Convergence rate of Markov chains and hybrid numerical schemes to jump-diffusions with application to the Bates model

MAYA BRIANI*
LUCIA CARAMELLINIO†
GIULIA TERENZI‡

Abstract

We study the rate of weak convergence of Markov chains to diffusion processes under suitable but quite general assumptions. We give an example in the financial framework, applying the convergence analysis to a multiple jumps tree approximation of the CIR process. Then, we combine the Markov chain approach with other numerical techniques in order to handle the different components in jump-diffusion coupled models. We study the speed of convergence of this hybrid approach and we provide an example in finance, applying our results to a tree-finite difference approximation in the Heston or Bates model.

Keywords: jump-diffusion processes; weak convergence; tree methods; finite-difference; stochastic volatility; European options.

2000 MSC: 60H35, 65C20, 91G60.

Contents

1 Introduction ................................................................. 2
2 Notation ........................................................................... 3
3 First order weak convergence of Markov chains to diffusions ........ 4
   3.1 An example: a first order weak convergent binomial tree for the CIR process . . . . 7
4 Hybrid schemes for jump-diffusions and convergence rate ........... 11
   4.1 The hybrid procedure ...................................................... 12
   4.2 Convergence speed of the scheme (4.10) ................................ 13
   4.3 An application: finite difference schemes ............................. 14
      4.3.1 Convergence in $l_2$-norm ........................................... 15
      4.3.2 Convergence in $l_\infty$-norm ........................................ 20

---

*Istituto per le Applicazioni del Calcolo, CNR Roma - m.briani@iac.cnr.it
†Dipartimento di Matematica, Università di Roma Tor Vergata, and INDAM-GNAMPA - caramell@mat.uniroma2.it
‡Dipartimento di Matematica, Università di Roma Tor Vergata and Université Paris-Est, Laboratoire d’Analyse et de Mathématiques Appliqués (UMR 8050), UPEM, UPEC, CNRS, Projet Mathrisk INRIA, F-77454, Marne-la-Vallée, France - terenzi@mat.uniroma2.it
The paper is devoted to the study of the weak convergence rate of numerical schemes allowing one to handle specific jump-diffusion processes. These include the well known stochastic volatility models by Heston [22] and by Bates [10]. Since these dynamics involve the square root process for the volatility, a special numerical treatment has to be considered. When dealing with European options, i.e. solutions to Partial (Integral) Differential Equation (hereafter P(I)DE) problems, numerical approaches involve tree methods [1, 31], Monte Carlo procedures [2, 3, 4, 5, 37], finite-difference numerical schemes [16, 24, 35] or quantization algorithms [32]. 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. We consider a numerical procedure which combines a tree method for the volatility process with a different numerical approach for the asset price process, for instance finite-difference. Such a hybrid method has been developed and numerically studied in [12, 13, 14] for the computation of European and American options in the stochastic volatility context. In this paper we study the rate of convergence. As a result, we can consider the Heston or the Bates model in the full parameter regime, differently from many other approaches. 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 [3, 4, 11, 37], all 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 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

\[
\frac{dY_t}{Y_t} = \mu_Y(Y_t)dt + \sigma_Y(Y_t)dB_t.
\]

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. [23]). 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 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 [6] for the CIR process. Recall that the CIR process [18] is a square root process, that is,

\[
\frac{dY_t}{\sqrt{Y_t}} = \kappa(\theta - Y_t)dt + \sigma \sqrt{Y_t}dB_t,
\]

with \(\kappa, \theta, \sigma > 0\). Recall also that this process lives in \([0, +\infty)\) and under the Feller condition \(2\kappa \theta > \sigma^2\) it never hits 0. Several trees are considered in the literature, see e.g. [17, 23, 34], but generally some numerical problems arise when the Feller condition fails. Our result for the tree in [6] (Theorem 3.2) works in any parameter regime. Recall that in equity markets, one often requires large values for the vol-vol \(\sigma\) whereas in interest rates context, \(\sigma\) is markedly lower (see e.g. the calibration results in [19] and in [15] p. 115, respectively). So, a result in the full parameter regime is actually essential.
Let us mention that our general convergence Theorem 3.1 may in principle be applied to more general trees constructed through the multiple jumps approach by Nelson and Ramaswamy [30], on which the tree in [6] 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 which the tree in [6] is based – to our knowledge, a theoretical study of the rate of convergence for general trees constructed through the multiple jumps approach by Nelson and Ramaswamy [30], on which the tree in [6] is based.

In the second part (Section 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 whose coefficients only depend on the process (Y_t)_{t\in[0,T]}:

\[ dX_t = \mu_X(Y_t)dt + \sigma_X(Y_t)dW_t + \gamma_X(Y_t)dH_t, \]

where \( H \) is a compound Poisson process independent of the 2-dimensional Brownian motion \( (W, B) \). 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 \( \mathbb{E}[f(X_T, Y_T)] \) through a generalization of the hybrid method introduced in [12], [13], [14] (Section 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 4.2), in Theorem 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 paper. 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 to be said 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 [9] or the recent [4] for the Heston model) or on a noise having at least a “good piece of absolutely continuous part” (Doebllin’s condition, see [8]).

We then consider two possible finite-difference schemes (Section 4.3) to handle the (local) PIDE related to the component \( X \): an implicit in time/centered in space scheme (Section 4.3.1) and an implicit in time/upwind in space scheme (Section 4.3.2). 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 4.1 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 (Section 5).

**Acknowledgment.** LC acknowledges the MIUR Excellence Department Project awarded to the Department of Mathematics, University of Rome Tor Vergata, CUP E83C180001 00006.

## 2 Notation

In this section we establish the notation which will be used later on. Let \( d \in \mathbb{N}^* = \mathbb{N} \setminus \{0\} \).

- For a multi-index \( l = (l_1, \ldots, l_d) \in \mathbb{N}^d \) we define \( |l| = \sum_{j=1}^d l_j \) and for \( y \in \mathbb{R}^d \), we define \( \partial^l_y = \partial_{y_{l_1}} \cdots \partial_{y_{l_d}} \) 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.

- \( L^p(\mathbb{R}^d, \mathfrak{m}) \) 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 \( \sigma \)-algebra on \( \mathbb{R}^d \), and we set \( \cdot \mid_{L^p(\mathbb{R}^d, \mathfrak{m})} \) the associated norm. The Lebesgue measure is denoted through \( dx \).
• Let \( D \subseteq \mathbb{R}^d \) be a domain (possibly closed) and \( q \in \mathbb{N} \). \( C^q(D) \) is the set of all functions on \( D \) which are \( q \)-times continuously differentiable. We set \( C^q_{\text{pol}}(D) \) the set of functions \( g \in C^q(D) \) such that there exist \( C, a > 0 \) for which

\[
|\partial_y^l g(y)| \leq C(1 + |y|^a), \quad y \in D, \quad |l| \leq q.
\]

We set \( C^q_{\text{pol},T}(D) \) the set of functions \( v \in C^{[q/2].q}([0, T] \times D) \) such that there exist \( C, a > 0 \) for which

\[
\sup_{t < T} |\partial_t^k \partial_y^l v(t, y)| \leq C(1 + |y|^a), \quad y \in D, \quad 2k + |l| \leq q.
\]

For brevity, we set \( C(D) = C^0(D) \), \( C_{\text{pol}}(D) = C^0_{\text{pol}}(D) \) and \( C_{\text{pol},T}(D) = C^0_{\text{pol},T}(D) \). We also need another functional space, that we call \( C^{p,q}(D) \), \( p \in [1, \infty] \), \( q \in \mathbb{N} \), \( m \in \mathbb{N}^* \): \( g = g(x, y) \in C^{p,q}(D) \) if \( g \in C^{p}_{\text{pol}}(\mathbb{R}^m \times D) \) and there exist \( C, a > 0 \) such that

\[
|\partial_x^k \partial_y^l g(x, y)|_{L^p(\mathbb{R}^m, dx)} \leq C(1 + |y|^a), \quad |l'| + |l| \leq q.
\]

Similarly as above, we set \( C^{p,q}_{\text{pol},T}(D) \) the set of the function \( v \in C^{p,q}_{\text{pol},T}(\mathbb{R}^m \times D) \) such that

\[
\sup_{t < T} |\partial_t^k \partial_x^l \partial_y^p v(t, x, y)|_{L^p(\mathbb{R}^m, dx)} \leq C(1 + |y|^a), \quad 2k + |l'| + |l| \leq q.
\]

• 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 \( l_p(\mathcal{X}) \) as the discrete \( l_p \)-space of the functions \( \varphi : \mathcal{X} \rightarrow \mathbb{R} \) 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 \( \Gamma : l_p(\mathcal{X}) \rightarrow l_p(\mathcal{X}) \), the induced norm is denoted by \( |\Gamma|_p = \sup_{|\varphi|_p \leq 1} |\Gamma \varphi|_p \). And for a function \( g : \mathbb{R}^d \rightarrow \mathbb{R} \), we set \( |g|_{l_p} \) 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 \( L^p(\Omega) \).

### 3 First order weak convergence of Markov chains to diffusions

Let \( d \in \mathbb{N}^* \) and \( D \subseteq \mathbb{R}^d \) 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

\[
dY_t = \mu_Y(Y_t) dt + \sigma_Y(Y_t) dW_t, \quad Y_0 \in D, \quad (3.1)
\]

where \( W \) is a \( \ell \)-dimensional standard Brownian motion. From now on, we set \( a_Y = \sigma_Y \sigma_Y^* \), the notation \( \star \) denoting transpose. We recall that the associated infinitesimal generator is given by

\[
\mathcal{A} = \frac{1}{2} \text{Tr}(a_Y D_y^2) + \mu_Y \cdot \nabla_y, \quad (3.2)
\]

where \( \text{Tr} \) denotes the matrix trace, \( D_y^2 \) and \( \nabla_y \) 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 \( T > 0 \), \( f : D \rightarrow \mathbb{R} \) and we define

\[
u(t, y) = \mathbb{E}[f(Y_T^{t,y})], \quad (t, y) \in [0, T] \times D, \quad (3.3)
\]

where \( Y_T^{t,y} \) denotes the solution to the SDE in \((3.1)\) 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) $\mu_Y$ 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 (3.1) such that $\mathbb{P}(\forall s \in [t, T], Y_{s,t}^{y} \in \mathcal{D}) = 1$;
(c) the function $u$ in (3.3) solves the PDE
\[
\begin{aligned}
\frac{\partial u}{\partial t} + A u &= 0, \quad \text{in } [0, T) \times \mathcal{D}, \\
u(T, y) &= f(y), \quad \text{in } \mathcal{D}.
\end{aligned}
\] (3.4)

The above properties (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 [33]. We will see in Section 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 $(Y_n^h)_{n=0,\ldots,N}$ denote a Markov chain, whose state space, at time-step $n$, is given by $Y_n^h \subset \mathcal{D}$. In our mind, $(Y_n^h)_{n=0,\ldots,N}$ 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 $n = 0, \ldots, N$. Of course, we assume that $Y_0^h = Y_0$, that is, $\mathbb{Y}_0^h = \{Y_0\}$. Without loss of generality, we may assume that $(Y_n^h)_{n=0,\ldots,N}$ is defined in $(\Omega, \mathcal{F}, \mathbb{P})$.

In order to study the rate of the weak convergence of $(Y_n^h)_{n=0,\ldots,N}$ to $Y$, we need to stress the requests that are usually done in order to merely prove the convergence (see e.g., [33]). In particular, we need the following assumption.

**Assumption $\mathcal{H}_1$.** There exists $\bar{h} > 0$ such that, for every $h < \bar{h}$, the first three local moments satisfy
\[
\begin{align*}
\mathbb{E}(Y_{n+1}^h - Y_n^h \mid Y_n^h) &= \mu_Y(Y_n^h) h + f_h(Y_n^h), \\
\mathbb{E}((Y_{n+1}^h - Y_n^h)(Y_{n+1}^h - Y_n^h)^* \mid Y_n^h) &= a_Y(Y_n^h) h + g_h(Y_n^h), \\
\mathbb{E}(Y_{n+1}^h - Y_n^h)^l \mid Y_n^h) &= j_{h,l}(Y_n^h), \quad l \in \mathbb{N}^d, |l| = 3,
\end{align*}
\]
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
\[
\begin{align*}
\sup_{h \leq \bar{h}} \sup_{n=0,\ldots,N} \|f_h(Y_n^h)\|_p &\leq Ch^2, \\
\sup_{h \leq \bar{h}} \sup_{n=0,\ldots,N} \|g_h(Y_n^h)\|_p &\leq C h^2, \\
\sup_{h \leq \bar{h}} \sup_{n=0,\ldots,N} \|j_{h,l}(Y_n^h)\|_p &\leq C h^2, \quad |l| = 3.
\end{align*}
\] (3.8) (3.9) (3.10)

We also need the following behavior of the moments.

**Assumption $\mathcal{H}_2$.** There exists $\bar{h} > 0$ such that for every $p > 1$ there exists $C_p > 0$ for which
\[
\begin{align*}
\sup_{h \leq \bar{h}} \sup_{0 \leq n \leq N} \|Y_n^h\|_p &\leq C_p, \\
\sup_{h \leq \bar{h}} \sup_{0 \leq n \leq N} \frac{1}{\sqrt{h}} \|Y_{n+1}^h - Y_n^h\|_p &\leq C_p.
\end{align*}
\] (3.11) (3.12)
We can now state the following first order weak convergence result.

**Theorem 3.1.** Let assumptions \( \mathcal{H}_1 \) and \( \mathcal{H}_2 \) hold and assume that \( u \in C^4_{\text{pol,T}}(D) \), \( u \) being defined in (3.8). Then there exist \( \bar{h} > 0 \) and \( C > 0 \) such that for every \( h < \bar{h} \) one has

\[
|\mathbb{E}[f(Y_{n}^h)] - \mathbb{E}[f(Y_T)]| \leq CTh.
\]

**Proof.** The proof is quite standard. Since \( \mathbb{E}[f(Y_{n}^h)] = \mathbb{E}[u(T, Y_T^h)] \) and \( \mathbb{E}[f(Y_T)] = u(0, Y_0) \), we have

\[
\mathbb{E}[f(Y_{n}^h)] - \mathbb{E}[f(Y_T)] = \mathbb{E}[u(T, Y_T^h) - u(0, Y_0)] = \sum_{n=0}^{N-1} \mathbb{E}[u((n+1)h, Y_{n+1}^h) - u(nh, Y_n^h)].
\]

Since \( u \in C^4_{\text{pol,T}}(D) \), 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 \( Y_n^h \) up to order 3 and 1 respectively. We obtain

\[
u((n+1)h, Y_{n+1}^h) = \sum_{0 \leq |l| + 2l' \leq 3} \partial_{\bar{y}}^l \partial_{\bar{y}}^{l'} u(nh, Y_n^h) \frac{h^l (Y_{n+1}^h - Y_n^h)^l}{|l|!l'|} + R_1(n, h, Y_{n+1}^h),
\]

where the remaining term \( R_1 \) is given by

\[
R_1(n, h, Y_{n+1}^h) = h^2 \int_0^1 (1 - \tau) \partial_{\bar{y}}^2 u(t + \tau h, Y_{n+1}^h) dt + h \sum_{|k|=2} (Y_{n+1}^h - Y_n^h)^k \int_0^1 (1 - \xi) \partial_{\bar{y}}^k \partial_t u(nh, Y_n^h + \xi(Y_{n+1}^h - Y_n^h)) d\xi + \sum_{|k|=4} \frac{(Y_{n+1}^h - Y_n^h)^k}{3!} \int_0^1 (1 - \xi)^3 \partial_{\bar{y}}^k u(nh, Y_n^h + \xi(Y_{n+1}^h - Y_n^h)) d\xi.
\]

We now pass to the conditional expectation w.r.t. \( Y_n^h \) in (3.13) and use (3.5) and (3.6). By rearranging the terms we obtain

\[
\mathbb{E}[u((n+1)h, Y_{n+1}^h) - u(nh, Y_n^h)] = h \mathbb{E} \left[ \partial_t u(nh, Y_n^h) + \mu_Y(Y_n^h) \cdot \nabla_{\bar{y}} u(nh, Y_n^h) + \frac{1}{2} \text{Tr}(a_Y D_{\bar{y}}^2 u(nh, Y_n^h)) \right] + R^1_n(h) + R^2_n(h) + R^3_n(h) + R^4_n(h) + R^5_n(h),
\]

in which

\[
R^1_n(h) = \mathbb{E}[R_1(n, h, Y_{n+1}^h)], \quad R^2_n(h) = h \mathbb{E}[(\mu_Y(Y_n^h)h + f_h(Y_n^h)) \cdot \nabla_{\bar{y}} \partial_t u(nh, Y_n^h)],
\]

\[
R^3_n(h) = \mathbb{E}[f_h(Y_n^h) \cdot \nabla_{\bar{y}} u(nh, Y_n^h)], \quad R^4_n(h) = \frac{1}{2} \mathbb{E}[(\text{Tr}(g_h(Y_n^h) D_{\bar{y}}^2 u(nh, Y_n^h))],
\]

\[
R^5_n(h) = \frac{1}{6} \sum_{|k|=3} \mathbb{E}[(\partial_{\bar{y}}^k u(nh, Y_n^h)) j_{h,k}(Y_n^h)].
\]

Thanks to (3.4), the first term in (3.14) is null, so

\[
|\mathbb{E}[u((n+1)h, Y_{n+1}^h) - u(nh, Y_n^h)]| \leq \sum_{i=1}^{5} |R^i_n(h)|.
\]

We now prove that \( |R^i_n(h)| \leq Ch^2 \), for every \( i = 1, \ldots, 5 \). Let \( \bar{h} > 0 \) such that both assumptions \( \mathcal{H}_1 \) and \( \mathcal{H}_2 \) hold and let \( h < \bar{h} \). Since the derivatives of \( u \) have polynomial growth, one has

\[
|\partial_t u(nh, Y_n^h)| \leq C \left( 1 + |Y_{n+1}^h| + |Y_n^h|^a \right) \left[ h^2 + h|Y_{n+1}^h - Y_n^h|^2 + |Y_{n+1}^h - Y_n^h|^4 \right],
\]

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, (3.11) and (3.12), we get

\[
|R^1_n(h)| \leq C \left| \left( 1 + |Y_{n+1}^h| + |Y_n^h|^a \right) \right|_2 \left[ h^2 + h(Y_{n+1}^h - Y_n^h)^2 + (Y_{n+1}^h - Y_n^h)^4 \right] \leq Ch^2.
\]
As regards $R_n^2(h)$, we use the polynomial growth of $\nabla_y \partial_t u$, the Cauchy-Schwarz inequality and the Hölder inequality, so that
\[
|R_n^2(h)| \leq C\mathbb{E}[\left(1 + |Y_n^h|^a\right) |\mu_Y(Y_n^h)|] h^2 + C\mathbb{E}[\left(1 + |Y_n^h|^a\right) |f_h(Y_n^h)|] h^2
\leq C\left[1 + |Y_n^h|^a\right]_2 \|\mu_Y(Y_n^h)\|_2 h^2 + C\left[1 + |Y_n^h|\right]_q \|f_h(Y_n^h)\|_p,
\]
where $p$ is given in (3.8) and $q$ is its conjugate exponent. Since $\mu_Y$ has polynomial growth, by (3.8) and (3.11) we get
\[
|R_n^2(h)| \leq Ch^2.
\]
The remaining terms $R_n^3(h)$, $R_n^4(h)$ and $R_n^5(h)$ can be handled similarly, so the statement follows. \hfill \square

### 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 well known CIR process $(Y_t)_{t \in [0,T]}$ solution to the SDE
\[
dY_t = \kappa(\theta - Y_t)dt + \sigma \sqrt{Y_t} dB_t, \quad Y_0 \geq 0.
\]
We assume that $\theta, \kappa, \sigma > 0$ and we stress that we never require the Feller condition $2\kappa \theta \geq \sigma^2$, ensuring that the process $Y$ does not hit 0. Therefore, the process $Y$ can reach 0.

The CIR process is widely used in finance to model interest rates or the volatility process in stochastic volatility models and there is a large literature on numerical methods to approximate it, see e.g. [2] [17] [23] [36]. We consider here the “multiple jumps” tree approximation for the CIR process developed in [6]. We first recall how the tree works and then, as an application of Theorem 3.1, we study the rate of convergence.

For $n = 0, 1, \ldots, N$ consider the lattice
\[
Y_n^h = \{y_k^n\}_{k=0,1,\ldots,n} \quad \text{with} \quad y_k^n = \left(\sqrt{Y_0} + \frac{\sigma}{2}(2k-n)\sqrt{h}\right)^2 1_{\{\sqrt{Y_0} + \frac{\sigma}{2}(2k-n)\sqrt{h} > 0\}}.
\]
Note that $Y_0^h = \{Y_0\}$. Moreover, the lattice is binomial recombining and, for $n$ large, the “small” points degenerate at 0. For each fixed node $(n, k) \in \{0, 1, \ldots, N - 1\} \times \{0, 1, \ldots, n\}$, the “up” jump $k_u(n, k)$ and the “down” jump $k_d(n, k)$ from $y_k^n \in Y_n^h$ are defined as
\[
\begin{align*}
k_u(n, k) &= \min\{k^* : k + 1 \leq k^* \leq n + 1 \text{ and } y_k^n + \mu_Y(y_k^n)h \leq y_{k^*}^{n+1}\}, \\
k_d(n, k) &= \max\{k^* : 0 \leq k^* \leq k \text{ and } y_k^n + \mu_Y(y_k^n)h \geq y_{k^*}^{n+1}\},
\end{align*}
\]
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 (3.16), resp. (3.17), is empty. This is called the “multiple jump approach”: the up jump can be larger than the closest up node, and similarly, the down jump can be smaller than the closest down node. This is as opposed to the “single jump approach”, where typically $k_d(n, k) = k$ and $k_u(n, k) = k + 1$.

The multiple jumps have been smartly introduced in [30] and are very useful because they allow one to define the transition probabilities such that the local first moment is asymptotically best fit. 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
\[
p_u(n, k) = 0 \vee \frac{\mu_Y(y_k^n)h + y_k^n - y_{k_u(n, k)}^{n+1}}{y_{k_u(n, k)}^{n+1} - y_{k_d(n, k)}^{n+1}} \land 1 \quad \text{and} \quad p_d(n, k) = 1 - p_u(n, k)
\]
respectively. We will see in next Proposition 3.3 that for $h$ small enough the parts “$\lor$” and “$\land$” can be omitted.

We call $(Y_n^h)_{n=0,1,\ldots,N}$ the Markov chain governed by the above jump probabilities. As an application of Theorem 3.1, we shall prove the following result.
Theorem 3.2. Let $f \in C_{pol}^b(\mathbb{R}_+)$. Then, there exist $\bar{h} > 0$ and $C > 0$ such that for every $h < \bar{h}$,

$$|E[f(Y^h_n)] - E[f(Y_T)]| \leq C Th,$$

that is, the tree approximation $(Y^h_n)_{n=0,...,N}$ is first order weak convergent.

In order to discuss the assumptions $\mathcal{H}_1$ and $\mathcal{H}_2$ of Theorem 3.1, we need some preliminary results which pave the way to the analysis of the convergence.

Proposition 3.3. There exist $\theta_*, \theta^*, C_*, \bar{h} > 0$ such that for any $h < \bar{h}$ the following properties hold.

1. If $\theta_* h \leq y^n_k \leq \theta^*/h$, then $k_u(n,k) = k + 1$, $k_d(n,k) = k$. Moreover,

$$y_{k_u(n,k)}^{n+1} = y^n_k + \frac{\sigma^2}{4} h + \sigma \sqrt{y^n_k h} \quad \text{and} \quad y_{k_d(n,k)}^{n+1} = y^n_k + \frac{\sigma^2}{4} h - \sigma \sqrt{y^n_k h}.$$  

2. If $y^n_k < \theta_* h$, then $k_d(n,k) = k$. Moreover,

$$0 \leq y_{k_u(n,k)}^{n+1} - y^n_k \leq C_* h. \tag{3.19}$$

3. If $y^n_k > \theta^*/h$, then $k_u(n,k) = k + 1$.

4. The jump probabilities are

$$p_u(n,k) = \frac{\mu_Y(y^n_k) h + y^n_k - y_{k_d(n,k)}^{n+1}}{y_{k_u(n,k)}^{n+1} - y_{k_d(n,k)}^{n+1}}, \quad p_d(n,k) = \frac{y_{k_u(n,k)}^{n+1} - y^n_k - \mu_Y(y^n_k) h}{y_{k_u(n,k)}^{n+1} - y_{k_d(n,k)}^{n+1}}. \tag{3.20}$$

The proof of Proposition 3.3 relies in a boring study of the properties of the lattice, so we postpone it in Appendix A. This is all we need to prove that $\mathcal{H}_2$ holds:

Proposition 3.4. The CIR approximating tree $\{Y^h_n\}_{n=0,...,N}$ satisfies Assumption $\mathcal{H}_2$.

Proof. Step 1: proof of (3.11). We use a technique firstly developed in [2] 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 $n = 0, \ldots, N - 1$ one has

$$0 \leq Y^h_{n+1} \leq (1 + bh)Y^n_n + Ch + \sigma \sqrt{Y^n_n h} W^h_{n+1}, \tag{3.21}$$

where $W^h_{n+1}$ is a r.v. such that

$$P(W^h_{n+1} = 2p_d(n,k)|Y^n_n = y^n_k) = p_u(n,k) = 1 - P(W^h_{n+1} = -2p_u(n,k)|Y^n_n = y^n_k). \tag{3.22}$$

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 (A.1)) that

$$y_{k+1}^{n+1} \leq y^n_k + \frac{\sigma^2}{4} h + \sigma \sqrt{y^n_k h}, \quad y_{k+1}^{n+1} \leq y^n_k + \frac{\sigma^2}{4} h - \sigma \sqrt{y^n_k h}.$$
Finally, if \( y_k^n \leq \theta_s h \), we have \( y_{k+1}^{n+1} = y_k^n \), while the up jump can be multiple but we can always write
\[
y_{k+1}^{n+1} \leq y_k^n + C_s h \leq y_k^n + C_s h + \sqrt{y_k^n h}.
\]
Summing up, if we set \( \bar{C} = \max \left( C_s, \frac{\sigma^2}{\theta} \right) \), for every \( h \) small we can write
\[
0 \leq Y_{n+1}^h \leq Y_n^h + \bar{C} h + \sigma \sqrt{Y_n^h Z_{n+1}^h},
\]
where \( Z_{n+1}^h \) is a random variable such that \( \mathbb{P}(Z_{n+1}^h = 1) = p_u(n,k) \) and \( \mathbb{P}(Z_{n+1}^h = -1) = p_d(n,k) \). 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 (3.22). We also define the function \( P_u(y_k^n) = p_u(n,k) \). Therefore,
\[
0 \leq Y_{n+1}^h \leq Y_n^h + \bar{C} h + \sigma \sqrt{Y_n^h h} (2P_u(Y_n^h) - 1) + \sigma \sqrt{Y_n^h h} W_{n+1}^h
\]
\[
\leq Y_n^h + \bar{C} h + \sigma \sqrt{Y_n^h h} (2P_u(Y_n^h) - 1) \mathbf{1}_{\{Y_n^h \leq \theta^* \}} + \sigma \sqrt{Y_n^h h} (2P_u(Y_n^h) - 1) \mathbf{1}_{\{Y_n^h < \theta^* \}},
\]
Now, if \( Y_n^h \geq \theta^* \) then \( \sqrt{Y_n^h h} \sigma \leq \theta \) and, since \( P_u \in [0,1] \), we have \( |2P_u(Y_n^h) - 1| \leq 1 \). Then, we have
\[
0 \leq Y_{n+1}^h \leq (1 + bh) Y_n^h + \bar{C} h + \sigma \sqrt{Y_n^h h} (2P_u(Y_n^h) - 1) \mathbf{1}_{\{Y_n^h \leq \theta^* \}} + \sigma \sqrt{Y_n^h h} W_{n+1}^h,
\]
where \( b = \frac{\sigma}{\sqrt{\theta^*}} \). Let us study the quantity
\[
\sigma \sqrt{Y_n^h h} (2P_u(Y_n^h) - 1) \mathbf{1}_{\{Y_n^h \leq \theta^* \}}.
\]
If \( \theta_s h > y_k^n < \theta^* h \), by using (3.20) and point 1. of Proposition 3.3 we can explicitly write
\[
\sigma \sqrt{y_k^n h} (2P_u(y_k^n) - 1) = \sqrt{y_k^n h} \left( 2 \left( \frac{1}{2} + \frac{4\mu Y(v_k^n) - \sigma^2}{8\sigma \sqrt{y_k^n}} \right) h - 1 \right) = \mu Y(v_k^n) h - \frac{\sigma^2}{4} h \leq \kappa h.
\]
If instead \( y_k^n \leq \theta_s h \), then by using 2. in Proposition 3.3 we have
\[
\sigma \sqrt{y_k^n h} (2P_u(y_k^n) - 1) \leq \sigma \sqrt{y_k^n h} \frac{2\mu Y(v_k^n) h + 2\mu Y(v_k^n) - y_k^n - y_k^{n+1}}{y_k^{n+1} - y_k^n} \leq \sigma \sqrt{y_k^n h} \frac{2\mu h + 2\mu h}{2\sigma \sqrt{y_k^n}} = (\kappa \theta + \theta_s) h.
\]
So, by inserting, for every \( n \leq N - 1 \) we get
\[
0 \leq Y_{n+1}^h \leq (1 + bh) Y_n^h + \bar{C} h + \sigma (\kappa \theta + \theta_s) h + \sigma \sqrt{Y_n^h h} W_{n+1}^h
\]
and (3.21) is proved.

Now, by using (3.21) and (3.22), we can repeat step by step the proof of Lemma 2.6 in [2] and we get (3.11).

Step 2: proof of (3.12). We can write
\[
|Y_{n+1}^h - Y_n^h|^p \leq 3^{p-1} \left( \frac{\sigma^2}{\theta} + \sigma \sqrt{Y_n^h h} Z_{n+1}^h \right)^p \mathbf{1}_{\{\theta_h < \theta^* \theta_h \}} + 3^{p-1} |Y_{n+1}^h - Y_n^h|^p \mathbf{1}_{\{Y_n^h \leq \theta_h \}} + 3^{p-1} |Y_{n+1}^h - Y_n^h|^p \mathbf{1}_{\{Y_n^h \geq \theta^* \theta_h \}} =: 3^{p-1} (I_1 + I_2 + I_3),
\]
where we have used that, on the set \( \{ \theta_s h < Y_n^h < \theta^*/h \} \), we have \( Y_{n+1}^h = Y_n^h + \frac{\sigma^2}{4} h + \sigma \sqrt{Y_n^h} Z_{n+1}^h \), with \( \mathbb{P}(Z_{n+1}^h = 1) = P_d(Y_n^h) \) and \( \mathbb{P}(Z_{n+1}^h = -1) = P_d(Y_n^h) \). Now, by using (3.11), Proposition 3.3, the Cauchy-Schwartz and the Markov inequality,

\[
I_1 \leq \mathbb{E}\left[\left(\frac{\sigma^2}{4} + \sigma \sqrt{Y_n^h} h\right)^p\right] \leq \left(2^{p-1} \left(\frac{\sigma^2}{4}\right)^p + \sigma^p \mathbb{E}[Y_n^h]^{p/2}\right) h^{p/2} \leq 2^{p-1} \left(\frac{\sigma^2}{4}\right)^p + \sigma^p \mathbb{E}[Y_n^h]^{p/2},
\]

\[
I_2 \leq C_p h^p,
\]

\[
I_3 \leq \mathbb{E}(Y_{n+1}^h - Y_n^h)^{2p} \mathbb{E}\left[Y_n^h > \frac{\theta^*}{h}\right]^{1/2} \leq 2^{p-1} \left(\frac{\sigma^2}{4}\right)^p + \sigma^p \mathbb{E}[C_p] h^{p/2},
\]

and (3.12) follows.

**Proposition 3.5.** The CIR approximating tree \( \{ Y_n^h \}_{n=0,\ldots,N} \) satisfies Assumption \( \mathcal{H}_1 \).

*Proof.* Straightforward computations give \( \mathbb{E}[Y_{n+1}^h - Y_n^h | Y_n^h] = \mu_f(Y_n^h)h \), so (3.5) and (3.8) immediately follow. As for (3.6),

\[
\mathbb{E}(Y_{n+1}^h - Y_n^h)^2 | Y_n^h = y_k^n = \mathbb{E}(Y_{n+1}^h - Y_n^h)^2 | Y_n^h = y_k^n 1_{\{y_k^n \leq \theta_s h\}} + \mathbb{E}(Y_{n+1}^h - Y_n^h)^2 | Y_n^h = y_k^n 1_{\{\theta_s h < y_k^n \leq \theta^*/h\}} + \mathbb{E}(Y_{n+1}^h - Y_n^h)^2 | Y_n^h = y_k^n 1_{\{y_k^n > \theta^*/h\}}.
\]

We study separately the first two terms of the above r.h.s. If \( y_k^n < \theta_s h \), Proposition 3.3 gives \( |y_{k+1}^n - y_k^n| \leq C_s h \) and \( |y_{k+1}^n - y_k^n| \leq | y_{k}^n \leq \theta_s h \}

\[
\mathbb{E}[Y_{n+1}^h - Y_n^h)^2 | Y_n^h = y_k^n 1_{\{y_k^n \leq \theta_s h\}} = \varphi_1(y_k^n) h^2 1_{\{ y_k^n \leq \theta_s h \}},
\]

with \( \varphi_1 \) such that \( |\varphi_1(y)| \leq C_s^2 \). If instead \( \theta_s h \leq y_k^n \leq \theta^*/h \), by using (3.20) we get

\[
(y_{k+1}^n - y_k^n)^2 p_u(n,k) + (y_{k+1}^n - y_k^n)^2 p_d(n,k) = \sigma^2 y_k^n h + \frac{\sigma^2}{2} (\kappa - y_k^n - \frac{\sigma^2}{2} h^2).
\]

So,

\[
\mathbb{E}[Y_{n+1}^h - Y_n^h)^2 | Y_n^h = y_k^n 1_{\{\theta_s h < y_k^n \leq \theta^*/h\}} = \left(\sigma^2 y_k^n h + \varphi_2(y_k^n) h^2\right) 1_{\{ \theta_s h < y_k^n \leq \theta^*/h \}},
\]

with \( \varphi_2 \) such that \( |\varphi_2(y)| \leq C_s^2 \). By inserting, (3.6) follows with \( g_n \) satisfying

\[
|g_n(Y_n^h)| \leq c_1 (1 + Y_n^h) h^2 + \mathbb{E}(Y_{n+1}^h - Y_n^h)^2 + \sigma h Y_n^h | Y_n^h 1_{\{ Y_n^h \geq \theta^*/h \}},
\]

c_1 denoting a suitable constant. By Proposition 3.4 and the Markov inequality, (3.9) follows.

Finally, for (3.7), we write

\[
\mathbb{E}[Y_{n+1}^h - Y_n^h)^3 | Y_n^h = y_k^n = \mathbb{E}[Y_{n+1}^h - Y_n^h)^3 | Y_n^h = y_k^n 1_{\{y_k^n \leq \theta_s h\}} + \mathbb{E}[Y_{n+1}^h - Y_n^h)^3 | Y_n^h = y_k^n 1_{\{y_k^n > \theta^*/h\}}.
\]

Now, if \( y_k^n \leq \theta_s h \) then \( Y_{n+1}^h - y_k^n \beta \leq C_s^3 h^3 \). If instead \( \theta_s h < y_k^n < \theta^*/h \), by (3.20) one obtains

\[
(y_{k+1}^n - y_k^n)^3 p_u(n,k) + (y_{k+1}^n - y_k^n)^3 p_d(n,k) = \mu_f(y_k^n) h^2 \left(\sigma^2 y_k^n h + 3\sigma d_4 h^2\right) + \left(\frac{\sigma^4}{2} y_k^n + \frac{\sigma^4}{16} h^2\right) h^2.
\]

Therefore,

\[
|g_n(Y_n^h)| \leq c_2 h^2 (1 + (Y_n^h)^2) + \mathbb{E}[Y_{n+1}^h - Y_n^h)^3 + \sigma h Y_n^h | Y_n^h 1_{\{ Y_n^h \geq \theta^*/h \}},
\]

c_2 denoting a suitable constant, and again by Proposition 3.4 and the Markov inequality, (3.10) follows.

We are finally ready for the

**Proof of Theorem 3.2.** By Theorem 4.1 in [2] (or Corollary 5.5), one has that if \( f \in C_p^{4}(\mathbb{R}_+) \) then \( u \in C_p^{4}(\mathbb{T}^{4}(\mathbb{R}_+)) \). Since Assumption \( \mathcal{H}_1 \) and \( \mathcal{H}_2 \) both hold, the statement follows as an application of Theorem 3.1.
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 3. More precisely, we consider the stochastic system

\[
\begin{align*}
    dX_t &= \mu_X(Y_t) dt + \sigma_X(Y_t) dB_t + \gamma_X(Y_t) dH_t, & X_0 \in \mathbb{R}^m, \\
    dY_t &= \mu_Y(Y_t) dt + \sigma_Y(Y_t) dW_t, & Y_0 \in \mathcal{D}, \\
\end{align*}
\]

where \( B \) is a \( \ell_1 \)-dimensional Brownian motion independent of \( W \) and \( H \) is a \( \ell_2 \)-dimensional compound Poisson process with intensity \( \lambda \) and i.i.d. jumps \( \{J_k\}_k \) taking values in \( \mathbb{R}^{\ell_2} \), that is,

\[
H_t = \sum_{k=1}^{K_t} J_k, \quad (4.2)
\]

\( K \) denoting a Poisson process with intensity \( \lambda \). We assume that the Poisson process \( K \), the jump amplitudes \( \{J_k\}_k \) and the Brownian motion \((B,W)\) 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:

\[
\nu(dx) = \nu(x)dx = \lambda p_{J_1}(x)dx.
\]

We denote by \( \mathcal{L} \) the infinitesimal generator associated with the diffusion pair \((X,Y)\):

\[
\mathcal{L}g(x,y) = \frac{1}{2} \text{Tr}(a(y) D^2_{x,y} g(x,y)) + \mu(y) \cdot \nabla_x g(x,y) + \int (g(x + \gamma_X(Y(t)) \zeta, y) - g(x, y)) \nu(d\zeta),
\]

where \( \mu(y) = (\mu_X(y), \mu_Y(y))^* \) and \( a(y) = \sigma \sigma^*(y) \), 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 \((x,y)\). 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 paper are still valid in the time-depending case under non restrictive classical assumptions.

Hereafter, we fix \( T > 0 \), \( f : \mathbb{R}^m \times \mathcal{D} \rightarrow \mathbb{R} \) and we define

\[
\begin{align*}
    u(t,x,y) &= \mathbb{E} \left[ f(X_T^{t,x,y}, Y_T^{t,y}) \right], \quad (t,x,y) \in [0,T] \times \mathbb{R}^m \times \mathcal{D},
\end{align*}
\]

where \((X^{t,x,y}_s, Y^{t,y}_s)_{s \in [t,T]}\) is the solution of \((4.1)\) with starting condition \((X_t, Y_t) = (x,y)\). We do not enter in specific assumptions but from now on, the following requests (1) and (2) are assumed to hold:

1. there exists a unique weak solution of \((4.1)\) and \( \mathbb{P}((X_t, Y_t) \in \mathbb{R}^m \times \mathcal{D} \ \forall t) = 1; \)
2. the function \( u \) in \((4.4)\) solves the PIDE

\[
\begin{align*}
    \partial_t u(t,x,y) + \mathcal{L} u(t,x,y) &= 0, & (t,x,y) \in [0,T) \times \mathbb{R}^m \times \mathcal{D}, \\
    u(T,x,y) &= f(x,y), & \text{in } \mathbb{R}^m \times \mathcal{D},
\end{align*}
\]

\( \mathcal{L} \) being given in \((4.3)\).

We are interested in computing \( u(0,X_0,Y_0) = \mathbb{E} \left[ f(X_T, Y_T) \right] \). This is a problem of interest in a large number of applications. For example, in finance \( X \) can represent the asset 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, so \( u(t,x,y) \) represents the value function at time \( t \) of a European option with maturity \( T \) and (discounted) payoff \( f \). In next Section 5 we give an application to the Heston model \([22]\) and the Bates model \([10]\).
4.1 The hybrid procedure

Let \( u \) be given in (4.4). We study here the computation of \( u(0, X_0, Y_0) \) by a backward hybrid procedure developed in [12, 13, 14]. 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 describe the approximation of \( u \).

We start from the representation of \( u(t, x, y) \) at times \( nh, h = T/N \) and \( n = 0, \ldots, N \), by the usual dynamic programming principle: for \( (x, y) \in \mathbb{R}^m \times \mathcal{D} \),

\[
\begin{cases}
  u(T, x, y) = f(x, y) & \text{and as } n = N - 1, \ldots, 0, \\
  u(nh, x, y) = \mathbb{E} \left[ u \left( (n + 1)h, X_{(n+1)h}^{nh,x,y}, Y_{(n+1)h}^{nh,y} \right) \right].
\end{cases}
\]  

(4.6)

So, the central issue is to have a good approximation of the expectations in (4.6).

As a first step, let \( (Y^n_h)_{n=0,\ldots,N} \) be a Markov chain which approximates \( Y \). Of course, we assume that \( (Y^n_h)_{n=0,\ldots,N} \) is independent of the noises \( B \) (Brownian motion) and \( H \) (compound Poisson process) driving \( X \) in (4.1). Then, at each step \( n = 0, 1, \ldots, N - 1 \), for every \( y \in Y_n^h \subseteq \mathcal{D} \) (the state space of \( Y_n^h \)) we write

\[
\mathbb{E} \left[ u \left( (n + 1)h, X_{(n+1)h}^{nh,x,y}, Y_{(n+1)h}^{nh,y} \right) \right] \approx \mathbb{E} \left[ u \left( (n + 1)h, X_{(n+1)h}^{nh,x,y}, \hat{Y}_{n+1}^h \right) \right].
\]

As a second step, we approximate the component \( X \) on \([nh, (n + 1)h] \) by freezing the coefficients in (4.1) at the observed position \( Y_n^h = y \), that is, for \( t \in [nh, (n + 1)h] \),

\[
X_{t}^{nh,x,y} \approx \hat{X}_{t}^{nh,x}(y) = x + \mu_X(y)(t - nh) + \sigma_X(y)(B_t - B_{nh}) + \gamma_X(y)(H_t - H_{nh}).
\]

Therefore, by using that the Markov chain, \( B \) and \( H \) are all independent, we write

\[
\mathbb{E} \left[ u \left( (n + 1)h, X_{(n+1)h}^{nh,x,y}, Y_{(n+1)h}^{nh,y} \right) \right] \approx \mathbb{E} \left[ u \left( (n + 1)h, \hat{X}_{(n+1)h}^{nh,x}(y), \hat{Y}_{n+1}^h \right) \right]
\]

\[
= \mathbb{E} \left[ \phi(\hat{Y}_{n+1}^h; x, y) | \hat{Y}_{n}^h = y \right],
\]

where

\[
\phi(\zeta; x, y) = \mathbb{E} \left[ u \left( (n + 1)h, \hat{X}_{(n+1)h}^{nh,x}(y), \zeta \right) \right].
\]  

(4.7)

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

\[
\partial_t v + \mathcal{L}^{(y)}v = 0, \quad \text{in } [nh, (n + 1)h) \times \mathbb{R}^m, \quad v((n + 1)h, x; y, \zeta) = u((n + 1)h, x, \zeta), \quad x \in \mathbb{R}^m,
\]  

(4.8)

where

\[
\mathcal{L}^{(y)}g(x) = \mu_X(y) \cdot \nabla_x g(x) + \frac{1}{2} \text{Tr}(a_X(y)D_x^2 g(x)) + \int \left( g(x + \gamma_X(y) \zeta) - g(x) \right) \nu(\zeta) d\zeta
\]  

(4.9)

is an integro-differential operator, acting on the functions \( g = g(x) \). Here \( a_X(y) = \sigma_X(y)\sigma_X^*(y) \), while \( \nabla_x \) and \( D_x^2 \) are, respectively, the gradient vector and the Hessian matrix with respect to \( x \in \mathbb{R}^m \).

Recall that in (1.3)–(1.9), \( y \in \mathcal{D} \) is just a parameter, so \( \mathcal{L}^{(y)} \) has constant coefficients.

Consider now a numerical solution of the PIDE (1.8). Let \( \Delta x = (\Delta x_1, \ldots, \Delta x_m) \) denote a fixed spatial step and set \( \mathcal{X} \) a grid on \( \mathbb{R}^m \) given by \( \mathcal{X} = \{ x : x = ((X_0)_1 + i_1 \Delta x_1, \ldots, (X_0)_m + i_m \Delta x_m), (i_1, \ldots, i_m) \in \mathbb{Z}^m \} \). For \( y \in \mathcal{D} \), let \( \Pi_{\Delta x}^h(y) \) be a linear operator (acting on suitable functions
on \( \mathcal{X} \) which gives the approximating solution to the PIDE (4.8) at time \( nh \). Then, as \( x \in \mathcal{X} \), we get the numerical approximation

\[
E\left[u((n+1)h, X_{(n+1)h}^{nh,x,y}, Y_{(n+1)h}^{nh,y})\right] \approx E\left[\Pi_{\Delta x}^h(y)u((n+1)h, \cdot, \hat{Y}_{n+1}^h(x)) \mid \hat{Y}_n^h = y\right].
\]

Therefore, by inserting in (4.6), the hybrid numerical procedure works as follows: the function \( x \mapsto u(0, x, Y_0) \), \( x \in \mathcal{X} \), is approximated by \( u_0^h(x, Y_0) \) backwardly defined as

\[
\begin{align*}
    u_N^h(x, y) &= f(x, y), \quad (x, y) \in \mathcal{X} \times Y_N^h, \quad \text{and as } n = N-1, \ldots, 0: \\
    u_n^h(x, y) &= E[\Pi_{\Delta x}^h(y)u_{n+1}(\cdot, \hat{Y}_{n+1}^h(x)) \mid \hat{Y}_n^h = y], \quad (x, y) \in \mathcal{X} \times Y_n^h.
\end{align*}
\]

### 4.2 Convergence speed of the scheme (4.10)

We introduce the following assumption on the linear operator \( \Pi_{\Delta x}^h(y) \) in (4.10) (recall the notation \( l_p(\mathcal{X}) \) in Section 2).

**Assumption \( K(p, c, \mathcal{E}) \).** 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 \( K(p, c, \mathcal{E}) \) if

\[
|\Pi_{\Delta x}^h(y)|_p \leq 1 + c(y)h \tag{4.11}
\]

and, \( u \) being defined in (4.3), for every \( n = 0, \ldots, N - 1 \), one has

\[
E\left[\Pi_{\Delta x}^h(\hat{Y}_n^h)u((n+1)h, \cdot, \hat{Y}_{n+1}^h(x)) \mid \hat{Y}_n^h\right] = u(nh, x, \hat{Y}_n^h) + R_n^h(x, \hat{Y}_n^h), \tag{4.12}
\]

where the remainder \( R_n^h(x, \hat{Y}_n^h) \) satisfies the following property: there exist \( \tilde{h}, C > 0 \) such that for every \( h < \tilde{h}, \Delta x < 1 \) and \( n \leq N = \lfloor T/h \rfloor \) one has

\[
\left\| e^{\sum_{i=1}^n c(\hat{Y}_i^h)h} |R_n^h(\cdot, \hat{Y}_n^h)|_p \right\|_p \leq Ch\mathcal{E}(h, \Delta x), \quad \text{if } p \in [1, \infty]. \tag{4.13}
\]

Assumption \( K(p, c, \mathcal{E}) \) is inspired by the Lax-Richtmeyer’s convergence theorem [27]. In fact, recall that the numerical procedure (4.10) aims to solve the multidimensional equation

\[
\partial_t u(t, x, y) + Lu(t, x, y) = 0, \quad (t, x, y) \in [0, T) \times \mathbb{R}^m \times \mathcal{D}.
\]

Being dependent on \( y \), the coefficients of \( L \) (see 4.3) are not constant as required by the Lax-Richtmeyer’s result. But at each time step \( n \), the hybrid scheme isolates the component \( y \) and applies the discrete operator \( \Pi_{\Delta x}^h(y) \) to numerically solve the PIDE

\[
\partial_t v(t, x) + L(y)v(t, x) = 0, \quad (t, x) \in [nh, (n + 1)h) \times \mathbb{R}^m.
\]

Here, \( y \) is just a parameter (the current position of the Markov chain), so the coefficients of \( L(y) \) (see 4.4) are indeed constant. That’s why the Lax-Richtmeyer technique can be adapted, as it follows in the next result.

**Theorem 4.1.** Assume that \( \Pi_{\Delta x}^h(y) \), \( y \in \mathcal{D} \), satisfies Assumption \( K(p, c, \mathcal{E}) \). Let \( u \) be defined in (4.4) and \( u^h \) be the approximation through the scheme (4.10). Then, there exist \( \tilde{h}, C > 0 \) such that for every \( h < \tilde{h} \) and \( \Delta x < 1 \) one has

\[
|u(0, \cdot, Y_0) - u_0^h(\cdot, Y_0)|_p \leq CTE(h, \Delta x). \tag{4.14}
\]
In particular, (b) allows us to define the measure $\nu$ from an appropriate change of variable allowing to set up the quadrature rules (see Remark 4.4 below). Let us stress that (a) is necessary in order to control suitable remaining terms, whereas (b) follows.

Moreover, hereafter we assume that the coefficients in (4.1) satisfy:

$$\text{inf}_{n=0}^{\infty} \sum_{l=0}^{n-1} \mathbb{E}\left[ \left( \prod_{l=0}^{n-1} \Pi_{\Delta x}(\hat{Y}_l^{h}) \right) \mathcal{R}_{n}^{h}(\cdot, \hat{Y}_n^{h}) \right] \leq TCE(h, \Delta x).$$

The case $p = \infty$ follows the same lines.

**4.3 An application: finite difference schemes**

We specify here some settings ensuring that the assumptions of Theorem 4.1 are satisfied. In particular, we choose the operator $\Pi_{\Delta x}(y)$ in (4.10) by means of two different finite difference schemes: the first one allows us to study the convergence in the $l_2$-norm (Section 4.3.1), while the second one in the $l_\infty$-norm (Section 4.3.2). For the sake of readability, we consider the case $m = d = \ell = \ell_1 = \ell_2 = 1$. Moreover, hereafter we assume that the coefficients in (4.1) satisfy:

(a) $\mu = (\mu_X, \mu_Y)^*$ and $\sigma_X$ have polynomial growth;

(b) either $\gamma_X \equiv 0$ (no jumps) or there exists $\varepsilon > 0$ such that $\inf_{y \in \mathcal{D}} |\gamma_X(y)| \geq \varepsilon$ (uniform ellipticity condition).

Let us stress that (a) is necessary in order to control suitable remaining terms, whereas (b) follows from an appropriate change of variable allowing to set up the quadrature rules (see Remark 4.4 below). In particular, (b) allows us to define the measure $\nu_g$ as follows:

$$\nu_g(x) = \begin{cases} 0 & \text{if } \gamma_X \equiv 0, \\ \frac{1}{|\gamma_X(y)|} \nu\left( \frac{x}{\gamma_X(y)} \right) & \text{otherwise, } y \in \mathcal{D}, \end{cases} \text{ (4.15)}$$

$\nu$ denoting the density of the Lévy measure.

**Proposition 4.2.** If $\frac{\gamma'}{\nu}, \frac{\mu'}{\nu} \in L^1(\mathbb{R}, d\nu)$, there exists $c_\nu \geq 0$ such that

$$\sum_{l \in \mathbb{Z}} \nu_g(l\Delta x) \Delta x \leq \lambda c_\nu, \quad \forall y \in \mathcal{D}. \text{ (4.16)}$$

**Proof.** The proof follows from the technical Lemma 4.3 below: if $\gamma_X$ is non null, (i) applied to $g(x) = \nu_g(x)$ gives $\sum_{l \in \mathbb{Z}} \nu_g(l\Delta x) \Delta x \leq \int_{\mathbb{R}} \nu(x) dx + \frac{\Delta x^2}{2|\gamma_X(y)|^2} \int_{\mathbb{R}} |\mu''(x)| dx$. Now we use the “uniformity” condition $\inf_{y \in \mathcal{D}} \gamma_X(y) \geq \varepsilon$, and the statement holds.

---

Proof. Let $err_n^h(\cdot, \hat{Y}_n^h)$ be the error at time $nh$, defined by

$$err_n^h(\cdot, \hat{Y}_n^h) = u(nh, \cdot, \hat{Y}_n^h) - u_n^h(\cdot, \hat{Y}_n^h).$$

Note that $err_n^h(\cdot, \hat{Y}_n^h) = 0$, because the final condition is the same. By (4.12) and (4.10), we can write

$$err_n^h(\cdot, \hat{Y}_n^h) = \mathbb{E}\left[ \prod_{l=0}^{n-1} \Pi_{\Delta x}(\hat{Y}_l^{h}) \mathcal{R}_{n}^{h}(\cdot, \hat{Y}_n^{h}) \right],$$

and, by iterating,

$$err_0^h(\cdot, Y_0) = - \sum_{n=0}^{N-1} \mathbb{E}\left[ \left( \prod_{l=0}^{n-1} \Pi_{\Delta x}(\hat{Y}_l^{h}) \right) \mathcal{R}_{n}^{h}(\cdot, \hat{Y}_n^{h}) \right],$$

in which we use the convention $\prod_{l=0}^{N-1} \Pi_{\Delta x}(\cdot) = \text{Id}$. We now apply (4.13). For $p \neq \infty$,

$$|err_n^h(\cdot, Y_0)|_p \leq \sum_{n=0}^{N-1} \mathbb{E}\left[ \left( \prod_{l=0}^{n-1} \Pi_{\Delta x}(\hat{Y}_l^{h}) \right) \mathcal{R}_{n}^{h}(\cdot, \hat{Y}_n^{h}) \right]_p^{1/p} \leq \sum_{n=0}^{N-1} \mathbb{E}\left[ e^{\sum_{l=0}^{n-1} pc(\hat{Y}_l^{h})} |\mathcal{R}_{n}^{h}(\cdot, \hat{Y}_n^{h})| \right]^{1/p} \leq \sum_{n=0}^{N-1} hC\mathcal{E}(h, \Delta x) \leq TCE(h, \Delta x).$$

The case $p = \infty$ follows the same lines. \qed
Lemma 4.3. Let $g \in C^2(\mathbb{R})$.

(i) If $g, g', g'' \in L^1(\mathbb{R}, dx)$ then

$$\left| \sum_{l \in \mathbb{Z}} g(l \Delta x) \Delta x - \int_\mathbb{R} g(x) dx \right| \leq \frac{\Delta x^2}{12} |g''|_{L^1(\mathbb{R}, dx)}. \quad (4.17)$$

(ii) If $g, g', g'' \in L^2(\mathbb{R}, dx)$ then

$$|g|^2 \leq |g|_{L^2(\mathbb{R}, dx)}^2 + \frac{\Delta x^2}{6} \left( |g'|_{L^2(\mathbb{R}, dx)}^2 + |g|_{L^2(\mathbb{R}, dx)} |g''|_{L^2(\mathbb{R}, dx)} \right). \quad (4.18)$$

Proof. We first recall the Poisson summation formula. It is worldwide famous but is usually written on the Schwartz space, we use here the following version (Appendix B contains the detailed proof): if $\varphi \in C^2(\mathbb{R})$ with $\varphi, \varphi', \varphi'' \in L^1(\mathbb{R}, dx)$ then

$$\sum_{n \in \mathbb{Z}} \varphi(n) = \int_\mathbb{R} \varphi(x) dx + \sum_{n \in \mathbb{Z}, n \neq 0} \int_\mathbb{R} \varphi(x) e^{-2\pi i nx} dx. \quad (4.19)$$

(i) We apply (4.19) to $\varphi(x) = g(x \Delta x)$. So,

$$\sum_{n \in \mathbb{Z}} g(n \Delta x) \Delta x - \int_\mathbb{R} g(x) dx = \sum_{n \in \mathbb{Z}, n \neq 0} \int_\mathbb{R} g(x) e^{-2\pi i nx/\Delta x} dx$$

the latter inequality coming from the integration by parts formula. The statement holds by recalling that $\sum_{n \geq 1} \frac{1}{n^2} = \frac{\pi^2}{6}$.

(ii) (4.18) immediately follows by applying (4.17) to the function $x \mapsto g^2(X_0 + x)$. This statement will be used to handle the error in $l_2$-norm coming from suitable Taylor’s expansions and from the quadrature approximation. \hfill \square

4.3.1 Convergence in $l_2$-norm

We study here the hybrid procedure introduced in [14] for the Bates model. Recall that, for $y \in \mathcal{D}$, $\Pi_{\Delta x}^h(y)$ gives the numerical solution on $\mathcal{X} = \{ x_i = X_0 + i \Delta x \}_{i \in \mathbb{Z}}$ a time $nh$ to the PIDE (1.8), the operator $\mathcal{L}^{(y)}$ therein being given in (1.9). It is clear that the solution $v$ of (1.8) depends on $y$ and $\zeta$ as well, but these are just parameters (and not variables of the PIDE), so for simplicity we drop here such dependence. So, we split the operator $\mathcal{L}^{(y)} v(t, x) = \mathcal{L}^{(y)}_{\text{diff}} v(t, x) + \mathcal{L}^{(y)}_{\text{int}} v(t, x)$ in its differential and integral part:

$$\mathcal{L}^{(y)}_{\text{diff}} v(t, x) = \mu \chi(y) \partial_x v(t, x) + \frac{1}{2} \sigma^2 \chi(y) \partial_{xx}^2 v(t, x), \quad (4.20)$$

$$\mathcal{L}^{(y)}_{\text{int}} v(t, x) = \int (v(t, x + \gamma \chi(y) z) - v(t, x)) \nu(z) dz. \quad (4.21)$$

We use the central finite difference scheme to solve $\mathcal{L}^{(y)}_{\text{diff}} v$ and the trapezoidal rule in order to approximate the integral term $\mathcal{L}^{(y)}_{\text{int}} v$. Applying an implicit-explicit method in time, we obtain an approximating solution $v^n = (v^n_j)_{j \in \mathbb{Z}} : \mathcal{X} \to \mathbb{R}$ to the PIDE (1.8) given by

$$A_{\Delta x}^h(y) v^n = B_{\Delta x}^h(y) v^{n+1}, \quad (4.22)$$

15
where the linear operators $A_{\Delta x}^h(y)$ is defined as

$$
(A_{\Delta x}^h)_{ij}(y) = \begin{cases} 
\alpha_{\Delta x}^h(y) - \beta_{\Delta x}^h(y), & \text{if } i = j + 1, \\
1 + 2\beta_{\Delta x}^h(y), & \text{if } i = j, \\
-\alpha_{\Delta x}^h(y) - \beta_{\Delta x}^h(y), & \text{if } i = j - 1, \\
0, & \text{if } |i - j| > 1
\end{cases}
$$

(4.23)

with

$$
\alpha_{\Delta x}^h(y) = \frac{h}{2\Delta x} \mu_X(y), \quad \beta_{\Delta x}^h(y) = \frac{h}{2\Delta x^2} \lambda_{\Delta x}(y).
$$

(4.24)

Moreover, we choose the approximation $B_{\Delta x}^h(y)$ for $L_{\text{int}}$ in order to work on the same numerical grid $X$. This is achievable by using the change of variable in $L_{\text{int}}^y$: $L_{\text{int}}^y v(t, x) = \int (v(t, x + \zeta) - v(t, x)) \nu_y(\zeta) d\zeta$, $\nu_y$ being defined in (4.15). Then we get

$$
(B_{\Delta x}^h)_{ij}(y) = \begin{cases} 
\nu_y((j - i)\Delta x) \Delta x & \text{if } j \neq i, \\
1 + h \left( \nu_y(0) \Delta x - \sum_{l \in \mathbb{Z}} \nu_y(l\Delta x) \Delta x \right) & \text{if } i = j.
\end{cases}
$$

(4.25)

Note that $B_{\Delta x}^h(y) = \text{Id}$ if $\gamma_X \equiv 0$.

**Remark 4.4.** The above construction for $B_{\Delta x}^h(y)$ justifies the “uniformly ellipticity” requirement for $\gamma_X$. One could drop this assumption by avoiding the change of variable in $L_{\text{int}}$. But this would bring to the use of a numerical grid depending on $y$ and therefore, the introduction of suitable interpolations. As a consequence, one would have a complication of the numerical scheme, the introduction of technical details and further notations.

The operators $A_{\Delta x}^h(y)$ and $B_{\Delta x}^h(y)$ in (4.22) and (4.25) respectively, have the following properties.

**Lemma 4.5.** For every $y \in \mathcal{D}$, $A_{\Delta x}^h(y)$ : $l_2(\mathcal{X}) \rightarrow l_2(\mathcal{X})$ is invertible and moreover,

$$
\sup_{y \in \mathcal{D}} |(A_{\Delta x}^h)^{-1}(y)|_2 \leq 1.
$$

And if $\frac{\nu}{\nu} \in L^1(\mathbb{R}, d\nu)$ then $\sup_{y \in \mathcal{D}} |B_{\Delta x}^h(y)|_2 \leq 1 + 2\lambda c_v h$, $c_v$ being defined in (4.10).

**Proof.** Let $\mathcal{F} : l_2(\mathcal{X}) \rightarrow L^2([0, 2\pi), dx)$ denote the Fourier transform: $\mathcal{F}(\varphi)(\theta) = \frac{\Delta x}{\sqrt{2\pi}} \sum_{j \in \mathbb{Z}} \varphi_j e^{-ij\Delta x \theta}$, $\theta \in [0, 2\pi)$, $\varphi \in l_2(\mathcal{X})$.

Fix $y \in \mathcal{D}$ and $w \in l_2(\mathcal{X})$. $v \in l_2(\mathcal{X})$ satisfies $A_{\Delta x}^h(y)v = w$ iff $\mathcal{F}(A_{\Delta x}^h(y)v) = \mathcal{F}(w)$. Straightforward computations give (see e.g. the proof of Theorem 5.1 in [14]) $\mathcal{F}(A_{\Delta x}^h(y)v) = \psi \times \mathcal{F}(v)$, with $\psi(\theta) = (\alpha_{\Delta x}^h(y) - \beta_{\Delta x}^h(y)) e^{-i\theta \Delta x} + 1 + 2\beta_{\Delta x}^h(y) - (\alpha_{\Delta x}^h(y) + \beta_{\Delta x}^h(y)) e^{i\theta \Delta x}$. It can be easily seen that $|\psi(\theta)| \geq 1 + 2\beta_{\Delta x}^h(y)(1 - \cos(\theta \Delta x)) \geq 1$. Hence $\mathcal{F}(v) = \mathcal{F}(w)/\psi \in L^2([0, 2\pi), dx)$ and its inverse Fourier transform uniquely defines the solution $v \in l_2(\mathcal{X})$ to $A_{\Delta x}^h(y)v = w$. Thus $A_{\Delta x}^h$ is invertible. Moreover, we obtain $|\mathcal{F}(v)|_{L^2([0, 2\pi), dx)} \leq |\mathcal{F}(w)|_{L^2([0, 2\pi), dx)}$. By the Parseval identity we get $|(A_{\Delta x}^h)^{-1}(y)w|_2 \leq |w|_2$, so $|(A_{\Delta x}^h)^{-1}(y)w|_2$.

Finally, for $w \in l_2(\mathcal{X})$ straightforward computations give

$$
\mathcal{F}(B_{\Delta x}^h(y)w)(\theta) = \left(1 + h \Delta x \sum_i \nu_y(l\Delta x)(e^{ih\theta} - 1)\right) \mathcal{F}(w)(\theta).
$$

Then, $|\mathcal{F}(B_{\Delta x}^h(y)w)|_{L^2([0, 2\pi), dx)} \leq (1 + 2\lambda c_v h)|\mathcal{F}(w)|_{L^2([0, 2\pi), dx)}$ because (4.10) holds. By the Parseval relation, $|B_{\Delta x}^h(y)w|_2 \leq (1 + 2\lambda c_v h)|w|_2$, which concludes the proof. 

\(\square\)
We can now state the convergence result in $l_2(\mathcal{X})$, saying that 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 $\hat{Y}^h$ (see Theorem 3.1).

**Theorem 4.6.** Let $u$ be defined in (4.4) and $(u_n^h)_{n=0,...,N}$ be given by (4.10) with the choice

$$\Pi_n^h(y) = (A_n^h)^{-1}B_n^h(y),$$

$A_n^h(x)$ and $B_n^h(x)$ being given in (4.23) and (4.25) respectively. Assume that

- $\frac{\partial^2 u}{\partial y^2}, \frac{\partial u}{\partial y} \in L^2(\mathbb{R}, d\nu)$;
- the Markov chain $(\hat{Y}^h_n)_{n=0,...,N}$ satisfies assumptions $\mathcal{H}_1$ and $\mathcal{H}_2$;
- $u \in C^{2,6}_{\text{pol,T}}(\mathbb{R}, \mathcal{D})$.

Then, there exist $\tilde{h}, C > 0$ such that for every $h < \tilde{h}$ and $\Delta x < 1$ one has

$$|u(0, \cdot, Y_0) - u_0^h(\cdot, Y_0)|_2 \leq CT(h + \Delta x^2).$$

**Proof.** The proof follows from Theorem 4.1 once we prove that Assumption $\mathcal{K}(2, 2\lambda c_y, h + \Delta x^2)$ holds.

First, Lemma 4.5 gives $|\Pi_n^h(x)|_2 \leq |(A_n^h)^{-1}(y)|_2|B_n^h(y)|_2 \leq 1 + 2\lambda c_y$, so (4.11) holds with $c(y) = 2\lambda c_y$. We prove now (4.13) with $p = 2$ and $\mathcal{E}(h, \Delta x) = h + \Delta x^2$. We first note that (4.12) equals to

$$\mathbb{E}\left[B_n^h(\hat{Y}^h_n) u((n + 1)h, \cdot, \hat{Y}^h_{n+1})(x) \mid \hat{Y}^h_n\right] = A_n^h(\hat{Y}^h_n) u(nh, \cdot, \hat{Y}^h_n)(x) + A_n^h(\hat{Y}^h_n) \mathcal{R}^h_n(\cdot, \hat{Y}^h_n)(x).$$

**Step 1. Taylor expansion of the l.h.s. of (4.27).** We set

$$I_1 = B_n^h(\hat{Y}^h_n) u((n + 1)h, \cdot, \hat{Y}^h_{n+1})(x_i)$$

$$= u((n + 1)h, x_i, \hat{Y}^h_{n+1}) + h \sum_i \nu\gamma_n^h(l\Delta x)\left(\sum_i u((n + 1)h, x_i + l\Delta x, \hat{Y}^h_{n+1}) - u((n + 1)h, x_i, \hat{Y}^h_{n+1})\right) \Delta x.$$ 

In the first term of the above r.h.s. we apply several Taylor’s expansion: of $t \mapsto u(t, x_i, \hat{Y}^h_n)$ around $nh$ up to order 1, of $y \mapsto u(nh, x_i, y)$ around $\hat{Y}^h_n$ up to order 3 and of $y \mapsto \partial_y u(nh, x_i, y)$ around $\hat{Y}^h_n$ up to order 1. Rearranging the terms we obtain

$$u((n + 1)h, x_i, \hat{Y}^h_{n+1}) = u(nh, x_i, \hat{Y}^h_n)$$

$$+ \partial_t u(nh, x_i, \hat{Y}^h_n) h + \partial_y u(nh, x_i, \hat{Y}^h_n)(\hat{Y}^h_{n+1} - \hat{Y}^h_n) + \frac{1}{2} \partial^2_y u(nh, x_i, \hat{Y}^h_n)(\hat{Y}^h_{n+1} - \hat{Y}^h_n)^2$$

$$+ \partial_y\partial_t u(nh, x_i, \hat{Y}^h_n) h(\hat{Y}^h_{n+1} - \hat{Y}^h_n) + \frac{1}{6} \partial^3_y u(nh, x_i, \hat{Y}^h_n)(\hat{Y}^h_{n+1} - \hat{Y}^h_n)^3$$

$$+ R_1(n, h, x_i, \hat{Y}^h_n, \hat{Y}^h_{n+1}),$$

where $R_1$ is given by

$$R_1(n, h, x_i, \hat{Y}^h_n, \hat{Y}^h_{n+1}) = h^2 \int_0^1 (1 - \tau) \partial^2_y u(nh + \tau h, x_i, \hat{Y}^h_{n+1}) d\tau$$

$$+ \frac{(\hat{Y}^h_{n+1} - \hat{Y}^h_n)^4}{6} \int_0^1 (1 - \tau)^3 \partial^4_y u(nh, x_i, \hat{Y}^h_n + \tau(\hat{Y}^h_{n+1} - \hat{Y}^h_n)) d\tau$$

$$+ h(\hat{Y}^h_{n+1} - \hat{Y}^h_n)^2 \int_0^1 (1 - \zeta) \partial_t \partial^3_y u(nh, x_i, \hat{Y}^h_n + \zeta(\hat{Y}^h_{n+1} - \hat{Y}^h_n)) d\zeta.$$ 

(4.29)
For the second term in the r.h.s. of (4.28), we stop the Taylor expansion of \( t \mapsto u((n + 1)h, x_i + l\Delta x, \hat{Y}_n^h) \) around \( nh \) at order 0 and of \( y \mapsto u(nh, x_i + l\Delta x, y) \) around \( \hat{Y}_n^h \) at order 1, obtaining

\[
\begin{align*}
& h \sum_l \nu_{h,n}^l (l\Delta x) \left[ u((n + 1)h, x_i + l\Delta x, \hat{Y}_n^{h+1}) - u((n + 1)h, x_i, \hat{Y}_n^h) \right] \Delta x \\
& = h \sum_l \nu_{h,n}^l (l\Delta x) \left[ u(nh, x_i + l\Delta x, \hat{Y}_n^h) - u(nh, x_i, \hat{Y}_n^h) \right] \Delta x \\
& + h \left( \hat{Y}_{n+1}^h - \hat{Y}_n^h \right) \sum_l \nu_{h,n}^l (l\Delta x) \left[ \partial_y u(nh, x_i + l\Delta x, \hat{Y}_n^h) - \partial_y u(nh, x_i, \hat{Y}_n^h) \right] \Delta x \\
& + R_2(n, h, x_i, \hat{Y}_n^h, \hat{Y}_{n+1}^h),
\end{align*}
\]

where \( R_2 \) contains the integral terms:

\[
\begin{align*}
R_2(n, h, x_i, \hat{Y}_n^h, \hat{Y}_{n+1}^h) &= h^2 \sum_l \nu_{h,n}^l (l\Delta x) \Delta x \times \\
& \times \int_0^1 (1 - \tau) \left[ \partial_t u(nh + \tau h, x_i + l\Delta x, \hat{Y}_n^h) - \partial_t u(nh, x_i, \hat{Y}_n^h) \right] d\tau \\
& + h \left( \hat{Y}_{n+1}^h - \hat{Y}_n^h \right)^2 \sum_l \nu_{h,n}^l (l\Delta x) \Delta x \times \\
& \times \int_0^1 (1 - \zeta) \left[ \partial_y u(nh, x_i + l\Delta x, \hat{Y}_n^h + \zeta (\hat{Y}_{n+1}^h - \hat{Y}_n^h)) - \partial_y u(nh, x_i, \hat{Y}_n^h) \right] d\zeta.
\end{align*}
\]

By resuming, we obtain

\[
I_1 = u(nh, x_i, \hat{Y}_n^h) + \partial_t u(nh, x_i, \hat{Y}_n^h)h + \partial_y u(nh, x_i, \hat{Y}_n^h)(\hat{Y}_{n+1}^h - \hat{Y}_n^h) \\
+ \frac{1}{2} \partial_y^2 u(nh, x_i, \hat{Y}_n^h)(\hat{Y}_{n+1}^h - \hat{Y}_n^h)^2 \\
+ h \Delta x \sum_l \nu_{h,n}^l (l\Delta x) \left[ u(nh, x_i + l\Delta x, \hat{Y}_n^h) - u(nh, x_i, \hat{Y}_n^h) \right] \\
+ \sum_{i=1}^2 R_i(n, h, x_i, \hat{Y}_n^h, \hat{Y}_{n+1}^h) + S(n, h, x_i, \hat{Y}_n^h, \hat{Y}_{n+1}^h),
\]

where

\[
\begin{align*}
S(n, h, x_i, \hat{Y}_n^h, \hat{Y}_{n+1}^h) &= \partial_y \partial_t u(nh, x_i, \hat{Y}_n^h) h(\hat{Y}_{n+1}^h - \hat{Y}_n^h) + \frac{1}{6} \partial_y^3 u(nh, x_i, \hat{Y}_n^h)(\hat{Y}_{n+1}^h - \hat{Y}_n^h)^3 \\
& + h \left( \hat{Y}_{n+1}^h - \hat{Y}_n^h \right) \sum_l \nu_{h,n}^l (l\Delta x) \left[ \partial_y u(nh, x_i + l\Delta x, \hat{Y}_n^h) - \partial_y u(nh, x_i, \hat{Y}_n^h) \right] \Delta x.
\end{align*}
\]

**Step 2.** Taylor expansion of the first addendum in the r.h.s. of (4.27). We set

\[
\begin{align*}
I_2 &= A \Delta x^2 u(nh, x_i, \hat{Y}_n^h)(x_i) \\
& = (\alpha \Delta x(\hat{Y}_n^h) - \beta \Delta x(\hat{Y}_n^h))u(nh, x_{i-1}, \hat{Y}_n^h) \\
& + (1 + 2\beta \Delta x(\hat{Y}_n^h))u(nh, x_i, \hat{Y}_n^h) - (\alpha \Delta x(\hat{Y}_n^h) + \beta \Delta x(\hat{Y}_n^h))u(nh, x_{i+1}, \hat{Y}_n^h).
\end{align*}
\]

We expand with Taylor \( x \mapsto u(nh, x, \hat{Y}_n^h) \) around \( x_i \) up to order 3 and we insert the values of \( \alpha \Delta x \) and \( \beta \Delta x \) in (4.24). Rearranging the terms we get

\[
I_2 = u(nh, x_i, \hat{Y}_n^h) - h\mu_X(\hat{Y}_n^h)\partial_x u(nh, x_i, \hat{Y}_n^h) - \frac{1}{2} h\sigma_X(\hat{Y}_n^h) \partial_x^2 u(nh, x_i, \hat{Y}_n^h) \\
+ R_3(n, h, x_i, \hat{Y}_n^h, \hat{Y}_{n+1}^h),
\]

where

\[
R_3(n, h, x_i, \hat{Y}_n^h, \hat{Y}_{n+1}^h) = \frac{\alpha \Delta x(\hat{Y}_n^h) - \beta \Delta x(\hat{Y}_n^h)}{12} h \Delta x^2 \int_0^1 (1 - \eta)^3 \partial_x^4 u(nh, x_i - \eta \Delta x, \hat{Y}_n^h) d\eta \\
- \frac{\alpha \Delta x(\hat{Y}_n^h) + \beta \Delta x(\hat{Y}_n^h)}{12} h \Delta x^2 \int_0^1 (1 - \eta)^3 \partial_x^4 u(nh, x_i + \eta \Delta x, \hat{Y}_n^h) d\eta \\
- \frac{1}{2} h \Delta x^2 \mu_X(\hat{Y}_n^h) \partial_x^2 u(nh, x_i, \hat{Y}_n^h).
\]

**Step 3.** Rearranging the terms. By resuming, from (4.31) and (4.33) we have

\[
I_1 - I_2 \\
= h \partial_t u(nh, x_i, \hat{Y}_n^h) + (\hat{Y}_{n+1}^h - \hat{Y}_n^h) \partial_y u(nh, x_i, \hat{Y}_n^h) + h\mu_X(\hat{Y}_n^h)\partial_x u(nh, x_i, \hat{Y}_n^h) \\
+ \frac{1}{2} \left[ (\hat{Y}_n^{h+1} - \hat{Y}_n^h)^2 \partial_x^2 u(nh, x_i, \hat{Y}_n^h) + h \sigma_X(\hat{Y}_n^h) \partial_x^2 u(nh, x_i, \hat{Y}_n^h) \right] \\
+ h \int (u(t, x + \gamma \hat{Y}_n^h, \zeta, \hat{Y}_n^h) - u(t, x, \hat{Y}_n^h)) \nu(\zeta) d\zeta \\
+ \sum_{i=1}^2 R_i(n, h, x_i, \hat{Y}_n^h, \hat{Y}_{n+1}^h) + S(n, h, \hat{Y}_n^h, \hat{Y}_{n+1}^h),
\]
in which we have used the change of variable giving
\[
\int (u(t, x + z, \hat{Y}_n^h) - u(t, x, \hat{Y}_n^h)) \nu_{\hat{Y}_n^h}(z) dz = \int (u(t, x + \gamma x (\hat{Y}_n^h) \zeta, \hat{Y}_n^h) - u(t, x, \hat{Y}_n^h)) \nu(\zeta) d\zeta
\]
and where
\[
R_4(n, h, x_i, \hat{Y}_n^h) = h \sum_{i=1}^4 \left[ u(t, x_i + l \Delta x, \hat{Y}_n^h) - u(t, x_i, \hat{Y}_n^h) \right] \nu_{\hat{Y}_n^h}(l \Delta x) \Delta x
\]
By passing to the conditional expectation and by using formulas (3.5), (3.6) and (3.7) for the local moments of order 1, 2 and 3, we obtain
\[
\tilde{R}_n^h(x_i, \hat{Y}_n^h) := \mathbb{E}[I_1 - I_2 \mid \hat{Y}_n^h] = h(\partial_t u(nh, x_i, \hat{Y}_n^h) + Lu(nh, x_i, \hat{Y}_n^h))
\]
Here we have used the following facts: \( u \) solves (4.27); \( \mathbb{E}[S(n, h, x_i, \hat{Y}_n^{h+1}) \mid \hat{Y}_n^h] = \sum_{i=1}^2 S_i(n, h, x_i, \hat{Y}_n^h) \), with (recall the definition of \( S \) in (4.32) and of the local moments \( f_h, g_h \) and \( j_h \) in (3.5), (3.6) and (3.7))
\[
S_1(n, h, x_i, \hat{Y}_n^h) = f_h(\hat{Y}_n^h) \partial_y u(nh, x_i, \hat{Y}_n^h) + \frac{1}{6} g_h(\hat{Y}_n^h) \partial_{y}^3 u(nh, x_i, \hat{Y}_n^h) + \frac{1}{6} j_h(\hat{Y}_n^h) \partial_{y}^3 u(nh, x_i, \hat{Y}_n^h) + \partial_y \partial_t u(nh, x_i, \hat{Y}_n^h) (h(\mu(y) \partial_t (\hat{Y}_n^h) + f_h(\hat{Y}_n^h)),
\]
\[
S_2(n, h, x_i, \hat{Y}_n^h) = h(h \mu(y) \partial_t (\hat{Y}_n^h) + f_h(\hat{Y}_n^h)) \times \sum_{i=1}^2 \nu_{\hat{Y}_n^h}(l \Delta x) [\partial_y u(nh, x_i + l \Delta x, \hat{Y}_n^h) - \partial_y u(nh, x_i, \hat{Y}_n^h)] \Delta x.
\]

**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 (4.27), we have to study \( \tilde{R}_n^h(\cdot, \hat{Y}_n^h) = (A_{\Delta x}^n)^{-1}(\hat{Y}_n^h) \tilde{R}_n^h(\cdot, \hat{Y}_n^h) \). By Lemma 4.5 it follows that \( \sup_{y \in \mathcal{D}} |(A_{\Delta x}^n)^{-1}(y)|_2 \leq 1 \), so
\[
(4.35)
\]
\[
\frac{1}{4} \sum_{i=1}^4 \mathbb{E}[\left| R_4(n, h, x_i, \hat{Y}_n^h) \right|^2] \leq C \sum_{i=1}^4 \mathbb{E}[\left| R_i(n, h, \cdot, \hat{Y}_n^h, \hat{Y}_n^{h+1}) \right|^2] + \sum_{i=1}^2 \mathbb{E}[\left| S_i(n, h, \cdot, \hat{Y}_n^h) \right|^2].
\]
Hence it suffices to prove that the above 6 terms are all upper bounded by \( Ch^2(h + \Delta x^2)^2 \). The inequalities studied in (ii) of Lemma 1.3 now come on.

Consider first \( R_1 \) in (4.29) and in particular, the first addendum therein. Set
\[
g_n(x) = h^2 \int_0^1 (1 - \tau) \partial_t^2 u(n \tau h, x, \hat{Y}_n^h) d\tau.
\]
Since \( u \in C_{pol,T}^1(\mathbb{R}, \mathcal{D}), \partial_x^k g_n \in L^2(\mathbb{R}, dx) \) for every \( k = 0, 1, 2 \) and \( |\partial_x^k g_n|_{L^2} \leq Ch^2(1 + |\hat{Y}_n^{h+1}|^a) \). So, by using (4.18),
\[
|g_n|_2^2 \leq Ch^2(1 + |\hat{Y}_n^{h+1}|^a)^2.
\]
Similar estimates hold for the other terms in \( R_1 \), so we can write
\[
|R_1(n, h, \cdot, \hat{Y}_n^h, \hat{Y}_n^{h+1})|_2 \leq C \left[ h^4(1 + |\hat{Y}_n^h|)^2 + |\hat{Y}_n^{h+1} - \hat{Y}_n| \right. \left. (1 + |\hat{Y}_n^h|)^2 + h^2 |\hat{Y}_n^{h+1} - \hat{Y}_n| (1 + |\hat{Y}_n^h|)^2 \right].
\]

19
By using the increment estimates (3.11), the moment estimates (3.12) and the Cauchy-Schwartz inequality, we obtain
\[ \mathbb{E}[|R_1(n,h,\cdot,\hat{Y}_n^h,\hat{Y}_{n+1}^h)|^2] \leq Ch^4. \]

The same arguments can be developed for \( R_3 \) in (4.34) and \( S_1 \) in (4.36). These give
\[ \mathbb{E}[|R_3(n,h,\cdot,\hat{Y}_n^h,\hat{Y}_{n+1}^h)|^2] \leq Ch^2 \Delta x^4 \quad \text{and} \quad \mathbb{E}[|S_1(n,h,\cdot,\hat{Y}_n^h)|^2] \leq Ch^4. \]

In order to study \( R_2 \) in (4.30), consider the first term and set
\[ g_n(x) = h^2 \sum \nu_{Y_n^h}(l\Delta x)\Delta x \times \int_0^1 (1-\tau)[\partial_t u(nh+\tau h, x+l\Delta x, \hat{Y}_{n+1}^h) - \partial_t u(nh+\tau h, x, \hat{Y}_{n+1}^h)] d\tau. \]

We notice that \( g_n \in C^2 \). By the Cauchy-Schwarz inequality for the (discrete) finite measure \( \nu_{Y_n^h}(l\Delta x)\Delta x \), \( l \in \mathbb{Z} \), we have
\[
|\partial_x g_n(x)|^2 \leq Ch^4 \sum \nu_{Y_n^h}(l\Delta x)\Delta x \times \int_0^1 (1-\tau)^2 \left( |\partial_x^2 \partial_t u(nh+\tau h, x+l\Delta x, \hat{Y}_{n+1}^h)|^2 + |\partial_x^2 \partial_t u(nh+\tau h, x, \hat{Y}_{n+1}^h)|^2 \right) d\tau.
\]

This gives \( |\partial_x g_n|_{L^2} \leq Ch^2 (1+|\hat{Y}_{n+1}^h|^a) \) and, by (4.15), \( |g_n|^2 \leq Ch^4 (1+|\hat{Y}_{n+1}^h|^a)^2 \). By developing the same arguments to the other terms in \( R_2 \), we obtain
\[ |R_2(n,h,\cdot,\hat{Y}_n^h,\hat{Y}_{n+1}^h)|^2 \leq C \left[ h^4 (1+|\hat{Y}_n^h|^a)^2 + h^2 |\hat{Y}_{n+1} - \hat{Y}_n|^4 (1+|\hat{Y}_n|^a) \right]. \]

And by passing to the expectation, we get \( \mathbb{E}[|R_2(n,h,\cdot,\hat{Y}_n^h,\hat{Y}_{n+1}^h)|^2] \leq Ch^4 \). A similar approach can be used to handle \( R_4 \) in (4.35) and in \( S_2 \) in (4.37), giving
\[ \mathbb{E}[|R_4(n,h,\cdot,\hat{Y}_n^h)|^2] \leq Ch^2 \Delta x^4 \quad \text{and} \quad \mathbb{E}[|S_2(n,h,\cdot,\hat{Y}_n^h)|^2] \leq Ch^4. \]

The proof is now completed.

\[ \square \]

### 4.3.2 Convergence in \( l_\infty \)-norm

We consider here a different finite difference scheme for equation (4.8): we still approximate (explicit in time) the integral term \( \mathcal{L}_{int}^y v \) in (4.21) with a trapezoidal rule, but we use an upwind first order scheme to approximate (implicit in time) the differential part \( \mathcal{L}_{diff}^y v \) in (4.20). 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. Hence, the resulting scheme is
\[ A_{\Delta x}^h(y) v^n = B_{\Delta x}^h(y) v^{n+1}, \quad (4.38) \]
where \( A_{\Delta x}^h(y) \) is the linear operator given by
\[
(A_{\Delta x}^h)_{ij}(y) = \begin{cases} 
-\beta_{\Delta x}^h(y) - |\alpha_{\Delta x}^h(y)| 1_{\alpha_{\Delta x}^h(y)<0}, & \text{if } i = j + 1, \\
1 + 2\beta_{\Delta x}^h(y) + |\alpha_{\Delta x}^h(y)|, & \text{if } i = j, \\
-\beta_{\Delta x}^h(y) - |\alpha_{\Delta x}^h(y)| 1_{\alpha_{\Delta x}^h(y)>0}, & \text{if } i = j - 1, \\
0, & \text{if } |i-j| > 1
\end{cases}
\]
(4.39)

with
\[
\alpha_{\Delta x}^h(y) = \frac{h}{\Delta x} \mu_X(y), \quad \beta_{\Delta x}^h(y) = \frac{h}{2\Delta x^2} \sigma_X^2(y),
\]
and \( B_{\Delta x}^h(y) \) is the linear operator defined in (4.25). Then we have:
Lemma 4.7. For every $y \in \mathcal{D}$, the operator $A^h_{\Delta x}(y) : l_\infty(\mathcal{X}) \rightarrow l_\infty(\mathcal{X})$ is invertible and
\[
\sup_{y \in \mathcal{D}} |(A^h_{\Delta x})^{-1}(y)|_{\infty} \leq 1
\]
And if $\nu', \nu'' \in L^1(\mathbb{R}, d\nu)$ then $\sup_{y \in \mathcal{D}} |B^h_{\Delta x}(y)|_{\infty} \leq 1 + 2\lambda c_\nu$, $c_\nu$ being defined in (4.16).

Proof. We write $A^h_{\Delta x}(y) = (1 + \eta(y))Id$ if $y_h(x) + \alpha^h_{\Delta x}(x) \geq 0$ and $P_{ij}(y) = 0$ if $|i-j| \neq 1$ and $P_{ij} = -(A^h_{\Delta x})_{ij}$ if $|i-j| = 1$. It easily follows that $|P(y)|_{\infty} \leq \eta(y)$. Moreover, it is easy to see that the operator $A^h_{\Delta x}(y) : l_\infty(\mathcal{X}) \rightarrow l_\infty(\mathcal{X})$ is invertible with inverse
\[
(A^h_{\Delta x})^{-1}(y) = ((1 + \eta(y))Id - P(y))^{-1} = \frac{1}{1 + \eta(y)} \sum_{k=0}^{\infty} P(y)^k.
\]
This gives $|(A^h_{\Delta x})^{-1}(y)|_{\infty} \leq 1$. The assertion for $B^h_{\Delta x}(y)$ follows from (4.25) and (4.16). \qed

We can now state the convergence result in $l_\infty(\mathcal{X})$.

Theorem 4.8. Let $u$ be defined in (4.4) and $(u^h_n)_{n=0,\ldots,N}$ be given by (4.10) 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 (4.39) and (4.25) respectively. Assume that:

- $\nu', \nu'' \in L^1(\mathbb{R}, d\nu)$;
- the Markov chain $(\hat{Y}^h_n)_{n=0,\ldots,N}$ satisfies assumptions $H_1$ and $H_2$;
- $u \in C^{\infty,4}_{pol,T}(\mathbb{R}, \mathcal{D})$.

Then, there exist $\tilde{h}, C > 0$ such that for every $h < \tilde{h}$ and $\Delta x < 1$ one has
\[
|u(0, \cdot, Y_0) - u^h_0(\cdot, Y_0)|_{\infty} \leq C(h + \Delta x).
\]

Proof. The statement follows by applying Theorem 4.1 once it is proved that $K(\infty, 2\lambda c_\nu, h + \Delta x)$ holds. This is just a rewriting of the proof of Theorem 4.6 in terms of the norm in $l_\infty(\mathcal{X})$. We only notice that, for handling the remaining terms, in $l_\infty$-norm we do not need to apply (4.18), so we do not need more regularity for $u$. That’s why the class $C^{\infty,4}_{pol,T}(\mathbb{R}, \mathcal{D})$ is enough. \qed

5 The hybrid procedure for the Heston or Bates model

As an application in finance, we consider the Heston [22] and the Bates [10] model. In this framework, $u(t, x, y)$ is in fact related to the value function at time $t$ of a European option with maturity $T$ and (discounted) payoff $f$.

Recall that under the Heston or Bates model, the asset price process $S$ and the volatility process $Y$ evolve following the stochastic differential system
\[
\begin{align*}
\frac{dS_t}{S_t} &= (r - \delta)dt + \mu \sqrt{V_t} dZ^1_t + \gamma d\tilde{H}_t, \\
dY_t &= \kappa(\theta - Y_t)dt + \sigma \sqrt{V_t} dZ^2_t,
\end{align*}
\]
where $S_0 > 0$, $Y_0 > 0$, $Z = (Z^1, Z^2)$ 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 $\lambda$ and i.i.d. jumps $\{J_k\}_k$ as in (4.2). Here, $\gamma = 1$.
(Bates model) or $\gamma = 0$ (Heston model). $r$ and $\delta$ are the interest rate and the dividend interest rate respectively. We assume, as usual, that the Poisson process $K$, the jump amplitudes $\{J_k\}_k$ and the correlated Brownian motion $(Z^1, Z^2)$ are independent.

With a simple transformation, we can reduce the model (5.1) to our reference model (4.1). To get rid of the correlated Brownian motion, we set $\tilde{\rho} = \sqrt{1 - \rho^2}$, $Z^2 = W$ and $Z^1 = \rho Z^2 + \tilde{\rho} B$, $(B, W)$ denoting a standard 2-dimensional Brownian motion. Moreover, considering the process $X_t = \log S_t - 2 W_t$, the pair $(X, Y)$ satisfies

$$
dX_t = \mu X(Y_t) dt + \tilde{\rho} \sqrt{Y_t} dB_t + \gamma dH_t,
\quad dY_t = \kappa (\theta - Y_t) dt + \sigma \sqrt{Y_t} dW_t,
$$

(5.2)

where $\mu X(y) = r - \delta - \frac{\rho^2}{2} - \frac{\rho}{2} \kappa (\theta - y)$, $H_t$ is the compound Poisson process written through the Poisson process $K$, with intensity $\lambda$, and the i.i.d. jumps $J_k = \log (1 + J_k)$. 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 $\nu$ satisfies the requests in Theorem 4.6 or Theorem 4.8. For example, these properties hold for the mixture of exponential laws used by Kou.

We consider the approximating Markov chain for the CIR process discussed in Section 3.1 and the two possible finite difference operator discussed in sections 4.3.1 and 4.3.2. As an application of Theorem 4.6 and Theorem 4.8, we get the following convergence rate result of the hybrid method.

**Theorem 5.1.** Let $(X, Y)$ be the solution to (5.2) and let $(\tilde{Y}^h_n)_{n=0,...,N}$ be the Markov chain introduced in Section 5.1 for the approximation of the CIR process $Y$. Let $u(t, x, y) = \mathbb{E}(f(X^h_t, Y^h_t))$ be as in (4.1) and $(u^h_n)_{n=0,...,N}$ be given by (4.10) 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 (4.23) and (4.25) respectively;
- $\nu^x, \nu^y \in L^2(\mathbb{R}, d\nu)$ and $\nu$ has finite moments of any order;
- $\partial^2_j f \in C^{2,6-j}_{\text{pol}}(\mathbb{R}, \mathbb{R}_+)$ for every $j = 0, \ldots, 6$.

Then, there exist $\bar{h}, C > 0$ such that for every $h < \bar{h}$ and $\Delta x < 1$ one has

$$
|u(0, \cdot, Y_0) - u^h_0(\cdot, Y_0)|_2 \leq CT(h + \Delta x^2).
$$

(ii) [Convergence in $l_\infty(\mathcal{X})$] Suppose that

- $A^h_{\Delta x}(y)$ and $B^h_{\Delta x}(y)$ are defined in (4.39) and (4.25) respectively;
- $\nu^x, \nu^y \in L^1(\mathbb{R}, d\nu)$ and $\nu$ has finite moments of any order;
- $\partial^2_j f \in C^{\infty,4-j}_{\text{pol}}(\mathbb{R}, \mathbb{R}_+)$ for every $j = 0, \ldots, 4$.

Then, there exist $\bar{h}, C > 0$ such that for every $h < \bar{h}$ and $\Delta x < 1$ one has

$$
|u(0, \cdot, Y_0) - u^h_0(\cdot, Y_0)|_\infty \leq CT(h + \Delta x).
$$

**Proof.** We apply Theorem 4.6 for (i) and Theorem 4.8 for (ii). Following Theorem 3.1, the assumptions $\mathcal{H}_1$ and $\mathcal{H}_2$ hold (see also Proposition 3.1). So, we need only to prove that if $\partial^2_j f \in C^{2,6-j}_{\text{pol}}(\mathbb{R}, \mathbb{R}_+)$ as $j = 0, 1, \ldots, 6$, resp. $\partial^2_j f \in C^{\infty,4-j}_{\text{pol}}(\mathbb{R}, \mathbb{R}_+)$ as $j = 0, 1, \ldots, 4$, then $u \in C^{2,6}_{\text{pol}, T}(\mathbb{R}, \mathbb{R}_+)$, resp. $u \in C^{\infty,4}_{\text{pol}, T}(\mathbb{R}, \mathbb{R}_+)$. This is proved in next Proposition 5.3 (set $\rho = 0$, $a = r - \delta - \frac{\rho}{2} \kappa$, and $b = \frac{\rho}{2} \kappa - \frac{\rho}{2}$ therein), the whole Section 5.1 being devoted to.
Remark 5.2. Another example of interest in finance is the Bates-Hull-White model \cite{14}, which is a Bates model coupled with a stochastic interest rate. The dynamics follows (5.1) in which \( r \) is not constant but given by the Vasicek model

\[
dr_t = \kappa_r (\theta_r - r_t) dt + \sigma r dZ_t^3,
\]

\( Z^3 \) 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 \cite{14} (Section 4.1), acting on each time interval \([nh, (n + 1)h]\).

5.1 A regularity result for the Heston PDE/Bates PIDE

We deal here with a slightly more general model: we consider the SDE

\[
\begin{align*}
\frac{dX_t}{a + b Y_t} & dt + \sqrt{1} dW^1_t + \gamma X dH_t, \\
\frac{dY_t}{\kappa (\theta - Y_t)} & dt + \sigma \sqrt{Y_t} dW^2_t,
\end{align*}
\]

(5.3)

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 \( \lambda \) and Lévy measure \( \nu \), which is assumed hereafter to have finite moments of any order. Here, \( a, b, \kappa, \theta \) denote constant parameters. Note that when \( a = r - \delta, b = -\frac{1}{\kappa} \) and \( \gamma = 0 \) (resp. \( \gamma = 1 \)), then \( (X, Y) \) is the standard Heston (resp. Bates) model for the log-price and volatility. When instead \( \rho = 0, a = r - \delta - \frac{1}{\kappa} \), we recover the equation (5.2) discussed in Theorem 5.1.

Let \( \mathcal{L} \) denote the infinitesimal generator associated to (5.3), that is,

\[
\mathcal{L}u = \frac{3}{2} (\partial_x^2 u + 2 \rho \sigma \partial_x \partial_y u + \sigma^2 \partial_y^2 u) + (a + b y) \partial_x u + \kappa (\theta - y) \partial_y u + \mathcal{L}_{\text{int}} u,
\]

(5.4)

where, hereafter, we set \( \mathcal{L}_{\text{int}} u(t, x, y) = \int [u(t, x + \gamma y \zeta, y) - u(t, x, y)] \nu(\zeta) d\zeta \).

So, the present section is devoted to the proof of the following result.

Proposition 5.3. Let \( p \in [1, \infty], q \in \mathbb{N} \) and suppose that \( \partial^{2j}_x f \in C^{p,q-j}_{\text{pol}}(\mathbb{R}, \mathbb{R}_+) \) for every \( j = 0, 1, \ldots, q \). Set

\[
\begin{align*}
\partial^n_x \partial^n_y u(t, x, y) = \mathbb{E} \left[ e^{-n(T-t)} \partial^n_x \partial^n_y f(X^n_{T,y}, Y^n_{T,x,y}) \right]
+ n \mathbb{E} \left[ \int_t^T \left[ \frac{1}{2} \partial_x^{n-2} \partial_y^{n-1} u + b \partial_x^{n+1} \partial_y^{n-1} u \right] (s, X^n_{s,t,x,y}, Y^n_{s,t,x,y}) ds \right],
\end{align*}
\]

(5.5)

where \( \partial_x^n \partial_y^{n-1} u := 0 \) when \( n = 0 \) and \( (X^n_{T,y}, Y^n_{T,x,y}), n \geq 0 \), denotes the solution starting from \( (x, y) \) at time \( t \) to the SDE (5.3) with parameters

\[
\begin{align*}
\rho_n = \rho, \quad a_n = a + n \rho \sigma, \quad b_n = b, \quad \kappa_n = \kappa, \quad \theta_n = \theta + \frac{n \gamma^2}{2 \kappa}, \quad \sigma_n = \sigma.
\end{align*}
\]

(5.6)

In particular, if \( q \geq 2 \) then \( u \in C^{1,2}([0, T] \times \bar{\mathcal{O}}), \bar{\mathcal{O}} = \mathbb{R} \times \mathbb{R}_+ \), solves the PIDE

\[
\begin{align*}
\begin{cases}
\partial_t u(t, x, y) + \mathcal{L} u(t, x, y) = 0, \quad (t, x, y) \in [0, T] \times \bar{\mathcal{O}}, \\
\partial_t u(T, x, y) = f(x, y), \quad (x, y) \in \mathcal{O}.
\end{cases}
\end{align*}
\]

(5.7)
Lemma 5.6. For the second property in (5.8), we refer, for example, to [2], whereas the first one follows from Proposition 4.1 of [2].

Remark 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 5.3 shows that the result holds also when one is not interested in the latter \(L^p\) condition. In this case, Proposition 5.3 reads: for \(q \in \mathbb{N}\), if \(\partial_x^{2j} f \in C^q_{\text{pol}}(\mathbb{R} \times \mathbb{R}_+)\) for every \(j = 0, 1, \ldots, q\) then \(u \in C^q_{\text{pol}, T}(\mathbb{R} \times \mathbb{R}_+)\). Moreover, the stochastic representation (5.7) holds and, if \(q \geq 2\), \(u\) solves PIDE (5.7).

As an immediate consequence of Proposition 5.3, we obtain the already known regularity result for the CIR process which has been already proved in Proposition 4.1 of [2].

Corollary 5.5. Assume that \(f = f(y)\) and set \(u(t, y) = \mathbb{E}[f(Y_T^{t,y})]\). If \(f \in C^q_{\text{pol}}(\mathbb{R}_+)\), then \(u \in C^q_{\text{pol}, T}(\mathbb{R}_+)\). Moreover, for \(n \leq q\),

\[
\partial^n_y u(t, y) = \mathbb{E} \left[ e^{-n\kappa(T-t)} \partial^n_y f(Y_T^{t,y}) \right],
\]

where \(Y_T^{t,y}\) denotes a CIR process starting from \(y\) at time \(t\) which solves the CIR dynamics with parameters \(\kappa = \kappa, \theta = \theta + \frac{na^2}{2\kappa}, \sigma = \sigma\). In particular, if \(q \geq 2\) then \(u \in C^q_{\text{pol}, T}(\mathbb{R}_+)\) solves the PDE

\[
\begin{cases}
  \partial_t u + Au = 0, & (t, y) \in [0, T] \times \mathbb{R}_+, \\
  u_n(T, y) = \partial^n_y f(y), & y \in \mathbb{R}_+
\end{cases}
\]

where \(A\) is the CIR infinitesimal generator given in (5.2).

We first need some preliminary results. First of all, recall that \(X\) and \(Y\) have uniformly bounded moments: for every \(T > 0\) and \(a \geq 1\) there exist \(A > 0\) such that for every \(t \in [0, T],\)

\[
\sup_{s \in [t, T]} \mathbb{E}[|X_s^{t,x,y}|^a] \leq A(1 + |x|^a + y^a) \quad \text{and} \quad \sup_{s \in [t, T]} \mathbb{E}[|Y_s^{t,y}|^a] \leq A(1 + y^a).
\]

(5.8)

For the second property in (5.8), we refer, for example, to [2], whereas the first one follows from standard techniques.

Lemma 5.6. Let \(p \in [0, \infty), g \in C^p \mathbb{R}^+(\mathbb{R}, \mathbb{R}_+), h \in C^p \mathbb{R}^+(\mathbb{R}, \mathbb{R}_+)\) and consider the function

\[
u(t, x, y) = \mathbb{E} \left[ e^{p(T-t)} g(X_T^{t,x,y}, Y_T^{t,y}) - \int_t^T e^{p(s-t)} h(s, X_s^{t,x,y}, Y_s^{t,y}) ds \right],
\]

(5.9)

where \(p \in \mathbb{R}\). Then \(u \in C^p \mathbb{R}^+(\mathbb{R}, \mathbb{R}_+)\).

Proof. We set

\[
u_1(t, x, y) = \mathbb{E} \left[ e^{p(T-t)} g(X_T^{t,x,y}, Y_T^{t,y}) \right], \quad \nu_2(t, x, y) = \mathbb{E} \left[ \int_t^T e^{p(s-t)} h(s, X_s^{t,x,y}, Y_s^{t,y}) ds \right]
\]

and we show that, for \(i = 1, 2, u_i \in C^p \mathbb{R}^+(\mathbb{R}, \mathbb{R}_+)\). 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 \(s \geq t_n \lor t\), \((X_s^{t_n,x_n,y_n}, Y_s^{t_n,y_n}) \to (X_s^{t,x,y}, Y_s^{t,y})\) in probability. We write \(u_2\) as

\[
u_2(t, x, y) = \int_0^T 1_{s > t} e^{p(s-t)} \mathbb{E} \left[ h(s, X_s^{t,x,y}, Y_s^{t,y}) \right] ds.
\]
Since $h$ is continuous, for $s > t_n \vee t$ the sequence $(h(s, X_{t_n}^{x,y}, Y_{t_n}^{x,y}))_n$ converges in probability to $h(s, X_t^{x,y}, Y_t^{x,y})$. By the polynomial growth of $h$ and (5.8), for $p > 1$ we have
\[
sup_n \mathbb{E}[|h(X_{t_n}^{x,y}, Y_{t_n}^{x,y})|^p] \leq C \sup_n \mathbb{E}[1 + |X_{t_n}^{x,y}|^p + (Y_{t_n}^{x,y})^p] < \infty. \tag{5.10}
\]
Thus, $(h(X_{t_n}^{x,y}, Y_{t_n}^{x,y}))_n$ is uniformly integrable, so $\{h(X_{t_n}^{x,y}, Y_{t_n}^{x,y})\}_n$ converges to $h(X_t^{x,y}, Y_t^{x,y})$ in $L^1(\Omega)$ and
\[
1_{s > t_n} \mathbb{E} \left[ e^{\theta(s-t)} h(s, X_{t_n}^{x,y}, Y_{t_n}^{x,y}) \right] \to 1_{s > t} \mathbb{E} \left[ e^{\theta(s-t)} h(s, X_t^{x,y}, Y_t^{x,y}) \right],
\]
a.e. $s \in [0, T]$. By (5.10), $u_2(t_n, x_n, y_n) \to u_2(t, x, y)$ thanks to the Lebesgue’s dominated convergence and moreover, $u_2$ grows polynomially. So, $u_2 \in C_{\text{pol}}(\mathbb{R} \times \mathbb{R}_+)$.

The case $p = \infty$ follows the same lines.

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 5.7.** Let $g \in C_{\text{pol}}(\bar{\mathcal{O}})$ and $h \in C_{\text{pol},T}(\bar{\mathcal{O}})$ 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 (5.9). Then, $u \in C([0, T] \times \bar{\mathcal{O}}) \cap C^{1,2}([0, T) \times \mathcal{O})$ and solves the PIDE
\[
\begin{aligned}
\partial_t u + Lu + gu &= h, \quad \text{in } [0, T) \times \mathcal{O}, \\
u(T, z) &= g(z), \quad \text{in } \mathcal{O}.
\end{aligned} \tag{5.11}
\]
Moreover, if the Feller condition $2\kappa \theta \geq \sigma^2$ holds then $u$ is the unique solution to (5.11) in the class $C_{\text{pol},T}(\bar{\mathcal{O}})$.

The proof employs standard techniques, see e.g. Proposition 3.2 in [20] with the use of classical results in parabolic PIDEs theory from [21, 29]. The uniqueness of the solution under the Feller condition follows from the fact that the CIR process never hits 0. So, we omit this proof.

**Lemma 5.8.** Let $u$ be defined in (5.9), with $g$ and $h$ such that, as $j = 0, 1$, $\partial^2_{x,y} g \in C^{1-j}_{\text{pol}}(\bar{\mathcal{O}})$ and $\partial^2_{x,y} h \in C^{1-j}_{\text{pol},T}(\bar{\mathcal{O}})$. Then $u \in C^1(\mathcal{O})$. Moreover, $\partial^2_{x,y} u \in C_{\text{pol},T}(\bar{\mathcal{O}})$ and one has
\[
\partial^m_{x,y} u(t, x, y) = \mathbb{E}^{t,x,y} \left[ e^{\theta(T-t)} \partial^m_{x,y} g(X_T, Y_T) - \int_t^T e^{\theta(s-t)} \partial^m_{x,y} h(s, X_s, Y_s) ds \right], \quad m = 1, 2, \tag{5.12}
\]
\[
\partial_y u(t, x, y) = \mathbb{E}^{t,x,y} \left[ e^{\theta(T-t)} \partial_y g(X_T, Y_T^*) + \int_t^T e^{\theta(T-s)} \partial_y h(s, X_s^*, Y_s^*) ds \right], \tag{5.13}
\]
where $(X_t^*, Y_t^*)$ solves (5.3) with new parameters $\rho_* = \rho, \ a_* = a + \rho \sigma, \ b_* = b, \ \kappa_* = \kappa, \ \theta_* = \theta + \frac{\sigma^2}{2\kappa}, \ \sigma_* = \sigma$.
Proof. First, the stochastic flow w.r.t. \( x \) is differentiable (here, \( (X^*)_s^{t,x,y} = x + Z^{t,y}_s \) and \( Z^{t,y}_s \) does not depend on \( x \)). Hence, by using the polynomial growth hypothesis, by \((5.9)\) one gets \((5.12)\). Let us prove \((5.13)\).

By Lemma \((5.7)\) \( u \) solves \((5.11)\). So, setting \( v = \partial_y u \), by derivating \((5.11)\) one has

\[
\begin{align*}
\partial_t v + \mathcal{L}_s v + g_s v &= h_s, \quad \text{in } [0, T) \times \mathcal{O}, \\
v(T, z) &= g_s(z), \quad \text{in } \mathcal{O}.
\end{align*}
\]

where \( \mathcal{L}_s \) is the infinitesimal generator of \((X^*, Y^*)\) and \( g_s = g - \kappa, \quad h_s = \partial_y h - b \partial_x u - \frac{1}{2} \partial^2_{yy} u, \quad g_s = \partial_y g. \)

By using \((5.12)\) and Lemma \((5.6)\) \( h_s \in C_{\text{pol}, T}(\mathcal{O}). \) Moreover, the Feller condition \( 2\kappa_\ast \theta_s \geq \sigma^2 \) holds, and by Lemma \((5.7)\) the unique solution with polynomial growth in \((x, y)\) to the above PIDE is

\[
\tilde{v}(t, x, y) = E_t^{t,x,y}[e^{\theta(T-t)}g(x_T, Y_T^*) - \int_t^T e^{\theta(s-t)}h(s, X^*_s, Y^*_s)ds].
\]

In order to identify \( \tilde{v} \) with \( v = \partial_y u \), one should know that \( \partial_y u \in C_{\text{pol}, T}(\mathcal{O}) \). 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 \([20]\).

For \( k \geq 1 \), let \( \varphi_k \) be a \( C^\infty(\mathbb{R}) \) approximation of \( \sqrt{|y|} \) such that \( \varphi_k(y) \geq 1/k, \quad \varphi_k(y) \to \sqrt{|y|} \) uniformly on the compact sets of \([0, +\infty)\) and \( \varphi_k \) 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

\[
\begin{align*}
\{ &dX^k_t = (a + bY^k_t) dt + \varphi_k(Y^k_t)dB_t + dH_t, \\
\} &dY^k_t = \kappa(\theta - Y^k_t)dt + \sigma \varphi_k(Y^k_t)dW_t,
\end{align*}
\]

whose generator is

\[
\mathcal{L}^k u = \frac{\varphi_k^2(y)}{2}(\partial^2_y u + 2\rho \partial_x \partial_y u + \sigma^2 \partial^2_y u) + (a + by) \partial_x u + \kappa(\theta - y) \partial_y u + \mathcal{L}_{\text{int}} u.
\]

Set

\[
\begin{align*}
U^k(t, x, y) &= E_t^{t,x,y}[e^{\theta(T-t)}g(X^k_T, Y^k_T) - \int_t^T e^{\theta(s-t)}h(s, X^k_s, Y^k_s)ds].
\end{align*}
\]

Let us first show that \( \partial_y U^k \in C_{\text{pol}, T}(\mathcal{O}). \) 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

\[
\begin{align*}
\partial_y U^k(t, x, y) &= \mathbb{E}[e^{\theta(T-t)} \langle \nabla_{x,y} g(X^{k,t,x,y}_T, Y^{k,t,x,y}_T, Z^{k,t,x,y}_T) \rangle] \\
&\quad - \int_t^T e^{\theta(s-t)} \mathbb{E}[\langle \nabla_{x,y} h(s, X^{k,t,x,y}_s, Y^{k,t,x,y}_s, Z^{k,t,x,y}_s) \rangle] ds.
\end{align*}
\]

The functions \( g, h \) and their derivatives have polynomial growth, so

\[
|\partial_y U^k(t, x, y)| \leq \mathbb{E}[C(1 + |X^{k,t,x,y}_T|^a + |Y^{k,t,x,y}_T|^a)|Z^{k,t,x,y}_T|] \\
+ \int_t^T e^{\theta(s-t)} \mathbb{E}[C(1 + |X^{k,t,x,y}_s|^a + |Y^{k,t,x,y}_s|^a)|Z^{k,t,x,y}_s|] ds
\]

and the usual \( L^p \)-estimates give

\[
\sup_{t \leq T} \left| \partial_y U^k(t, x, y) \right| \leq C_k(1 + |x|^a_k + y^{a_k})
\]

for suitable constants \( C_k, a_k > 0 \). Moreover, from the standard theory of parabolic PIDEs, \( u^k \) is a solution to

\[
\begin{align*}
\{ &\partial_t u^k + \mathcal{L}_s u^k + g u^k = h, \quad \text{in } [0, T) \times \mathcal{O}, \\
u^k(T, z) = g(z), \quad \text{in } \mathcal{O}.
\end{align*}
\]
By differentiating, \( v^k = \partial_y u^k \) solves the problem

\[
\begin{cases}
\partial_t v^k + L_{k,*} v^k + g_* u^k = h_{k,*}, & \text{in } [0,T) \times \mathcal{O}, \\
v^k(T, z) = g_*(z), & \text{in } \mathcal{O},
\end{cases}
\]

where

\[
L_{k,*} v = \frac{\sigma^2(y)}{2} \left( \partial_x^2 v + 2 \rho \sigma \partial_x \partial_y v + \sigma^2 \partial_y^2 v \right) + (a + by + 2 \rho \varphi_k \varphi'_k(y)) \partial_x v + (\kappa(\theta - y) + \sigma^2 \varphi_k \varphi'_k(y)) \partial_y v + L_{on} v
\]

and \( h_{k,*} = \partial_y h - b \partial_x u^k - \varphi_k \varphi'_k(y) \partial_x^2 u^k \). By developing the same arguments as before, we get \( h_{k,*} \in C_{\text{pol},T}(\mathcal{O}) \). The PIDE for \( v^k \) has a unique solution in \( C_{\text{pol},T}(\mathcal{O}) \) (recall that, by construction, the second order operator is uniformly elliptic). Thus, the Feynman-Kac formula gives

\[
\partial_y u^k(t, x, y) = \mathbb{E}^t, x, y \left[ e^{(T-t)} g_*(X_T^{k,*}, Y_T^{k,*}) - \int_t^T e^{(s-t)} h_{k,*}(s, X_s^{k,*}, Y_s^{k,*}) ds \right],
\]

where \((X^{k,*}, Y^{k,*})\) is the diffusion with infinitesimal generator given by \( L_{k,*} \). Now, the standard \( L^p \) estimates for \((X^{k,*}, Y^{k,*})\) hold uniformly in \( k \) (recall that \( \varphi_k \) 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 \( \sup_k \sup_{t \leq T} |u^k(t, x, y)| + \sup_k \sup_{t \leq T} |\partial_y u^k(t, x, y)| \leq C(1 + |x|^a + |y|^a) \), for suitable \( C, a > 0 \) (possibly different from the ones above). Moreover, the stability results in \([7]\) give \( \lim_{k \to \infty} u^k(t, x, y) = u(t, x, y) \) and \( \lim_{n \to \infty} \partial_y u^k(t, x, y) = v(t, x, y) \) 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

\[
\int_v(t, x, y) dxdy = \int \lim_k \partial_y u^k(t, x, y) \phi(x, y) dxdy = - \int \lim_k u^k(t, x, y) \partial_y \phi(x, y) dxdy = - \int u(t, x, y) \partial_y \phi(x, y) dxdy.
\]

Therefore, \( v(t, x, y) = \partial_y u(t, x, y) \) in \([0, T) \times \mathcal{O}\).  

We can now prove the result which this section is devoted to.

**Proof of Proposition [5.5]** We follow an induction on \( q \). If \( q = 0 \), Lemma [5.6] gives the result. Suppose the statement is true up to \( q - 1 \geq 0 \) and let us prove it for \( q \).

Take \( f \) such that \( \partial_x^j f \in C^{\rho, q-j}(\mathbb{R}, \mathbb{R}_+) \) for every \( j = 0, 1, \ldots, q \). Then, by induction, \( \partial_x^l \partial_x^m \partial_y^n u \in C^{\rho, 0}(\mathbb{R}, \mathbb{R}_+) \) when \( 2l + m + n \leq q - 1 \). So, we just need to prove that \( \partial_x^l \partial_x^m \partial_y^n u \in C^{\rho, 0}(\mathbb{R}, \mathbb{R}_+) \) for any \( l, m, n \) such that \( 2l + m + n = q \).

Assume first \( l = 0 \). For \( n = 0 \), we use that \( X_T^{l, x, y} = x + Z_T^{l, x, y} \) and we get \( \partial_x^m \partial_y^n u(t, x, y) = \mathbb{E}^t, x, y \left[ \partial_x^m \partial_y^n f(X_T, Y_T) \right] \). Since \( \partial_x^m f \in C^{\rho, 0}(\mathbb{R}, \mathbb{R}_+) \) for any \( m \leq 2q \), by Lemma 5.6 we obtain \( \partial_x^n u \in C^{\rho, 0, 0}(\mathbb{R}, \mathbb{R}_+) \) for every \( m \leq 2q \).

Fix now \( n > 0 \) and \( m \geq 0 \). Recursively applying Lemma 5.8, we get formula (5.5). Let us stress that, because of the presence of the derivatives \( \partial_x^{m+2} \partial_y^{-1} u \) and \( \partial_x^{m+1} \partial_y^{-1} u \) in (5.5), the recursively application of Lemma 5.8 gives the constraint \( m + 2n \leq q \). Then, by Lemma 5.6, it follows that \( \partial_x^n \partial_y^m u \in C^{\rho, 0}(\mathbb{R}, \mathbb{R}_+) \) 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 (5.6), Lemma 5.4 ensures that if \( m + 2n \leq 2q \) then \( u_{m,n} = \partial_x^m \partial_y^n u \) solves

\[
\begin{cases}
\partial_t u_{m,n} + L_n u_{m,n} - n ku_{m,n} = -n [\frac{1}{2} u_{m+2,n-1} + bu_{m+1,n-1}], & \text{in } [0, T) \times \mathcal{O}, \\
u_{m,n}(T, x, y) = \partial_x^n \partial_y^m f(x, y), & \text{in } \mathcal{O},
\end{cases}
\]

(5.15)
where \( L_n \) is the generator in (5.4) with the (new) parameters in (5.6). Therefore, the general case concerning \( \partial_t \partial_x^m \partial_y^n u \) with \( 2l + m + n = q \) follows by an iteration on \( l \): by (5.15),
\[
\partial_t \partial_x^m \partial_y^n u = -L_n \partial_t^{l-1} \partial_x^m \partial_y^n u + n \kappa \partial_t^{l-1} \partial_x^m \partial_y^n u - n \left[ \frac{1}{2} \partial_t^{l-1} \partial_x^m \partial_y^{n+2} \partial_y^{n-2} u + b \partial_t^{l-1} \partial_x^m \partial_y^{n-1} u \right].
\]

\( \square \)

\section*{A Lattice properties of the CIR approximating tree}

The aim of this section is to prove Proposition 3.3. For later use, let us first give some (trivial) properties of the lattice. First, by construction, \( k_d(n, k) \leq k < k_u(n, k) \), so that \( y_{k_d(n,k)}^{n+1} \leq y_k^{n+1} \leq y_k^n \leq y_{k+1}^n \leq y_{k_u(n,k)}^n \). Moreover for every \( n \) and \( k \), it is easy to see that
\[
y_k^n \leq y_k^{n+1}, \quad y_k^{n+1} \leq y_k^{n} \leq y_{k+1}^{n+1},
\]
\[
y_k^n \leq y_{k-1}^{n} + \sigma^2 h + 2\sigma \sqrt{y_{k-1}^{n} h}, \quad y_k^{n+1} \leq y_k^{n} + \frac{\sigma^2}{4} h - \sigma \sqrt{y_k^n h}. \tag{A.1}
\]

\textit{Proof of Proposition 3.3.} 1. The statement is an immediate consequence of the following facts:
\[
\text{if } k_u(n, k) \geq k + 2, \text{ then } y_k^n \leq \theta^*_h, \tag{A.2}
\]
\[
\text{if } k_d(n, k) \leq k - 1, \text{ then } y_k^n > \theta^*_h, \tag{A.3}
\]
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 \( h = 1/\kappa \), one has \( y_k^n + \mu_Y(y_k^n)h > 0 \). Moreover, as a direct consequence of (3.16) and (3.17) and of (A.1), we have that, if \( \mu_Y(y_k^n)h > 0 \), then \( k_d(n, k) = k \), and if \( \mu_Y(y_k^n)h < 0 \), then \( k_u(n, k) = k + 1 \).

Concerning (A.2), we obviously assume \( y_k^n > 0 \), so that \( y_{k+1}^{n+1} > 0 \). Now that, from (3.16),
\[
y_k^n + \mu_Y(y_k^n)h > y_{k_d(n,k)-1}^{n+1} \geq y_{k+1}^{n+1} = y_k^n + \frac{\sigma^2}{4} h + \sigma \sqrt{y_k^n h}.
\]

Since \( \mu_Y(y_k^n)h \leq \kappa \theta h \), we get \( \kappa \theta h > \frac{\sigma^2}{4} h + \sigma \sqrt{y_k^n h} \geq \sigma \sqrt{y_k^n h} \), from which \( y_k^n < \left( \frac{\kappa \theta}{\sigma} \right)^2 \) and (A.2) holds.

We prove now (A.3). First of all observe that, if \( y_k^n \leq \theta^*_h \), then \( \mu_Y(y_k^n)h > 0 \) and so \( k_d(n, k) = k \). Then we have \( y_k^n > \theta^*_h \) and from (3.15) we can assume \( y_{k+1}^{n+1} > 0 \) up to take \( h < (2\sqrt{\theta^*_h}/\sigma)^2 \). Now, by (3.17) we get
\[
y_k^n + \mu_Y(y_k^n)h < y_{k_d(n,k)+1}^{n+1} \leq y_{k+1}^{n+1} = y_k^n + \frac{\sigma^2}{4} h - \sigma \sqrt{y_k^n h},
\]
so that
\[
\kappa(\theta - y_k^n)h < \frac{\sigma^2}{4} h - \sigma \sqrt{y_k^n h}.
\]

This gives \( \kappa y_k^n h > \sigma \sqrt{y_k^n h} - \frac{\sigma^2}{4} h + \kappa \theta h \) and, for \( h \) small enough, \( y_k^n h > \frac{\sigma^2}{4} \), that is, (A.3) holds.

2. If \( y_k^n \leq \theta^*_h \), (A.3) 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 (3.19), 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
\[
y_{k_u(n,k)}^{n+1} - y_k^n = \frac{\sigma^2}{4} h + \sigma \sqrt{y_k^n h} \leq \left( \frac{\sigma^2}{4} + \sigma \sqrt{\theta^*_h} \right) h \leq C \theta^*_h,
\]
and the statement holds. If instead \( k_u(n,k) \geq k+2 \), then by (3.10) we have

\[
y_{k_u(n,k)}^{n+1} - y_{k_u(n,k)}^n < \mu y(y_k^n)h.
\]

We apply the third inequality in (A.1) (with \( n \) replaced by \( n+1 \) and \( k = k_u(n,k) \)) and we get

\[
0 \leq y_{k_u(n,k)}^{n+1} - y_{k_u(n,k)}^n \leq y_{k_u(n,k)}^{n+1} - 2\sigma \sqrt{y_{k_u(n,k)}^{n+1} - y_{k_u(n,k)}^n} + \sigma^2 h - y_{k_u(n,k)}^n
\]

\[
\leq \mu(y_k^n)h + 2\sigma \sqrt{y_k^n + \mu y(y_k^n)h}h + \sigma^2 h \leq (\kappa \theta + 2\sigma \sqrt{\theta* + \kappa \theta} + \sigma^2)h \leq C_* h.
\]

3. The statement follows from (A.2).

4. Formula (3.20) is proved once we show that the sets \( K_u(n,k) = \{k^* : k+1 \leq k^* \leq n+1 \text{ and } y_k^0 + \mu y(y_k^n)h \leq y_{k_u(n,k)}^{n+1} \} \) and \( K_d(n,k) = \{k^* : 0 \leq k^* \leq k \text{ and } y_k^0 + \mu y(y_k^n)h \geq y_{k_d(n,k)}^{n+1} \} \) are nonempty. Indeed, if \( y_k^0 > \theta_* h \) then \( k_u = k+1 \), so \( K_u(n,k) \neq \emptyset \). And if \( y_k^0 < \theta_* h \),

\[
y_{n+1}^{n+1} - y_k^n - \mu y(y_k^n)h \geq y_0 - \theta_* h - \kappa \theta h = y_0 - (\theta_* + \kappa \theta)h > 0
\]

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 \( y_k^0 < \theta_* h \) then \( k_d(n,k) = k \) by Proposition 3.3 so that \( K_d(n,k) \neq \emptyset \). If instead \( y_k^0 \geq \theta_* h \), then

\[
y_{n+1}^{n+1} - y_k^n - \mu y(y_k^n)h \leq y_0 - \theta_* h - \kappa \theta h + \kappa y_k^n h \leq y_0 - \frac{\theta_* h}{h} + \kappa y_k^n h.
\]

Recalling that \( h = \frac{\theta_*}{C} \), we note that there exists \( C > 0 \) such that

\[
y_k^n h \leq y_N h = \left(\sqrt{y_0 + \frac{\sigma}{2} Nh} \right)^2 = \left(\sqrt{y_0 V h + \frac{\sigma}{2} T} \right)^2 \leq C.
\]

Therefore

\[
y_{n+1}^{n+1} - y_k^n - \mu y(y_k^n)h \leq y_0 - \frac{\theta_* h}{h} + \kappa C < 0
\]

for \( h < \frac{\theta_*}{\theta_* + \kappa C} \). So, \( K_d(n,k) \neq \emptyset \). Now, by (3.16) and (3.17), since \( K_u(n,k) \neq \emptyset \) and \( K_d(n,k) \neq \emptyset \),

\[
0 \leq \mu(y_k^n)h + y_k^n - y_{k_u(n,k)}^{n+1} \left(\frac{y_{k_u(n,k)}^{n+1} - y_{k_u(n,k)}^n}{y_{k_u(n,k)}^{n+1} - y_{k_u(n,k)}^n} \right) = 1 + \frac{\mu y(y_k^n)h + y_k^n - y_{k_u(n,k)}^{n+1}}{y_{k_u(n,k)}^{n+1} - y_{k_u(n,k)}^n} \leq 1.
\]

\[\square\]

B Proof of (4.3)

For \( x \in \mathbb{R} \), let \( \lfloor x \rfloor = \sup\{k \in \mathbb{Z} : k \leq x\} \) denote the integer part. For \( N \in \mathbb{N} \), straightforward computations give

\[
\sum_{|n| \leq N} \varphi(n) = \frac{1}{2}(\varphi(N) + \varphi(-N)) + \int_{-N}^{N} \varphi(x)dx + \int_{-N}^{N} (x - \lfloor x \rfloor - \frac{1}{2}) \varphi'(x)dx.
\]

We recall that \( \varphi(\pm N) \to 0 \) as \( N \to \infty \) (because \( \varphi, \varphi' \in L^1(\mathbb{R},dx) \)). Moreover, the Fourier series representation gives

\[
x - \lfloor x \rfloor - \frac{1}{2} = \sum_{n \in \mathbb{Z}, n \neq 0} \frac{e^{-2\pi inx}}{2\pi in}, \quad x \in \mathbb{R}.
\]
So,

\[ \sum_{n \in \mathbb{Z}} \varphi(n) = \int_{\mathbb{R}} \varphi(x) dx + \int_{\mathbb{R}} \sum_{n \in \mathbb{Z}, n \neq 0} \frac{e^{-2\pi inx}}{2\pi in} \varphi'(x) dx. \]

With \( \mathcal{F}[\cdot] \) denoting the Fourier transform, we have

\[ \int_{\mathbb{R}} e^{-2\pi inx} \varphi'(x) dx = \mathcal{F}[\varphi](2\pi n) = 2\pi in \mathcal{F}[\varphi](2\pi n) \]

and

\[ |\mathcal{F}[\varphi'](2\pi n)| \leq \left| \frac{3|\varphi''|(2\pi n)}{2\pi n} \right| \leq \frac{M}{n} \]

because \( \varphi'' \in L^1(\mathbb{R}, dx) \). Thus, we can put the sum outside the integral and the statement holds.

References

[1] E. Akyıldırım, Y. Dolinsky, H.M. Soner, Approximating stochastic volatility by recombinant trees, *Ann. Appl. Probab.*, **24** (2014), pp. 2176–2205.

[2] A. Alfonsi, On the discretization schemes for the CIR (and Bessel squared) processes, *Monte Carlo Methods Appl.*, **11** (2005), pp. 355–467.

[3] A. Alfonsi, High order discretization schemes for the CIR process: Application to affine term structure and Heston models, *Math. Comp.*, **79** (2010), pp. 209–237.

[4] M. Altmayer, A. Neuenkirch, Discretising the Heston model: an analysis of the weak convergence rate, *IMA J. Numer. Anal.*, **37** (2017), pp. 1930–1960.

[5] L. Andersen, Simple and efficient simulation of the Heston stochastic volatility model, *J. Comput. Finance*, **11** (2008), pp. 1–42.

[6] E. Appolloni, L. Caramellino, A. Zanette, A robust tree method for pricing American options with CIR stochastic interest rate, *IMA J. Manag. Math.*, **26** (2015), pp. 345–375.

[7] K. Bahlali, B. Mezerdi, Y. Ouknine, Pathwise uniqueness and approximation of solutions of stochastic differential equations. in: Séminaire de Probabilités XXXII, Lecture Notes in Math., vol. 1686, Springer, Berlin 1998, pp. 166–187.

[8] V. Bally, C. Rey, Approximation of Markov semigroups in total variation distance, *Electron. J. Probab.*, **21** (2016), no. 12, 44 pp.

[9] V. Bally, D. Talay, The law of the Euler scheme for stochastic differential equations. I. Convergence rate of the distribution function, *Probab. Theory Related Fields*, **104** (1996), pp. 43–60.

[10] D.S. Bates, Jumps and stochastic volatility: exchange rate processes implicit in Deutsch mark options, *Rev. Fin.*, **9** (1996), pp. 69–107.

[11] M. Bossy, H. Olivero, Strong convergence of the symmetrized Milstein scheme for some CEV-like SDEs, *Bernoulli*, **24** (2018), pp. 1995–2042.

[12] M. Briani, L. Caramellino, A. Zanette, A hybrid approach for the implementation of the Heston model, *IMA J. Manag. Math.*, **28** (2017), pp. 467–500.

[13] M. Briani, L. Caramellino, A. Zanette, A hybrid tree/finite-difference approach for Heston-Hull-White type models. *J. Comput. Finance*, **21** (2017), pp. 1–45.

[14] M. Briani, L. Caramellino, G. Terenzi, A. Zanette, On a hybrid method using trees and finite-difference for pricing options in complex models, Preprint *ArXiv*:1603.07225, 2017.

[15] D. Brigo, F. Mercurio, *Interest Rate Models: Theory and Practice*, Springer 2001.

[16] C. Chiarella, B. Kang, G. Meyer, A. Ziogas, The evaluation of American option prices under stochastic volatility and jump-diffusion dynamics using the method of lines, *Int. J. Theor. Appl. Finan.*, **12** (2009), pp. 393–425.
[17] M. Costabile, M. Gaudenzi, I. Massabò, A. Zanette, Evaluating fair premiums of equity-linked policies with surrender option in a bivariate model, *Insurance Math. Econom.*, 45 (2009), pp. 286–295.

[18] J.C. Cox, J.E. Ingersoll, S.A. Ross, A Theory of the Term Structure of Interest Rates, *Econometrica*, 53 (1985), pp. 385–407.

[19] D. Duffie, J. Pan, K. Singleton, Transform analysis and asset pricing for affine jump-diffusions, *Econometrica*, 68 (2000), pp. 1343–1376.

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

[21] M.G. Garroni, J.L. Menaldi, *Green Functions for Second Order Parabolic Integro-Differential Problems*, Pitman Research Notes in Mathematics Series, vol. 275, 1993.

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

[23] J.E. Hilliard, A.L. Schwartz, A.L. Tucker, Bivariate binomial pricing with generalized interest rate processes, *J. Financ. Res.*, XIX-4 (1996), pp. 585–602.

[24] A. Itkin, Efficient Solution of Backward Jump-Diffusion PIDEs with Splitting and Matrix Exponentials, *J. Comput. Finance*, 19 (2016), pp. 29–70.

[25] S.G. Kou, A Jump-Diffusion Model for Option Pricing, *Management Science*, 48 (2002), pp. 1086–1101.

[26] D. Lamberton, G. Terenzi, Variational formulation of American option prices in the Heston model. *ArXiv:1711.11311*, 2017

[27] P.D. Lax, R.D. Richtmyer, Survey of the stability of linear finite difference equations, *Commun. Pure Appl. Math.*, 9 (1956), 267–293.

[28] R.C. Merton, Option pricing when underlying stock returns are discontinuous, *J. Financial Econom.*, 3 (1976), pp. 125–144.

[29] R. Mikulevicius, H. Pragarauskas, On Cauchy-Dirichlet problem in half-space for linear integro-differential equations in weighted Hölder spaces, *Electron. J. Probab.*, 10 (2005), pp. 1398–1416.

[30] D.B. Nelson, K. Ramaswamy, Simple binomial processes as diffusion approximations in financial models. *The Review of Financial Studies*, 3 (1990), pp. 393–430.

[31] H. Nieuwenhuis, M. Vellekoop, A tree-based method to price American Options in the Heston Model, *J. Comput. Finance*, 13 (2009), 1–21.

[32] G. Pagès, J. Printems, Functional quantization for numerics with an application to option pricing, *Monte Carlo Methods Appl.*, 11 (2005), 407–446.

[33] D.W. Stroock, S.R.S. Varadhan, *Multidimensional Diffusion Processes*, Springer, Berlin 1979.

[34] Y. Tian, A reexamination of lattice procedures for interest rate-contingent claims, *Adv. Futures Options Res.*, 7 (1994), pp. 87–110.

[35] J. Toivanen, A Componentwise Splitting Method for Pricing American Options Under the Bates Model, *Applied and Numerical Partial Differential Equations*, 15 (2010), pp. 213–227.

[36] J. Wei, Valuing American equity options with a stochastic interest rate: a note. *J. Financ. Eng.*, 2 (1996), pp. 195–206.

[37] C. Zheng, Weak convergence rate of a time-discrete scheme for the Heston stochastic volatility model, *SIAM J. Numer. Anal.*, 55 (2017), pp. 1243–1263.