2.3.9), we get that $u'' - u'$ satisfies

$$\begin{cases} -\frac{\partial (u'' - u')}{\partial t} - \mathcal{A}^n(u'' - u') + \zeta_\varepsilon'(u - \psi)(u'' - u') = \zeta_\varepsilon'(u - \psi)(\psi'' - \psi') - \zeta_\varepsilon''(u - \psi)(u' - \psi')^2, \\ u''(T) - u'(T) = \psi'' - \psi'. \end{cases} \tag{2.3.10}$$

Recall that $\psi'' - \psi' \geq 0$ by assumption and that ζ_ε is increasing and concave. Then,

$$\zeta_\varepsilon'(u - \psi)(\psi'' - \psi') - \zeta_\varepsilon''(u - \psi)(u' - \psi')^2 \geq 0, \quad u''(T) - u'(T) = \psi'' - \psi' \geq 0,$$

hence, by using again the Comparison principle, we deduce that $u'' - u' \geq 0$ which concludes the proof. □

The proof of Theorem 2.3.1 is now almost immediate.

Proof of Theorem 2.3.1. Thanks to Proposition 2.3.4, the function u^n is increasing in the y variable for all $n \in \mathbb{N}$. Then, the assertion follows by using Proposition 2.3.3. □

2.4 The American put price

From now on we focus our attention on the standard put option with strike price K and maturity T, that is we fix $\varphi(s) = (K - s)_+$ and we study the properties of the function

$$P(t, s, y) = \sup_{\tau \in \mathcal{T}_{t,T}} \mathbb{E}[e^{-r(\tau - t)}(K - S_\tau^{t,s,y})_+]. \tag{2.4.11}$$

The following result easily follows from (2.4.11).

Proposition 2.4.1. *The price function P satisfies:*

(i) $(t, s, y) \mapsto P(t, s, y)$ *is continuous and positive;*

(ii) $t \mapsto P(t, s, y)$ *is nonincreasing;*

(iii) $y \mapsto P(t, s, y)$ *is nondecreasing;*

(iv) $s \mapsto P(t, s, y)$ *is nonincreasing and convex.*

Proof. The proofs of 1. and 2. are classical and straightforward. As regards 3., we note that φ is convex and the function $\psi(x) = (K - e^x)_+$ belongs to the space $W^{1,2}(\mathbb{R}, e^{-\gamma|x|})$ for a $\gamma > 1$ but it is not regular enough to apply Proposition 2.3.1. However, we can proceed with an approximation procedure. Indeed, thanks to density results and [66, Lemma 3.3], we can approximate the function ψ with a sequence of functions $\psi_n \in W^{2,2}(\mathbb{R}, e^{-\gamma|x|}) \cap C^2(\mathbb{R})$ such that $\psi'' - \psi' \geq 0$, so the assertion easily follows passing to the limit. 4. follows from the fact that $\varphi(s) = (K - s)_+$ is nonincreasing and convex. □

Moreover, thanks to the Lipschitz continuity of the payoff function, we have the following result.

Proposition 2.4.2. *The function $x \mapsto u(t, x, y)$ is Lipschitz continuous while the function $y \mapsto u(t, x, y)$ is Hölder continuous. If $2\kappa\theta \geq \sigma^2$ the function $y \mapsto u(t, x, y)$ is locally Lipschitz continuous on* $(0, \infty)$.

Proof. By using standard techniques for the CIR process, as we have done in the proof of Lemma 2.3.2, we can prove that, for every fixed $t \geq 0$ and $y, y' \geq 0$,

$$\mathbb{E}\left|Y_t^y - Y_t^{y'}\right| \leq |y - y'|. \tag{2.4.12}$$

Then, for $(x,y),(x',y') \in \mathbb{R}\times[0,\infty)$ we have

$$
\begin{aligned}
|u(t,x,y)-u(t,x',y')| &= \left|\sup_{\theta\in\mathcal{T}_{t,T}} \mathbb{E}[e^{-r(\theta-t)}(K-e^{X_\theta^{t,x,y}})_+] - \sup_{\theta\in\mathcal{T}_{t,T}} \mathbb{E}[e^{-r(\theta-t)}(K-e^{X_\theta^{t,x',y'}})_+]\right| \\
&\le \sup_{\theta\in\mathcal{T}_{t,T}} \left|\mathbb{E}\Big[e^{-r(\theta-t)}(K-e^{X_\theta^{t,x,y}})_+ - e^{-r(\theta-t)}(K-e^{X_\theta^{t,x',y'}})_+\Big]\right| \\
&\le C\mathbb{E}\left[\sup_{u\in[t,T]}|X_u^{t,x,y}-X_u^{t,x',y'}|\right] \\
&\le C\left(|x-x'| + \int_t^T \mathbb{E}[|Y_u^{t,y}-Y_y^{t,y'}|]du + \mathbb{E}\left[\sup_{s\in[t,T]}\left|\int_t^s(\sqrt{Y_u^{t,y}}-\sqrt{Y_u^{t,y'}})dW_u\right|\right]\right) \\
&\le C\left(|x-x'| + \int_t^T \mathbb{E}[|Y_u^{t,y}-Y_y^{t,y'}|]du + \left(\mathbb{E}\left[\sup_{s\in[t,T]}\left|\int_t^s(\sqrt{Y_u^{t,y}}-\sqrt{Y_u^{t,y'}})dW_u\right|\right]^2\right)^{\frac{1}{2}}\right) \\
&\le C\left(|x-x'| + \int_t^T \mathbb{E}[|Y_u^{t,y}-Y_u^{t,y'}|]du + \left(\mathbb{E}\left[\int_t^T (Y_u^{t,y}-Y_u^{t,y'})du\right]\right)^{\frac{1}{2}}\right) \\
&\le C_T(|x-x'|+\sqrt{|y-y'|}).
\end{aligned}
$$

Now, recall that, if $2\kappa\theta \ge \sigma^2$, the volatility process Y is strictly positive so we can apply Itô's Lemma to the square root function and the process Y_t in the open set $(0,\infty)$. We get

$$
\begin{aligned}
\sqrt{Y_t^y} &= \sqrt{y} + \int_0^t \frac{1}{2\sqrt{Y_u^y}}dY_u^y - \frac{1}{2}\int_0^t \frac{1}{4(Y_u^y)^{\frac{3}{2}}}\sigma^2 Y_u^y du \\
&= \sqrt{y} + \left(\frac{\kappa\theta}{2}-\frac{\sigma^2}{8}\right)\int_0^t \frac{1}{\sqrt{Y_u^y}}du - \frac{\kappa}{2}\int_0^t \sqrt{Y_u^y}du + \frac{\sigma}{2}W_t.
\end{aligned}
$$

Differentiating with respect to y (see also [81]) we deduce that

$$
\frac{\dot{Y}_t^y}{2\sqrt{Y_t^y}} = \frac{1}{2\sqrt{y}} + \left(\frac{\kappa\theta}{2}-\frac{\sigma^2}{8}\right)\int_0^t -\frac{\dot{Y}_u^y}{2(Y_u^y)^{\frac{3}{2}}}du - \frac{\kappa}{2}\int_0^t \frac{\dot{Y}_u^y}{2\sqrt{Y_u^y}}du \le \frac{1}{2\sqrt{y}}, \qquad a.s. \qquad (2.4.13)
$$

since $\kappa\theta \ge \sigma^2/2 \ge \sigma^2/4$ and $Y_t^y > 0$, $\dot{Y}_t^y \ge 0$ (see Theorem 3.7 in [85]).

Therefore, let us consider $y, y' \geq a$. Repeating the same calculations as before

$$\begin{aligned}
&|u(t,x,y)-u(t,x,y')|\\
&\leq C\left(\int_t^T \mathbb{E}[|Y_u^{t,y}-Y_u^{t,y'}|]du+\left(\mathbb{E}\left[\sup_{s\in[t,T]}\left|\int_t^s(\sqrt{Y_u^{t,y}}-\sqrt{Y_u^{t,y'}})dW_u\right|\right]^2\right)^{\frac{1}{2}}\right)\\
&\leq C\left(\int_t^T \mathbb{E}[|Y_u^{t,y}-Y_u^{t,y'}|]du+\left(\mathbb{E}\left[\int_t^T(\sqrt{Y_u^{t,y}}-\sqrt{Y_u^{t,y'}})^2du\right]\right)^{\frac{1}{2}}\right)\\
&= C\left(\int_t^T \mathbb{E}[|Y_s^{t,y}-Y_s^{t,y'}|]du+\left(\mathbb{E}\left[\int_t^T du\left(\int_y^{y'}\frac{\dot{Y}_u^{t,w}}{2\sqrt{Y_u^{t,w}}}dw\right)^2\right]\right)^{\frac{1}{2}}\right)\\
&\leq C_T\left(|y-y'|+\left(\mathbb{E}\left[\int_t^T\left(\frac{1}{2\sqrt{a}}|y-y'|\right)^2du\right]\right)^{\frac{1}{2}}\right)\\
&\leq C_T|y-y'|,
\end{aligned}$$

which completes the proof. □

Remark 2.4.3. *Studying the properties of the put price also clarifies the behaviour of the call price since it is straightforward to extend to the Heston model the symmetry relation between call and put prices. In fact, let us highlight the dependence of the prices with respect to the parameters K, r, δ, ρ, that is let us write*

$$P(t,x,y;K,r,\delta,\rho)=\sup_{\tau\in\mathcal{T}_{t,T}}\mathbb{E}[e^{-r(\tau-t)}(K-S_\tau^{t,s,y})_+],$$

for the put option price and

$$C(t,s,y;K,r,\delta,\rho)=\sup_{\tau\in\mathcal{T}_{t,T}}\mathbb{E}[e^{-r(\tau-t)}(S_\tau^{t,s,y}-K)_+],$$

for the call option. Then, we have $C(t,s,y;K,r,\delta,\rho)=P(t,K,y;x,\delta,r,-\rho)$.

In fact, for every $\tau\in\mathcal{T}_{t,T}$, we have

$$\begin{aligned}
&\mathbb{E}e^{-r(\tau-t)}\left(se^{\int_t^\tau\left(r-\delta-\frac{Y_s^{t,y}}{2}\right)ds+\int_t^\tau\sqrt{Y_s^{t,y}}dB_s}-K\right)_+\\
&\quad=\mathbb{E}e^{-\delta(\tau-t)}e^{\int_t^\tau\sqrt{Y_s^{t,y}}dB_s-\int_t^\tau\frac{Y_s^{t,y}}{2}ds}\left(x-Ke^{\int_t^\tau\left(\delta-r+\frac{Y_s^{t,y}}{2}\right)ds-\int_t^\tau dB_s}\right)_+\\
&\quad=\mathbb{E}e^{-\delta(\tau-t)}e^{\int_t^T\sqrt{Y_s^{t,y}}dB_s-\int_t^T\frac{Y_s^{t,y}}{2}ds}\left(x-Ke^{\int_t^\tau\left(\delta-r+\frac{Y_s^{t,y}}{2}\right)ds-\int_t^\tau\sqrt{Y_s^{t,y}}dB_s}\right)_+,
\end{aligned}$$

where the last equality follows from the fact that $(e^{\int_t^s \sqrt{Y_s^{t,y}}dB_s - \int_t^s \frac{Y_s^{t,y}}{2}ds})_{s\in[t,T]}$ *is a martingale. Then, note that the process* $\hat{B}_t = B_t - \sqrt{Y_t^{t,y}}t$ *is a Brownian motion under the probability measure* $\hat{P}$ *which has density* $d\hat{\mathbb{P}}/d\mathbb{P} = e^{\int_t^T \sqrt{Y_s^{t,y}}dB_s - \int_t^T \frac{Y_s^{t,y}}{2}ds}$*. Therefore*

$$\mathbb{E}e^{-r(\tau-t)}\left(se^{\int_t^\tau\left(r-\delta-\frac{Y_s^{t,y}}{2}\right)ds+\int_t^\tau\sqrt{Y_s^{t,y}}dB_s}-K\right)_+ = \hat{\mathbb{E}}e^{-\delta(\tau-t)}\left(x-Ke^{\int_t^\tau\left(\delta-r-\frac{Y_s^{t,y}}{2}\right)ds-\int_t^\tau\sqrt{Y_s^{t,y}}dB_s}\right)_+.$$

Under the probability $\hat{\mathbb{P}}$*, the process* $(-\hat{B}, W)$ *is a Brownian motion with correlation coefficient* $-\rho$ *so that the assertion follows.*

2.4.1 The exercise boundary

Let us introduce the so called continuation region

$$\mathcal{C} = \{(t,s,y) \in [0,T) \times (0,\infty) \times [0,\infty) : P(t,s,y) > \varphi(s)\}$$

and its complement, the exercise region

$$\mathcal{E} = \mathcal{C}^c = \{(t,s,y) \in [0,T) \times (0,\infty) \times [0,\infty) : P(t,s,y) = \varphi(s)\}.$$

Note that, since P and φ are both continuous, $\mathcal{C}$ is an (relative) open set while $\mathcal{E}$ is a closed set.

Generalizing the standard definition given in the Black and Scholes type models, we consider the *critical exercise price* or *free exercise boundary*, defined as

$$b(t,y) = \inf\{s > 0 | P(t,s,y) > (K-s)_+\}, \qquad (t,y) \in [0,T) \times [0,\infty).$$

We have $P(t,s,y) = \varphi(s)$ for $s \in [0, b(t,y))$ and also for $s = b(t,y)$, due to the continuity of P and φ. Moreover, since P is convex, we can write

$$\mathcal{C} = \{(t,s,y) \in [0,T) \times (0,\infty) \times [0,\infty) : s > b(t,y)\}$$

and

$$\mathcal{E} = \{(t,s,y) \in [0,T) \times (0,\infty) \times [0,\infty) : s \le b(t,y)\}.$$

We now study some properties of the free boundary $b : [0,T) \times [0,\infty) \to \mathbb{R}$. First of all, we have the following simple result.

Proposition 2.4.4. *We have:*

(i) for every fixed $y \in [0,\infty)$, the function $t \mapsto b(t,y)$ is nondecreasing and right continuous;

(ii) for every fixed $t \in [0,T)$, the function $y \mapsto b(t,y)$ is nonincreasing and left continuous.

Proof. 1. Recalling that the map $t \mapsto P(t,s,y)$ is nonincreasing, we directly deduce that $t \mapsto b(t,y)$ is nondecreasing. Then, fix $t \in [0,T)$ and let $(t_n)_{n\geq 1}$ be a decreasing sequence such that $\lim_{n\to\infty} t_n = t$. The sequence $(b(t_n,y))_n$ is nonincreasing so that $\lim_{n\to\infty} b(t_n,y)$ exists and we have $\lim_{n\to\infty} b(t_n,y) \geq b(t,y)$. On the other hand, we have

$$P(t_n, b(t_n,y), y) = \varphi(b(t_n,y)) \qquad n \geq 1,$$

and, by the continuity of P and φ,

$$P(t_n, \lim_{n\to\infty} b(t_n,y), y) = \varphi(\lim_{n\to\infty} b(t_n,y)).$$

We deduce by the definition of b that $\lim_{n\to\infty} b(t_n,y) \leq b(t,y)$ which concludes the proof.

2. The second assertion can be proved with the same arguments, this time recalling that $y \mapsto P(t,s,y)$ is a nondecreasing function. □

Note that, since $P > 0$, we have $b(t,y) \in [0,K)$. Indeed, we can prove the positivity of the function.

Proposition 2.4.5. *We have $b(t,y) > 0$ for every $(t,y) \in [0,T) \times [0,\infty)$.*

Proof. Without loss of generality we can assume that $0 < t < T$, since T is arbitrary and the put price is a function of $T - t$. Suppose that $b(t^*,y^*) = 0$ for some $(t^*,y^*) \in (0,T) \times [0,\infty)$. Since $b(t,y) \geq 0$, $t \mapsto b(t,y)$ is nondecreasing and $y \mapsto b(t,y)$ is nonincreasing, we have $b(t,y) = 0$ for $(t,y) \in (0,t^*) \times (y^*,\infty)$, so that

$$P(t,s,y) > \varphi(s), \qquad (t,s,y) \in (0,t^*) \times (0,\infty) \times (y^*,\infty).$$

To simplify the calculations, we pass to the logarithm in the space variable and we consider the functions $u(t,x,y) = P(t,e^x,y)$ and $\psi(x) = \varphi(e^x)$. We have $u(t,x,y) > \psi(x)$ and

$$(\partial_t + \tilde{\mathcal{L}} - r)u = 0 \qquad \text{on } (0,t^*) \times \mathbb{R} \times (y^*,\infty),$$

where $\tilde{\mathcal{L}}$ was defined in (2.2.4). Since $t \mapsto u(t,x,y)$ is nondecreasing, we deduce that, for $t \in (0,t^*)$, $(\tilde{\mathcal{L}} - r)u = \partial_t u \geq 0$ in the sense of distributions. Therefore, for any nonnegative and C^∞ test functions θ, ψ and ζ which have support respectively in $(0,t^*)$, $(-\infty,\infty)$ and (y^*,∞), we have

$$\int_0^{t^*} \theta(t)dt \int_{-\infty}^{\infty} dx \int_{y^*}^{\infty} dy \tilde{\mathcal{L}}u(t,x,y)\phi(x)\zeta(y) \geq r \int_0^{t^*} \theta(t)dt \int_{-\infty}^{\infty} dx \int_{y^*}^{\infty} dy (K - e^x)\phi(x)\zeta(y),$$

or equivalently, by the continuity of the integrands in t,

$$\int_{-\infty}^{\infty} dx \int_{y^*}^{\infty} dy \tilde{\mathcal{L}}u(t,x,y)\phi(x)\zeta(y) \geq r \int_{-\infty}^{\infty} dx \int_{y^*}^{\infty} dy (K-e^x)\phi(x)\zeta(y). \tag{2.4.14}$$

Let χ_1 and χ_2 be two nonnegative C^∞ functions such that $\operatorname{supp}\chi_1 \subseteq [-1,0]$, $\operatorname{supp}\chi_2 \subseteq [0,1]$ and $\int \chi_1(x)dx = \int \chi_2(x)dx = 1$. Let us apply (2.4.14) with $\phi(x) = \lambda\chi_1(\lambda x)$ and $\zeta(y) = \sqrt{\lambda}\chi_2(\sqrt{\lambda}(y - y^*))$, with $\lambda > 0$. For the right hand side of (2.4.14), we have

$$r \int_{-\infty}^{\infty} dx \int_{y^*}^{\infty} dy (K-e^x)\phi(x)\zeta(y) = rK - r\int_{-\infty}^{\infty} e^{\frac{x}{\lambda}}\chi_1(x)dx.$$

Since $\operatorname{supp}\chi_1 \subset [-1,0]$, $\lim_{\lambda\to 0}\int e^{\frac{x}{\lambda}}\chi_1(x)dx = 0$, so that

$$\lim_{\lambda\to 0} r\int_{\mathbb{R}} dx \int_{-\infty}^{y^*} dy(K-e^x)\phi(x)\zeta(y) = rK > 0. \tag{2.4.15}$$

As regards the left hand side of (2.4.14), we have

$$\begin{aligned}
&\int_{-\infty}^{+\infty} dx \int_{y^*}^{\infty} \tilde{\mathcal{L}}u(t,x,y)\phi(x)\zeta(y)dy \\
&= \int_{-\infty}^{+\infty} dx \int_{y^*}^{\infty} \frac{y}{2}\left(\frac{\partial^2 u}{\partial x^2}(t,x,y) + 2\rho\sigma\frac{\partial^2 u}{\partial x\partial y}(t,x,y) + \sigma^2\frac{\partial^2 u}{\partial y^2}(t,x,y)\right)\lambda\chi_1(\lambda x)\sqrt{\lambda}\chi_2(\sqrt{\lambda}(y-y^*))dy \\
&+ \int_{-\infty}^{+\infty} dx \int_{y^*}^{\infty} \left(\left(r-\delta-\frac{y}{2}\right)\frac{\partial u}{\partial x}(t,x,y) + \kappa(\theta-y)\frac{\partial u}{\partial y}(t,x,y)\right)\lambda\chi_1(\lambda x)\sqrt{\lambda}\chi_2(\sqrt{\lambda}(y-y^*))dy.
\end{aligned}$$

We first study the second order derivatives term. Integrating by parts two times we have

$$\begin{aligned}
&\int_{-\infty}^{+\infty} dx \int_{y^*}^{\infty} \frac{y}{2}\frac{\partial^2}{\partial x^2}u(t,x,y)\lambda\chi_1(\lambda x)\sqrt{\lambda}\chi_2(\sqrt{\lambda}(y-y^*))dy \\
&\quad = \int_{-\infty}^{+\infty} dx \int_{y^*}^{\infty} \frac{y}{2}u(t,x,y)\lambda^3\chi_1''(\lambda x)\sqrt{\lambda}\chi_2(\sqrt{\lambda}(y-y^*))dy \\
&\quad = \lambda^{\frac{3}{2}}\int_{-\infty}^{+\infty} dx \int_0^{\infty} \frac{1}{2}\left(y+\sqrt{\lambda}y^*\right)u\left(t,\frac{x}{\lambda},\frac{y}{\sqrt{\lambda}}+y^*\right)\chi_1''(x)\chi_2(y)dy.
\end{aligned}$$

Since u is bounded and χ_2 has support in $[0,1]$, the last term goes to 0 as λ tends to 0. For the mixed derivative term, since $\chi_2(0) = 0$,

$$\begin{aligned}
&\int_{-\infty}^{+\infty} dx \int_{y^*}^{\infty} \rho\sigma y \frac{\partial^2}{\partial x \partial y} u(t,x,y) \lambda\chi_1(\lambda x)\sqrt{\lambda}\chi_2(\sqrt{\lambda}(y-y^*))dy \\
&= -\rho\sigma \int_{-\infty}^{+\infty} dx \int_{y^*}^{\infty} y \frac{\partial}{\partial y} u(t,x,y)\lambda^2 \chi_1'(\lambda x)\sqrt{\lambda}\chi_2(\sqrt{\lambda}(y-y^*))dy \\
&= \rho\sigma \int_{-\infty}^{+\infty} dx \int_{y^*}^{\infty} u(t,x,y)\lambda^2 \chi_1'(\lambda x)\sqrt{\lambda}\chi_2(\sqrt{\lambda}(y-y^*))dy \\
&\quad + \rho\sigma \int_{-\infty}^{+\infty} dx \int_{y^*}^{\infty} u(t,x,y)\lambda^2 \chi_1'(\lambda x)\lambda\chi_2'(\sqrt{\lambda}(y^*-y))dy \\
&= \lambda\rho\sigma \int_{-\infty}^{+\infty} dx \int_{0}^{\infty} u\left(t, \frac{x}{\lambda}, \frac{y}{\sqrt{\lambda}} + y^*\right) \chi_1'(x)\chi_2(y)dy \\
&\quad + \lambda^{\frac{3}{2}}\rho\sigma \int_{-\infty}^{+\infty} dx \int_{0}^{\infty} u\left(t, \frac{x}{\lambda}, \frac{y}{\sqrt{\lambda}} + y^*\right) \chi_1'(x)\chi_2'(y)dy,
\end{aligned}$$

which goes to 0 as λ tends to 0 with the same arguments as before.

Moreover, integrating by parts two times, we have

$$\begin{aligned}
&\int_{-\infty}^{+\infty} dx \int_{y^*}^{\infty} \frac{y}{2}\sigma^2 \frac{\partial^2}{\partial y^2} u(t,x,y)\lambda\chi_1(\lambda x)\sqrt{\lambda}\chi_2(\sqrt{\lambda}(y-y^*))dy \\
&= -\int_{-\infty}^{+\infty} dx \int_{y^*}^{\infty} \frac{\sigma^2}{2}\frac{\partial}{\partial y} u(t,x,y)\lambda\chi_1(\lambda x)\left(\sqrt{\lambda}\chi_2(\sqrt{\lambda}(y-y^*)) + y\lambda\chi_2'(\sqrt{\lambda}(y-y^*))\right)dy \\
&= \int_{-\infty}^{+\infty} dx \int_{y^*}^{\infty} \frac{\sigma^2}{2} u(t,x,y)\left(2\lambda\chi_1(\lambda x)\lambda\chi_2'(\sqrt{\lambda}(y-y^*))\right)dy \\
&= \sqrt{\lambda}\sigma^2 \int_{-\infty}^{+\infty} dx \int_{0}^{\infty} u\left(t, \frac{x}{\lambda}, \frac{y}{\sqrt{\lambda}} + y^*\right)\chi_1(x)\left(\lambda\chi_2'(y) + \frac{1}{2}\lambda^{\frac{3}{2}}\left(y + \sqrt{\lambda}y^*\right)\chi_2''(y)\right)dy
\end{aligned}$$

which again tends to 0 as λ goes to 0. We now study the terms in (2.4.14) which contains the first order derivatives of u. First, note that

$$\begin{aligned}
&\int_{-\infty}^{+\infty} dx \int_{y^*}^{\infty} \left(r - \delta - \frac{y}{2}\right)\frac{\partial}{\partial x} u(t,x,y)\lambda\chi_1(\lambda x)\sqrt{\lambda}\chi_2(\sqrt{\lambda}(y-y^*))dy \\
&= -\int_{-\infty}^{+\infty} dx \int_{y^*}^{\infty} \left(r - \delta - \frac{y}{2}\right) u(t,x,y)\lambda^2\chi_1'(\lambda x)\sqrt{\lambda}\chi_2(\sqrt{\lambda}(y-y^*))dy \\
&= -\sqrt{\lambda}\int_{-\infty}^{+\infty} dx \int_{0}^{\infty} \left(\sqrt{\lambda}r - \sqrt{\lambda}\delta - \frac{1}{2}\left(y + \sqrt{\lambda}y^*\right)\right) u\left(t, \frac{x}{\lambda}, \frac{y}{\sqrt{\lambda}} + y^*\right)\chi_1'(x)\chi_2(y)dy.
\end{aligned}$$

Again, passing to the limit, the last term tends to 0. On the other hand,

$$\begin{aligned}
&\int_{-\infty}^{+\infty} dx \int_{y^*}^{\infty} \kappa(\theta - y)\frac{\partial}{\partial y}u(t,x,y)\lambda\chi_1(\lambda x)\sqrt{\lambda}\chi_2(\sqrt{\lambda}(y-y^*))dy \\
&\quad = \int_{-\infty}^{+\infty} dx \int_{y^*}^{\infty} \kappa\theta\frac{\partial}{\partial y}u(t,x,y)\lambda\chi_1(\lambda x)\sqrt{\lambda}\chi_2(\sqrt{\lambda}(y-y^*))dy \\
&\qquad - \int_{-\infty}^{+\infty} dx \int_{y^*}^{\infty} \kappa y\frac{\partial}{\partial y}u(t,x,y)\lambda\chi_1(\lambda x)\sqrt{\lambda}\chi_2(\sqrt{\lambda}(y-y^*))dy.
\end{aligned}$$

Integrating by parts and doing the usual change of variables we have

$$\begin{aligned}
&\int_{-\infty}^{+\infty} dx \int_{y^*}^{\infty} \kappa\theta\frac{\partial}{\partial y}u(t,x,y)\lambda\chi_1(\lambda x)\sqrt{\lambda}\chi_2(\sqrt{\lambda}(y-y^*))dy \\
&\quad = -\sqrt{\lambda}\int_{-\infty}^{+\infty} dx \int_{0}^{\infty} \kappa\theta u\left(t, \frac{x}{\lambda}, \frac{y}{\sqrt{\lambda}} + y^*\right)\chi_1(x)\chi_2'(y)dy,
\end{aligned}$$

which tends to 0 as λ tends to 0, while

$$-\int_{-\infty}^{+\infty} dx \int_{y^*}^{\infty} \kappa y\frac{\partial}{\partial y}u(t,x,y)\lambda\chi_1(\lambda x)\sqrt{\lambda}\chi_2(\sqrt{\lambda}(y-y^*))dy,$$

which is nonpositive, since u is nondecreasing in y. We finally deduce that

$$\limsup_{\lambda\to 0}\int_{-\infty}^{+\infty} dx \int_{y^*}^{\infty} dy\mathcal{L}u(t,x,y)\phi(x)\zeta(y) \leq 0, \tag{2.4.16}$$

which, together with (2.4.15), contradicts (2.4.14). Then, the assertion follows. □

As regards the regularity of the free boundary, we can prove the following result.

Proposition 2.4.6. *For any $t \in [0,T)$ there exists a countable set $\mathcal{N} \subseteq (0,\infty)$ such that*

$$b(t^-, y) = b(t,y), \qquad y \in (0,\infty)\setminus\mathcal{N}.$$

Proof. Without loss of generality we pass to the logarithm in the $s-$variable and we prove the assertion for the function $\tilde{b}(t,y) = \ln b(t,y)$. Fix $t \in [0,T)$ and recall that $y \mapsto \tilde{b}(t,y)$ is a nonincreasing function, so it has at most a countable set of discontinuity points. Let $y^* \in (0,\infty)$ be a continuity point for the maps $y \mapsto \tilde{b}(t,y)$ and $y \mapsto \tilde{b}(t^-,y)$ and assume that

$$\tilde{b}(t^-, y^*) < \tilde{b}(t,y^*). \tag{2.4.17}$$

Set $\epsilon = \frac{\tilde{b}(t,y^*)-\tilde{b}(t^-,y^*)}{2}$. By continuity, there exist $y_0, y_1 > 0$ such that for any $y \in (y_0,y_1)$ we have

$$\tilde{b}(t,y) > \tilde{b}(t,y^*) - \frac{\epsilon}{4}, \qquad \text{and} \qquad \tilde{b}(t^-,y) < \tilde{b}(t^-,y^*) + \frac{\epsilon}{4}.$$

Therefore, by using (2.4.17), we get, for any $y \in (y_0, y_1)$,

$$\tilde{b}(t,y) > \tilde{b}(t,y^*) - \frac{\epsilon}{4} > \tilde{b}(t^-,y^*) + \frac{3}{4}\epsilon > \tilde{b}(t^-,y^*) + \frac{\epsilon}{4} > \tilde{b}(t^-,y).$$

Now, set $b^- = \tilde{b}(t^-,y^*) + \frac{\epsilon}{4}$ and $b^+ = \tilde{b}(t^-,y^*) + \frac{3}{4}$ and let $(s,x,y) \in (0,t) \times (b^-,b^+) \times (y_0,y_1)$. Since $t \mapsto \tilde{b}(t,\cdot)$ is nondecreasing, we have $x > \tilde{b}(t^-,y) > \tilde{b}(s,y)$, so that $u(s,x,y) > \psi(x)$. Therefore, on the set $(0,t) \times (b^-,b^+) \times (y_0,y_1)$ we have

$$(\tilde{\mathcal{L}} - r)u(s,x,y) = -\frac{\partial u}{\partial t}(s,x,y) \geq 0.$$

This means that, for any nonnegative and C^∞ test functions θ, ψ and ζ which have support respectively in $(0,t)$, (b^-,b^+) and (y_0,y_1) we can write

$$\int_0^t \theta(t)dt \int_{-\infty}^{\infty} dx \int_{y^*}^{\infty} dy (\tilde{\mathcal{L}} - r)u(t,x,y)\phi(x)\zeta(y) \geq 0.$$

By the continuity of the integrands in t, we deduce that $(\tilde{\mathcal{L}} - r)u(t,\cdot,\cdot) \geq 0$ in the sense of distributions on the set $(b^-,b^+) \times (y_0,y_1)$.

On the other hand, for any $(s,x,y) \in (t,T) \times (b^-,b^+) \times (y_0,y_1)$, we have $x \leq \tilde{b}(t,y) \leq \tilde{b}(s,y)$, so that $u(s,x,y) = \psi(x)$. Therefore, it follows from $\frac{\partial u}{\partial t} + (\tilde{\mathcal{L}} - r)u \leq 0$ and the continuity of the integrands that $(\tilde{\mathcal{L}} - r)u(t\cdot,\cdot) = (\tilde{\mathcal{L}} - r)\psi(\cdot) \leq 0$ in the sense of distributions on the set $(b^-,b^+) \times (y_0,y_1)$.

We deduce that $(\tilde{\mathcal{L}} - r)\psi = 0$ on the set $(b^-,b^+) \times (y_0,y_1)$, but it is easy to see that $(\tilde{\mathcal{L}} - r)\psi(x) = (\tilde{\mathcal{L}} - r)(K - e^x) = \delta e^x - rK$ and thus cannot be identically zero in a nonempty open set. □

Remark 2.4.7. *It is worth observing that the arguments used in [95] in order to prove the continuity of the exercise price of American options in a multidimensional Black and Scholes model can be easily adapted to our framework. In particular, if we consider the t-sections of the exercise region, that is*

$$\begin{aligned} \mathcal{E}_t &= \{(s,y) \in (0,\infty) \times [0,\infty) : P(t,s,y) = \varphi(s)\}, \\ &= \{(s,y) \in (0,\infty) \times [0,\infty) : s \leq b(t,y)\}, \qquad t \in [0,T), \end{aligned} \tag{2.4.18}$$

we can easily prove that

$$\mathcal{E}_t = \bigcap_{u>t} \mathcal{E}_u, \qquad \mathcal{E}_t = \overline{\bigcup_{u<t} \mathcal{E}_u}. \tag{2.4.19}$$

However, unlike the case of an American option on several assets, in our case (2.4.19) *is not sufficient to deduce the continuity of the function* $t \mapsto b(t,y)$.

2.4.2 Strict convexity in the continuation region

We know that P is convex in the space variable (see Proposition 2.4.1). In [93] it is also proved that, in the case of non-degenerate stochastic volatility models, P is strictly convex in the continuation region but the proof follows an analytical approach which cannot be applied in our degenerate model. In this section we extend this result to the Heston model by using purely probabilistic techniques.

We will need the following Lemma, whose proof can be found in the Appendix.

Lemma 2.4.8. *For every continuous function $s : [0,T] \to \mathbb{R}$ such that $s(0) = S_0$ and for every $\epsilon > 0$ we have*

$$\mathbb{P}\left(\sup_{t\in[0,T]} |S_t - s(t)| < \epsilon, \sup_{t\in[0,T]} |Y_t - Y_0| < \epsilon\right) > 0.$$

Theorem 2.4.9. *The function $s \mapsto P(t,s,y)$ is strictly convex in the continuation region.*

Proof. Without loss of generality we can assume $t = 0$. We have to prove that, if (s_1, y), $(s_2, y) \in (0,\infty) \times [0,\infty)$ are such that $(0, s_1, y), (0, s_2, y) \in \mathcal{C}$, then

$$P(0, \theta s_1 + (1-\theta)s_2, y) < \theta P(0, s_1, y) + (1-\theta)P(0, s_2, y). \tag{2.4.20}$$

Let us rewrite the price process as $S_t^{s,y} = s e^{\int_0^t \left(r-\delta-\frac{Y_u}{2}\right)du + \int_0^t \sigma\sqrt{Y_u}dB_u} := sM_t^y$, where $M_t^y = S_t^{1,y}$ and assume that, for example, $s_1 > s_2$. We claim that it is enough to prove that, for $\varepsilon > 0$ small enough,

$$\mathbb{P}\Big((\theta s_1 + (1-\theta)s_2)M_t^y > b(t, Y_t)\, \forall t \in [0,T) \,\&\, (\theta s_1 + (1-\theta)s_2)M_T^y \in (K-\varepsilon, K+\varepsilon)\Big) > 0. \tag{2.4.21}$$

In fact, let τ^* be the optimal stopping time for $P(0, \theta s_1 + (1-\theta)s_2, y)$. If $(\theta s_1 + (1-\theta)s_2)M_t^y > b(t, Y_t)$ for every $t \in [0,T)$, then we are in the continuation region for all $t \in [0,T)$, hence $\tau^* = T$. Then, the condition $(\theta s_1 + (1-\theta)s_2)M_T^y \in (K-\varepsilon, K+\varepsilon)$ for $\varepsilon > 0$ small enough ensures on one hand that $s_1 M_{\tau^*}^y > K$, since

$$\begin{aligned} s_1 M_{\tau^*}^y &= (\theta s_1 + (1-\theta)s_2)M_{\tau^*}^y + (1-\theta)(s_1 - s_2)M_{\tau^*}^y \\ &> K - \varepsilon + \frac{(1-\theta)(s_1-s_2)(K-\varepsilon)}{\theta s_1 + (1-\theta)s_2} > K, \end{aligned}$$

for ε small enough. On the other hand, it also ensures that $s_2 M_{\tau^*}^y < K$, which can be proved with similar arguments. Therefore, we get

$$\mathbb{P}\left((K - s_1 M_{\tau^*}^y)_+ = 0 \,\&\, (K - s_2 M_{\tau^*}^y)_+ > 0\right) > 0,$$

which, from a closer look at the graph of the function $x \mapsto (K-x)_+$, implies that

$$\mathbb{E}[e^{-r\tau^*}(K-(\theta s_1+(1-\theta)s_2)M^y_{\tau^*})_+] < \theta\mathbb{E}[e^{-r\tau^*}(K-s_1M^y_{\tau^*})_+] + (1-\theta)\mathbb{E}[e^{-r\tau^*}(K-s_2M^y_{\tau^*})_+],$$

and, as a consequence, (2.4.20).

So, the rest of the proof is devoted to prove that (2.4.21) is actually satisfied.

With this aim, we first consider a suitable continuous function $m : [0,T] \to \mathbb{R}$ constructed as follows. In order to simplify the notation, we set $s = \theta s_1 + (1-\theta)s_2$. Note that, for $\varepsilon > 0$ small enough, we have $s = \theta s_1 + (1-\theta)s_2 > b(0,y)+\varepsilon$ since $(0,s_1,y)$ and $(0,s_2,y)$ are in the continuation region $\mathcal{C}$, that is $s_1, s_2 \in (b(0,y),\infty)$. By the right continuity of the map $t \mapsto b(t,y)$, we know that there exists $\bar{t} \in (0,T)$ such that $s > b(t,y)+\frac{\varepsilon}{2}$ for any $t \in [0,\bar{t}]$. Moreover the function $y \mapsto b(\bar{t},y)$ is left continuous and nonincreasing, so there exists $\eta_\varepsilon > 0$ such that $s > b(\bar{t},z)+\frac{\varepsilon}{4}$ for any $z \geq y-\eta_\varepsilon$. Now, set

$$m(t) = \begin{cases} 1+\frac{t}{\bar{t}}\left(\frac{K+\frac{\varepsilon}{2}}{s}-1\right), & 0 \leq t \leq \bar{t}, \\ \frac{K+\frac{\varepsilon}{2}}{s}, & \bar{t} \leq t \leq T. \end{cases}$$

Note that m is continuous, $m(0)=1$ and, recalling that $t \mapsto b(t,y)$ is nondecreasing and $b(t,y) < K$,

$$sm(t) = \begin{cases} s+\frac{t}{\bar{t}}\left(K+\frac{\varepsilon}{2}-s\right) \geq s > b(\bar{t},y-\eta_\varepsilon)+\frac{\varepsilon}{4}, & 0 \leq t \leq \bar{t}, \\ K+\frac{\varepsilon}{2} \geq b(t,y-\eta_\varepsilon), & \bar{t} \leq t \leq T. \end{cases}$$

Moreover, by Lemma 2.4.8, we know that, for any $\epsilon > 0$,

$$\mathbb{P}\left(\sup_{t\in[0,T]} |sM^y_t - sm(t)| < \epsilon,\ \sup_{t\in[0,T]} |Y_t - y| < \epsilon\right) > 0.$$

Therefore, by applying Lemma 2.4.8 with $\epsilon = \min\left\{\frac{\varepsilon}{8}, \eta_\varepsilon\right\}$, we have that, with positive probability,

$$sM^y_t > sm(t) - \frac{\varepsilon}{8} \geq b(t,y-\eta_\varepsilon)+\frac{\varepsilon}{8} \geq b(t,Y_t).$$

and

$$sM^y_T \leq sm(T)+\frac{\varepsilon}{8} \leq K+\varepsilon, \qquad sM^y_T \geq sm(T)-\frac{\varepsilon}{8} \geq K-\varepsilon,$$

which proves (2.4.21) concluding the proof.

□

2.4.3 Early exercise premium

We now extend to the stochastic volatility Heston model a well known result in the Black and Scholes world, the so called *early exercise premium formula*. It is an explicit formulation of the quantity $P - P_e$, where $P_e = P_e(t,s,y)$ is the European put price with the same strike price K and maturity T of the American option with price function $P = P(t,s,y)$. Therefore, it represents the additional price you have to pay for the possibility of exercising before maturity.

Proposition 2.4.10. *Let $P_e(0,S_0,Y_0)$ be the European put price at time 0 with maturity T and strike price K. Then, one has*

$$P(0,S_0,Y_0) = P_e(0,S_0,Y_0) - \int_0^T e^{-rs}\mathbb{E}[(\delta S_s - rK)\mathbf{1}_{\{S_s \le b(s,Y_s)\}}]ds.$$

The proof of Proposition 2.4.10 relies on purely probabilistic techniques and is based on the results first introduced in [65]. Let $U_t = e^{-rt}P(t,S_t,Y_t)$ and $Z_t = e^{-rt}\varphi(S_t)$. Since U_t is a supermartingale, we have the Snell decomposition

$$U_t = M_t - A_t, \tag{2.4.22}$$

where M is a martingale and A is a nondecreasing predictable process with $A_0 = 0$, continuous with probability 1 thanks to the continuity of φ. On the other hand,

$$\begin{aligned} Z_t = e^{-rt}(K - S_t)_+ &= Z_0 - r\int_0^t e^{-rs}(K - S_s)_+ ds - \int_0^t e^{-rs}\mathbf{1}_{(-\infty,K]}dS_s + \int_0^t e^{-rs}dL_s^K(S) \\ &= m_t + a_t, \end{aligned}$$

where $L_t^K(S)$ is the local time of S in K,

$$m_t = Z_0 - \int_0^t e^{-rs}\mathbf{1}_{(-\infty,K]}S_s\sqrt{Y_s}dB_s$$

is a local martingale, and

$$a_t = -r\int_0^t e^{-rs}(K - S_s)_+ ds - \int_0^t e^{-rs}\mathbf{1}_{(-\infty,K]}S_s(r - \delta)ds + \int_0^t e^{-rs}dL_s^K(S)$$

is a predictable process with finite variation and $a_0 = 0$. Recall that a_t can be written as the sum of an increasing and a decreasing component, that is $a_t = a_t^+ + a_t^-$. Since $(L_t^K)_t$ is increasing, we deduce that the decreasing process $(a_t^-)_t$ is absolutely continuous with respect to the Lebesgue measure, that is

$$da_t^- \ll dt.$$

We denote by $k_t = k(t,S_t,Y_t)$ the density of a_t^- w.r.t. dt.

We now define

$$\zeta_t = U_t - Z_t \geq 0.$$

Thanks to Tanaka's formula,

$$\zeta_t = \zeta_t^+ = \zeta_0 + \int_0^t \mathbf{1}_{\{\zeta_s > 0\}} d\zeta_s + \frac{1}{2} L_t^0(\zeta),$$

where $L_t^0(\zeta)$ is the local time of ζ in 0. Therefore,

$$\begin{aligned}\zeta_t &= \zeta_0 + \int_0^t \mathbf{1}_{\{\zeta_s > 0\}} d(U_s - Z_s) + \frac{1}{2} L_t^0(\zeta) \\ &= \zeta_0 + \int_0^t \mathbf{1}_{\{\zeta_s > 0\}} dM_s - \int_0^t \mathbf{1}_{\{\zeta_s > 0\}} dm_s - \int_0^t \mathbf{1}_{\{\zeta_s > 0\}} da_s + \frac{1}{2} L_t^0(\zeta),\end{aligned}$$

where the last equality follows from the fact that the process A_t only increases on the set $\{\zeta_t = 0\}$. Then, we can write

$$U_t = U_0 + \bar{M}_t - \int_0^t \mathbf{1}_{\{\zeta_s > 0\}} da_s + \frac{1}{2} L_t^0(\zeta) + a_t = U_0 + \bar{M}_t + \int_0^t \mathbf{1}_{\{\zeta_s = 0\}} da_s + \frac{1}{2} L_t^0(\zeta),$$

where $\bar{M}_t = \int_0^t \mathbf{1}_{\{\zeta_s > 0\}} d(M_s - m_s) + m_t$ is a local martingale. Thanks to the continuity of U_t we have the uniqueness of the decompositions, so

$$-A_t = \int_0^t \mathbf{1}_{\{\zeta_s = 0\}} da_s + \frac{1}{2} L_t^0(\zeta). \tag{2.4.23}$$

This means in particular that $\int_0^t \mathbf{1}_{\{\zeta_s = 0\}} da_s + \frac{1}{2} L_t^0(\zeta)$ is decreasing, but $L_t^0(\zeta)$ is increasing so $-\int_0^t \mathbf{1}_{\{\zeta_s = 0\}} da_s$ must be an increasing process and

$$\frac{1}{2} dL_t^0(\zeta) \ll \mathbf{1}_{\{\zeta_t = 0\}} da_t^- \ll dt.$$

We define μ_t the density of $\frac{1}{2} L_t^0(\zeta)$ w.r.t. a_t^- and, by Motoo Theorem (see [41]), we can write $\mu_t = \mu(t, S_t, Y_t)$. Moreover, let us consider the t-sections of the exercise region defined in (2.4.18). We can easily prove the following Lemma.

Lemma 2.4.11. *For any $t \in [0, T)$ we have*

$$\mathcal{E}_t = \overline{\mathring{\mathcal{E}}_t},$$

and $\mathring{\mathcal{E}}_t = \{(s, y) \in (0, \infty) \times [0, \infty) : 0 < s < b(t, y^+)\} \neq \emptyset$, *where* $b(t, y^+) = \lim_{y \to y^+} b(t, y)$.

The proof is given in the Appendix for the sake of completeness. Now, let us prove the following preliminary result.

Lemma 2.4.12. *The local time $L^0_t(\zeta)$ is indistinguishable from 0.*

Proof. First of all, note that $L^0_t(\zeta)$ only increases on the set $\{(t,S_t,Y_t)\in\partial\mathcal{E}\}$. In fact, recall that $L^a_t=\int_0^t \mathbf{1}_{\{U_s-Z_s=a\}}dL^a_s$ for every $a>0$ and $t>0$, so that

$$\int_0^t \mathbf{1}_{\{(s,S_s,Y_s)\in\mathring{\mathcal{E}}\}}dL^a_s=0.$$

Moreover it is well known that, for any $t>0$, $L^0_t=\lim_{a\to 0}L^a_t$, which implies that the measures L^a_t weakly converge to L^0_t as $a\to 0$. Then, we can deduce that

$$L^0_t(\{(s,S_s,Y_s)\in\mathring{\mathcal{E}}\})\le\liminf L^a_t(\{(s,S_s,Y_s)\in\mathring{\mathcal{E}}\})=0.$$

Moreover, thanks to Lemma 2.4.11, we have

$$\partial\mathcal{E}_t=\{(t,b(t,y),y):(t,y)\in[0,T)\times[0,\infty)\}\cup\bigcup_{y\in\mathcal{D}_t}\{(s,y):b(t,y)<s\le b(t,y^+)\},$$

where $\mathcal{D}_t$ is the (numerable) set of the discontinuity points of $y\mapsto b(t,y)$ and the union is disjoint. Thanks to the continuity of P and φ, it is easy to show that $Leb\{(t,b(t,y),y):(t,y)\in[0,T)\times[0,\infty)\}=0$, so that $Leb\,\partial\mathcal{E}_t=0$ for any $t\in[0,T]$. Therefore,

$$\begin{aligned}\mathbb{E}[L^0_t(\zeta)]&=\mathbb{E}[\int_0^t \mathbf{1}_{\{U_s-Z_s=0\}}dL^0_s]=\mathbb{E}[\int_0^t \mathbf{1}_{\{(s,S_s,Y_s)\in\partial\mathcal{E}\}}\mu(s,S_s,Y_s)k(s,S_s,Y_s)ds]\\&=\int_0^t ds\int_{\partial\mathcal{E}_s}dxdy\;\mu(s,x,y)k(s,x,y)p(s,x,y)=0.\end{aligned}$$

□

We can now prove Proposition 2.4.10.

Proof of Proposition 2.4.10. Thanks to (2.4.23) and Proposition 2.4.12 we can rewrite (2.4.22) as

$$U_t=M_t+\int_0^t \mathbf{1}_{\{U_s=Z_s\}}da_s=M_t+\int_0^t e^{-rs}(\mathcal{L}-r)\varphi(S_s)\mathbf{1}_{\{S_s\le b(s,Y_s)\}}ds,$$

where the last equality derives from the application of the Itô formula to the discounted payoff Z. In particular, we have

$$\begin{aligned}U_0=M_0=\mathbb{E}[M_T]&=\mathbb{E}[U_T]-\mathbb{E}\left[\int_0^T e^{-rs}(\mathcal{L}-r)\varphi(S_s)\mathbf{1}_{\{S_s\le b(s,Y_s)\}}ds\right]\\&=\mathbb{E}[U_T]-\int_0^T e^{-rs}\mathbb{E}[(\delta S_s-rK)\mathbf{1}_{\{S_s\le b(s,Y_s)\}}]ds.\end{aligned}$$

The assertion follows recalling that $U_0 = P(0, S_0, Y_0)$ and $\mathbb{E}[U_T] = \mathbb{E}[Z_T] = \mathbb{E}[e^{-rT}(K - S_T)_+]$, which corresponds to the price $P_e(0, S_0, Y_0)$ of an European put with maturity T and strike price K. □

2.4.4 Smooth fit

In this section we analyse the behaviour of the derivatives of the value function with respect to the s and y variables on the boundary of the continuation region. In other words, we prove a weak formulation of the so called smooth fit principle.

In order to do this, we need two technical lemmas whose proofs can be found in the appendix. The first one is a general result about the behaviour of the trajectories of the CIR process.

Lemma 2.4.13. *For all $y \geq 0$ we have, with probability one,*

$$\limsup_{t \downarrow 0} \frac{Y_t^y - y}{\sqrt{2t \ln \ln(1/t)}} = -\liminf_{t \downarrow 0} \frac{Y_t^y - y}{\sqrt{2t \ln \ln(1/t)}} = \sigma \sqrt{y}.$$

The second one is a result about the behaviour of the trajectories of a standard Brownian motion.

Lemma 2.4.14. *Let $(B_t)_{t \geq 0}$ be a standard Brownian motion and let $(t_n)_{n \in \mathbb{N}}$ be a deterministic sequence of positive numbers with $\lim_{n \to \infty} t_n = 0$. We have, with probability one,*

$$\liminf_{n \to \infty} \frac{B_{t_n}}{\sqrt{t_n}} = -\infty \tag{2.4.24}$$

We are now in a position to prove the following smooth fit result.

Proposition 2.4.15. *For any $(t, y) \in [0, T) \times [0, \infty)$ we have $\frac{\partial}{\partial s} P(t, b(t, y), y) = \varphi'(b(t, y))$.*

Proof. The general idea of the proof goes back to [18] for the Brownian motion (see also [83, Chapter 4]). Without loss of generality we can fix $t = 0$. Note that, for $h > 0$, since $b(0, y) - h \leq b(0, y)$, we have

$$\frac{P(0, b(0, y) - h, y) - P(0, b(0, y), y)}{h} = \frac{\varphi(b(0, y) - h) - \varphi(b(0, y))}{h},$$

so that, since φ is continuously differentiable near $b(0, y)$, $\frac{\partial^-}{\partial s} P(0, b(0, y), y) = \varphi'(b(0, y))$.

On the other hand, for $h > 0$ small enough, since $P \geq \varphi$ and $P(0, b(0, y), y) = \varphi(b(0, y))$, we get

$$\frac{P(0, b(0, y) + h, y) - P(0, b(0, y), y)}{h} \geq \frac{\varphi(b(0, y) + h) - \varphi(b(0, y))}{h},$$

so that

$$\liminf_{h \downarrow 0} \frac{P(0, b(0, y) + h, y) - P(0, b(0, y), y)}{h} \geq \varphi'(b(0, y)).$$

Now, for the other inequality, we consider the optimal stopping time related to $P(0, b(0, y) + h, y)$, i.e.

$$\tau_h = \inf\{t \in [0,T) \mid S_t^{0,b(0,y)+h,y} < b(t, Y_t^y)\} \wedge T = \inf\left\{t \in [0,T) \mid M_t^y \leq \frac{b(t, Y_t^y)}{b(0,y)+h}\right\} \wedge T,$$

where $M_t^y = S_t^{1,y}$. Recall that $P(0, b(0, y), y) \geq \mathbb{E}\big(e^{-r\tau_h}\varphi(b(0, y)M_{\tau_h}^y)\big)$, so we can write

$$\begin{aligned}\frac{P(0, b(0,y)+h, y) - P(0, b(0,y), y)}{h} &= \frac{\mathbb{E}\,(e^{-r\tau_h}\varphi((b(0,y)+h)M_{\tau_h}^y) - P(0, b(0,y), y)}{h} \\ &\leq \mathbb{E}\left(e^{-r\tau_h}\frac{\varphi\left((b(0,y)+h)M_{\tau_h}^y\right) - \varphi\left(b(0,y)M_{\tau_h}^y\right)}{h}\right).\end{aligned}$$

Assume for the moment that

$$\lim_{h\to 0}\tau_h = 0, \quad a.s. \tag{2.4.25}$$

so we have

$$\lim_{h\downarrow 0}\frac{\varphi((b(0,y)+h)M_{\tau_h}^y)\big) - \varphi(b(0,y)M_{\tau_h}^y)}{h} = \varphi'(b(0,y)).$$

Moreover, recall that $M_{\tau_h}^y \leq \frac{b(t,Y_t^y)}{b(0,y)+h} \leq \frac{K}{b(0,y)}$ if $\tau_h < T$ and $M_{\tau_h}^y = M_T^y$ if $\tau_h = T$. Therefore, by using the fact that φ is Lipschitz continuous and the dominated convergence, we obtain

$$\limsup_{h\downarrow 0}\frac{P(0, b(0,y)+h, y) - P(0, b(0,y), y)}{h} \leq \varphi'(b(0,y))$$

and the assertion is proved.

It remains to prove (2.4.25). Since $t \mapsto b(t, y)$ is nondecreasing, if $M_t^y < \frac{b(0,y)}{b(0,y)+h}$ and $Y_t^y = y$, we have

$$M_t^y < \frac{b(0,y)}{b(0,y)+h} \leq \frac{b(t, Y_t^y)}{b(0,y)+h},$$

so that

$$\tau_h \leq \inf\left\{t \geq 0 \mid M_t^y < \frac{b(0,y)}{b(0,y)+h} \;\&\; Y_t^y = y\right\}. \tag{2.4.26}$$

We now show that we can find a sequence $t_n \downarrow 0$ such that $Y_{t_n}^y = 0$ and $M_{t_n}^y < 1$. First, recall that with a standard transformation we can write

$$\begin{cases}\frac{dS_t}{S_t} = (r-\delta)dt + \sqrt{Y_t}(\sqrt{1-\rho^2}d\bar{W}_t + \rho dW_t), & S_0 = s > 0,\\ dY_t = \kappa(\theta - Y_t)dt + \sigma\sqrt{Y_t}dW_t, & Y_0 = y \geq 0,\end{cases} \tag{2.4.27}$$

where $\bar{W}$ is a standard Brownian motion independent of W. Set $\Lambda_t^y = \ln M_t^y$. We deduce from Lemma 2.4.13 that there exists a sequence $t_n \downarrow 0$ such that $Y_{t_n}^y = y$ $\mathbb{P}_y$-a.s. . Therefore, from (2.4.27) we can write $\int_0^{t_n}\sqrt{Y_s^y}dW_s = -\frac{\kappa}{\sigma}\int_0^{t_n}(\theta - Y_s^y)ds$ for all $n \in \mathbb{N}$. So, we have

$$\Lambda_{t_n}^y = (r-\delta)t_n - \int_0^{t_n}\frac{Y_s^y}{2}ds + \sqrt{1-\rho^2}\int_0^{t_n}\sqrt{Y_s^y}d\bar{W}_s - \frac{\rho\kappa}{\sigma}\int_0^{t_n}(\theta - Y_s^y)ds.$$

Conditioning with respect to W we have

$$\begin{aligned}\liminf_{n\to\infty}\Lambda^y_{t_n}&=\liminf_{n\to\infty}\frac{(r-\delta)t_n}{\sqrt{\int_0^{t_n}Y^y_s ds}}-\frac{\int_0^{t_n}\frac{Y^y_s}{2}ds}{\sqrt{\int_0^{t_n}Y^y_s ds}}+\frac{\sqrt{1-\rho^2}\int_0^{t_n}\sqrt{Y^y_s}d\bar W_s}{\sqrt{\int_0^{t_n}Y^y_s ds}}-\frac{\frac{\rho\kappa}{\sigma}\int_0^{t_n}(\theta-Y^y_s)ds}{\sqrt{\int_0^{t_n}Y^y_s ds}}\\&=\liminf_{n\to\infty}\frac{(r-\delta)t_n}{\sqrt{\int_0^{t_n}Y^y_s ds}}-\frac{\int_0^{t_n}\frac{Y^y_s}{2}ds}{\sqrt{\int_0^{t_n}Y^y_s ds}}+\frac{\sqrt{1-\rho^2}\tilde W_{\int_0^{t_n}Y^y_s ds}}{\sqrt{\int_0^{t_n}Y^y_s ds}}-\frac{\frac{\rho\kappa}{\sigma}\int_0^{t_n}(\theta-Y^y_s)ds}{\sqrt{\int_0^{t_n}Y^y_s ds}}=-\infty,\end{aligned}$$

where we have used the Dubins-Schwartz Theorem and we have applied Lemma 2.4.14 to the standard Brownian motion $\tilde W$ and the sequence $\sqrt{\int_0^{t_n}Y^y_s ds}$ which can be considered deterministic.

We deduce that, up to extract a subsequence of t_n, we have $\Lambda^y_{t_n}<0$ and, as a consequence, $M^y_{t_n}<1$. Therefore, for any any fixed n, there exists h small enough such that $M^y_{t_n}<\frac{b(0,y)}{b(0,y)+h}$ so that, by definition, $\tau_h\le t^n$. We conclude the proof passing to the limit as n goes to infinity.

□

As regards the derivative with respect to the y variable, we have the following result.

Proposition 2.4.16. *If $2\kappa\theta\ge\sigma^2$, for any $(t,y)\in[0,T)\times(0,\infty)$ we have $\frac{\partial}{\partial y}P(t,b(t,y),y)=0$.*

Proof. Again we fix $t=0$ with no loss of generality. Since $y\to P(t,s,y)$ in nondecreasing, for any $h>0$ we have $P(0,b(0,y),y-h)\le P(0,b(0,y),y)=\varphi(b(0,y))$ so that $P(0,b(0,y),y-h)=\varphi(b(0,y))$. Therefore,

$$\frac{P(0,b(0,y),y-h)-P(0,b(0,y),y)}{h}=0,$$

hence $\frac{\partial^-}{\partial y}P(0,b(0,y),y)=0$. On the other hand, since $y\mapsto P(t,x,y)$ is nondecreasing, for any $h>0$ we have

$$\liminf_{h\downarrow 0}\frac{P(0,b(0,y),y+h)-P(0,b(0,y),y)}{h}\ge 0,$$

To prove the other inequality, we consider the stopping time related to $P(0,b(0,y),y+h)$, that is

$$\tau_h=\inf\left\{t\in[0,T)\mid S^{0,b(0,y),y+h}_t<b(t,Y^{y+h}_t)\right\}\wedge T=\inf\left\{t\in[0,T)\mid M^{y+h}_t<\frac{b(t,Y^{y+h}_t)}{b(0,y)}\right\}\wedge T$$

and we assume for the moment that

$$\lim_{h\to 0}\tau_h=0. \tag{2.4.28}$$

We have

$$\begin{aligned}\frac{P(0,b(0,y),y+h)-P(0,b(0,y),y)}{h} &= \frac{\mathbb{E}\left(e^{-r\tau_h}\varphi\left(b(0,y)M^{y+h}_{\tau_h}\right)\right)-P(0,b(0,y),y)}{h}\\ &\le \mathbb{E}\left[e^{-r\tau_h}\frac{\varphi\left(b(0,y)M^{y+h}_{\tau_h}\right)-\varphi(b(0,y)M^{y}_{\tau_h})}{h}\right]\\ &\le K\frac{\mathbb{E}\left[\left|M^{y+h}_{\tau_h}-M^{y}_{\tau_h}\right|\right]}{h},\end{aligned} \tag{2.4.29}$$

where the last inequality follows from the fact that φ is Lipschitz continuous and $b(0,y)\le K$.

Now, if the Feller condition $2\kappa\theta\ge\sigma^2$ is satisfied, we can write

$$M^{y+h}_t-M^y_t=\int_y^{y+h}\left(\int_0^t\frac{\dot{Y}^\zeta_s}{2\sqrt{Y^\zeta_s}}dB_s-\frac{1}{2}\int_0^t\dot{Y}^\zeta_s ds\right)e^{(r-\delta)t-\int_0^t\frac{Y^\zeta_s}{2}ds+\int_0^t\sqrt{Y^\zeta_s}dB_s}d\zeta.$$

The exponential process $e^{-\int_0^t\frac{Y^\zeta_s}{2}ds+\int_0^t\sqrt{Y^\zeta_s}dB_s}$ satisfies the assumptions of the Girsanov Theorem, so we can introduce a new probability measure $\hat{\mathbb{P}}$ under which the process $\hat{W}_t=W_t-\int_0^t\sqrt{Y_s}ds$ is a standard Brownian motion. If we denote by $\hat{\mathbb{E}}$ the expectation under the probability $\hat{\mathbb{P}}$, substituting in (2.4.29) and using (2.4.13) we get

$$\frac{P(0,b(0,y),y+h)-P(0,b(0,y),y)}{h}\le\frac{e^{(r-\delta)T}K}{h}\int_y^{y+h}d\zeta\hat{\mathbb{E}}\left[\left|\int_0^{\tau_h}\frac{\dot{Y}^\zeta_s}{2\sqrt{Y^\zeta_s}}d\hat{W}_s\right|\right]$$

$$\le\frac{e^{(r-\delta)T}K}{h}\int_y^{y+h}d\zeta\left(\hat{\mathbb{E}}\left[\int_0^{\tau_h}\left(\frac{\dot{Y}^\zeta_s}{2\sqrt{Y^\zeta_s}}\right)^2ds\right]\right)^{1/2}\le\frac{e^{(r-\delta)T}K}{h}\int_y^{y+h}\frac{1}{2\sqrt{\zeta}}\hat{\mathbb{E}}[\sqrt{\tau_h}]d\zeta$$

which tends to 0 as h tends to 0.

Therefore, as in the proof of Proposition 2.4.15, it remains to prove that $\lim_{h\downarrow0}\tau_h=0$. In order to do this, we can proceed as follows. Again, set

$$\Lambda^y_t=\ln(M^y_t)=(r-\delta)t-\frac{1}{2}\int_0^tY^y_sds+\int_0^t\sqrt{Y^y_s}dW_s,$$

so that

$$\tau_h=\inf\left\{t\in[0,T)\mid\Lambda^{y+h}_t\le\ln\left(\frac{b(t,Y^{y+h}_t)}{b(0,y)}\right)\right\}\wedge T.$$

We deduce from Lemma (2.4.13) that, almost surely, there exist two sequences $(t_n)_n$ and $(\hat{t}_n)_n$ which converge to 0 with $0 < t_n < \hat{t}_n$ and such that

$$Y^y_{t_n} = y, \qquad \text{and, for } t \in (t_n, \hat{t}_n), \quad Y_t < y.$$

In fact, it is enough to consider a sequence $(\hat{t}_n)_n$ such that $\lim_{h\downarrow 0} \hat{t}_n = 0$ and $Y_{t_n} < y$ and define $t_n = \sup\{t \in [0, \hat{t}_n) \mid Y^y_t = y\}$.

Proceeding as in the proof of Proposition 2.4.15, up to extract a subsequence we can assume

$$\Lambda^y_{t_n} < 0.$$

On the other hand, up to extract a subsequence of h converging to 0, we can assume that, almost surely,

$$\lim_{h\downarrow 0} \sup_{t\in[0,T]} \left| Y^{y+h}_t - Y^y_t \right| = \lim_{h\downarrow 0} \sup_{t\in[0,T]} \left| \Lambda^{y+h}_t - \Lambda^y_t \right| = 0.$$

Now, let us fix $n \in \mathbb{N}$. For h small enough, there exists $\delta > 0$ such that

$$\Lambda^{y+h}_t < 0, \qquad t \in (t_n - \delta, t_n + \delta).$$

Then, for any $\tilde{t}_n \in (t_n - \delta, t_n + \delta) \cap (t_n, \hat{t}_n)$, we have at the same time $\Lambda^{y+h}_{\tilde{t}_n} < 0$ and, since $Y^y_{\tilde{t}_n} < y$, $Y^{y+h}_{\tilde{t}_n} < y$ for h small enough. Recalling that $t \mapsto b(t, y)$ is nondecreasing and $y \mapsto b(t, y)$ is nonincreasing, we deduce that

$$b(\tilde{t}_n, Y^{y+h}_{\tilde{t}_n}) \geq b(0, Y^{y+h}_{\tilde{t}_n}) \geq b(0, y).$$

Therefore

$$\Lambda^{y+h}_{\tilde{t}_n} \leq \ln \left(\frac{b(\tilde{t}_n, Y^{y+h}_{\tilde{t}_n})}{b(0, y)} \right)$$

and, as a consequence, $\tau_h \leq \tilde{t}_n \leq \hat{t}_n$ so (2.4.28) follows.

□

2.5 Appendix: some proofs

We devote the appendix to the proof of some technical results used in this chapter.

2.5.1 Proofs of Section 2.3

Proof of Lemma 2.3.2. Consider $1 > a_1 > a_2 > \cdots > a_m > \cdots > 0$ defined by

$$\int_{a_1}^{1} \frac{1}{u} du = 1, \ldots, \int_{a_m}^{a_{m-1}} \frac{1}{u} du = m, \ \ldots .$$

We have that a_m tends to 0 as m tends to infinity. Let $(\eta_m)_{m\geq 1}$, be a family of continuous functions such that

$$\operatorname{supp} \eta_m \subseteq (a_m, a_{m-1}), \quad 0 \leq \eta_m(u) \leq \frac{2}{um}, \quad \int_{a_m}^{a_{m-1}} \eta_m(u) du = 1.$$

Moreover, we set

$$\phi_m(x) := \int_0^{|x|} dy \int_0^y \eta_m(u) du, \qquad x \in \mathbb{R}.$$

It is easy to see that $\phi_m \in C^2(\mathbb{R})$, $|\phi_m'| \leq 1$ and $\phi_m(x) \uparrow |x|$ as $m \to \infty$. Fix $t \in [0, T]$. Applying Itô's formula and passing to the expectation we have, for any $m \in \mathbb{N}$,

$$\begin{aligned}\mathbb{E}[\phi_m(Y_t^n - Y_t)] = \kappa \int_0^t \mathbb{E}\left[\phi_m'(Y_s^n - Y_s)(Y_s - f_n^2(Y_s^n))\right] ds \\ + \frac{\sigma^2}{2} \int_0^t \mathbb{E}\left[\phi_m''(Y_s^n - Y_s)(f_n(Y_s^n) - \sqrt{Y_s})^2\right] ds\end{aligned} \tag{2.5.30}$$

Let us analyse the right hand term in (2.5.30). Since $|\phi_m'| \leq 1$, we have

$$\left|\kappa \int_0^t \mathbb{E}\left[\phi_m'(Y_s^n - Y_s)(Y_s - f_n^2(Y_s^n))\right] ds\right| \leq \kappa \int_0^t \mathbb{E}\left[|f_n^2(Y_s^n) - Y_s^n|\right] ds + \kappa \int_0^t \mathbb{E}\left[|Y_s^n - Y_s|\right] ds$$

On the other hand,

$$\begin{aligned}&\left|\frac{\sigma^2}{2} \int_0^t \mathbb{E}\left[\phi_m''(Y_s^n - Y_s)(f_n(Y_s^n) - \sqrt{Y_s})^2\right] ds\right| \\ &\leq \sigma^2 \int_0^t \mathbb{E}\left[|\phi_m''(Y_s^n - Y_s)|(f_n(Y_s^n) - \sqrt{Y_s^n})^2] ds\right] + \sigma^2 \int_0^t \mathbb{E}\left[|\phi_m''(Y_s^n - Y_s)|(\sqrt{Y_s^n} - \sqrt{Y_s})^2\right] ds \\ &\leq \sigma^2 \int_0^t \mathbb{E}\left[\frac{2}{m|Y_s^n - Y_s|}(f_n(Y_s^n) - \sqrt{Y_s^n})^2 \mathbf{1}_{\{a_m \leq Y_s^n - Y_s \leq a_{m-1}\}}] ds\right] \\ &\quad + \sigma^2 \int_0^t \mathbb{E}\left[\frac{2}{m|Y_s^n - Y_s|}|Y_s^n - Y_s|\right] ds \\ &\leq \frac{2\sigma^2}{ma_m} \int_0^t \mathbb{E}\left[(f_n(Y_s^n) - \sqrt{Y_s^n})^2] ds\right] + \frac{2\sigma^2 t}{m}.\end{aligned}$$

Observe that, if $|x| \geq a_{m-1}$,

$$\phi_m(x) \geq \int_{a_{m-1}}^{|x|} dy = |x| - a_{m-1}.$$

Therefore, for any m large enough,

$$\mathbb{E}[|Y_t^n - Y_t|] \le \kappa \int_0^t \mathbb{E}[|Y_s^n - Y_s|]ds + \kappa \int_0^t \mathbb{E}\left[|f_n^2(Y_s^n) - Y_s^n|\right] ds$$
$$+ \frac{2\sigma^2}{m a_m} \int_0^t \mathbb{E}\left[(f_n(Y_s^n) - \sqrt{Y_s^n})^2]ds\right] + \frac{2\sigma^2 t}{m} + a_{m-1}.$$

Recall that $f_n(y) \to f(y) \equiv y$ locally uniformly and that Y^n has continuous paths. Moreover, since $f_n^2(x) \le A(|x|+1)$ with A independent of n, it is easily to see that for any $p > 1$ there exists $C > 0$ independent of n such that

$$\mathbb{E}\left[\sup_{t\in[0,T]} |Y_t^n|^p\right] \le C. \tag{2.5.31}$$

Therefore, by using Lebesgue's Theorem and recalling that $\lim_{m\to\infty} a_m = 0$, we deduce that for any $\delta > 0$ it is possible to choose $\bar{n}$ such that for every $n \ge \bar{n}$

$$\mathbb{E}[|Y_t^n - Y_t|] < C \int_0^t \mathbb{E}[|Y_s^n - Y_s|] + \delta.$$

We can now apply Gronwall's inequality and we deduce that $\mathbb{E}[|Y_t^n - Y_t|] < \delta e^{Ct}$, so that

$$\lim_{n\to\infty} \mathbb{E}[|Y_t^n - Y_t|] = 0 \tag{2.5.32}$$

from the arbitrariness of δ.

Now, note that

$$\sup_{t\in[0,T]} |Y_t^n - Y_t| \le \kappa \int_0^T |Y_s - Y_s^n| ds + \sup_{t\in[0,T]} \left| \int_0^t (\sqrt{Y_s} - f_n(Y_s^n)) dW_s \right| \tag{2.5.33}$$

The first term in the right hand side of (2.5.33) converges to 0 in probability thanks to (2.5.32), so it is enough to prove that the second term converges to 0. We have

$$\mathbb{E}\left[\sup_{t\in[0,T]} \left| \int_0^t (\sqrt{Y_s} - f_n(Y_s^n)) dW_s \right| \right] \le \left(\int_0^T \mathbb{E}[|\sqrt{Y_s} - f_n(Y_s^n)|^2] ds \right)^{\frac{1}{2}} \tag{2.5.34}$$

and

$$\mathbb{E}\left[|\sqrt{Y_s} - f_n(Y_s^n)|^2\right] \le 2\mathbb{E}\left[|\sqrt{Y_s} - \sqrt{Y_s^n}|^2\right] + 2\mathbb{E}\left[|\sqrt{Y_s^n} - f_n(Y_s^n)|^2\right]$$
$$\le 2\mathbb{E}\left[|Y_s - Y_s^n|\right] + 2\mathbb{E}\left[|\sqrt{Y_s^n} - f_n(Y_s^n)|^2\right].$$

Therefore, we can conclude that (2.5.34) tends to 0 as n goes to infinity by using (2.5.32) and the Lebesgue Theorem so that (2.5.37) is proved.

As regards (2.3.6), for every $n \in \mathbb{N}$ we have

$$X_t^n = x + \int_0^t \left(r - \delta - \frac{f_n^2(Y_s^n)}{2} \right) ds + \int_0^t f_n(Y_s^n) dB_s,$$

so that

$$\sup_{t\in[0,T]} |X_t^n - X_t| \le \frac{1}{2} \int_0^T |f_n^2(Y_s^n) - Y_s| ds + \sup_{t\in[0,T]} \left| \int_0^t (f_n(Y_s^n) - \sqrt{Y_s}) dB_s \right|. \tag{2.5.35}$$

It is enough to show that the two terms in the right hand side of (2.5.35) converge to 0 in probability.

Concerning the first term, note that, since Y has continuous paths, for every $\omega \in \Omega$, $Y_{[0,T]}(\omega)$ is a compact set and $K := \{x | d(x, Y_{[0,T]}) \le 1\}$ is compact as well. For n large enough, Y^n lies in K, so

$$\int_0^T |f_n^2(Y_s^n) - f^2(Y_s)| ds \le \int_0^T |f_n^2(Y_s^n) - f^2(Y_s^n)| ds + \int_0^T |f^2(Y_s^n) - f^2(Y_s)| ds,$$

which goes to 0 as n tends to infinity, since $f_n^2 \to f^2$ locally uniformly and f^2 is a continuous function.

On the other hand, for the second term in the right hand side of (2.5.35), we have

$$\mathbb{E}\left[\sup_{t\in[0,T]} \left| \int_0^t f(Y_s^n) - \sqrt{Y_s} dW_s \right| \right] \le \left(\int_0^T \mathbb{E}[(f(Y_s^n) - \sqrt{Y_s})^2] ds \right)^{\frac{1}{2}}$$

and we can prove with the usual arguments that the last term goes to 0.

□

2.5.2 Proofs of Section 2.4

Proofs of Lemma 2.4.8. To simplify the notation we pass to the logarithm and we prove the assertion for the pair (X, Y). We can get rid of the correlation between the Brownian motions with a standard transformation, getting

$$\begin{cases} dX_t = (r - \delta - \frac{1}{2} Y_t) dt + \sqrt{Y_t} (\sqrt{1-\rho^2} d\bar{W}_t + \rho dW_t), & X_0 \in \mathbb{R}, \\ dY_t = \kappa(\theta - Y_t) dt + \sigma \sqrt{Y_t} dW_t, & Y_0 \ge 0, \end{cases}$$

where $\bar{W}$ is a standard Brownian motion independent of W. Moreover, from the SDE satisfied by Y we deduce $\int_0^t \sqrt{Y_s} dW_s = \frac{1}{\sigma} \left(Y_t - Y_0 - \int_0^t \kappa(\theta - Y_s) ds \right)$. Conditioning with respect to Y, we reduce to prove that, for every continuous function $m : [0,T] \to \mathbb{R}$ such that $m(0) = X_0$ and for every $\epsilon > 0$ we have

$$\mathbb{P}\left(\sup_{t\in[0,T]} |X_t - m(t)| < \epsilon \mid Y \right) > 0, \tag{2.5.36}$$

and

$$\mathbb{P}\left(\sup_{t\in[0,T]}|Y_t - Y_0| < \epsilon\right) > 0. \tag{2.5.37}$$

As regards (2.5.36), by using the Dubins-Schwartz Theorem, there exists a Brownian motion $\tilde{W}$ such that

$$\mathbb{P}\left(\sup_{t\in[0,T]}\left|x + \int_0^t \left(r - \delta - \frac{Y_s}{2} - \frac{\rho\kappa}{\sigma}(\theta - Y_s)\right)ds + \frac{\rho}{\sigma}(Y_t - y) + \sqrt{1-\rho^2}\int_0^t \sqrt{Y_s}d\bar{W}_s - m(t)\right| < \epsilon \mid Y\right)$$
$$= \mathbb{P}\left(\sup_{t\in[0,T]}\left|\sqrt{1-\rho^2}\int_0^t \sqrt{Y_s}d\bar{W}_s - \tilde{m}(t)\right| < \epsilon \mid Y\right)$$
$$= \mathbb{P}\left(\sup_{t\in[0,T]}\left|\sqrt{1-\rho^2}\tilde{W}_{\int_0^t Y_s ds} - \tilde{m}(t)\right| < \epsilon \mid Y\right),$$

where $\tilde{m}(t) = m(t) - x - \int_0^t \left(r - \delta - \frac{Y_s}{2} - \frac{\rho\kappa}{\sigma}(\theta - Y_s)\right) ds - \frac{\rho}{\sigma}(Y_t - y)$ is a continuous function which, conditioning w.r.t. Y, can be considered deterministic. Then, (2.5.36) follows by the support theorem for Brownian motions.

In order to prove (2.5.37), we distinguish two cases. Assume first that $Y_0 = y_0 > 0$ and, for $a \geq 0$, define the stopping time

$$T_a = \inf\{t > 0 \mid Y_t = a\}.$$

Moreover, let us consider the function

$$\eta(y) = \begin{cases} \sqrt{y}, & \text{if } y > \frac{y_0}{2}, \\ \frac{\sqrt{y_0}}{2} & \text{if } y \leq \frac{y_0}{2}, \end{cases}$$

and the process $(\tilde{Y}_t)_{t\in[0,T]}$, solution to the uniformly elliptic SDE

$$d\tilde{Y}_t = \kappa(\theta - \tilde{Y}_t)dt + \sigma\eta(\tilde{Y}_t)dW_t, \qquad \tilde{Y}_0 = Y_0.$$

It is clear that $Y_t = \tilde{Y}_t$ on the set $\left\{t \leq T_{\frac{y_0}{2}}\right\}$ so we have, if $\epsilon < \frac{y_0}{2}$,

$$\mathbb{P}\left(\sup_{t\in[0,T]}|Y_t - Y_0| < \epsilon\right) = \mathbb{P}\left(\sup_{t\in[0,T]}|\tilde{Y}_t - Y_0| < \epsilon\right),$$

where the last inequality follows from the classical Support Theorem for uniformly elliptic diffusions (see, for example, [88]).

On the other hand, if we assume $Y_0 = 0$, then we can write

$$\mathbb{P}\left(\sup_{t\in[0,T]} Y_t < \epsilon\right) = \mathbb{P}\left(T_{\frac{\epsilon}{2}} \geq T\right) + \mathbb{P}\left(T_{\frac{\epsilon}{2}} < T, \forall t \in \left[T_{\frac{\epsilon}{2}}, T\right] Y_t < \epsilon\right).$$

Now, if $\mathbb{P}\left(T_{\frac{\epsilon}{2}} < T\right) > 0$, we can deduce that the second term in the right hand side is positive using the strong Markov property and the same argument we have used before in the case with $Y_0 \neq 0$. Otherwise, $\mathbb{P}\left(T_{\frac{\epsilon}{2}} \geq T\right) = 1$ which concludes the proof. $\square$

Proof of Lemma 2.4.11. Let us define $\tilde{\mathcal{E}}_t = \{(s,y) \in (0,\infty) \times [0,\infty) : s < b(t,y^+)\}$. Note that $\tilde{\mathcal{E}}_t \neq \emptyset$ since $b > 0$. We first show that $\overline{\tilde{\mathcal{E}}_t} = \mathcal{E}_t$. If $(s,y) \in \tilde{\mathcal{E}}_t$, then $s < b(t,y^+) \leq b(t,y)$, since $y \mapsto b(t,y)$ is nonincreasing. Therefore, $\tilde{\mathcal{E}}_t \subseteq \mathcal{E}_t$ so that, since $\mathcal{E}_t$ is closed, $\overline{\tilde{\mathcal{E}}_t} \subseteq \mathcal{E}_t$.

On the other hand, let $(s,y) \in \mathcal{E}_t$ and consider the sequence $((s_n,y_n))_n = ((s-1/n, y-1/n))_n$. Then, $(s_n,y_n) \to (s,y)$ and we prove that $(s_n,y_n) \in \tilde{\mathcal{E}}_t$, so that $(s,y) \in \overline{\tilde{\mathcal{E}}_t}$. In fact, for each $n \in \mathbb{N}$, we can consider the sequence $((s_{n,k},y_{n,k}))_{k>n}, = \left(\left(s - \frac{1}{n} + \frac{1}{k}, y - \frac{1}{n} + \frac{1}{k}\right)\right)_{k>n}$. We have

$$s_{n,k} = s - \frac{1}{n} + \frac{1}{k} < s \leq b(t,y) \leq b\left(t, y - \frac{1}{n} + \frac{1}{k}\right) = b\left(t, y_{n,k}\right).$$

Letting k tends to infinity, we get

$$s_n < s \leq b(t,y_n^+),$$

hence $(s_n,y_n) \in \tilde{\mathcal{E}}_t$, and the assertion is proved.

Then, we show that $\tilde{\mathcal{E}}_t = \mathring{\mathcal{E}}_t$. Note that $\tilde{\mathcal{E}}_t$ is an open set, since the function $(s,y) \mapsto b(t,y^+) - s$ is lower semicontinuous. Therefore $\tilde{\mathcal{E}}_t \subseteq \mathring{\mathcal{E}}_t$. Let us now consider an open set $A \subseteq \mathcal{E}_t$. Fix $(s,y) \in A$, then $\left(s + \frac{1}{n}, y + \frac{1}{n}\right) \in A$ for n large enough. Therefore,

$$s < s + \frac{1}{n} \leq b\left(t, y + \frac{1}{n}\right) \leq b(t,y^+),$$

hence $(s,y) \in \tilde{\mathcal{E}}_t$. $\square$

Proof of Lemma 2.4.13. We have

$$\begin{aligned} Y_t^y - y &= \kappa \int_0^t (\theta - Y_s^y) ds + \sigma \int_0^t \sqrt{Y_s^y} dW_s \\ &= \sigma\sqrt{y} W_t + \kappa \int_0^t (\theta - Y_s^y) ds + \sigma \int_0^t \left(\sqrt{Y_s^y} - \sqrt{y}\right) dW_s, \end{aligned}$$

so it is enough to prove that, if $(H_t)_{t\geq 0}$ is a predictable process such that $\lim_{t\downarrow 0} H_t = 0$ a.s., we have

$$\lim_{t\downarrow 0} \frac{\int_0^t H_s dWs}{\sqrt{2t \ln\ln(1/t)}} = 0 \text{ p.s.}$$

This follows by using standard arguments, we include a proof for the sake of completeness. By using Dubins-Schwartz inequality we deduce that, if $f(t) = \sqrt{2t \ln\ln(1/t)}$, for t near to 0 we have

$$\left|\int_0^t H_s dWs\right| \leq Cf\left(\int_0^t H_s^2 ds\right).$$

Let us consider $\varepsilon > 0$. For t small enough, we have $\int_0^t H_s^2 ds \leq \varepsilon t$ and, since f increases near 0,

$$\left| \int_0^t H_s dW s \right| \leq C f(\varepsilon t).$$

We have

$$\frac{f^2(\varepsilon t)}{f^2(t)} = \frac{\varepsilon t \ln\ln(1/\varepsilon t)}{t \ln\ln(1/t)} = \varepsilon \frac{\ln\left(\ln(1/t) + \ln(1/\varepsilon)\right)}{\ln\ln(1/t)}$$
$$\leq \varepsilon \frac{\ln\left(\ln(1/t)\right) + \frac{\ln(1/\varepsilon)}{\ln(1/t)}}{\ln\ln(1/t)} = \varepsilon \left(1 + \frac{\ln(1/\varepsilon)}{\ln(1/t)\ln\ln(1/t)}\right),$$

where we have used the inequality $\ln(x+h) \leq \ln(x) + \frac{h}{x}$ (for $x, h > 0$). Therefore $\limsup_{t\downarrow 0} \frac{f(\varepsilon t)}{f(t)} \leq \sqrt{\varepsilon}$ and the assertion follows. □

Proof of Lemma 2.4.14. With standard inversion arguments, we can reduce to prove that, for a sequence t_n such that $\lim_{n\to\infty} t_n = \infty$, we have, with probability one,

$$\limsup_{n\to\infty} \frac{B_{t_n}}{\sqrt{t_n}} = +\infty. \tag{2.5.38}$$

The assertion is equivalent to

$$\mathbb{P}\left(\limsup_{n\to\infty} \frac{B_{t_n}}{\sqrt{t_n}} \leq c\right) = 0, \qquad c > 0,$$

that is

$$\mathbb{P}\left(\bigcup_{m\geq 1}\bigcap_{n\geq m} \left\{\frac{B_{t_n}}{\sqrt{t_n}} \leq c\right\}\right) = 0, \qquad c > 0.$$

Therefore, it is sufficient to prove that $\mathbb{P}\left(\bigcap_{n\geq m}\left\{\frac{B_{t_n}}{\sqrt{t_n}} \leq c\right\}\right) = 0$ for every $m \in \mathbb{N}$ and $c > 0$. Take, for example, $m = 1$ and consider the random variables $\frac{B_{t_1}}{\sqrt{t_1}}$ and $\frac{B_{t_n}}{\sqrt{t_n}}$, for some $n > 1$. Then,

$$\frac{B_{t_1}}{\sqrt{t_1}}, \frac{B_{t_n}}{\sqrt{t_n}} \sim \mathcal{N}(0,1),$$

where $\mathcal{N}(0,1)$ is the standard Gaussian law and

$$\mathrm{Cov}\left(\frac{B_{t_1}}{\sqrt{t_1}}, \frac{B_{t_n}}{\sqrt{t_n}}\right) = \frac{t_1 \wedge t_n}{\sqrt{t_1 t_n}} < \sqrt{\frac{t_1}{t_n}},$$

which tends to 0 as n tends to infinity. We deduce that

$$\mathbb{P}\left(\frac{B_{t_1}}{\sqrt{t_1}} \leq c, \frac{B_{t_n}}{\sqrt{t_n}} \leq c\right) \to \mathbb{P}(Z_1 \leq c, Z_2 \leq c) = \mathbb{P}(Z_1 \leq c)^2,$$

where Z_1 and Z_2 are independent with Z_1, $Z_2 \sim \mathcal{N}(0,1)$.

Take now $m_n \in \mathbb{N}$ such that $t_{m_n} > nt_n$. Then, we have

$$\frac{B_{t_1}}{\sqrt{t_1}}, \frac{B_{t_n}}{\sqrt{t_n}}, \frac{B_{t_{m_n}}}{\sqrt{t_{m_n}}} \sim \mathcal{N}(0,1)$$

and

$$\mathrm{Cov}\left(\frac{B_{t_1}}{\sqrt{t_1}}, \frac{B_{t_{m_n}}}{\sqrt{t_{m_n}}}\right), \mathrm{Cov}\left(\frac{B_{t_n}}{\sqrt{t_n}}, \frac{B_{t_{m_n}}}{\sqrt{t_{m_n}}}\right) \leq \sqrt{\frac{t_n}{t_{m_n}}}.$$

which again tends to 0 ad n tends to infinity. Therefore, we have

$$\mathbb{P}\left(\frac{B_{t_1}}{\sqrt{t_1}} \leq c, \frac{B_{t_n}}{\sqrt{t_n}} \leq c, \frac{B_{t_{m_n}}}{\sqrt{t_{m_n}}} \leq c\right) \to \mathbb{P}(Z_1 \leq c)^3$$

with $Z_1 \sim \mathcal{N}(0,1)$. Iterating this procedure, we can find a subsequence $(t_{n_k})_{k\in\mathbb{N}}$ such that $t_{n_k} \to \infty$ and

$$\mathbb{P}\left(\bigcap_{k\geq 1}\left\{\frac{B_{t_{n_k}}}{\sqrt{t_{n_k}}} \leq c\right\}\right) = 0$$

which proves that $\limsup_{n\to\infty} \frac{B_{t_n}}{\sqrt{t_n}} = +\infty$.

□

Part II

Hybrid schemes for pricing options in jump-diffusion stochastic volatility models

Chapter 3

Hybrid Monte Carlo and tree-finite differences algorithm for pricing options in the Bates-Hull-White model

3.1 Introduction

In this chapter, which is extracted from [27], we focus on the so called Bates-Hull-White model. Following the previous work in [24, 25], we further develop and study the hybrid tree/finite-difference approach and the hybrid Monte Carlo technique in order to numerically evaluate option prices.

The Bates model [17] is a stochastic volatility model with price jumps: the dynamics of the underlying asset price is driven by both a Heston stochastic volatility [58] and a compound Poisson jump process of the type originally introduced by Merton [77]. Such a model was introduced by Bates in the foreign exchange option market in order to tackle the well-known phenomenon of the volatility smile behavior. Here, we assume a possibly stochastic interest rate following the Vasicek model, and we call the full model as Bates-Hull-White. In the case of plain vanilla European options, Fourier inversion methods [33] lead to closed-form formulas to compute the price under the Bates model. Nevertheless, in the American case the numerical literature is limited. Typically, numerical methods are based on the use of the dynamic programming principle to which one applies either deterministic schemes from numerical analysis and/or from tree methods or Monte

Carlo techniques.

The option pricing hybrid tree/finite-difference approach we deal with, derives from applying an efficient recombining binomial tree method in the direction of the volatility and the interest rate components, whereas the asset price component is locally treated by means of a one-dimensional partial integro-differential equation (PIDE), to which a finite-difference scheme is applied. Here, the numerical treatment of the nonlocal term coming from the jumps involves implicit-explicit techniques, as well as numerical quadratures.

The existing literature on numerical schemes for the option pricing problem in this framework is quite poor. Tree methods are available only for the Heston model, see [94], but they are not really efficient when the Feller condition does not hold. Another approach is given by the dicretization of partial differential problems. When the jumps are not considered, namely for the Heston and the Heston-Hull-White models, available references are widely recalled in [24, 25]. In the standard Bates model, that is, presence of jumps but no randomness in the interest rate, the finite-difference methods for solving the 2-dimensional PIDE associated with the option pricing problems can be based on implicit, explicit or alternating direction implicit schemes. The implicit scheme requires to solve a dense sparse system at each time step. Toivanen [92] proposes a componentwise splitting method for pricing American options. The linear complementarity problem (LCP) linked to the American option problem is decomposed into a sequence of five one-dimensional LCP's problems at each time step. The advantage is that LCP's need the use of tridiagonal matrices. Chiarella *et al.* [34] developed a method of lines algorithm for pricing and hedging American options again under the standard Bates dynamics. More recently Itkin [64] proposes a unified approach to handle PIDE's associated with Lévy's models of interest in Finance, by solving the diffusion equation with standard finite-difference methods and by transforming the jump integral into a pseudo-differential operator. But to our knowledge, no deterministic numerical methods are available in the literature for the Bates-Hull-White model, that is, when the the interest rate is assumed to be stochastic.

From the simulation point of view, the main problem consists in the treatment of the CIR dynamics for the volatility process. It is well known that the standard Euler-Maruyama discretization does not work in this framework. As far as we know, the most accurate simulation schemes for the CIR process have been introduced by Alfonsi [4]. Other methods are available in the literature, see e.g. [7], but in this chapter the Alfonsi technique is the one we compare with. In fact, in our numerical experiments we also apply a hybrid Monte Carlo technique: we couple the simulation of the approximating tree for the volatility and the interest rate components with a standard simulation of the underlying asset price, which uses Brownian increments and a straightforward treatment of the jumps. In the case of American option, this is associated with the Longstaff and Schwartz algorithm [76], allowing to treat the dynamic programming principle.

As already observed in [24, 25], roughly speaking our methods consist in the application of the most efficient method whenever this is possible: a recombining binomial tree for the volatility and the interest rate, a standard PIDE approach or a standard simulation technique in the direction of the asset price. The results of the numerical tests again support the accuracy of our hybrid methods and besides, we also justify the good behavior of the methods from the theoretical point of view (see also Chapter 4).

This chapter is devoted to present in detail the hybrid procedures introduced in [27] to compute functionals of the Bates jump model with stochastic interest rate. In particular, we consider a hybrid tree-finite differences procedure which uses a tree method in the direction of the volatility and the interest rate and a finite-difference approach in order to handle the underlying asset price process. We also propose hybrid simulations for the model, following a binomial tree in the direction of both the volatility and the interest rate, and a space-continuous approximation for the underlying asset price process coming from a Euler-Maruyama type scheme. As regards the theoretical analysis of the algorithm, we study here the stability properties of the procedure and we refer to Chapter 4 for an analysis of the rate of convergence of a generalization of this algorithm under quite general assumptions. We provide numerical experiments which show the reliability and the efficiency of the algorithms.

The chapter is organized as follows. In Section 3.2, we introduce the Bates-Hull-White model. In Section 3.3 we describe the tree procedure for the volatility and the interest rate pair (Section 3.3.1), we illustrate our discretization of the log-price process (Section 3.3.2) and the hybrid Monte Carlo simulations (Section 3.3.3). Section 3.4 is devoted to the hybrid tree/finite-difference method: we first set the numerical scheme for the associated local PIDE problem (Section 3.4.1), then we apply it to the solution of the whole pricing scheme (Section 3.4.2) and analyze the numerical stability of the resulting tree/finite-difference method (Section 3.4.3). Section 3.5 refers to the practical use of our methods and numerical results and comparisons are widely discussed.

3.2 The Bates-Hull-White model

We recall that in the Bates-Hull-White model the volatility is assumed to follow the CIR process and the underlying asset price process contains a further noise from a jump as introduced by Merton. Moreover, the interest rate follows a stochastic model, which we assume to be described by a generalized Ornstein-Uhlenbeck (hereafter OU) process. More precisely, the dynamics under the risk neutral measure of the share price S, the volatility process Y and the interest rate r, are given by the following jump-diffusion model:

$$\begin{aligned}
&\frac{dS_t}{S_{t^-}} = (r_t - \delta)dt + \sqrt{Y_t}\, dZ_t^S + dH_t, \\
&dY_t = \kappa_Y(\theta_Y - Y_t)dt + \sigma_Y \sqrt{Y_t}\, dZ_t^Y, \\
&dr_t = \kappa_r(\theta_r(t) - r_t)dt + \sigma_r dZ_t^r,
\end{aligned} \tag{3.2.1}$$

where δ denotes the continuous dividend rate, $S_0, Y_0, r_0 > 0$, Z^S, Z^Y and Z^r are correlated Brownian motions and H is a compound Poisson process with intensity λ and i.i.d. jumps $\{J_k\}_k$, that is

$$H_t = \sum_{k=1}^{K_t} J_k, \tag{3.2.2}$$

K denoting a Poisson process with intensity λ. We assume that the Poisson process K, the jump amplitudes $\{J_k\}_k$ and the 3-dimensional correlated Brownian motion (Z^S, Z^Y, Z^r) are independent. As suggested by Grzelak and Oosterlee in [55], the significant correlations are between the noises governing the pairs (S, Y) and (S, r). So, as done in [25], we assume that the couple (Z^Y, Z^r) is a standard Brownian motion in $\mathbb{R}^2$ and Z^S is a Brownian motion in $\mathbb{R}$ which is correlated both with Z^Y and Z^r:

$$d\langle Z^S, Z^Y\rangle_t = \rho_1 dt \quad \text{and} \quad d\langle Z^S, Z^r\rangle_t = \rho_2 dt.$$

We recall that the volatility process Y follows a CIR dynamics with mean reversion rate κ_Y, long run variance θ_Y and σ_Y denotes the vol-vol (volatility of the volatility). We assume that $\theta_Y, \kappa_Y, \sigma_Y > 0$ and we stress that we never require in this chapter that the CIR process satisfies the Feller condition $2\kappa_Y\theta_Y \geq \sigma_Y^2$, ensuring that the process Y never hits 0. So, we allow the volatility Y to reach 0. The interest rate r_t is described by a generalized OU process, in particular θ_r is time-dependent but deterministic and fits the zero-coupon bond market values, for details see [30]. We write the process r as follows:

$$r_t = \sigma_r R_t + \varphi_t \tag{3.2.3}$$

where

$$R_t = -\kappa_r \int_0^t R_s\, ds + Z_t^r \quad \text{and} \quad \varphi_t = r_0 e^{-\kappa_r t} + \kappa_r \int_0^t \theta_r(s) e^{-\kappa_r (t-s)} ds. \tag{3.2.4}$$

From now on we set

$$Z^Y = W^1, \quad Z^r = W^2, \quad Z^S = \rho_1 W^1 + \rho_2 W^2 + \rho_3 W^3,$$

where $W = (W^1, W^2, W^3)$ is a standard Brownian motion in $\mathbb{R}^3$ and the correlation parameter ρ_3 is given by

$$\rho_3 = \sqrt{1 - \rho_1^2 - \rho_2^2}, \quad \rho_1^2 + \rho_2^2 \leq 1.$$

By passing to the logarithm $X = \ln S$ in the first component, by taking into account the above mentioned correlations and by considering the process R as in (3.2.3)-(3.2.4), we reduce to the triple (X, Y, R) given by

$$\begin{aligned} dX_t &= \mu_X(Y_t, R_t, t)dt + \sqrt{Y_t}\left(\rho_1 dW_t^1 + \rho_2 dW_t^2 + \rho_3 dW_t^3\right) + dN_t, \quad X_0 = \ln S_0 \in \mathbb{R}, \\ dY_t &= \mu_Y(Y_t)dt + \sigma_Y \sqrt{Y_t}\, dW_t^1, \quad Y_0 > 0, \\ dR_t &= \mu_R(R_t)dt + dW_t^2, \quad R_0 = 0, \end{aligned} \tag{3.2.5}$$

where

$$\mu_X(y, r, t) = \sigma_r r + \varphi_t - \delta - \frac{1}{2} y, \tag{3.2.6}$$

$$\mu_Y(y) = \kappa_Y(\theta_Y - y), \tag{3.2.7}$$

$$\mu_R(r) = -\kappa_r r, \tag{3.2.8}$$

and N_t is the compound Poisson process with intensity λ and the i.i.d. jumps $\{\log(1 + J_k)\}_k$, that is

$$N_t = \sum_{k=1}^{K_t} \log(1 + J_k),$$

K being a Poisson process with intensity λ. Recall that K, the jump amplitudes $\{\log(1+J_k)\}_k$ and the 3-dimensional standard Brownian motion (W^1, W^2, W^3) are all independent. We also recall that the Lévy measure associated with N is given by

$$\nu(dx) = \lambda \mathbb{P}(\log(1 + J_1) \in dx),$$

and whenever $\log(1 + J_1)$ is absolutely continuous then ν has a density as well:

$$\nu(dx) = \nu(x)dx = \lambda p_{\log(1+J_1)}(x)dx, \tag{3.2.9}$$

$p_{\log(1+J_1)}$ denoting the probability density function of $\log(1+J_1)$. For example, in the Merton model [77] it is assumed that $\log(1+J_1)$ has a normal distribution, that is

$$\log(1+J_1) \sim N(\mu, \eta^2).$$

This is the choice we will do in our numerical experiments, as done in Chiarella *et al.* [34]. But other jump-amplitude measures can be selected. For instance, in the Kou model [69] the law of $\log(1+J_1)$ is a mixture of exponential laws:

$$p_{\log(1+J_1)}(x) = p\lambda_+ e^{-\lambda_+ x}\,\mathbb{1}_{\{x>0\}} + (1-p)\lambda_- e^{\lambda_- x}\,\mathbb{1}_{\{x<0\}},$$

$\mathbb{1}_A$ denoting the indicator function of A. Here, the parameters $\lambda_\pm > 0$ control the decrease of the distribution tails of negative and positive jumps respectively, and p is the probability of a positive jump.

Given this framework, our aim is to numerically compute the price of options with maturity T and payoff given by a function of the underlying asset price process S. By passing to the transformation $X = \ln S$, we assume that the payoff is a function of the log-price process:

$$\begin{aligned} &\text{European payoff:} \quad \Psi(X_T), \\ &\text{American payoff:} \quad (\Psi(X_t))_{t\in[0,T]}, \end{aligned}$$

where $\Psi \geq 0$. The option price function $P(t,x,y,r)$ is then given by

$$\begin{aligned} &\text{European price:} \quad P(t,x,y,r) = \mathbb{E}\Big(e^{-\int_t^T (\sigma_r R_s^{t,r} + \varphi_s)ds}\,\Psi(X_T^{t,x,y,r})\Big), \\ &\text{American price:} \quad P(t,x,y,r) = \sup_{\tau\in\mathcal{T}_{t,T}} \mathbb{E}\Big(e^{-\int_t^\tau (\sigma_r R_s^{t,r} + \varphi_s)ds}\,\Psi(X_\tau^{t,x,y,r})\Big), \end{aligned} \tag{3.2.10}$$

where $\mathcal{T}_{t,T}$ denotes the set of all stopping times taking values on $[t,T]$. Note that we have used the relation between the interest rate $(r_t)_t$ and the process $(R_t)_t$, see (3.2.3) and (3.2.4). Hereafter, $(X^{t,x,y,r}, Y^{t,y}, R^{t,r})$ denotes the solution of the jump-diffusion dynamic (3.2.5) starting at time t in the point (x,y,r).

3.3 The dicretized process

We first set up the discretization of the triple (X,Y,R) we will take into account.

3.3.1 The 2-dimensional tree for (Y, R)

We consider an approximation for the pair (Y, R) on the time-interval $[0, T]$ by means of a 2-dimensional computationally simple tree. This means that we construct a Markov chain running over a 2-dimensional recombining bivariate lattice and, at each time-step, both components of the Markov chain can jump only upwards or downwards. We consider the "multiple-jumps" approach by Nelson and Ramaswamy [79]. A detailed description of this procedure and of the benefits of its use, can be found in [10, 24, 25]. Here, we limit the reasoning to the essential ideas and to the main steps in order to set-up the whole algorithm. We start by considering a discretization of the time-interval $[0, T]$ in N subintervals $[nh, (n+1)h]$, $n = 0, 1, \dots, N$, with $h = T/N$.

For the CIR volatility process Y, we consider the binomial tree procedure firstly introduced in [10]. For $n = 0, 1, \dots, N$, consider the lattice

$$\mathcal{Y}_n = \{y_k^n\}_{k=0,1,\dots,n} \quad \text{with} \quad y_k^n = \Big(\sqrt{Y_0} + \frac{\sigma_Y}{2}(2k-n)\sqrt{h}\Big)^2 \mathbb{1}_{\{\sqrt{Y_0}+\frac{\sigma_Y}{2}(2k-n)\sqrt{h}>0\}}. \tag{3.3.11}$$

Note that $y_0^0 = Y_0$, so that $\mathcal{Y}_0^h = \{Y_0\}$. Moreover, the lattice is binomial recombining and, for n large, the "small" points degenerate at 0. Let us briefly recall how this lattice arises (see [10] for all the details). The idea is to reduce to a process with a constant diffusion coefficient. So, let us consider the process $\hat{Y}_t = \sqrt{Y_t}$. If we (heuristically) apply Itô formula, we get that the dynamics of $\hat{Y}_t$ is given by

$$d\hat{Y}_t = \mu_{\hat{Y}}(\hat{Y}_t)dt + \frac{\sigma}{2}dZ_t^Y,$$

for a suitable drift coefficient $\mu_{\hat{Y}} = \mu_{\hat{Y}}(y)$. The term $\frac{\sigma}{2}dB_t$ gives the foremost contribution to the local movement of $\hat{Y}_t$. The standard binomial recombining tree for the Brownian motion lives on the lattice

$$\frac{\sigma}{2}(2k-n)\sqrt{h}, \qquad 0 \le k \le n \le N.$$

Coming back to Y, we get the lattice in (3.3.11). Note that the term $\mathbb{1}_{\{\sqrt{Y_0}+\frac{\sigma_Y}{2}(2k-n)\sqrt{h}>0\}}$ is inserted in order to deal with invertible functions.

We now define the multiple "up" and "down" jumps: the discretized process can jump just on two nodes which in turn are not necessarily the closest ones to the starting node. In particular, for each fixed $y_k^n \in \mathcal{Y}_n$, we define the "up" and "down" jump by $y_{k_u(n,k)}^{n+1}$ and $y_{k_d(n,k)}^{n+1}$, $k_u(n,k)$ and $k_d(n,k)$ being respectively defined as

$$k_u(n,k) = \min\{k^* \,:\, k+1 \le k^* \le n+1 \text{ and } y_k^n + \mu_Y(y_k^n)h \le y_{k^*}^{n+1}\}, \tag{3.3.12}$$

$$k_d(n,k) = \max\{k^* \,:\, 0 \le k^* \le k \text{ and } y_k^n + \mu_Y(y_k^n)h \ge y_{k^*}^{n+1}\} \tag{3.3.13}$$

where μ_Y is the drift of Y, defined in (3.2.6), and with the understanding $k_u(n,k) = n+1$, respectively $k_d(n,k) = 0$, if the set in the r.h.s. of (3.3.12), respectively (3.3.13), is empty. The transition probabilities are defined as follows: starting from the node (n,k) the probability that the process jumps to $k_u(n,k)$ and $k_d(n,k)$ at time-step $n+1$ are set as

$$p_u^Y(n,k) = 0 \vee \frac{\mu_Y(y_k^n)h + y_k^n - y_{k_d(n,k)}^{n+1}}{y_{k_u(n,k)}^{n+1} - y_{k_d(n,k)}^{n+1}} \wedge 1 \quad \text{and} \quad p_d^Y(n,k) = 1 - p_u^Y(n,k) \tag{3.3.14}$$

respectively. We recall that the multiple jumps and the transition probabilities are set in order to best fit the local first moment of the diffusion Y. We will see in Chapter 4 that this property will be crucial in order to study the theoretical convergence of the procedure.

We follow the same approach for the binomial tree for the process R. For $n = 0,1,\ldots,N$ consider the lattice

$$\mathcal{R}_n = \{r_j^n\}_{j=0,1,\ldots,n} \quad \text{with} \quad r_j^n = (2j-n)\sqrt{h}. \tag{3.3.15}$$

Notice that $r_{0,0} = 0 = R_0$. For each fixed $r_j^n \in \mathcal{R}_n$, we define the "up" and "down" jump by means of $j_u(n,j)$ and $j_d(n,j)$ defined by

$$j_u(n,j) = \min\{j^* : j+1 \le j^* \le n+1 \text{ and } r_j^n + \mu_R(r_j^n)h \le r_{j^*}^{n+1}\}, \tag{3.3.16}$$

$$j_d(n,j) = \max\{j^* : 0 \le j^* \le j \text{ and } r_j^n + \mu_R(r_j^n)h \ge r_{j^*}^{n+1}\}, \tag{3.3.17}$$

μ_R being the drift of the process R, see (3.2.8). As before, $j_u(n,j) = n+1$, respectively $j_d(n,j) = 0$, if the set in the r.h.s. of (3.3.16), respectively (3.3.17), is empty and the transition probabilities are as follows: starting from the node (n,j), the probability that the process jumps to $j_u(n,j)$ and $j_d(n,j)$ at time-step $n+1$ are set as

$$p_u^R(n,j) = 0 \vee \frac{\mu_R(r_j^n)h + r_j^n - r_{j_d(n,j)}^{n+1}}{r_{j_u(n,j)}^{n+1} - r_{j_d(n,j)}^{n+1}} \wedge 1 \quad \text{and} \quad p_d^R(n,j) = 1 - p_u^R(n,j) \tag{3.3.18}$$

respectively.

Figure 3.1 shows a picture of the lattices $\mathcal{Y}_n$ (left) and $\mathcal{R}_n$ (right), together with possible instances of the up and down jumps.

The whole tree procedure for the pair (Y,R) is obtained by joining the trees built for Y and for R. Namely, for $n = 0,1,\ldots,N$, consider the lattice

$$\mathcal{Y}_n \times \mathcal{R}_n = \{(y_k^n, r_j^n)\}_{k,j=0,1,\ldots,n}. \tag{3.3.19}$$

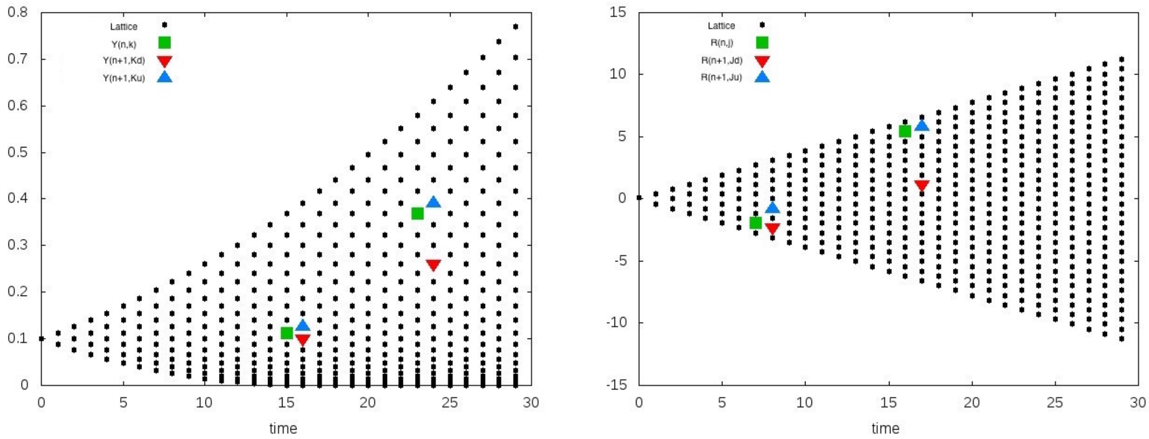

Figure 3.1: The tree for the process Y (left) and for R (right), showing as the trees may be visited.

Starting from the node (n, k, j), which corresponds to the position $(y_k^n, r_j^n) \in \mathcal{Y}_n \times \mathcal{R}_n$, we define the four possible jumps by means of the following four nodes at time $n + 1$:

$$\begin{aligned}
&(n+1, k_u(n,k), j_u(n,j)) \quad \text{with probability} \quad p_{uu}(n,k,j) = p_u^Y(n,k)p_u^R(n,j),\\
&(n+1, k_u(n,k), j_d(n,j)) \quad \text{with probability} \quad p_{ud}(n,k,j) = p_u^Y(n,k)p_d^R(n,j),\\
&(n+1, k_d(n,k), j_u(n,j)) \quad \text{with probability} \quad p_{du}(n,k,j) = p_d^Y(n,k)p_u^R(n,j),\\
&(n+1, k_d(n,k), j_d(n,j)) \quad \text{with probability} \quad p_{dd}(n,k,j) = p_d^Y(n,k)p_d^R(n,j),
\end{aligned} \tag{3.3.20}$$

where the above nodes $k_u(n,k)$, $k_d(n,k)$, $j_u(n,j)$, $j_d(n,j)$ and the above probabilities $p_u^Y(n,k)$, $p_d^Y(n,k)$, $p_u^R(n,j)$, $p_d^R(n,j)$ are defined in (3.3.12)-(3.3.13), (3.3.16)-(3.3.17), (3.3.14) and (3.3.18). The factorization of the jump probabilities in (3.3.20) follows from the orthogonality property of the noises driving the two processes. This procedure gives rise to a Markov chain $(\hat{Y}_n^h, \hat{R}_n^h)_{n=0,\dots,N}$ that weakly converges, as $h \to 0$, to the diffusion process $(Y_t, R_t)_{t\in[0,T]}$ solution to

$$\begin{aligned}
dY_t &= \mu_Y(Y_t)dt + \sigma_Y \sqrt{Y_t}\, dW_t^1, \quad Y_0 > 0,\\
dR_t &= \mu_R(R_t)\, dt + dW_t^2, \quad R_0 = 0.
\end{aligned}$$

This can be seen by using standard results (see e.g. the techniques in [79]) and the convergence of the chain approximating the volatility process proved in [10]. And this holds independently of the validity of the Feller condition $2\kappa_Y\theta_Y \geq \sigma_Y^2$.

Details and remarks on the extension of this procedure to more general cases can be found in [25]. In particular, if the correlation between the Brownian motions driving (Y, R) was not null, one could define the jump probabilities by matching the local cross-moment (see Remark 3.1 in [25]).

3.3.2 The approximation on the X-component

We describe here how we manage the X-component in (3.2.5) by taking into account the tree procedure given for the pair (Y,R). We go back to (3.2.5): by isolating $\sqrt{Y_t}dW_t^1$ in the second line and dW_t^2 in the third one, we obtain

$$dX_t = \mu(Y_t,R_t,t)dt + \rho_3\sqrt{Y_t}\,dW_t^3 + \frac{\rho_1}{\sigma_Y}dY_t + \rho_2\sqrt{Y_t}dR_t + dN_t \tag{3.3.21}$$

with

$$\begin{aligned}\mu(y,r,t) &= \mu_X(y,r,t) - \tfrac{\rho_1}{\sigma_Y}\mu_Y(y) - \rho_2\sqrt{y}\,\mu_R(r)\\ &= \sigma_r r + \varphi_t - \delta - \tfrac{1}{2}\,y - \tfrac{\rho_1}{\sigma_Y}\,\kappa_Y(\theta_Y - y) + \rho_2\kappa_r r\sqrt{y}\end{aligned} \tag{3.3.22}$$

(μ_X, μ_Y and μ_R are defined in (3.2.6), (3.2.7) and (3.2.8) respectively). To numerically solve (3.3.21), we mainly use the fact that the noises W^3 and N are independent of the processes Y and R. So, we first take the approximating tree $(\hat{Y}_n^h, \hat{R}_n)_{n=0,1,\dots,N-1}$ discussed in Section 3.3.1 and we set $(\bar{Y}_t^h, \bar{R}_t^h)_{t\in[0,T]} = (\hat{Y}^h_{\lfloor t/h\rfloor} + 1, \hat{R}^h_{\lfloor t/h\rfloor} + 1)_{t\in[0,T]}$ the associated time-continuous càdlàg approximating process for (Y,R). Then, we insert the discretization $(\bar{Y}^h, \bar{R}^h)$ for (Y,R) in the coefficients of (3.3.21). Therefore, the final process $\bar{X}^h$ approximating X is set as follows: $\bar{X}_0^h = X_0$ and for $t \in (nh,(n+1)h]$ with $n = 0,1,\dots,N-1$

$$\begin{aligned}\bar{X}_t^h = \;& \bar{X}_{nh}^h + \mu(\bar{Y}_{nh}^h, \bar{R}_{nh}^h, nh)(t-nh) + \rho_3\sqrt{\bar{Y}_t^h}(W_t^3 - W_{nh}^3)\\ &+ \frac{\rho_1}{\sigma_Y}(\bar{Y}_t^h - \bar{Y}_{nh}^h) + \rho_2\sqrt{\bar{Y}_t^h}(\bar{R}_t^h - \bar{R}_{nh}^h) + (N_t - N_{nh}).\end{aligned} \tag{3.3.23}$$

3.3.3 The Monte Carlo approach

Let us show how one can simulate a single path by using the tree approximation (3.3.19) for the couple (Y,R) and the Euler scheme (3.3.23) for the X-component.

Let $(\hat{X}_n)_{n=0,1,\dots,N}$ be the sequence approximating X at times nh, $n = 0,1,\dots,N$, by means of the scheme in (3.3.23): $\hat{X}_0^h = X_0$ and for $t \in [nh,(n+1)h]$ with $n = 0,1,\dots,N-1$ then

$$\begin{aligned}\hat{X}_{n+1}^h = \;& \hat{X}_n^h + \mu(\hat{Y}_n^h, \hat{R}_n^h, nh)h + \rho_3\sqrt{h\hat{Y}_n^h}\Delta_{n+1}\\ &+ \frac{\rho_1}{\sigma_Y}(\hat{Y}_{n+1}^h - \hat{Y}_n^h) + \rho_2\sqrt{\hat{Y}_n^h}(\hat{R}_{n+1}^h - \hat{R}_n^h) + (N_{(n+1)h} - N_{nh}),\end{aligned}$$

where μ is defined in (3.3.22) and $\Delta_1, \dots, \Delta_N$ denote i.i.d. standard normal r.v.'s, independent of the noise driving the chain $(\hat{Y}, \hat{R})$. The simulation of $N_{(n+1)h} - N_{nh}$ is straightforward: one first generates a Poisson r.v. K_h^{n+1} of parameter λh and if $K_h^{n+1} > 0$ then also the log-amplitudes $\log(1+J_k^{n+1})$ for $k = 1, \dots, K_h^{n+1}$ are simulated. Then, the observed jump of the compound Poisson process is written as the sum of the simulated log-amplitudes, so that

$$\begin{aligned} \hat{X}_{n+1}^h = \quad & \hat{X}_n^h + \mu(\hat{Y}_n^h, \hat{R}_n^h, nh)h + \rho_3 \sqrt{h \hat{Y}_n^h} \Delta_{n+1} \\ & + \frac{\rho_1}{\sigma_Y}(\hat{Y}_{n+1}^h - \hat{Y}_n^h) + \rho_2 \sqrt{\hat{Y}_n^h}(\hat{R}_{n+1}^h - \hat{R}_n^h) + \sum_{k=1}^{K_h^{n+1}} \log(1 + J_k^{n+1}), \end{aligned} \tag{3.3.24}$$

in which the last sum is set equal to 0 if $K_h^{n+1} = 0$.

The above simulation scheme is plain: at each time step $n \geq 1$, one lets the pair (Y, R) evolve on the tree and simulate the process X by using (3.3.24). We will refer to this procedure as *hybrid Monte Carlo algorithm*, the word "hybrid" being related to the fact that two different noise sources are considered: we simulate a continuous process in space (the component X) starting from a discrete process in space (the tree for (Y, R)).

The simulations just described will be used in Section 3.5 in order to set-up a Monte Carlo procedure for the computation of the option price function (3.2.10). In the case of American options, the simulations are coupled with the Monte Carlo algorithm by Longstaff and Schwartz in [76].

3.4 The hybrid tree/finite difference approach

The price-function $P(t, x, y, r)$ in (3.2.10) is typically computed by means of the standard backward dynamic programming algorithm. So, consider a discretization of the time interval $[0, T]$ into N subintervals of length $h = T/N$. Then the price $P(0, X_0, Y_0, R_0)$ is numerically approximated through the quantity $P_h(0, X_0, Y_0, R_0)$ backwardly given by

$$\begin{cases} P_h(T, x, y, r) = \Psi(x) \quad \text{and as } n = N-1, \dots, 0, \\ P_h(nh, x, y, r) = \max\Big\{\widehat{\Psi}(x), e^{-(\sigma_r r + \varphi_{nh})h} \mathbb{E}\Big(P_h\big((n+1)h, X_{(n+1)h}^{nh,x,y,r}, Y_{(n+1)h}^{nh,y}, R_{(n+1)h}^{nh,r}\big)\Big)\Big\}, \end{cases} \tag{3.4.25}$$

for $(x, y, r) \in \mathbb{R} \times \mathbb{R}_+ \times \mathbb{R}$, in which

$$\widehat{\Psi}(x) = \begin{cases} 0 & \text{in the European case,} \\ \Psi(x) & \text{in the American case.} \end{cases}$$

So, what is needed is a good approximation of the expectations appearing in the above dynamic programming principle. This is what we first deal with, starting from the dicretized process $(\bar{Y}^h, \bar{Y}^h, \bar{R}^h)$ introduced in Section 3.3.

3.4.1 The local 1-dimensional partial integro-differential equation

Let $\bar{X}^h$ denote the process in (3.3.23). If we set

$$\bar{Z}^h_t = \bar{X}^h_t - \frac{\rho_1}{\sigma_Y}(\bar{Y}^h_t - \bar{Y}^h_{nh}) - \rho_2\sqrt{\bar{Y}^h_{nh}}(\bar{R}^h_t - \bar{R}_{nh}), \quad t \in [nh, (n+1)h] \tag{3.4.26}$$

then we have

$$\begin{aligned} d\bar{Z}^h_t &= \mu(\bar{Y}^h_{nh}, \bar{R}^h_{nh}, nh)dt + \rho_3\sqrt{\bar{Y}^h_{nh}}\, dW^3_t, + dN_t \quad t \in (nh, (n+1)h], \\ \bar{Z}^h_{nh} &= \bar{X}^h_{nh}, \end{aligned} \tag{3.4.27}$$

that is, $\bar{Z}^h$ solves a jump-diffusion stochastic equation with constant coefficients and at time nh it starts from $\bar{Y}^h_{nh}$. Take now a function f: we are interested in computing

$$\mathbb{E}(f(X_{(n+1)h}) \mid X_{nh} = x, Y_{nh} = y, R_{nh} = r).$$

We actually need a function f of all variables (x, y, r) but at the present moment the variable x is the most important one, we will see later on that the introduction of (y, r) is straightforward. So, we numerically compute the above expectation by means of the one done on the approximating processes, that is,

$$\begin{aligned} &\mathbb{E}\big(f(\bar{X}^h_{(n+1)h}) \mid \bar{X}^h_{nh} = x, \bar{Y}^h_{nh} = y, \bar{R}^h_{nh} = r\big) \\ &= \mathbb{E}\big(f(\bar{Z}^h_{(n+1)h} + \frac{\rho_1}{\sigma_Y}(\bar{Y}^h_{(n+1)h} - \bar{Y}^h_{nh}) + \rho_2\sqrt{\bar{Y}^h_{nh}}(\bar{R}^h_{(n+1)h} - \bar{R}^h_{nh})) \mid \bar{Z}^h_{nh} = x, \bar{Y}^h_{nh} = y, \bar{R}^h_{nh} = r\big), \end{aligned}$$

in which we have used the process $\bar{Z}^h$ in (3.4.26). Since $(\bar{Y}^h, \bar{R}^h)$ is independent of the Brownian noise W^3 and on the compound Poisson process N driving $\bar{Z}^h$ in (3.4.27), we have the following: we set

$$\Psi_f(\zeta; x, y, r) = \mathbb{E}(f(\bar{Z}^h_{(n+1)h} + \zeta) \mid \bar{Z}^h_{nh} = x, \bar{Y}^h_{nh} = y, \bar{R}^h_{nh} = r) \tag{3.4.28}$$

and we can write

$$\begin{aligned}&\mathbb{E}(f(\bar{X}^h_{(n+1)h}) \mid \bar{X}^h_{nh}=x, \bar{Y}^h_{nh}=y, \bar{R}^h_{nh}=r)\\ &\quad=\mathbb{E}\Big(\Psi_f\Big(\tfrac{\rho_1}{\sigma_Y}(\bar{Y}^h_{(n+1)h}-\bar{Y}^h_{nh})+\rho_2\sqrt{y}(\bar{R}^h_{(n+1)h}-\bar{R}^h_{nh});x,y,r\Big)\,\Big|\,\bar{Y}^h_{nh}=y, \bar{R}^h_{nh}=r\Big).\end{aligned} \tag{3.4.29}$$

Now, in order to compute the quantity $\Psi_f(\zeta)$ in (3.4.28), we consider a generic function g and set

$$u(t,x;y,r)=\mathbb{E}(g(\bar{Z}^h_{(n+1)h}) \mid \bar{Z}^h_t=x, \bar{Y}^h_t=y, \bar{R}^h_t=r), \quad t\in[nh,(n+1)h].$$

By (3.4.27) and the Feynman-Kac representation formula we can state that, for every fixed $r\in\mathbb{R}$ and $y\geq 0$, the function $(t,x)\mapsto u(t,x;y,r)$ is the solution to

$$\begin{cases}\partial_t u(t,x;y,r)+\mathcal{L}^{(y,r)}u(t,x;y,r)=0 & y\in\mathbb{R}, t\in[nh,(n+1)h),\\ u((n+1)h,x;y,r)=g(y) & x\in\mathbb{R},\end{cases} \tag{3.4.30}$$

where $\mathcal{L}^{(y,r)}$ is the integro-differential operator

$$\begin{aligned}\mathcal{L}^{(y,r)}u(t,x;y,r)= \quad &\mu(y,r)\partial_x u(t,x;y,r)+\tfrac{1}{2}\rho_3^2 y\partial^2_{xx}u(t,x;y,r)\\ &+\int_{-\infty}^{+\infty}[u(t,x+\xi;y,r)-u(t,x;y,r)]\,\nu(\xi)d\xi,\end{aligned} \tag{3.4.31}$$

where μ is given in (3.3.22) and ν is the Lévy measure associated with the compound Poisson process N, see (3.2.9). We are assuming here that the Lévy measure is absolutely continuous (in practice, we use a Gaussian density), but it is clear that the procedure we are going to describe can be straightforwardly extended to other cases.

Finite-difference and numerical quadrature

In order to numerically compute the solution to the PIDE (3.4.30) at time nh, we generalize the approach already developed in [24, 25]: we apply a one-step finite-difference algorithm to the differential part of the problem coupled now with a quadrature rule to approximate the integral term.

We start by fixing an infinite grid on the x-axis $\mathcal{X}=\{x_i=X_0+i\Delta x\}_{i\in\mathbb{Z}}$, with $\Delta x=x_i-x_{i-1}$, $i\in\mathbb{Z}$. For fixed n and given $r\in\mathbb{R}$ and $y\geq 0$, we set $u^n_i=u(nh,x_i;y,r)$ the discrete solution of (3.4.30) at time nh on the point x_i of the grid $\mathcal{X}$ – for simplicity of notations, in the sequel we do not stress in u^n_i the dependence on (y,r).

First of all, to numerically compute the integral term in (3.4.31) we need to truncate the infinite integral domain to a bounded interval $\mathcal{I}$, to be taken large enough in order that

$$\int_{\mathcal{I}} \nu(\xi) d\xi \approx \lambda. \tag{3.4.32}$$

In terms of the process, this corresponds to truncate the large jumps. We assume that the tails of ν rapidly decrease – this is not really restrictive since applied models typically require that the tails of ν decrease exponentially. Hence, we take $L \in \mathbb{N}$ large enough, set $\mathcal{I} = [-L\Delta y, +L\Delta y]$ and apply to (3.4.32) the trapezoidal rule on the grid $\mathcal{X}$ with the same step Δx previously defined. Then, for $\xi_l = l\Delta x$, $l = -L, \ldots, L$, we have

$$\int_{-L\Delta y}^{+L\Delta y} [u(t, x + \xi) - u(t, x)] \, \nu(\xi) d\xi \approx \Delta x \sum_{l=-L}^{L} (u(t, x + \xi_l) - u(t, x)) \, \nu(\xi_l). \tag{3.4.33}$$

We notice that $x_i + \xi_l = X_0 + (i + l)\Delta x \in \mathcal{X}$, so the values $u(t, x_i + \xi_l)$ are well defined on the numerical grid $\mathcal{X}$ for any i, l. These are technical settings and can be modified and calibrated for different Lévy measures ν.

But in practice one cannot solve the PIDE problem over the whole real line. So, we have to choose artificial bounds and impose numerical boundary conditions. We take a positive integer $M > 0$ and we define a finite grid $\mathcal{X}_M = \{x_i = X_0 + i\Delta x\}_{i \in \mathcal{J}_M}$, with $\mathcal{J}_M = \{-M, \ldots, M\}$, and we assume that $M > L$. Notice that for $x = x_i \in \mathcal{X}_M$ then the integral term in (3.4.33) splits into two parts: one part concerning nodes falling into the numerical domain $\mathcal{X}_M$ and another part concerning nodes falling out of $\mathcal{X}_M$. As an example, at time $t = nh$ we have

$$\sum_{l=-L}^{L} u(nh, x_i + \xi_l) \nu(\xi_l) \approx \sum_{l=-L}^{L} u^n_{i+l} \nu(\xi_l) = \sum_{l \,:\, |l| \le L, |i+l| \le M} u^n_{i+l} \; \nu(\xi_l) + \sum_{l \,:\, |l| \le L, |i+l| > M} \tilde{u}^n_{i+l} \; \nu(\xi_l),$$

where $\tilde{u}^n_\cdot$ stands for (unknown) values that fall out of the finite numerical domain $\mathcal{X}_M$. This implies that we must choose some suitable artificial boundary conditions. In a financial context, in [39] it has been shown that a good choice for the boundary conditions is the payoff function. Although this is the choice we will do in our numerical experiments, for the sake of generality we assume here the boundary values outside $\mathcal{X}_M$ to be settled as $\tilde{u}^n_i = b(nh, x_i)$, where $b = b(t, x)$ is a fixed function defined in $[0, T] \times \mathbb{R}$.

Going back to the numerical scheme to solve the differential part of the equation (3.4.30), as already done in [25], we apply an implicit in time approximation. However, to avoid to solve at each time step a linear system with a dense matrix, the non-local integral term needs anyway an explicit in time approximation. We then obtain an implicit-explicit (hereafter IMER) scheme as proposed in [39] and [28]. Notice that more sophisticated IMER methods may be applied, see for instance [29, 87]. Let us stress that these techniques could be used in our framework, being more accurate but expensive.

As done in [25], to achieve greater precision we use the centered approximation for both first and second order derivatives in space. The discrete solution u^n at time nh is then computed in terms of the known value u^{n+1} at time $(n+1)h$ by solving the following discrete problem: for all $i \in \mathcal{J}_M$,

$$\frac{u_i^{n+1} - u_i^n}{h} + \tilde{\mu}_X(y,r)\frac{u_{i+1}^n - u_{i-1}^n}{2\Delta x} + \frac{1}{2}\rho_3^2\, y\, \frac{u_{i+1}^n - 2u_i^n + u_{i-1}^n}{\Delta x^2} + \Delta x \sum_{l=-R}^{R} \big(u_{i+l}^{n+1} - u_i^{n+1}\big)\, \nu(\xi_l) = 0. \tag{3.4.34}$$

We then get the solution $u^n = (u_{-M}^n, \dots, u_M^n)^T$ by solving the following linear system

$$A\, u^n = B u^{n+1} + d, \tag{3.4.35}$$

where $A = A(y,r)$ and B are $(2M+1) \times (2M+1)$ matrices and d is a $(2M+1)$-dimensional boundary vector defined as follows.

► **The matrix** A**.** From (3.4.34), we set A as the tridiagonal real matrix given by

$$A = \begin{pmatrix} 1+2\beta & -\alpha-\beta & & & \\ \alpha-\beta & 1+2\beta & -\alpha-\beta & & \\ & \ddots & \ddots & \ddots & \\ & & \alpha-\beta & 1+2\beta & -\alpha-\beta \\ & & & \alpha-\beta & 1+2\beta \end{pmatrix}, \tag{3.4.36}$$

with

$$\alpha = \frac{h}{2\Delta x}\, \mu(nh, y, r) \quad \text{and} \quad \beta = \frac{h}{2\Delta x^2}\, \rho_3^2 y, \tag{3.4.37}$$

μ being defined in (3.3.22). We emphasize that at each time step n, the quantities v and x are constant and known values (defined by the tree procedure for (Y,R)) and then α and β are constant parameters.

▶ **The matrix** B**.** Again from (3.4.34), B is the $(2M+1)\times(2M+1)$ real matrix given by

$$B = I + h\Delta x \begin{pmatrix} \nu(0)-\Lambda & \nu(\Delta x) & \dots & \nu(L\Delta x) & 0 \\ \nu(-\Delta x) & \nu(0)-\Lambda & \nu(\Delta x) & \dots & \nu(L\Delta x) \\ & \ddots & \ddots & \ddots & \\ 0 & \nu(-L\Delta x) & \dots & \nu(-\Delta x) & \nu(0)-\Lambda \end{pmatrix}, \qquad (3.4.38)$$

where I is the identity matrix and

$$\Lambda = \sum_{l=-L}^{L} \nu(\xi_l).$$

▶ **The boundary vector** d**.** The vector $d \in \mathbb{R}^{2M+1}$ contains the numerical boundary values:

$$d = a_b^n + a_b^{n+1}, \qquad (3.4.39)$$

with

$$a_b^n = ((\beta-\alpha)b_{-M-1}^n, 0, \dots, 0, (\beta+\alpha)b_{M+1}^n)^T \in \mathbb{R}^{2M+1}$$

and $a_b^{n+1} \in \mathbb{R}^{2M+1}$ is such that

$$(a_b^{n+1})_i = \begin{cases} h\Delta x \displaystyle\sum_{l=-L}^{-M-i-1} \nu(x_l)\ b_{i+l}^{n+1}, & \text{for } i=-M,\dots,-M+L-1, \\ 0 & \text{for } i=-M+L,\dots,M-L, \\ h\Delta x \displaystyle\sum_{l=M-i+1}^{L} \nu(x_l)\ b_{i+l}^{n+1}, & \text{for } i=M-L+1,\dots,M-1, \end{cases}$$

where we have used the standard notation $b_i^n = b(nh, x_i)$, $i \in \mathcal{J}_M$.

In practice, we numerically solve the linear system (3.4.35) with an efficient algorithm (see next Remark 3.5.1). We notice here that a solution to (3.4.35) really exists because for $\beta \neq |\alpha|$, the matrix $A = A(y,r)$ is invertible (see e.g. Theorem 2.1 in [31]). Then, at time nh, for each fixed $y \geq 0$ and $r \in \mathbb{R}$, we approximate the solution $x \mapsto u(nh, x; y, r)$ of (3.4.30) on the points x_i's of the grid in terms of the discrete solution $u^n = \{u_i^n\}_{i\in\mathcal{J}_M}$, which in turn is written in terms of the value $u^{n+1} = \{u_i^{n+1}\}_{i\in\mathcal{J}_M}$ at time $(n+1)h$. In other words, we set

$$u(nh, x_i; y, r) \approx u_i^n,\ i \in \mathcal{J}_M, \text{ where } u^n = (u_i^n)_{i\in\mathcal{J}_M} \text{ solves (3.4.35)} \qquad (3.4.40)$$

The final local finite-difference approximation

We are now ready to tackle our original problem: the computation of the function $\Psi_f(\zeta; x, y, r)$ in (3.4.28) allowing one to numerically compute the expectation in (3.4.29). So, at time step n, the pair (y, r) is chosen on the lattice $\mathcal{Y}_n \times \mathcal{R}_n$: $y = y^n_k$, $r = r^n_j$ for $0 \le k, j \le n$. We call $A^n_{k,j}$ the matrix A in (3.4.36) when evaluated in (y^n_k, r^n_j) and d^n the boundary vector in (3.4.39) at time-step n. Then, (3.4.40) gives

$$\Psi_f(\zeta; x_i, y^n_k, r^n_j) \simeq u^n_{i,k,j}, \text{ where } u^n_{\cdot,k,j} = (u^n_{i,k,j})_{i\in\mathcal{J}_M} \text{ solves the linear system}$$
$$A^n_{k,j} u^n_{\cdot,k,j} = Bf(x_\cdot + \zeta) + d^n.$$

Therefore, by taking the expectation w.r.t. the tree-jumps, the expectation in (3.4.29) is finally computed on $\mathcal{X}_M \times \mathcal{Y}_n \times \mathcal{R}_n$ by means of the above approximation:

$$\mathbb{E}(f(\bar X^h_{(n+1)h}) \mid \bar X^h_{nh} = x_i, \bar Y^h_{nh} = y^n_k, \bar R^h_{nh} = r^n_j) \simeq u^n_{i,k,j},$$

where $u^n_{\cdot,k,j} = (u^n_{i,k,j})_{i\in\mathcal{J}_M}$ solves the linear system

$$A^n_{k,j} u^n_{\cdot,k,j} = \sum_{a,b\in\{u,d\}} p_{ab}(n,k,j) Bf\Big(x_\cdot + \frac{\rho_1}{\sigma_Y}(y^{n+1}_{k_a(n,k)} - y^n_k) + \rho_2\sqrt{y}(r^{n+1}_{j_b(n,j)} - r^n_j)\Big) + d^n.$$

Finally, if f is a function on the whole triple (x, y, r), by using standard properties of the conditional expectation one gets

$$\begin{aligned}
&\mathbb{E}(f(\bar X^h_{(n+1)h}, \bar Y^h_{(n+1)h}, \bar R^h_{(n+1)h}) \mid \bar X^h_{nh} = x_i, \bar Y^h_{nh} = y^n_k, \bar R^h_{nh} = r^n_j) \simeq u^n_{i,k,j},\\
&\text{where } u^n_{\cdot,k,j} = (u^n_{i,k,j})_{i\in\mathcal{J}_M} \text{ solves the linear system}\\
&A^n_{k,j} u^n_{\cdot,k,j}\\
&= \sum_{a,b\in\{u,d\}} p_{ab}(n,k,j) Bf\Big(x_\cdot + \frac{\rho_1}{\sigma_Y}(y^{n+1}_{k_a(n,k)} - y^n_k) + \rho_2\sqrt{y}(r^{n+1}_{j_b(n,j)} - r^n_j), y^{n+1}_{k_a(n,k)}, r^{n+1}_{j_b(n,j)}\Big) + d^n.
\end{aligned} \tag{3.4.41}$$

3.4.2 Pricing European and American options

We are now ready to approximate the function P_h solution to the dynamic programming principle (3.4.25). We consider the discretization scheme $(\bar X^h, \bar Y^h, \bar R^h)$ discussed in Section 3.4.1 and we use the approximation (3.4.41) for the conditional expectations that have to be computed at each time step n. So, for every point $(x_i, y^n_k, r^n_j) \in \mathcal{X}_M \times \mathcal{Y}_n \times \mathcal{R}_n$, by (3.4.41) we have

$$\mathbb{E}\Big(P_h\big((n+1)h, X^{nh,x_i,y^n_k,r^n_j}_{(n+1)h}, Y^{nh,y^n_k}_{(n+1)h}, R^{nh,r^n_j}_{(n+1)h}\big)\Big) \simeq u^n_{i,k,j}$$

where $u^n_{\cdot,k,j} = (u^n_{i,k,j})_{i\in\mathcal{J}_M}$ solves the linear system

$$A^n_{k,j}u^n_{\cdot,k,j} = B\sum_{a,b\in\{u,d\}} p_{ab}(n,k,j)\times$$
$$\times P_h\Big((n+1)h, y_\cdot + \frac{\rho_1}{\sigma_Y}(y^{n+1}_{k_a(n,k)} - y^n_k) + \rho_2\sqrt{y}(r^{n+1}_{j_b(n,j)} - r^n_j, y^n_k, r^n_j), y^{n+1}_{k_a(n,k)}, r^{n+1}_{j_b(n,j)}\Big) + d^n. \tag{3.4.42}$$

We then define the approximated price $\tilde{P}_h(nh, x, y, r)$ for $(x, y, r) \in \mathcal{X}_M \times \mathcal{Y}_n \times \mathcal{R}_n$ and $n = 0, 1, \dots, N$ as

$$\begin{cases} \tilde{P}_h(T, x_i, y^N_k, r^N_j) = \Psi(x_i) \quad \text{and as } n = N-1, \dots, 0: \\ \tilde{P}_h(nh, x_i, y^n_k, r^n_j) = \max\Big\{\widehat{\Psi}(x_i), e^{-(\sigma_r r^n_j + \varphi_{nh})h}\tilde{u}^n_{i,k,j}\Big\} \end{cases} \tag{3.4.43}$$

in which $\tilde{u}^n_{\cdot,k,j} = (\tilde{u}^n_{i,k,j})_{i\in\mathcal{J}_M}$ is the solution to the system in (3.4.42) with P_h replaced by $\tilde{P}_h$.

Note that the system in (3.4.42) requires the knowledge of the function $y \mapsto \tilde{P}_h((n+1)h, x, y, r)$ in points x's that do not necessarily belong to the grid $\mathcal{X}_M$. Therefore, in practice we compute such a function by means of linear interpolations, working as follows. For fixed n, k, j, a, b, we set $I_{n,k,j,a,b}(i)$, $i \in \mathcal{J}_M$, as the index such that

$$x_i + \frac{\rho_1}{\sigma_Y}(y^{n+1}_{k_a(n,k)} - y^n_k) + \rho_2\sqrt{y}(r^{n+1}_{j_b(n,j)} - r^n_j) \in [x_{I_{n,k,j,a,b}(i)}, x_{I_{n,k,j,a,b}(i)+1}),$$

with $I_{n,k,j,a,b}(i) = -M$ if $x_i + \frac{\rho_1}{\sigma_Y}(y^{n+1}_{k_a(n,k)} - y^n_k) + \rho_2\sqrt{y}(r^{n+1}_{j_b(n,j)} - r^n_j) < -M$ and $I_{n,k,j,a,b}(i) + 1 = M$ if $x_i + \frac{\rho_1}{\sigma_Y}(y^{n+1}_{k_a(n,k)} - y^n_k) + \rho_2\sqrt{y}(r^{n+1}_{j_b(n,j)} - r^n_j) > M$. We set

$$q_{n,k,j,a,b}(i) = \frac{x_i + \frac{\rho_1}{\sigma_Y}(y^{n+1}_{k_a(n,k)} - y^n_k) + \rho_2\sqrt{y}(r^{n+1}_{j_b(n,j)} - r^n_j) - x_{I_{n,k,j,a,b}(i)}}{\Delta x}.$$

Note that $q_{n,k,j,a,b}(i) \in [0, 1)$. We define

$$(\mathfrak{I}_{a,b}\tilde{P}_h)((n+1)h, x_i, y^{n+1}_{k_a(n,k)}, r^{n+1}_{j_b(n,j)}) = \tilde{P}_h((n+1)h, x_{I_{n,k,j,a,b}(i)}, y^{n+1}_{k_a(n,k)}, r^{n+1}_{j_b(n,j)})\,(1 - q_{n,k,j,a,b}(i))$$
$$+ \tilde{P}_h((n+1)h, x_{I_{n,k,j,a,b}(i)+1}, y^{n+1}_{k_a(n,k)}, r^{n+1}_{j_b(n,j)})\, q_{n,k,j,a,b}(i)$$

and we set

$$\tilde{P}_h\Big((n+1)h, x_i + \frac{\rho_1}{\sigma_Y}(y^{n+1}_{k_a(n,k)} - y^n_k) + \rho_2\sqrt{y}(r^{n+1}_{j_b(n,j)} - r^n_j), y^{n+1}_{k_a(n,k)}, r^{n+1}_{j_b(n,j)}\Big)$$
$$= (\mathfrak{I}_{a,b}\tilde{P}_h)((n+1)h, x_i, y^{n+1}_{k_a(n,k)}, r^{n+1}_{j_b(n,j)}).$$

Therefore, starting from (3.4.42), in practice the function $\tilde{u}^n_{\cdot,k,j} = (\tilde{u}^n_{i,k,j})_{i\in\mathcal{J}_M}$ in (3.4.43) is taken as the solution to the linear system

$$A^n_{k,j}\tilde{u}^n_{\cdot,k,j} = B\sum_{a,b\in\{u,d\}} p_{ab}(n,k,j)(\mathfrak{I}_{a,b}\tilde{P}_h)((n+1)h, x_\cdot, y^{n+1}_{k_a(n,k)}, r^{n+1}_{j_b(n,j)}) + d^n. \tag{3.4.44}$$

We can then state our final numerical procedure:

$$\begin{cases} \tilde{P}_h(T, x_i, y_k^N, r_j^N) = \Psi(x_i) \quad \text{and as } n = N-1, \dots, 0: \\ \tilde{P}_h(nh, x_i, y_k^n, r_j^n) = \max\left\{ \widehat{\Psi}(x_i), e^{-(\sigma_r r_j^n + \varphi_{nh})h} \tilde{u}_{i,k,j}^n \right\} \end{cases} \tag{3.4.45}$$

$\tilde{u}^n_{\cdot,k,j} = (\tilde{u}^n_{i,k,j})_{i \in \mathcal{J}_M}$ being the solution to the system (3.4.44).

Remark 3.4.1. *In the case of an infinite grid, that is $M = +\infty$, $i \mapsto I_{n,k,j,a,b}(i)$ is a translation: $I_{n,k,j,a,b}(i) = I_{n,k,j,a,b}(0) + i$. So, $x_i \mapsto (\mathfrak{I}_{a,b}\tilde{P}_h)((n+1)h, x_i, y^{n+1}_{k_a(n,k)}, r^{n+1}_{j_b(n,j)})$ is just a linear convex combination of translations of $x_i \mapsto \tilde{P}_h((n+1)h, x_i, y^{n+1}_{k_a(n,k)}, r^{n+1}_{j_b(n,j)})$.*

3.4.3 Stability analysis of the hybrid tree/finite-difference method

We analyze here the stability of the resulting tree/finite-difference scheme. To this purpose, we consider a norm, defined on functions of the variables (x, y, r), which is the uniform norm with respect to the volatility and the interest rate components (y, r) and coincides with the standard l_2 norm with respect to the direction x (see next (3.4.51)). The choice of the l_2 norm allows one to perform a von Neumann analysis in the component x on the infinite grid $\mathcal{X} = \{x_i = X_0 + i\Delta x\}_{i\in\mathbb{Z}}$, that is, without truncating the domain and without imposing boundary conditions. Therefore, our stability analysis does not take into account boundary effects. This approach is extensively used in the literature, see e.g. [45], and yields good criteria on the robustness of the algorithm independently of the boundary conditions.

Let us first write down explicitly the scheme (3.4.45) on the infinite grid $\mathcal{X} = \{x_i\}_{i\in\mathbb{Z}}$. For a fixed function $f = f(t, x, y, r)$, we set $g = f$ (in the case of American options) or $g = 0$ (in the case of European options) and we consider the numerical scheme given by

$$\begin{cases} F_h(T, x_i, y_k^N, r_j^N) = f(T, x_i, y_k^N, r_j^N) \quad \text{and as } n = N-1, \dots, 0: \\ F_h(nh, x_i, y_k^n, r_j^n) = \max\left\{ g(nh, x_i, y_k^n, r_j^n), e^{-(\sigma_r r_j^n + \varphi_{nh})h} u_{i,k,j}^n \right\} \end{cases} \tag{3.4.46}$$

where $u^n_{\cdot,k,j} = (u^n_{i,k,j})_{i\in\mathbb{Z}}$ is the solution to

$$\begin{aligned} &(\alpha_{n,k,j} - \beta_{n,k})u^n_{i-1,k,j} + (1 + 2\beta_{n,k})u^n_{i,k,j} - (\alpha_{n,k,j} + \beta_{n,k})u^n_{i+1,k,j} \\ &= \sum_{a,b\in\{d,u\}} p_{ab}(n,k,j) \times \Big[(\mathfrak{I}_{a,b}F_h)((n+1)h, x_i, y^{n+1}_{k_a(n,k)}, r^{n+1}_{j_b(n,j)}) + \\ &\quad + h\Delta x \sum_l \nu(\xi_l) \big((\mathfrak{I}_{a,b}F_h)((n+1)h, x_{i+l}, y^{n+1}_{k_a(n,k)}, r^{n+1}_{j_b(n,j)}) \\ &\quad - (\mathfrak{I}_{a,b}F_h)((n+1)h, x_i, y^{n+1}_{k_a(n,k)}, r^{n+1}_{j_b(n,j)}) \big) \Big], \end{aligned} \tag{3.4.47}$$

in which $\alpha_{n,k,j}$ and $\beta_{n,k,j}$ are the coefficients α and β defined in (3.4.37) when evaluated in the pair (y^n_k, r^n_j). Note that (3.4.47) is simply the linear system (3.4.44) on the infinite grid, with $d^n \equiv 0$ (no boundary conditions are needed). Let us stress that in next Remark 3.4.3 we will see that, since $\beta_{n,k} \geq 0$, a solution to (3.4.47) does exist, at least for "nice" functions f. It is clear that the case $g = f$ is linked to the American algorithm whereas the case $g = 0$ is connected to the European one: (3.4.46) gives our numerical approximation of the function

$$F(t,x,y,r) = \begin{cases} \mathbb{E}\Big(e^{-(\sigma_r \int_t^T R^{t,r}_s ds + \int_t^T \varphi_s ds)} f(T, X^{t,x,y,r}_T, Y^{t,y}_T, R^{t,r}_T)\Big) & \text{if } g = 0, \\ \sup\limits_{\tau \in \mathcal{T}_{t,T}} \mathbb{E}\Big(e^{-(\sigma_r \int_t^\tau R^{t,r}_s ds + \int_t^\tau \varphi_s ds)} f(\tau, X^{t,x,y,r}_\tau, Y^{t,y}_\tau, R^{t,r}_\tau)\Big) & \text{if } g = f, \end{cases} \tag{3.4.48}$$

at times nh and in the points of the grid $\mathcal{X} \times \mathcal{Y}_n \times \mathcal{R}_n$.

The "discount truncated scheme" and its stability

In our stability analysis, we consider a numerical scheme which is a slight modification of (3.4.46): we fix a (possibly large) threshold $\vartheta > 0$ and we consider the scheme

$$\begin{cases} F^\vartheta_h(T, x_i, y^N_k, r^N_j) = f(T, x_i, y^N_k, r^N_j) \quad \text{and as } n = N-1, \dots, 0: \\ F^\vartheta_h(nh, x_i, y^n_k, r^n_j) = \max\Big\{ g(nh, x_i, y^n_k, r^n_j), e^{-(\sigma_r r^n_j \mathbb{1}_{\{r^n_j > -\vartheta\}} + \varphi_{nh})h} u^n_{i,k,j} \Big\} \end{cases} \tag{3.4.49}$$

with $g = f$ or $g = 0$, where $u^n_{\cdot,k,j} = (u^n_{i,k,j})_{i \in \mathbb{Z}}$ is the solution to (3.4.47), with $(\Im_{a,b} F_h)$ replaced by $(\Im_{a,b} F^\vartheta_h)$. Let us stress that the above scheme (3.4.46) really differs from (3.4.49) only when $\sigma_r > 0$ (stochastic interest rate). And in this case, in the discounting factor of (3.4.49) we do not allow r^n_j to run everywhere on its grid: in the original scheme (3.4.46), the exponential contains the term r^n_j whereas in the present scheme (3.4.49) we put $r^n_j \mathbb{1}_{\{r^n_j > -\vartheta\}}$, so we kill the points of the grid $\mathcal{R}_n$ below the threshold $-\vartheta$. And in fact, (3.4.49) aims to numerically compute the function

$$F^\vartheta(t,x,y,r) = \begin{cases} \mathbb{E}\Big(e^{-(\sigma_r \int_t^T R^{t,r}_s \mathbb{1}_{\{R^{t,r}_s > -\vartheta\}} ds + \int_t^T \varphi_s ds)} f(T, X^{t,x,y,r}_T, Y^{t,y}_T, R^{t,r}_T)\Big) & \text{if } g = 0, \\ \sup\limits_{\tau \in \mathcal{T}_{t,T}} \mathbb{E}\Big(e^{-(\sigma_r \int_t^\tau R^{t,r}_s \mathbb{1}_{\{R^{t,r}_s > -\vartheta\}} ds + \int_t^\tau \varphi_s ds)} f(\tau, X^{t,x,y,r}_\tau, Y^{t,y}_\tau, R^{t,r}_\tau)\Big) & \text{if } g = f, \end{cases} \tag{3.4.50}$$

at times nh and in the points of the grid $\mathcal{X} \times \mathcal{Y}_n \times \mathcal{R}_n$. Recall that in practice h is small but fixed, so that the implemented scheme incorporates a threshold (see for instance the tree given in Figure 3.1). And actually, in our numerical experiments we observe a real stability. However, we will discuss later on how much one can lose with respect to the solution of (3.4.46).

For $n = N, \dots, 0$, the scheme (3.4.49) returns a function in the variables $(x, y, r) \in \mathcal{X} \times \mathcal{Y}_n \times \mathcal{R}_n$. Note that $\mathcal{Y}_n \times \mathcal{R}_n \subset I_n^Y \times I_n^R$, where

$$I_n^Y = [y_0^n, y_n^n] \quad \text{and} \quad I_n^R = [r_0^n, r_n^n],$$

that is, the intervals between the smallest and the biggest node at time-step n:

$$y_0^n = \Big(\sqrt{Y_0} - \frac{\sigma_Y}{2}\, n\sqrt{h}\Big)^2 \mathbb{1}_{\{\sqrt{Y_0} - \frac{\sigma_Y}{2}\, n\sqrt{h} > 0\}}, \qquad y_n^n = \Big(\sqrt{Y_0} + \frac{\sigma_Y}{2}\, n\sqrt{h}\Big)^2,$$

$$r_0^n = -n\sqrt{h}, \qquad r_n^n = n\sqrt{h}.$$

As n decreases to 0, the intervals I_n^Y and I_n^R are becoming smaller and smaller and at time 0 they collapse to the single point $y_0^0 = Y_0$ and $r_0^0 = R_0 = 0$ respectively. So, the norm we are going to define takes into account these facts: at time nh we consider for $\phi = \phi(t, x, y, r)$ the norm

$$\|\phi(nh, \cdot)\|_n = \sup_{(y,r) \in I_n^Y \times I_n^R} \|\phi(nh, \cdot, y, r)\|_{l_2(\mathcal{X})} = \sup_{(y,r) \in I_n^Y \times I_n^R} \Big(\sum_{i \in \mathbb{Z}} |\phi(nh, x_i, y, r)|^2 \Delta y\Big)^{\frac{1}{2}}. \qquad (3.4.51)$$

In particular,

$$\|\phi(0, \cdot)\|_0 = \|\phi(0, \cdot, Y_0, R_0)\|_{l_2(\mathcal{X})} = \Big(\sum_{i \in \mathbb{Z}} |\phi(x_i, Y_0, R_0)|^2 \Delta y\Big)^{1/2} \quad \text{and}$$

$$\|\phi(T, \cdot)\|_N \le \sup_{(y,r) \in \mathbb{R}_+ \times \mathbb{R}} \|\phi(x_i, y, r)\|_{l_2(\mathcal{X})} = \sup_{(y,r) \in \mathbb{R}_+ \times \mathbb{R}} \Big(\sum_{i \in \mathbb{Z}} |\phi(x_i, y, r)|^2 \Delta y\Big)^{1/2}.$$

We are now ready to give our stability result.

Theorem 3.4.2. *Let $f \ge 0$ and, in the case $g = f$, suppose that*

$$\sup_{t \in [0,T]} |f(t, x, y, r)| \le \gamma_T |f(T, x, y, r)|,$$

for some $\gamma_T > 0$. Then, for every $\vartheta > 0$ the numerical scheme (3.4.49) *is stable with respect to the norm* (3.4.51)*:*

$$\|F_h^\vartheta(0, \cdot)\|_0 \le C_T^{N,\vartheta} \|F_h^\vartheta(T, \cdot)\|_N = C_T^{N,\vartheta} \|f(T, \cdot)\|_N, \quad \forall h, \Delta y,$$

where

$$C_T^{N,\vartheta} = \begin{cases} e^{2\lambda c T + \sigma_r \vartheta T - \sum_{n=1}^N \varphi_{nh} h} \stackrel{N \to \infty}{\longrightarrow} C_T^\vartheta = e^{2\lambda c T + \sigma_r \vartheta T - \int_0^T \varphi_t dt} & \text{if } g = 0, \\ \max\Big\{\gamma_T, e^{2\lambda c T + \sigma_r \vartheta T - \sum_{n=1}^N \varphi_{nh} h}\Big\} \stackrel{N \to \infty}{\longrightarrow} C_T^\vartheta = \max\Big\{\gamma_T, e^{2\lambda c T + \sigma_r \vartheta T - \int_0^T \varphi_t dt}\Big\} & \text{if } g = f, \end{cases}$$

in which $c > 0$ is such that $\sum_l \nu(\xi_l) \Delta x \le \lambda c$. In the standard Bates model, that is $\sigma_r = 0$ and deterministic interest rate $r_t = \varphi_t$, the discount truncated scheme (3.4.49) *coincides with the standard scheme* (3.4.45) *and the stability follows for* (3.4.45).

Proof. In order to simplify the notation, we set $g^n_{i,k,j} = g(nh, x_i, y^n_k, r^n_j)$ and, similarly, $F^n_{i,k,j} = F^\vartheta_h(nh, x_i, y^n_k, r^n_j)$, $(\mathfrak{I}_{a,b}F^{n+1}_h)_{i,k_a,j_b} = (\mathfrak{I}_{a,b}F^\vartheta_h)((n+1)h, x_i, y^{n+1}_{k_a(n,k)}, r^{n+1}_{j_b(n,j)})$ (we have also dropped the dependence on ϑ). The scheme (3.4.49) says that, at each time step $n < N$ and for each fixed $0 \le k, j \le n$,

$$F^n_{i,k,j} = \max\Big\{g^n_{i,k,j}, e^{-(\sigma_r r^n_j \mathbb{1}_{\{r^n_j > -\vartheta\}} + \varphi_{nh})h} u^n_{i,k,j}\Big\}, \tag{3.4.52}$$

where, according to (3.4.47), $u^n_{i,k,j}$ solves

$$\begin{aligned}&(\alpha_{n,k,j} - \beta_{n,k})u^n_{i-1,k,j} + (1 + 2\beta_{n,k})u^n_{i,k,j} - (\alpha_{n,k,j} + \beta_{n,k})u^n_{i+1,k,j}\\ &= \sum_{a,b\in\{d,u\}} p_{ab}(n,k,j)\Big((\mathfrak{I}_{a,b}F^{n+1})_{i,k_a,j_b} + h\Delta x \sum_l \nu(\xi_l)\big[(\mathfrak{I}_{a,b}F^{n+1})_{i+l,k_a,j_b} - (\mathfrak{I}_{a,b}F^{n+1})_{i,k_a,j_b}\big]\Big).\end{aligned} \tag{3.4.53}$$

Let $\mathfrak{F}\varphi$ denote the Fourier transform of $\varphi \in l_2(\mathcal{X})$, that is,

$$\mathfrak{F}\varphi(\theta) = \frac{\Delta x}{\sqrt{2\pi}} \sum_{s\in\mathbb{Z}} \varphi_s e^{-\mathbf{i}\, s\Delta y\theta}, \quad \theta \in \mathbb{R},$$

$\mathbf{i}$ denoting the imaginary unit. We get from (3.4.53)

$$\begin{aligned}&\Big((\alpha_{n,k,j} - \beta_{n,k})e^{-\mathbf{i}\,\theta\Delta x} + 1 + 2\beta_{n,k} - (\alpha_{n,k,j} + \beta_{n,k})e^{\mathbf{i}\,\theta\Delta x}\Big)\mathfrak{F}u^n_{k,j}(\theta)\\ &= \Big(1 + h\Delta x \textstyle\sum_l \nu(\xi_l)(e^{\mathbf{i}\,l\theta\Delta x} - 1)\Big) \textstyle\sum_{a,b\in\{d,u\}} p_{ab}(n,k,j)\mathfrak{F}(\mathfrak{I}_{a,b}F^{n+1})_{k_a,j_b}(\theta).\end{aligned} \tag{3.4.54}$$

Note that

$$\begin{aligned}|(\alpha_{n,k,j} - \beta_{n,k})e^{-\mathbf{i}\,\theta\Delta x} + 1 + 2\beta_{n,k} - (\alpha_{n,k,j} + \beta_{n,k})e^{\mathbf{i}\,\theta\Delta x}|&\\ \ge \big|\mathfrak{Re}\big[(\alpha_{n,k,j} - \beta_{n,k})e^{-\mathbf{i}\,\theta\Delta x} + 1 + 2\beta_{n,k} - (\alpha_{n,k,j} + \beta_{n,k})e^{\mathbf{i}\,\theta\Delta x}\big]\big|&\\ = 1 + 2\beta_{n,k}(1 - \cos(\theta\Delta x)) \ge 1,&\end{aligned}$$

for every $\theta \in [0, 2\pi)$ (recall that $\beta_{n,k} \ge 0$). And since $\sum_l \nu(\xi_l)\Delta x \le \lambda c$, we obtain

$$\begin{aligned}|\mathfrak{F}u^n_{k,j}(\theta)| &\le \Big(1 + h\Delta x \sum_{l\in\mathbb{Z}} |e^{\mathbf{i}\,l\theta\Delta x} - 1|\nu(\xi_l)\Big) \sum_{a,b\in\{d,u\}} p_{ab}(n,k,j)|\mathfrak{F}(\mathfrak{I}_{a,b}F^{n+1})_{k_a,j_b}(\theta)|\\ &\le (1 + 2\lambda ch) \sum_{a,b\in\{d,u\}} p_{ab}(n,k,j)|\mathfrak{F}(\mathfrak{I}_{a,b}F^{n+1})_{k_a,j_b}(\theta)|.\end{aligned}$$

Therefore,

$$\|\mathfrak{F}u^n_{k,j}\|_{L^2([0,2\pi),\mathrm{Leb})} \le (1 + 2\lambda ch) \sum_{a,b\in\{d,u\}} p_{ab}(n,k,j)\|\mathfrak{F}(\mathfrak{I}_{a,b}F^{n+1})_{k_a,j_b}\|_{L^2([0,2\pi),\mathrm{Leb})}.$$

We use now the Parseval identity $\|\mathfrak{F}\varphi\|_{L^2([0,2\pi),\mathrm{Leb})} = \|\varphi\|_{l_2(\mathcal{X})}$ and we get

$$\begin{aligned}\|u^n_{\cdot,k,j}\|_{l^2(\mathcal{X})} &\leq (1+2\lambda ch) \sum_{a,b\in\{d,u\}} p_{ab}(n,k,j)\|(\mathfrak{I}_{a,b}F^{n+1})_{\cdot,k_a,j_b}\|_{l^2(\mathcal{X})}\\ &= (1+2\lambda ch) \sum_{a,b\in\{d,u\}} p_{ab}(n,k,j)\|F^{n+1}_{\cdot,k_a,j_b}\|_{l^2(\mathcal{X})},\end{aligned}$$

the first equality following from the fact that $i \mapsto (\mathfrak{I}_{a,b}F^{n+1})_{i,k_a,j_b}$ is a linear convex combination of translations of $i \mapsto F^{n+1}_{i,k_a,j_b}$ (see Remark 3.4.1). This gives

$$\sup_{0\leq k,j\leq n} \|e^{-(\sigma_r r^n_j \mathbb{1}_{\{r^n_j>-\vartheta\}}+\varphi_{nh})h} u^n_{\cdot,k,j}\|_{l_2(\mathcal{X})} \leq (1+2\lambda ch)e^{\sigma_r\vartheta h-\varphi_{nh}h} \sup_{0\leq k,j\leq n+1} \|F^{n+1}_{\cdot,k,j}\|_{l_2(\mathcal{X})}$$

and from (3.4.52), we obtain

$$\sup_{0\leq k,j\leq n} \|F^n_{\cdot,k,j}\|_{l_2(\mathcal{X})} \leq \max\Big(\sup_{0\leq k,j\leq n} \|g^n_{\cdot,k,j}\|_{l_2(\mathcal{X})}, (1+2\lambda ch)e^{\sigma_r\vartheta h-\varphi_{nh}h} \sup_{0\leq k,j\leq n+1} \|F^{n+1}_{\cdot,k,j}\|_{l_2(\mathcal{X})}\Big).$$

We now continue assuming that $g = f$, the case $g = 0$ following in a similar way. So,

$$\sup_{0\leq k,j\leq n} \|F^n_{\cdot,k,j}\|_{l_2(\mathcal{X})} \leq \max\Big(\gamma_T\|f(T,\cdot)\|_N, (1+2\lambda ch)e^{\sigma_r\vartheta h-\varphi_{nh}h} \sup_{0\leq k,j\leq n+1} \|F^{n+1}_{\cdot,k,j}\|_{l_2(\mathcal{X})}\Big).$$

For $n = N-1$ we then obtain

$$\sup_{0\leq k,j\leq n} \|F^{N-1}_{\cdot,k,j}\|_{l_2(\mathcal{X})} \leq \max\Big(\gamma_T\|f(T,\cdot)\|_N, (1+2\lambda ch)e^{\sigma_r\vartheta h-\varphi_{(N-1)h}h}\|f(T,\cdot)\|_N\Big)$$

and by iterating the above inequalities, we finally get

$$\|F^0\|_0 = \|F^0_{\cdot,0,0}\|_{l_2(\mathcal{X})} \leq \max\Big(\gamma_T\|f(T,\cdot)\|_N, (1+2\lambda ch)^N e^{N\sigma_r Lh-\sum_{n=1}^N \varphi_{nh}h}\|f(T,\cdot)\|_N\Big).$$

□

Remark 3.4.3. *We have incidentally proved that, as n varies, the solution* $u^n_{\cdot,k,j}$ *to the infinite linear system* (3.4.47) *actually exists and is unique if* $\|f(T,\cdot)\|_N < \infty$*. In fact, starting from equality* (3.4.54)*, we define the function* $\psi_{k,j}(\theta)$*,* $\theta \in [0,2\pi)$*, by*

$$\begin{aligned}&\Big((\alpha_{n,k,j}-\beta_{n,k})e^{-\mathbf{i}\,\theta\Delta x}+1+2\beta_{n,k}-(\alpha_{n,k,j}+\beta_{n,k})e^{\mathbf{i}\,\theta\Delta x}\Big)\psi_{k,j}(\theta)\\ &= \Big(1+h\Delta x\textstyle\sum_l \nu(\xi_l)(e^{\mathbf{i}\,l\theta\Delta x}-1)\Big)\textstyle\sum_{a,b\in\{d,u\}} p_{ab}(n,k,j)\mathfrak{F}(\mathfrak{I}_{a,b}F^{n+1})_{k_a,j_b}(\theta).\end{aligned}$$

As noticed in the proof of Proposition 3.4.2, the factor multiplying $\psi_{k,j}(\theta)$ *is different from zero because* $\beta_{n,k} \geq 0$. *So, the definition of* $\psi_{k,j}$ *is well posed and moreover,* $\psi_{k,j} \in L^2([0, 2\pi,), \mathrm{Leb})$. *We now set* $u^n_{\cdot,k,j}$ *as the inverse Fourier transform of* $\psi_{k,j}$*, that is,*

$$u^n_{l,k,j} = \frac{1}{\Delta y\sqrt{2\pi}} \int_0^{2\pi} \psi_{k,j}(\theta) e^{\mathbf{i}\, l\theta\Delta y} d\theta, \quad l \in \mathbb{Z}.$$

Straightforward computations give that $u^n_{\cdot,k,j}$ *fulfils the equation system* (3.4.47).

Of course, Theorem 3.4.2 gives a stability property for the scheme introduced in [25] for the Heston-Hull-White model: just take $\lambda = 0$ (no jumps are considered).

Back to the original scheme (3.4.46)

Let us now discuss what may happen when one introduces the threshold ϑ. We recall that the original scheme (3.4.46) gives the numerical approximation of the function F in (3.4.48) whereas the discount truncated scheme (3.4.49) aims to numerically compute the function F^ϑ in (3.4.50). Proposition 3.4.4 below shows that, under standard hypotheses, F^ϑ tends to F as $\vartheta \to \infty$ very fast. This means that, in practice, we lose very few in using (3.4.49) in place of (3.4.46).

Proposition 3.4.4. *Suppose that* $f = f(t, x, y, r)$ *has a polynomial growth in the variables* (x, y, r)*, uniformly in* $t \in [0, T]$. *Let* F *and* F^ϑ*, with* $\vartheta > 0$*, be defined in* (3.4.48) *and* (3.4.50) *respectively. Then there exist positive constants* c_T *and* $C_T(x, y, r)$ *(depending on* (x, y) *in a polynomial way and on* r *in an exponential way) such that for every* $\vartheta > 0$

$$|F(t, x, y, r) - F^\vartheta(t, x, y, r)| \leq \sigma_r C_T(x, y, r) e^{-c_T|\vartheta + x e^{-\kappa_r(T-t)}|^2},$$

for every $t \in [0, T]$ *and* $(x, y, r) \in \mathbb{R} \times \mathbb{R}_+ \times \mathbb{R}$.

Proof. In the following, C denotes a positive constant, possibly changing from line to line, which depends on (x, y, r) polynomially in (x, y) and exponentially in r. We have

$$\begin{aligned} &|F(t, x, y, r) - F^\vartheta(t, x, y, r)| \\ &\leq C\mathbb{E}\left(\sup_{t \leq u \leq T} |f(u, X^{t,x,y,r}_u, Y^{t,y}_u, R^{t,r}_u)| \times e^{-\sigma_r \int_t^u R^{t,r}_s \mathbb{1}_{\{R^{t,r}_s > -\vartheta\}} ds} \times \left(e^{-\sigma_r \int_t^u R^{t,r}_s \mathbb{1}_{\{R^{t,r}_s < -\vartheta\}} ds} - 1 \right) \right). \end{aligned} \tag{3.4.55}$$

Set now

$$\tau^{t,r}_{-\vartheta} = \inf\{s \geq t \,:\, R^{t,r}_s \leq -\vartheta\}.$$

Notice that $\{R_s < -\theta\} \subseteq \{\tau_{-\theta} < s\} \subseteq \{\tau_{-\theta} < T\}$. Therefore, one has $\mathbb{1}_{\{R_s^{t,r} < -\vartheta\}} \leq \mathbb{1}_{\{\tau_{-\vartheta}^{t,r} < T\}}$ and

$$-\sigma_r \int_t^u R_s^{t,r} \mathbb{1}_{\{R_s^{t,r} < -\vartheta\}} ds = \int_t^u |\sigma_r R_s^{t,r}| \mathbb{1}_{\{R_s^{t,r} < -\vartheta\}} ds \leq \sigma_r \mathbb{1}_{\{\tau_{-\vartheta}^{t,r} < T\}} \int_t^u |R_s^{t,r}| ds.$$

So we can write

$$0 \leq e^{-\sigma_r \int_t^u R_s^{t,r} \mathbb{1}_{\{R_s^{t,r} < -\vartheta\}} ds} - 1 \leq e^{\sigma_r \mathbb{1}_{\{\tau_{-\vartheta}^{t,r} < T\}} \int_t^u |R_s^{t,r}| ds} - 1 = \left(e^{\sigma_r \int_t^u |R_s^{t,r}| ds} - 1 \right) \mathbb{1}_{\{\tau_{-\vartheta}^{t,r} < T\}}$$

Substituting in (3.4.55) and applying Hölder inequality, we get

$$\begin{aligned}
&|F(t,x,y,r) - F^{\vartheta}(t,x,y,r)| \\
&\leq C\mathbb{E}\left(\sup_{t \leq u \leq T} |f(u, X_u^{t,x,y,r}, Y_u^{t,y}, R_u^{t,r})| e^{-\sigma_r \int_t^u R_s^{t,r} \mathbb{1}_{\{R_s^{t,r} > -\vartheta\}} ds} \left(e^{\sigma_r \int_t^u |R_s^{t,r}| ds} - 1 \right) \mathbb{1}_{\{\tau_{-\vartheta}^{t,r} < T\}} \right) \\
&\leq C\mathbb{E}\left(\sup_{t \leq u \leq T} |f(u, X_u^{t,x,y,r}, Y_u^{t,y}, R_u^{t,r})|^2 e^{2\sigma_r \int_t^u |R_s^{t,r}| ds} \left(e^{\sigma_r \int_t^u |R_s^{t,r}| ds} - 1 \right)^2 \right)^{1/2} \times \\
&\qquad \mathbb{P}\left(\mathbb{1}_{\{\tau_{-\vartheta}^{t,r} < T\}} \right)^{1/2} \\
&\leq C\mathbb{E}\left(\sup_{t \leq u \leq T} |f(u, X_u^{t,x,y,r}, Y_u^{t,y}, R_u^{t,r})|^2 \times e^{4\sigma_r \int_t^T |R_s^{t,r}| ds} \right)^{1/2} \times \mathbb{P}\left(\mathbb{1}_{\{\tau_{-\vartheta}^{t,r} < T\}} \right)^{1/2} \\
&\leq C\mathbb{E}\left(\sup_{t \leq u \leq T} |f(u, X_u^{t,x,y,r}, Y_u^{t,y}, R_u^{t,r})|^4 \right)^{1/4} \times \mathbb{E}\left(e^{8\sigma_r \int_t^T |R_s^{t,r}| ds} \right)^{1/4} \times \mathbb{P}\left(\mathbb{1}_{\{\tau_{-\vartheta}^{t,r} < T\}} \right)^{1/2}. \qquad (3.4.56)
\end{aligned}$$

The first term in the left hand side of (3.4.56) is finite since f has polynomial growth in the space variables, uniformly in the time variable, and by using standard estimates. Also the second term in (3.4.56) is finite. This is because, for every $c > 0$,

$$\mathbb{E}\left(e^{c \sup_{t \leq s \leq T} |R_s^{t,r}|} \right) < \infty. \qquad (3.4.57)$$

In fact, recalling that that $R_s^{t,r} = re^{-\kappa_r(s-t)} + \int_t^s e^{-\kappa_r(s-u)} dW_u^2$, (3.4.57) follows from the fact that, for a Brownian motion W, $\sup_{0 \leq s \leq T} |W_s|$ has finite exponential moments of any order, for every $T > 0$. This is true since $\sup_{0 \leq s \leq T} |W_s| \leq \sup_{0 \leq s \leq T} W_s + \sup_{0 \leq s \leq T}(-W_s)$ and $\mathbb{E}(e^{p \sup_{0 \leq s \leq T} W_s})$ $< \infty$ for every $p > 0$. As regards the third term in (3.4.56), note that

$$\begin{aligned}
\mathbb{P}(\tau_{-\vartheta}^{t,r} \leq T) &= \mathbb{P}(\inf_{s \in [t,T]} R_s^{t,r} < -\vartheta) = \mathbb{P}\Big(\inf_{s \in [t,T]} \left(re^{-\kappa_r(s-t)} + \int_t^s e^{-\kappa_r(s-u)} dW_u^2 \right) < -\vartheta \Big) \\
&\leq \mathbb{P}\Big(\sup_{s \in [t,T]} \Big| \int_t^s e^{\kappa_r u} dW_u^2 \Big| > \vartheta + re^{-\kappa_r(T-t)} \Big) \leq 2 \exp\Big(- \frac{|\vartheta + re^{-\kappa_r(T-t)}|^2}{2 \int_t^T e^{2\kappa_r u} du} \Big).
\end{aligned}$$

By inserting the above estimates in (3.4.56), we get the result.

□

Further remarks

As already stressed, the introduction of the threshold $-\vartheta$ allows one to handle the discount term. In order to get rid of the discount, a possible approach consists in the use of a transformed function, as developed by several authors (see e.g. Haentjens and in't Hout [56] and references therein). This is a nice fact for European options (PIDE problem), being on the contrary a non definitive tool when dealing with American options (obstacle PIDE problem). Let us see why.

First of all, let us come back to the model for the triple (X, Y, R), see (3.2.5). The infinitesimal generator is

$$\begin{aligned}\mathcal{L}_t u = \ & \Big(\sigma_r r + \varphi_t - \delta - \frac{1}{2}y\Big)\partial_x u + \kappa_Y(\theta_Y - y)\partial_y u - \kappa_r r \partial_r u \\ & + \frac{1}{2}\Big(y\partial^2_{xx}u + \sigma_Y^2 y \partial^2_{yy}u + \partial^2_{rr}u + 2\rho_1\sigma_Y y \partial^2_{xy}u + 2\rho_2\sqrt{y}\,\partial^2_{xr}u\Big) \\ & + \int_{-\infty}^{+\infty}\left[u(t, x+\xi; y, r) - u(t, x; y, r)\right]\nu(\xi)d\xi.\end{aligned} \tag{3.4.58}$$

We set

$$G(t, r) = \mathbb{E}\Big(e^{-\sigma_r \int_t^T R_s^{t,r} ds}\Big)$$

and we recall several known facts: one has (see e.g. [72])

$$G(t, r) = e^{-r\sigma_r\Lambda(t,T) - \frac{\sigma_r^2}{2\kappa_r^2}(\Lambda(t,T) - T + t) - \frac{\sigma_r^2}{4\kappa_r}\Lambda^2(t,T)}, \quad \Lambda(t, T) = \frac{1 - e^{-\kappa_r(T-t)}}{\kappa_r} \tag{3.4.59}$$

and moreover, G solves the PDE

$$\begin{aligned}& \partial_t G - \kappa_r x \partial_x G + \frac{1}{2}\partial^2_{rr}G - \sigma_r r G = 0, \quad t \in [0, T), r \in \mathbb{R}, \\ & G(T, r) = 1.\end{aligned} \tag{3.4.60}$$

Lemma 3.4.5. *Let* $\mathcal{L}_t$ *denote the infinitesimal generator in* (3.4.58). *Set* $\overline{u} = u \cdot G^{-1}$. *Then*

$$\partial_t u + \mathcal{L}_t u - ru = G\big(\partial_t \overline{u} + \overline{\mathcal{L}}_t \overline{u}\big),$$

where

$$\overline{\mathcal{L}}_t = \mathcal{L}_t - \sigma_r \frac{1 - e^{-\kappa_r(T-t)}}{\kappa_r}\big[\rho_2\sqrt{y}\partial_x\overline{u} + \partial_r\overline{u}\big].$$

Proof. Since G depends on t and r only, straightforward computations give

$$\partial_t u + \mathcal{L}_t u - xu = G\big[\partial_t\overline{u} + \mathcal{L}_t\overline{u}\big] + \partial_r G(t, r)\big[\rho_2\sqrt{y}\partial_x\overline{u} + \partial_r\overline{u}\big] + \overline{u}\big[\partial_t G - \kappa_r r \partial_r G + \frac{1}{2}\partial^2_{rr}G - \sigma_r r G\big].$$

By (3.4.60), the last term is null. The statement now follows by observing that $\partial_r \ln G(t,r) = -\sigma_r \frac{1-e^{-\kappa_r(T-t)}}{\kappa_r}$. $\square$

We notice that the operator $\overline{L}_t$ in Lemma 3.4.5 is the infinitesimal generator of the jump-diffusion process $(\overline{X}, \overline{Y}, \overline{R})$ which solves the stochastic differential equation as in (3.2.5), with the same diffusion coefficients and jump-terms but with the new drift coefficients

$$\mu_{\overline{X}}(t,y,r) = \mu_X(y,r) - \sigma_r \frac{1-e^{-\kappa_r(T-t)}}{\kappa_r}\rho_2\sqrt{y}, \qquad \mu_{\overline{Y}}(y) \equiv \mu_Y(y),$$

$$\mu_{\overline{R}}(r) = \mu_R(t,r) - \sigma_r \frac{1-e^{-\kappa_r(T-t)}}{\kappa_r}.$$

Let us first discuss the scheme (3.4.46) with $g = 0$ (European options), which gives the numerical approximation for the function F in (3.4.48). By passing to the associated PIDE, Lemma 3.4.5 says that

$$F(t,x,y,r) = G(t,r)\overline{F}(t,x,y,r),$$

where

$$\overline{F}(t,x,y,r) = \mathbb{E}(e^{-\int_t^T \varphi_s ds} f(T, \overline{X}_T^{t,x,y,r}, \overline{Y}_T^{t,y}, \overline{R}_T^{t,r})).$$

Therefore, in practice one has to numerically evaluate the function $\overline{F}$. By using our hybrid tree/finite-difference approach, this means to consider the scheme in (3.4.49), with the new coefficient $\overline{\alpha}_{n,k,j}$ (written starting from the new drift coefficients) but with a discount depending on the (deterministic) function φ only, that is, with $e^{-(\sigma_r r_j^n \mathbb{1}_{\{r_j^n > -L\}} + \varphi_{nh})h}$ replaced by $e^{-\varphi_{nh}h}$. And the proof of the Proposition 3.4.2 shows that one gets

$$\|\overline{F}_h(0,\cdot)\|_0 \leq \max\big(\gamma_T, e^{2\lambda c T - \sum_{n=0}^N \varphi_{nh} h}\big)\|f(T,\cdot)\|_N.$$

In other words, by using a suitable transformation, the European scheme is always stable and no thresholds are needed.

Let us discuss now the American case, that is, the scheme (3.4.46) with $g = f$, giving an approximation of the function F in (3.4.48). One could think to use the above transformation in order to get rid of the exponential depending on the process R. Set again

$$\overline{F}(t,x,y,r) = G(t,r)^{-1}F(t,x,y,r).$$

By using the associated obstacle PIDE problem, Lemma 3.4.5 suggests that

$$\overline{F}(t,x,y,r) = \sup_{\tau \in \mathcal{T}_{t,T}} \mathbb{E}(e^{-\int_t^\tau \varphi_s ds} \overline{f}(\tau, \overline{X}_\tau^{t,x,y,r}, \overline{Y}_\tau^{t,y}, \overline{R}_\tau^{t,r})),$$

with $\overline{f}(t,x,y,r) = G^{-1}(t,r)f(t,x,y,r)$. So, in order to numerically compute $\overline{F}$, one needs to set up the scheme (3.4.49) with the new coefficient $\overline{\alpha}_{n,k,j}$, with f replaced by $\overline{f}$, $g = \overline{f}$ and with the discounting factor $e^{-(\sigma_r r_j^n \mathbb{1}_{\{r_j^n > -L\}} + \varphi_{nh})h}$ replaced by $e^{-\varphi_{nh} h}$. So, again one is able to cancel the unbounded part of the discount. Nevertheless, the unpleasant point is that even if $\|f(T,\cdot)\|_N$ has a bound which is uniform in N then $\|\overline{f}(T,\cdot)\|_N$ may not have because $G^{-1}(t,r)$ has an exponential containing r, see (3.4.59). In other words, the unboundedness problem appears now in the obstacle.

3.5 The hybrid Monte Carlo and tree/finite-difference approach algorithms in practice

The present section is devoted to our numerical experiments. We first summarise the main steps of our algorithms and then we present several numerical tests.

3.5.1 A schematic sketch of the main computational steps in our algorithms

In short, we outline here the main computational steps of the two proposed algorithms.

First, the procedures need the following preprocessing steps, concerning the construction of the bivariate tree:

(T1) define a discretization of the time-interval $[0,T]$ in N subintervals $[nh,(n+1)h]$, $n = 0,\dots,N-1$, with $h = T/N$;

(T2) for the process Y, set the binomial tree y_k^n, $0 \le k \le n \le N$, by using (3.3.15), then compute the jump nodes $k_a(n,k)$ and the jump probabilities $p_a^Y(n,k)$, $a \in \{u,d\}$, by using (3.3.12)-(3.3.13) and (3.3.14);

(T3) for the process R, set the binomial tree r_j^n, $0 \le j \le N$, by using (3.3.15), then compute the jump nodes $j_b(n,j)$ and the jump probabilities $p_b^R(n,j)$, $b \in \{u,d\}$, by using (3.3.16)-(3.3.17) and (3.3.18);

(T4) for the 2-dimensional process (Y,R), merge the binomial trees in the bivariate tree (y_k^n, r_j^n), $0 \le k,j \le n \le N$, by using (3.3.19), then compute the jump-nodes $(k_a(n,k), j_b(n,j))$ and the transitions probabilities $p_{ab}(n,k,j)$, $(a,b) \in \{d,u\}$, by using (3.3.20).

The bivariate tree for (Y,R) is now settled. Our hybrid tree/finite-difference algorithm can be resumed as follows:

(FD1) set a mesh grid x_i for the solution of all the PIDE's;

(FD2) for each node (y_k^N, r_j^N), $0 \le k, j \le N$, compute the option prices at maturity for each x_i, $i \in \mathcal{X}_M$, by using the payoff function;

(FD3) for $n = N-1, \dots 0$: for each (y_k^n, r_j^n), $0 \le k, j \le n$, compute the option prices for each $x_i \in \mathcal{X}_M$, by solving the linear system (3.4.44).

Notice that, at each time step n, we need only the one-step PIDE solution in the time interval $[nh, (n+1)h]$. Moreover, both the (constant) PIDE coefficients and the Cauchy final condition change according to the position of the volatility and the interest rate components on the bivariate tree at time step n.

Remark 3.5.1. *We observe that in order to compute the option price by the hybrid tree/finite-difference procedure, in step* (FD3) *we need to solve many times the tridiagonal system* (3.4.44)*. This is typically solved by the LU-decomposition method in $O(M)$ operations (recall that the total number of the grid values $x_i \in \mathcal{X}_M$ is $2M+1$). However, due to the approximation of the integral term* (3.4.33)*, at each time step $n < N$ we have to compute the sum*

$$\sum \tilde{u}_{i+l}^{n+1} \nu(\xi_l), \tag{3.5.61}$$

which is the most computationally expensive step of this part of the algorithm: when applied directly, it requires $O(M^2)$ operations. Following the Premia software implementation [84], in our numerical tests we use the Fast Fourier Transform to compute the term (3.5.61) *and the computational costs of this step reduce to $O(M \log M)$.*

We conclude by briefly recalling the main steps of the hybrid Monte Carlo method:

(MC1) let the chain $(\hat{Y}_n^h, \hat{R}_n^h)$ evolve for $n = 1, \dots, N$, following the probability structure in (T4);

(MC2) generate $\Delta_1, \dots, \Delta_N$ i.i.d. standard normal r.v.'s independent of the noise driving the chain $(\hat{Y}^h, \hat{R}^h)$;

(MC3) generate $K_h^1, \dots, K_h^N$ i.i.d. positive Poisson r.v.'s of parameter λh, independent of both the chain $(\hat{Y}^h, \hat{R}^h)$ and the Gaussian r.v.'s $\Delta_1, \dots, \Delta_N$, and for every $n = 1, \dots, N$, if $K_h^n > 0$ simulate the corresponding amplitudes $\log(1 + J_1^n), \dots, \log(1 + J_{K_h^n}^n)$;

(MC4) starting from $\hat{X}_0^h = X_0$, compute the approximate values $\hat{X}_n^h$, $1 \le n \le N$, by using (3.3.24);

(MC5) following the desired Monte Carlo method (European or Longstaff-Schwartz algorithm [76] in the case of American options), repeat the above simulation scheme and compute the option price.

Remark 3.5.2. *In Section 3.5.2 we develop numerical experiments in order to study the behavior of our hybrid methods. Our tests involve also the standard Bates model, that is without any randomness in the interest rate. Recall that in the standard Bates model the dynamic reduces to*

$$\begin{aligned} \frac{dS_t}{S_{t^-}} &= (r-\delta)dt + \sqrt{Y_t}\, dZ_t^S + dH_t, \\ dY_t &= \kappa_Y(\theta_Y - Y_t)dt + \sigma_Y \sqrt{Y_t}\, dZ_t^Y, \end{aligned} \tag{3.5.62}$$

with $S_0 > 0$, $Y_0 > 0$ and $r \geq 0$ constant parameters. We assume a correlation between the two Brownian noises:

$$d\langle Z^S, Z^Y\rangle_t = \rho dt, \quad |\rho| < 1.$$

Finally, H_t is the compound Poisson process already introduced in Section 3.2, see (3.2.2)*. We can apply our hybrid approach to this case as well: it just suffices to follow the computational steps listed above except for the construction of the binomial tree for the process R. Consequently, we do not need the bivariate tree for (Y, R), specifically we omit steps* (T3)-(T4) *and we replace step* (MC1) *with*

(MC1') *let the chain $\hat{Y}^h_n$ evolve for $n = 1, \dots, N$, following the probability structure in* (T2)*.*

And of course, in all computations we set equal to 0 the parameters involved in the dynamics for r, except for the starting value r_0. In particular, we have $\sigma_r = 0$ and $\varphi_t = r_0$ for every t.

3.5.2 Numerical results

We develop several numerical results in order to assess the efficiency and the robustness of the hybrid tree/finite-difference method and the hybrid Monte Carlo method in the case of plain vanilla options. The Monte Carlo results derive from our hybrid simulations and, for American options, the use of the Monte Carlo algorithm by Longstaff and Schwartz in [76].

We first provide results for the standard Bates model (see Remark 3.5.2) and secondly, for the case in which the interest rate process is assumed to be stochastic, see (3.2.1).

Following Chiarella *et al.* [34], in our numerical tests we assume that the jumps for the log-returns are normal, that is,

$$\log(1 + J_1) \sim N\Big(\gamma - \frac{1}{2}\eta^2, \eta^2\Big), \tag{3.5.63}$$

N denoting the Gaussian law (we also notice that the results in [34] correspond to the choice $\gamma = 0$). In Section 3.5.2, we first compare our results with the ones provided in Chiarella *et al.* [34]. Then in Section 3.5.2 we study options with large maturities and when the Feller condition is not fulfilled. Finally, Section 3.5.2 is devoted to test experiments for European and American options in the Bates model with stochastic interest rate. The codes have been written by using the C++ language and the computations have all been performed in double precision on a PC 2,9 GHz Intel Core I5 with 8 Gb of RAM.

The standard Bates model

We refer here to the standard Bates model as in (3.5.62). In the European and American option contracts we are dealing with, we consider the following set of parameters, already used in the numerical results provided in Chiarella *et al.* [34]:

- initial price $S_0 = 80, 90, 100, 110, 120$, strike price $K = 100$, maturity $T = 0.5$;
- (constant) interest rate $r = 0.03$, dividend rate $\delta = 0.05$;
- initial volatility $Y_0 = 0.04$, long-mean $\theta_Y = 0.04$, speed of mean-reversion $\kappa_Y = 2$, vol-vol $\sigma_Y = 0.4$, correlation $\rho = -0.5, 0.5$;
- intensity $\lambda = 5$, jump parameters $\gamma = 0$ and $\eta = 0.1$ (recall (3.5.63)).

It is known that the case $\rho > 0$ may lead to moment explosion, see. e.g. [9]. Nevetheless, we report here results for this case as well, for the sake of comparisons with the study in Chiarella *et al.* [34].

In order to numerically solve the PIDE using the finite difference scheme, we first localize the variables and the integral term to bounded domains. We use for this purpose the estimates for the localization domain and the truncation of large jumps given by Yoltchkova and Tankov [96]. For example, for the previous model parameters the PIDE problem is solved in the finite interval $[\ln S_0 - 1.59, \ln S_0 + 1.93]$.

The numerical study of the hybrid tree/finite-difference method **HTFD** is split into two cases:

- **HTFDa**: time steps $N_t = 50$ and varying mesh grid $\Delta x = 0.01$, 0.005, 0.0025, 0.00125;
- **HTFDb**: time steps $N_t = 100$ and varying mesh grid $\Delta x = 0.01$, 0.005, 0.0025, 0.00125.

Concerning the Monte Carlo method, we compare the results by using the hybrid simulation scheme in Section 3.3.3, that we call **HMC**. We compare our hybrid simulation scheme with the accurate third-order Alfonsi [4] discretization scheme for the CIR stochastic volatility process and by using an exact scheme for the interest rate. In addition, we simulate the jump component in the standard way. The resulting Monte Carlo scheme is here called **AMC**. In both Monte Carlo methods, we consider varying number of Monte Carlo iterations N_{MC} and two cases for the number of time discretization steps iterations:

- **HMCa** and **AMCa**: $N_t = 50$ and $N_{\text{MC}} = 10000, 50000, 100000, 200000$;

- **HMCb** and **AMCb**: $N_t = 100$ and $N_{\text{MC}} = 10000, 50000, 100000, 200000$.

All Monte Carlo results include the associated 95% confidence interval.

Table 3.1 reports European call option prices. Comparisons are given with a benchmark value obtained using the Carr-Madan pricing formula **CF** in [33] that applies Fast Fourier Transform methods (see the Premia software implementation [84]).

In Table 3.2 we provide results for American call option prices. In this case we compare with the values obtained by using the method of lines in [35], called **MOL**, with mesh parameters 200 time-steps, 250 volatility lines, 2995 asset grid points, and the **PSOR** method with mesh parameters $1000, 3000, 6000$ that Chiarella *et al.* [34] used as the true solution. Moreover, we consider the Longstaff-Schwartz [76] Monte Carlo algorithm both for **AMC** and **HMC**. In particular

- **HMCLSa** and **AMCLSa**: 10 exercise dates, $N_t = 50$ and $N_{\text{MC}} = 10000, 50000, 100000, 200000$;

- **HMCLSb** and **AMCLSb**: 20 exercise dates, $N_t = 100$ and $N_{\text{MC}} = 10000, 50000, 100000, 200000$.

Tables 3.3 and 3.4 refer to the computational time cost (in seconds) of the various algorithms for $\rho = -0.5$ in the European and American case respectively.

In order to make some heuristic considerations about the speed of convergence of our approach **HTFD**, we consider the convergence ratio proposed in [40], defined as

$$\text{ratio} = \frac{P_{\frac{N}{2}} - P_{\frac{N}{4}}}{P_N - P_{\frac{N}{2}}}, \tag{3.5.64}$$

where P_N denotes here the approximated price obtained with $N = N_t$ number of time steps. Recall that $P_N = O(N^{-\alpha})$ means that ratio $= 2^\alpha$. Table 3.5 suggests that the convergence ratio for

HTDFb is approximatively linear. The analysis of the convergence in Chapter 4 will confirm this heuristic deduction.

We notice that the above argument does not formally allow to state the speed of convergence of a method knowing its ratio. We will come back on this topic in the next chapter of this thesis. However, we anticipate here that our theoretical analysis of the convergence confirms the first order in time rate of convergence of the procedure.

The numerical results in Table 3.1-3.4 show that **HTFD** is accurate, reliable and efficient for pricing European and American options in the Bates model. Moreover, our hybrid Monte Carlo algorithm **HMC** appears to be competitive with **AMC**, that is the one from the accurate simulations by Alfonsi [4]: the numerical results are similar in term of precision and variance but **HMC** is definitely better from the computational times point of view. Additionally, because of its simplicity, **HMC** represents a real and interesting alternative to **AMC**.

As a further evidence of the accuracy of our hybrid methods, in Figure 3.2 and 3.3 we study the shapes of implied volatility smiles across moneyness $\frac{K}{S_0}$ and maturities T using **HTFDa** with $N_t = 50$ and $\Delta y = 0.005$, **HMCa** with $N_t = 50$ and $N_{\mathrm{MC}} = 50000$ and we compare the graphs with the results from the benchmark values **CF**.

(a)

$\rho = -0.5$	Δx	HTFDa	HTFDb	CF	N_{MC}	HMCa	HMCb	AMCa	AMCb
	0.01	1.1302	1.1302		10000	1.08±0.09	1.11±0.09	1.00±0.09	1.08±0.09
	0.005	1.1293	1.1294		50000	1.12±0.04	1.17±0.04	1.07±0.04	1.10±0.04
$S_0 = 80$	0.0025	1.1291	1.1292	1.1293	100000	1.14±0.03	1.14±0.03	1.13±0.03	1.13±0.03
	0.00125	1.1291	1.1292		200000	1.13±0.02	1.14±0.02	1.11±0.02	1.12±0.02
	0.01	3.3331	3.3312		10000	3.27±0.17	3.27±0.17	3.19±0.16	3.22±0.16
	0.005	3.3315	3.3301		50000	3.32±0.08	3.40±0.08	3.24±0.07	3.26±0.0
$S_0 = 90$	0.0025	3.3311	3.3298	3.3284	100000	3.34±0.05	3.34±0.05	3.32±0.05	3.33±0.05
	0.00125	3.3310	3.3297		200000	3.32±0.04	3.35±0.04	3.28±0.04	3.31±0.04
	0.01	7.5245	7.5239		10000	7.46±0.25	7.46±0.25	7.37±0.24	7.36±0.25
	0.005	7.5236	7.5224		50000	7.53±0.11	7.62±0.11	7.40±0.11	7.43±0.11
$S_0 = 100$	0.0025	7.5231	7.5221	7.5210	100000	7.54±0.08	7.52±0.08	7.53±0.08	7.52±0.08
	0.00125	7.5230	7.5220		200000	7.50±0.06	7.54±0.06	7.46±0.06	7.50±0.06
	0.01	13.6943	13.6940		10000	13.69±0.34	13.69±0.34	13.52±0.33	13.48±0.33
	0.005	13.6923	13.6924		50000	13.71±0.15	13.81±0.15	13.55±0.15	13.58±0.15
$S_0 = 110$	0.0025	13.6918	13.6921	13.6923	100000	13.72±0.11	13.69±0.11	13.67±0.11	13.70±0.11
	0.00125	13.6917	13.6920		200000	13.64±0.08	13.71±0.08	13.63±0.07	13.69±0.08
	0.01	21.3173	21.3185		10000	21.40±0.41	21.40±0.41	21.08±0.40	21.03±0.41
	0.005	21.3156	21.3168		50000	21.35±0.18	21.46±0.19	21.17±0.18	21.21±0.18
$S_0 = 120$	0.0025	21.3152	21.3164	21.3174	100000	21.36±0.13	21.32±0.13	21.29±0.13	21.33±0.13
	0.00125	21.3152	21.3163		200000	21.25±0.09	21.33±0.09	21.26±0.09	21.33±0.09

(b)

$\rho = 0.5$	Δx	HTFDa	HTFDb	CF	N_{MC}	HMCa	HMCb	AMCa	AMCb
	0.01	1.4732	1.4751		10000	1.42±0.12	1.40±0.12	1.37±0.12	1.35±0.12
	0.005	1.4724	1.4744		50000	1.49±0.06	1.47±0.05	1.40±0.05	1.42±0.05
$S_0 = 80$	0.0025	1.4723	1.4742	1.4760	100000	1.48±0.04	1.46±0.04	1.46±0.04	1.49±0.04
	0.00125	1.4722	1.4741		200000	1.47±0.03	1.48±0.03	1.48±0.03	1.48±0.03
	0.01	3.6849	3.6859		10000	3.63±0.19	3.63±0.19	3.48±0.19	3.49±0.19
	0.005	3.6836	3.6849		50000	3.70±0.09	3.70±0.09	3.57±0.09	3.60±0.09
$S_0 = 90$	0.0025	3.6832	3.6847	3.6862	100000	3.67±0.06	3.67±0.06	3.66±0.06	3.71±0.06
	0.00125	3.6832	3.6847		200000	3.66±0.04	3.70±0.04	3.69±0.04	3.68±0.04
	0.01	7.6247	7.6245		10000	7.58±0.28	7.58±0.28	7.35±0.28	7.36±0.27
	0.005	7.6238	7.6232		50000	7.66±0.13	7.65±0.13	7.47±0.12	7.52±0.12
$S_0 = 100$	0.0025	7.6234	7.6229	7.6223	100000	7.61±0.09	7.59±0.09	7.58±0.09	7.66±0.09
	0.00125	7.6233	7.6228		200000	7.58±0.06	7.64±0.06	7.62±0.06	7.61±0.06
	0.01	13.4863	13.4835		10000	13.48±0.36	13.48±0.36	13.21±0.36	13.19±0.36
	0.005	13.4842	13.4818		50000	13.55±0.17	13.49±0.16	13.27±0.16	13.35±0.16
$S_0 = 110$	0.0025	13.4837	13.4814	13.4791	100000	13.47±0.12	13.41±0.12	13.44±0.12	13.54±0.12
	0.00125	13.4836	13.4813		200000	13.42±0.08	13.49±0.08	13.47±0.08	13.48±0.08
	0.01	20.9678	20.9661		10000	21.04±0.44	21.04±0.44	20.67±0.44	20.64±0.43
	0.005	20.9659	20.9642		50000	21.05±0.20	20.98±0.20	20.71±0.20	20.81±0.20
$S_0 = 120$	0.0025	20.9655	20.9636	20.9616	100000	20.96±0.14	20.87±0.14	20.92±0.14	21.04±0.14
	0.00125	20.9654	20.9635		200000	20.88±0.10	20.96±0.10	20.97±0.10	20.98±0.10

Table 3.1: *Standard Bates model. Prices of European call options. Test parameters:* $K = 100$, $T = 0.5$, $r = 0.03$, $\delta = 0.05$, $Y_0 = 0.04$, $\theta_Y = 0.04$, $\kappa_Y = 2$, $\sigma_Y = 0.4$, $\lambda = 5$, $\gamma = 0$, $\eta = 0.1$, $\rho = -0.5, 0.5$.

(a)

$\rho = -0.5$	Δx	HTFDa	HTFDb	PSOR	MOL	N_{MC}	HMCLSa	HMCLSb	AMCLSa	AMCLSb
	0.01	1.1365	1.1365			10000	1.03±0.08	1.14±0.09	1.06±0.09	1.03±0.09
	0.005	1.1356	1.1358			50000	1.19±0.04	1.14±0.04	1.18±0.04	1.12±0.04
$S_0 = 80$	0.0025	1.1354	1.1356	1.1359	1.1363	100000	1.15±0.03	1.13±0.03	1.13±0.03	1.13±0.03
	0.00125	1.1353	1.1355			200000	1.14±0.02	1.14±0.02	1.14±0.02	1.14±0.02
	0.01	3.3579	3.3563			10000	3.39±0.15	3.44±0.16	3.38±0.15	3.48±0.16
	0.005	3.3564	3.3551			50000	3.46±0.07	3.33±0.07	3.46±0.07	3.32±0.07
$S_0 = 90$	0.0025	3.3560	3.3548	3.3532	3.3530	100000	3.35±0.05	3.35±0.05	3.33±0.05	3.36±0.05
	0.00125	3.3559	3.3547			200000	3.35±0.03	3.33±0.03	3.35±0.03	3.34±0.03
	0.01	7.6010	7.6006			10000	7.68±0.23	7.88±0.24	7.63±0.23	7.80±0.24
	0.005	7.6001	7.5992			50000	7.75±0.11	7.59±0.10	7.76±0.10	7.53±0.10
$S_0 = 100$	0.0025	7.5997	7.5989	7.5970	7.5959	100000	7.56±0.07	7.61±0.07	7.56±0.07	7.61±0.07
	0.00125	7.5996	7.5989			200000	7.58±0.05	7.55±0.05	7.58±0.05	7.57±0.05
	0.01	13.8853	13.8854			10000	13.90±0.29	14.28±0.30	13.84±0.29	14.10±0.29
	0.005	13.8836	13.8842			50000	14.05±0.13	13.89±0.12	14.07±0.13	13.86±0.12
$S_0 = 110$	0.0025	13.8832	13.8839	13.8830	13.8827	100000	13.80±0.09	13.91±0.09	13.84±0.09	13.89±0.09
	0.00125	13.8831	13.8838			200000	13.86±0.06	13.84±0.06	13.87±0.06	13.83±0.06
	0.01	21.7180	21.7199			10000	21.83±0.34	22.07±0.33	21.71±0.30	22.04±0.34
	0.005	21.7168	21.7187			50000	21.91±0.15	21.76±0.13	21.90±0.15	21.72±0.13
$S_0 = 120$	0.0025	21.7166	21.7184	21.7186	21.7191	100000	21.59±0.10	21.78±0.10	21.64±0.10	21.72±0.10
	0.00125	21.7165	21.7183			200000	21.68±0.07	21.65±0.07	21.68±0.07	21.67±0.07

(b)

$\rho = 0.5$	Δx	HTFDa	HTFDb	PSOR	MOL	N_{MC}	HMCLSa	HMCLSb	AMCLSa	AMCLSb
	0.01	1.4817	1.4837			10000	1.32±0.11	1.03±0.09	1.51±0.13	0.66±0.08
	0.005	1.4809	1.4830			50000	1.51±0.05	1.31±0.05	1.54±0.05	1.47±0.05
$S_0 = 80$	0.0025	1.4807	1.4828	1.4843	1.4848	100000	1.50±0.04	1.50±0.04	1.51±0.04	1.48±0.04
	0.00125	1.4807	1.4828			200000	1.50±0.03	1.49±0.02	1.49±0.03	1.47±0.02
	0.01	3.7134	3.7148			10000	3.83±0.19	3.79±0.17	3.89±0.19	3.95±0.19
	0.005	3.7121	3.7139			50000	3.81±0.08	3.70±0.08	3.84±0.08	3.69±0.08
$S_0 = 90$	0.0025	3.7118	3.7137	3.7145	3.7146	100000	3.69±0.06	3.75±0.06	3.72±0.06	3.70±0.06
	0.00125	3.7118	3.7137			200000	3.70±0.04	3.71±0.04	3.72±0.04	3.70±0.04
	0.01	7.7044	7.7051			10000	7.74±0.26	7.85±0.25	7.96±0.26	7.99±0.26
	0.005	7.7036	7.7039			50000	7.85±0.12	7.68±0.11	7.87±0.12	7.68±0.11
$S_0 = 100$	0.0025	7.7033	7.7036	7.7027	7.7018	100000	7.66±0.08	7.75±0.08	7.65±0.08	7.73±0.08
	0.00125	7.7032	7.7036			200000	7.69±0.06	7.67±0.05	7.68±0.06	7.69±0.05
	0.01	13.6770	13.6756			10000	13.57±0.32	13.98±0.31	13.88±0.32	14.12±0.33
	0.005	13.6752	13.6742			50000	13.83±0.14	13.67±0.13	13.89±0.14	13.64±0.13
$S_0 = 110$	0.0025	13.6747	13.6739	13.6722	13.6715	100000	13.56±0.09	13.74±0.10	13.58±0.10	13.71±0.10
	0.00125	13.6747	13.6738			200000	13.65±0.07	13.65±0.07	13.64±0.07	13.64±0.07
	0.01	21.3668	21.3671			10000	21.45±0.32	21.60±0.35	21.39±0.33	21.84±0.34
	0.005	21.3655	21.3658			50000	21.54±0.15	21.40±0.14	21.61±0.16	21.40±0.13
$S_0 = 120$	0.0025	21.3653	21.3655	21.3653	21.3657	100000	21.26±0.10	21.43±0.10	21.27±0.10	21.38±0.10
	0.00125	21.3652	21.3653			200000	21.31±0.07	21.33±0.07	21.31±0.07	21.31±0.07

Table 3.2: *Standard Bates model. Prices of American call options. Test parameters:* $K = 100$, $T = 0.5$, $r = 0.03$, $\delta = 0.05$, $Y_0 = 0.04$, $\theta_Y = 0.04$, $\kappa_Y = 2$, $\sigma_Y = 0.4$, $\lambda = 5$, $\gamma = 0$, $\eta = 0.1$, $\rho = -0.5, 0.5$.

Δx	HTFDa	HTDFb	N_{MC}	HMCa	HMCb	AMCa	AMCb	CF
0.01	0.09	0.34	10000	0.007	0.16	0.16	0.30	
0.005	0.18	0.72	50000	0.36	0.72	0.79	1.51	
0.0025	0.46	1.62	100000	0.71	1.44	1.57	3.12	0.001
0.00125	0.84	3.53	200000	1.45	2.95	3.14	6.17	

Table 3.3: *Standard Bates model. Computational times (in seconds) for European call options in Table 3.1 for* $S_0 = 100$, $\rho = -0.5$.

Δx	HTFDa	HTDFb	N_{MC}	HMCLSa	HMCLSb	AMCLSa	AMCLSb
0.01	0.10	0.37	10000	0.09	0.23	0.20	0.45
0.005	0.19	0.77	50000	0.47	1.11	1.01	2.25
0.0025	0.48	1.77	100000	1.07	2.25	2.01	4.57
0.00125	0.95	3.61	200000	1.94	4.55	4.05	8.98

Table 3.4: *Standard Bates model. Computational times (in seconds) for American call options in Table 3.2 for* $S_0 = 100$, $\rho = -0.5$.

N	$S_0 = 80$	$S_0 = 90$	$S_0 = 100$	$S_0 = 110$	$S_0 = 120$
200	1.919250	1.961063	1.894156	2.299666	2.109026
400	2.172836	2.209762	2.556021	1.673541	1.996332
800	1.544849	1.851932	1.463712	2.935697	2.106880

Table 3.5: *Standard Bates model. HTFDb-ratio* (3.5.64) *for the price of American call options as the starting point* S_0 *varies with fixed space step* $\Delta x = 0.0025$. *Test parameters:* $T = 0.5$, $r = 0.03$, $\delta = 0.05$, $Y_0 = 0.04$, $\theta = 0.04$, $\kappa = 2$, $\sigma = 0.4$, $\lambda = 5$, $\gamma = 0$, $\eta = 0.1$, $\rho = -0.5$.

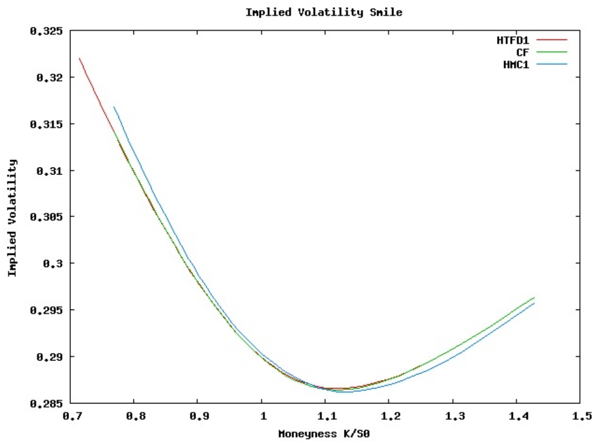

Figure 3.2: *Standard Bates model. Moneyness vs implied volatility for European call options. Test parameters:* $T = 0.5$, $r = 0.03$, $\delta = 0.05$, $Y_0 = 0.04$, $\theta_Y = 0.04$, $\kappa_Y = 2$, $\sigma_Y = 0.4$, $\lambda = 5$, $\gamma = 0$, $\eta = 0.1$, $\rho = -0.5$.

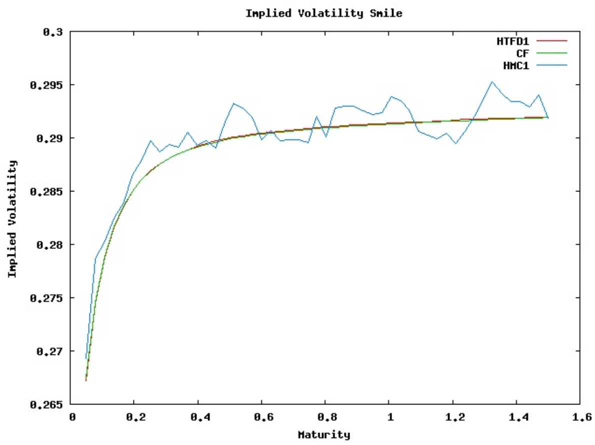

Figure 3.3: *Standard Bates model. Maturity vs implied volatility for European call options. Test parameters:* $S_0 = 100$, $K = 100$, $r = 0.03$, $\delta = 0.05$, $Y_0 = 0.04$, $\theta_Y = 0.04$, $\kappa_Y = 2$, $\sigma_Y = 0.4$, $\lambda = 5$, $\gamma = 0$, $\eta = 0.1$, $\rho = -0.5$.

Options with large maturity in the standard Bates model

In order to verify the robustness of the proposed algorithms we consider experiments when the Feller condition $2\kappa_Y\theta_Y \geq \sigma_Y^2$ is not fulfilled for the CIR volatility process. We additionally stress our tests by considering large maturities. For this purpose we consider the parameters from Chiarella *et al.* [34] already used in Section 3.5.2 with $\rho = -0.5$, except for the maturity and the vol-vol, which are modified as follows: $T = 5$ and $\sigma_Y = 0.7$ respectively.

Table 3.6 reports European call option prices, which are compared with the true values (**CF**). In Table 3.7 we provide results for American call option prices. The settings for the experiments **HTFDa-b**, **HMCa-b** and **AMCa-b** are the same as described at the beginning of Section 3.5.2. The settings for the experiments in the American case **HMCLSa-b** and **AMCLSa-b** are changed

- **HMCLSa** and **AMCLSa**: 20 exercise dates, $N_t = 100$ and $N_{\mathrm{MC}} = 10000, 50000, 100000, 200000$;

- **HMCLSb** and **AMCLSb**: 40 exercise dates, $N_t = 200$ and $N_{\mathrm{MC}} = 10000, 50000, 100000, 200000$.

In the American case the benchmark values **B-AMC** are obtained by the Longstaff-Schwartz [76] Monte Carlo algorithm with 300 exercise dates, combined with the accurate third-order Alfonsi method with 3000 discretization time steps and 1 million iterations.

The numerical results suggest that large maturities bring to a slight loss of accuracy for **HTFD** and **HMC**, even if both methods provide a satisfactory approximation of the true option prices, being in turn mostly compatible with the results from the Alfonsi Monte Carlo method. It is worth noticing that for long maturity $T = 5$ we have developed experiments with the same number of steps both in time (N_t) and space step (Δx) as for $T = 0.5$. So, the numerical experiments are not slower, and it is clear that one could achieve a better accuracy for larger values of N_t.

$\rho = -0.5$	Δx	HTFDa	HTFDb	CF	N_{MC}	HMCa	HMCb	AMCa	AMCb
	0.01	9.0085	8.9457		10000	9.21±0.55	9.09±0.55	8.69±0.53	8.56±0.51
	0.0050	9.0032	8.9405		50000	9.13±0.25	8.92±0.24	8.81±0.24	9.04±0.24
$S_0 = 80$	0.0025	9.0020	8.9392	8.9262	100000	9.01±0.17	8.81±0.17	8.92±0.17	8.88±0.17
	0.00125	9.0016	8.9389		200000	8.99±0.12	8.92±0.12	8.95±0.12	8.90±0.12
	0.01	12.7405	12.6520		10000	12.95±0.67	12.95±0.67	12.29±0.65	12.15±0.6
	0.0050	12.7342	12.6458		50000	12.87±0.30	12.64±0.29	12.49±0.29	12.76±0.3
$S_0 = 90$	0.0025	12.7327	12.6442	12.6257	100000	12.72±0.21	12.50±0.21	12.63±0.21	12.58±0.21
	0.00125	12.7323	12.6438		200000	12.71±0.15	12.61±0.15	12.66±0.15	12.61±0.15
	0.01	17.0324	16.9176		10000	17.24±0.80	17.24±0.80	16.43±0.77	16.29±0.75
	0.0050	17.0254	16.9106		50000	17.18±0.36	16.91±0.35	16.73±0.35	17.03±0.35
$S_0 = 100$	0.0025	17.0237	16.9089	16.8855	100000	17.00±0.25	16.74±0.25	16.91±0.25	16.84±0.25
	0.00125	17.0232	16.9084		200000	16.99±0.18	16.86±0.18	16.94±0.18	16.88±0.18
	0.01	21.8149	21.6741		10000	22.04±0.93	22.04±0.93	21.06±0.93	20.91±0.88
	0.0050	21.8067	21.6659		50000	21.96±0.42	21.67±0.41	21.43±0.41	21.82±0.41
$S_0 = 110$	0.0025	21.8047	21.6639	21.6364	100000	21.76±0.29	21.47±0.29	21.69±0.29	21.59±0.29
	0.00125	21.8042	21.6634		200000	21.76±0.21	21.59±0.20	21.70±0.20	21.63±0.20
	0.01	27.0196	26.8539		10000	27.26±1.05	27.26±1.05	26.12±1.03	25.94±1.01
	0.0050	27.0108	26.8452		50000	27.17±0.47	26.86±0.46	26.56±0.46	27.02±0.47
$S_0 = 120$	0.0025	27.0086	26.8430	26.8121	100000	26.94±0.33	26.63±0.33	26.89±0.33	26.78±0.33
	0.00125	27.0081	26.8425		200000	26.95±0.23	26.75±0.23	26.89±0.23	26.81±0.23

Table 3.6: *Standard Bates model. Prices of European call options. Test parameters:* $K = 100$, $T = 5$, $r = 0.03$, $\delta = 0.05$, $Y_0 = 0.04$, $\theta_Y = 0.04$, $\kappa_Y = 2$, $\sigma_Y = 0.7$, $\lambda = 5$, $\gamma = 0$, $\eta = 0.1$, $\rho = -0.5$. *Case* $2\kappa_Y\theta_Y < \sigma_Y^2$.

$\rho = -0.5$	Δy	HTFDa	HTFDb	B-AMC	N_{MC}	HMCLSa	HMCLSb	AMCLSa	AMCLSb
	0.01	9.8335	9.7978		10000	10.15±0.46	10.20±0.46	10.47±0.47	9.80±0.42
	0.0050	9.8283	9.7927		50000	9.93±0.20	9.86±0.20	9.89±0.19	9.78±0.19
$S_0 = 80$	0.0025	9.8271	9.7914	9.7907± 0.04	100000	9.76±0.14	9.69±0.13	9.74±0.14	9.76±0.13
	0.00125	9.8267	9.7911		200000	9.79±0.10	9.70±0.09	9.73±0.10	9.72±0.09
	0.01	14.0801	14.0318		10000	14.58±0.56	14.46±0.55	14.94±0.58	14.08±0.51
	0.0050	14.0741	14.0258		50000	14.13±0.24	14.14±0.24	14.19±0.23	14.12±0.23
$S_0 = 90$	0.0025	14.0726	14.0244	14.0030± 0.05	100000	13.98±0.16	13.87±0.16	13.94±0.16	13.89±0.16
	0.00125	14.0722	14.0240		200000	13.93±0.12	13.91±0.11	13.94±0.12	13.96±0.11
	0.01	19.0658	19.0075		10000	19.59±0.66	19.44±0.63	19.88±0.66	19.13±0.59
	0.0050	19.0594	19.0011		50000	19.10±0.27	19.06±0.27	19.26±0.26	19.01±0.26
$S_0 = 100$	0.0025	19.0578	18.9995	18.9632± 0.05	100000	18.92±0.19	18.88±0.18	18.85±0.19	18.90±0.18
	0.00125	19.0574	18.9991		200000	18.80±0.13	18.84±0.13	18.85±0.13	18.92±0.13
	0.01	24.7434	24.6788		10000	25.02±0.74	24.84±0.72	25.32±0.72	24.78±0.67
	0.0050	24.7364	24.6719		50000	24.79±0.30	24.57±0.29	24.94±0.29	24.72±0.29
$S_0 = 110$	0.0025	24.7347	24.6701	24.6289± 0.06	100000	24.53±0.21	24.47±0.20	24.50±0.21	24.51±0.20
	0.00125	24.7343	24.6697		200000	24.42±0.14	24.45±0.14	24.50±0.15	24.53±0.14
	0.01	31.0646	30.9983		10000	30.88±0.74	31.15±0.75	31.18±0.74	31.04±0.71
	0.0050	31.0577	30.9914		50000	31.10±0.32	30.94±0.31	31.32±0.32	30.98±0.32
$S_0 = 120$	0.0025	31.0559	30.9896	30.9052±0.07	100000	30.89±0.23	30.72±0.22	30.70±0.22	30.72±0.22
	0.00125	31.0555	30.9892		200000	30.72±0.16	30.73±0.16	30.77±0.16	30.89±0.15

Table 3.7: *Standard Bates model. Prices of American call options. Test parameters:* $K = 100$, $T = 5$, $r = 0.03$, $\delta = 0.05$, $Y_0 = 0.04$, $\theta_Y = 0.04$, $\kappa_Y = 2$, $\sigma_Y = 0.7$, $\lambda = 5$, $\gamma = 0$, $\delta = 0.1$, $\rho = -0.5$. *Case* $2\kappa_Y\theta_Y < \sigma_Y^2$.

Bates model with stochastic interest rate

We consider now the case of Bates model associated with the Vasiceck model for the stochastic interest rate. For the Bates model we consider the parameters from Chiarella *et al.* [34] already used in Section 3.5.2. Moreover, for the interest rate parameter we fix the following parameters:

- initial interest rate $r_0 = 0.03$, speed of mean-reversion $\kappa_r = 1$, interest rate volatility $\sigma_r = 0.2$;
- time-varying long-term mean $\theta_r(t)$ fitting the theoretical bond prices to the yield curve observed on the market, here set as $P_r(0, T) = e^{-0.03T}$.

We study the cases

$$\rho_1 = \rho_{SY} = -0.5 \quad \text{and} \quad \rho_2 = \rho_{Sr} = -0.5, 0.5.$$

No correlation is assumed to exist between r and Y. We consider the mesh grid $\Delta y = 0.02$, 0.01, 0.005, 0.0025, the case $\Delta y = 0.00125$ being removed because it requires huge computational times. The numerical results are labeled **HTFDa-b**, **HMCa-b**, **AMCa-b**, **HMCLSa-b**, **AMCLSa-b**, their settings being given at the beginning of Section 3.5.2.

When the interest rate is assumed to be stochastic, no references are available in the literature. Therefore, we propose benchmark values obtained by using a Monte Carlo method in which the CIR paths are simulated through the accurate third-order Alfonsi [4] discretization scheme and the interest rate paths are generated by an exact scheme. For these benchmark values, called **B-AMC**, the number of Monte Carlo iterations and of the discretization time steps are set as $N_{\text{MC}} = 10^6$ and $N_t = 300$ respectively. In the American case, **B-AMC** is evaluated through the Longstaff-Schwartz [76] algorithm with 20 exercise dates. All Monte Carlo results report the 95% confidence intervals.

European and American call option prices are given in tables 3.8 and 3.9 respectively. Tables 3.10 and 3.11 refer to the computational time cost (in seconds) of the different algorithms in the European Call case and American Call case respectively. The numerical results confirm the good numerical behavior of **HTFD** and **HMC** in the Bates-Hull-White model as well.

(a)

$\rho_{Sr} = -0.5$	Δx	HTFDa	HTFDb	B-AMC	N_{MC}	HMCa	HMCb	AMCa	AMCb
	0.02	1.0169	1.0079		10000	1.00±0.09	0.96±0.09	1.00±0.09	1.06±0.10
	0.01	1.0201	1.0188		50000	1.02±0.04	0.97±0.04	0.98±0.04	1.01±0.04
$S_0 = 80$	0.0050	1.0199	1.0194	1.0153±0.01	100000	1.00±0.03	1.00±0.03	1.01±0.03	1.03±0.03
	0.0025	1.0197	1.0193		200000	1.01±0.02	1.01±0.02	1.02±0.02	1.00±0.02
	0.01	3.1172	3.1032		10000	3.05±0.16	3.05±0.16	3.07±0.16	3.14±0.17
	0.01	3.1186	3.1137		50000	3.10±0.07	3.03±0.07	3.02±0.07	3.09±0.07
$S_0 = 90$	0.0050	3.1174	3.1135	3.1008±0.02	100000	3.07±0.05	3.08±0.05	3.09±0.05	3.14±0.05
	0.0025	3.1174	3.1136		200000	3.09±0.04	3.10±0.04	3.11±0.04	3.08±0.04
	0.02	7.2528	7.2472		10000	7.17±0.24	7.17±0.24	7.20±0.24	7.24±0.25
	0.01	7.2528	7.2479		50000	7.21±0.11	7.18±0.11	7.12±0.11	7.21±0.11
$S_0 = 100$	0.0050	7.2528	7.2480	7.2315±0.02	100000	7.18±0.08	7.24±0.08	7.20±0.08	7.27±0.08
	0.0025	7.2528	7.2480		200000	7.22±0.05	7.25±0.05	7.24±0.05	7.20±0.05
	0.02	13.4553	13.4565		10000	13.30±0.32	13.30±0.32	13.41±0.33	13.39±0.33
	0.01	13.4465	13.4440		50000	13.37±0.15	13.40±0.15	13.27±0.15	13.38±0.15
$S_0 = 110$	0.0050	13.4435	13.4407	13.4256±0.03	100000	13.35±0.10	13.46±0.10	13.38±0.10	13.48±0.10
	0.0025	13.4432	13.4404		200000	13.40±0.07	13.47±0.07	13.43±0.07	13.39±0.07
	0.02	21.1320	21.1356		10000	20.89±0.40	20.89±0.40	21.08±0.40	20.99±0.41
	0.01	21.1243	21.1239		50000	21.03±0.18	21.09±0.18	20.92±0.18	21.03±0.18
$S_0 = 120$	0.0050	21.1222	21.1214	21.1070±0.04	100000	21.01±0.13	21.17±0.13	21.04±0.13	21.17±0.13
	0.0025	21.1215	21.1207		200000	21.06±0.09	21.16±0.09	21.12±0.09	21.06±0.09

(b)

$\rho_{Sr} = 0.5$	Δx	HTFDa	HTFDb	B-AMC	N_{MC}	HMCa	HMCb	AMCa	AMCb
	0.02	1.3459	1.3379		10000	1.29±0.11	1.28±0.11	1.32±0.10	1.41±0.11
	0.01	1.3482	1.3471		50000	1.34±0.05	1.30±0.05	1.32±0.05	1.35±0.05
$S_0 = 80$	0.0050	1.3479	1.3475	1.3446±0.01	100000	1.32±0.03	1.31±0.03	1.34±0.03	1.34±0.03
	0.0025	1.3477	1.3473		200000	1.33±0.02	1.34±0.02	1.35±0.02	1.32±0.02
	0.01	3.7320	3.7233		10000	3.62±0.18	3.62±0.18	3.64±0.18	3.76±0.19
	0.01	3.7323	3.7304		50000	3.69±0.08	3.65±0.08	3.64±0.18	3.76±0.19
$S_0 = 90$	0.0050	3.7311	3.7298	3.7263±0.02	100000	3.66±0.06	3.68±0.06	3.71±0.06	3.73±0.06
	0.0025	3.7311	3.7299		200000	3.69±0.04	3.72±0.04	3.73±0.04	3.68±0.04
	0.02	8.0100	8.0073		10000	7.83±0.26	7.83±0.26	7.82±0.26	8.00±0.27
	0.01	8.0112	8.0102		50000	7.92±0.12	7.93±0.12	7.93±0.12	7.97±0.12
$S_0 = 100$	0.0050	8.0114	8.0107	8.0069±0.03	100000	7.91±0.08	7.97±0.08	7.99±0.08	8.02±0.08
	0.0025	8.0114	8.0107		200000	7.95±0.06	8.02±0.06	8.00±0.06	7.95±0.06
	0.02	14.1482	14.1505		10000	13.89±0.35	13.89±0.35	13.88±0.35	14.07±0.36
	0.01	14.1413	14.1414		50000	14.01±0.16	14.05±0.16	14.03±0.16	14.09±0.16
$S_0 = 110$	0.0050	14.1388	14.1388	14.1323±0.03	100000	14.01±0.11	14.10±0.11	14.12±0.11	14.14±0.11
	0.0025	14.1386	14.1386		200000	14.06±0.08	14.17±0.08	14.13±0.08	14.07±0.08
	0.02	21.6737	21.6772		10000	21.37±0.42	21.37±0.42	21.35±0.42	21.51±0.43
	0.01	21.6670	21.6674		50000	21.50±0.19	21.55±0.19	21.52±0.19	21.60±0.19
$S_0 = 120$	0.0050	21.6651	21.6653	21.6501±0.04	100000	21.52±0.13	21.63±0.13	21.64±0.13	21.68±0.14
	0.0025	21.6645	21.6646		200000	21.57±0.10	21.71±0.10	21.65±0.10	21.58±0.09

Table 3.8: *Bates-Hull-White model. Prices of European call options. Test parameters:* $K = 100$, $T = 0.5$, $\delta = 0.05$, , $r_0 = 0.03$, $\kappa_r = 1$, $\sigma_r = 0.2$, $Y_0 = 0.04$, $\theta_Y = 0.04$, $\kappa_Y = 2$, $\sigma_Y = 0.4$, $\lambda = 5$, $\gamma = 0$, $\eta = 0.1$, $\rho_{SY} = -0.5$,$\rho_{Sr} = -0.5, 0.5$.

(a)

$\rho_{Sr} = -0.5$	Δx	HTFDa	HTFDb	B-AMC	N_{MC}	HMCLSa	HMCLSb	AMCLSa	AMCLSb
	0.02	1.0561	1.0470		10000	0.76±0.07	0.56±0.06	0.95±0.08	0.82±0.08
	0.01	1.0598	1.0588		50000	1.08±0.04	0.91±0.04	1.01±0.04	0.96±0.04
$S_0 = 80$	0.0050	1.0597	1.0596	1.0544±0.01	100000	1.07±0.03	1.03±0.03	1.07±0.03	1.04±0.03
	0.0025	1.0596	1.0595		200000	1.05±0.02	1.04±0.02	1.07±0.02	1.05±0.02
	0.01	3.2511	3.2364		10000	3.28±0.15	3.39±0.16	3.35±0.16	3.07±0.15
	0.01	3.2537	3.2493		50000	3.33±0.07	3.21±0.07	3.25±0.07	3.30±0.07
$S_0 = 90$	0.0050	3.2528	3.2494	3.2273±0.01	100000	3.23±0.05	3.24±0.05	3.27±0.05	3.25±0.05
	0.0025	3.2528	3.2495		200000	3.22±0.03	3.23±0.03	3.25±0.03	3.24±0.03
	0.02	7.6012	7.5952		10000	7.64±0.22	7.99±0.23	7.80±0.23	7.68±0.22
	0.01	7.6020	7.5976		50000	7.72±0.10	7.58±0.09	7.61±0.10	7.65±0.10
$S_0 = 100$	0.0050	7.6022	7.5980	7.5589±0.02	100000	7.54±0.07	7.62±0.07	7.61±0.07	7.54±0.07
	0.0025	7.6022	7.5980		200000	7.54±0.05	7.54±0.05	7.56±0.05	7.60±0.05
	0.02	14.1510	14.1524		10000	14.22±0.28	14.61±0.29	14.35±0.29	14.07±0.28
	0.01	14.1443	14.1425		50000	14.25±0.13	14.11±0.12	14.16±0.12	14.17±0.13
$S_0 = 110$	0.0050	14.1420	14.1401	14.0909±0.03	100000	14.03±0.09	14.18±0.09	14.10±0.09	14.06±0.09
	0.0025	14.1419	14.1399		200000	14.05±0.06	14.04±0.06	14.07±0.06	14.13±0.06
	0.02	22.2466	22.2505		10000	22.38±0.32	22.84±0.33	22.46±0.32	22.15±0.32
	0.01	22.2412	22.2419		50000	22.35±0.15	22.27±0.14	22.24±0.14	22.28±0.14
$S_0 = 120$	0.0050	22.2398	22.2402	22.1736±0.03	100000	22.12±0.10	22.27±0.10	22.19±0.10	22.17±0.10
	0.0025	22.2394	22.2397		100000	22.12±0.10	22.27±0.10	22.19±0.10	22.17±0.10

(b)

$\rho_{Sr} = 0.5$	Δx	HTFDa	HTFDb	B-AMC	N_{MC}	HMCLSa	HMCLSb	AMCLSa	AMCLSb
	0.02	1.3551	1.3470		10000	1.18±0.09	1.29±0.10	1.12±0.09	0.80±0.08
	0.01	1.3576	1.3566		50000	1.35±0.05	1.17±0.04	1.33±0.05	1.25±0.05
$S_0 = 80$	0.0050	1.3573	1.3570	1.3559±0.01	100000	1.33±0.03	1.30±0.03	1.33±0.03	1.27±0.03
	0.0025	1.3571	1.3569		200000	1.35±0.02	1.31±0.02	1.38±0.02	1.34±0.02
	0.01	3.7696	3.7606		10000	3.72±0.17	3.78±0.17	3.82±0.18	3.72±0.17
	0.01	3.7705	3.7688		50000	3.86±0.08	3.71±0.08	3.80±0.08	3.81±0.08
$S_0 = 90$	0.0050	3.7694	3.7685	3.7633±0.02	100000	3.75±0.06	3.74±0.05	3.76±0.05	3.74±0.05
	0.0025	3.7694	3.7686		200000	3.75±0.04	3.74±0.04	3.80±0.04	3.79±0.04
	0.02	8.1285	8.1249		10000	8.12±0.24	8.52±0.26	8.25±0.26	8.15±0.25
	0.01	8.1308	8.1301		50000	8.25±0.11	8.08±0.11	8.15±0.11	8.18±0.11
$S_0 = 100$	0.0050	8.1311	8.1308	8.1122±0.03	100000	8.07±0.08	8.16±0.08	8.11±0.08	8.10±0.08
	0.0025	8.1312	8.1309		200000	8.08±0.06	8.07±0.06	8.14±0.06	8.16±0.06
	0.02	14.4455	14.4468		10000	14.48±0.32	14.84±0.33	14.43±0.32	14.51±0.32
	0.01	14.4409	14.4414		50000	14.60±0.15	14.40±0.14	14.45±0.14	14.47±0.14
$S_0 = 110$	0.0050	14.4389	14.4395	14.3884±0.03	100000	14.34±0.10	14.47±0.10	14.39±0.10	14.38±0.10
	0.0025	14.4388	14.4394		200000	14.35±0.07	14.37±0.07	14.38±0.07	14.48±0.07
	0.02	22.2859	22.2893		10000	22.23±0.36	22.87±0.39	22.45±0.36	22.29±0.35
	0.01	22.2815	22.2827		50000	22.50±0.17	22.29±0.16	22.27±0.16	22.28±0.16
$S_0 = 120$	0.0050	22.2802	22.2813	22.2039±0.04	100000	22.17±0.12	22.31±0.12	22.24±0.12	22.22±0.12
	0.0025	22.2798	22.2808		200000	22.17±0.08	22.17±0.08	22.17±0.08	22.32±0.08

Table 3.9: *Bates-Hull-White model. Prices of American call options. Test parameters:* $K = 100$, $T = 0.5$, $\delta = 0.05$, $r_0 = 0.03$, $\kappa_r = 1$, $\sigma_r = 0.2$, $Y_0 = 0.04$, $\theta_Y = 0.04$, $\kappa_Y = 2$, $\sigma_Y = 0.4$, $\lambda = 5$, $\gamma = 0$, $\eta = 0.1$, $\rho_{SY} = -0.5$,$\rho_{Sr} = -0.5, 0.5$.

Δx	HTFDa	HTDFb	$N_{\rm MC}$	HMCa	HMCb	AMCa	AMCb
0.02	2.77	22.95	10000	0.13	0.25	0.36	0.48
0.01	6.15	48.17	50000	0.66	1.35	1.11	2.48
0.005	12.12	99.19	100000	1.37	2.56	1.82	4.99
0.0025	27.61	204.88	200000	2.56	5.08	3.70	9.96

Table 3.10: *Bates-Hull-White model. Computational times (in seconds) for European call options in Table 3.8 for* $S_0 = 100$, $\rho_{Sr} = -0.5$.

Δx	HTFDa	HTDFb	$N_{\rm MC}$	HMCLSa	HMCLSb	AMCLSa	AMCLSb
0.02	2.77	23.10	10000	0.28	0.43	0.40	0.62
0.01	6.39	48.65	50000	0.80	1.79	1.30	2.72
0.005	12.50	99.85	100000	1.91	3.89	3.02	6.15
0.0025	27.92	205.60	200000	4.03	8.11	5.20	10.75

Table 3.11: *Bates-Hull-White model. Computational times (in seconds) for American call options in Table 3.9 for* $S_0 = 100$, $\rho_{Sr} = -0.5$.

Chapter 4

Weak convergence rate of Markov chains and hybrid numerical schemes for jump-diffusion processes

4.1 Introduction

This chapter is devoted to the study of the weak convergence rate of numerical schemes allowing one to handle specific jump-diffusion processes which include the Heston and Bates models in the full parameters regime. We generalize the hybrid tree- finite difference method described in Chapter 3 for the computation of European and American options in the stochastic volatility context and we study the rate of convergence. Let us mention that, under these models, the literature is rich in numerical methods but, as far as we know, poor in results on the rate of convergence, with the exception of the papers [4, 6, 23, 98], all of them either dealing with schemes written on Brownian increments or requiring restrictions on the Heston diffusion parameters. So, we first study the convergence rate of tree methods and then we tackle the hybrid procedure.

Tree methods rely heavily on Markov chains. So, in the first part (Section 4.3) we study the rate at which a sequence of Markov chains weakly converges to a diffusion process $(Y_t)_{t\in[0,T]}$ solution to

$$dY_t = \mu_Y(Y_t)dt + \sigma_Y(Y_t)dW_t.$$

In this framework, the weak convergence is well known to be governed by the behaviour of the local moments up to order 3 or 4 (see e.g. [89]). In order to get the speed of convergence, we need to stress such requests, making further but quite general assumptions on the behaviour of the moments, and in Theorem 4.3.1 we prove a first order weak convergence result. As an application, we give an example from the financial framework: we theoretically study the convergence rate of the tree approximation proposed in [10] for the CIR process (and described in Section 3.3.1). Several trees are considered in the literature, see e.g. [36, 59, 91], but all of them work poorly from the numerical point of view when the Feller condition fails. Our result for the tree in [10] (Theorem 4.3.2) works in any parameter regime. Recall that in equity markets, one often requires large values for the vol-vol σ whereas in interest rates context, σ is markedly lower (see e.g. the calibration results in [44] and in [30] p. 115, respectively). So, a result in the full parameter regime is actually essential. We stress that our convergence Theorem 4.3.1 is completely general and may in principle be applied to more general trees constructed through the multiple jumps approach by Nelson and Ramaswamy [79] or also to other cases, e.g. the recent tree method developed in [2].

In the second part (Section 4.4), we link to $(Y_t)_{t\in[0,T]}$ a jump-diffusion process $(X_t)_{t\in[0,T]}$ which evolves according to a stochastic differential equation whose coefficients only depend on the process $(Y_t)_{t\in[0,T]}$:

$$dX_t = \mu_X(Y_t)dt + \sigma_X(Y_t)dB_t + \gamma_X(Y_t)dH_t,$$

where H is a compound Poisson process independent of the 2-dimensional Brownian motion (B, W). So, the pair $(X_t, Y_t)_{t\in[0,T]}$ evolves following a Stochastic Differential Equation (hereafter SDE) with jumps. Given a function f, we consider the numerical computation of $\mathbb{E}[f(X_T, Y_T)]$ or $\sup_{\tau\in\mathcal{T}_{0,T}} \mathbb{E}[f(X_\tau, Y_\tau)]$ through a method (Section 4.4.1), which works backwardly by approximating the process Y with a Markov chain and by using a different numerical scheme for solving a (local) PIDE allowing us to work in the direction of the process X. Then (Section 4.4.2), in Theorem 4.4.1 we give a general result on the rate of convergence of the hybrid approach. We stress that the approximating algorithm is not directly written on a Markov approximation, so one cannot extend the convergence result provided in the first part of the chapter. We then study the stability and the consistency of the hybrid method, but in a sense that allows us to exploit the probabilistic properties of the Markov chain approximating the process Y.

It is worth mentioning that the test functions on which we study the rate of convergence are smooth. In fact, there is a strict connection between such hybrid schemes and the use of a discrete noise in the approximation procedure. This means that we cannot use regularizing arguments *à la Malliavin* in order to relax the smoothness requests, as it can be done when the approximation algorithm is based on the Brownian noise (see the seminal paper [16] or the recent [6] for the

Heston model) or on a noise having at least a "good piece of absolutely continuous part" (Doeblin's condition, see [14]).

We then consider two possible finite-difference schemes (Section 4.4.3) to handle the (local) PIDE related to the component X: an implicit in time/centered in space scheme (Section 4.4.3) and an implicit in time/upwind in space scheme (Section 4.4.3). In both cases, the numerical treatment of the nonlocal term coming from the jumps involves implicit-explicit techniques, as well as numerical quadratures. We apply the convergence Theorem 4.4.1 and we obtain that the hybrid algorithm has a rate of convergence of the first order in time and of a order in space according to the chosen numerical scheme. As an application, we give the weak convergence rate of the hybrid procedure written on the Heston and on the Bates model for pricing European options (Section 4.5). Finally, in Section 4.6 we give a theoretical result on the convergence rate in the case of American options.

4.2 Notation

In this section we establish the notation which will be used in this chapter. Let $d \in \mathbb{N}^* = \mathbb{N} \setminus \{0\}$.

• For a multi-index $l = (l_1, \dots, l_d) \in \mathbb{N}^d$ we define $|l| = \sum_{j=1}^d l_j$ and for $y \in \mathbb{R}^d$, we define $\partial_y^l = \partial_{y_1}^{l_1} \cdots \partial_{y_d}^{l_d}$ and $y^l = y_1^{l_1} \cdots y_d^{l_d}$. Moreover, we denote by $|y|$ the standard Euclidean norm in $\mathbb{R}^d$ and for any linear operator $A : \mathbb{R}^d \to \mathbb{R}^d$, we denote by $|A| = \sup_{|y|=1} |Ay|$ the induced norm.

• $L^p(\mathbb{R}^d, d\mathfrak{m})$ denotes the standard L^p-space w.r.t. the measure $\mathfrak{m}$ on $(\mathbb{R}^d, \mathcal{B}_d)$, $\mathcal{B}_d$ denoting the Borel σ-algebra on $\mathbb{R}^d$, and we set $| \cdot |_{L^p(\mathbb{R}^d, d\mathfrak{m})}$ the associated norm. The Lebesgue measure is denoted through dx.

• Let $\mathcal{D} \subseteq \mathbb{R}^d$ be a domain (possibly closed) and $q \in \mathbb{N}$. $C^q(\mathcal{D})$ is the set of all functions on $\mathcal{D}$ which are q-times continuously differentiable. We set $C^q_{\mathbf{pol}}(\mathcal{D})$ the set of functions $g \in C^q(\mathcal{D})$ such that there exist $C, a > 0$ for which

$$|\partial_y^l g(y)| \leq C(1 + |y|^a), \qquad y \in \mathcal{D},\ |l| \leq q.$$

For $[a, b] \subseteq \mathbb{R}^+$, we set $C^q_{\mathbf{pol},[a,b]}(\mathcal{D})$ the set of functions $v = v(t, y)$ such that $v \in C^{\lfloor q/2 \rfloor, q}([a, b) \times \mathcal{D})$ and there exist $C, c > 0$ for which

$$\sup_{t \in [a,b)} |\partial_t^k \partial_y^l v(t, y)| \leq C(1 + |y|^c), \qquad y \in \mathcal{D},\ 2k + |l| \leq q.$$

For brevity, we set $C(\mathcal{D}) = C^0(\mathcal{D})$, $C_{\mathbf{pol}}(\mathcal{D}) = C^0_{\mathbf{pol}}(\mathcal{D})$ and $C_{\mathbf{pol},[a,b]}(\mathcal{D}) = C^0_{\mathbf{pol},[a,b]}(\mathcal{D})$. We also need another functional space, that we call $C^{p,q}_{\mathbf{pol}}(\mathbb{R}^m, \mathcal{D})$, $p \in [1, \infty]$, $q \in \mathbb{N}$, $m \in \mathbb{N}^*$: $g = g(x, y) \in$

$C^{p,q}_{\mathbf{pol}}(\mathbb{R}^m, \mathcal{D})$ if $g \in C^q_{\mathbf{pol}}(\mathbb{R}^m \times \mathcal{D})$ and there exist $C, c > 0$ such that

$$|\partial_x^{l'} \partial_y^l g(\cdot, y)|_{L^p(\mathbb{R}^m, dx)} \leq C(1 + |y|^c), \quad |l'| + |l| \leq q.$$

Similarly as above, we set $C^{p,q}_{\mathbf{pol},[a,b]}(\mathbb{R}^m, \mathcal{D})$ the set of the function $v \in C^q_{\mathbf{pol},[a,b]}(\mathbb{R}^m \times \mathcal{D})$ such that

$$\sup_{t \in [a,b)} |\partial_t^k \partial_x^{l'} \partial_y^l v(t, \cdot, y)|_{L^p(\mathbb{R}^m, dx)} \leq C(1 + |y|^c), \quad 2k + |l'| + |l| \leq q.$$

If $[a, b] = [0, T]$, to simplify the notation, we set $C^q_{\mathbf{pol},[0,T]}(\mathcal{D}) = C^q_{\mathbf{pol},T}(\mathcal{D})$ and $C^{p,q}_{\mathbf{pol},[0,T]}(\mathcal{D}) = C^{p,q}_{\mathbf{pol},T}(\mathcal{D})$.

• For fixed $X_0 = (X_{01}, \ldots, X_{0d}) \in \mathbb{R}^d$ and $\Delta x = (\Delta x_1, \ldots, \Delta x_d) \in (0, +\infty)^d$ (spatial step), $\mathcal{X} = \{x = (X_{01} + i_1 \Delta x_1, \ldots, X_{0d} + i_d \Delta x_d)\}_{i \in \mathbb{Z}^d}$ denotes a discrete grid in $\mathbb{R}^d$. For $p \in [1, \infty]$, we set $l_p(\mathcal{X})$ the discrete l_p-space of the functions $\varphi : \mathcal{X} \to \mathbb{R}$ with the norm $|\varphi|_p = (\sum_{x \in \mathcal{X}} |\varphi(x)|^p \Delta x_1 \cdots \Delta x_d)^{1/p}$ if $p \in [1, \infty)$ and $|\varphi|_\infty = \sup_{x \in \mathcal{X}} |\varphi(x)|$ if $p = \infty$. Moreover, for a linear operator $\Gamma : l_p(\mathcal{X}) \to l_p(\mathcal{X})$, the induced norm is denoted by $|\Gamma|_p = \sup_{|\varphi|_p \leq 1} |\Gamma \varphi|_p$. And for a function $g : \mathbb{R}^d \to \mathbb{R}$, we set $|g|_p$ the $l_p(\mathcal{X})$ norm of the restriction of g on $\mathcal{X}$. When $d = 1$, we identify $(\varphi(x))_{x \in \mathcal{X}}$ with $(\varphi_i)_{i \in \mathbb{Z}}$ through $\varphi_i = \varphi(X_0 + i\Delta x)$, $i \in \mathbb{Z}$.

• $L^p(\Omega)$ is the short notation for the standard L^p-space on the probability space $(\Omega, \mathcal{F}, \mathbb{P})$, on which the expectation is denoted by $\mathbb{E}$. We set $\| \cdot \|_p$ the norm in $L^p(\Omega)$.

4.3 First order weak convergence of Markov chains to diffusions

Let $d \in \mathbb{N}^*$ and $\mathcal{D} \subseteq \mathbb{R}^d$ be a convex domain or a closure of it. On a probability space $(\Omega, \mathcal{F}, \mathbb{P})$, we consider a d-dimensional diffusion process driven by

$$dY_t = \mu_Y(Y_t)dt + \sigma_Y(Y_t)dW_t, \qquad Y_0 \in \mathcal{D}, \tag{4.3.1}$$

where W is a ℓ-dimensional standard Brownian motion. From now on, we set $a_Y = \sigma_Y \sigma_Y^\star$, the notation $\star$ denoting transpose. We recall that the associated infinitesimal generator is given by

$$\mathcal{A} = \frac{1}{2}\mathrm{Tr}(a_Y D_y^2) + \mu_Y \cdot \nabla_y, \tag{4.3.2}$$

where Tr denotes the matrix trace, D_y^2 and ∇_y are, respectively, the Hessian and the gradient operator w.r.t. the space variable y and the notation "$\cdot$" stands for the scalar product.

Hereafter, we fix $T > 0$, $f : \mathcal{D} \to \mathbb{R}$ and we define

$$u(t, y) = \mathbb{E}[f(Y_T^{t,y})], \quad (t, y) \in [0, T] \times \mathcal{D}, \tag{4.3.3}$$

where $Y^{t,y}$ denotes the solution to the SDE in (4.3.1) that starts at t in the position y. We do not enter in specific requests for the diffusion coefficients or for f, we just ask that the following properties are met:

(a) μ_Y has polynomial growth;

(b) for every $(t, y) \in [0, T] \times \mathcal{D}$ there exists a unique weak solution $(Y_s^{t,y})_{s\in[t,T]}$ of (4.3.1) such that $\mathbb{P}(\forall s \in [t, T],\, Y_s^{t,y} \in \mathcal{D}) = 1$;

(c) the function u in (4.3.3) solves the PDE

$$\begin{cases} \frac{\partial u}{\partial t} + \mathcal{A}u = 0, & \text{in } [0, T) \times \mathcal{D}, \\ u(T, y) = f(y), & \text{in } \mathcal{D}. \end{cases} \tag{4.3.4}$$

The above proverties (a), (b) and (c) will be assumed to hold throughout this section.

We are interested in the numerical evaluation of $u(0, Y_0) = \mathbb{E}(f(Y_T))$. A widely used and computationally convenient method is by computing the above expectation on an approximation of the process Y. Here, we consider an approximation through a Markov chain that weakly converges to the diffusion process Y, see e.g. the classical references [89]. We will see in Section 4.3.1 an application to tree methods, that is, when the process Y is approximated by means of a computationally simple Markov chain. Here, our aim is to study, under suitable but quite general assumptions, the order of weak convergence.

So, let $N \in \mathbb{N}^*$ and set $h = T/N$. The parameters N and h are fixed once for all. Let $(Y_n^h)_{n=0,\dots,N}$ denote a Markov chain, whose state space, at time-step n, is given by $\mathcal{Y}_n^h \subset \mathcal{D}$. In our mind, $(Y_n^h)_{n=0,\dots,N}$ is a Markov process which is a discrete weak approximation in time (and possibly in space) of the d-dimensional diffusion Y, namely, Y_n^h approximates Y at times nh, for every $n = 0, \dots, N$. Of course, we assume that $Y_0^h = Y_0$, that is, $\mathcal{Y}_0^h = \{Y_0\}$. Without loss of generality, we may assume that $(Y_n^h)_{n=0,\dots,N}$ is defined in $(\Omega, \mathcal{F}, \mathbb{P})$.

In order to study the rate of the weak convergence of $(Y_n^h)_{n=0,\dots,N}$ to Y, we need to stress the requests that are usually done in order to merely prove the convergence (see e.g. [89]). In particular, we need the following assumption.

Assumption $\mathcal{A}_1$. *There exists $\bar{h} > 0$ such that, for every $h < \bar{h}$, the first three local moments* *satisfy*

$$\mathbb{E}[Y^h_{n+1} - Y^h_n \mid Y^h_n] = \mu_Y(Y^h_n)h + f_h(Y^h_n), \tag{4.3.5}$$

$$\mathbb{E}[(Y^h_{n+1} - Y^h_n)(Y^h_{n+1} - Y^h_n)^\star \mid Y^h_n] = a_Y(Y^h_n)h + g_h(Y^h_n), \tag{4.3.6}$$

$$\mathbb{E}[(Y^h_{n+1} - Y^h_n)^l \mid Y^h_n] = j_{h,l}(Y^h_n), \qquad l \in \mathbb{N}^d,\ |l| = 3, \tag{4.3.7}$$

where $f_h : \mathcal{D} \to \mathbb{R}^d$*,* $g_h : \mathcal{D} \to \mathbb{R}^{d\times d}$ *and* $j_{h,l} : \mathcal{D} \to \mathbb{R}$ *satisfy the following properties: there exist* $p > 1$ *and* $C > 0$ *such that*

$$\sup_{h\le\bar h}\ \sup_{n=0,\dots,N} \|f_h(Y^h_n)\|_p \le Ch^2, \tag{4.3.8}$$

$$\sup_{h\le\bar h}\ \sup_{n=0,\dots,N} \|g_h(Y^h_n)\|_p \le Ch^2, \tag{4.3.9}$$

$$\sup_{h\le\bar h}\ \sup_{n=0,\dots,N} \|j_{h,l}(Y^h_n)\|_p \le Ch^2, \quad |l| = 3. \tag{4.3.10}$$

We also need the following behavior of the moments.

Assumption $\mathcal{A}_2$. *There exists* $\bar h > 0$ *such that for every* $p > 1$ *there exists* $C_p > 0$ *for which*

$$\sup_{h<\bar h}\ \sup_{0\le n\le N} \|Y^h_n\|_p \le C_p, \tag{4.3.11}$$

$$\sup_{h<\bar h}\ \sup_{0\le n\le N} \frac{1}{\sqrt{h}}\|Y^h_{n+1} - Y^h_n\|_p \le C_p. \tag{4.3.12}$$

We can now state the following first order weak convergence result.

Theorem 4.3.1. *Let assumptions* $\mathcal{A}_1$ *and* $\mathcal{A}_2$ *hold and assume that* $u \in C^4_{\mathbf{pol},T}(\mathcal{D})$*,* u *being defined in* (4.3.3)*. Then there exist* $\bar h > 0$ *and* $C > 0$ *such that for every* $h < \bar h$ *one has*

$$|\mathbb{E}[f(Y^h_N)] - \mathbb{E}[f(Y_T)]| \le CTh.$$

Proof. The proof is quite standard. Since $\mathbb{E}[f(Y^h_N)] = \mathbb{E}[u(T, Y^h_T)]$ and $\mathbb{E}[f(Y_T)] = u(0, Y_0)$, we have

$$\mathbb{E}[f(Y^h_T)] - \mathbb{E}[f(Y_T)] = \mathbb{E}[u(T, Y^h_T) - u(0, Y_0)] = \sum_{n=0}^{N-1} \mathbb{E}[u((n+1)h, Y^h_{n+1}) - u(nh, Y^h_n)].$$

Since $u \in C^4_{\mathbf{pol},T}(\mathcal{D})$, we can apply Taylor's formula to $t \mapsto u(t, y)$ around nh up to order 1 and to the functions $y \mapsto u(t, y)$ and $y \mapsto \partial_t u(t, y)$ around Y^h_n up to order 3 and 1 respectively. We obtain

$$u((n+1)h, Y^h_{n+1}) = \sum_{0\le|l|+2l'\le 3} \partial^l_y \partial^{l'}_t u(nh, Y^h_n)\frac{h^{l'}(Y^h_{n+1} - Y^h_n)^l}{|l|!l'!} + R_1(n, h, Y^h_n, Y^h_{n+1}), \tag{4.3.13}$$

where the remaining term R_1 is given by

$$\begin{aligned}R_1(n,h,Y_n^h,Y_{n+1}^h) &= h^2\int_0^1(1-\tau)\partial_t^2u(t+\tau h,Y_{n+1}^h)d\tau\\&+h\sum_{|k|=2}(Y_{n+1}^h-Y_n^h)^k\int_0^1(1-\xi)\partial_y^k\partial_t u(nh,Y_n^h+\xi(Y_{n+1}^h-Y_n^h))d\xi\\&+\sum_{|k|=4}\frac{(Y_{n+1}^h-Y_n^h)^k}{3!}\int_0^1(1-\xi)^3\partial_y^k u(nh,Y_n^h+\xi(Y_{n+1}^h-Y_n^h))d\xi.\end{aligned}$$

We now pass to the conditional expectation w.r.t. Y_n^h in (4.3.13) and use (4.3.5) and (4.3.6). By rearranging the terms we obtain

$$\begin{aligned}&\mathbb{E}[u((n+1)h,Y_{n+1}^h)-u(nh,Y_n^h)]\\&\quad=h\mathbb{E}\left[\partial_t u(nh,Y_n^h)+\mu_Y(Y_n^h)\cdot\nabla_y u(nh,Y_n^h)+\frac{1}{2}\mathrm{Tr}(a_Y D_y^2u(nh,Y_n^h))\right]+\sum_{i=1}^5R_n^i(h),\end{aligned}\tag{4.3.14}$$

in which

$$\begin{aligned}R_n^1(h)&=\mathbb{E}[R_1(n,h,V_n^h,V_{n+1}^h)], & R_n^2(h)&=h\mathbb{E}[(\mu_Y(Y_n^h)h+f_h(Y_n^h))\cdot\nabla_y\partial_t u(nh,Y_n^h)],\\R_n^3(h)&=\mathbb{E}[f_h(Y_n^h)\cdot\nabla_y u(nh,Y_n^h)], & R_n^4(h)&=\frac{1}{2}\mathbb{E}[\mathrm{Tr}(g_h(Y_n^h)D_y^2u(nh,Y_n^h))],\\R_n^5(h)&=\frac{1}{6}\sum_{|k|=3}\mathbb{E}[\partial_y^k u(nh,Y_n^h)j_{h,k}(Y_n^h)].\end{aligned}$$

Thanks to (4.3.4), the first term in (4.3.14) is null, so

$$|\mathbb{E}[u((n+1)h,Y_{n+1}^h)-u(nh,Y_n^h)]|\le\sum_{i=1}^5|R_n^i(h)|.$$

We now prove that $|R_n^i(h)|\le Ch^2$, for every $i=1,\dots 5$. Let $\bar h>0$ such that both assumptions $\mathcal{A}_1$ and $\mathcal{A}_2$ hold and let $h<\bar h$. Since the derivatives of u have polynomial growth, one has

$$|R_1(n,h,Y_n^h,Y_{n+1}^h)|\le C\Big(1+|Y_n^h|+|Y_{n+1}^h|\Big)^a\Big[h^2+h|Y_{n+1}^h-Y_n^h|^2+|Y_{n+1}^h-Y_n^h|^4\Big],$$

where $C,a>0$ denote constants that are independent of h and, from now on, may change from a line to another. Then, by using the Cauchy-Schwarz inequality, (4.3.11) and (4.3.12), we get

$$|R_n^1(h)|\le C\big\|(1+|Y_{n+1}^h|+|Y_n^h|)^a\big\|_2\,\big\|h^2+h(Y_{n+1}^h-Y_n^h)^2+(Y_{n+1}^h-Y_n^h)^4\big\|_2\le Ch^2.$$

As regards $R^2_n(h)$, we use the polynomial growth of $\nabla_y \partial_t u$, the Cauchy-Schwarz inequality and the Hölder inequality, so that

$$\begin{aligned}|R^2_n(h)| &\leq C\mathbb{E}[\big(1+|Y^h_n|^a\big)|\mu_Y(Y^h_n)|]\,h^2 + C\mathbb{E}[\big(1+|Y^h_n|^a\big)|f_h(Y^h_n)|]\\ &\leq C\big\|1+|Y^h_n|^a\big\|_2\,\big\|\mu_Y(Y^h_n)\big\|_2\,h^2 + C\big\|1+|Y^h_n|^a\big\|_q\,\big\|f_h(Y^h_n)\big\|_p,\end{aligned}$$

where p is given in (4.3.8) and q is its conjugate exponent. Since μ_Y has polynomial growth, by (4.3.8) and (4.3.11) we get

$$|R^2_n(h)| \leq Ch^2.$$

The remaining terms $R^3_n(h)$, $R^4_n(h)$ and $R^5_n(h)$ can be handled similarly, so the statement follows. □

4.3.1 An example: a first order weak convergent binomial tree for the CIR process

We now fix $d=1$ and $\mathcal{D}=\mathbb{R}_+=[0,\infty)$. We consider the CIR process $(Y_t)_{t\in[0,T]}$ solution to the SDE

$$dY_t = \kappa(\theta - Y_t)dt + \sigma\sqrt{Y_t}\,dW_t, \quad Y_0 \geq 0.$$

We assume that $\theta,\kappa,\sigma>0$ and we do not require the Feller condition. Therefore, the process Y can reach 0.

We consider here the "multiple jumps" tree approximation for the CIR process described in Section 3.3.1. We first briefly recall how the tree works and then, as an application of Theorem 4.3.1, we study the rate of convergence.

Recall that, for $n=0,1,\dots,N$ we have the lattice

$$\mathcal{Y}^h_n = \{y^n_k\}_{k=0,1,\dots,n} \quad \text{with} \quad y^n_k = \Big(\sqrt{Y_0} + \frac{\sigma}{2}(2k-n)\sqrt{h}\Big)^2 \mathbb{1}_{\{\sqrt{Y_0}+\frac{\sigma}{2}(2k-n)\sqrt{h}>0\}}. \tag{4.3.15}$$

Note that $\mathcal{Y}^h_0=\{Y_0\}$. For each fixed node $(n,k)\in\{0,1,\dots,N-1\}\times\{0,1,\dots,n\}$, the "up" jump $k_u(n,k)$ and the "down" jump $k_d(n,k)$ from $y^n_k\in\mathcal{Y}^h_n$ are defined as

$$k_u(n,k) = \min\{k^* \,:\, k+1\leq k^*\leq n+1 \text{ and } y^n_k + \mu_Y(y^n_k)h \leq y^{n+1}_{k^*}\}, \tag{4.3.16}$$

$$k_d(n,k) = \max\{k^* \,:\, 0\leq k^*\leq k \text{ and } y^n_k + \mu_Y(y^n_k)h \geq y^{n+1}_{k^*}\}, \tag{4.3.17}$$

where $\mu_Y(y) = \kappa(\theta - y)$ and with the understanding $k_u(n,k) = n+1$, resp. $k_d(n,k) = 0$, if the set in (4.3.16), resp. (4.3.17), is empty. In fact, starting from the node (n,k) the probability that the process jumps to $k_u(n,k)$ and $k_d(n,k)$ at time-step $n+1$ are set as

$$keyp_u(n,k) = 0 \vee \frac{\mu_Y(y^n_k)h + y^n_k - y^{n+1}_{k_d(n,k)}}{y^{n+1}_{k_u(n,k)} - y^{n+1}_{k_d(n,k)}} \wedge 1 \quad \text{and} \quad p_d(n,k) = 1 - p_u(n,k)$$

respectively. We will see in next Proposition 4.3.3 that for h small enough the parts "$0\vee$" and "$\wedge 1$" can be omitted.

We call $(Y^h_n)_{n=0,1,\dots,N}$ the Markov chain governed by the above jump probabilities. As an application of Theorem 4.3.1, we shall prove the following result.

Theorem 4.3.2. *Let $f \in C^4_{\mathbf{pol}}(\mathbb{R}_+)$. Then, there exist $\bar h > 0$ and $C > 0$ such that for every $h < \bar h$,*

$$|\mathbb{E}[f(Y^h_N)] - \mathbb{E}[f(Y_T)]| \leq CTh,$$

that is, the tree approximation $(Y^h_n)_{n=0,\dots,N}$ is first order weak convergent.

In order to discuss the assumptions $\mathcal{A}_1$ and $\mathcal{A}_2$ of Theorem 4.3.1, we need some preliminary results which pave the way to the analysis of the convergence.

Proposition 4.3.3. *There exist $\theta_*, \theta^*, C_*, \bar h > 0$ such that for any $h < \bar h$ the following properties hold.*

(i) If $\theta_* h \leq y^n_k \leq \theta^*/h$, then $k_u(n,k) = k+1$, $k_d(n,k) = k$. Moreover,

$$y^{n+1}_{k_u(n,k)} = y^n_k + \frac{\sigma^2}{4}h + \sigma\sqrt{y^n_k h} \quad \textit{and} \quad y^{n+1}_{k_d(n,k)} = y^n_k + \frac{\sigma^2}{4}h - \sigma\sqrt{y^n_k h}.$$

(ii) If $y^n_k < \theta_* h$, then $k_d(n,k) = k$. Moreover,

$$0 \leq y^{n+1}_{k_u(n,k)} - y^n_k \leq C_* h. \tag{4.3.18}$$

(iii) If $y^n_k > \theta^*/h$, then $k_u(n,k) = k+1$.

(iv) The jump probabilities are

$$p_u(n,k) = \frac{\mu_Y(y^n_k)h + y^n_k - y^{n+1}_{k_d(n,k)}}{y^{n+1}_{k_u(n,k)} - y^{n+1}_{k_d(n,k)}}, \qquad p_d(n,k) = \frac{y^n_{k_u(n,k)} - y^n_k - \mu_Y(y^n_k)h}{y^{n+1}_{k_u(n,k)} - y^{n+1}_{k_d(n,k)}}. \tag{4.3.19}$$

The proof of Proposition 4.3.3 relies on a boring study of the properties of the lattice, so we postpone it in Appendix 4.7.1. This is all we need to prove that $\mathcal{A}_2$ holds:

Proposition 4.3.4. *The CIR approximating tree $\{Y_n^h\}_{n=0,\dots,N}$ satisfies Assumption $\mathcal{A}_2$.*

Proof. **Step 1: proof of** (4.3.11)**.** We use a technique firstly developed in [3] for a CIR discretization scheme based on Brownian increments. The key point is the proof of a monotonicity property allowing one to control the moments of the tree: there exist $b, C, \bar h > 0$ such that for every $h < \bar h$ and $n = 0, \dots, N-1$ one has

$$0 \le Y_{n+1}^h \le (1+bh)Y_n^h + Ch + \sigma\sqrt{Y_n^h h}\, W_{n+1}^h, \tag{4.3.20}$$

where W_{n+1}^h is a r.v. such that

$$\mathbb{P}(W_{n+1}^h = 2p_d(n,k)|Y_n^h = y_k^n) = p_u(n,k) = 1 - \mathbb{P}(W_{n+1}^h = -2p_u(n,k)|Y_n^h = y_k^n). \tag{4.3.21}$$

To this purpose, fix a node (n,k). For the sake of simplicity, we write k_u, resp. k_d, in place of $k_u(n,k)$, resp. $k_d(n,k)$. We have (see (4.7.94)) that

$$y_{k+1}^{n+1} \le y_k^n + \frac{\sigma^2}{4}h + \sigma\sqrt{y_k^n h}, \qquad y_k^{n+1} \le y_k^n + \frac{\sigma^2}{4}h - \sigma\sqrt{y_k^n h}.$$

By Proposition 4.3.3, for $h < \bar h$, if $\theta_* h < y_k^n < \theta^*/h$ the up and down jumps are both single, hence $y_{k_u}^{n+1} = y_{k+1}^{n+1}$ and $y_{k_d}^{n+1} = y_k^{n+1}$ On the other hand, if $y_k^n \ge \theta^*/h$ the up jump is single, that is $y_{k_u}^{n+1} = y_{k+1}^{n+1}$, while the down jump can be multiple but, in every case, is still true that

$$y_{k_d}^{n+1} \le y_k^{n+1} = y_k^n + \frac{\sigma^2}{4}h - \sigma\sqrt{y_k^n h}.$$

Finally, if $y_k^n \le \theta_* h$, we have $y_{k_d}^{n+1} = y_k^{n+1}$, while the up jump can be multiple but we can always write

$$y_{k_u}^{n+1} \le y_k^n + C_* h \le y_k^n + C_* h + \sigma\sqrt{y_k^n h}.$$

Summing up, if we set $\bar C = \max\left(C_*, \frac{\sigma^2}{4}\right)$, for every h small we can write

$$0 \le Y_{n+1}^h \le Y_n^h + \bar C h + \sigma\sqrt{Y_n^h h}\, Z_{n+1}^h,$$

where Z_{n+1}^h is a random variable such that $\mathbb{P}(Z_{n+1}^h = +1|Y_n^h = y_k^n) = p_u(n,k)$ and $\mathbb{P}(Z_{n+1}^h = -1|Y_n^h = y_k^n) = p_d(n,k)$. Note that $\mathbb{E}(Z_{n+1}^h|Y_n^h = y_k^n) = p_u(n,k) - p_d(n,k) = 2p_u(n,k) - 1$. Then, the random variable

$$W_{n+1}^h = Z_{n+1}^h - \mathbb{E}[Z_{n+1}^h|Y_n^h]$$

has exactly the law given in (4.3.21). We also define the function $P_u(y^n_k) = p_u(n,k)$. Therefore,

$$
\begin{aligned}
0 \leq & Y^h_{n+1} \leq Y^h_n + \bar{C}h + \sigma\sqrt{Y^h_n h}\,(2P_u(Y^h_n) - 1) + \sigma\sqrt{Y^h_n h}\,W^h_{n+1} \\
\leq & Y^h_n + \bar{C}h + \sigma\sqrt{\theta^*}\sqrt{\frac{Y^h_n h}{\theta^*}}\,\big|2P_u(Y^h_n) - 1\big|\mathbb{1}_{\{Y^h_n \geq \frac{\theta^*}{h}\}} + \sigma\sqrt{Y^h_n h}\,\big(2P_u(Y^h_n) - 1\big)\mathbb{1}_{\{Y^h_n < \frac{\theta^*}{h}\}} \\
& + \sigma\sqrt{Y^h_n h}\,W^h_{n+1}.
\end{aligned}
$$

Now, if $Y^h_n \geq \frac{\theta^*}{h}$ then $\sqrt{\frac{Y^h_n h}{\theta^*}} \leq \frac{Y^h_n h}{\theta^*}$ and, since $P_u \in [0,1]$, we have $|2P_u(Y^h_n) - 1| \leq 1$. Then,

$$
0 \leq Y^h_{n+1} \leq (1 + bh)Y^h_n + \bar{C}h + \sigma\sqrt{Y^h_n h}\,\big(2P_u(Y^h_n) - 1\big)\mathbb{1}_{\{Y^h_n < \frac{\theta^*}{h}\}} + \sigma\sqrt{Y^h_n h}\,W^h_{n+1},
$$

where $b = \frac{\sigma}{\sqrt{\theta^*}}$. Let us study the quantity $\sigma\sqrt{Y^h_n h}\,(2P_u(Y^h_n) - 1)\mathbb{1}_{\{Y^h_n < \frac{\theta^*}{h}\}}$. If $\theta_* h < y^n_k < \theta^*/h$, by using (4.3.19) and point 1. of Proposition 4.3.3, we can explicitly write

$$
\sigma\sqrt{y^n_k h}\,(2P_u(y^n_k) - 1) = \sigma\sqrt{y^n_k h}\,\Big(2\Big(\frac{1}{2} + \frac{4\mu_Y(v^n_k) - \sigma^2}{8\sigma\sqrt{y^n_k h}}\Big)h - 1\Big) = \mu_Y(v^n_k)h - \frac{\sigma^2}{4}h \leq \kappa\theta h.
$$

If instead $y^n_k \leq \theta_* h$, then by using 2. in Proposition 4.3.3 we have

$$
\begin{aligned}
\sigma\sqrt{y^n_k h}\,(2P_u(y^n_k) - 1) &= \sigma\sqrt{y^n_k h}\,\frac{2\mu_Y(y^n_k)h + 2y^n_k - y^{n+1}_{k_d(n,k)} - y^{n+1}_{k_u(n,k)}}{y^{n+1}_{k_u(n,k)} - y^{n+1}_{k_d(n,k)}} \\
&\leq \sigma\sqrt{y^n_k h}\,\frac{2\mu_Y(y^n_k)h + 2y^n_k}{y^{n+1}_{k+1} - y^{n+1}_k} \leq \sigma\sqrt{y^n_k h}\,\frac{2\kappa\theta h + 2\theta_* h}{2\sigma\sqrt{y^n_k h}} = (\kappa\theta + \theta_*)h.
\end{aligned}
$$

So, by inserting, for every $n \leq N - 1$ we get

$$
0 \leq Y^h_{n+1} \leq (1 + bh)Y^h_n + \bar{C}h + \sigma(\kappa\theta + \theta_*)h + \sigma\sqrt{Y^h_n h}\,W^h_{n+1}
$$

and (4.3.20) is proved.

Now, we repeat step by step the proof of Lemma 2.6 in [3] in order to get (4.3.11). We use induction on p. For $p = 1$, by definition one has $\mathbb{E}[Y^h_{n+1}|Y^h_n] = Y^h_n + \mu_Y(Y^h_n)h$ and, by passing to the expectation, $\mathbb{E}[Y^h_{n+1}] = \mathbb{E}[Y^h_n] + \mathbb{E}[\mu_Y(Y^h_n)h] \leq \mathbb{E}[Y^h_n] + \kappa\theta h$, from which we obtain $\mathbb{E}[Y^h_{n+1}] \leq Y_0 + \kappa\theta(n+1)h \leq Y_0 + \kappa\theta T$ and the case $p = 1$ is proved. So, assume that (4.3.11) holds for $p - 1$ and let us prove its validity for p. Using (4.3.20), we have

$$
\mathbb{E}[(Y^h_{n+1})^p] \leq \sum_{l_1 + l_2 + l_3 = p} \frac{p!}{l_1! l_2! l_3!}(1 + bh)^{l_1}\sigma^{l_2}C^{l_3}\mathbb{E}\left[(Y^h_n)^{l_1 + \frac{l_2}{2}}h^{l_3 + \frac{l_2}{2}}(W^h_{n+1})^{l_2}\right].
$$

So, it is sufficient to control $\mathcal{E}(l_1, l_2, l_3) = \mathbb{E}\left[(Y^h_n)^{l_1 + \frac{l_2}{2}}h^{l_3 + \frac{l_2}{2}}(W^h_{n+1})^{l_2}\right]$ for $l_1 + l_2 + l_3 = p$.

Assume first that $l_1 + \frac{l_2}{2} \leq p - \frac{3}{2}$, a case giving $l_3 + \frac{l_2}{2} \geq \frac{3}{2}$. Without loss of generality we can assume $C_{p-1} \geq 1$. Moreover, recall that $|W^h_{n+1}| \leq 2$. By using the Hölder's inequality with $\alpha = \frac{p-1}{l_1+\frac{l_2}{2}}$, we get

$$\mathcal{E}(l_1,l_2,l_3) \leq |\mathcal{E}(l_1,l_2,l_3)| \leq \mathbb{E}\left[(Y^h_n)^{l_1+\frac{l_2}{2}}\right] 2^{l_2} h^{l_3+\frac{l_2}{2}} \leq C_{p-1}2^{l_2}h^{\frac{3}{2}}.$$

Therefore

$$\sum_{\substack{l_1+l_2+l_3=p\\ l_1+l_2/2\leq p-3/2}} \frac{p!}{l_1!l_2!l_3!}(1+bh)^{l_1}\sigma^{l_2}C^{l_3}\mathcal{E}(l_1,l_2,l_3) \leq C_{p-1}h^{\frac{3}{2}} \sum_{l_1+l_2+l_3=p} \frac{p!}{l_1!l_2!l_3!}(1+bh)^{l_1}(2\sigma)^{l_2}C^{l_3}$$
$$\leq C_{p-1}h^{\frac{3}{2}}(1+b+2\sigma+C)^p.$$

The case $l_1 + \frac{l_2}{2} > p - \frac{3}{2}$ gives 4 further contributions, namely $(l_1,l_2,l_3) = (p,0,0)$, $(p-1,0,1)$, $(p-1,1,0)$ and $(p-2,2,0)$. So, we get

$$\mathbb{E}[(Y^h_{n+1})^p] \leq C_{p-1}(1+b+2\sigma+C)^p h^{\frac{3}{2}} + (1+bh)^p\mathbb{E}[(Y^h_n)^p] + p(1+bh)^{p-1}Ch\mathbb{E}[(Y^h_n)^{p-1}]$$
$$+ p(1+bh)^{p-1}\sigma Ch^{1/2}\mathbb{E}[(Y^h_n)^{p-1/2}W^h_{n+1}] + \frac{p(p-1)}{2}(1+bh)^{p-2}\sigma^2 h\mathbb{E}[(Y^h_n)^{p-1}(W^h_{n+1})^2].$$

Consider the last two terms above. For the first, we note that

$$\mathbb{E}[(Y^h_n)^{p-1/2}W^h_{n+1}] = \mathbb{E}[(Y^h_n)^{p-1/2}\mathbb{E}[W^h_{n+1}|^hY_n]] = 0$$

and for the second, we recall that $|W^h_{n+1}| \leq 2$. So, we easily obtain

$$\mathbb{E}[(Y^h_{n+1})^p] \leq C_{p-1}h(1+b+2\sigma+C)^p\Big[1+p+\frac{p(p-1)}{2}\Big] + (1+bh)^p\mathbb{E}[(Y^h_n)^p].$$

By recursion on n, we get

$$\mathbb{E}[(Y^h_{n+1})^p] \leq C_{p-1}h(1+b+2\sigma+C)^p\,\frac{p^2+p+2}{2}\sum_{j=0}^{n}(1+bh)^{jp} + Y_0^p(1+bh)^{(n+1)p}$$

and 4.3.11 now follows.

Step 2: proof of (4.3.12). We can write

$$|Y^h_{n+1} - Y^h_n|^p \leq 3^{p-1}\Big|\frac{\sigma^2}{4}h + \sigma\sqrt{Y^h_n h}Z^h_{n+1}\Big|^p \mathbb{1}_{\{\theta_* h < Y^h_n < \theta^*/h\}} + 3^{p-1}|Y^h_{n+1} - Y^h_n|^p\mathbb{1}_{\{Y^h_n \leq \theta_* h\}}$$
$$+ 3^{p-1}|Y^h_{n+1} - Y^h_n|^p\mathbb{1}_{\{Y^h_n \geq \theta^*/h\}} =: 3^{p-1}(I_1+I_2+I_3),$$

where we have used that, on the set $\{\theta_* h < Y^h_n < \theta^*/h\}$, we have $Y^h_{n+1} = Y^h_n + \frac{\sigma^2}{4}h + \sigma\sqrt{Y^h_n h}Z^h_{n+1}$, with $\mathbb{P}(Z^h_{n+1} = 1 \mid Y^h_{n+1}) = P_u(Y^h_n)$ and $\mathbb{P}(Z^h_{n+1} = -1 \mid Y^h_{n+1}) = P_d(Y^h_n)$. Now, by using (4.3.11), Proposition 4.3.3, the Cauchy-Swartz and the Markov inequality,

$$I_1 \leq \mathbb{E}\Big[\Big(\frac{\sigma^2}{4}h + \sigma\sqrt{Y^h_n h}\Big)^p\Big] \leq 2^{p-1}\Big(\Big(\frac{\sigma^2}{4}\Big)^p + \sigma^p \mathbb{E}[(Y^h_n)^p]^{1/2}\Big)h^{p/2} \leq 2^{p-1}\Big(\Big(\frac{\sigma^2}{4}\Big)^p + \sigma^p\sqrt{C_p}\Big)h^{p/2},$$

$$I_2 \leq C_*^p h^p,$$

$$I_3 \leq \mathbb{E}[(Y^h_{n+1} - Y^h_n)^{2p}]^{1/2}\mathbb{P}\Big(Y^h_n > \frac{\theta^*}{h}\Big)^{1/2} \leq 2^p\sqrt{\frac{C_{2p}C_p}{(\theta^*)^p}}\,h^{p/2},$$

and (4.3.12) follows. □

Proposition 4.3.5. *The CIR approximating tree $\{Y^h_n\}_{n=0,\dots,N}$ satisfies Assumption $\mathcal{A}_1$.*

Proof. Straightforward computations give $\mathbb{E}[Y^h_{n+1} - Y^h_n \mid Y^h_n] = \mu_Y(Y^h_n)h$, so (4.3.5) and (4.3.8) immediately follow. As for (4.3.6),

$$\begin{aligned}\mathbb{E}[(Y^h_{n+1} - Y^h_n)^2 \mid Y^h_n = y^n_k] &= \mathbb{E}[(Y^h_{n+1} - Y^h_n)^2 \mid Y^h_n = y^n_k]\mathbb{1}_{\{y^n_k \leq \theta_* h\}}\\ &+ \mathbb{E}[(Y^h_{n+1} - Y^h_n)^2 \mid Y^h_n = y^n_k]\mathbb{1}_{\{\theta_* h \leq y^n_k \leq \theta^*/h\}} + \mathbb{E}[(Y^h_{n+1} - Y^h_n)^2 \mid Y^h_n = y^n_k]\mathbb{1}_{\{y^n_k > \theta^*/h\}}.\end{aligned}$$

We study separately the first two terms of the above r.h.s. If $y^n_k < \theta_* h$, Proposition 4.3.3 gives $|y^{n+1}_{k_u} - y^n_k| \leq C_* h$ and $|y^{n+1}_{k_d} - y^n_k| \leq C_* h$ so that

$$\mathbb{E}[(Y^h_{n+1} - Y^h_n)^2 \mid Y^h_n = y^n_k]\mathbb{1}_{\{y^n_k \leq \theta_* h\}} = \varphi_1(y^n_k)h^2\mathbb{1}_{\{y^n_k \leq \theta_* h\}},$$

with φ_1 such that $|\varphi_1(y)| \leq C_*^2$. If instead $\theta_* h \leq y^n_k \leq \theta^*/h$, by using (4.3.19) we get

$$(y^{n+1}_{k_u} - y^n_k)^2 p_u(n,k) + (y^{n+1}_{k_d} - y^n_k)^2 p_d(n,k) = \sigma^2 y^n_k h + \frac{\sigma^2}{2}\Big(\kappa(\theta - y^n_k) - \frac{\sigma^2}{8}\Big)h^2.$$

So,

$$\mathbb{E}[(Y^h_{n+1} - Y^h_n)^2 \mid Y^h_n = y^n_k]\mathbb{1}_{\{\theta_* h \leq y^n_k \leq \theta^*/h\}} = \big(\sigma^2 y^n_k h + \varphi_2(y^n_k)h^2\big)\mathbb{1}_{\{\theta_* h \leq y^n_k \leq \theta^*/h\}},$$

with φ_2 such that $|\varphi_2(y)| \leq \frac{\sigma^2}{2}\Big(\kappa(\theta + y) + \frac{\sigma^2}{8}\Big)$. By inserting, (4.3.6) follows with g_h satisfying

$$|g_h(Y^h_n)| \leq c_1(1 + Y^h_n)h^2 + \mathbb{E}((Y^h_{n+1} - Y^h_n)^2 + \sigma h Y^h_n \mid Y^h_n)\mathbb{1}_{\{Y^h_n \geq \theta^*/h\}},$$

c_1 denoting a suitable constant. By Proposition 4.3.4 and the Markov inequality, (4.3.9) follows.

Finally, for (4.3.7), we write

$$\begin{aligned}\mathbb{E}[(Y^h_{n+1} - Y^h_n)^3 \mid Y^h_n = y^n_k] &= \mathbb{E}[(Y^h_{n+1} - Y^h_n)^3 \mid Y^h_n = y^n_k]\mathbb{1}_{\{y^n_k \leq \theta_* h\}}\\ &+ \mathbb{E}[(Y^h_{n+1} - Y^h_n)^3 \mid Y^h_n = y^n_k]\mathbb{1}_{\{\theta_* h < y^n_k < \theta^*/h\}} + \mathbb{E}[(Y^h_{n+1} - Y^h_n)^3 \mid Y^h_n = y^n_k]\mathbb{1}_{\{y^n_k \geq \theta^*/h\}}.\end{aligned}$$

Now, if $y_k^n \le \theta_* h$ then $|Y_{n+1}^h - y_k^n|^3 \le C_*^3 h^3$. If instead $\theta_* h < y_k^n < \theta^*/h$, by (4.3.19) one obtains

$$(y_{k_u}^{n+1} - y_k^n)^3 p_u(n,k) + (y_{k_d}^{n+1} - y_k^n)^3 p_d(n,k) = \mu_Y(y_k^n) h^2 \Big(\sigma^2 y_k^n + \frac{3\sigma^4}{16}\, h\Big) + \Big(\frac{\sigma^4}{2}\, y_k^n + \frac{\sigma^4}{16}\, h\Big) h^2.$$

Therefore,

$$|j_h(Y_n^h)| \le c_2 h^2 (1 + (Y_n^h)^2) + \mathbb{E}(|Y_{n+1}^h - Y_n^h|^3 + \sigma h Y_n^h \mid Y_n^h) \mathbb{1}_{\{Y_n^h \ge \theta^*/h\}},$$

c_2 denoting a suitable constant, and again by Proposition 4.3.4 and the Markov inequality, (4.3.10) follows. □

We are finally ready for the

Proof of Theorem 4.3.2. By Theorem 4.1 in [3] (or Corollary 4.5.5), one has that if $f \in C^4_{\mathbf{pol}}(\mathbb{R}_+)$ then $u \in C^4_{\mathbf{pol},T}(\mathbb{R}_+)$. Since Assumption $\mathcal{A}_1$ and $\mathcal{A}_2$ both hold, the statement follows as an application of Theorem 4.3.1. □

4.4 Hybrid schemes for jump-diffusions and convergence rate

We now introduce a m-dimensional jump-diffusion $(X_t)_{t\in[0,T]}$ whose dynamics is given by coefficients depending on the process $(Y_t)_{t\in[0,T]}$ discussed in Section 4.3. More precisely, we consider the stochastic system

$$\begin{cases} dX_t = \mu_X(Y_t)dt + \sigma_X(Y_t)\, dB_t + \gamma_X(Y_t) dH_t, & X_0 \in \mathbb{R}^m, \\ dY_t = \mu_Y(Y_t)dt + \sigma_Y(Y_t)\, dW_t, & Y_0 \in \mathcal{D}, \end{cases} \tag{4.4.22}$$

where B is a ℓ_1-dimensional Brownian motion and H is a ℓ_2- dimensional compound Poisson process with intensity λ and i.i.d. jumps $\{J_k\}_k$, that is

$$H_t = \sum_{k=1}^{K_t} J_k, \tag{4.4.23}$$

K denoting a Poisson process with intensity λ. We assume that the Poisson process K, the jump amplitudes $\{J_k\}_k$ and the Brownian motions B and W are independent. Moreover, we ask that J_1 has a density p_{J_1}, so that the Lévy measure associated with H has a density as well:

$$\nu(dx) = \nu(x)dx = \lambda p_{J_1}(x) dx.$$

Hereafter, we denote by $\mathcal{L}$ the infinitesimal generator associated with the diffusion pair (X, Y), i.e.

$$\begin{aligned}\mathcal{L}g(x,y) = \; & \frac{1}{2}\mathrm{Tr}(a(y)D^2_{x,y}g(x,y)) + \mu(y)\cdot\nabla_{x,y}g(x,y) \\ & + \gamma_X(y)\int (g(x+\zeta,y) - g(x,y))\nu(d\zeta),\end{aligned} \tag{4.4.24}$$

where $\mu(y) = (\mu_X(y), \mu_Y(y))^\star$ and $a(y) = \sigma\sigma^\star(y)$, where

$$\sigma(y) = \begin{pmatrix} \sigma_X(y) & 0_{m\times d} \\ 0_{d\times m} & \sigma_Y(y) \end{pmatrix}.$$

Here, $D^2_{x,y}$ and $\nabla_{x,y}$ are respectively the Hessian and the gradient operator w.r.t. the space variables x and y. We assume that the coefficients of X do not depend on the time variable just to simplify the notation, but all the proofs in this chapter are still valid in the time-depending case under non restrictive classical assumptions.

Let $(X^{t,x,y}_s, Y^{t,x}_s)_{s\in[t,T]}$ be the solution of (4.4.22) with starting condition $(X_t, Y_t) = (x, y)$. Hereafter, we fix $T > 0$ and $f : \mathbb{R}^m \times \mathcal{D} \to \mathbb{R}$. We are interested in computing the quantity $u(0, X_0, Y_0)$, where, as specified from time to time, u is given by

$$u(t,x,y) = \mathbb{E}\Big[f(X^{t,x,y}_T, Y^{t,y}_T)\Big], \qquad (t,x,y) \in [0,T]\times\mathbb{R}^m\times\mathcal{D}, \tag{4.4.25}$$

or

$$u(t,x,y) = \sup_{\tau\in\mathcal{T}_{t,T}} \mathbb{E}\Big[f(X^{t,x,y}_\tau, Y^{t,y}_\tau)\Big], \qquad (t,x,y) \in [0,T]\times\mathbb{R}^m\times\mathcal{D}, \tag{4.4.26}$$

where $\mathcal{T}_{t,T}$ denotes the set of all stopping times taking values on $[t, T]$.

This can be, in general, a problem of interest in a large number of applications. Of course, the immediate application in this thesis is in the financial world, where X can represent the log-price (or a transformation of it) and Y can be interpreted as a random source such as a stochastic volatility and/or a stochastic interest rate. In this framework, the function defined in (4.4.25) is the price value at time t of a European option with maturity T and (discounted) payoff f, while the function u as defined in (4.4.26) is the value function of the corresponding American option. Therefore, from now on we will refer to the European case when u is defined as in (4.4.25) and to the American case where u is given by (4.4.26).

From now on, the following assumptions (1), (2) and (3) will be in force throughout this chapter:

(1) there exists a unique weak solution of (4.4.22) such that $\mathbb{P}((X_t, Y_t) \in \mathbb{R}^m \times \mathcal{D}\ \forall t) = 1$;

(2) μ and σ have polynomial growth;

(3) the function u in (4.4.25) solves the PDE

$$\begin{cases} \partial_t u(t,x,y) + \mathcal{L}u(t,x,y) = 0 & (t,x,y) \in [0,T) \times \mathbb{R}^m \times \mathcal{D}, \\ u(T,x,y) = f(x,y), & \text{in } \mathbb{R}^m \times \mathcal{D}. \end{cases} \tag{4.4.27}$$

4.4.1 The hybrid procedure

The European case

Let u be given in (4.4.25). We study here the computation of $u(0, X_0, Y_0)$ by a backward hybrid algorithm which generalizes the procedure developed in [24, 25, 27] and described in Chapter 3. Roughly speaking, one uses a Markov chain in order to approximate the process Y and a different numerical procedure to handle the jump-diffusion component X. Let us briefly recall the main ideas and set up the approximation of u.

We start from the representation of $u(t,x,y)$ at times nh, $h = T/N$ and $n = 0, \dots, N$, by the usual (backward) dynamic programming principle: for $(x,y) \in \mathbb{R}^m \times \mathcal{D}$,

$$\begin{cases} u(T,x,y) = f(x,y) \quad \text{and as } n = N-1, \dots, 0, \\ u(nh,x,y) = \mathbb{E}\Big[u\big((n+1)h, X^{nh,x,y}_{(n+1)h}, Y^{nh,y}_{(n+1)h}\big)\Big]. \end{cases} \tag{4.4.28}$$

So, the central issue is to have a good approximation of the expectations in (4.4.28).

As a first step, let $(Y^h_n)_{n=0,\dots,N}$ be the Markov chain discussed in Section 4.4.2 which approximates Y. Of course, we assume that $(Y^h_n)_{n=0,\dots,N}$ is independent of the Brownian motion B and the compound Poisson process H driving X in (4.4.22). Then, at each step $n = 0, 1, \dots, N-1$, for every $y \in \mathcal{Y}^h_n$ we write

$$\mathbb{E}\Big[u\big((n+1)h, X^{nh,x,y}_{(n+1)h}, Y^{nh,y}_{(n+1)h}\big)\Big] \approx \mathbb{E}\Big[u\big((n+1)h, X^{nh,x,y}_{(n+1)h}, Y^h_{n+1}\big) \big| Y^h_n = y\Big].$$

Recall that $\mathcal{Y}^h_n \subseteq \mathcal{D}$ is the state space of Y^h_n and that $\mathcal{Y}^h_0 = \{Y_0\}$.

As a second step, we approximate the component X on $[nh, (n+1)h]$ by freezing the coefficients in (4.4.22) at the observed position $Y^h_n = y$, that is, for $t \in [nh, (n+1)h]$,

$$X^{nh,x,y}_t \overset{\text{law}}{\approx} \widehat{X}^{nh,x}_t(y) = x + \mu_X(y)(t-nh) + \sigma_X(y)\,(B_t - B_{nh}) + \gamma_X(y)(H_t - H_{nh}).$$

Therefore, by using that the Markov chain, B and H are all independent, we write

$$\begin{aligned}\mathbb{E}\Big[u\big((n+1)h, X^{nh,x,y}_{(n+1)h}, Y^{nh,y}_{(n+1)h}\big)\Big] &\approx \mathbb{E}\Big[u\big((n+1)h, \widehat{X}^{nh,x}_{(n+1)h}(y), Y^h_{n+1}\big)\big|Y^h_n = y\Big]\\ &= \mathbb{E}\big[\phi(Y^h_{n+1}; x, y)\big|Y^h_n = y\big],\end{aligned}$$

where

$$\phi(\zeta; x, y) = \mathbb{E}\big[u((n+1)h, \widehat{X}^{nh,x}_{(n+1)h}(y), \zeta)\big]. \tag{4.4.29}$$

From the Feynman-Kac formula, one gets $\phi(\zeta; x, y) = v(nh, x; y, \zeta)$, where $(t, x) \mapsto v(t, x; y, \zeta)$ is the solution at time nh of the parabolic PIDE Cauchy problem

$$\begin{aligned}&\partial_t v + \mathcal{L}^{(y)} v = 0, && \text{in } [nh, (n+1)h) \times \mathbb{R}^m,\\ &v((n+1)h, x; y, \zeta) = u((n+1)h, x, \zeta), && x \in \mathbb{R}^m,\end{aligned} \tag{4.4.30}$$

where $\mathcal{L}^{(y)}$ is the integro-differential operator acting on the functions $g = g(x)$ given by

$$\mathcal{L}^{(y)} g(x) = \mu_X(y) \cdot \nabla_x g(x) + \frac{1}{2}\,\mathrm{Tr}(a_X(y) D^2_x g(x)) + \gamma_X(y) \cdot \int \big(g(x+\zeta) - g(x)\big)\nu(\zeta)d\zeta. \tag{4.4.31}$$

Here $a_X(y) = \sigma_X(y)\sigma^\star_X(y)$, while ∇_x and D^2_x are the m dimensional gradient vector and the Hessian matrix with respect to the x variable respectively. Recall that here y is just a parameter and that for each fixed $y \in \mathcal{D}$, $\mathcal{L}^{(y)}$ has constant coefficients.

We consider now a numerical solution of the PIDE (4.4.30). Let $\Delta x = (\Delta x_1, \dots, \Delta x_m)$ denote a fixed spatial step and set $\mathcal{X}$ denote a grid on $\mathbb{R}^m$ given by $\mathcal{X} = \{x \,:\, x = ((X_0)_1 + i_1\Delta x_1, \dots, (X_0)_m + i_m\Delta x_m), (i_1, \dots, i_m) \in \mathbb{Z}^m\}$. For $y \in \mathcal{D}$, let $\Pi^h_{\Delta x}(y)$ be a linear operator (acting on suitable functions on $\mathcal{X}$) which gives the approximating solution to the PIDE (4.4.30) at time nh. Then we get the numerical approximation

$$\mathbb{E}\Big[u\big((n+1)h, X^{nh,x,y}_{(n+1)h}, Y^{nh,y}_{(n+1)h}\big)\Big] \approx \mathbb{E}\Big[\Pi^h_{\Delta x}(y) u\big((n+1)h, \cdot, Y^h_{n+1}\big)(x)\big|Y^h_n = y\Big], \quad x \in \mathcal{X}.$$

Therefore, by inserting in (4.4.28), the hybrid numerical procedure works as follows: the function $x \mapsto u(0, x, Y_0)$, $x \in \mathcal{X}$, is approximated by $u^h_0(x, Y_0)$ backwardly defined as

$$\begin{cases} u^h_N(x, y) = f(x, y), \quad (x, y) \in \mathcal{X} \times \mathcal{Y}^h_N, \quad \text{and as } n = N-1, \dots, 0:\\ u^h_n(x, y) = \mathbb{E}[\Pi^h_{\Delta x}(y) u^h_{n+1}(\cdot, Y^h_{n+1})(x) \mid Y^h_n = y], \quad (x, y) \in \mathcal{X} \times \mathcal{Y}^h_n. \end{cases} \tag{4.4.32}$$

The American case

Let us now consider the function u defined in (4.4.26). Again, we want an approximation of the quantity $u(0, X_0, Y_0)$. In practice, at times nh, the function u is approximated by the function $\tilde{u}^h_n$ defined through the backward programming dynamic principle, that is,

$$\begin{cases} \tilde{u}^h_N(x,y) = f(x,y) \quad \text{and as } n = N-1,\dots,0 \\ \tilde{u}^h_n(x,y) = \max\Big\{f(x,y), \mathbb{E}\Big[\tilde{u}^h_{n+1}\big(X^{nh,x,y}_{(n+1)h}, Y^{nh,y}_{(n+1)h}\big)\Big]\Big\}. \end{cases} \tag{4.4.33}$$

In financial terms, $\tilde{u}^h_0$ corresponds to approximate the original continuous time American option price at $t = 0$ by the price of an option which can be exercised only at the discrete times nh, $n = 0,\dots,N$ (Bermudean option).

Now, at each step of (4.4.33), we can use the procedure described in Section 4.4.1 in order to compute the conditional expectations therein. Therefore, the hybrid numerical procedure becomes: for $n = 0, 1, \dots, N$ and $(x,y) \in \mathcal{X} \times \mathcal{Y}^h_n$, $\tilde{u}^h_n(x,y)$ is approximated by $u^h_n(x,y)$ defined as

$$\begin{cases} u^h_N(x,y) = f(x,y), \quad \text{and as } n = N-1,\dots,0\text{:} \\ u^h_n(x,y) = \max\Big\{f(x,y), \mathbb{E}[\Pi^h_{\Delta x}(y) u^h_{n+1}(\cdot, \bar{Y}^{nh,y}_{(n+1)h})(x)]\Big\}. \end{cases} \tag{4.4.34}$$

The general hybrid procedure

As we have done in Chapter 3, it is useful to put together in a unique formulation the numerical procedures described respectively in Section 4.4.1 for the European case and in Section 4.4.1 for the American case. In both cases we have to consider at time nh the function $\tilde{u}^h_n$ defined as

$$\begin{cases} \tilde{u}^h_N(x,y) = f(x,y) \quad \text{and as } n = N-1,\dots,0 \\ \tilde{u}^h_n(x,y) = \max\Big\{g(x,y), \mathbb{E}\Big[\tilde{u}^h_{n+1}\big(X^{nh,x,y}_{(n+1)h}, Y^{nh,y}_{(n+1)h}\big)\Big]\Big\}, \end{cases} \tag{4.4.35}$$

where

$$g(x,y) = \begin{cases} 0, & \text{in the European case;} \\ f(x,y), & \text{in the American case.} \end{cases}$$

We stress that, in the European case, the function $\tilde{u}^h_n$ coincides with the function u defined in (4.4.25) at time nh, while, in the American case, it is the Bermudean approximation of the (continuous monitored) American option value given in (4.4.33).

Then, for $n = 0, 1, \dots, N$ and $(x, y) \in \mathcal{X} \times \mathcal{Y}_n^h$, we approximate the function $\tilde{u}_n^h$ by the function u_n^h defined as

$$\begin{cases} u_N^h(x,y) = f(x,y), \quad \text{and as } n = N-1, \dots, 0: \\ u_n^h(x,y) = \max\left\{ g(x,y), \mathbb{E}\left[\Pi_{\Delta x}^h(y) u_{n+1}^h(\cdot, \bar{Y}_{(n+1)h}^{nh,y})(x)\right]\right\}. \end{cases} \tag{4.4.36}$$

Our aim is to study the speed of convergence of the scheme (4.4.36) that is, we give a quantitative estimate for

$$|\tilde{u}_0^h(x,y) - u_0^h(x,y)|, \qquad (x,y) \in \mathcal{X} \times \mathcal{Y}_0^h.$$

As regards the American case, we recognize two types of error. The first one is the error induced by the approximation of the function $u(0, \cdot)$ in (4.4.26) with the function $\tilde{u}_0^h(\cdot)$ in the backward programming principle (4.4.33). In the standard hypotheses on the model, that is, for sublinear and Lipschitz continuous diffusion coefficients and standard semiconvex payoff function, this error is known to be of the first order in h (we refer, for example, to Theorem 2 in [13]). The degenerate models such as the Heston model do not satisfy such requests, so we might just argue a first order error in time. The second type of error is the one related to the approximation of $\tilde{u}_0^h$ with the function u_0^h defined in (4.4.34). Here, we focus on studying the latter one.

4.4.2 Convergence speed of the hybrid scheme

The idea is to follow the hybrid nature of the procedure by using numerical techniques, that is, an analysis of the stability and of the consistency of the method. This will be done in a sense that allows us to exploit the probabilistic properties of the Markov chain approximating the process Y.

We introduce the following assumption on the linear operator $\Pi_{\Delta x}^h(y)$ in (4.4.32) (recall the notation $l_p(\mathcal{X})$ in Section 4.2).

Assumption $\mathcal{B}(p, c, \mathcal{E})$. *Let $p \in [1, \infty]$, $c = c(y) \geq 0$, $y \in \mathcal{D}$ and $\mathcal{E} = \mathcal{E}(h, \Delta x) \geq 0$ such that $\lim_{(h,\Delta x)\to 0} \mathcal{E}(h, \Delta x) = 0$. We say that the linear operator $\Pi_{\Delta x}^h(y) : l_p(\mathcal{X}) \to l_p(\mathcal{X})$, $y \in \mathcal{D}$, satisfies Assumption $\mathcal{B}(p, c, \mathcal{E})$ if*

$$|\Pi_{\Delta x}^h(y)|_p \leq 1 + c(y)h \tag{4.4.37}$$

and, $\tilde{u}_n^h$ being defined in (4.4.35), for every $n = 0, \dots, N-1$, one has

$$\mathbb{E}\Big[\Pi_{\Delta x}^h(Y_n^h)\tilde{u}_h^{n+1}(\cdot, Y_{n+1}^h)(x) \,\Big|\, Y_n^h = y\Big] = \mathbb{E}[\tilde{u}_n^h(X^{nh,x,y}, Y_n^{nh,y})] + \mathcal{R}_n^h(x,y), \tag{4.4.38}$$

where the remainder $\mathcal{R}_n^h(x,y)$, $(x,y)\in\mathcal{X}\times\mathcal{Y}_n^h$ *satisfies the following property: there exist* $\bar h<1$ *and* $C>0$ *such that for every* $n\in\mathbb{N}$, $h<\bar h$, $|\Delta x|<1$ *and* $n\le N=\lfloor T/h\rfloor$ *one has*

$$\begin{aligned}\left\| e^{\sum_{l=1}^{n} c(Y_l^h)h} |\mathcal{R}_n^h(\cdot,Y_n^h)|_p \right\|_p &\le Ch\mathcal{E}(h,\Delta x), \qquad \text{if } p\in[1,\infty),\\ \left\| e^{\sum_{l=1}^{n} c(Y_l^h)h} |\mathcal{R}_n^h(\cdot,Y_n^h)|_\infty \right\|_1 &\le Ch\mathcal{E}(h,\Delta x), \qquad \text{if } p=\infty.\end{aligned} \tag{4.4.39}$$

Assumption $\mathcal{B}(p,c,\mathcal{E})$ is inspired by the Lax-Richtmeyer's convergence theorem [75]. In fact, recall that at each time step n, the hybrid scheme isolates the component y and applies the discrete operator $\Pi_{\Delta x}^h(y)$ for solving (one step in time) the PIDE

$$\partial_t v(t,x)+\mathcal{L}^{(y)}v(t,x)=0, \qquad (t,x)\in[nh,(n+1)h)\times\mathbb{R}^m.$$

Here, y is just a parameter (the current position of the Markov chain), so the coefficients of $\mathcal{L}^{(y)}$ (see (4.4.31)) are indeed constant. That's why the Lax-Richtmeyer technique can be adapted, as it follows in the next result.

Theorem 4.4.1. *Assume that* $\Pi_{\Delta x}^h(y)$, $y\in\mathcal{D}$, *satisfies Assumption* $\mathcal{B}(p,c,\mathcal{E})$. *Let* $\tilde u_n^h$ *be the function defined in*(4.4.35) *and* u_n^h *be the approximation through the scheme* (4.4.36). *Then, there exist* $\bar h\in(0,1)$ *and* $C>0$ *such that for every* $h<\bar h$ *and* $\Delta x<1$ *one has*

$$|\tilde u_0^h(\cdot,Y_0)-u_0^h(\cdot,Y_0)|_p\le CT\mathcal{E}(h,\Delta x). \tag{4.4.40}$$

Proof. Set $\mathrm{err}_n^h(\cdot,Y_n^h)=\tilde u_n^h(\cdot,Y_n^h)-u_n^h(\cdot,Y_n^h)$. By using the relation $|\max\{(a,b)\}-\max\{(a',b')\}|\le \max\{|a-a'|,|b-b'|\}$ we get

$$\begin{aligned}|\mathrm{err}_n^h(x,Y_n^h)| &\le \left| \mathbb{E}\left[\tilde u_{n+1}^h(X_{n+1}^{nh,x,y},Y_{(n+1)h}^{nh,y})\right]\Big|_{y=Y_n^h} - \mathbb{E}\left[\Pi_{\Delta x}^h(Y_n^h)u_{n+1}^h(\cdot,Y_{n+1}^h)(x)\big|Y_n^h\right]\right|\\ &\le \left|\mathbb{E}[\Pi_{\Delta x}^h(Y_n^h)\mathrm{err}_{n+1}^h(\cdot,Y_{n+1}^h)(x)|Y_n^h]\right| + |\mathcal{R}_n^h(x,Y_n^h)|,\end{aligned}$$

in which we have used (4.4.38). Since $\mathrm{err}_n^h(x_i,Y_N^h)=0$, by iterating one gets

$$|\mathrm{err}_0^h(\cdot,Y_0)|\le\sum_{n=0}^{N-1}\mathbb{E}\left[\left|\left(\prod_{l=0}^{n-1}\Pi_{\Delta x}^h(Y_l^h)\right)\mathcal{R}_n^h(\cdot,Y_n^h)\right|\right],$$

in which we use the convention $\prod_{l=0}^{-1}(\cdot) = \mathrm{Id}$. We use now (4.4.39). For $p \neq \infty$,

$$
\begin{aligned}
|\mathrm{err}_h^0(\cdot, Y_0)|_p &\leq \sum_{n=0}^{N-1} \Big|\mathbb{E}\Big[\Big(\prod_{l=0}^{n-1} \Pi_{\Delta x}^h(Y_l^h)\Big)\mathcal{R}_n^h(\cdot, Y_n^h)\Big]\Big|_p \leq \sum_{n=0}^{N-1} \mathbb{E}\Big[\Big|\Big(\prod_{l=0}^{n-1} \Pi_{\Delta x}^h(Y_l^h)\Big)\mathcal{R}_n^h(\cdot, Y_n^h)\Big|_p^p\Big]^{1/p} \\
&\leq \sum_{n=0}^{N-1} \Big(\mathbb{E}\big[e^{\sum_{l=1}^n pc(Y_l^h)h}|\mathcal{R}_n^h(\cdot, Y_n^h)|_p^p\big]\Big)^{\frac{1}{p}} \leq \sum_{n=0}^{N-1} hC\mathcal{E}(h, \Delta x) \leq TC\mathcal{E}(h, \Delta x).
\end{aligned}
$$

The case $p = \infty$ follows the same lines. □

Remark 4.4.2. *In Assumption $\mathcal{B}(p, c, \mathcal{E})$ we have required that the constant C and the function $\mathcal{E}$ in* (4.4.39) *do not depend on h and n. A closer look at the proof of Theorem 4.4.1 shows that this assumption can be relaxed. In fact, we can replace C and $\mathcal{E}$ in* (4.4.39) *by $C_{h,n}$ and $\mathcal{E}_{h,n}$ which depend on h and n but such that $\lim_{(h,\Delta x)\to(0,0)} \sum_{n=0}^{N-1} hC_{h,n}\mathcal{E}_{h,n}(h, \Delta x) = 0$. However, in this case we do not get information about the rate of convergence of the method.*

4.4.3 An example: finite difference schemes

We specify here some settings ensuring that the assumptions of Theorem 4.4.1 are satisfied. In particular, we choose the operator $\Pi_{\Delta x}^h(y)$ in (4.4.32) by means of two different finite difference schemes: the first one is a generalization of the procedure described in Chapter 3 and allows us to study the convergence in the l_2-norm, while the second one works l_∞. For the sake of readability, we consider the case $m = d = \ell = \ell_1 = \ell_2 = 1$.

As regards the Markov chain $(Y_n^h)_{n=0,\dots,N}$, in addition to Assumption $\mathcal{A}_1$ and $\mathcal{A}_2$ (see Section 4.3), we will need also the following:

Assumption $\mathcal{A}_3(g)$ *Let $g = g(y) \geq 0$, $y \in \mathcal{D}$. $(Y_n^h)_{n=0,\dots,N}$ satisfies Assumption $\mathcal{A}_3(g)$ if*

$$\mathbb{E}\left[e^{\sum_{l=1}^N g(Y_l^h)}\right] < \infty.$$

Moreover, we assume hereafter that the Lévy measure ν satisfies the following property: there exists $c_\nu > 0$ such that for every $\Delta x < 1$ one has

$$\sum_{l\in\mathbb{Z}} \nu(l\Delta x)\Delta x \leq \lambda c_\nu, \tag{4.4.41}$$

where λ is the intensity of the Poisson process K in the definition of the coumpound Poisson process H in (3.2.2).

Convergence in l_2-norm

We study here a hybrid procedure which generalizes the one introduced in [27] and described in Chapter 3 for the Bates model. For $y \in \mathcal{D}$, $\Pi^h_{\Delta x}(y)$ gives the numerical solution on $\mathcal{X} = \{x_i = X_0 + i\Delta x\}_{i\in\mathbb{Z}}$ a time nh to the PIDE (4.4.30), the operator $\mathcal{L}^{(y)}$ therein being given in (4.4.31). It is clear that the solution v of (4.4.30) depends on y and ζ as well, but these are just parameters (and not variables of the PIDE), so for simplicity we drop here such dependence. We split the operator $\mathcal{L}^{(y)} = \mathcal{L}^{(y)}_{\mathrm{diff}} + \mathcal{L}^{(y)}_{\mathrm{int}}$ in its differential and integral part:

$$\mathcal{L}^{(y)}_{\mathrm{diff}} v(x) = \mu_X(y)\partial_x v(x) + \frac{1}{2}\sigma^2_X(y)\partial^2_x v(x), \tag{4.4.42}$$

$$\mathcal{L}^{(y)}_{\mathrm{int}} v(x) = \gamma_X(y) \int \big(v(x+z) - v(x)\big)\nu(z)dz. \tag{4.4.43}$$

We now apply the trapezoidal rule in order to approximate the integral term $\mathcal{L}^{(y)}_{\mathrm{int}} v$ and we use the central finite difference scheme to solve $\mathcal{L}^{(y)}_{\mathrm{diff}} v$. Applying an implicit-explicit method in time, we obtain an approximating solution $v^n = (v^n_j)_{j\in\mathbb{Z}}$ to the PIDE (4.4.30) given by the solution of the linear equation

$$A^h_{\Delta x}(y) v^n = B^h_{\Delta x}(y) v^{n+1} \tag{4.4.44}$$

(recall that v^{n+1} is known). Here $A^h_{\Delta x}(y)$ is the linear operator given by

$$(A^h_{\Delta x})_{ij}(y) = \begin{cases} \alpha^h_{\Delta x}(y) - \beta^h_{\Delta x}(y), & \text{if } i = j+1, \\ 1 + 2\beta^h_{\Delta x}(y), & \text{if } i = j, \\ -\alpha^h_{\Delta x}(y) - \beta^h_{\Delta x}(y), & \text{if } i = j-1, \\ 0, & \text{if } |i-j| > 1, \end{cases} \tag{4.4.45}$$

with

$$\alpha^h_{\Delta x}(y) = \frac{h}{2\Delta x}\mu_X(y), \qquad \beta^h_{\Delta x}(y) = \frac{h}{2\Delta x^2}\sigma^2_X(y), \tag{4.4.46}$$

and $B^h_{\Delta x}(y)$ is the linear operator defined as

$$(B^h_{\Delta x})_{ij}(y) = \begin{cases} \gamma_X(y) h \Delta x \nu((j-i)\Delta x), & \text{if } j \neq i, \\ 1 + h\Delta x \gamma_X(y)\Big(\nu(0) - \sum_{l\in\mathbb{Z}^*} \nu(l\Delta x)\Big) & \text{if } i = j. \end{cases} \tag{4.4.47}$$

Then we have

Lemma 4.4.3. *For every $y \in \mathcal{D}$, the operator $A^h_{\Delta x}(y) : l_2(\mathcal{X}) \to l_2(\mathcal{X})$ is invertible and $|(A^h_{\Delta x})^{-1}(y)|_2 \leq 1$. Moreover $|B^h_{\Delta x}(y)|_2 \leq 1 + 2\lambda c_\nu |\gamma_X(y)| h$, where c_ν is defined in* (4.4.41).

Proof. Fix $y \in \mathcal{D}$ and $w \in l_2(\mathcal{X})$. Then $A^h_{\Delta x}(y)v = w$, for some $v \in l_2(\mathcal{X})$, if and only if

$$(\alpha^h_{\Delta x}(y) - \beta^h_{\Delta x}(y))v_{j-1} + (1 + 2\beta^h_{\Delta x}(y))v_j - (\alpha^h_{\Delta x}(y) + \beta^h_{\Delta x}(y))v_{j+1} = w_j, \qquad j \in \mathbb{Z}, \quad (4.4.48)$$

$\alpha^h_{\Delta x}$ and $\beta^h_{\Delta x}$ being given in (4.4.46). Let $\hat{\varphi}$ denote the Fourier transform of $\varphi \in l_2(\mathcal{X})$, that is, $\hat{\varphi}(\theta) = \frac{\Delta x}{\sqrt{2\pi}} \sum_{j \in \mathbb{Z}} \varphi_j e^{-\mathbf{i} j \Delta x \theta}$, $\theta \in [0, 2\pi)$, $\mathbf{i}$ denoting the imaginary unit. We define the function $\psi(\theta)$, $\theta \in [0, 2\pi)$, by

$$\Big((\alpha^h_{\Delta x}(y) - \beta^h_{\Delta x}(y))e^{-i\theta\Delta x} + 1 + 2\beta^h_{\Delta x}(y) - (\alpha^h_{\Delta x}(y) + \beta^h_{\Delta x}(y))e^{i\theta\Delta x}\Big)\psi(\theta) = \hat{w}(\theta). \qquad (4.4.49)$$

Note that

$$\begin{aligned}
&|(\alpha^h_{\Delta x}(y) - \beta^h_{\Delta x}(y))e^{-\mathbf{i}\,\theta\Delta x} + 1 + 2\beta^h_{\Delta x}(y) - (\alpha^h_{\Delta x}(y) + \beta^h_{\Delta x}(y))e^{\mathbf{i}\,\theta\Delta x}| \\
&\qquad \geq \big|\mathfrak{Re}\big[(\alpha^h_{\Delta x}(y) - \beta^h_{\Delta x}(y))e^{-\mathbf{i}\,\theta\Delta x} + 1 + 2\beta^h_{\Delta x}(y) - (\alpha^h_{\Delta x}(y) + \beta^h_{\Delta x}(y))e^{\mathbf{i}\,\theta\Delta x}\big]\big| \\
&\qquad = 1 + 2\beta^h_{\Delta x}(y)(1 - \cos(\theta\Delta x)) \geq 1,
\end{aligned}$$

for every $\theta \in [0, 2\pi)$. So, $\psi \in L^2([0, 2\pi), dx)$ and we can define $v.$ as its inverse Fourier transform:

$$v_j = \frac{1}{\Delta x\sqrt{2\pi}} \int_0^{2\pi} \psi(\theta) e^{\mathbf{i} j\theta\Delta x} d\theta, \qquad j \in \mathbb{Z}.$$

Straightforward computations give that v is the unique solution to (4.4.48), hence $A^h_{\Delta x}$ is invertible. Moreover, from (4.4.49) we obtain $|\psi(\theta)| \leq |\hat{w}(\theta)|$, so that $|\psi(\theta)|_{L^2([0,2\pi),dx)} \leq |\hat{w}(\theta)|_{L^2([0,2\pi),dx)}$. We use now the Parseval identity $|\hat{\varphi}|_{L^2([0,2\pi),dx)} = |\varphi|_2$ and we get $|(A^h_{\Delta x})^{-1}(y)w|_2 \leq |w|_2$, which gives $|(A^h_{\Delta x})^{-1}(y)|_2 \leq 1$. Finally, for $w \in l_2(\mathcal{X})$ we have

$$(B^h_{\Delta x}(y)w)_j = w_j + h\Delta x \gamma_X(y)\Big(\sum_l \nu(l\Delta x)w_{j+l} - \sum_l \nu(l\Delta x)w_j\Big),$$

so that

$$\widehat{B^h_{\Delta x}(y)}w(\theta) = \Big(1 + h\Delta x\gamma_X(y)\sum_l \nu(l\Delta x)(e^{\mathbf{i}l\theta} - 1)\Big)\hat{w}(\theta).$$

Then,

$$|\widehat{B^h_{\Delta x}(y)}w|_{L^2([0,2\pi),dx)} \leq (1 + 2\lambda c_\nu|\gamma_X(y)|h)|\hat{w}|_{L^2([0,2\pi),dx)},$$

because $|e^{\mathbf{i}l\theta} - 1| \leq 2$ and $\sum_l \nu(l\Delta x)\Delta x \leq \lambda c_\nu$. By the Parseval relation, $|B^h_{\Delta x}(y)w|_2 \leq (1 + 2\lambda c_\nu|\gamma_X(y)|h)|w|_2$, which concludes the proof. □

In the following we will use functions $v \in C^{p,q}_{\mathbf{pol},[nh,(n+1)h]}(\mathbb{R}, \mathcal{D})$ a.e. uniformly in n and h. This means that $v \in C^{\lfloor q/2 \rfloor, q}([a, b), \mathbb{R} \times \mathcal{D})$ a.e. and there exist $C, c > 0$ independent of n and h such that

$$\sup_{t \in [nh,(n+1)h)} |\partial^k_t \partial^{l'}_x \partial^l_y v(t, \cdot, y)|_{L^p(\mathbb{R}^m,dx)} \leq C(1 + |y|^c), \quad 2k + |l'| + |l| \leq q.$$

We can now state the convergence result.

Theorem 4.4.4. *Let $\tilde{u}_n^h$ be defined in* (4.4.35) *and u_n^h be given by* (4.4.36) *with the choice*

$$\Pi^h_{\Delta x}(y) = (A^h_{\Delta x})^{-1} B^h_{\Delta x}(y),$$

$A^h_{\Delta x}(y)$ and $B^h_{\Delta x}(y)$ being given in (4.4.45) *and* (4.4.47) *respectively. Moreover, for $n = 0, \dots, N$, consider the function*

$$v_n^h(t,x,y) = \mathbb{E}\left[\tilde{u}^h_{n+1}(X^{t,x,y}_{(n+1)h}, Y^{t,y}_{(n+1)h})\right], \qquad t \in [nh,(n+1)h]. \tag{4.4.50}$$

Assume that

- $\frac{\nu'}{\nu}, \frac{\nu''}{\nu} \in L^2(\mathbb{R}, d\nu)$*;*
- *the Markov chain $(Y^h_n)_{n=0,\dots,N}$ satisfies assumptions $\mathcal{A}_1$, $\mathcal{A}_2$ and $\mathcal{A}_3(4\lambda c_\nu |\gamma_X|)$;*
- $v_n^h \in C^{2,6}_{\mathbf{pol},[nh,(n+1)h]}(\mathbb{R},\mathcal{D})$ *a.e. and uniformly in n and h.*

Then, there exist $\bar{h}, C > 0$ such that for every $h < \bar{h}$ and $\Delta x < 1$ one has

$$|\tilde{u}^h_0(\cdot, Y_0) - u^h_0(\cdot, Y_0)|_2 \le CT(h + \Delta x^2). \tag{4.4.51}$$

We stress that, from (4.4.51), the rate of convergence is of the second order in space, because of the choice of a second order finite difference scheme, and of first order in time, as it is natural also for the presence of the approximating Markov chain Y^h (see Theorem 4.3.1).

Theorem 4.4.4 is a direct consequence of Theorem 4.4.1 once we prove that Assumption $\mathcal{B}(p,c,\mathcal{E})$ holds with $p = 2$, $c(y) = 2\lambda c_\nu |\gamma_X|(y)$ and $\mathcal{E}(h, \Delta x) = h + \Delta x^2$. To this purpose, we first need two technical lemmas which allow us to handle the error coming from suitable Taylor's expansions and from the quadrature approximation. We postpone the proofs to Appendix 4.7.2.

Lemma 4.4.5. (i) *Let $g \in C^2(\mathbb{R})$ be such that $g, g', g'' \in L^1(\mathbb{R}, dx)$. Then*

$$\Big|\sum_{l\in\mathbb{Z}} g(x_l)\Delta x - \int_{\mathbb{R}} g(x)dx\Big| \le \frac{\Delta x^2}{12}\, |g''|_{L^1(\mathbb{R},dx)}. \tag{4.4.52}$$

(ii) *Let $g \in C^2(\mathbb{R})$ be such that $g, g', g'' \in L^2(\mathbb{R}, dx)$. Then*

$$\sum_{l\in\mathbb{Z}} g^2(x_l)\Delta x \le |g|^2_{L^2(\mathbb{R},dx)} + \frac{\Delta x^2}{6}\big(|g'|^2_{L^2(\mathbb{R},dx)} + |g|_{L^2(\mathbb{R},dx)} \times |g''|_{L^2(\mathbb{R},dx)}\big). \tag{4.4.53}$$

Remark 4.4.6. *In our convergence result Theorem 4.4.4 or also in the following Theorem 4.4.10, we require that $\frac{\nu'}{\nu}, \frac{\nu''}{\nu} \in L^1(\mathbb{R}, d\nu)$ (recall that ν is a finite positive measure), and this implies that $\nu, \nu', \nu'' \in L^1(\mathbb{R}, dx)$. By using* (4.4.52), (4.4.41) *holds with $\lambda c_\nu = \lambda + |\nu''|_{L^1(\mathbb{R},dx)}$.*

Lemma 4.4.7. *Let $g : [0,T] \times \mathbb{R} \times \mathcal{D} \to \mathbb{R}$ be such that*

$$\exists\, a, A > 0 : \quad \sup_{t\in[0,T)} |\partial_x^k g(t,\cdot,y)|_{L^2(\mathbb{R},dx)} \le A(1+|y|^a), \quad k=0,1,2 \tag{4.4.54}$$

and suppose that

$$\frac{\nu'}{\nu},\ \frac{\nu''}{\nu} \in L^2(\mathbb{R}, d\nu). \tag{4.4.55}$$

For fixed $h < T$, $\Delta x > 0$ and $\gamma \ge 0$, consider the functions defined by

$$\Psi_1(t,x,y) = \sum_l \nu(l\Delta x)\big[g(t,x+l\Delta x,y) - g(t,x,y)\big]\Delta x, \quad (t,x,y) \in [0,T]\times\mathbb{R}\times\mathcal{D},$$
$$\Psi_2(t,x,y) = \int_0^1 (1-\tau)^\gamma g(t+\tau h, ,y)d\tau, \quad (t,x,y)\in[0,T-h]\times\mathbb{R}\times\mathcal{D},$$
$$\Psi_3(t,x,y) = \int_0^1 (1-\eta)^\gamma g(t,x+\eta\Delta x,y)d\eta, \quad (t,x,y)\in[0,T]\times\mathbb{R}\times\mathcal{D},$$
$$\Psi_4(t,x,y,z) = \int_0^1 (1-\zeta)^\gamma g(t,x,y+\zeta(z-y))d\zeta, \quad (t,x,y,z)\in[0,T]\times\mathbb{R}\times\mathcal{D}\times\mathcal{D}.$$

Then there exists $C > 0$ such that

$$\sup_{t\in[0,T]} |\Psi_n(t,\cdot,y)|_2 \le C(1+|y|^a), \quad n=1,2,3, \tag{4.4.56}$$
$$\sup_{t\in[0,T]} |\Psi_4(t,\cdot,y,z)|_2 \le C(1+|y|^a+|z|^a). \tag{4.4.57}$$

Moreover, set

$$\Psi_5(t,x,y) = \int g(t,x+\xi,y)\nu(\xi)d\xi - \sum_l g(t,x+l\Delta x,y)\nu(l\Delta x)\Delta x, \quad (t,x,y)\in[0,T]\times\mathbb{R}\times\mathcal{D}.$$

If (4.4.54) *holds also with $k = 3, 4$, there exists $C > 0$ such that*

$$\sup_{t\in[0,T]} |\Psi_5(t,\cdot,y)|_2 \le \lambda C(1+|y|^a)\,\Delta x^2. \tag{4.4.58}$$

We can now prove the following key result.

Proposition 4.4.8. *Set $\Pi^h_{\Delta x}(y) = (A^h_{\Delta x})^{-1}B^h_{\Delta x}(y)$, with $A^h_{\Delta x}(y)$ and $B^h_{\Delta x}(y)$ given in* (4.4.45) *and* (4.4.47)*. For all $n = 0,\dots,N-1$, let v^h_n be the function defined in* (4.4.50)*. Suppose that*

- $\frac{\nu'}{\nu}, \frac{\nu''}{\nu} \in L^2(\mathbb{R}, d\nu)$;
- $(Y^h_n)_{n=0,\dots,N}$ *satisfies Assumptions $\mathcal{A}_1$, $\mathcal{A}_2$ and $\mathcal{A}_3(4\lambda c_\nu|\gamma_X|)$;*

- $v_n^h \in C^{2,6}_{\mathbf{pol},[nh,(n+1)h]}(\mathbb{R},\mathcal{D})$ *a.e. and uniformly in* n *and* h.

Then $\Pi^h_{\Delta x}(y)$ *satisfies Assumption* $\mathcal{B}(2, 2\lambda c_\nu|\gamma_X|, h+\Delta x^2)$.

Proof. Lemma 4.4.3 gives $|\Pi^h_{\Delta x}(y)|_2| \le |(A^h_{\Delta x})^{-1}(y)|_2|B^h_{\Delta x}(y)|_2 \le 1 + 2\lambda c_\nu|\gamma_X(y)|h$, so (4.4.37) holds with $c(y) = 2\lambda c_\nu|\gamma_X(y)|$. We prove now (4.4.39) with $p = 2$ and $\mathcal{E}(h,\Delta x) = h + \Delta x^2$. We first recall that $v_n^h(t,x,y) = \mathbb{E}\left[\tilde{u}^h_{n+1}(X^{t,x,y}_{(n+1)h}, Y^{t,y}_{(n+1)h})\right]$ for $t \in [nh,(n+1)h]$ so that (4.4.38) equals to

$$\mathbb{E}\Big[\Pi^h_{\Delta x}(y)v_h^n((n+1)h,\cdot,Y^h_{n+1})(x)\,\Big|\,Y^h_n = y\Big] = v_n^h(nh,x,y) + \mathcal{R}^h_n(x,y), \tag{4.4.59}$$

which can be rewritten as

$$\mathbb{E}\big[B^h_{\Delta x}(Y^h_n)v^h_n((n+1)h,\cdot,Y^h_{n+1})(x) \mid Y^h_n\big] = A^h_{\Delta x}(Y^h_n)v^h_n(nh,\cdot,Y^h_n)(x) + A^h_{\Delta x}(Y^h_n)\mathcal{R}^h_n(\cdot,Y^h_n)(x). \tag{4.4.60}$$

Step 1. Taylor expansion of the l.h.s. of (4.4.60)**.** We set

$$\begin{aligned} I_1 = B^h_{\Delta x}(Y^h_n)v^h_n((n+1)h,\cdot,Y^h_{n+1})(x_i) = v^h_n((n+1)h,x_i,Y^h_{n+1}) \\ + h\gamma_X(Y^h_n)\sum_l \nu(x_l)\Big(v^h_n((n+1)h,x_{i+l},Y^h_{n+1}) - v^h_n((n+1)h,x_i,Y^h_{n+1})\Big)\Delta x. \end{aligned} \tag{4.4.61}$$

As regard the first term in the r.h.s. above, we first apply Taylor's expansion to $t \mapsto v^h_n(t,x_i,Y^h_{n+1})$ around nh up to order 1 and, then, we consider the Taylor expansion of $y \mapsto v^h_n(nh,x_i,y)$ around Y^h_n up to order 3 and of $y \mapsto \partial_t v^h_n(nh,x_i,y)$ around Y^h_n up to order 1. Rearranging the terms we obtain

$$\begin{aligned} &v^h_n((n+1)h,x_i,Y^h_{n+1}) = v^h_n(nh,x_i,Y^h_n) \\ &+ \partial_t v^h_n(nh,x_i,Y^h_n)h + \partial_y v^h_n(nh,x_iY^h_n)(Y^h_{n+1}-Y^h_n) + \frac{1}{2}\partial^2_y v^h_n(nh,x_i,Y^h_n)(Y^h_{n+1}-Y^h_n)^2 \\ &+ \partial_y\partial_t v^h_n(nh,x_i,Y^h_n)\,h(Y^h_{n+1}-Y^h_n) + \frac{1}{6}\partial^3_y v^h_n(nh,x_i,Y^h_n)(Y^h_{n+1}-Y^h_n)^3 + R_1(n,h,x_i,Y^h_n,Y^h_{n+1}), \end{aligned}$$

where R_1 is given by

$$\begin{aligned} R_1(n,h,x_i,Y^h_n,Y^h_{n+1}) &= h^2\int_0^1 (1-\tau)\partial^2_t v^h_n(nh+\tau h,x_i,Y^h_{n+1})d\tau \\ &+\frac{(Y^h_{n+1}-Y^h_n)^4}{6}\int_0^1 (1-\zeta)^3\partial^4_y v^h_n(nh,x_i,Y^h_n+\zeta(Y^h_{n+1}-Y^h_n))d\zeta \\ &+h(Y^h_{n+1}-Y^h_n)^2\int_0^1 (1-\zeta)\partial_t\partial^2_y v^h_n(nh,x_i,Y^h_n+\zeta(Y^h_{n+1}-Y^h_n))d\zeta. \end{aligned} \tag{4.4.62}$$

For the second term in the right hand side of (4.4.61), we stop the Taylor expansion of $t \mapsto v_n^h((n+1)h, x_{i+l}, Y_{n+1}^h)$ around nh at order 0 and of $y \mapsto v_n^h(nh, x_{i+l}, y)$ around Y_h^n at order 1, obtaining

$$
\begin{aligned}
& h\gamma_X(Y_n^h)\Big(\sum_l \nu(x_l)v_n^h((n+1)h, x_{i+l}, Y_{n+1}^h) - \sum_l \nu(x_l)v_n^h((n+1)h, x_i, Y_{n+1}^h)\Big)\Delta x \\
& = h\gamma_X(Y_n^h)\sum_l \nu(x_l)\big[v_n^h(nh, x_{i+l}, Y_n^h) - v_n^h(nh, x_i, Y_n^h)\big]\Delta x \\
& + h\gamma_X(Y_n^h)(Y_{n+1}^h - Y_n^h)\sum_l \nu(x_l)\big[\partial_y v_n^h(nh, x_{i+l}, Y_n^h) - \partial_y v_n^h(nh, x_i, Y_n^h)\big]\Delta x \\
& + R_2(n, h, x_i, Y_n^h, Y_{n+1}^h),
\end{aligned}
$$

where the remaining term R_2 contains the integral terms:

$$
\begin{aligned}
& R_2(n, h, x_i, Y_n^h, Y_{n+1}^h) = \\
& h^2\gamma_X(Y_n^h)\sum_l \nu(x_l)\Delta x \int_0^1 (1-\tau)\big[\partial_t v_n^h(nh+\tau h, x_{i+l}, Y_{n+1}^h) - \partial_t v_n^h(nh+\tau h, x_i, Y_{n+1}^h)\big]d\tau \\
& + h\gamma_X(Y_n^h)(Y_{n+1}^h - Y_n^h)^2 \sum_l \nu(x_l)\Delta x \times \\
& \times \int_0^1 (1-\zeta)\big[\partial_y v_n^h(nh, x_{i+l}, Y_n^h + \zeta(Y_{n+1}^h - Y_n^h)) - \partial_y v_n^h(nh, x_i, Y_n^h + \zeta(Y_{n+1}^h - Y_n^h))\big]d\zeta.
\end{aligned}
\tag{4.4.63}
$$

By resuming, we obtain

$$
\begin{aligned}
I_1 = \; & v_n^h(nh, x_i, Y_n^h) + \partial_t v_n^h(nh, x_i, Y_n^h)h + \partial_y v_n^h(nh, x_i, Y_n^h)(Y_{n+1}^h - Y_n^h) \\
& + \frac{1}{2}\partial_y^2 v_n^h(nh, x_i, Y_n^h)(Y_{n+1}^h - Y_n^h)^2 + \partial_y\partial_t v_n^h(nh, x_i, Y_n^h)\, h(Y_{n+1}^h - Y_n^h) \\
& + \frac{1}{6}\partial_y^3 v_n^h(nh, x_i, Y_n^h)(Y_{n+1}^h - Y_n^h)^3 + h\Delta x\gamma_X(Y_n^h)\sum_l \nu(x_l)\big[v_n^h(nh, x_{i+l}, Y_n^h) - v_n^h(nh, x_i, Y_n^h)\big] \\
& + \sum_{i=1}^3 R_i(n, h, x_i, Y_n^h, Y_{n+1}^h),
\end{aligned}
\tag{4.4.64}
$$

where

$$
R_3(n, h, x_i, Y_n^h, Y_{n+1}^h) = h(Y_{n+1}^h - Y_n^h)\gamma_X(Y_n^h)\sum_l \nu(x_l)\big[\partial_y v_n^h(nh, x_{i+l}, Y_n^h) - \partial_y v_n^h(nh, x_i, Y_n^h)\Big]\Delta x.
\tag{4.4.65}
$$

Step 2. Taylor expansion of the first addendum in the r.h.s. of (4.4.60). We set

$$I_2 = A^h_{\Delta x} v^h_n(nh,\cdot,Y^h_n)(x_i) = (\alpha^h_{\Delta x}(Y^h_n) - \beta^h_{\Delta x}(Y^h_n))v^h_n(nh,x_{i-1},Y^h_n) \\ + (1+2\beta^h_{\Delta x}(Y^h_n))v^h_n(nh,x_i,Y^h_n) - (\alpha^h_{\Delta x}(Y^h_n) + \beta^h_{\Delta x}(Y^h_n))v^h_n(nh,x_{i+1},Y^h_n).$$

We expand with Taylor $x \mapsto v^h_n(nh,x,Y^h_n)$ around x_i up to order 3 and we insert the values of $\alpha^h_{\Delta x}$ and $\beta^h_{\Delta x}$ in (4.4.46). Rearranging the terms we get

$$\begin{aligned} I_2 =& v^h_n(nh,x_i,Y^h_n) - h\mu_X(Y^h_n)\partial_x v^h_n(nh,x_i,Y^h_n) - \frac{1}{2}\, h\sigma^2_X(Y^h_n)\partial^2_x v^h_n(nh,x_i,Y^h_n) \\ &+ R_4(n,h,x_i,Y^h_n,Y^h_{n+1}) \end{aligned} \tag{4.4.66}$$

where

$$\begin{aligned} R_4(n,h,x_i,Y^h_n,Y^h_{n+1}) =& \frac{\Delta x\mu_X(Y^h_n) - \sigma^2_X(Y^h_n)}{12}\, h\Delta x^2 \int_0^1 (1-\eta)^3 \partial^4_x v^h_n(nh,x_i-\eta\Delta x,Y^h_n)d\eta \\ &- \frac{\Delta x\mu_X(Y^h_n) + \sigma^2_X(Y^h_n)}{12}\, h\Delta x^2 \int_0^1 (1-\eta)^3 \partial^4_x v^h_n(nh,x_i+\eta\Delta x,Y^h_n)d\eta \\ &- \frac{1}{6}\, h\Delta x^2 \mu_X(Y^h_n)\partial^3_x v^h_n(nh,x_i,Y^h_n). \end{aligned} \tag{4.4.67}$$

Step 3. Rearranging the terms. By resuming, from (4.4.64) and (4.4.66) we have

$$\begin{aligned} I_1 - I_2 =& h\partial_t v^h_n(nh,x_i,Y^h_n) + (Y^h_{n+1} - Y^h_n)\partial_y v^h_n(nh,x_i,Y^h_n) + h\mu_X(Y^h_n)\partial_x v^h_n(nh,x_i,Y^h_n) \\ &+ \frac{1}{2}\big[(Y^h_{n+1} - Y^h_n)^2\partial^2_y v^h_n(nh,x_i,Y^h_n) + h\,\sigma^2_X(Y^h_n)\partial^2_x v^h_n(nh,x_i,Y^h_n)\big] \\ &+ h\gamma_X(Y^h_n)\int (v^h_n(t,x+\zeta,Y^h_n) - v^h_n(t,x,Y^h_n))\nu(d\zeta) + \partial_y\partial_t v^h_n(nh,x_i,Y^h_n)\, h(Y^h_{n+1} - Y^h_n) \\ &+ \frac{1}{6}\partial^3_y v^h_n(nh,x_i,Y^h_n)(Y^h_{n+1} - Y^h_n)^3 + \sum_{i=1}^5 R_i(n,h,x_i,Y^h_n,Y^h_{n+1}) \end{aligned}$$

where

$$\begin{aligned} R_5(n,h,x_i,Y^h_n) =& h\gamma_X(Y^h_n)\sum_l \big[v^h_n(t,x_{i+l},Y^h_n) - v^h_n(t,x_i,Y^h_n)\big]\nu(l\Delta x)\Delta x \\ &- h\gamma_X(Y^h_n)\int \big[v^h_n(t,x_i+z,Y^h_n) - v^h_n(t,x_i,Y^h_n)\big]\nu(dz). \end{aligned} \tag{4.4.68}$$

Now, note that, by the Feynman-Kac formula, the function $v^h_n(t,x,y) = \mathbb{E}\left[\tilde u^h_{n+1}(X^{t,x,y}_{(n+1)h}, Y^{t,y}_{(n+1)h})\right]$ solves the PIDE

$$\begin{cases} \partial_t v^h_n(t,x,y) + \mathcal{L}v^h_n(t,x,y) = 0, & (t,x,y) \in [nh,(n+1)h) \times \mathbb{R}^m \times \mathcal{D}, \\ v^h_n((n+1)h,x,y) = \tilde u^h_{n+1}(x,y), & \text{in } \mathbb{R}^m \times \mathcal{D}. \end{cases}$$

Then, by passing to the conditional expectation and by using formulas (4.3.5), (4.3.6) and (4.3.7) for the local moments of order 1, 2 and 3, we obtain

$$\begin{aligned}\widetilde{\mathcal{R}}^h_n(x_i,Y^h_n) :=& \mathbb{E}[I_1 - I_2 \mid Y^h_n] = h(\partial_t[v^h_n(nh,x_i,Y^h_n) + \mathcal{L}[v^h_n(nh,x_i,Y^h_n))\\ &+ \mathbb{E}\Big[\partial_y\partial_t[v^h_n(nh,x_i,Y^h_n)\,h(Y^h_{n+1}-Y^h_n) + \frac{1}{6}\partial^3_y[v^h_n(nh,x_i,Y^h_n)(Y^h_{n+1}-Y^h_n)^3 \mid Y^h_n\Big]\\ &+ \sum_{i=1}^{5}\mathbb{E}[R_i(n,h,x_i,Y^h_n,Y^h_{n+1}) \mid Y^h_n]\\ &= \sum_{i=1}^{6}\mathbb{E}[R_i(n,h,x_i,Y^h_n,Y^h_{n+1}) \mid Y^h_n]\end{aligned}$$

where we have set

$$\begin{aligned}R_6(n,h,x_i,Y^h_n,Y^h_{n+1}) = f_h(Y^h_n)\partial_y v^h_n(nh,x_i,Y^h_n) + \frac{1}{2}g_h(Y^h_n)\partial^2_y v^h_n(nh,x_i,Y^h_n)\\ + \frac{1}{6}j_h(Y^h_n)\partial^3_y v^h_n(nh,x_i,Y^h_n),\end{aligned} \tag{4.4.69}$$

f_h, g_h and j_h being defined in (4.3.5),(4.3.6) and (4.3.7).

Step 4. Estimate of the remainder. Hereafter, C denotes a positive constant which may vary from a line to another and is independent of $n, h, \Delta x$.

By (4.4.60), the remaining we have to study is $\mathcal{R}^h_n(\cdot,Y^h_n) = (A^h_{\Delta x})^{-1}(Y^h_n)\widetilde{\mathcal{R}}^h_n(\cdot,Y^h_n)$. By Lemma 4.4.3, $|(A^h_{\Delta x})^{-1}(y)|_2 \leq 1$, so $|\mathcal{R}^h_n(\cdot,Y^h_n)|_2 \leq |\widetilde{\mathcal{R}}^h_n(\cdot,Y^h_n)|_2$. Now, by applying the Cauchy-Schwarz inequality and by using Assumption $\mathcal{A}_3(4\lambda c|\gamma_X|)$,

$$\begin{aligned}\mathbb{E}\big[e^{\sum_{l=1}^n 2\lambda c\gamma_X(Y^h_l)h}|\mathcal{R}^h_n(\cdot,Y^h_n)|^2_2\big] &\leq \mathbb{E}\big[e^{\sum_{l=1}^n 4\lambda c\gamma_X(Y^h_l)h}\big]^{1/2}\mathbb{E}\big[|\mathcal{R}^h_n(\cdot,Y^h_n)|^4_2\big]^{1/2}\\ &\leq \mathbb{E}\big[|\widetilde{\mathcal{R}}^h_n(\cdot,Y^h_n)|^4_2\big]^{1/2} \leq C\sum_{i=1}^{6}\mathbb{E}\big[|R_i(n,h,\cdot,Y^h_n,Y^h_{n+1})|^4_2\big]^{1/2}.\end{aligned}$$

So, we study the above 6 terms: we prove in fact that each one is upper bounded by $C(h^2+h\Delta x^2)^2$. The inequalities studied in Lemma 4.4.7 now come on.

Consider first R_1 in (4.4.62). By applying (4.4.56) for Ψ_2 and Ψ_4, we get

$$\begin{aligned}&|R_1(n,h,\cdot,Y^h_n,Y^h_{n+1})|^4_2\\ &\quad\leq C\big[h^8(1+(Y^h_n)^a)^4 + |Y_{n+1}-Y_n|^{16}(1+|Y^h_n|^a+|Y^h_{n+1}|^a)^4 + h^4|Y_{n+1}-Y_n|^8(1+|Y^h_n|^a)^4\big].\end{aligned}$$

So, by using the increment estimates (4.3.11), the moment estimates (4.3.12) and the Cauchy-Schwartz inequality, we obtain

$$\mathbb{E}\big[|R_1(n,h,\cdot,Y^h_n,Y^h_{n+1})|^4_2\big]^{1/2} \leq Ch^4.$$

R_4 in (4.4.67) can be handled in a similar way: recalling that μ_X and σ_X have polynomial growth, we apply now (4.4.56) for Ψ_3 and we get

$$\mathbb{E}\big[|R_4(n,h,\cdot,Y_n^h,Y_{n+1}^h)|_2^4\big]^{1/2} \le Ch^2\Delta x^4.$$

The same approach can be used for R_6 in (4.4.69): we use first (4.4.53), then the Hölder inequality and (4.3.8), (4.3.9), (4.3.10). Thus, with simple calculations

$$\mathbb{E}\big[|R_6(n,h,\cdot,Y_n^h,Y_{n+1}^h)|_2^4\big]^{1/2} \le Ch^4.$$

In order to study R_2 in (4.4.63), let us first set

$$g(t,x,Y_{n+1}^h) = \int_0^1 (1-\tau)\partial_t v_n^h(t+\tau h,x,Y_{n+1}^h)d\tau.$$

Then, for $k=0,1,2$,

$$|\partial_y^k g(t,\cdot,Y_{n+1}^h)|_{L^2(\mathbb{R},dx)}^2 \le \int_0^1 (1-\tau)^2|\partial_y^k v_n^h(nh+\tau h,\cdot,Y_{n+1}^h)|_{L^2(\mathbb{R},dx)}^2 d\tau \le C(1+|Y_{n+1}^h|^a)^2,$$

so, by (4.4.56) for Ψ_1, we obtain

$$\begin{aligned}\Big|\gamma_X(Y_n^h)\sum_l \nu(l\Delta x)\big[g(nh,\cdot+l\Delta x,Y_{n+1}^h)-g(nh,\cdot,Y_{n+1}^h)\big]\Delta x\Big|_2 &\le C|\gamma_X(Y_n^h)|(1+|Y_{n+1}^h|^a)\\ &\le C(1+|Y_n^h|)(1+|Y_{n+1}^h|^a),\end{aligned}$$

the latter because γ_X has sublinear growth. And if we define

$$g(t,y,Y_n^h,Y_{n+1}^h) = \int_0^1 (1-\zeta)\partial_y u(t,y,Y_n^h+\zeta(Y_{n+1}^h-Y_n^h))d\zeta,$$

the same reasonings give

$$\begin{aligned}&\Big|\gamma_X(Y_n^h)\sum_l \nu(l\Delta x)\big[g(nh,\cdot+l\Delta x,Y_n^h,Y_{n+1}^h)-g(nh,\cdot,Y_n^h,Y_{n+1}^h)\big]\Delta x\Big|_2\\ &\qquad\le C|\gamma_X(Y_n^h)|(1+|Y_n^h|^a+|Y_{n+1}^h|^a) \le C(1+|Y_n^h|)(1+|Y_n^h|^a+|Y_{n+1}^h|^a).\end{aligned}$$

Therefore, by the Cauchy-Schwartz inequality, (4.3.12) and (4.3.11), we finally obtain

$$\mathbb{E}\big[|R_2(n,h,\cdot,Y_n^h,Y_{n+1}^h)|_2^4\big]^{1/2} \le Ch^4.$$

R_3 in (4.4.65) can be estimated analogously, so we get

$$\mathbb{E}\big[|R_3(n,h,\cdot,Y_n^h,Y_{n+1}^h)|_2^4\big]^{1/2} \le Ch^4.$$

Finally, for R_5 in (4.4.68), (4.4.58) gives that $|R_5(n,h,\cdot,Y_n^h)|_2 \le Ch(1+|Y_n^h|^a)\,\Delta x^2$ and by passing to the expectation, (4.3.11) gives

$$\mathbb{E}\big[|R_5(n,h,\cdot,Y_n^h)|_2^4\big]^{1/2} \le Ch^2\Delta x^4.$$

Putting all the above estimates together, the statement holds. □

Proof of Theorem 4.4.4. The proof is a straightforward application of Proposition 4.4.8 and Theorem 4.4.1. □

Convergence in l_∞-norm

We consider here a different finite difference scheme for equation (4.4.30): we still approximate (explicit in time) the integral term $\mathcal{L}^{(y)}_{\text{int}}v$ in (4.4.43) with a trapezoidal rule, but we use an upwind first order scheme to approximate (implicit in time) the differential part $\mathcal{L}^{(y)}_{\text{diff}}v$ in (4.4.42). As usually done in convection-diffusion problems, we distinguish the cases in which $\mu_X(y)$ is positive or negative in order to take into account the asymmetry given by the convection term and we use one sided difference in the appropriate direction. Specifically, if $\mu_X(y)\ge 0$, we approximate $\mathcal{L}^{(y)}_{\text{diff}}u$ by using the scheme

$$\frac{v_i^{n+1}-v_i^n}{h}+\mu_X(y)\frac{v_{i+1}^n-v_i^n}{\Delta x}+\frac{1}{2}\sigma_X^2(y)\frac{v_{i+1}^n-2v_i^n+v_{i-1}^n}{\Delta x^2},$$

while, if $\mu_X(y)\le 0$, we use the approximation

$$\frac{v_i^{n+1}-v_i^n}{h}+\mu_X(y)\frac{v_i^n-v_{i-1}^n}{\Delta x}+\frac{1}{2}\sigma_X^2(y)\frac{v_{i+1}^n-2v_i^n+v_{i-1}^n}{\Delta x^2}.$$

The resulting scheme is

$$A^h_{\Delta x}(y)v^n = B^h_{\Delta x}(y)v^{n+1}, \tag{4.4.70}$$

where $A^h_{\Delta x}(y)$ is the linear operator given by

$$(A^h_{\Delta x})_{ij}(y)=\begin{cases}-\beta^h_{\Delta x}(y)-|\alpha^h_{\Delta x}(y)|\mathbb{1}_{\alpha^h_{\Delta x}(y)<0}, & \text{if } i=j+1,\\ 1+2\beta^h_{\Delta x}(y)+|\alpha^h_{\Delta x}(y)|, & \text{if } i=j,\\ -\beta^h_{\Delta x}(y)-|\alpha^h_{\Delta x}(y)|\mathbb{1}_{\alpha^h_{\Delta x}(y)>0}, & \text{if } i=j-1,\\ 0, & \text{if } |i-j|>1,\end{cases} \tag{4.4.71}$$

with

$$\alpha^h_{\Delta x}(y)=\frac{h}{\Delta x}\mu_X(y),\qquad \beta^h_{\Delta x}(y)=\frac{h}{2\Delta x^2}\sigma_X^2(y),$$

and $B^h_{\Delta x}(y)$ is the linear operator defined in (4.4.47). Then we have:

Lemma 4.4.9. *For every $y \in \mathcal{D}$, the operator $A^h_{\Delta x}(y) : l_\infty(\mathcal{X}) \to l_\infty(\mathcal{X})$ is invertible and $|(A^h_{\Delta x})^{-1}(y)|_\infty \leq 1$. Moreover, $|B^h_{\Delta x}(y)|_\infty \leq 1 + 2\lambda c_\nu |\gamma_X(y)|$. Finally, if $\gamma_X \equiv 1$, $\Pi^h_{\Delta x}(y) = (A^h_{\Delta x})^{-1} B^h_{\Delta x}(y)$ is a stochastic operator, that is,*

$$(\Pi^h_{\Delta x})_{ij}(y) \geq 0, \quad i, j \in \mathbb{Z}, \qquad\qquad \sum_{j \in \mathbb{Z}} (\Pi^h_{\Delta x})_{ij}(y) = 1, \quad j \in \mathbb{Z}.$$

Proof. We write $A^h_{\Delta x}(y) = \eta(y) I - P(y)$, where $\eta(y) = 1 + 2\beta^h_{\Delta x}(y) + |\alpha^h_{\Delta x}(y)|$, I is the identity operator and $P_{ij}(y) = 0$ if $|i-j| \neq 1$ and $P_{ij} = -(A^h_{\Delta x})_{ij}$ if $|i-j| = 1$. So, it is easy to see that the operator $A^h_{\Delta x}(y) : l_\infty(\mathcal{X}) \to l_\infty(\mathcal{X})$ is invertible with inverse

$$(A^h_{\Delta x})^{-1}(y) = (\eta(y) I - P)^{-1} = \frac{1}{\eta} \sum_{k=0}^{\infty} \frac{P^k}{\eta^k}.$$

The assertion for $B^h_{\Delta x}(y)$ immediately follows from (4.4.47). Finally, $(A^h_{\Delta x})^{-1}_{ij}(y) \geq 0$ for all i, j because all entries of $P(y)$ are non negative and $(B^h_{\Delta x})_{ij}(y) \geq 0$ if $\mu_X \equiv 1$. Moreover, $\Pi^h_{\Delta x}(y) 1 = 1$ because, by construction, $A^h_{\Delta x}(y) 1 = 1$ and $B^h_{\Delta x}(y) 1 = 1$ when $\mu_X \equiv 1$. □

We can now state the convergence result.

Theorem 4.4.10. *Let $\tilde{u}^h_n$ be defined in (4.4.35) and u^h_n be given by (4.4.36) with the choice*

$$\Pi^h_{\Delta x}(y) = (A^h_{\Delta x})^{-1} B^h_{\Delta x}(y),$$

$A^h_{\Delta x}(y)$ and $B^h_{\Delta x}(y)$ being given in (4.4.71) and (4.4.47) respectively. Moreover, for $n = 0, \ldots, N$, consider the function

$$v^h_n(t, x, y) = \mathbb{E}\left[\tilde{u}^h_{n+1}(X^{t,x,y}_{(n+1)h}, Y^{t,y}_{(n+1)h})\right], \qquad t \in [nh, (n+1)h].$$

Assume that

- $\frac{\nu'}{\nu}, \frac{\nu''}{\nu} \in L^1(\mathbb{R}, d\nu)$;
- *the Markov chain $(Y^h_n)_{n=0,\ldots,N}$ satisfies assumptions $\mathcal{A}_1$, $\mathcal{A}_2$ and $\mathcal{A}_3(4\lambda c_\nu |\gamma_X|)$;*
- $v^h_n \in C^{\infty,4}_{\mathbf{pol},[nh,(n+1)h]}(\mathbb{R}, \mathcal{D})$ *a.e. and uniformly in n and h.*

Then, there exist $\bar{h}, C > 0$ such that for every $h < \bar{h}$ and $\Delta x < 1$ one has

$$|\tilde{u}^h_0(\cdot, Y_0) - u^h_0(\cdot, Y_0)|_\infty \leq CT(h + \Delta x^2).$$

Proof. By rewriting the proof of Proposition 4.4.8 in terms of the norm in $l_\infty(\mathcal{X})$, one gets that $\Pi^h_{\Delta x}(y)$ satisfies $\mathcal{B}(\infty, 2\lambda c_\nu|\gamma_X|, h+\Delta x)$. The statement now follows by applying Theorem 4.4.1. We only notice that here one applies (4.4.52) to the remaining term R_5 in (4.4.68). Since this term contains just v^h_n, one does not need more regularity for v^h_n, that's why we do not need that $v^h_n \in C^{\infty,6}_{\mathbf{pol},T}(\mathbb{R},\mathcal{D})$ and the class $C^{\infty,4}_{\mathbf{pol},T}(\mathbb{R},\mathcal{D})$ is enough. □

It is natural to look for conditions on the function f which ensure that the regularity assumptions on the function v^h_n for $n=0,\dots,N$, which are required In Theorem 4.4.10, are actually satisfied. Of course, these conditions depend on the regularity of the model. In Sections 4.5 and 4.6 we will study the case of the degenerate Heston or Bates model.

4.5 The European case in the Heston/Bates model

As an application in finance, in this section we apply our convergence results to to a tree-finite difference procedure for pricing European options in the Heston ([58]) or Bates ([17]) model: the asset price process S and the volatility process Y evolve following the stochastic differential system

$$\begin{aligned} \frac{dS_t}{S_{t^-}} &= (r-\delta)dt + \mu\sqrt{Y_t}\,dZ^1_t + \gamma d\tilde{H}_t,\\ dY_t &= \kappa(\theta - Y_t)dt + \sigma\sqrt{Y_t}\,dZ^2_t, \end{aligned} \tag{4.5.72}$$

where $S_0>0$, $Y_0\geq 0$, $Z=(Z^1,Z^2)$ is a correlated Brownian motions with $d\langle Z^1,Z^2\rangle_t=\rho dt$, $|\rho|<1$, $\tilde{H}$ is a compound Poisson process with intensity λ and i.i.d. jumps $\{\tilde{J}_k\}_k$ as in (4.4.23). Here, $\gamma=1$ (Bates model) or $\gamma=0$ (Heston model). The above quantities r and δ are the interest rate and the dividend interest rate respectively. We assume, as usual, that the Poisson process K, the jump amplitudes $\{\tilde{J}_k\}_k$ and the correlated Brownian motion (Z^1,Z^2) are independent.

With a simple transformation, we can reduce the model (4.5.72) to our reference model (4.4.22). To get rid of the correlated Brownian motion, we set

$$\bar\rho = \sqrt{1-\rho^2} \quad \text{and} \quad Z^2 = W, \quad Z^1 = \rho Z^2 + \bar\rho B,$$

in which (B,W) denotes a standard 2-dimensional Brownian motion. Moreover, considering the process $X_t = \log S_t - \frac{\rho}{\sigma}Y_t$, we reduce to the jump-diffusion pair (X,Y), which evolves according to

$$\begin{aligned} dX_t &= \mu_X(Y_t)dt + \bar\rho\sqrt{Y_t}\,dB_t + \gamma dH_t,\\ dY_t &= \kappa(\theta - Y_t)dt + \sigma\sqrt{Y_t}\,dW_t, \end{aligned} \tag{4.5.73}$$

where

$$\mu_X(y) = r - \delta - \frac{y}{2} - \frac{\rho}{\sigma}\kappa(\theta - y),$$

H_t is the compound Poisson process written through the Poisson process K, with intensity λ, and the i.i.d. jumps $J_k = \log(1+\tilde{J}_k)$. The standard Bates model requires that J_1 has a normal law. But it is clear that the convergence result holds for other laws such that the Lévy measure ν satisfies the requests in Theorem 4.4.4 or Theorem 4.4.10. For example, these properties hold for the mixture of exponential laws used by Kou [69].

In this section we focus on European options. Recall that, in this case, the function $\tilde{u}^h_n(\cdot)$ defined in (4.4.35) is nothing but the European price value at time nh, that is $u(nh, \cdot)$ where u is defined in (4.4.25). Moreover, we can easily see that, for any $n = N-1, \dots$, the function v^h_n defined in (4.4.50) satisfies

$$v^h_n(t, x, y) = u(t, x, y), \qquad t \in [nh, (n+1)h].$$

We consider the approximating Markov chain for the CIR process discussed in Section 4.3.1 and the two possible finite difference operator discussed in Section 4.4.3 and 4.4.3. As an application, we get the following convergence rate result of the hybrid method.

Theorem 4.5.1. *Let (X, Y) be the solution to (4.5.73) and let $(Y^h_n)_{n=0,\dots,N}$ be the Markov chain introduced in Section 4.3.1 for approximating the CIR process Y. Let $u(t, x, y) = \mathbb{E}(f(X^{t,x,y}_T, Y^{t,y}_T))$ be as in (4.4.25) and $(u^h_n)_{n=0,\dots,N}$ be given by (4.4.32) with the choice*

$$\Pi^h_{\Delta x}(y) = (A^h_{\Delta x})^{-1} B^h_{\Delta x}(y).$$

(*i*) [Convergence in $l_2(\mathcal{X})$] *Suppose that*

- *$A^h_{\Delta x}(y)$ and $B^h_{\Delta x}(y)$ are defined in (4.4.45) and (4.4.47) respectively;*
- *$\frac{\nu'}{\nu}, \frac{\nu''}{\nu} \in L^2(\mathbb{R}, d\nu)$ and ν has finite moments of any order;*
- *$\partial^{2j}_x f \in C^{2,6-j}_{\mathbf{pol}}(\mathbb{R}, \mathbb{R}_+)$ for every $j = 0, \dots, 6$.*

Then, there exist $\bar{h}, C > 0$ such that for every $h < \bar{h}$ and $\Delta x < 1$ one has

$$|u(0, \cdot, Y_0) - u^h_0(\cdot, Y_0)|_2 \le CT(h + \Delta x^2).$$

(*ii*) [Convergence in $l_\infty(\mathcal{X})$] *Suppose that*

- *$A^h_{\Delta x}(y)$ and $B^h_{\Delta x}(y)$ are defined in (4.4.71) and (4.4.47) respectively;*

- $\frac{\nu'}{\nu}, \frac{\nu''}{\nu} \in L^1(\mathbb{R}, d\nu)$ *and* ν *has finite moments of any order;*
- $\partial_x^{2j} f \in C^{\infty,4-j}_{\mathbf{pol}}(\mathbb{R}, \mathbb{R}_+)$ *for every* $j = 0, \dots, 4$.

Then, there exist $\bar h, C > 0$ *such that for every* $h < \bar h$ *and* $\Delta x < 1$ *one has*

$$|u(0, \cdot, Y_0) - u_0^h(\cdot, Y_0)|_\infty \le CT(h + \Delta x).$$

Proof. We apply Theorem 4.4.4 for (i) and Theorem 4.4.10 for (ii). The validity of assumptions $\mathcal{A}_1$ and $\mathcal{A}_2$ is proved in Proposition 4.3.4 and since here $\gamma_X = \gamma \in \{0,1\}$, $\mathcal{A}_3(4\lambda c_\nu|\gamma_X|)$ trivially holds. So, we need only to prove that if $\partial_x^{2j} f \in C^{2,6-j}_{\mathbf{pol}}(\mathbb{R}, \mathbb{R}_+)$ as $j = 0, 1, \dots, 6$, resp. $\partial_x^{2j} f \in C^{\infty,4-j}_{\mathbf{pol}}(\mathbb{R}, \mathbb{R}_+)$ as $j = 0, 1, \dots, 4$, then $u \in C^{2,6}_{\mathbf{pol},T}(\mathbb{R}, \mathbb{R}_+)$, resp. $u \in C^{\infty,4}_{\mathbf{pol},T}(\mathbb{R}, \mathbb{R}_+)$. This is proved in next Proposition 4.5.3 (set $\rho = 0$, $\mathfrak{a} = r - \delta - \frac{\rho}{\sigma}\kappa\theta$ and $\mathfrak{b} = \frac{\rho}{\sigma}\kappa - \frac{1}{2}$ therein), the whole Section 4.5.1 being devoted to. □

Remark 4.5.2. *In Chapter 3 we have considered the Bates-Hull-White model [27], which is a Bates model coupled with a stochastic interest rate. Recall that the dynamics follows* (4.5.72) *in which* r *is not constant but given by the Vasicek model*

$$dr_t = \kappa_r(\theta_r - r_t)dt + \sigma_r dZ^3_t,$$

Z^3 *being a Brownian motion correlated with* Z^1 *(and possibly* Z^2*). Here, there is no global transformation allowing one to reduce to our reference model. Nevertheless, a similar convergence result can be proved by means of the local transformation introduced in Section 3.4.1, acting on each time interval* $[nh, (n+1)h]$.

4.5.1 A regularity result for the Heston PDE/Bates PIDE

We deal here with a slightly more general model: we consider the SDE

$$\begin{aligned} dX_t &= (\mathfrak{a} + \mathfrak{b} Y_t)\, dt + \sqrt{Y_t}\, dW^1_t + \gamma_X dH_t, \\ dY_t &= \kappa(\theta - Y_t)dt + \sigma\sqrt{Y_t}\, dW^2_t, \end{aligned} \tag{4.5.74}$$

where W^1, W^2 are correlated Brownian motions with $d\langle W^1, W^2\rangle_t = \rho dt$ and H is a compound Poisson process with intensity λ and Lévy measure ν, which is assumed hereafter to have finite moments of any order. Here, $\mathfrak{a}, \mathfrak{b} \in \mathbb{R}$ and $\gamma_X \in \{0,1\}$ denote constant parameters. Note that when $\mathfrak{a} = r - \delta$ (interest rate minus dividend rate), $\mathfrak{b} = -\frac{1}{2}$ and $\gamma_X = 0$ (resp. $\gamma_X = 1$), then (X, Y) is the standard Heston (resp. Bates) model for the log-price and volatility. When instead $\rho = 0$, $\mathfrak{a} = r - \delta - \frac{\rho}{\sigma}\kappa\theta$ and $\mathfrak{b} = \frac{\rho}{\sigma}\kappa - \frac{1}{2}$, we recover the equation (4.5.73) discussed in Theorem 4.5.1.

Let $\mathcal{L}$ denote the infinitesimal generator associated to (4.5.74), that is,

$$\mathcal{L}u = \frac{y}{2}\left(\partial_x^2 u + 2\rho\sigma\partial_x\partial_y u + \sigma^2\partial_y^2 u\right) + (\mathfrak{a} + \mathfrak{b}y)\,\partial_x u + \kappa(\theta - y)\partial_y u + \mathcal{L}_{\text{int}} u, \tag{4.5.75}$$

where, hereafter, we set

$$\mathcal{L}_{\text{int}} u(t,x,y) = \gamma_X \int \big[u(t,x+\zeta,y) - u(t,x,y)\big]\nu(\zeta)d\zeta.$$

So, the present section is devoted to the proof of the following result.

Proposition 4.5.3. *Let $p \in [1,\infty]$, $q \in \mathbb{N}$ and suppose that $\partial_x^{2j} f \in C^{p,q-j}_{\mathbf{pol}}(\mathbb{R},\mathbb{R}_+)$ for every $j = 0,1,\dots,q$. Set*

$$u(t,x,y) = \mathbb{E}\big[f(X_T^{t,x,y}, Y_T^{t,y})\big].$$

Then $u \in C^{p,q}_{\mathbf{pol},T}(\mathbb{R},\mathbb{R}_+)$. Moreover, the following stochastic representation holds: for $m + 2n \le 2q$,

$$\begin{aligned}\partial_x^m\partial_y^n u(t,x,y) &= \mathbb{E}\left[e^{-n\kappa(T-t)}\partial_x^m\partial_y^n f(X_T^{n,t,x,y}, Y_T^{n,t,x,y})\right]\\ &+ n\,\mathbb{E}\left[\int_t^T \left[\frac{1}{2}\partial_x^{m+2}\partial_y^{n-1}u + \mathfrak{b}\partial_x^{m+1}\partial_y^{n-1}u\right](s, X_s^{n,t,x,y}, Y_s^{n,t,x,y})ds\right],\end{aligned} \tag{4.5.76}$$

where $\partial_x^m\partial_y^{n-1}u := 0$ when $n = 0$ and $(X^{n,t,x,y}, Y^{n,t,x,y})$, $n \ge 0$, denotes the solution starting from (x,y) at time t to the SDE (4.5.74) with parameters

$$\rho_n = \rho, \quad \mathfrak{a}_n = \mathfrak{a} + n\rho\sigma, \quad \mathfrak{b}_n = \mathfrak{b}, \quad \kappa_n = \kappa, \quad \theta_n = \theta + \frac{n\sigma^2}{2\kappa}, \quad \sigma_n = \sigma. \tag{4.5.77}$$

In particular, if $q \ge 2$ then $u \in C^{1,2}([0,T]\times\bar{\mathcal{O}})$, $\bar{\mathcal{O}} = \mathbb{R}\times\mathbb{R}_+$, solves the PIDE

$$\begin{cases}\partial_t u(t,x,y) + \mathcal{L}u(t,x,y) = 0, & t \in [0,T),\ (x,y) \in \bar{\mathcal{O}},\\ u(T,x,y) = f(x,y), & (x,y) \in \bar{\mathcal{O}}.\end{cases} \tag{4.5.78}$$

Remark 4.5.4. *For our purposes, we need both the polynomial growth condition for $(x,y) \mapsto u(t,x,y)$ and the L^p property for $x \mapsto u(t,x,y)$, and similarly for the derivatives. A closer look to the proof of Proposition 4.5.3 shows that the result holds also when one is not interested in the latter L^p condition. In this case, Proposition 4.5.3 reads: for $q \in \mathbb{N}$, if $\partial_x^{2j} f \in C^{q-j}_{\mathbf{pol}}(\mathbb{R}\times\mathbb{R}_+)$ for every $j = 0,1,\dots,q$ then $u \in C^{q}_{\mathbf{pol},T}(\mathbb{R}\times\mathbb{R}_+)$. Moreover, the stochastic representation* (4.5.76) *holds and, if $q \ge 2$, u solves PIDE* (4.5.78)*.*

As an immediate consequence of Proposition 4.5.3, we obtain the already known regularity result for the CIR process which has been already proved in Proposition 4.1 of [3].

Corollary 4.5.5. *Assume that $f = f(y)$ and set $u(t,y) = \mathbb{E}\big[f(Y_T^{t,y})\big]$. If $f \in C^q_{\mathbf{pol}}(\mathbb{R}_+)$, then $u \in C^q_{\mathbf{pol},T}(\mathbb{R}_+)$. Moreover, for $n \le q$,*

$$\partial_y^n u(t,y) = \mathbb{E}\left[e^{-n\kappa(T-t)}\partial_y^n f(Y_T^{n,t,y})\right],$$

where $Y^{n,t,y}$ denotes a CIR process starting from y at time t which solves the CIR dynamics with parameters $\kappa_n = \kappa$, $\theta_n = \theta + \frac{n\sigma^2}{2\kappa}$, $\sigma_n = \sigma$. In particular, if $q \ge 2$ then $u \in C^2_{\mathbf{pol}}(\mathbb{R}_+)$ solves the PDE

$$\begin{cases} \partial_t u + \mathcal{A}u = 0, & (t,y) \in [0,T) \times \mathbb{R}_+, \\ u_n(T,y) = \partial_y^n f(y), & y \in \mathbb{R}_+, \end{cases}$$

where $\mathcal{A}$ is the CIR infinitesimal generator (see (4.3.2)).

We first need some preliminary results. First of all, recall that X and Y have uniformly bounded moments: for every $T > 0$ and $a \ge 1$ there exist $A > 0$ such that for every $t \in [0,T]$,

$$\sup_{s\in[t,T]} \mathbb{E}[|X_s^{t,x,y}|^a] \le A(1 + |x|^a + y^a) \text{ and } \sup_{s\in[t,T]} \mathbb{E}[|Y_s^{t,y}|^a] \le A(1 + y^a). \tag{4.5.79}$$

For the second property in (4.5.79), we refer, for example, to [3], whereas the first one follows from standard techniques.

Lemma 4.5.6. *Let $p \in [0,\infty]$, $g \in C^{p,0}_{\mathbf{pol}}(\mathbb{R},\mathbb{R}_+)$, $h \in C^{p,0}_{\mathbf{pol},T}(\mathbb{R},\mathbb{R}_+)$ and consider the function*

$$u(t,x,y) = \mathbb{E}\left[e^{\varrho(T-t)}g(X_T^{t,x,y}, Y_T^{t,y}) - \int_t^T e^{\varrho(s-t)}h(s, X_s^{t,x,y}, Y_s^{t,y})ds\right], \tag{4.5.80}$$

where $\varrho \in \mathbb{R}$. Then $u \in C^{p,0}_{\mathbf{pol},T}(\mathbb{R},\mathbb{R}_+)$.

Proof. We set

$$u_1(t,x,y) = \mathbb{E}\left[e^{\varrho(T-t)}g(X_T^{t,x,y}, Y_T^{t,y})\right], \qquad u_2(t,x,y) = \mathbb{E}\left[\int_t^T e^{\varrho(s-t)}h(s, X_s^{t,x,y}, Y_s^{t,y})ds\right]$$

and we show that, for $i = 1,2$, $u_i \in C^{p,0}_{\mathbf{pol},T}(\mathbb{R},\mathbb{R}_+)$. We prove it for $i = 2$, the case $i = 1$ being similar and easier.

Fix $(t,x,y) \in [0,T] \times \mathbb{R} \times \mathbb{R}_+$ and let $(t_n,x_n,y_n)_n \subset [0,T] \times \mathbb{R} \times \mathbb{R}_+$ be such that $(t_n,x_n,y_n) \to (t,x,y)$ as $n \to \infty$. One can easily prove that, for every fixed $s \ge t_n \vee t$, $(X_s^{t_n,x_n,y_n}, Y_s^{t_n,y_n}) \to (X_s^{t,x,y}, Y_s^{t,y})$ in probability. We write u_2 as

$$u_2(t,x,y) = \int_0^T \mathbb{1}_{s>t} e^{\varrho(s-t)}\mathbb{E}\left[h(s, X_s^{t,x,y}, Y_s^{t,y})\right] ds$$

Since h is continuous, for $s > t_n \vee t$ the sequence $(h(s, X_s^{t_n,x_n,y_n}, Y_s^{t_n,y_n}))_n$ converges in probability to $h(s, X_s^{t,x,y}, Y_s^{t,y})$. By the polynomial growth of h and (4.5.79), for $p > 1$ we have

$$\sup_n \mathbb{E}[|h(X_T^{t_n,x_n,y_n}, Y_T^{t_n,y_n})|^p] \le \sup_n C\mathbb{E}[1 + |X_T^{t_n,y_n}|^{ap} + (Y_T^{t_n,y_n})^{ap}] < \infty. \tag{4.5.81}$$

Thus, $(h(X_T^{t_n,x_n,y_n}, Y_T^{t_n,y_n}))_n$ is uniformly integrable, so $h(X_T^{t_n,x_n,y_n}, Y_T^{t_n,y_n}) \to h(X_T^{t,x,y}, Y_T^{t,y})$ in L^1 and

$$\mathbb{1}_{s>t_n}\mathbb{E}\left[e^{\varrho(s-t_n)}h(s, X_s^{t_n,x_n,y_n}, Y_s^{t_n,y_n})\right] \to \mathbb{1}_{s>t}\mathbb{E}\left[e^{\varrho(s-t)}h(s, X_s^{t,x,y}, Y_s^{t,y})\right],$$

a.e. $s \in [0,T]$. By (4.5.81), $u_2(t_n, x_n, y_n) \to u_2(t, x, y)$ thanks to the Lebesgue's dominated convergence and moreover, u_2 grows polynomially. So, $u_2 \in \mathcal{C}_{\mathbf{pol},T}(\mathbb{R} \times \mathbb{R}_+)$.

Fix now $p \neq \infty$. We have

$$\begin{aligned}
\sup_{t\le T} \|u_2(t,\cdot,y)\|_{L^p(\mathbb{R},dx)} &= \sup_{t\le T}\left\|\mathbb{E}\left[\int_t^T e^{\varrho(s-t)}h(s, X_s^{t,\cdot,y}, Y_s^{t,y})ds\right]\right\|_{L^p(\mathbb{R},dx)} \\
&\le C\sup_{t\le T}\mathbb{E}\left[\int_t^T \left\|h(s, X_s^{t,\cdot,y}, Y_s^{t,y})\right\|^p_{L^p(\mathbb{R},dx)}\right]^{1/p} = C\sup_{t\le T}\mathbb{E}\left[\int_t^T \left\|h(s, \cdot + H_s^{t,y}, Y_s^{t,y})\right\|^p_{L^p(\mathbb{R},dx)}\right]^{1/p} \\
&= C\sup_{t\le T}\mathbb{E}\left[\int_t^T \left\|h(s, \cdot, Y_s^{t,y})\right\|^p_{L^p(\mathbb{R},dx)}\right]^{1/p} \le CT\sup_{t\le s\le T}(1 + \mathbb{E}[(Y_s^{t,y})^{pa}])^{1/p}
\end{aligned}$$

in which we have used twice the Cauchy-Schwarz inequality. Then, by using (4.5.79), we have $u_2 \in C^{p,0}_{\mathbf{pol},T}(\mathbb{R}, \mathbb{R}_+)$. The case $p = \infty$ follows the same lines. $\square$

To simplify the notation, from now on we set $\mathbb{E}^{t,x,y}[\cdot] = \mathbb{E}[\cdot | X_t = x, Y_t = y]$ and $\mathcal{O} = \mathbb{R} \times (0,\infty)$..

Lemma 4.5.7. *Let $g \in \mathcal{C}_{\mathbf{pol}}(\bar{\mathcal{O}})$ and $h \in C_{\mathbf{pol},T}(\bar{\mathcal{O}})$ be such that $\mathcal{O} \ni z \mapsto h(t,z)$ is locally Hölder continuous uniformly on the compact sets of $[0,T)$. Let u be defined in* (4.5.80)*. Then, $u \in \mathcal{C}([0,T] \times \bar{\mathcal{O}}) \cap \mathcal{C}^{1,2}([0,T) \times \mathcal{O})$ and solves the PIDE*

$$\begin{cases} \partial_t u + \mathcal{L}u + \varrho u = h, & \text{in } [0,T) \times \mathcal{O}, \\ u(T,z) = g(z), & \text{in } \mathcal{O}. \end{cases} \tag{4.5.82}$$

Moreover, if the Feller condition holds, that is, $2\kappa\theta \ge \sigma^2$, then u is the unique solution to (4.5.82) *in the class $C_{\mathbf{pol},T}(\bar{\mathcal{O}})$.*

Proof. Let $S \in [0,T)$, $\mathcal{R} = \mathbb{R} \times (\epsilon, \infty)$, $\epsilon > 0$, $Q = [0,S) \times \mathcal{R}$ and consider the PIDE problem

$$\begin{cases} \partial_t v + \mathcal{L}v + \varrho v = h, & \text{in } Q, \\ v = u, & \text{in } \partial_0 Q, \end{cases}$$

$\partial_0 Q$ denoting the parabolic boundary of Q. The coefficients satisfy in Q all the classical assumptions (see e.g. [53, 78]), so a unique (bounded) solution $v \in C^{1,2}([0,T) \times \mathcal{R}) \cap C([0,T] \times \bar{\mathcal{R}})$ actually exists (and have Hölder continuous derivatives v_t, $\nabla_z v$ and $D_z^2 v$ in $\bar{Q}$). As a consequence,

$$Z_s := e^{\varrho s} v(s, X_s, Y_s) - \int_t^s e^{\varrho r} h(r, X_r, Y_r) dr$$

is a martingale over $[t, S \wedge \tau_{\mathcal{R}}]$, where $\tau_{\mathcal{R}}$ denotes the exit time of (X, Y) from $\mathcal{R}$. Then,

$$\begin{aligned} e^{\varrho t} v(t,x,y) &= \mathbb{E}^{t,x,y}(Z_t) = \mathbb{E}^{t,x,y}(Z_{S\wedge\tau_{\mathcal{R}}}) \\ &= \mathbb{E}^{t,x,y}\Big[e^{\varrho S\wedge\tau_{\mathcal{R}}} u(S\wedge\tau_{\mathcal{R}}, X_{S\wedge\tau_{\mathcal{R}}}, Y_{S\wedge\tau_{\mathcal{R}}}) - \int_t^{S\wedge\tau_{\mathcal{R}}} e^{\varrho r} h(r, X_r, Y_r) dr\Big]. \end{aligned}$$

Now, by the strong Markov property,

$$e^{\varrho S\wedge\tau_{\mathcal{R}}} u(S\wedge\tau_{\mathcal{R}}, X_{S\wedge\tau_{\mathcal{R}}}, Y_{S\wedge\tau_{\mathcal{R}}}) = \mathbb{E}\Big[e^{\rho T} g(X_T, Y_T) - \int_{S\wedge\tau_{\mathcal{R}}}^T e^{\varrho r} h(r, X_r, Y_r) dr \,\Big|\, \mathcal{F}_{S\wedge\tau_{\mathcal{R}}}\Big].$$

By replacing above, it follows that $v \equiv u$ in Q. Whence, the first assertion is proved. Suppose now that $2\kappa\theta \geq \sigma^2$ and that g has polynomial growth. Let $w \in \mathcal{C}([0,T] \times \bar{\mathcal{O}})$ denote a solution to (4.5.82) with polynomial growth. We prove that $w = u$. Let $S_n < T$ and let $\mathcal{R}_n$ denote a sequence rectangles as before such that $Q_n = [0, S_n) \times \mathcal{R}_n \uparrow [0,T) \times \mathcal{O}$. Let w_n the unique solution to

$$\begin{cases} \partial_t w_n + \mathcal{L} w_n + \varrho w_n = h, & \text{in } Q_n, \\ w_n = w, & \text{in } \partial_0 Q_n. \end{cases}$$

Since w trivially solves the above PIDE problem, we get $w_n = w$ and

$$e^{\varrho t} w(t,x,y) = \mathbb{E}^{t,x,y}\Big[e^{\varrho S_n\wedge\tau_{\mathcal{R}_n}} w(S_n\wedge\tau_{\mathcal{R}_n}, X_{S_n\wedge\tau_{\mathcal{R}_n}}, Y_{S_n\wedge\tau_{\mathcal{R}_n}}) - \int_t^{S_n\wedge\tau_{\mathcal{R}_n}} e^{\varrho r} h(r, X_r, Y_r) dr\Big].$$

Now, as $n \to \infty$, one has $\tau_{\mathcal{R}_n} \uparrow \infty$ because, by the Feller condition, $\mathbb{P}^{t,y}(Y_s > 0 \,\forall s) = 1$. Then, we pass to the limit and since w is continuous and has polynomial growth, we easily obtain $w \equiv u$. $\square$

Lemma 4.5.8. *Let u be defined in* (4.5.80)*, with g and h such that, as $j = 0, 1$, $\partial_x^{2j} g \in C^{1-j}_{\mathbf{pol}}(\bar{\mathcal{O}})$ and $\partial_x^{2j} h \in \mathcal{C}^{1-j}_{\mathbf{pol},T}(\bar{\mathcal{O}})$. Then $u \in \mathcal{C}^1_{\mathbf{pol},T}(\bar{\mathcal{O}})$. Moreover, $\partial_x^2 u \in \mathcal{C}_{\mathbf{pol},T}(\bar{\mathcal{O}})$ and one has*

$$\partial_x^m u(t,x,y) = \mathbb{E}^{t,x,y}\left[e^{\varrho(T-t)} \partial_x^m g(X_T, Y_T) - \int_t^T e^{\varrho(s-t)} \partial_x^m h(s, X_s, Y_s) ds\right], \quad m = 1, 2, \tag{4.5.83}$$

$$\partial_y u(t,x,y) = \mathbb{E}^{t,x,y}\left[e^{(\varrho-\kappa)(T-t)} \partial_y g(X_T^*, Y_T^*) + \int_t^T e^{(\varrho-\kappa)(T-s)} \Big[\partial_y h + \frac{1}{2}\partial_x^2 u + \mathfrak{b}\partial_x u\Big](s, X_s^*, Y_s^*) ds\right], \tag{4.5.84}$$

where (X_t^*, Y_t^*) solves (4.5.74) *with new parameters $\rho_* = \rho$, $\mathfrak{a}_* = \mathfrak{a} + \rho\sigma$, $\mathfrak{b}_* = \mathfrak{b}$, $\kappa_* = \kappa$, $\theta_* = \theta + \frac{\sigma^2}{2\kappa}$, $\sigma_* = \sigma$.*

Proof. First, the stochastic flow w.r.t. x is differentiable (here, $(X^*)^{t,x,y}_s = x + Z^{t,y}_s$ and $Z^{t,y}_s$ does not depend on x). Hence, by using the polynomial growth hypothesis, by (4.5.80) one gets (4.5.83). Let us prove (4.5.84).

By Lemma 4.5.7 u solves (4.5.82). So, setting $v = \partial_y u$, by derivating (4.5.82) one has

$$\begin{cases} \partial_t v + \mathcal{L}_* v + \varrho_* v = h_*, & \text{in } [0,T) \times \mathcal{O}, \\ v(T,z) = g_*(z), & \text{in } \mathcal{O}. \end{cases}$$

where $\mathcal{L}_*$ is the infinitesimal generator of (X^*, Y^*) and $\varrho_* = \varrho - \kappa$, $h_* = \partial_y h - \mathfrak{b}\partial_x u - \frac{1}{2}\partial^2_x u$, $g_* = \partial_y g$. By using (4.5.83) and Lemma 4.5.6, $h_* \in C_{\mathbf{pol},T}(\bar{\mathcal{O}})$. Moreover, the Feller condition $2\kappa_*\theta_* \geq \sigma_*^2$ holds, and by Lemma 4.5.7 the unique solution with polynomial growth in (x,y) to the above PIDE is

$$\bar{v}(t,x,y) = \mathbb{E}^{t,x,y}\left[e^{\varrho(T-t)} g_*(X^*_T, Y^*_T) - \int_t^T e^{\varrho(s-t)} h_*(s, X^*_s, Y^*_s) ds \right].$$

In order to identify $\bar{v}$ with $v = \partial_y u$ we would need to know that $\partial_y u \in C_{\mathbf{pol},T}(\mathcal{O})$. If the diffusion coefficient of Y^* was more regular, one could use arguments from the stochastic flow. But this is not the case, hence we use a density argument inspired by [47].

For $k \geq 1$, let φ_k be a $C^\infty(\mathbb{R})$ approximation of $\sqrt{|y|}$ such that $\varphi_k(y) \geq 1/k$, $\varphi_k(y) \to \sqrt{|y|}$ uniformly on the compact sets of $[0,+\infty)$ and φ_k^2 is Lipschitz continuous uniformly in k (which means that $\varphi_k \varphi_k'$ is bounded uniformly in k). Consider the diffusion process (X^k, Y^k) defined by

$$\begin{cases} dX^k_t = \left(\mathfrak{a} + \mathfrak{b} Y^k_t\right) dt + \varphi_k(Y^k_t) dB_t + dH_t, \\ dY^k_t = \kappa(\theta - Y^k_t) dt + \sigma \varphi_k(Y^k_t) dW_t, \end{cases} \tag{4.5.85}$$

whose generator is

$$\mathcal{L}_k u = \frac{\varphi_k^2(y)}{2}\left(\partial_x^2 u + 2\rho\sigma\partial_x\partial_y u + \sigma^2\partial_y^2 u\right) + (\mathfrak{a} + \mathfrak{b} y)\,\partial_x u + \kappa(\theta - y)\partial_y u + \mathcal{I}u.$$

Set

$$u^k(t,x,y) = \mathbb{E}^{t,x,y}\left[e^{\varrho(T-t)} g(X^k_T, Y^k_T) - \int_t^T e^{\varrho(s-t)} h(s, X^k_s, Y^k_s) ds \right].$$

Le us first show that $\partial_y u^k \in C_{\mathbf{pol},T}(\mathcal{O})$. Since the diffusion coefficients associated to (X^k, Y^k) are good enough, we can consider the first variation process: by calling $Z^{k,t,x,y}_s = (\partial_y X^{k,t,x,y}_s, \partial_y Y^{k,t,x,y}_s)$, we get

$$\begin{aligned} \partial_y u^k(t,x,y) =& \mathbb{E}\left[e^{\varrho(T-t)} \left\langle \nabla_{x,y} g(X^{k,t,x,y}_T, Y^{k,t,x,y}_T), Z^{k,t,x,y}_T \right\rangle \right] \\ &- \int_t^T e^{\varrho(s-t)} \mathbb{E}\left[\left\langle \nabla_{x,y} h(s, X^{k,t,x,y}_s, Y^{k,t,x,y}_s), Z^{k,t,x,y}_s \right\rangle \right] ds. \end{aligned}$$

The functions g, h and their derivatives have polynomial growth, so

$$\begin{aligned}\left|\partial_y u^k(t,x,y)\right| \leq & \mathbb{E}\left[C(1+|X_T^{k,t,x,y}|^a+|Y_T^{k,t,x,y}|^a)|Z_T^{k,t,x,y}|\right] \\ & + \int_t^T e^{\varrho(s-t)}\mathbb{E}\left[C(1+|X_s^{k,t,x,y}|^a+|Y_s^{k,t,x,y}|^a)|Z_s^{k,t,x,y}|\right] ds\end{aligned}$$

and the usual L^p-estimates give

$$\sup_{t<T}\left|\partial_y u^k(t,x,y)\right| \leq C_k(1+|x|^{a_k}+y^{a_k}),$$

for suitable constants $C_k, a_k > 0$. Moreover, from the standard theory of parabolic PIDEs, u^k is a solution to

$$\begin{cases}\partial_t u^k + \mathcal{L}_k u^k + \varrho u^k = h, & \text{in } [0,T)\times\mathcal{O}, \\ u^k(T,z) = g(z), & \text{in } \mathcal{O}.\end{cases}$$

By differentiating, $v^k = \partial_y u^k$ solves the problem

$$\begin{cases}\partial_t v^k + \mathcal{L}_{k,*} v^k + \varrho_* v^k = h_{k,*}, & \text{in } [0,T)\times\mathcal{O}, \\ v^k(T,z) = g_*(z), & \text{in } \mathcal{O}.\end{cases}$$

where

$$\begin{aligned}\mathcal{L}_{k,*}v = & \frac{\varphi_k^2(y)}{2}\left(\partial_x^2 v + 2\rho\sigma\partial_x\partial_y v + \sigma^2\partial_y^2 v\right) \\ & + \left(\mathfrak{a}+\mathfrak{b}y+2\rho\sigma\varphi_k\varphi_k'(y)\right)\partial_x v + \left(\kappa(\theta-y)+\sigma^2\varphi_k\varphi_k'(y)\right)\partial_y v + \mathcal{I}v\end{aligned}$$

and $h_{k,*} = \partial_y h - \mathfrak{b}\partial_x u^k - \varphi_k\varphi_k'(y)\partial_x^2 u^k$. By developing the same arguments as before, we get $h_{k,*} \in C_{\mathbf{pol},T}(\bar{\mathcal{O}})$. The PIDE for v^k has a unique solution in $C_{\mathbf{pol},T}(\mathcal{O})$ (recall that, by construction, the second order operator is uniformly elliptic). Thus, the Feynman-Kac formula gives

$$\partial_y u^k(t,x,Y) = \mathbb{E}^{t,x,y}\left[e^{\varrho(T-t)}g_*(X_T^{k,*},Y_T^{k,*}) - \int_t^T e^{\varrho(s-t)}h_{k,*}(s,X_s^{k,*},Y_s^{k,*})ds\right],$$

where $(X^{k,*},Y^{k,*})$ is the diffusion with infinitesimal generator given by $\mathcal{L}_{k,*}$. Now, the standard L^p estimates for (X^k,Y^k) and $(X^{k,*},Y^{k,*})$ hold uniformly in k (recall that φ_k is sublinear uniformly in k and $\varphi_k\varphi_k'$ is bounded uniformly in k): for every $p\geq 1$ there exist $C, a > 0$ such that

$$\sup_k\sup_{t\leq T}\mathbb{E}^{t,x,y}\left(|X_t^k|^p+|Y_t^k|^p\right) + \sup_k\sup_{t\leq T}\mathbb{E}^{t,x,y}\left(|X_t^{k,*}|^p+|Y_t^{k,*}|^p\right) \leq C(1+|x|^a+|y|^a).$$

This gives that

$$\sup_k\sup_{t<T}|u^k(t,x,y)| + \sup_k\sup_{t<T}\left|\partial_y u^k(t,x,y)\right| \leq C(1+|x|^a+|y|^a),$$

for suitable $C, a > 0$ (possibly different from the ones above). Moreover, using the stability results of [12] one obtains

$$\lim_{n\to\infty} u^k(t,x,y) = u(t,x,y) \quad \text{and} \quad \lim_{n\to\infty} \partial_y u^k(t,x,y) = v(t,x,y)$$

for every $(t,x,y) \in [0,T) \times \mathcal{O}$. And thanks to the above uniform polynomial bounds for u^k and $\partial_y u^k$, for every $\phi \in C^\infty(\mathcal{O})$ with compact support we easily get

$$\int v(t,x,y)\phi(x,y)dxdy = \int \lim_k \partial_y u^k(t,x,y)\phi(x,y)dxdy$$
$$= -\int \lim_k u^k(t,x,y)\partial_y\phi(x,y)dxdy = -\int u(t,x,y)\partial\phi(x,y)dxdy.$$

Therefore, $v(t,x,y) = \partial_y u(t,x,y)$ in $[0,T) \times \mathcal{O}$. The statement now follows.

□

We can now prove the result which this section is devoted to.

Proof of Proposition 4.5.3. We follow an induction on q. If $q = 0$, Lemma 4.5.6 gives the result. Suppose the statement is true up to $q-1 \geq 1$ and let us prove it for q.

Take f such that $\partial_x^{2j} f \in C^{p,q-j}_{\mathbf{pol}}(\mathbb{R},\mathbb{R}_+)$ for every $j = 0,1,\dots,q$. Then, by induction, $\partial_t^l\partial_x^m\partial_y^n u \in C^{p,0}_{\mathbf{pol},T}(\mathbb{R},\mathbb{R}_+)$ when $2l+m+n \leq q-1$. So, we just need to prove that $\partial_t^l\partial_x^m\partial_y^n u \in C^{p,0}_{\mathbf{pol},T}(\mathbb{R},\mathbb{R}_+)$ for any l,m,n such that $2l+m+n = q$.

Assume first $l = 0$. For $n = 0$, we use that $X_T^{t,x,y} = x + Z_T^{t,y}$ and we get $\partial_x^m u(t,x,y) = \mathbb{E}^{t,x,y}\big[\partial_x^m f(X_T,Y_T)\big]$. Since $\partial_x^m f \in C^{p,0}_{\mathbf{pol}}(\mathbb{R},\mathbb{R}_+)$ for any $m \leq 2q$, by Lemma 4.5.6 we obtain $\partial_x^m u \in C^{p,0}_{\mathbf{pol},T}(\mathbb{R},\mathbb{R}_+)$ for every $m \leq 2q$.

Fix now $n > 0$ and $m \geq 0$. Recursively applying Lemma 4.5.8, we get formula (4.5.76). Let us stress that, because of the presence of the derivatives $\partial_x^{m+2}\partial_y^{n-1}u$ and $\partial_x^{m+1}\partial_y^{n-1}u$ in (4.5.76), the recursively application of Lemma 4.5.8 gives the constraint $m+2n \leq q$. Then, by Lemma 4.5.6, it follows that $\partial_x^m\partial_y^n u \in C^{p,0}_{\mathbf{pol},T}(\mathbb{R},\mathbb{R}_+)$ for every $m,n \in \mathbb{N}$ such that $m+2n \leq 2q$, and in particular when $m+n = q$.

Consider now the case $l > 0$. By (4.5.76), Lemma 4.5.7 ensures that if $m+2n \leq 2q$ then $u_{n,m} = \partial_x^m\partial_y^n u$ solves

$$\begin{cases} \partial_t u_{m,n} + \mathcal{L}_n u_{m,n} - n\kappa u_{m,n} = -n\big[\frac{1}{2}u_{m+2,n-1} + \mathfrak{b}u_{m+1,n-1}\big] & \text{in } [0,T)\times\mathcal{O}, \\ u_{m,n}(T,x,y) = \partial_x^m\partial_y^n f(x,y) & \text{in } \mathcal{O}, \end{cases} \tag{4.5.86}$$

where $\mathcal{L}_n$ is the generator in (4.5.75) with the (new) parameters in (4.5.77). Therefore, the general case concerning $\partial_t^l \partial_x^m \partial_y^n u$ with $2l + m + n = q$ follows by an iteration on l: by (4.5.86),

$$\partial_t^l \partial_x^m \partial_y^n u = -\mathcal{L}_n \partial_t^{l-1} \partial_x^m \partial_y^n u + n\kappa \partial_t^{l-1} \partial_x^m \partial_y^n u - n\Big[\frac{1}{2}\partial_t^{l-1} \partial_x^{m+2} \partial_y^{n-1} u + \mathfrak{b}\partial_t^{l-1} \partial_x^{m+1} \partial_y^{n-1} u\Big].$$

□

4.6 The American case in the Heston/Bates model

In this section we focus on the American case. We first prove a simple lemma which better specifies the behaviour of the moments in the Heston and Bates model.

Lemma 4.6.1. *For every $p \geq 2$ there exists $C > 0$ (depending on p and on the model parameters) such that*

$$\sup_{t \in [nh,(n+1)h]} \mathbb{E}[|X^{t,x,y}_{(n+1)h}|^p] \leq (1 + Ch)(1 + |x|^p + y^p), \tag{4.6.87}$$

$$\sup_{t \in [nh,(n+1)h]} \mathbb{E}[(Y^{t,y}_{(n+1)h})^p] \leq (1 + Ch)(1 + y^p). \tag{4.6.88}$$

Proof. It can be easily proved that there exists $C > 0$ such that

$$\sup_{t\in[0,T]} \mathbb{E}[|X_t|^p] \leq C(1 + |x|^p + y^p), \qquad \sup_{t\in[0,T]} \mathbb{E}[(Y^{t,y}_t)^p] \leq C(1 + y^p). \tag{4.6.89}$$

We start by proving (4.6.88). Let us fix $p \geq 1$. By using Itô's Lemma, for any $t \in [nh,(n+1)h]$ we have

$$(Y^{t,y}_{(n+1)h})^p = y^p + p\int_t^{(n+1)h} \Big(\Big(\kappa\theta - \frac{p-1}{2}\sigma^2\Big)(Y^{t,y}_s)^{p-1} - \kappa(Y^{t,y}_s)^p\Big)\, ds + p\sigma \int_t^{(n+1)h} (Y^{t,y}_s)^{p-\frac{1}{2}} dW_s.$$

Passing to the expectation and using (4.6.89), we can find $C > 0$ (depending on p and on the coefficients of the model) such that

$$\sup_{t\in[nh,(n+1)h]} \mathbb{E}[(Y^{t,y}_{(n+1)h})^p] \leq y^p + hC(1 + y^{p-1} + y^p) \leq (1 + 2Ch)(1 + y^p),$$

from which (4.6.88) follows. As regards (4.6.87), again by Itô's Lemma, for $t \in [nh,(n+1)h]$ we get

$$\begin{aligned}|X^{t,x,y}_{(n+1)h}|^{2p} = x^p + &\int_t^{(n+1)h} \Big[2p\mu_X(Y^{t,y}_s)(X^{t,x,y}_{s^-})^{2p-1} + p(2p-1)\sigma_X^2(Y^{t,y}_s)(X^{t,x,y}_{s^-})^{2p-2}\Big]\, ds \\ &+ \int_t^{(n+1)h} (X^{t,x,y}_{s^-} + J_{N_s})^{2p} - (X^{t,x,y}_{s^-})^{2p} dK_s + \int_t^{(n+1)h} 2p\sigma_X(Y^{t,y}_s)(X^{t,x,y}_{s^-})^{2p-1} dB_s,\end{aligned}$$

K denoting the Poisson process driving the compound Poisson process H, whose associated Lévy measure is ν. Passing to the expectation, and using the martingale properties (which hold thanks to (4.6.89)) we get

$$\mathbb{E}[|X^{t,x,y}_{(n+1)h}|^{2p}] = x^{2p} + \int_t^{(n+1)h} \left[\mathbb{E}[2p\mu_X(Y^{t,y}_s)(X^{t,x,y}_s)^{2p-1} + p(2p-1)\sigma^2_X(Y^{t,y}_s)(X^{t,x,y}_s)^{2p-2}]\right] ds$$
$$+ \int_t^{(n+1)h} ds \int \mathbb{E}[(X^{t,x,y}_s + z)^{2p} - (X^{t,x,y}_s)^{2p}]\nu(dz).$$

(4.6.87) now follows by using Hölder inequality, the estimate (4.6.89) and the existence of all moments under ν. □

Again, we approximate the CIR process with the Markov chain discussed in Section 4.3.1 and we consider the two finite difference operators introduced in Section 4.4.3 and 4.4.3. Therefore, we get the following convergence rate result.

Theorem 4.6.2. *Let* (X, Y) *be the solution to* (4.5.73) *and let* $(Y^h_n)_{n=0,\dots,N}$ *be the Markov chain introduced in Section 4.3.1 for the approximation of the CIR process* Y*. Let* $\tilde{u}^h_n$ *be defined in* (4.4.33) *and* u^h_h *be given by* (4.4.34) *with the choice*

$$\Pi^h_{\Delta x}(y) = (A^h_{\Delta x})^{-1} B^h_{\Delta x}(y).$$

(*i*) [Convergence in $l_2(\mathcal{X})$] *Suppose that*

- $A^h_{\Delta x}(y)$ *and* $B^h_{\Delta x}(y)$ *are defined in* (4.4.45) *and* (4.4.47) *respectively;*
- $\frac{\nu'}{\nu}, \frac{\nu''}{\nu} \in L^2(\mathbb{R}, d\nu)$ *and* ν *has finite moments of any order;*
- $f \in C^\infty_{\mathbf{pol}}(\mathbb{R} \times \mathcal{D})$ *is such that there exist* $C, a > 0$ *with*

$$|\partial^{l'}_x \partial^l_y f(\cdot, y)|_{L^2(\mathbb{R},dx)} \le C(1 + y^a), \qquad l', l \in \mathbb{N}.$$

Then, there exist $\bar{h}, C > 0$ *such that for every* $h < \bar{h}$ *and* $\Delta x < 1$ *one has*

$$|u(0, \cdot, Y_0) - u^h_0(\cdot, Y_0)|_2 \le CT(h + \Delta x^2).$$

(*ii*) [Convergence in $l_\infty(\mathcal{X})$] *Suppose that*

- $A^h_{\Delta x}(y)$ *and* $B^h_{\Delta x}(y)$ *are defined in* (4.4.71) *and* (4.4.47) *respectively;*
- $\frac{\nu'}{\nu}, \frac{\nu''}{\nu} \in L^1(\mathbb{R}, d\nu)$ *and* ν *has finite moments of any order;*

- *$f \in C^{\infty}_{\mathbf{pol}}(\mathbb{R}\times\mathcal{D})$ is such that there exist $C, a > 0$ with*

$$|\partial_x^{l'}\partial_y^l f(\cdot,y)|_{L^\infty(\mathbb{R},dx)} \le C(1+y^a), \qquad l', l \in \mathbb{N}.$$

Then, there exist $\bar h, C > 0$ such that for every $h < \bar h$ and $\Delta x < 1$ one has

$$|u(0,\cdot,Y_0) - u_0^h(\cdot,Y_0)|_\infty \le CT(h+\Delta x).$$

Proof. We prove (i), (ii) following in the same way. The validity of assumptions $\mathcal{A}_1$ and $\mathcal{A}_2$ is proved in Proposition 4.3.4 and since $\gamma_X \equiv 1$ or $\gamma_X \equiv 0$, $\mathcal{A}_3(4\lambda c_\nu |\gamma_X|)$ trivially holds. So, as in the European case, in order to apply Theorem 4.4.4 it is enough to prove that the function v_n^h defined in (4.4.50) belongs to the space $C^{2,6}_{\mathbf{pol},[nh,(n+1)h]}(\mathbb{R},\mathcal{D})$ a.e. and uniformly in n and h.

Let us consider a function $f \in C^{\infty}_{\mathbf{pol}}(\mathbb{R}\times\mathcal{D})$ such that for any $l, l' \in \mathbb{N}$ there exist $C_{l',l}$, $a_{l,l'} > 0$ such that

$$|\partial_x^{l'}\partial_y^l f(\cdot,y)|_{L^2(\mathbb{R},dx)} \le C_{l',l}(1+y^{a_{l,l'}}), \qquad y \in \mathcal{D}. \tag{4.6.90}$$

We point out that in the statement of the theorem we actually require that there exist $C, a > 0$ such that $C_{l',l} \le C$ and $a_{l',l} \le a$ for any $l, l' \in \mathbb{N}$. We will use this strong assumption only at the end of the proof, when it will be clear why we need it in order to get the assertion.

We proceed by a backward iteration. For $n = N-1$ we have $v_{N-1}^h(t,x,y) = \mathbb{E}\big[f(X_T^{t,x,y}, Y_T^{t,y})\big]$. By the proof of Proposition 4.5.3 and by using (4.6.87) and (4.6.88), we deduce that, if $l = 0$, by using (4.6.87)-(4.6.88) we have

$$\sup_{t\in[(N-1)h,T)} |\partial_x^{l'} v_{N-1}^h(t,\cdot,y)|_{L^2(\mathbb{R},dx)} \le C_{l',0}(1+C_0h)(1+y^{a_{l',0}}).$$

On the other hand, again from the proof of Proposition 4.5.3, we have that, for $t \in [(N-1)h, T)$,

$$\begin{aligned}\partial_x^{l'}\partial_y^l v_{N-1}^h(t,x,y) &= \mathbb{E}\left[e^{-l\kappa(T-t)}\partial_x^{l'}\partial_y^l f(X_T^{l,t,x,y}, Y_T^{l,t,x,y})\right]\\ &+ l\,\mathbb{E}\left[\int_t^T \left[\frac{1}{2}\partial_x^{l'+2}\partial_y^{l-1}v_{N-1}^h + b\partial_x^{l'+1}\partial_y^{l-1}v_{N-1}^h\right](s, X_s^{l,t,x,y}, Y_s^{l,t,x,y})ds\right],\end{aligned} \tag{4.6.91}$$

where $b = \frac{\rho}{\sigma}\kappa - \frac{1}{2}$ and (X^l, Y^l) is the solution of the Heston/Bates model with new coefficients $r_l = r + l\rho\sigma$, $\kappa_l = \kappa$, $\theta_l = \theta + \frac{l\sigma^2}{2\kappa}$, $\sigma_l = \sigma$. Denote by C_l the constant such that

$$\sup_{t\in[(N-1)h,T)} \mathbb{E}^{t,y}[(Y^l_{(n+1)h})^p] \le (1+y^p)(1+C_lh).$$

Then, if $l = 1$, by (4.6.91) we get

$$\begin{aligned}\sup_{t\in[(N-1)h,T)} |\partial_x^{l'}\partial_y v_{N-1}^h(t,\cdot,y)| &\le C_{l',1}(1+C_1h)(1+y^{a_{l',1}})\\ &+ h\left(\frac{1}{2}C_{l'+2,0}(1+C_1h)(1+y^{a_{l'+2,0}}) + |b|C_{l'+1,0}(1+C_1h)(1+y^{a_{l'+1,0}})\right).\end{aligned}$$

Without loss of generality we can assume that $\frac{1}{2}+|b|\le C_1$, $C_i\le C_{i+1}$ and that the constants $C_{l,l'}$ and $a_{l,l'}$ are nondecreasing in both l and l'. Then, we easily deduce that

$$\sup_{t\in[(N-1)h,T)} |\partial_x^{l'}\partial_y v_{N-1}^h(t,\cdot,y)|_{L^2(\mathbb{R},dx)} \le C_{l'+2,1}(1+C_1h)^2(1+y^{a_{l'+2,1}}).$$

With the same arguments, if $l=2$, we get

$$\sup_{t\in[N-1)h,T)} |\partial_x^{l'}\partial_y^2 v_{N-1}^h(t,\cdot,y)|_{L^2(\mathbb{R},dx)} \le C_{l'+4,1}(1+C_2h)^3(1+y^{a_{l'+4,1}}).$$

By iterating, it can be easily seen that

$$\sup_{t\in[N-1)h,T)} |\partial_x^{l'}\partial_y^l v_{N-1}^h(t,\cdot,y)|_{L^2(\mathbb{R},dx)} \le C_{l',l}^{(h,N-1)}\left(1+y^{a_{l',l}^{(N-1)}}\right),$$

where

$$C_{l',l}^{(h,N-1)} = C_{l'+2l,l}(1+C_lh)^{l+1}, \qquad a_{l',l}^{(N-1)} = a_{l'+2l,l}.$$

As regard the derivatives w.r.t. the time variable, again from the proof of Proposition 4.5.3, we have

$$\begin{aligned}\partial_t^{l''}\partial_x^{l'}\partial_y^l v_{N-1}^h = &-\mathcal{L}_l\partial_t^{l-1}\partial_x^{l'}\partial_y^l v_{N-1}^h + l\kappa\partial_t^{l''-1}\partial_x^{l'}\partial_y^l v_{N-1}^h\\ &- l\Big[\frac{1}{2}\partial_t^{l-1}\partial_x^{l'+2}\partial_y^{l-1}v_{N-1}^h + b\partial_t^{l''-1}\partial_x^{l'+1}\partial_y^{l-1}v_{N-1}^h\Big],\end{aligned}$$

so that

$$\sup_{t\in[nh,(n+1)h)} |\partial_t^{l''}\partial_x^{l'}\partial_y^l v_{N-1}^h(t,\cdot,y)|_{L^2(\mathbb{R},dx)} \le clC_{l'+2,l+2}^{(h,N-1)}\left(1+y^{a_{l',l}^{(N-1)}+l''}\right), \tag{4.6.92}$$

where c is a constant which depends on the coefficient of the model.

Therefore,

$$\tilde{u}_{N-1}^h(x,y) = \max\{f(x,y), v_{N-1}^h((N-1)h,x,y)\}$$

is a continuous function, whose derivatives, of any order, a.e. continuously exist and for every l',l,

$$|\partial_x^{l'}\partial_y^l\tilde{u}_{N-1}^h(\cdot,y)|_{L^2(\mathbb{R},dx)} \le C_{l',l}^{(h,N-1)}\left(1+y^{a_{l',l}^{(N-1)}}\right) \quad \text{a.e..} \tag{4.6.93}$$

Note that the estimates (4.6.92) on the time derivatives of v_{N-1}^h are not involved in the estimate (4.6.93) and, as a consequence, in the iterative procedure.

At time step $n=N-2$ the function v_{N-2}^h is defined by

$$v_{N-2}^h(t,x,y) = \mathbb{E}\big[\tilde{u}_{N-1}^h(X_{(N-1)h}^{t,x,y}, Y_{(N-1)h}^{t,y})\big], \qquad t\in[(N-2)h,(N-1)h].$$

By developing arguments already done for $n = N - 1$, we get

$$\sup_{t\in[N-1)h,T)} |\partial_x^{l'} \partial_y v_{N-2}^h(t,\cdot,y)|_{L^2(\mathbb{R},dx)} \leq C_{l',l}^{(h,N-2)} \left(1 + y^{a_{l',l}^{(N-2)}}\right),$$

where

$$C_{l',l}^{(h,N-2)} = C_{l'+2l,l}^{(h,N-1}(1+C_l h)^{l+1} = C_{l'+4l,l}(1+C_l h)^{2(l+1)}, \qquad a_{l',l}^{(N-2)} = a_{l'+4l,l}.$$

Moreover

$$\sup_{t\in[nh,(n+1)h)} |\partial_t^{l''} \partial_x^{l'} \partial_y^l v_{N-2}^h(t,\cdot,y)|_{L^2(\mathbb{R},dx)} \leq clC_{l'+,l+2}^{h,N-2} \left(1 + y^{a_{l',l}^{N-2}+l''}\right).$$

Therefore, the function

$$\tilde{u}_{N-2}^h(x,y) = \max\{f(x,y), v_{N-2}^h((N-2)h,x,y)\}$$

is a continuous function, whose derivatives, of any order, a.e. continuously exist and for every l', l,

$$|\partial_x^{l'} \partial_y^l \tilde{u}_{N-2}^h(\cdot,y)|_{L^2(\mathbb{R},dx)} \leq C_{l',l}^{h,N-2} \left(1 + y^{a_{l',l}^{N-2}+l''}\right) \quad \text{a.e.},$$

By iterating, we get that, at time step $n = N - k$, the function v_{N-k}^h satisfies

$$|\partial_x^{l'} \partial_y^l v_{N-k}^h(\cdot,y)|_{L^2(\mathbb{R},dx)} \leq C_{l',l}^{(h,N-k)} \left(1 + y^{a_{l',l}^{(N-k)}+l''}\right) \quad \text{a.e.},$$

where

$$C_{l',l}^{(h,N-k)} = C_{l'+2kl,l}(1+C_l h)^{k(l+1)}, \qquad a_{l',l}^{(N-k)} = a_{l'+2kl,l}.$$

Again

$$\sup_{t\in[nh,(n+1)h)} |\partial_t^{l''} \partial_x^{l'} \partial_y^l v_{N-k}^h(t,\cdot,y)|_{L^2(\mathbb{R},dx)} \leq clC_{l'+2,l+2}^{(h,N-k)} \left(1 + y^{a_{l',l}^{(N-k)}+l''}\right).$$

In order to have $v_n^h \in C^{2,6}_{\mathbf{pol},[nh,(n+1)h]}(\mathbb{R},\mathcal{D})$ a.e. and uniformly in n and h, we need estimates of the derivatives $\partial_x^{l'} \partial_y^l v_n^h$ for $l + l' \leq 6$ which are uniform in n and h. It is clear that for each $k \leq N$, since $h = T/N$ and $l \leq 6$,

$$(1+C_l h)^{k(l+1)} \leq e^{C_l hN(l+1)} \leq e^{7TC_6}.$$

Moreover, the assumption that there exist $C, a > 0$ such that $C_{l',l} \leq C$ and $a_{l',l} \leq a$ for any $l, l' \in \mathbb{N}$ now comes in. Thanks to this, we can deduce that $v_n^h \in C^{2,6}_{\mathbf{pol},[nh,(n+1)h]}(\mathbb{R},\mathcal{D})$ a.e. and uniformly in n and h, so by Theorem 4.4.4 we get the result. □

Remark 4.6.3. *In Theorem 4.6.2 we require really strong regularity and boundedness assumptions on the test function f. On the other hand, let us stress that our algorithm is strongly based on numerical analysis techniques. When these procedures are used, as far as we know, literature is missing in results on the rate of convergence of numerical schemes for obstacle problems.*

Let us mention that, in some particular cases, different approaches could in principle be followed. For example, let us consider the scheme introduced in Section 4.4.3, where the linear operator is given by

$$\Pi^h_{\Delta x}(y) = (A^h_{\Delta x})^{-1} B^h_{\Delta x}(y),$$

$A^h_{\Delta x}(y)$ *and* $B^h_{\Delta x}(y)$ *being defined in* (4.4.71) *and* (4.4.47) *respectively. Here, we have proved in Lemma 4.4.9 that* $\Pi^h_{\Delta x}(y)$ *is a stochastic operator. From a probabilistic point of view, this means that the algorithm can be written through a Markov chain (see [24]). Then, one could apply purely probabilistic methods to prove the convergence of the procedure, for example by developing techniques similar to the ones introduced in [13]. On the other hand, in this case,* $\Pi^h_{\Delta x}(y)$ *is a monotone linear operator, so another possible way to proceed is to use the theory introduced by Barles [15], which uses viscosity solutions. In order to do this, we need a comparison principle for viscosity solutions of Heston-type degenerate parabolic problems (note that in Section 1.3 we have proved such a result in the case of weak solutions). However, both the mentioned approaches give in principle just the convergence, that is, no information about the rate of convergence is provided.*

4.7 Appendix

4.7.1 Lattice properties of the CIR approximating tree

The aim of this section is to prove Propostition 4.3.3. For later use, let us first give some (trivial) properties of the lattice. First, by construction, $k_d(n,k) \le k < k_u(n,k)$, so that $y^{n+1}_{k_d(n,k)} \le y^{n+1}_k \le y^n_k \le y^{n+1}_{k+1} \le y^{n+1}_{k_u(n,k)}$. Moreover for every n and k, it is easy to see that

$$\begin{aligned} y^n_k \le y^n_{k+1}, \quad y^{n+1}_k \le y^n_k \le y^{n+1}_{k+1}, \\ y^n_k \le y^n_{k-1} + \sigma^2 h + 2\sigma\sqrt{v^n_{k-1} h}, \quad y^{n+1}_k \le y^n_k + \frac{\sigma^2}{4}\, h - \sigma\sqrt{y^n_k h}. \end{aligned} \tag{4.7.94}$$

Proof of Proposition 4.3.3. 1. The statement is an immediate consequence of the following facts:

$$\text{if } k_u(n,k) \ge k+2, \text{ then } y^n_k < \theta_* h, \tag{4.7.95}$$

$$\text{if } k_d(n,k) \le k-1, \text{ then } y^n_k > \theta^*/h, \tag{4.7.96}$$

which we now prove.

First of all, note that $y_k^n + \mu_Y(y_k^n)h = \kappa\theta h + y_k^n(1-\kappa h)$, so by choosing $\bar{h} = 1/\kappa$, one has $y_k^n + \mu_Y(y_k^n)h > 0$. Moreover, as a direct consequence of (4.3.16)–(4.3.17) and of (4.7.94), we have that, if $\mu_Y(y_k^n) > 0$, then $k_d(n,k) = k$, and if $\mu_Y(y_k^n) < 0$, then $k_u(n,k) = k+1$.

Concerning (4.7.95), we obviously assume $y_k^n > 0$, so that $y_{k+1}^{n+1} > 0$. Note that, from (4.3.16),

$$y_k^n + \mu_Y(y_k^n)h > y_{k_u(n,k)-1}^{n+1} \geq y_{k+1}^{n+1} = y_k^n + \frac{\sigma^2}{4}h + \sigma\sqrt{y_k^n h}.$$

Since $\mu_Y(y_k^n) \leq \kappa\theta$, we get

$$\kappa\theta h > \frac{\sigma^2}{4}h + \sigma\sqrt{y_k^n h} > \sigma\sqrt{y_k^n h},$$

from which

$$y_k^n < \Big(\frac{\kappa\theta}{\sigma}\Big)^2 h = \theta_* h.$$

We prove now (4.7.96). First of all observe that, if $y_k^n \leq \theta$, then $\mu_Y(y_k^n) > 0$ and so $k_d(n,k) = k$. Then we have $y_k^n > \theta$ and from (4.3.15) we can assume $y_k^{n+1} > 0$ up to take $h < (2\sqrt{\theta}/\sigma)^2$. Now, by (4.3.17) we get

$$y_k^n + \mu_Y(y_k^n)h < y_{k_d(n,k)+1}^{n+1} \leq y_k^{n+1} = y_k^n + \frac{\sigma^2}{4}h - \sigma\sqrt{y_k^n h},$$

so that

$$\kappa(\theta - y_k^n)h < \frac{\sigma^2}{4}h - \sigma\sqrt{y_k^n h}.$$

This gives $\kappa y_k^n h > \sigma\sqrt{v_k^n h} - \frac{\sigma^2}{4}h + \kappa\theta h$ and, for h small enough, one gets $y_k^n h > \frac{\sigma^2}{4\kappa^2}$.

2. If $y_k^n \leq \theta_* h$, (4.7.96) gives $k_d(n,k) = k$. As regards the up jump, the case $y_{k_u(n,k)}^{n+1} = 0$ is trivial so we consider $y_{k_u(n,k)}^{n+1} > 0$. In order to prove (4.3.18), we consider two possible cases: $k_u(n,k) = k+1$ and $k_u(n,k) \geq k+2$. In the first case, we have

$$y_{k_u(n,k)}^{n+1} - y_k^n = \frac{\sigma^2}{4}h + \sigma\sqrt{y_k^n h} \leq \Big(\frac{\sigma^2}{4} + \sigma\sqrt{\theta_*}\Big)h \leq C_* h,$$

and the statement holds. If instead $k_u(n,k) \geq k+2$, then by (4.3.16) we have

$$y_{k_u(n,k)-1}^{n+1} - y_k^n < \mu_Y(y_k^n)h.$$

We apply the third inequality in (4.7.94) (with n replaced by $n+1$ and $k = k_u(n,k)$) and we get

$$\begin{aligned} 0 \leq y_{k_u(n,k)}^{n+1} - y_k^n &\leq y_{k_u(n,k)-1}^{n+1} + 2\sigma\sqrt{y_{k_u(n,k)-1}^{n+1} h} + \sigma^2 h - y_k^n \\ &\leq \mu_Y(y_k^n)h + 2\sigma\sqrt{(y_k^n + \mu_Y(y_k^n)h)h} + \sigma^2 h \\ &\leq (\kappa\theta + 2\sigma\sqrt{\theta_* + \kappa\theta} + \sigma^2)h \leq C_* h. \end{aligned}$$

3. The statement follows from (4.7.95).

4. Formula (4.3.19) follows from the fact that the sets $K_u(n,k)$ and $K_d(n,k)$ are nonempty. Indeed, if $y_k^n > \theta_* h$ then $k_u = k+1$, so $K_u(n,k) \neq \emptyset$. And if $y_k^n < \theta_* h$,

$$y_{n+1}^{n+1} - y_k^n - \mu_Y(y_k^n)h \geq Y_0 - \theta_* h - \kappa\theta h = Y_0 - (\theta_* + \kappa\theta)h > 0$$

for $h < Y_0/(\theta_* + \kappa\theta)$, which gives $k_u(n,k) < n+1$. Therefore $K_u(n,k) \neq \emptyset$ for every (n,k).

As regards $K_d(n,k)$, if $y_k^n < \theta^*/h$ then $k_d(n,k) = k$ by Proposition 4.3.3, so that $K_d(n,k) \neq \emptyset$. If instead $y_k^n \geq \theta^*/h$, then

$$y_0^{n+1} - y_k^n - \mu_Y(y_k^n)h \leq Y_0 - \frac{\theta^*}{h} - \kappa\theta h + \kappa y_k^n h \leq Y_0 - \frac{\theta^*}{h} + \kappa y_k^n h.$$

Recalling that $h = T/N$, we note that there exists $C > 0$ such that

$$y_k^n h \leq y_N^N h = \Big(\sqrt{Y_0} + \frac{\sigma}{2} N\sqrt{h}\Big)^2 h = \Big(\sqrt{Y_0}\sqrt{\frac{T}{N}} + \frac{\sigma}{2}T\Big)^2 \leq C.$$

Therefore

$$y_0^{n+1} - y_k^n - \mu_Y(y_k^n)h \leq Y_0 - \frac{\theta^*}{h} + \kappa C < 0$$

for $h < \frac{\theta^*}{Y_0 + \kappa C}$. So, $K_d(n,k) \neq \emptyset$.

Now, by (4.3.17) and (4.3.16), since $K_d(n,k) \neq \emptyset$ and $K_u(n,k) \neq \emptyset$,

$$\frac{\mu_Y(y_k^n)h + y_k^n - y_{k_d(n,k)}^{n+1}}{y_{k_u(n,k)}^{n+1} - y_{k_d(n,k)}^{n+1}} \geq 0, \qquad \frac{\mu_Y(y_k^n)h + y_k^n - y_{k_d(n,k)}^{n+1}}{y_{k_u(n,k)}^{n+1} - y_{k_d(n,k)}^{n+1}} = 1 + \frac{\mu_Y(y_k^n)h + y_k^n - y_{k_u(n,k)}^{n+1}}{y_{k_u(n,k)}^{n+1} - y_{k_d(n,k)}^{n+1}} \leq 1.$$

□

4.7.2 Proof of Lemma 4.4.5 and Lemma 4.4.7

We first recall the Poisson summation formula. It is worldwide famous but is usually written on the Schwartz space. We propose here the following version.

Proposition 4.7.1. *If $\varphi \in C^2(\mathbb{R})$ with $\varphi, \varphi', \varphi'' \in L^1(\mathbb{R}, dx)$ then*

$$\sum_{n\in\mathbb{Z}} \varphi(n) = \int_{\mathbb{R}} \varphi(x)dx + \sum_{n\in\mathbb{Z}, n\neq 0} \int_{\mathbb{R}} \varphi(x) e^{-2\pi \mathbf{i} n x} dx. \tag{4.7.97}$$

Proof. For $x \in \mathbb{R}$, let $\lfloor x \rfloor = \sup\{k \in \mathbb{Z} : k \le x\}$ denote the integer part. For $N \in \mathbb{N}$, straightforward computations give

$$\sum_{|n|\le N} \varphi(n) = \frac{1}{2}(\varphi(N)+\varphi(-N)) + \int_{-N}^{N} \varphi(x)dx + \int_{-N}^{N} \Big(x - \lfloor x \rfloor - \frac{1}{2}\Big) g'(x)dx.$$

We recall that $\varphi(\pm N) \to 0$ as $N \to \infty$ (because $\varphi, \varphi' \in L^1(\mathbb{R}, dy)$). Moreover, the Fourier series representation gives

$$x - \lfloor x \rfloor - \frac{1}{2} = \sum_{n\in\mathbb{Z}, n\neq 0} \frac{e^{-2\pi \mathbf{i} n x}}{2\pi \mathbf{i} n}, \quad x \in \mathbb{R}.$$

So,

$$\sum_{n\in\mathbb{Z}} \varphi(n) = \int_{\mathbb{R}} \varphi(x)dx + \int_{\mathbb{R}} \sum_{n\in\mathbb{Z}, n\neq 0} \frac{e^{-2\pi \mathbf{i} n x}}{2\pi \mathbf{i} n} \varphi'(x)dx.$$

Let $\mathfrak{F}[\cdot]$ denote the Fourier transform. Then, $\int_{\mathbb{R}} e^{-2\pi \mathbf{i} n x}\varphi'(x)dx = \mathfrak{F}[\varphi'](2\pi n) = 2\pi \mathbf{i} n \mathfrak{F}[\varphi]\ (2\pi n)$. We also have $|\mathfrak{F}[\varphi'](2\pi n)| \le |\frac{\mathfrak{F}[\varphi''](2\pi n)}{2\pi n}| \le \frac{M}{n}$. Thus, we can put the sum outside the integral and the statement holds.

□

Proof of Lemma 4.4.5l. (i) We apply (4.7.97) to $\varphi(x) = g(x_0 + x\Delta x)$. So,

$$\begin{aligned}\sum_{n\in\mathbb{Z}} f(x_n)\Delta x - \int_{\mathbb{R}} g(x)dx =& \sum_{n\in\mathbb{Z}, n\neq 0} e^{2\pi \mathbf{i} n x_0/\Delta x} \int_{\mathbb{R}} g(x) e^{-2\pi \mathbf{i} n x/\Delta x} dx \\ =& \Delta x^2 \sum_{n\in\mathbb{Z}, n\neq 0} \frac{e^{2\pi \mathbf{i} n x_0/\Delta x}}{(2\pi \mathbf{i} n)^2} \int_{\mathbb{R}} g''(x) e^{-2\pi \mathbf{i} n x/\Delta x} dx,\end{aligned}$$

the latter inequality coming from the integration by parts formula. The statement now follows by recalling that $\sum_{n\ge 1} \frac{1}{n^2} = \frac{\pi^2}{6}$.

(ii) We apply (4.4.52) to the function g^2. Note that if $g, g', g'' \in L^2(\mathbb{R}, dx)$ then g^2 and its derivatives up to order 2 belong to $L^1(\mathbb{R}, dx)$. Moreover, $\int_{\mathbb{R}} g^2(x)dx = |g|_{L^2}^2$ and $|(g^2)''|_{L^1} \le 2|g'|_{L^2}^2 + 2|g|_{L^2}|g''|_{L^2}$, and (4.4.53) immediately follows. □

Proof of Lemma 4.4.7. Hereafter, $C > 0$ denotes a constant which can vary from line to line.

As regard Ψ_1, we recall that $i \mapsto \nu(i\Delta x)\Delta x / \sum_l \nu(l\Delta x)\Delta x$ is a probability measure on $\mathcal{X}$ and $\sum_l \nu(l\Delta x)\Delta x \le c\lambda$. Then,

$$\begin{aligned}|\Psi_1|_2^2 =& \sum_i \Big(\sum_l \nu(l\Delta x)[g(t, x_{i+l}, y) - g(t, x_i, y)]\Delta x\Big)^2 \Delta x \\ \le& 2c\lambda \sum_i \sum_l \nu(l\Delta x)[g^2(t, x_{i+l}, y) + g^2(t, x_i, y)]\Delta x^2 \le 2c^2\lambda^2 |g|_2^2.\end{aligned}$$

By (ii) of Lemma 4.4.5 and (4.4.54), we can write

$$|\Psi_1|_2^2 \leq 2c^2\lambda^2\Big(|g|^2_{L^2(\mathbb{R},dx)} + \frac{\Delta x^2}{6}\left(|\partial_y g|^2_{L^2(\mathbb{R},dx)} + |g|_{L^2(\mathbb{R},dx)} \times |\partial_y^2 g|_{L^2(\mathbb{R},dx)}\right)\Big) \leq C(1+|y|^a)^2.$$

Concerning Ψ_2, by using again (ii) of Lemma 4.4.5 we have

$$\begin{aligned}|\Psi_2|_2^2 &\leq \int_0^1 (1-\tau)^{2\gamma}\Big(\sum_i g^2(t+\tau h, x_i, y)\Delta x\Big)d\tau \\ &\leq \int_0^1 (1-\tau)^{2\gamma}\Big[|g(t+\tau h,\cdot,y)|^2_{L^2(\mathbb{R},dx)} + \frac{\Delta x^2}{6}\Big(|\partial_y g(t+\tau h,\cdot,y)|^2_{L^2(\mathbb{R},dx)} \\ &\quad + |g(t+\tau h,\cdot,y)|^2_{L^2(R)} \times |\partial_y^2 g(t+\tau h,\cdot,y)|^2_{L^2(\mathbb{R},dx)}\Big)\Big]d\tau \leq C(1+|y|^a)^2.\end{aligned}$$

For Ψ_3 and Ψ_4 the assertion follows in a similar way. Finally, again from (ii) of Lemma (4.4.5),

$$|\Psi_5|_2^2 \leq |\Psi_5|^2_{L^2(\mathbb{R},dx)} + \frac{\Delta x^2}{6}\left(|\Psi_5'|^2_{L^2(\mathbb{R},dx)} + |\Psi_5|_{L^2(\mathbb{R},dx)} \times |\Psi_5''|_{L^2(\mathbb{R},dx)}\right). \tag{4.7.98}$$

Now, by (i) of Lemma 4.4.5,

$$\begin{aligned}|\Psi_5|^2_{L^2(\mathbb{R},dx)} &= \int \Big|\int g(t,\zeta+x,y)\nu(x)dx - \sum_l g(t,\zeta+l\Delta x,y)\nu(l\Delta x)\Delta x\Big|^2 d\zeta \\ &\leq \frac{\Delta x^4}{144}\int\Big(\int\big|\partial_y^2(g(t,\zeta+x,y)\nu(x))\big|dx\Big)^2 d\zeta \\ &\leq \frac{\Delta x^4}{36}\int d\zeta\int\Big(|\partial_y^2 g(t,\zeta+x,y)|^2 + |\partial_y g(t,\zeta+x,y)|^2\Big|\frac{\nu'(x)}{\nu(x)}\Big|^2 + |g(t,\zeta+x,y)|^2\Big|\frac{\nu''(x)}{\nu(x)}\Big|^2\Big)\nu(x)dx \\ &= \frac{\Delta x^4}{36}\int \nu(x)dx\int\Big(|\partial_y^2 g(t,\zeta+y,y)|^2 + |\partial_y g(t,\zeta+y,y)|^2\Big|\frac{\nu'(x)}{\nu(x)}\Big|^2 + |g(t,\zeta+y,y)|^2\Big|\frac{\nu''(x)}{\nu(x)}\Big|^2\Big)d\zeta \\ &= \frac{\Delta x^4}{36}\Big(|\partial_y^2 g(t,\cdot,y)|^2_{L^2(\mathbb{R},dx)}|\nu| + |\partial_y g(t,\cdot,y)|^2_{L^2(\mathbb{R},dx)}\Big|\frac{\nu'}{\nu}\Big|^2_{L^2(\mathbb{R},d\nu)} + |g(t,\cdot,y)|^2_{L^2(\mathbb{R},dx)}\Big|\frac{\nu''}{\nu}\Big|^2_{L^2(\mathbb{R},d\nu)}\Big) \\ &\leq C\lambda\Delta x^4(1+|y|^a)^2,\end{aligned}$$

last inequality following from (4.4.54) and (4.4.55). Similar calculations allow one to bound the terms $|\Psi_5'|_{L^2(\mathbb{R},dx)}$ and $|\Psi_5''|_{L^2(\mathbb{R},dx)}$ in (4.7.98). □

Acknowledgments

This thesis was written at the Dipartimento di Matematica of the Università di Roma Tor Vergata and at the Laboratoire d' Analyse et des Mathématiques Appliquées of the Université Paris-Est Marne-la-Valle. I had two advisors, two laboratories, two places to call *home*.

I cannot find the words to say how much last three years have meant to me, both from a professional and a personal point of view. What is for sure is that I would never been here without the help and the support of some people. Therefore, I want to dedicate this last part of my thesis, maybe the most difficult to write, to say them thanks.

My first, heartfelt, thanks go to my advisors, Lucia Caramellino and Damien Lamberton.

Grazie Lucia, for having supported me in all these years, for having me taught much more than maths, for your friendship. Merci Damien, because if now I know some French I owe it to you, for your kindness and irony, for always welcoming me with a smile. Thank you both for having me taught more or less everything I can do now, for your necessary help, for all the time you have dedicated to me with infinite patience.

I heartily thank Maya Briani, who is like a third advisor to me: thank you Maya for your kindness, your advice and for all the time we have spent together discussing about maths and talking as good friends do. I am also grateful to Antonino Zanette for having given me the opportunity to work with him.

I would like to express my gratitude to Prof. A. Neuenkirch and to Prof. A. Pascucci, for accepting to referee my work, and to Prof. A. Alfonsi and Prof. F. Antonelli, for accepting to be part of the committee.

I was lucky enough to find warm and stimulating environments in both my laboratories. Many thanks to the Probability groups both in Tor Vergata and in UPEM. In particular, thank to Prof. V. Bally and to Prof. P. Baldi, for his help during my stays in France and for all the discussions we had about Paris and maths.

I am sincerely grateful to the coordinator of the Ph.D. School in Rome, Prof. A. Braides, and to Simonetta De Nicola, Christiane Lafargue, Audrey Patout and Sylvie Cach: your availability and patience were precious during all these years.

I met great people and I had fantastic Ph.D. colleagues in both my universities. In France, I wish to especially thank Huong and Yushun, my first friends in UPEM. As regards Tor Vergata, I have a long list of friends to thank. First of all, thanks to the members of my beloved 1225, especially to Duccio, Rossana and Gianluca, who have been by my side from the beginning. I also wish to thank the other Ph.D. students and researchers (mathematicians, of course, but computer scientists too!) with whom I have shared fruitful discussions, lunches, laughs and beer during all these years. Among all, a special thank goes to Michele: I could not have a better colleague, flatmate, friend. Thanks, a lot.

It's the turn to thank my friends, who have been standing close to me during all these years.

Thanks to all the fantastic people I met in Paris, with whom I have shared unforgettable moments. In particular, thanks to Davide, Greta, Graziano, Matteo, Michela, Valeria and, of course, to Danilo and Vanessa, best Roman presents Paris could have given me.

Thanks to Anna, Anna Paola, Martina and Maurizia, my favourite, amazing, women in maths: your support helped me in many situations, thanks! Thanks also to my old university colleagues, who are now dear friends of mine.

Thanks to Federica and to my *Sardinia girls*, who were very close to me in the last year, and to my neighbour and friend Giulia.

Finally, my heartfelt thanks to Beatrice, Isabella, Liliana, Manlio, Michela, Michele, Simone and Simone: you have been my certainty in all these years, and I feel so lucky for this. A very special thank to Beatrice who always believed in me, even when I did not do it.

L'ultimo e più grande ringraziamento va a tutta la mia famiglia, in particolare ai miei genitori e ai miei fratelli Marco e Simone. Siete e sarete sempre il mio punto di riferimento più grande.

Bibliography

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

[2] E. Akyildirim, Y. Dolinsky, H.M. Soner (2014): Approximating stochastic volatility by recombinant trees. *Ann. Appl. Probab.* 24, 2176–2205.

[3] A. Alfonsi (2005): On the discretization schemes for the CIR (and Bessel squared) processes. *Monte Carlo Methods Appl.* 11, 355–467.

[4] A. Alfonsi (2010): High order discretization schemes for the CIR process: Application to affine term structure and Heston models. *Math. Comp.* 79, 209–237.

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

[6] M. Altmayer, A. Neuenkirch (2017): Discretising the Heston model: an analysis of the weak convergence rate. *IMA J. Numer. Anal.* 37, 1930–1960.

[7] L. Andersen (2006): Efficient Simulation of the Heston Stochastic Yolatility Model. Preprint available at `http://www.ressources-actuarielles.net/`.

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

[9] L.B.G. Andersen, Y.Y. Piterbarg (2007): Moment explosions in stochastic volatility models. *Finance Stoch.*, 11, 29-50.

[10] E. Appolloni, L. Caramellino, A. Zanette (2015): A robust tree method for pricing American options with CIR stochastic interest rate. *IMA J. Manag. Math.*, 26, 345-375.

[11] S. Assing, S.D. Jacka, A. Ocejo (2014): Monotonicity of the value function for a two-dimensional optimal stopping problem. *Ann. Appl. Probab.* 24(4), 1554-1584.

[12] K. Bahlali, B. Mezerdi, Y. Ouknine (1686): Pathwise uniqueness and approximation of solutions of stochastic differential equations. Séminaire de Probabilités, XXXII, Lecture Notes in Math., 1686, Springer, Berlin, 166-187.

[13] V. Bally, G. Pagès (2003): Error analysis of the optimal quantization algorithm for obstacle problems. *Stoch. Processes App.* 106, 1–40.

[14] V. Bally, C. Rey (2016): Approximation of Markov semigroups in total variation distance. *Electron. J. Probab.* 21, no. 12, 44 pp.

[15] G. Barles (1997): Convergence of Numerical Schemes for Degenerate Parabolic Equations. *Arising in Finance Theory. In L. Rogers & D. Talay (Eds.), Numerical Methods in Finance (Publications of the Newton Institute, pp. 1-21).* Cambridge University Press.

[16] V. Bally, D. Talay (1996): The law of the Euler scheme for stochastic differential equations. I. Convergence rate of the distribution function. *Probab. Theory Relat. Fields* 104, 43–60.

[17] D.S. Bates (1996): Jumps and stochastic volatility: exchange rate processes implicit in Deutsch mark options. *Rev. Fin.* 9, 69–107

[18] J. Bather (1970): Optimal stopping problems for Brownian motion. *Adv. in Appl. Probab.*, 2, 259–286.

[19] A. Bensoussan, J.L. Lions (1982): *Applications of variational inequalities in stochastic control*, Studies in Mathematics and its Applications, 12, North-Holland Publishing Co., Amsterdam-New York. Translated from the French.

[20] F. Black (1988): The holes in Black-Scholes. *Risk* 1(4), 30–33.

[21] F. Black, M. Scholes (1973): The pricing of options and corporate liabilities. *J. Polit. Econ.* 81, 637-654.

[22] L. Bergomi (2016): *Stochastic volatility modeling.* Chapman & Hall/CRC Financial Mathematics Series, CRC Press, Boca Raton, FL.

[23] M. Bossy, H. Olivero (2018) Strong convergence of the symmetrized Milstein scheme for some CEV-like SDEs. *Bernoulli*, 24, 1995–2042.

[24] M. Briani, L. Caramellino, A. Zanette (2017): A hybrid approach for the implementation of the Heston model. *IMA J. Manag. Math.*, 28, 467–500.

[25] M. Briani, L. Caramellino, A. Zanette (2016): A hybrid tree/finite-difference approach for Heston-Hull-White type models. *J. Comput. Finance*, 21, 1–45.

[26] M. Briani, L. Caramellino, G. Terenzi (2018): Convergence rate od Markov chains and hybrid numerical schemes to jump-diffusions with application to the Bates model. Preprint, `arXiv:1809.10545`.

[27] M. Briani, L. Caramellino, G. Terenzi, A. Zanette (2017): On a hybrid method using trees and finite-difference for pricing options in complex models. Preprint, `ArXiv:1603.07225`.

[28] M. Briani, C. La Chioma, R. Natalini (2004): Convergence of numerical schemes for viscosity solutions to integro-differential degenerate parabolic problems arising in financial theory. *Numer. Math.*, 98(4), 607–646.

[29] M. Briani, R. Natalini, G. Russo (2007): Implicit-Explicit Numerical Schemes for Jump-Diffusion Processes. *Calcolo*, 44, 33-57.

[30] D. Brigo, F. Mercurio (2006): *Interest Rate Models - Theory and Practice.* Springer, Berlin.

[31] L. Brugnano, D. Trigiante (1992): Tridiagonal matrices: Invertibility and conditioning, *Linear Algebra Appl.*, **166**, 131-150.

[32] A. Canale, R.M. Mininni, A. Rhandi (2017): Analytic approach to solve a degenerate parabolic PDE for the Heston model. *Math. Methods Appl. Sci.*, 40(13), 4982–4992.

[33] P. Carr, D. Madan (1999): Option valuation using the Fast Fourier Transform. *J. Comput. Finance*, 3, 463-520.

[34] C. Chiarella, B. Kang, G. Meyer, A. Ziogas (2009): The evaluation of American option prices under stochastic volatility and jump-diffusion dynamics using the method of lines. *Int. J. Theor. Appl. Finan.*, 12, 393.

[35] C. Chiarella, B. Kang, G.H. Meyer (2012): The evaluation of barrier option prices under stochastic volatility. *Comput. Math. Appl.*, **64**, 2034-2048.

[36] M. Costabile, M. Gaudenzi, I. Massabò, A. Zanette (2009) Evaluating fair premiums of equity-linked policies with surrender option in a bivariate model. *Insurance Math. Econom.* 45, 286–295.

[37] J. C. Cox (1975): Notes on option pricing I: constant elasticity of variance diffusion. *Working paper*, Stanford University, Stanford CA.

[38] J. C. Cox, J. Ingersoll, S. Ross (1985): A theory of the term structure of interest rates, *Econometrica*, **53**, 385-407.

[39] R. Cont, E. Voltchkova (2005): A finite difference scheme for option pricing in jump-diffusion and exponential Lévy models. *SIAM J. Numer. Anal.*, 43(4), 1596–1626.

[40] V. D'Halluin, P.A. Forsyth, G. Labahn (2005): A semi-Lagrangian Approach for American Asian options under jump-diffusion, *SIAM J. Sci. Comp.* **27**, 315-345.

[41] C. Dellacherie, P.A. Meyer (1975): *Probabilités et potentiel,* vol. IV. Hermann, Paris.

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

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

[44] D. Duffie, J. Pan, K. Singleton (2000) Transform analysis and asset pricing for affine jump-diffusions. *Econometrica* 68, 1343–1376.

[45] D.J. Duffy (2006): *Finite difference methods in financial engineering. A partial differential equation approach.* Wiley Finance Series.

[46] B. Dupire (1997): Pricing and hedging with smiles. *Mathematics of derivative securities* (Cambridge, 1995), vol. 15 of Publ. Newton Inst. Cam- bridge Univ. Press, Cambridge, 103–111.

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

[48] 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.

[49] 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.

[50] W. Feller (1951): Two singular diffusion problems. *Ann. of Math.* (2), 54: 173–182.

[51] P. Foschi, A. Pascucci (2008): Path dependent volatility. *Decis. Econ. Finance 31*, 1 , 13-32.

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

[53] M.G. Garroni, J.L. Menaldi (1993) *Green Functions for Second Order Parabolic Integro-Differential Problems.* Pitman Research Notes in Mathematics Series, 275.

[54] J. Gatheral, T. Jaisson, M. Rosenbaum (2014): Volatility is rough. Preprint, *arXiv:1410.3394*.

[55] A.L. Grzelak, C.W. Oosterlee (2011): On the Heston model with stochastic interest rates. *SIAM J. Fin. Math.* 2, 255-286.

[56] T. Haentjens, K.J. in't Hout (2012): Alternating direction implicit finite difference schemes for the Heston-Hull-White partial differential equation. *J. Comp. Finan.* 16, 83–110.

[57] P. Hagan, D. Kumar, A. Lesniewski, D. Woodward (2014): Arbitrage free sabr. *Wilmott*, (69): 60-75.

[58] S. L. Heston (1993): A Closed-Form Solution for Options with Stochastic Volatility with Applications to Bond and Currency Options, *Rev. Financial Stud.*, 6, 327-3q43.

[59] J. E. Hilliard, A. L. Schwartz, A.L. Tucker (1996): Bivariate binomial pricing with generalized interest rate processes. *J. Financ. Res.* XIX-4, 585–602.

[60] D. G. Hobson, L. C. G. Rogers (1998): Complete models with stochastic volatility. *Math. Finance 8*, 1 , 27-48.

[61] J. Hull, A. White (1987): The pricing of options on assets with stochastic volatilities. *J. Finance*, 42, 281–300.

[62] J. Hull, A. White (1994): Numerical procedures for implementing term structure models I. *Journal of Derivatives* 2(1), 7-16.

[63] N. Ikeda, S. Watanabe (1981): *Stochastic Differential Equations and Diffusion Processes.* North Holland Publ. Co., Amsterdam -Oxford -New York.

[64] A. Itkin (2016): Efficient Solution of Backward Jump-Diffusion PIDEs with Splitting and Matrix Exponentials. *J. Comput. Finance*, 19, 29-70.

[65] S. D. Jacka (1993): Local times, optimal stopping and semimartingales. *Ann. Appl. Probab.*, 21(1), 329-339.

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

[67] M. Keller Ressel (2011): Moment explosions and long-term behavior of affine stochastic volatility models. *Math. Finance*,21, 23-98.

[68] D. Kinderlehrer, G. Stampacchia (1980): *An introduction to variational inequalities and their applications.* Volume 31 of *Classic in Applied Mathematics.* Society for Industrial and Applied Mathematics (SIAM), Philadelphia.

[69] S.G. Kou (2002): A Jump-Diffusion Model for Option Pricing. *Management Science*, 48, 1086-1101.

[70] O. A. Ladyženskaja, V.A. Solonnikov, N.N. Ural'ceva (1968): *Linear and quasilinear equations of parabolic type.* Translated from the Russian by S. Smith. Translations of Mathematical Monographs, Vo. 23,. American Mathematical Society, Providence, R.I.

[71] D. Lamberton (1998): Error estimates for the binomial approximation of American put options. *Ann. Appl. Probab.*, 8(1), 206-233.

[72] D. Lamberton, B. Lapeyre (2008): Introduction to stochastic calculus applied to finance. Second edition. Chapman & Hall/CRC Financial Mathematics Series.

[73] D. Lamberton, G. Terenzi (2018): Variational formulation of American option prices in the Heston model. *SIAM J. Financial Math.*, to appear.

[74] D. Lamberton, G. Terenzi (2018): American option price properties in Heston-type models. Working paper.

[75] P.D. Lax, R.D. Richtmyer (1956): Survey of the stability of linear finite difference equations. *Commun. Pure Appl. Math.*, 9, 267–293.

[76] F.A. Longstaff, E.S. Schwartz (2001): Valuing American options by simulations: a simple least squares approach. *Rev. Financ. Stud.*, 14, 113-148.

[77] R.C. Merton (1976): Option pricing when underlying stock returns are discontinuous. *J. Financial Econom.*, 3, 125–144.

[78] R. Mikulevicius, H. Pragarauskas (2004): On Cauchy-Dirichlet problem in half-space for linear integro-differential equations in weighted Hölder spaces. *Electron. J. Probab.*, 10, 1398–1416.

[79] D.B. Nelson, K. Ramaswamy (1990): Simple binomial processes as diffusion approximations in financial models. *Rev. Financ. Stud.*, 3, 393-430.

[80] H. Nieuwenhuis, M. Vellekoop (2009): A tree-based method to price American Options in the Heston Model. *J. Comput. Finance* 13, 1–21.

[81] S. M. Ould Aly (2013): Monotonicity of prices in Heston model. *Int. J. Theor. Appl. Finance* 16(3), 1350016, 23 pp.

[82] G. Pagès, J. Printems (2005): Functional quantization for numerics with an application to option pricing. *Monte Carlo Methods Appl.* 11, 407–446.

[83] G. Peskir, A. Shiryaev (2006): *Optimal Stopping and Free-Boundary Problem.* Lectures in Mathematics, ETH Zurich. Birkhauser.

[84] Premia: An Option Pricer. `http://www.premia.fr`

[85] D. Revuz, M. Yor (1994): *Continuous martingales and Brownian motion.* volume 293 of Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences.] Springer-Verlag, Berlin, second edition.

[86] The rough volatility network: `https://sites.google.com/site/roughvol/home`

[87] S. Salmi, J. Toivanen (2014): IMEX schemes for pricing options under jump-diffusion models. *Appl. Numer. Math.*, 84, 33-45.

[88] D. W. Stroock, S.R.S. Varadhan (1972): On the support of diffusion processes with applications to the strong maximal principle. *Proc. of Sixth Berkeley Symp. Math. Statist. Prob.*, 333-359, Univ. California Press, Berkeley.

[89] D. W. Stroock, S.R.S. Varadhan (1979): *Multidimensional Diffusion Processes.* Springer, Berlin.

[90] E. Stein, J. Stein (1991): Stock price distributions with stochastic volatility: an analytic approach. *Rev. Financ. Stud.*, 4, 727–752.

[91] Y. Tian (1994): A reexamination of lattice procedures for interest rate-contingent claims. *Adv. Futures Options Res.* 7, 87–110.

[92] J. Toivanen (2010): A Componentwise Splitting Method for Pricing American Options Under the Bates Model. Applied and numerical partial differential equations, Vol. 15 of *Comput. Methods Appl. Sci.*, Springer, New York, 213–227.

[93] N. Touzi (1999): American options exercise boundary when the volatility changes randomly. *Appl. Math. Optim.* , 39(3), 411-422.

[94] M. Vellekoop, H. Nieuwenhuis (2009): A tree-based method to price American Options in the Heston Model. *J. Comput. Finance*, 13, 1–21.

[95] S. Villeneuve (1999): Exercise Regions of American Options on Several Assets. *Finance Stoch.*, 3(3), 295-322.

[96] E. Voltchkova, P. Tankov (2008): Deterministic methods for option pricing in exponential Lévy models. PREMIA documentation. Available online at: *http://www.premia.fr*

[97] J. Wei (1996): Valuing American equity options with a stochastic interest rate: a note. *J. Financ. Eng.* 2, 195–206.

[98] C. Zheng (2017): Weak convergence rate of a time-discrete scheme for the Heston stochastic volatility model. *SIAM J. Numer. Anal.* 55, 1243–1263.