A hybrid approach for the implementation of the Heston model

MAYA BRIANI

Istituto per le Applicazioni del Calcolo “M. Picone”, Consiglio Nazionale delle Ricerche, Rome, Italy

LUCIA CARAMELLINO

Dipartimento di Matematica, Università di Roma Tor Vergata, Roma, Italy

AND

ANTONINO ZANETTE*

Dipartimento di Scienze Economiche e Statistiche, Università di Udine, Udine, Italy

*Corresponding author. Email: antonino.zanette@uniud.it

[Received on 4 November 2014; accepted on 21 October 2015]

We propose an efficient hybrid tree/finite difference method in order to approximate the Heston model (and possibly other stochastic volatility models). We prove the convergence by embedding the procedure in a bivariate Markov chain and we study the approximation of European and American option prices. We finally provide numerical experiments that give accurate option prices in the Heston model, showing the reliability and the efficiency of the algorithm.

Keywords: tree methods; finite differences; Heston model; European and American options.

1. Introduction

The Black–Scholes model was the most popular model for derivative pricing and hedging, although it has shown several problems with capturing dramatic moves in financial markets. In fact, the assumption of a constant volatility in the Black–Scholes model over the lifetime of the derivative is not realistic. As an alternative to the Black–Scholes model, stochastic volatility models emerged. The Heston model (Heston, 1993) is perhaps the most popular stochastic volatility model, allowing one to obtain closed-formulae in the European case using Fourier transform. In the American option pricing case, the main algorithms turn out to be tree methods, Fourier-cosine methods and finite difference methods. Approximating trees for the Heston model have been considered in different papers; see e.g. Leisen (2000), Florescu & Viens (2005, 2008), Hilliard & Schwartz (1996) and Guan & Xiaoqiang (2000). The tree approach of Vellekoop & Nieuwenhuis (2009) actually provides to our knowledge, the best tree procedure in the literature. They use an approach which is based on a modification of an explicitly defined stock price tree where the number of nodes grows quadratically in the number of time steps. Fang & Oosterlee (2011) use a Fourier-cosine series expansion approach for pricing Bermudan options under the Heston model. As for finite difference methods for solving the parabolic partial differential equation (PDE) associated to the option pricing problems, they can be based on implicit, explicit or alternating direction implicit (ADI) schemes. The implicit scheme requires to solve a sparse system at each time step. Clarke & Parrott (1999) and Oosterlee (2003) formulate the American put pricing problem as a linear complementarity problem (LCP) and use an implicit finite difference scheme combined with a multigrid procedure, whereas Zvan et al. (1998) use a penalty method. The explicit scheme is a quick
approach although it requires small time steps to retain the stability. This leads to a large number of time steps and is not economic in computation. The ADI schemes are good alternative methods. For example, Hout & Foulon (2010) investigate four splitting schemes of the ADI type for solving the PDE Heston equation: the Douglas scheme, the Craigh–Sneyd scheme, the Modified Craigh–Sneyd scheme and the Hundsdorfer–Verwer scheme. Ikonen & Toivanen (2009) propose a componentwise splitting method for pricing American options in the Heston model. The LCP associated with the American option problem is decomposed into a sequence of five one-dimensional LCPs problems at each time step. The advantage is that LCP’s need the use of tridiagonal matrices. In Hout & Foulon (2010), the splitting method of Ikonen and Tovainen is combined with ADI schemes in order to obtain more efficient numerical results.

In this paper, we propose a new approach based both on tree and finite difference methods. Roughly speaking, our method approximates the CIR-type volatility process (Cox et al., 1985) through a tree approach already studied in Appolloni et al. (2014), which turns out to be very robust and reliable. And at each step, we make use of a suitable transformation of the asset price process, allowing one to take care of a new diffusion process with null correlation w.r.t. the volatility process. Then, by taking into account the conditional behaviour with respect to the evolution of the volatility process, we consider a finite difference method to deal with the evolution of the (transformed) underlying asset price process. We note that our procedure can be easily adapted to general stochastic volatility models, in particular to the Hull and White model (Hull & White, 1987) and the Stein and Stein model (Stein & Stein, 1991) (see next Remark 2.1). We also stress that jumps may be allowed in the dynamics for the underlying asset price process, but this shall be the subject of a further work.

The paper is organized as follows. In Section 2, we introduce the model, we study in detail the PDE associated with the pricing problem (Section 2.1) and then we set up our hybrid finite difference scheme (Section 2.2). In Section 3, we give the formal definition of the approximating algorithm and we study the convergence. Section 4 is devoted to the numerical results and to comparisons with other existing methods, showing the efficiency of the method in terms of the results and the computational time cost.

2. Construction of the method

In the Heston model (Heston, 1993), the dynamics under the risk neutral measure of the share price $S$ and the volatility process $V$ are governed by the stochastic differential equation system

\[
\frac{dS(t)}{S(t)} = (r - \delta) \, dt + \sqrt{V(t)} \, dZ_S(t),
\]

\[
dV(t) = \kappa(\theta - V(t)) \, dt + \sigma \sqrt{V(t)} \, dZ_V(t),
\]

with $S(0) = S_0 > 0$ and $V(0) = V_0 > 0$, where $Z_S$ and $Z_V$ are Brownian motions with correlation coefficient $\rho$: $d\langle Z_S, Z_V \rangle(t) = \rho \, dt$. Here $r$ is the risk-free rate of interest and $\delta$ the continuous dividend rate. We assume $\kappa, \theta > 0$ and we recall that the dynamics of $V$ follows a CIR process with mean reversion rate $\kappa$ and long-run variance $\theta$. The parameter $\sigma$ is called the volatility of the volatility.

From now on we set

\[
\tilde{\rho} = \sqrt{1 - \rho^2} \quad \text{and} \quad Z_V = W, \quad Z_S = \rho W + \tilde{\rho} Z,
\]
in which \((W, Z)\) denotes a standard two-dimensional Brownian motion. So, the dynamics can be written as
\[
\frac{dS(t)}{S(t)} = (r - \delta) \, dt + \sqrt{V(t)} \, (\rho \, dW(t) + \tilde{\rho} \, dZ(t)),
\]
(2.1)
\[
dV(t) = \kappa (\theta - V(t)) \, dt + \sigma \sqrt{V(t)} \, dW(t).
\]
(2.2)

We consider the diffusion pair \((Y, V)\), where
\[
Y_t = \log S_t - \frac{\rho}{\sigma} V_t.
\]
(2.3)

One has
\[
dY(t) = \left( r - \delta - \frac{1}{2} V_t - \frac{\rho}{\sigma} \kappa (\theta - V_t) \right) \, dt + \tilde{\rho} \sqrt{V(t)} \, dZ(t),
\]
(2.4)
\[
dV(t) = \kappa (\theta - V(t)) \, dt + \sigma \sqrt{V(t)} \, dW(t),
\]
(2.5)
(recall that \(W\) and \(Z\) are independent Brownian motions), with
\[
Y_0 = \log S_0 - \frac{\rho}{\sigma} V_0.
\]
(2.6)
In the following, we define \(\mu_Y\) and \(\mu_V\) to be the drift coefficient of \(Y_t\) and \(V_t\), respectively, i.e.
\[
\mu_Y(v) = r - \delta - \frac{1}{2} v - \frac{\rho}{\sigma} \kappa (\theta - v) \quad \text{and} \quad \mu_V(v) = \kappa (\theta - v).
\]
(2.7)
This means that any functional of the pair \((S_t, V_t)\) can be written as a suitable functional of the pair \((Y_t, V_t)\) by using the transformation (2.3), so \((Y_t, V_t)\) will be our underlying process of interest.

2.1 The associated pricing PDE in a small time interval

Let \(f = f(y, v)\) be a function of the space–variable pair \((y, v)\). For \(h\) small, we need to compute (an estimate for) the quantity
\[
u(t, y, v) = \mathbb{E}(f(Y_{t+h}^y, V_{t+h}^v)),
\]
in which \((Y_{t+h}^y, V_{t+h}^v)\) denotes the solution to (2.4) and (2.5) with the starting condition \((Y_t, V_t) = (y, v)\). Roughly speaking, the time instant \(t\) plays the role of a discretization instant in \([0, T]\), that is \(t = nh\), so \(t + h = (n + 1)h\) stands for the next discretizing time. To our purposes, we first note that
\[
\mathbb{E}(f(Y_{t+h}^{y,v}, V_{t+h}^{y,v})) = \mathbb{E}(\mathbb{E}(f(Y_{t+h}^{y,v}, V_{t+h}^{y,v}) \mid \mathcal{F}_{t+h}^W))
\]
where \(\mathcal{F}_{t+h}^W = \sigma(W_u : u \leq t + h)\). Now, conditionally to \(\mathcal{F}_{t+h}^W\), the volatility process \(V\) can be considered deterministic, so that \(Y\) has constant coefficients: setting
\[
U_{t+h}^{(g),y} = y + \int_t^{t+h} \mu_Y(g_s) \, ds + \tilde{\rho} \int_t^{t+h} \sqrt{g_s} \, dZ(s),
\]
(2.8)
where $g \in C([t, t + h])$ with $g \geq 0$ a.e. and $g_t = v$, then
\[
\mathbb{E}(f(Y_{t+h}^{y,v}, V_{t+h}^{v}) | \mathcal{F}_{t+h}) = \mathbb{E}(f(U_{t+h}^{(y),v}, g_{t+h}) | g = v).
\]
We define now
\[
\tilde{u}(t, y; g) = \mathbb{E}(f(U_{t+h}^{(y),v}, g_{t+h})),
\]
so that $\mathbb{E}(f(Y_{t+h}^{y,v}, V_{t+h}^{v}) | \mathcal{F}_{t+h}) = \tilde{u}(t, y; V_{t,h})$ and
\[
u(t, y, v) = \mathbb{E}(\tilde{u}(t, y; V_{t,h}^{v})).
\]
(2.10)

Now, from the Feynman–Kac formula, the function $\tilde{u}(t, y; g)$ in (2.9) can be linked to the solution to a parabolic PDE problem:
\[
\partial_s \tilde{u}(s, y; g) + \mu_Y(g_s) \partial_y \tilde{u}(s, y; g) + \frac{1}{2} \rho^2 g_s \partial_{yy}^2 \tilde{u}(s, y; g) = 0 \quad y \in \mathbb{R}, s \in [t, t + h),
\]
(2.11)
\[
\tilde{u}(t + h, y; g) = f(y, g_{t+h}) \quad y \in \mathbb{R},
\]
(to be formally correct, one should precise the right conditions on $f$ in order to get the solution of (2.11)—we do not enter in these arguments because we aim to give here only the main ideas that inspired the construction of our numerical procedure). Since we consider the case $h \simeq 0$, we have $g_s \simeq g_t = v$ and $\mu_Y (g_s) \simeq \mu_Y (g_t) = \mu_Y (v)$. So, our first approximation relies on freezing the coefficients of (2.11) at time $t$: we approximate $\tilde{u}(t, y; g)$ by the solution $\hat{u}^h(s, y; v, g_{t+h})$ to the PDE problem
\[
\partial_s \hat{u}^h(s, y; v, g_{t+h}) + \mu_Y(v) \partial_y \hat{u}^h(s, y; v, g_{t+h}) + \frac{1}{2} \rho^2 v \partial_{yy}^2 \hat{u}^h(s, y; v, g_{t+h}) = 0 \quad y \in \mathbb{R}, s \in (t, t + h),
\]
(2.12)
\[
\hat{u}^h(t + h, y; v, g_{t+h}) = f(y, g_{t+h}) \quad y \in \mathbb{R}.
\]

Let us remark that the solution to (2.12) actually depends on $g$ only through $v = g_t$ (appearing in the coefficients of the second-order operator) and $g_{t+h}$ (appearing in the Cauchy condition); that is why we used the notation $\hat{u}^h(s, y; v, g_{t+h})$. In contrast, the function solving (2.11) depends in principle on the whole trajectory $g$ over the time interval $[t, t + h]$.

Problem (2.12) can be solved by using a finite difference numerical method. Numerical reasonings suggest the use of an implicit approximation (in time) if $v$ is ‘far enough’ from zero; otherwise an explicit method should be considered—details are given in Sections 3.1.1 and 3.1.2. This means that one splits the real line by a space-grid $\mathcal{Y}^h = \{y_j = Y_0 + j \Delta y_h\}_{j \in \mathbb{Z}}$, $\Delta y_h$ denoting a fixed fixes a space step, and approximates the solution $\hat{u}^h(s, y; v, g_{t+h})$ to (2.12) on the grid $\mathcal{Y}^h$ by means of a linear operator (infinite-dimensional matrix) $\Pi^h(v) = (\Pi^h(v)_{ij})_{i,j \in \mathbb{Z}}$. In other words, one gets
\[
\hat{u}^h(s, y_i; v, g_{t+h}) \simeq \sum_{j \in \mathbb{Z}} \Pi^h(v)_{ij} f(y_j, g_{t+h}), \quad i \in \mathbb{Z}.
\]

Now, recalling (2.10) and the fact that $\tilde{u} \simeq \hat{u}^h$, on the grid $\mathcal{Y}^h$ the function $\nu$ is approximated through
\[
u(t, y, v) \simeq \mathbb{E}(\hat{u}^h(t, y_i; v, V_{t+h}^{v})) \simeq \sum_{j \in \mathbb{Z}} \Pi^h(v)_{ij} \mathbb{E}(f(y_j, V_{t+h}^{v})), \quad i \in \mathbb{Z}.
\]
(2.13)
We stress that the expectation on the r.h.s. above is now written in terms of the process \( V \) only, and this is the key point of our story because we can now use the tree method in Appolloni et al. (2014). But we will examine in depth this point in a moment.

In practice, one cannot solve the PDE problem over the whole real line. So, one takes a positive integer \( M_h > 0 \) such that \( M_h \Delta y_h \to +\infty \) as \( h \to 0 \) and considers a discretization of the (space) interval \([−M_h \Delta y_h + Y_0, Y_0 + M_h \Delta y_h]\) in \( 2M_h + 1 \) equally spaced points \( y_j = Y_0 + j \Delta y_h, j \in \mathcal{J}_{M_h} = \{-M_h, \ldots, M_h\} \). Then, the grid \( \mathcal{Y}_{M_h} = \{y_j = Y_0 + j \Delta y_h\}_{j \in \mathcal{J}_{M_h}} \) is finite and the approximation of \( \hat{u}^h_t(nh, y, \nu, g_{t+h}) \) is done by adding to (2.12) suitable boundary conditions. By calling again \( \Pi^h(v) \) the matrix (now, finite-dimensional) giving the solution from the finite difference approach, we still obtain

\[
\mathbb{E}(f(Y_{t+h}^h, V_{t+h}^h))|_{y = y_i} \simeq \sum_{j \in \mathcal{J}_{M_h}} \Pi^h(v)_{ij} \mathbb{E}(f(y_j, V_{t+h}^h)), \quad i \in \mathcal{J}_{M_h}.
\] (2.14)

### 2.2 The hybrid tree/finite difference approach

In order to compute the expectation in the r.h.s. of (2.14), we use the binomial tree method in Appolloni et al. (2014), which we can briefly recall as follows.

For \( n = 0, 1, \ldots, N \), consider the lattice

\[
\mathcal{Y}_n^h = \{v_{n,k}\}_{k=0,1,\ldots,n} \quad \text{with} \quad v_{n,k} = \left( \sqrt{V_0} + \frac{\sigma}{2} (2k - n) \sqrt{h} \right)^2 \mathbb{1}_\{V_0 + (\sigma/2) (2k - n) \sqrt{h} > 0\}
\] (2.15)

(note that \( v_{0,0} = V_0 \)) and for each fixed \( v_{n,k} \in \mathcal{Y}_n^h \), we define

\[
k^h_k(n, k) = \max \{k^\ast : 0 \leq k^\ast \leq k \text{ and } v_{n,k} + \mu_V(v_{n,k}) h \geq v_{n+1,k^\ast}\},
\] (2.16)

\[
k^h_k(n, k) = \min \{k^\ast : k + 1 \leq k^\ast \leq n + 1 \text{ and } v_{n,k} + \mu_V(v_{n,k}) h \leq v_{n+1,k^\ast}\}
\] (2.17)

with the understanding \( k^h_k(n, k) = 0 \) (respectively, \( k^h_k(n, k) = n + 1 \)) if the set in (2.16) (respectively, in (2.17)) is empty. The transition probabilities are defined as follows: starting from the node \((n, k)\), the probability that the process jumps to \( k^h_k(n, k) \) at time-step \( n + 1 \) is set as

\[
p^h_{k^h_k(n, k)} = \begin{cases} \ 
0 & \text{if } \frac{\mu_V(v_{n,k}) h + v_{n,k} - v_{n+1,k^h_k(n,k)}}{v_{n+1,k^h_k(n,k)} - v_{n+1,k^h_k(n,k)}} \wedge 1. \\
\end{cases}
\] (2.18)

And of course, the jump to \((n+1, k^h_k(n, k))\) happens with probability \( p^h_{k^h_k(n, k)} = 1 - p^h_{k^h_k(n, k)} \). This gives rise to a Markov chain \( (\tilde{V}_n^h)_{n=0,\ldots,N} \) that weakly converges, as \( h \to 0 \), to the diffusion process \((V_t)_{t\in[0,T]}\) and turns out to be a robust tree approximation for the CIR process \( V \). This means that we can approximate the expectation of suitable functionals of the diffusion \( V \) with the same expectation evaluated on the Markov chain \( \tilde{V}^h \):

\[
\mathbb{E}(\psi(\{V^h_{n+1,k} \mid n,k \}) \simeq \mathbb{E}(\psi(\tilde{V}_{n+1}^h | \tilde{V}_n^h = v_{n,k})) = \sum_{k^\ast \in \mathcal{K}_{n,k}} \psi(v_{n+1,k^\ast}) p^h_{k^\ast},
\]
with \( K_{n,k} = \{ k^h(n, k), k^d(n, k) \} \). So, by inserting this approximation in (2.19) we get
\[
\mathbb{E}(f(Y^{t,v}_{t+h}, V^{t,v}_{t+h}) | t = nh, y_{v}, v_{n} = v_{n,k}) \simeq \sum_{j \in \mathcal{J}_{M_h}} \prod_{k^*}^{h} (v_i)_{i,j} f(y_j, v_{n+1,k^*}) p_{k^*}, \quad i \in \mathcal{J}_{M_h}, \; v_{n,k} \in \mathcal{V}_n^h. \tag{2.19}
\]

In other words, the expectation \( \mathbb{E}(f(Y^{t,v}_{t+h}, V^{t,v}_{t+h}) \mid t = nh \) and \((y, v) = (y_i, v_{n,k}) \in \mathcal{V}_n^h \times \mathcal{V}_n^h \) is approximated with the integral w.r.t. the (finite) measure on \( \mathcal{V}_n^h \times \mathcal{V}_{n+1}^h \) defined as
\[
\mu^h(y_i, v_{n,k}; A \times B) = \sum_{k^* \in K_{n,k}} \sum_{j \in \mathcal{J}_{M_h}} \prod_{k^*}^{h} (v_{n,k})_{i,j} \delta_{y_j}(A) \delta_{v_{n+1,k^*}}(B) \tag{2.20}
\]
for every Borel sets \( A \) and \( B \), \( \delta_{a}(\cdot) \) denoting the Dirac mass in \( a \). Now, in Section 3 we shall be able to prove that, for small values of \( h \), \( \mu^h(v) \) is a stochastic matrix. This gives that \( \mu^h(y_i, v_{n,k}; \cdot) \) is actually a probability measure, that can be interpreted as a transition probability measure. Thus, we are doing expectations on a Markov chain \( (\bar{Y}^h_n, \bar{V}^h_n)_{n=0,1,...,N} \), whose state-space, at time step \( n \), is given by \( \mathcal{V}_n^h \times \mathcal{V}_n^h \) and whose transition probability measure at time step \( n \) is given by (2.20). Moreover, in Section 3.2 we will prove that, under appropriate conditions on \( \Delta y_h \) and \( M_h \) (see next (3.21) and (3.22)), the family of Markov chains \( (\bar{Y}^h, \bar{V}^h)_{h} \) weakly converges to the diffusion process \( (Y, V) \). And this gives the convergence of our hybrid tree/finite difference algorithm approximating the Heston model.

### 2.3 An example: pricing American options

Consider an American option with maturity \( T \) and payoff function \( \Phi(S_t, V_t) \) for \( t \in [0,T] \). First of all, by using (2.3), we replace the pair \((S, V)\) with the pair \((Y, V)\), so the obstacle will be given by
\[
\Psi(Y_t, V_t) = \Phi(e^{Y_{t+\sigma/\sigma}V_t}, V_t), \quad t \in [0, T].
\]

The price at time 0 of such an option is then approximated by a backward dynamic programming algorithm, working as follows. First, consider a discretization of the time interval \([0, T]\) into \( N \) subintervals of length \( h = T/N : [0, T] = \cup_{n=0}^{N-1} [nh, (n+1)h] \). Then the price \( P(0, Y_0, V_0) \) of such an American option is numerically approximated through the quantity \( P_h(0, Y_0, V_0) \) which is iteratively defined as follows: for \((y, v) \in \mathbb{R} \times \mathbb{R}_+ \),
\[
\begin{cases}
P_h(T, y, v) = \Psi(y, v) & \text{and as } n = N - 1, \ldots, 0, \\
P_h(nh, y, v) = \max\{\Psi(y, v), e^{-rh}E(P_h((n+1)h, Y_{(n+1)h}^v, V_{(n+1)h}^v))\}.
\end{cases}
\]

From the financial point of view, this means to allow the exercise at the fixed times \( nh, n = 0, \ldots, N \). Now, in order to compute the expectations, we use the approximating algorithm just described: by using (2.19), we numerically compute the above backward induction as follows: for \( n = 0, 1, \ldots, N \), we define \( \bar{P}_h(nh, y, v) \) for \((y, v) \in \mathcal{V}_n^h \times \mathcal{V}_n^h \) by
\[
\begin{cases}
\bar{P}_h(T, y_i, v_{n,k}) = \Psi(y_i, v_{n,k}) & i \in \mathcal{J}_{M_h} \text{ and } v_{n,k} \in \mathcal{V}_n^h, \text{ and as } n = N - 1, \ldots, 0, \\
\bar{P}_h(nh, y_i, v_{n,k}) = \max \left\{ \Psi(y_i, v), e^{-rh} \sum_{k^*} \prod_{k^*}^{h} (v_{n,k})_{i,j} \bar{P}_h((n+1)h, y_j, v_{n+1,k^*}) p_{k^*} \right\},
\end{cases} \tag{2.21}
\]

\( i \in \mathcal{J}_{M_h} \) and \( v_{n,k} \in \mathcal{V}_n^h \).
where the sum above is done for $k^* \in \{k^h_0(n,k), k^h_d(n,k)\}$ and $j \in J_{M_i}$. In next Section 3.3, we shall discuss the convergence of the backward induction (2.21) to the price of the associated American option written on the Heston model.

**Remark 2.1** Roughly speaking, our algorithm is a mixing of a tree method for the process $V$ and a finite difference method to handle the noise in $Y$ (which is independent of the noise driving $V$). This numerical procedure is strongly based on the fact that the process $V$ can be efficiently approximated by a tree method and the process $Y$ (to which we apply a finite difference method) is an Itô’s process whose coefficients depend on $V$ only and whose driving Brownian motion is independent of the Brownian noise in the stochastic differential equation for $V$, as in (2.4) and (2.5). It is worth observing that this situation is standard in all the well-known and studied stochastic volatility models via diffusion processes. In fact, consider a general stochastic volatility model, that is

$$\frac{dS_t}{S_t} = (r - \delta) \, dt + \eta(V_t) (\rho \, dW_t + \bar{\rho} \, dZ_t), \quad (2.22)$$

$$dV_t = \mu(V_t) \, dt + \sigma(V_t) \, dW_t, \quad (2.23)$$

where $\eta, \mu_V$ and $\sigma_V$ are suitable functions. We assume that there exist two functions $G$ and $\xi$ such that

$$\int_0^t \eta(V_s) \, dW_s = G(V_t) + \int_0^t \xi(V_s) \, ds. \quad (2.24)$$

Note that if one requires that $G$ be twice differentiable, by the Itô’s formula (2.24) is equivalent to the conditions

$$\eta(v) = G'(v)\sigma_V(v) \quad \text{and} \quad \xi(v) = -G'(v)\mu_V(v) - \frac{1}{2} G''(v)\sigma_V(v)^2. \quad (2.25)$$

Set now

$$Y_t = \ln S_t - \rho G(V_t),$$

so that $(S_t, V_t) = (e^{Y_t + \rho G(V_t)}, V_t)$. Then, by applying the Itô’s formula, one immediately gets

$$dY_t = \mu_Y(V_t) \, dt + \rho \eta(V_t) \, dZ_t \quad \text{and} \quad dV_t = \mu_V(V_t) \, dt + \sigma_V(V_t) \, dW_t,$$

in which $\mu_Y(v) = r - \delta - \frac{1}{2} \eta(v)^2 - \rho \xi(v)$. This is a situation similar to that described for the Heston model (in which $G(v) = v/\sigma$ and $\xi(v) = -(1/\sigma)\mu_V(v)$) and one can proceed similarly. Recall that such functions $G$ and $\xi$ actually exist also in the most famous stochastic volatility models, which are the following:

- Hull and White model (Hull & White, 1987): equations (2.22–2.23) are

$$\frac{dS_t}{S_t} = (r - \delta) \, dt + \sqrt{V_t} (\rho \, dW_t + \bar{\rho} \, dZ_t) \quad \text{and} \quad dV_t = \mu_V(V_t) \, dt + \sigma_V(V_t) \, dW_t,$$

and, here, $G(v) = (2/\sigma) \sqrt{v}$ and $\xi(v) = (-\mu/\sigma + \sigma/4) \sqrt{v}$ (recall that, $V$ being a geometric Brownian motion, one can restrict to the half-space $v > 0$);
Stein and Stein model (Stein & Stein, 1991): equations (2.22) and (2.23) are

\[
\frac{dS_t}{S_t} = (r - \delta)dt + V_t(\rho \, dW_t + \bar{\rho} \, dZ_t) \quad \text{and} \quad dV_t = \kappa(\theta - V_t)dt + \sigma \, dW_t
\]

and, here, \(G(v) = v^2/2\sigma\) and \(\xi(v) = -(1/\sigma)\kappa v(\theta - v) + \frac{1}{2}\sigma\).

3. The convergence of the algorithm

We first set up the finite difference method we take into account. Then, in Section 3.2, we formally define the approximating Markov chain and prove the weak convergence to the Heston model in the path space.

For ease of notation, for a while we drop the dependence on the time-step \(h\) for the space step \(\Delta y_h\) and the number \(M_h\) related to the points of the space-grid, so we simply write \(\Delta y\) and \(M\), as well as \(\mathcal{J}_M = \{-M, \ldots, M\}\) and \(\mathcal{Y}_M = \{y_i = Y_0 + i\Delta y\}_{i \in \mathcal{J}_M}\).

3.1 The finite difference scheme for the PDE problem (2.12)

As described in Section 2.2, at each time step \(n\) we need to numerically solve (2.12) for \(t = t_n = nh\). So, we briefly describe the finite difference method we apply to problem (2.12), outlining some key properties of the associated operator allowing us to prove the convergence. For further information on finite difference methods for PDEs we refer the reader for instance to Strikwerda (2004).

Let \(t = nh\), \(v\) and \(v^* = g_{nh+h}\) be fixed and let us set \(u^n_j = \tilde{u}^n_{n}(nh, y_j)\) the discrete solution of (2.12) at time \(nh\) on the point \(y_j\) of the grid \(\mathcal{Y}_M\) — for simplicity of notations, we do not stress on \(u^n_j\) the dependence on \(v\) (from the coefficients of the PDE), \(v^*\) (from the Cauchy problem) and \(h\).

It is well known that the behaviour of the solution to (2.12) changes with respect the magnitude of the rate between the diffusion coefficient \(\rho^2/v/2\) and the advection term \(\mu_y(v)\). To deal with these cases, we fix a small real threshold \(\epsilon > 0\) and in the following we shall describe how to solve both the case \(v < \epsilon\) and \(v > \epsilon\) by applying an explicit in time and an implicit in time approximation, respectively.

It is well known that, for a big enough diffusion coefficient, to avoid over-restrictive conditions on the grid steps, it is suggested to apply implicit finite differences to problem (2.12). In this case, the discrete solution \(\{u^n_j\}_{j \in \mathbb{Z}}\) at time \(nh\) will then be computed in terms of the solution \(\{u^{n+1}_j\}_{j \in \mathbb{Z}}\) at time \((n+1)h\) by solving the following discrete problem:

\[
\frac{u^{n+1}_j - u^n_j}{h} + \mu_y(v) \frac{u^n_{j+1} - u^n_{j-1}}{2\Delta y} + \frac{1}{2} \frac{\rho^2}{v} \frac{u^n_{j+1} - 2u^n_j + u^n_{j-1}}{\Delta y^2} = 0, \quad j \in \mathbb{Z},
\]

where \(\Delta y = y_j - y_{j-1}, \forall j \in \mathbb{Z}\).

On the other hand, when the diffusion coefficient is small compared with the reaction one, it is suggested to apply an explicit in time approximation coupled with a forward or backward finite difference for the first-order term \(u\), depending on the sign of the reaction coefficient.

Specifically, for \(v\) close to 0, that is \(v < \epsilon\), we solve the problem by the following approximation schemes: when \(\mu_y(v) > 0\),

\[
\frac{u^{n+1}_j - u^n_j}{h} + \mu_y(v) \frac{u^{n+1}_{j+1} - u^{n+1}_{j-1}}{\Delta y} + \frac{1}{2} \frac{\rho^2}{v} \frac{u^{n+1}_{j+1} - 2u^{n+1}_j + u^{n+1}_{j-1}}{\Delta y^2} = 0, \quad j \in \mathbb{Z},
\]
while, when $\mu_Y(v) < 0$,
\[
\frac{u_j^{n+1} - u_j^n}{h} + \mu_Y(v) \frac{u_j^{n+1} - u_{j-1}^{n+1}}{\Delta y} + \frac{1}{2} \rho^2 v \frac{u_j^{n+1} - 2u_{j+1}^{n+1} + u_{j-1}^{n+1}}{\Delta y^2} = 0, \quad j \in \mathbb{Z}. \tag{3.3}
\]

As previously mentioned at the end of Section 2.1, for the numerical tests one has to deal with a finite grid $Y^h_M = \{y_j\}_{j \in J_M}$ and problems (3.1) and (3.2) have to be coupled with suitable numerical boundary conditions. Here, we assume that the boundary of the domain values are defined by the following relations (Neumann-type boundary conditions): in the implicit case we set
\[
u_{M-1}^n = u_{M+1}^n, \quad u_{M}^n = u_{M-1}^n, \tag{3.4}
\]
whereas in the explicit case we set
\[
u_{M-1}^{n+1} = u_{M+1}^{n+1}, \quad u_{M}^{n+1} = u_{M-1}^{n+1}. \tag{3.5}
\]

Other conditions can surely be selected, for example the two boundary values $u_{M-1}^{n+1}$ and $u_{M}^{n+1}$ may be a priori fixed by a known constant (this typically appears in financial problems). Specifically, other types of boundary conditions can be handled by using Proposition A.1 (see Appendix A) and all arguments that follow apply as well (for details, see next Remark 3.11). So, hereafter we set
\[
\alpha = \frac{h}{2\Delta y} \mu_Y(v) \quad \text{and} \quad \beta = \frac{h}{2\Delta y^2} \rho^2 v. \tag{3.6}
\]

### 3.1.1 The case $v > \epsilon$

By applying implicit finite differences (3.1) coupled with boundary conditions (3.4), we get the solution $u^n = (u_{M-1}^n, \ldots, u_M^n)^\top$ by solving the following linear system:
\[
Au^n = u^{n+1}, \tag{3.7}
\]
where $A$ is the $(2M + 1) \times (2M + 1)$ tridiagonal real matrix given by
\[
A = \begin{pmatrix}
1 + 2\beta & -2\beta & & & \\
\alpha - \beta & 1 + 2\beta & -\alpha - \beta & & \\
& \ddots & \ddots & \ddots & \\
\alpha - \beta & 1 + 2\beta & -\alpha - \beta & & \\
& & -2\beta & 1 + 2\beta & \\
\end{pmatrix}. \tag{3.8}
\]

We immediately prove that the solution $u^n$ to (3.7) actually exists at least when $\beta > |\alpha|$ (we will see later that this is not a restrictive condition).

**Proposition 3.1** Assume that $\beta > |\alpha|$. Then $A$ is invertible and $A^{-1}$ is a stochastic matrix, that is all entries are non-negative and, for $1 = (1, \ldots, 1)^\top$, $A1 = 1$.

**Proof.** The matrix $A = (a_{ij})_{i,j \in J_M}$ satisfies
\[
(P1) \quad A1 = 1, \quad \text{i.e.} \quad \sum_{j=-M}^M a_{ij} = 1 \text{ for } i \in J_M
\]
and, for $\beta > |\alpha|$, one has also
(P2) \( a_{ii} > 0 \) for \( i \in J_M \) and for \( j \in J_M, j \neq i, a_{ij} \leq 0; \)

(P3) \( A \) is strict or irreducibly diagonally dominant, i.e. \( \sum_{j \in J_M, j \neq i} |a_{ij}| < a_{ii} \) for \( i \in J_M \).

(P2)–(P3) give that \( A \) is an invertible \( M \)-matrix (see for instance Berman & Plemmons, 1994), so that \( A^{-1} \) is non-negative (i.e. \( a_{ij}^{-1} \geq 0, i, j = -M, \ldots, M \)). Moreover, by (P1), \( I = A^{-1} I \).

For each \( l \in \mathbb{N} \) and \( y \in \mathcal{Y}_M \), we consider the polynomial \( (y - y_i)^l \) and we call \( \psi^l_i(y) \in \mathbb{R}^{2M+1} \) the associated (vector) function of \( y \in \mathcal{Y}_M \):

\[
(\psi^l_i(y))_k = (y_k - y_i)^l = \Delta^l(k - i), \quad k \in J_M.
\]

(3.9)

In Section 3.2, we need to deal with \( A^{-1}\psi^l_i(y) \) for \( l \leq 4 \) and \( i \in J_M \). So, we study such objects. By Proposition 3.1, for \( \beta > |\alpha| \) one has that \( A \) is invertible and we may then compute \( A^{-1}\psi^l_i(y) \). We also note that \( \psi^0_i(y) = I \), so that \( A^{-1}\psi^0_i(y) = A^{-1} I = I \).

In the following, the symbol \( [\cdot] \) will stand for the floor function and we use the understanding \( \sum_{k=1}^{0} [\cdot]_k = 0 \). Moreover, we let \( e_i \) denote the standard orthonormal basis: for \( i, k \in J_M \), \( (e_i)_k = 0 \) for \( k \neq i \) and \( (e_i)_k = 1 \) if \( k = i \).

**Lemma 3.2** Let \( \psi^l_i(y) \) be defined in (3.9). Then, for every \( l \in \mathbb{N} \) and \( i \in J_M \) one has

\[
A\psi^l_i(y) = \psi^l_i(y) - \sum_{j=0}^{l-1} \binom{l}{j} a_{l-j}\Delta^j(y^l - y^l_i) + b_{l+}M e_{-M} + b_{l-}M e_M,
\]

(3.10)

where

\[
a_n = (\beta - \alpha)(-1)^n + (\beta + \alpha), \quad n \in \mathbb{N},
\]

(3.11)

\[
b_{l+}M = \pm 2 \sum_{j=0, l-j \text{ odd}}^{l-1} \binom{l}{j} (\beta \pm \alpha) \Delta^j(y_{\pm M} - y_i)^j.
\]

(3.12)

Moreover, \( b_{l+}M \) can be bounded as follows:

\[
|b_{l+}M| \leq 2(\beta \pm \alpha)(\Delta y + |y_{\pm M} - y_i|)^j,
\]

(3.13)

**Proof.** For \( k \in J_M \) with \( k \neq \pm M \), we have

\[
A\psi^l_i(y))_k = -(\beta - \alpha)(y_k - y_i - \Delta y)^l + (1 + 2\beta)(y_k - y_i)^l - (\beta + \alpha)(y_k - y_i + \Delta y)^l
\]

\[
= (1 + 2\beta)(y_k - y_i)^l - \sum_{j=0}^{l-1} \binom{l}{j} (y_k - y_i)^j((\beta - \alpha)(-1)^{-j} + (\beta + \alpha))\Delta^j
\]

\[
= (y_k - y_i)^l - \sum_{j=0}^{l-1} \binom{l}{j} (y_k - y_i)^j a_{l-j}\Delta^j.
\]
Following similar reasonings for \( k = \pm M \), we easily obtain

\[
(A \psi^{\dagger}_i(y))_{\pm M} = (y_{\pm M} - y_i)^l - 2\beta \sum_{j=0}^{l-1} \binom{l}{j} (\mp \Delta y)^{j-l}(y_{\pm M} - y_i)^j
\]

and by using (3.12), we get the full form (3.10). Finally, estimate (3.13) immediately follows from Newton’s binomial formula. \( \square \)

We are now ready to characterize \( A^{-1} \psi^{\dagger}_i(y) \) for every polynomial \( \psi^{\dagger}_i(y) \) as in (3.9).

**Proposition 3.3** Suppose that \( \beta > |\alpha| \) and, for \( l \geq 1 \), let \( \gamma_{l,k}, k \leq l \), be iteratively (backwardly) defined as follows: \( \gamma_{1,0} = a_0 \) and

\[
\gamma_{l,k} = \binom{l}{k} a_{l-k} \Delta y^{l-k} + \sum_{j=k+1}^{l-1} \gamma_{l,j} \binom{j}{k} a_{j-k} \Delta y^{j-k}, \quad k = l - 1, \ldots, 0,
\]

where the coefficients \( a_n \) are given in (3.11). Then

\[
A^{-1} \psi^{\dagger}_i(y) = \psi^{\dagger}_i(y) + \sum_{j=0}^{l-1} \gamma_{l,j} \psi^{\dagger}_j(y) + \phi^{\dagger}_{l,M}(y),
\]

in which

\[
\phi^{\dagger}_{l,M}(y) = T_{l,i}^M A^{-1} e_{-M} + T_{l,i}^M A^{-1} e_M \quad \text{with} \quad T_{l,i}^{\pm M} = -b_{\pm M}^{l-i} - \sum_{j=0}^{l-1} \gamma_{l,j} b_{l,j}^{\pm M},
\]

the \( b_{l,j}^{\pm M} \)'s being given in (3.12). Moreover, if \( \Delta y \leq 1 \), \( M \Delta y \geq 1 \) and \( l^2(l+1)(\beta \Delta y^2 + |\alpha| \Delta y) \leq 1 \), the following estimate holds for \( T_{l,i}^{\pm M} \): for every \( i \in J_M \),

\[
|T_{l,i}^{\pm M}| \leq 4(\beta \pm \alpha) \Delta y^l (1 + 2M)^l.
\]

**Proof.** It is clear that \( A^{-1} \psi^{\dagger}_i(y) = \psi^{\dagger}_i(y) + \sum_{j=1}^{l-1} \gamma_{l,j} \psi^{\dagger}_j(y) + \phi^{\dagger}_i(y) \) if and only if

\[
\psi^{\dagger}_i(y) = A \psi^{\dagger}_i(y) + \sum_{j=0}^{l-1} \gamma_{l,j} A \psi^{\dagger}_j(y) + A \phi^{\dagger}_{l,M}(y).
\]

We call \((\ast)\) the r.h.s. above. By using Lemma 3.2 and setting \( \theta^{\dagger}_i = b_{l,i}^{-M} e_{-M} + b_{l,i}^M e_M \), one has

\[
(\ast) = \psi^{\dagger}_i(y) - \sum_{k=0}^{l-1} \binom{l}{k} a_{l-k} \Delta y^{l-k} \psi^{\dagger}_k(y) + \theta^{\dagger}_i
\]

\[
+ \sum_{j=0}^{l-1} \gamma_{l,j} \left( \psi^{\dagger}_j(y) - \sum_{k=0}^{j-1} \binom{j}{k} a_{j-k} \Delta y^{j-k} \psi^{\dagger}_k(y) + \theta^{\dagger}_j \right) + A \phi^{\dagger}_{l,M}(y)
\]
\[
\begin{aligned}
\psi^i_j(y) &= \psi^i_j(y) - \sum_{k=0}^{l-1} \binom{l}{k} a_{l-k} \Delta y^{l-k} \psi^i_k(y) + \sum_{j=0}^{l-1} \gamma_{l,j} \psi^i_j(y) \\
&- \sum_{j=0}^{l-1} \gamma_{l,j} \sum_{k=0}^{j-1} \binom{j}{k} a_{j-k} \Delta y^{j-k} \psi^i_k(y) + \sum_{j=0}^{l-1} \gamma_{l,j} \theta^j_i + A\phi^i_{l,M}(y) \\
&= \psi^i_j(y) + \sum_{k=0}^{l-1} \left( \binom{l}{k} a_{l-k} \Delta y^{l-k} + \gamma_{l,k} - \sum_{j=k+1}^{l-1} \gamma_{l,j} \binom{j}{k} a_{j-k} \Delta y^{j-k} \right) \psi^i_k(y) \\
&+ \theta^j_i + \sum_{j=0}^{l-1} \gamma_{l,j} \theta^j_i + A\phi^i_{l,M}(y).
\end{aligned}
\]

By the definition of the \(\gamma_{l,k}\)'s and \(\phi^i_{l,M}(y)\)'s, each coefficient in the above (first) sum is null and the sum of the last three terms is zero, so that \((\ast) = \psi^i_j(y)\). Let us discuss \((3.16)\). By using \((3.13)\) and the fact that \(|y_{\pm} - y_{i}| \leq 2M \Delta y\), since \(\Delta y(1 + 2M) \geq 1\), we can write

\[
|T^\pm_{l,j}(y)| \leq 2(\beta \pm \alpha) \Delta y^l (1 + 2M)^l \left( 1 + \sum_{j=0}^{l-1} |\gamma_{l,j}| \right).
\]

It remains to prove that, under our constraints, \(\sum_{j=0}^{l-1} |\gamma_{l,j}| \leq 1\). For every \(k = 0, 1, \ldots, l - 1\) and \(j = k + 1, \ldots, l - 1\) we consider the estimates

\[
\binom{l}{k} \leq 2^l \quad \text{and} \quad \binom{j}{k} \leq 2^j.
\]

By inserting in \((3.14)\), we can write

\[
|\gamma_{l,k}| \leq 2^l |a_{l-k}| \Delta y^{l-k} + 2^l \sum_{j=k+1}^{l-1} |\gamma_{l,j}| |a_{j-k}| \Delta y^{j-k}, \quad k = l - 1, \ldots, 0.
\]

We also note that, for \(m \geq 1\) and \(\Delta y \leq 1\),

\[
|a_m| \Delta y^m \leq 2(\beta \Delta y^2 + |\alpha| \Delta y)
\]

so that we get

\[
|\gamma_{l,k}| \leq 2^{l+1} (\beta \Delta y^2 + |\alpha| \Delta y) + 2^{l+1} (\beta \Delta y^2 + |\alpha| \Delta y) \sum_{j=k+1}^{l-1} |\gamma_{l,j}|, \quad k = l - 1, \ldots, 0.
\]

Now, if \(2^{l+2} (\beta \Delta y^2 + |\alpha| \Delta y) \leq 1\), we get

\[
|\gamma_{l,k}| \leq \frac{1}{2l} + \frac{1}{2l} \sum_{j=k+1}^{l-1} |\gamma_{l,j}|, \quad k = l - 1, \ldots, 0
\]
and by summing over \( k \),

\[
\sum_{k=0}^{l-1} |\gamma_{ik}| \leq \frac{1}{2} + \frac{1}{2l} \sum_{k=0}^{l-1} \sum_{j=k+1}^{l-1} |\gamma_{ij}| = \frac{1}{2} + \frac{1}{2l} \sum_{j=1}^{l-1} \sum_{k=0}^{l-1} |\gamma_{ij}| \leq \frac{1}{2} + \frac{1}{2} \sum_{j=0}^{l-1} |\gamma_{ij}|
\]

from which it follows that \( \sum_{k=0}^{l-1} |\gamma_{ik}| \leq 1 \) and the statement holds. \( \square \)

**Remark 3.4** For further use, we write down explicitly the vector \( A^{-1} \psi_i(y) \) for \( l = 1, 2, 4 \). Straightforward computations give the following:

\[
\begin{align*}
A^{-1} \psi_1(y) &= \psi_1(y) + 2\alpha \Delta y 1 + \phi_{1,M}(y), \\
A^{-1} \psi_2(y) &= \psi_2(y) + 4\alpha \Delta y \psi_1(y) + 2(\beta + 2\alpha) \Delta y^2 1 + \phi_{2,M}(y), \\
A^{-1} \psi_4(y) &= \psi_4(y) + 8\alpha \Delta y \psi_2(y) + 12(\beta + 4\alpha^2) \Delta y^2 \psi_2(y) + 8(\alpha + 12\alpha^2 + 18\alpha \beta) \Delta y^3 \psi_1(y) + 2(\beta + 16\alpha^2 + 96\alpha^3 + 12\beta^2 + 192\alpha^3 \beta) \Delta y^4 1 + \phi_{4,M}(y).
\end{align*}
\]

In particular, by replacing the \( \alpha \) and \( \beta \) expressions (3.6), we get the formulas

\[
\begin{align*}
(A^{-1} \psi_1(y))_i &= h \mu_Y(v) + (\phi_{1,M}^i)_i, \\
(A^{-1} \psi_2(y))_i &= h \rho^2 v + 2h \Delta y \mu_Y(v) + (\phi_{2,M}^i)_i, \\
(A^{-1} \psi_4(y))_i &= h \Delta y^3 \rho^2 v + 8h^2 \Delta y^2 \mu_Y(v)^2 + 24h^3 \mu_Y(v)^3 + 6h^2 \rho^4 v^2 + 24 \frac{h^4}{\Delta y} \rho^2 v \mu_Y(v)^3 + (\phi_{4,M}^i)_i. 
\end{align*}
\]

Furthermore, to deal with the numerical boundary conditions, as those given in (3.4), we need to study the behaviour of the \( i \)th component of the boundary term \( \phi_{i,M}^j(y) \) in (3.15) as \( i \) is ‘far from the boundary’ and \( l = 1, 2, 4 \). Here, we use a quite general result (allowing one to set up different boundary conditions) whose precise statement and proof are postponed in Appendix A.

**Proposition 3.5** Suppose that \( \beta > |\alpha| \). Let \( l \in \mathbb{N}, i \in J_M \) and let \( \phi_{i,M}^j(y) \) denote the boundary term in (3.15). Assume that \( \Delta y \leq 1, M \Delta y \geq 1 \) and \( l^2 + 2(\beta \Delta y^2 + |\alpha| \Delta y) \leq 1 \). Then one has

\[
|\phi_{i,M}^j(y)| \leq 8(\beta + |\alpha|) \Delta y^j (1 + 2M)^l \left( 1 - \frac{1}{1 + \beta + |\alpha|} \right)^{M-i}.
\]

**Proof.** Since \( \beta > |\alpha| \), \( A \) satisfies the requirements in Proposition A.1. So, we use such a result, which has been specialized to our type of matrix in Remark A.2: taking \( a = 1 + 2\beta, b = -\beta + \alpha \) and \( c = -\beta - \alpha \), we obtain

\[
|A^{-1} e_{\pm M})_i| \leq \left( \frac{\beta \pm \alpha}{\gamma_{\pm M}} \right)^{M-i}
\]

where

\[
\gamma_{\pm M} = \min \left( 1 + 2\beta - \frac{2\beta(\beta \mp \alpha)}{1 + 2\beta}, \frac{1 + 2\beta + \sqrt{(1 + 2\beta)^2 - 4(\beta^2 - \alpha^2)}}{2} \right).
\]
Straightforward computations give $\gamma_{\pm M}^* \geq 1 + \beta \pm \alpha$, so that

$$|(A^{-1}e_{\pm M})_i| \leq \left( \frac{\beta \pm \alpha}{1 + \beta \pm \alpha} \right)^{M-i}.$$

Now, since $\beta > |\alpha|$ we can write $(\beta \pm \alpha)/(1 + \beta \pm \alpha) = 1 - 1/(1 + \beta \pm |\alpha|)$, so that

$$|(A^{-1}e_{\pm M})_i| \leq \left( 1 - \frac{1}{1 + \beta \pm |\alpha|} \right)^{M-i}.$$

We insert now the above estimate and estimate (3.16) in (3.15), and we get the result. □

3.1.2 The case $v < \epsilon$

Here we go through our procedure for the explicit in time approximation. We recall that, for $v < \epsilon$, we consider the forward finite differences (3.2) or the backward finite differences (3.3) for the first-order term depending on the sign of the reaction coefficient: $\mu_Y(v) > 0$ or $\mu_Y(v) < 0$, respectively, and from (3.6) this means that $\alpha > 0$ or $\alpha < 0$, respectively. So, by considering also the case $\alpha = 0$ and by coupling with the boundary conditions (3.5), straightforward computations give that the solution $u^n$ of the explicit in time scheme satisfies the condition

$$u^n = Cu^{n+1}$$

where

$$C = \begin{pmatrix}
1 - 2\beta - 2|\alpha| & 2\beta + 2|\alpha| \\
\beta + 2|\alpha| & 1 - 2\beta - 2|\alpha| & \beta + 2|\alpha| \mathbb{I}_{|\alpha| < 0} \\
\vdots & \ddots & \ddots \\
\beta + 2|\alpha| \mathbb{I}_{|\alpha| < 0} & 1 - 2\beta - 2|\alpha| & \beta + 2|\alpha| \mathbb{I}_{|\alpha| > 0} \\
2\beta + 2|\alpha| & 1 - 2\beta - 2|\alpha|
\end{pmatrix}, \quad (3.18)$$

$\alpha$ and $\beta$ being given in (3.6) and $\mathbb{I}$ denoting the indicator function. We remark that $C$ is a stochastic matrix if and only if

$$2\beta + 2|\alpha| \leq 1. \quad (3.19)$$

We also note that if $\epsilon$ is small enough, such a condition is not restrictive, but we will discuss deeper this point later.

In Section 3.2, we need to deal with $C\psi^i_l(y)$ for $l \leq 4$ and $i \in J_M$, where the function $\psi^i_l(y) \in \mathbb{R}^{2M+1}$ are defined in (3.9). So, we obtain the following lemma.

**Lemma 3.6** Let $\psi^i_l(y)$ be defined in (3.9). Then, for every $l \in \mathbb{N}$ and $i \in J_M$ one has

$$C\psi^i_l(y) = \psi^i_l(y) + \sum_{j=0}^{l-1} \binom{l}{j} d_{l-j} \Delta y^{l-j} \psi^i_j(y) + c^{-M}_{l,i} e_{-M} + c^M_{l,i} e_M,$$

where

$$d_n = (-1)^n (\beta + 2|\alpha| \mathbb{I}_{|\alpha| < 0}) + \beta + 2|\alpha| \mathbb{I}_{|\alpha| > 0}, \quad n \in \mathbb{N},$$
that is \(d_n = 2(\beta + |\alpha|)\) if \(n\) is even and \(d_n = 2\alpha\) if \(n\) is odd, and \(c^{\pm}_i\) are given by

\[
c^{\pm}_i = 2(\beta + 2|\alpha|) \sum_{j=0, \ i-j \in \mathcal{A}_+}^{l-1} \binom{l}{j} \Delta y^{l-j}(y_{z+M} - y_i)^j
\]

where \(\mathcal{A}_+\) and \(\mathcal{A}_-\) denote the set of the even and the odd numbers, respectively.

The proof is straightforward (it suffices to follow the same arguments developed for Lemma 3.2), so we omit it.

**Remark 3.7** In the special case \(l = 1, 2, 4\), Lemma 3.6 gives

\[
C\psi_1(y) = \psi_1(y) + 2\alpha \Delta y 1 + c^M_{-1} e_{-M} + c_{-1} e_M,
\]

\[
C\psi_2(y) = \psi_2(y) + 4\alpha \Delta y \psi_1(y) + 2(\beta + |\alpha|) \Delta y^2 1 + c^M_{-2} e_{-M} + c_{-2} e_M
\]

\[
C\psi_4(y) = \psi_4(y) + 8\alpha \Delta y \psi_3(y) + 12(\beta + |\alpha|) \Delta y^2 \psi_2(y) + 8\alpha \Delta y^3 \psi_1(y) + 2(\beta + |\alpha|) \Delta y^4 1
\]

\[
+ c^M_{-4} e_{-M} + c_{-4} e_M.
\]

In particular, since \((\psi_i(y))_i = 0\) for every \(l \geq 1\) and assuming \(|i| < M\), the \(i\)th entries of the above sequences are given by

\[
(C\psi_1(y))_i = h\mu_Y(v),
\]

\[
(C\psi_2(y))_i = h\tilde{\rho}^2 v + h\Delta y |\mu_Y(v)|,
\]

\[
(C\psi_4(y))_i = h\Delta y^2 \tilde{\rho}^2 v + h\Delta y^3 |\mu_Y(v)|,
\]

in which we have inserted the formulas for \(\alpha\) and \(\beta\) in (3.6).

### 3.2 The associated Markov chain and the convergence of the hybrid algorithm

We denote, as in Section 2.2, \((\bar{V}^h)_n=0,\ldots, N\) with \(N h = T\), the Markov chain approximating the volatility process \(V\) introduced in Appolloni et al. (2014). We recall that the state-space at step \(n\) is given by \(\bar{V}^h_n\) defined in (2.15). We define now the \(Y\)-component of our Markov chain.

First, given the time-step \(h\), we set up the dependence on \(h\) for the space step \(\Delta y\), the number \(M\) giving the points of the grid \(Y_M = \{y_i = Y_0 + i\Delta y; i = -M, \ldots, M\}\) and the threshold \(\epsilon\) that allows us to use the explicit or the implicit finite difference method. So, we assume that

\[
\Delta y \equiv \Delta y_h = c_y h^p, \quad M \equiv M_h = c_M h^{-q}, \quad \epsilon \equiv \epsilon_h = c_\epsilon h^p
\]

where \(c_M > 0\) and the constants \(c_y, c_\epsilon, p, q > 0\) are chosen as follows:

\[
p < 1, \quad q > p, \quad \frac{2c_y}{\tilde{\rho}^2} \left| r - \delta - \frac{\rho}{\sigma} \kappa \theta \right| < c_\epsilon \quad \text{or}
\]

\[
p = 1, \quad q > p, \quad \frac{2c_y}{\tilde{\rho}^2} \left| r - \delta - \frac{\rho}{\sigma} \kappa \theta \right| < c_\epsilon < \left( \frac{1}{2} - \frac{1}{c_y} \left| r - \delta - \frac{\rho}{\sigma} \kappa \theta \right| \right) \frac{c_y^2}{\tilde{\rho}^2},
\]

\[
\text{with } c_y > 0, \quad \tilde{\rho} > 0, \quad \kappa > 0, \quad \sigma > 0, \quad \delta > 0,
\]

\[
\text{and } \rho < 0.
\]
in which the parameters $\kappa, \theta, r, \delta, \sigma$ and $\rho$ come from our model; see Section 2. Let us stress that the last constraint in (3.22) can be really satisfied, for example by choosing $c_\gamma > 4|r - \delta - (\rho/\sigma)\kappa\theta|$. We also note that (3.21) and (3.22) give $M_h \Delta y_h = O(h^{-q+p}) \to \infty$ as $h \to 0$, so that $\mathcal{V}_h = \{y_i = Y_0 + i \Delta y_h; i = -M_h, \ldots, M_h\} \uparrow \mathbb{R}$ as $h \to 0$. Moreover, one has the following proposition.

**Proposition 3.8** Let (3.21) and (3.22) hold and let $\beta = \beta_0$ and $\alpha = \alpha_h$ be given in (3.6) with the constraints (3.21) and (3.22). Then there exists $h_0 > 0$ such that, for every $h < h_0$, one has the following conditions:

(i) if $v > \epsilon_h$, then $\beta_h > |\alpha_h|$;

(ii) if $v \leq \epsilon_h$, then $2\beta_h + 2|\alpha_h| < 1$.

**Proof.** By formula (2.7), we write $|\mu_\varphi(v)| \leq a_\gamma + b_\gamma v$, with $a_\gamma = |r - \delta - (\rho/\sigma)\kappa\theta|$ and $b_\gamma = |(\rho/\sigma)\kappa| - 1$.

(i) One has $\beta_h = h^{1-2p}/2c_\gamma^2 \rho^2 v$ and $|\alpha_h| \leq (h^{1-p}/2c_\gamma)(a_\gamma + b_\gamma v)$, so $\beta_h > |\alpha_h|$ if $(h^{1-2p}/2c_\gamma^2)\rho^2 v > (h^{1-p}/2c_\gamma)(a_\gamma + b_\gamma v)$, that is if $v(\rho^2/c_\gamma - \rho^2 b_\gamma) > \alpha_h h^p$. Since $v > \epsilon_h = c_\gamma h^p$, we get that $\beta_h > |\alpha_h|$ if $c_\epsilon(\rho^2/c_\gamma - \rho^2 b_\gamma) > \alpha_h$, and this holds, for every $h$ small because, by (3.22), one has $c_\epsilon(\rho^2/c_\gamma) > 2\alpha_h$.

(ii) One has $2\beta_h + 2\alpha_h \leq (h^{1-2p}/2c_\gamma^2)\rho^2 c_\gamma h^p + (h^{1-p}/2c_\gamma)(a_\gamma + b_\gamma c_\gamma h^p) = h^{-p}(\rho^2/c_\gamma^2) c_\epsilon + a_\gamma/c_\gamma + (b_\gamma/c_\gamma)h^p$. If $p < 1$, $2\beta_h + 2\alpha_h < 1$ for every $h$ small. If instead $p = 1$, one has $2\beta_h + 2\alpha_h \leq (\rho^2/c_\gamma^2) c_\epsilon + a_\gamma/c_\gamma + (b_\gamma/c_\gamma)h$ and if (3.22) holds, then $2\beta_h + 2\alpha_h < 1 + (b_\gamma\rho^2/2a_\gamma c_\gamma)h < 1$ for every $h$ small enough. \[\square\]

Now, Proposition 3.8 ensures that there exists $h_0 > 0$ such that, for every $h < h_0$ and for every $v \in \cup_{n=0}^{N} \mathcal{V}_n$, the matrix $A^{-1}$ discussed in Section 3.1.1 and the matrix $C$ discussed in Section 3.1.2 are both well defined and are stochastic matrices. So, for $h$ small, that is $h < h_0$ with $h_0$ as in Proposition 3.8, we define $\Pi^\hbar(v)$ as follows:

— if $v > \epsilon_h$, $\Pi^\hbar(v)$ is the inverse of the matrix $A$ in (3.8);

— if $v \leq \epsilon_h$, $\Pi^\hbar(v)$ is the matrix $C$ in (3.18).

As a consequence, we can assert that, for every $v \in \cup_{n=0}^{N} \mathcal{V}_n$, $\Pi^\hbar(v) = (\Pi^\hbar(v)_{ij})_{i,j \in J_{M_h}}$ is a stochastic matrix. We now define the following transition probability law: at time-step $n \in \{0, 1, \ldots, N\}$, for $(y, v_{n,k}) \in \mathcal{V}_n \times \mathcal{V}_n$ we set $\mu^\hbar(y, v_{n,k}; \cdot)$ the probability law in $\mathbb{R}^2$ as in (2.20), that is

$$
\mu^\hbar(y_i, v_{n,k}; A \times B) = \sum_{k^* \in \{1, \ldots, K^\hbar(n,k)\}} \Pi^\hbar(v_{n,k})_{ij} \delta_{\{y_i\}}(A) \delta_{\{v_{n+1,k^*}\}}(B).
$$

So, we call $\tilde{X}^\hbar = (\tilde{X}^\hbar)_n = (0, 1, \ldots, N)$ the two-dimensional Markov chain having (3.23) as its transition probability law at time-step $n \in \{0, 1, \ldots, N\}$, that is

$$
\mathbb{P}(\tilde{X}_{n+1} = (y_j, v_{n+1,k^*}) | \tilde{X}_n = (y_i, v_{n,k})) = \begin{cases} 
\Pi^\hbar(v_{n,k})_{ij} | p^\hbar_{k^*}(n,k) & \text{if } k^* = k^\hbar(n,k) \\
\Pi^\hbar(v_{n,k})_{ij} | p^\hbar_{k^*}(n,k) & \text{if } k^* = k^\hbar(n,k) \\
0 & \text{otherwise},
\end{cases}
$$

for every $(y_i, v_{n,k}) \in \mathcal{V}_n \times \mathcal{V}_n$ and $(y_j, v_{n+1,k^*}) \in \mathcal{V}_n \times \mathcal{V}_{n+1}$. Since $\sum_j \Pi^\hbar(v)_{ij} = 1$, one gets that the second component of $\tilde{X}^\hbar$ is a Markov chain itself and it equals, in law, to $\tilde{V}^\hbar$. So, we write $\tilde{X}^\hbar = (\tilde{Y}^\hbar, \tilde{V}^\hbar)$.
and, for every function \( f : \mathbb{R}^2 \to \mathbb{R} \), we have

\[
\mathbb{E}(f(\tilde{Y}^h_{n+1}, \tilde{V}^h_{n+1}) | (\tilde{Y}^h_n, \tilde{V}^h_n) = (y_i, v_{n,k})) = \sum_{k^* \in \{K^0_n, K^1_n, K^2_n\}} \sum_{j \in J_{k^*}} f(y_j, v_{n+1,k^*}) \Pi^h(v_{n,k}, i, j | k^*). \tag{3.24}
\]

Coming back to the discussion in Section 2.2, by (3.24) we can assert that our algorithm actually approximates the diffusion pair \( X = (Y, V) \) with the Markov chain \( \tilde{X}^h = (\tilde{Y}^h, \tilde{V}^h) \). So, we set \( X^h = (Y^h, V^h) \) as the piecewise constant and càdlàg interpolation in time of \( \tilde{X}^h \), that is

\[
X^h_t = \tilde{X}^h_n, \quad t \in [nh, (n + 1)h), \quad n = 0, 1, \ldots, N - 1. \tag{3.25}
\]

We set \( D([0, T]; \mathbb{R}^2) \) the space of the \( \mathbb{R}^2 \)-valued and càdlàg functions on the interval \([0, T]\), which we assume to be endowed with the Skorohod topology (see e.g. Billingsley, 1968). Our main result is the following theorem.

**Theorem 3.9** Under (3.21) and (3.22), as \( h \to 0 \), the family \( \{X^h\}_h = \{Y^h, V^h\}_h \) defined through (3.25) and (3.24) weakly converges in the space \( D([0, T]; \mathbb{R}^2) \) to the diffusion process \( X = (Y, V) \) solution to (2.4) and (2.5).

**Proof.** The idea of the proof is standard; see e.g. Nelson & Ramaswamy (1990) or also classical books such as Billingsley (1968), Ethier & Kurtz (1986) or Stroock & Varadhan (1979).

To simplify the notations, let us set

\[
\mathcal{M}^{Y}_{n,i,k}(h; l) = \mathbb{E}((\tilde{Y}^h_{n+1} - y_i)_l | (\tilde{Y}^h_n, \tilde{V}^h_n) = (y_i, v_{n,k})), \quad l = 1, 2, 4,
\]

\[
\mathcal{M}^{V}_{n,i,k}(h; l) = \mathbb{E}((\tilde{V}^h_{n+1} - v_{n,k})_l | (\tilde{Y}^h_n, \tilde{V}^h_n) = (y_i, v_{n,k})), \quad l = 1, 2, 4,
\]

\[
\mathcal{M}^{Y,V}_{n,i,k}(h) = \mathbb{E}((\tilde{Y}^h_{n+1} - y_i)(\tilde{V}^h_{n+1} - v_{n,k}) | (\tilde{Y}^h_n, \tilde{V}^h_n) = (y_i, v_{n,k})).
\]

It is clear that \( \mathcal{M}^{Y}_{n,i,k}(h; l) \) is the local moment of order \( l \) at time \( nh \) related to \( Y \), as well as \( \mathcal{M}^{V}_{n,i,k}(h; l) \) is related to the component \( V \) and \( \mathcal{M}^{Y,V}_{n,i,k}(h) \) is the local cross-moment of the two components at the generic time step \( n \). So, the proof reduces to checking that, for fixed \( r_s, v_s > 0 \) and setting \( \Lambda_s = \{(n, i, k) : v_{n,k} \leq v_s, |y_i| \leq r_s\} \), the following properties (i)–(iii) hold:

(i) (convergence of the local drift):

\[
\lim_{h \to 0} \sup_{(n,i,k) \in \Lambda_s} \frac{1}{h} |\mathcal{M}^{Y}_{n,i,k}(h; 1) - (\mu_Y)_{n,k} h| = 0,
\]

\[
\lim_{h \to 0} \sup_{(n,i,k) \in \Lambda_s} \frac{1}{h} |\mathcal{M}^{V}_{n,i,k}(h; 1) - (\mu_V)_{n,k} h| = 0;
\]

where we have set \((\mu_Y)_{n,k} = \mu_Y(v_{n,k})\) and \((\mu_V)_{n,k} = \mu_V(v_{n,k})\);
(ii) (convergence of the local diffusion coefficient):

\[
\lim_{h \to 0} \sup_{(n,i,k) \in A_*} \frac{1}{h} |\mathcal{M}_{n,i,k}^V(h; 2) - \bar{\rho}^2 v_{n,k} h| = 0,
\]

\[
\lim_{h \to 0} \sup_{(n,i,k) \in A_*} \frac{1}{h} |\mathcal{M}_{n,i,k}^V(h; 2) - \sigma^2 v_{n,k} h| = 0
\]

\[
\lim_{h \to 0} \sup_{(n,i,k) \in A_*} \frac{1}{h} |\mathcal{M}_{n,i,k}^{V,V}(h)| = 0;
\]

(iii) (fast convergence to 0 of the fourth-order local moments)

\[
\lim_{h \to 0} \sup_{(n,i,k) \in A_*} \frac{1}{h} \mathcal{M}_{n,i,k}^V(h; 4) = 0,
\]

\[
\lim_{h \to 0} \sup_{(n,i,k) \in A_*} \frac{1}{h} \mathcal{M}_{n,i,k}^V(h; 4) = 0.
\]

We recall that the \(V\)-component of the two-dimensional Markov chain is a Markov chain itself and we have

\[
\mathcal{M}_{n,i,k}^V(h; l) \equiv \mathcal{M}_{n,k}^V(h; l) = \mathbb{E}((\bar{\nu}_{n+1}^h - v_{n,k})^l \mid \bar{\nu}_n^h = v_{n,k}), \quad l = 1, 2, 4.
\]

The limits in (i)–(iii) containing \(\mathcal{M}_{n,k}^V(h; l)\) for \(l = 1, 2, 4\), have been already proved in Appolloni et al. (2014) (see Theorem 7 therein), so we prove the validity of the remaining limits.

We set \(\psi_i^j(y)\) the vector in \(\mathbb{R}^{2M_k+1}\) given by \((\psi_i^j(y))_j = (y_j - y_i)^j, j \in J_{M_k}\). From (3.24) we get

\[
\mathcal{M}_{n,i,k}^V(h; l) = \sum_{y \in \mathbb{R}^n} \Pi^h(v_{n,k})_{ij}(\psi_i^j(y))_{ij} = (\Pi^h(v_{n,k})\psi_i^j(y))_{ij}
\]

and we note that the above quantity has been already discussed in the previous sections. We set \(A_* = A_{*,1,h} \cup A_{*,2,h}\), with

\[
A_{*,1,h} = \{(n,i,k) : \epsilon_h < v_{n,k} \leq v_*, |y_i| \leq r_*\},
\]

\[
A_{*,2,h} = \{(n,i,k) : v_{n,k} \leq \epsilon_h, |y_i| \leq r_*\}.
\]

For \((n,i,k) \in A_{*,1,h}\), \(\Pi^h(v_{n,k})\) is the inverse of the matrix \(A\) in (3.8). So, by using (3.17), we have

\[
\mathcal{M}_{n,i,k}^V(h; 1) = (A^{-1}\psi_1^i(y))_j = h(\mu_Y)_{n,k} + (\phi_{1,M_k}^i)_i,
\]

\[
\mathcal{M}_{n,i,k}^V(h; 2) = (A^{-1}\psi_2^i(y))_j = h\bar{\rho}^2 v_{n,k} + 2h\Delta y_h(\mu_Y)_{n,k} + (\phi_{2,M_k}^i)_i,
\]

\[
\mathcal{M}_{n,i,k}^V(h; 4) = (A^{-1}\psi_4^i(y))_j = h\Delta^2 y_h^2(\rho^2 v_{n,k} + 8h\Delta y_h(\mu_Y)_{n,k}^2 + 24h^3(\mu_Y)_{n,k}^3 + 6h^2\rho^4 v_{n,k}^3
\]

\[+ 24 \frac{h^4}{\Delta y_h} \rho^2 v_{n,k}(\mu_Y)_{n,k}^3 + (\phi_{3,M_k}^i)_i,
\]

\(\phi_{l,M_k}^i(y)\) being given in (3.15). In Lemma 3.10, we prove that, for \(l \leq 4\),

\[
\sup_{(n,i,k) \in A_{*,1,h}} \frac{1}{h} |(\phi_{l,M_k}^i(y))_i| \to 0 \quad \text{as} \ h \to 0.
\]
And since \((\mu_Y)_{n,k}\) is bounded on \(A_s\), the limits in (i)–(iii) associated with \(\mathcal{M}_{n,i,k}^Y(h;l), l = 1, 2, 4\), hold uniformly in \(A_{s,1,h}\). We prove the same on the set \(A_{s,2,h}\). For \((n, i, k) \in A_{s,2,h}\), the matrix \(\Pi^h(n_{i,k})\) to be taken into account is given by the matrix \(C\) in (3.18). Moreover, for \(h\) small enough, we have that if \((n, i, k) \in A_{s,2,h}\), then \(|i| < M_h\). Therefore, by (3.20) we obtain

\[
\begin{align*}
\mathcal{M}_{n,i,k}^Y(h; 1) &= (C \psi_1^i(y))_i = h(\mu_Y)_{n,k}, \\
\mathcal{M}_{n,i,k}^Y(h; 2) &= (C \psi_2^i(y))_i = h^2 \psi_2^i(y)_{n,k} + h\Delta y_h(\mu_Y)_{n,k}, \\
\mathcal{M}_{n,i,k}^Y(h; 4) &= (C \psi_4^i(y))_i = h\Delta y_h^2 \psi_4^i(y)_{n,k} + h\Delta y_h^3(\mu_Y)_{n,k}
\end{align*}
\]

and again the limits in (i)–(iii) concerning \(\mathcal{M}_{n,i,k}^Y(h;l), l = 1, 2, 4\), hold uniformly in \(A_{s,2,h}\). It remains to study the cross-moment. By using (3.24), it is given by

\[
\mathcal{M}_{n,i,k}^{Y,V}(h) = \mathcal{M}_{n,i,k}^Y(h; 1) \mathcal{M}_{n,k}^V(h; 1)
\]

and the convergence as in (ii) immediately follows from the already proved limits in (i). 

In order to conclude, we only need to prove the following lemma.

**Lemma 3.10** Assume that (3.21) and (3.22) both hold. Let \(v_s, r_s > 0\) and set

\[
A_{s,1,h} = A_s = \{(n, i, k) : \epsilon_h < v_{n,k} \leq v_s, |y_i| \leq r_s\}
\]

Then one has

\[
\lim_{h \to 0} \sup_{(n,i,k) \in A_{s,1,h}} \frac{1}{h} |(\phi_{i,M}^h(y))_i| = 0,
\]

for every \(l \leq 4\), where \(\phi_{i,M}^h(y)\) is defined in (3.15) with \(M = M_h\).

**Proof.** We use Proposition 3.5. Here, \(C\) denotes a positive constant that may vary from line to line.

Under (3.21) and (3.22), for \((n, i, k) \in A_{s,1,h}\), we have already observed that \(\beta_h > |\alpha_h|, \alpha_h, \beta_h\) being given in (3.6), and the constraints \(\Delta y_h \leq 1\) and \(M_h \Delta y_h \geq 1\) are trivially satisfied for \(h\) small. Moreover, on the set \(A_{s,1,h}\), there exists \(C > 0\) such that, for every \(l \leq 4\),

\[
l^{l+2}(\beta_h \Delta y_h^2 + |\alpha_h| \Delta y_h) \leq Ch \leq 1.
\]

for \(h\) small enough. And since \(\beta_h + |\alpha_h| \leq Ch^{1-2p}\), we also have

\[
1 - \frac{1}{1 + \beta + |\alpha|} \leq 1 - \frac{1}{1 + Ch^{1-2p}}.
\]

Then, by applying Proposition 3.5, we can write

\[
|(\phi_{i,M}^h(y))_i| \leq Ch^{-(q-p)l+1-2p} \left(1 - \frac{1}{1 + Ch^{1-2p}}\right)^{M_h - i}.
\]

Now, on the set \(A_{s,1,h}\) we have \(|y_i| \leq r_s\), so that \(|i| \leq (r_s + |Y_0|) / \Delta y_h \leq Ch^{-p}\). And by recalling that \(q > p\), we get \(M_h - i \geq C(h^{-q} - h^{-p}) \geq ch^{-q}\) for some \(c > 0\). So, we have proved that there exist \(C, c > 0\).
such that, for every $h$ close to 0,

$$\sup_{(n,i,k)\in\Lambda_0} \frac{1}{h} |(\phi_{i,M}^h (y))_i| \leq Ch^{-(q-p)l-2p} \left( 1 - \frac{1}{1 + Ch^{l-2p}} \right) =: CI_h \quad \text{for } l \leq 4.$$ 

We see now that, $I_h \rightarrow 0$ as $h \rightarrow 0$ under our constraints $q > p > 0$ and $p \leq 1$. In fact, if $1 - 2p \geq 0$, then $1 - 1/(1 + Ch^{l-2p}) \leq 1 - 1/(1 + C)$ for every $h$ small enough, and $I_h \rightarrow 0$. Otherwise, if $1 - 2p < 0$, then we write

$$\log I_h \simeq -((q - p)l + 2p) \log h - \frac{ch^{-q}}{1 + Ch^{l-2p}}$$

$$= h^{-(q-2p+1)} \left[ -((q - p)l + 2p) h^{q-2p+1} \log h - \frac{c}{h^{2p-1} + C} \right].$$

Since $q - 2p + 1 = q - p - p + 1 > -p + 1 \geq 0$, then $q - 2p + 1 > 0$ and $\log I_h \rightarrow -\infty$. The statement now holds. \qed

**Remark 3.11** Theorem 3.9 proves the convergence of the algorithm in the case we introduce suitable boundary conditions in the finite difference component of our procedure. We stress that our Neumann-type conditions (3.4) and (3.5) have brought to the matrices $A$ and $C$ given by (3.8) and (3.18), respectively. But these conditions may be replaced by other types of boundary conditions that can be handled by using Proposition A.1 (see Appendix A). For example, in the case of time-independent Dirichlet-type conditions, that is conditions that fix the value of the function at the end-points $y_{\pm M}$ of the grid (this typically appears in financial problems), one could set $u_{\pm M}^n = u_{\pm M}^{n+1}$ for every $n$. This means that the matrices $A$ and $C$ to be taken into account become

$$A = \begin{pmatrix} 1 & 0 & & & \alpha - \beta & 1 + 2\beta & -\alpha - \beta \\ & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots \\ & & & \alpha - \beta & 1 + 2\beta & -\alpha - \beta & 0 & 1 \end{pmatrix}, \quad \text{(3.27)}$$

and

$$C = \begin{pmatrix} 1 & 0 & & & \beta + 2|\alpha|\delta_{|\alpha|<0} & 1 - 2\beta - 2|\alpha| & \beta + 2|\alpha|\delta_{|\alpha|>0} \\ & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots \\ & & & & \beta + 2|\alpha|\delta_{|\alpha|<0} & 1 - 2\beta - 2|\alpha| & \beta + 2|\alpha|\delta_{|\alpha|>0} & 0 & 1 \end{pmatrix}, \quad \text{(3.28)}$$

respectively. So, again Proposition A.1 in Appendix A can be used—for details; see Remark A.3. As a consequence, the boundary contributions can be handled as in Lemma 3.10, and the proof of Theorem 3.9 can be easily reproduced.

**Remark 3.12** We also recall that the use of a boundary is a numerical requirement which is necessary to solve the problem in practice. However, one could prove the convergence result even in the whole grid $\gamma^h = \{y_i = Y_0 + i\Delta y_h\}_{i\in\mathbb{Z}}$, that is by considering the solutions to (3.1) and (3.2) without linking to these
equations any boundary condition. In fact, under the requirements of Theorem 3.9, one can prove that the inverse of the implicit difference linear operator $A$ associated with (3.1) remains a Markovian transition function, as well as the explicit difference linear operator $C$ related to (3.2). Moreover, formulas similar to (3.17) and (3.20) hold, with a simplification due to the fact that boundary contributions do not exist. This means that the arguments used to prove Theorem 3.9 can be applied also on the infinite grid $\mathcal{Y}^h = \{y_i = Y_0 + i\Delta y_i\}_{i \in \mathbb{Z}}$.

3.3 Convergence of American prices

Owing to Theorem 3.9, we can deal with the convergence of the price evaluated on our Markov chain to the one computed on the Heston model.

The problem of pricing European options is simple when the payoff is not too complicated. In fact, let $T$ denote the maturity date and $f$ be the payoff function, that is $f : D([0, T]; \mathbb{R}) \to \mathbb{R}_+$. First, we write the payoff in terms of the transformation (2.3), so we get the transformed payoff function $g(y_v) = f(e^{v\theta/\sigma^2})$, $(y_v) \in D([0, T]; \mathbb{R}^2)$, and the associated option prices on the continuous and the discrete model as seen at time 0 are given by

$$P_{Eu} = \mathbb{E}(\tilde{g}(Y, V)) \quad \text{and} \quad P_{Eu}^h = \mathbb{E}(\tilde{g}(Y^h, V^h)),$$

respectively, $\tilde{g}$ denoting the discounted payoff, i.e. $\tilde{g} = e^{-rT}g$ (it is clear that the writing $\mathbb{E}$ for both prices is an abuse of notations, since in principle one should use the notations $\mathbb{E}_p$ and $\mathbb{E}_{ph}$ related to the measures $\mathbb{P}$ and $\mathbb{P}^h$ of the probability space where the processes $(Y, V)$ and $(Y^h, V^h)$ are defined). Now, the weak convergence in Theorem 3.9 ensures the convergence $P_{Eu}^h \to P_{Eu}$ of the European price when the discounted payoff-function fulfills the following requests: $(y, v) \mapsto \tilde{g}(y, v)$ is continuous and there exists $a > 0$ and $h_*>0$ such that

$$\sup_{h<h_*} \mathbb{E}|(\tilde{g}(Y^h, V^h)|^{1+a}) < \infty.$$

This a consequence of standard results on the convergence of the expectations for sequences of random variables which are weakly convergent and satisfy uniform integrability properties.

As for American style options, even for simple payoffs things are more difficult because of the presence of optimal stopping times. However, due to the results in Amin & Khanna (1994) (but see also Lamberton & Pagès, 1990), we can deduce the convergence of the prices for suitable payoffs. In fact, let $f : [0, T] \times D([0, T]; \mathbb{R}) \to \mathbb{R}_+$ denote a payoff function. By passing to the pair $(Y, V)$ and by considering the resulting discounted payoff function $\tilde{g}(t, y, v) = e^{-\sigma^2f(t, e^{y\theta/\sigma^2})}$, the associated option prices on the continuous and the discrete model as seen at time 0 are given by

$$P_{Am} = \sup_{\tau \in \mathcal{G}_{0T}} \mathbb{E}(\tilde{g}(\tau, Y, V)) \quad \text{and} \quad P_{Am}^h = \sup_{\eta \in \mathcal{G}_{0T}^h} \mathbb{E}(\tilde{g}(\eta, Y^h, V^h)),$$

where $\mathcal{G}_{0T}$ and $\mathcal{G}_{0T}^h$ denote the stopping times in $[0, T]$ w.r.t. the filtration $\mathcal{F}_t = \sigma((Y_s, V_s) : s \leq t)$ and $\mathcal{F}^h_t = \sigma((Y^h_s, V^h_s) : s \leq t)$, respectively.

Consider the following assumptions:

(H1) $\tilde{g} : [0, T] \times D([0, T]; \mathbb{R}^2) \to \mathbb{R}_+$ is a continuous function (in the product topology) and, for every $\xi, \eta \in D([0, T]; \mathbb{R}^2)$ such that $\xi_s = \eta_s$ for every $s \in [0, t]$, then $\tilde{g}(t, \xi) = \tilde{g}(t, \eta)$;

(H2) there exist $a > 0$ and $h_*>0$ such that $\sup_{h<h_*} \mathbb{E}(\sup_{t \leq T} |\tilde{g}(t, Y^h, V^h)|^{1+a}) < \infty$.
Let $\eta^h$ denote the optimal stopping time related to the discrete problem, that is

$$P^h_{Am} = \mathbb{E}(\tilde{g}(\eta^h, Y^h, V^h)).$$

Then, by using the arguments in Amin & Khanna (1994), properties (i)–(iii) in the proof of Theorem 3.9 ensure the existence of $\eta \in G_{0,T}$ such that the triple $(\eta^h, Y^h, V^h)$ converges in law to $(\eta, Y, V)$—this is very roughly speaking: in the words by Amin and Khanna, $\eta$ is ‘in some appropriate sense a legitimate stopping time w.r.t. the filtration generated by $(Y, V)$’, and to be precise one should refer to a further probability space (for technical detail, see the discussion at pp. 299–300 and in particular Theorem 4.1 in Amin & Khanna, 1994). We stress that in Amin & Khanna (1994) this result is based on two starting assumptions: the (global) Lipschitz continuity and the sublinearity property for both the drift and the diffusion coefficient of the pair $(Y, V)$. Here, the Lipschitz continuity property does not hold because of the presence of the square-root function, so the diffusion coefficient is only Hölder continuous. Nevertheless, they use such a property just to ensure the existence of the strong solution, a condition that holds here, and use specifically the sublinearity property of the drift coefficients, which holds here. So, their arguments apply as well. Once the weak convergence of the triple is achieved, the plan to prove the convergence of the prices is the following. Under (H1) one gets that $\{\tilde{g}(\eta^h, Y^h, V^h)\}_h$ converges in law, as $h \to 0$, to $\tilde{g}(\eta, Y, V)$. Moreover, (H2) implies that the $\{\tilde{g}(\eta^h, Y^h, V^h)\}_h$ is a uniformly integrable family of random variables, and this suffices to get the convergence of the expectations. Finally, one gets $P_{Am} = \mathbb{E}(\tilde{g}(\eta, Y, V))$, and the convergence $P^h_{Am} \to P_{Am}$ of the American prices is achieved.

As an immediate consequence, both European and American put options can be numerically evaluated by means of the approximating algorithm described in Section 2.2 (and, due to the call-put parity formula, also European call options can be numerically priced with our method). However, in next Section 4, we apply our procedure to numerically price European/American barrier options written on the Heston model, that is options which are much more sophisticated and, for example, do not fulfil (H1). Nevertheless, the numerical experiments give a real evidence of the goodness (that is, convergence) of our algorithm also in this case.

4. Numerical results

In this section, we provide numerical results in order to assess the efficiency and the robustness of our hybrid tree/finite difference method in the case of plain vanilla options and barrier options. All the computations have been performed in double precision on a PC 1.7 GHz Intel Core I5 with a 4 Gb of RAM.

The hybrid tree/finite difference algorithm introduced in Section 2.2 is here called HTFD and it is split in two different variants:

- HTFD1: when the number $N_t$ of time steps and the number $N_S$ of space steps are different;
- HTFD2: when the number $N_t$ of time steps and the number $N_S$ of space steps are equal.

4.1 European and American vanilla options: comparison with tree methods

We compare here the performance of the hybrid tree/finite difference algorithm HTFD with the tree method of Vellekoop and Nieuwenhuis (called VN) in Vellekoop & Nieuwenhuis (2009) for the computation of European and American options in the Heston model. All the numerical results of this section (except for our method) are obtained by using the Premia software (PREMIA). In the European and American option contracts we are dealing with, we consider the following set of parameters: initial
Table 1 Prices of European put options. \( \sigma = 0.04, 0.5, 1. \) \( S_0 = 100, \) \( K = 100, \) \( T = 1, r = \log(1.1), \delta = 0, V_0 = 0.1, \theta = 0.1, \kappa = 2 \) and \( \rho = -0.5. \)

| \( \sigma \) | \( N_S \) | VN | HTFD1 | HTFD2 | CF |
|-------|-------|-----|------|------|----|
| 0.04  | 50    | 8.040982 | 7.934492 | 7.911034 |
|       | 100   | 8.021780 | 7.970437 | 7.970437 |
|       | 200   | 8.003938 | 7.978890 | 7.983188 | 7.994716 |
|       | 400   | 7.984248 | 7.980984 | 7.990825 |
| 0.50  | 50    | 8.148234 | 7.758954 | 7.746533 |
|       | 100   | 7.727191 | 7.804520 | 7.804520 |
|       | 200   | 7.813599 | 7.816749 | 7.821404 | 7.8318540 |
|       | 400   | 7.910909 | 7.818596 | 7.827805 |
| 1.00  | 50    | 6.586889 | 7.214303 | 7.247748 |
|       | 100   | 7.114225 | 7.225292 | 7.225292 |
|       | 200   | 7.964052 | 7.228235 | 7.229139 | 7.2313083 |
|       | 400   | 6.639931 | 7.224356 | 7.233742 |

In order to study the numerical robustness of the algorithms, we choose three different values for \( \sigma: \) we set \( \sigma = 0.04, 0.5, 1. \) We first consider the case \( \sigma = 0.04, \) that is \( \sigma \) close to zero (which implies that the Heston PDE is convection-dominated in the \( V \)-direction). Moreover, for \( \sigma = 1, \) we stress that the Feller condition \( 2\kappa\theta \geq \sigma^2 \) is not satisfied.

In the (pure) tree method VN, we fix the number of points in the \( V \) coordinate as \( N_V = 50, \) with varying number of time and space steps: \( N_t = N_S = 50, 100, 200, 400. \)

As already mentioned, the numerical study of the hybrid tree/finite difference method HTFD is split in two cases: HTFD1 refers to the (fixed) number of time steps \( N_t = 100 \) and varying number of space steps \( N_S = 50, 100, 200, 400; \) we add the situation HTFD2 where the number of time steps is equal to the number of space steps \( N_t = N_S = 50, 100, 200, 400. \)

Table 1 reports European put option prices. Comparisons are given with a benchmark value obtained using the Carr–Madan pricing formula CF in Carr & Madan (1999) that applies fast Fourier transform methods.

In Table 2, we provide results for American put option prices. In this case we use a benchmark from the Monte Carlo Longstaff–Schwartz algorithm, called MC–LS, as in Longstaff & Schwartz (2001), with a huge number of Monte Carlo simulation (1 million iterations) which are done by means of the accurate (Alfonsi, 2010) discretization scheme for the CIR process with \( M = 100 \) discretization time steps and Bermudan exercise dates. We recall that the Alfonsi method provides a Monte Carlo weak second-order scheme for the CIR process, without any restriction on the parameters.

Table 3 refers to the computational time cost (in seconds) of the different algorithms for \( \sigma = 0.5 \) in the European case.

The numerical results show that the hybrid tree/finite difference method is very accurate, reliable and efficient. For \( \sigma = 0.5 \) and \( \sigma = 1, \) Tables 1–3 definitely show that both HTFD1 and HTFD2 get more precise pricing results, and need much less CPU time w.r.t. VN. Figure 1 offers a speed precision efficiency graphs for HTFD1, HTFD2 and VN in the low volatility case \( \sigma = 0.04. \) We consider CPU times vs error. The pricing error is computed using the benchmark price obtained with the CF method.
Table 2  Prices of American put options. \( \sigma = 0.04, 0.5, 1 \). \( S_0 = 100, K = 100, T = 1, r = \log(1.1), \delta = 0, V_0 = 0.1, \theta = 0.1, \kappa = 2 \) and \( \rho = -0.5 \).

| \( N_S \) | VN | HTFD1 | HTFD2 | MC–LS |
|---|---|---|---|---|
| \( \sigma = 0.04 \) | 50 | 9.100312 | 8.966651 | 8.932445 |
| | 100 | 9.086233 | 9.016732 | 9.016732 |
| | 200 | 9.073722 | 9.028866 | 9.042581 | 9.074102 |
| | 400 | 9.063396 | 9.031881 | 9.054538 |
| \( \sigma = 0.50 \) | 50 | 9.150887 | 8.763369 | 8.731867 |
| | 100 | 8.982206 | 8.841776 | 8.841776 |
| | 200 | 8.981855 | 8.862606 | 8.878530 | 8.904514 |
| | 400 | 9.058313 | 8.866911 | 8.892583 |
| \( \sigma = 1.00 \) | 50 | 8.588392 | 8.185052 | 8.206052 |
| | 100 | 9.020989 | 8.263395 | 8.263395 |
| | 200 | 9.251595 | 8.281755 | 8.290371 | 8.277985 |
| | 400 | 9.102788 | 8.283214 | 8.304415 |

Table 3  Computational times (in seconds) for European put options for \( \sigma = 0.5 \).

| \( N_S \) | VN | HTFD1 | HTDF2 | CF |
|---|---|---|---|---|
| 50 | 0.11 | 0.02 | 0.007 |
| 100 | 0.42 | 0.04 | 0.040 |
| 200 | 1.73 | 0.08 | 0.380 | 0.018 |
| 400 | 7.06 | 0.16 | 3.040 |

The accuracy of HTFD is confirmed in Figs 2 and 3 where we compare the shapes of implied volatility smiles across moneyness \( K/S_0 \) and maturities \( T \) for different values of \( \sigma \) using HTFD2 and VN with \( N_S = 200 \). In the VN method, as the vol–vol increases we notice a much more oscillatory behaviour of the implied volatility across moneyness. This behaviour is not really surprising, since in Table 1 we observe an evident deterioration of the accuracy of the results from the VN algorithm when the volatility rises.

4.2  Greeks

We consider here the computation of different hedge ratios. In particular, we compute the following Greeks: Delta, Gamma, Theta and Vega.

Concerning Delta and Gamma, we can calculate the associated values by using the already built finite difference stencils. In particular, we use a unique HTFD run to compute price, Delta and Gamma, the latter two given by

\[
\text{Delta} = \frac{\text{Price}(0, y_{i+1}, V_0) - \text{Price}(0, y_{i-1}, V_0)}{y_{i+1} - y_{i-1}},
\]

\[
\text{Gamma} = 2 \frac{(\text{Price}(0, y_{i+1}, V_0) - \text{Price}(0, y_i, V_0))/y_{i+1} - y_i - (\text{Price}(0, y_i, V_0) - \text{Price}(0, y_{i-1}))/y_i - y_{i-1}}{y_{i+1} - y_{i-1}}.
\]
A HYBRID APPROACH FOR THE IMPLEMENTATION OF THE HESTON MODEL

Fig. 1. Error vs CPU time for $\sigma = 0.04$.

Fig. 2. Implied volatility vs maturity for $\sigma = 0.04$, $\sigma = 0.5$ and $\sigma = 1$.

Fig. 3. Implied volatility vs moneyness for $\sigma = 0.04$, $\sigma = 0.5$ and $\sigma = 1$. 

Downloaded from https://academic.oup.com/imaman/article-abstract/28/4/467/2884335 by guest on 28 January 2018
Table 4  Deltas of European put options. $\sigma = 0.04, 0.5, 1$. $S_0 = 100$, $K = 100$, $T = 1$, $r = \log(1.1)$, $\delta = 0$, $V_0 = 0.1$, $\theta = 0.1$, $\kappa = 2$ and $\rho = -0.5$

| $\sigma$ | $N_S$ | VN   | HTFD1 | HTFD2 | CF  |
|----------|-------|------|-------|-------|-----|
| 0.04     | 50    | -0.321134 | -0.320047 | -0.320149 |     |
|          | 100   | -0.320299 | -0.319982 | -0.319826 | -0.319751 |
|          | 200   | -0.320199 | -0.319959 | -0.319826 |     |
|          | 400   | -0.319819 | -0.319954 | -0.319803 |     |
| 0.50     | 50    | -0.286876 | -0.281549 | -0.281948 |     |
|          | 100   | -0.281648 | -0.281096 | -0.281096 |     |
|          | 200   | -0.281483 | -0.280987 | -0.280825 | -0.280650 |
|          | 400   | -0.280925 | -0.280940 | -0.280728 |     |
| 1.00     | 50    | -0.282386 | -0.247117 | -0.249431 |     |
|          | 100   | -0.302834 | -0.243152 | -0.243152 |     |
|          | 200   | -0.308659 | -0.242249 | -0.241908 | -0.241012 |
|          | 400   | -0.298714 | -0.241773 | -0.241501 |     |

Table 5  Gammas of European put options. $\sigma = 0.04, 0.5, 1$. $S_0 = 100$, $K = 100$, $T = 1$, $r = \log(1.1)$, $\delta = 0$, $V_0 = 0.1$, $\theta = 0.1$, $\kappa = 2$, $\rho = -0.5$

| $\sigma$ | $N_S$ | VN    | HTFD1 | HTFD2 | CF  |
|----------|-------|-------|-------|-------|-----|
| 0.04     | 50    | 0.011341 | 0.011400 | 0.011456 |     |
|          | 100   | 0.011298 | 0.011344 | 0.011344 |     |
|          | 200   | 0.011297 | 0.011331 | 0.011309 | 0.011282 |
|          | 400   | 0.011300 | 0.011327 | 0.011293 |     |
| 0.50     | 50    | 0.010732 | 0.011059 | 0.011121 |     |
|          | 100   | 0.011188 | 0.010967 | 0.010967 |     |
|          | 200   | 0.011176 | 0.010943 | 0.010919 | 0.010886 |
|          | 400   | 0.011330 | 0.010938 | 0.010899 |     |
| 1.00     | 50    | 0.014029 | 0.011618 | 0.011688 |     |
|          | 100   | 0.013080 | 0.011173 | 0.011173 |     |
|          | 200   | 0.011520 | 0.011103 | 0.011074 | 0.010998 |
|          | 400   | 0.012793 | 0.011074 | 0.011031 |     |

where $\text{Price}(t, y, v)$ denotes the price–function and the $y_i$’s are the points of the finite difference grid. As a consequence, the computational time cost does not increase.

In order to compute Theta and Vega, we use two different HTFD runs and the computational time cost is doubled. Theoretically, we could compute Theta and Vega by using a unique computation. For example, in the Vega case this would require the computation of $(\text{Price}(\Delta t, y_0, v_{1,1}) - \text{Price}(\Delta t, y_0, v_{1,0}))/ (v_{1,1} - v_{1,0})$. But in the high volatility case the increment $v_{1,1} - v_{1,0}$ may be huge and this would lead to inaccuracy. Concerning Theta, additional interpolations in time and space would be required. So, in practice it is more convenient to use double computations both for Vega and Theta.
Therefore, we consider the standard formulas

\[
\text{Theta} = \frac{\text{Price}(\delta, Y_0, V_0) - \text{Price}(0, Y_0, V_0)}{\delta},
\]

\[
\text{Vega} = \frac{\text{Price}(0, Y_0, V_0(1 + \delta)) - \text{Price}(0, Y_0, V_0)}{V_0 \delta},
\]

where in practice we set \(\delta = 10^{-5}\). Note that in order to evaluate \(\text{Price}(\delta, Y_0, V_0)\), we consider a tree for the volatility process in the time interval \([\delta, T]\). Similar computations are used for the tree method VN.
Tables 4–7 report European Greeks results using the same parameters of Table 1, and still confirm the better behavior and the robustness of HTFD method with respect to the VN method, especially as $\sigma$ increases.

4.3  American vanilla options: comparison with finite difference methods

We compare here the performance of the hybrid tree/finite difference algorithm HTFD with various finite difference methods’ pricing results given in Clarke & Parrott (1999), Ikonen & Toivanen (2009), Oosterlee (2003) and Zvan et al. (1998). The model parameter values are $\kappa = 5, \theta = 0.16, \sigma = 0.9, \rho = 0.1, r = 0.1, \delta = 0$ and $V_0 = 0.25$. The strike is $K = 10$ and different values of the initial stock prices $S_0 = 8, 9, 10, 11, 12$ are considered. Comparisons in the American cases are given with the reference values provided in ZFV (Zvan et al., 1998), IT-PSOR (Ikonen & Toivanen, 2009), OO (Oosterlee, 2003) and CP (Clarke & Parrott, 1999) methods. These values with those obtained by the HTFD algorithm are compared in Table 8. In order to study the convergence behaviour of HTFD2, in Table 9 we consider
Table 9  HTFD2-ratio (4.1) for the price of American put options as the starting point $S_0$ varies. Test parameters are $K = 10$, $T = 0.25$, $r = 0.1$, $\delta = 0$, $V_0 = 0.25$, $\theta = 0.16$, $\kappa = 5$, $\rho = 0.1$ and $\sigma = 0.9$

| $N$ | $S_0 = 8$ | $S_0 = 9$ | $S_0 = 10$ | $S_0 = 11$ | $S_0 = 12$ |
|-----|-----------|-----------|------------|------------|------------|
| 200 | 2.194914  | 2.653543  | 2.280589   | 3.181503   | 3.204384   |
| 400 | 2.115892  | 2.322611  | 2.380855   | 2.237965   | 2.423995   |
| 800 | 2.074822  | 2.140352  | 2.165029   | 2.178548   | 2.044881   |

the following convergence ratio proposed in D’Halluin et al. (2005):

$$\text{ratio} = \frac{P_{N/2} - P_{N/4}}{P_N - P_{N/2}}, \quad (4.1)$$

where $P_N$ denotes here the approximated price obtained with $N = N_t = N_S$ number of time steps. Table 9 suggests that the convergence ratio for HTDF2 is linear, as it is expected to be because of the tree contribution (whose error typically behaves linearly). Moreover, Table 9 shows that the numerically observed ratios are very stable, and this gives a strong evidence of the robustness of the method.

4.4 European and American barrier options

We study here the continuously monitored barrier options case and we compare our hybrid tree/finite difference algorithm with the numerical results of the method of lines provided in Chiarella et al. (2012). So, we consider European and American up-and-out call options with the following set of parameters: $K = 100$, $T = 0.5$, $r = 0.03$, $\delta = 0.05$, $V_0 = 0.1$, $\theta = 0.1$, $\kappa = 2$, $\sigma = 0.1$, $\rho = -0.5$. The up barrier is $H = 130$. We choose different values for $S_0$: $S_0 = 80, 100, 120$. We compare the HTFD results with the prices obtained by using the method of lines in Chiarella et al. (2012), called MOL, with mesh parameters 100, 200 and 6400.

Tables 10 and 12 report European and American up-and-out option prices, respectively, while Table 11 refers to the computational time cost (in seconds) of the various algorithms for the European barrier case. The computational time of the MOL method has been taken from Table 2 in Chiarella et al. (2012) and shows that HTDF needs much less runtime.

5. Conclusions

In this paper, we introduce a novel numerical method, a hybrid tree/finite difference method, to compute option prices in the Heston model. The convergence of the algorithm is studied using a Markov chain approach. The numerical results confirm the reliability of the method both in plain vanilla and barrier options cases. Moreover, the numerical results show that the hybrid tree/finite difference method is very efficient, precise and robust for option pricing in the Heston model, also in terms of the computational time cost. Finally, our method may be easily generalized to other existing stochastic volatility models, even in the presence of jumps.
Table 10 Prices of European call up-and-out options. Up barrier is $H = 130$. $K = 100$, $T = 0.5$, $r = 0.03$, $\delta = 0.05$, $V_0 = 0.1$, $\theta = 0.1$, $\kappa = 2$, $\rho = -0.5$ and $\sigma = 0.1$

| $S_0$ | $N_S$ | HTFD1 | HTFD2 | MOL  |
|-------|-------|-------|-------|------|
| 80    | 50    | 0.913861 | 0.875374 |     |
|       | 100   | 0.893484 | 0.893484 |     |
|       | 200   | 0.895127 | 0.900893 | 0.9029 |
|       | 400   | 0.897820 | 0.902770 |     |
| 100   | 50    | 2.635396 | 2.583568 |     |
|       | 100   | 2.606249 | 2.606249 |     |
|       | 200   | 2.597363 | 2.591857 | 2.5908 |
|       | 400   | 2.603679 | 2.594134 |     |
| 120   | 50    | 1.417225 | 1.438429 |     |
|       | 100   | 1.485704 | 1.485704 |     |
|       | 200   | 1.500692 | 1.482193 | 1.4782 |
|       | 400   | 1.504755 | 1.486212 |     |

Table 11 Computational times (in seconds) for European barrier options

| $N_S$ | HTFD1 | HTFD2 | MOL  |
|-------|-------|-------|------|
| 50    | 0.007 | 0.017 |     |
| 100   | 0.132 | 0.132 |     |
| 200   | 0.284 | 1.079 | 268  |
| 400   | 0.535 | 8.901 |     |

Table 12 Prices of American call up-and-out options. Up barrier is $H = 130$. $K = 100$, $T = 0.5$, $r = 0.03$, $\delta = 0.05$, $V_0 = 0.1$, $\theta = 0.1$, $\kappa = 2$, $\rho = -0.5$ and $\sigma = 0.1$

| $S_0$ | $N_S$ | HTFD1 | HTFD2 | MOL  |
|-------|-------|-------|-------|------|
| 80    | 50    | 1.199802 | 1.285959 |     |
|       | 100   | 1.369914 | 1.369914 |     |
|       | 200   | 1.400823 | 1.396628 | 1.4012 |
|       | 400   | 1.400710 | 1.401111 |     |
| 100   | 50    | 8.274116 | 8.269779 |     |
|       | 100   | 8.286667 | 8.286667 |     |
|       | 200   | 8.284054 | 8.294226 | 8.3003 |
|       | 400   | 8.283815 | 8.296745 |     |
| 120   | 50    | 21.943742 | 21.884228 |     |
|       | 100   | 21.820015 | 21.820015 |     |
|       | 200   | 21.785274 | 21.815989 | 21.8216 |
|       | 400   | 21.779648 | 21.804518 |     |
Appendix A. Boundary sensitivity for the implicit finite difference operator

We study here the behaviour of the solution \( x = (x_1, \ldots, x_N) \) of the two following linear systems:

\[
Ax = v_1 \quad \text{(A.1)}
\]
\[
Ax = v_N, \quad \text{(A.2)}
\]

where \( v_i, i = 1, \ldots, N \), denotes the standard orthonormal basis in \( \mathbb{R}^N \), i.e. \( (v_i)_k = 0 \) for \( k \neq i \) and \( (v_i)_i = 1, i = 1, \ldots, N \), and where \( A \) has the following general tridiagonal form:

\[
A = \begin{pmatrix}
a_1 & c_1 & & & \\
b & a & c & & \\
& b & a & c & \\
& & b & a & c \\
b_N & a_N & & & 
\end{pmatrix}.
\] (A.3)

The result we are going to present is due for matrices \( A \) as in (A.3) and that satisfy the hypotheses (P2)–(P3) in the proof of Proposition 3.1, ensuring that they are invertible \( M \)-matrices (see, for instance Berman & Plemmons, 1994).

**Proposition A.1** Suppose that the matrix \( A \) in (A.3) satisfies

\[
a, a_1, a_N > 0, \quad b, c \leq 0, \quad c_1, b_N \leq 0, \quad a > |b| + |c|, \quad a_1 > |c_1|, \quad a_N > |b_N|.
\] (A.4)

Assume, moreover, the following stability conditions on the ‘boundary’ values \( a_1, a_N, c_1 \) and \( b_N \):

\[
\frac{|bc_1|}{a_1} < z_+ \quad \text{and} \quad \frac{|b_N c|}{a_N} < z_+,
\] (A.5)

where \( z_+ = (a + \sqrt{a^2 - 4|bc|})/2 \). Then the solution \( x \) of (A.1) is defined by a sequence \( \{x_k\}_{k=1}^{N} \) of positive terms and there exists a positive value \( \gamma^* > |b| \) such that, for \( k = 2, \ldots, N - 1 \),

\[
x_{N+1-k} \leq x_1 \left( \frac{|b|}{\gamma^*} \right)^{N-k} \quad \text{and} \quad x_N \leq x_1 \frac{|b_N|}{a_N} \left( \frac{|b|}{\gamma^*} \right)^{N-2}.
\] (A.6)

Similarly, for the solution \( x \) of (A.2) it holds \( x_k > 0 \) for all \( k = 1, \ldots, N \), and there exists a positive value \( \gamma^* > |c| \) such that, for \( k = 2, \ldots, N - 1 \),

\[
x_k \leq x_N \left( \frac{|c|}{\gamma^*} \right)^{N-k} \quad \text{and} \quad x_1 \leq x_N \frac{|c_1|}{a_1} \left( \frac{|c|}{\gamma^*} \right)^{N-2}.
\] (A.7)

**Proof.** Let us start by first estimating the solution \( x \) of system (A.2). By applying the Thomas algorithm, also known as tridiagonal matrix algorithm (Thomas, 1949), the solution of (A.2) is given by back
substitutions:

\[ x_N = \frac{1}{\gamma_N}, \quad x_k = \frac{|c|}{\gamma_k} x_{k+1} \quad \text{for} \quad k = N - 1, \ldots, 2, \quad x_1 = \frac{|c_1|}{\gamma_1} x_2, \]

where the coefficients \( \gamma_k \) are recursively defined by

\[ \gamma_1 = a_1, \quad \gamma_2 = a - \frac{|bc_1|}{\gamma_1}, \quad \gamma_k = a - \frac{|bc|}{\gamma_{k-1}} \quad \text{for} \quad k = 3, \ldots, N - 1, \quad \gamma_N = a_N - \frac{|b_N c|}{\gamma_{N-1}}. \tag{A.8} \]

It is easy to verify that, under assumptions (A.4) for \( k = 3, \ldots, N - 1 \), the sequence \( \{\gamma_k\} \) has two strictly positive fixed points \( z_{\pm} = (a \pm \sqrt{a^2 - 4|bc|})/2 \). Moreover, \( z_- \) is an unstable fixed point, while \( z_+ \) is stable. By condition (A.5) and relation \( \gamma_2 = a - |bc_1|/a_1 \), we have that \( \gamma_2 > z_- \). So, starting from \( \gamma_2 \), the sequence converges to \( z_+ \) and we have that, for \( \gamma^* = \min \{\gamma_2, z_+\} \),

\[ \gamma_k \geq \gamma^*, \quad k = 2, \ldots, N - 1. \tag{A.9} \]

By assumptions (A.4), it is easy to verify that \( \gamma^* > |c| \) for both cases \( \gamma^* = \gamma_2 \) and \( \gamma^* = z_+ \). Moreover, in (A.8) the inequalities \( \gamma_{N-1} \geq \gamma^* > |c| \) imply that \( \gamma_N > 0 \).

Going back to the sequence \( \{x_k\}_{i=1, \ldots, N} \), we first note that since \( \gamma_N > 0 \), it follows that \( x_N > 0 \) and, accordingly, \( x_k \geq 0 \) for all \( k = 1, \ldots, N \). Moreover, from condition (A.9) we obtain (A.7). In fact, for \( k = 2, \ldots, N - 1 \) we have

\[ x_k = \frac{|c|}{\gamma_k} x_{k+1} \leq \frac{|c|}{\gamma^*} x_{k+1} \leq \cdots \leq \left(\frac{|c|}{\gamma^*}\right)^{N-k} x_N \]

and thus

\[ x_1 = \frac{|c_1|}{\gamma_1} x_2 \leq \frac{|c_1|}{a_1} \left(\frac{|c|}{\gamma^*}\right)^{N-2} x_N. \]

To obtain the estimate (A.6), we introduce the \( N \times N \) matrix \( U \) satisfying \( U v_i = v_{i+1} - i, i = 1, \ldots, N \), so

\[ U = \{u_{ij}\}_{i,j=1, \ldots, N} \]

such that, for \( i = 1, \ldots, N \), \( u_{ij} = 0 \) if \( j \neq N - (i - 1) \), and otherwise, \( u_{N-(i-1)} = 1 \). Since \( U v_N = v_1 \) and \( U U = I \) (i.e., \( U^{-1} = U \)), to compute (A.6) we use that \( A x = v_1 \Leftrightarrow \tilde{\Lambda} x = v_N \), where \( \tilde{\Lambda} = U A U \) and \( \tilde{x} = (x_N, x_{N-1}, \ldots, x_1)^T \). So, following the same reasoning as above, we get the estimate (A.6) with \( \tilde{\gamma}_2 = a - |b_N c|/a_N \) and \( \gamma^* = \min (\tilde{\gamma}_2, z_+) > |b| \).

**Remark A.2** Assume that \( A \) has the form (A.3) with \( a > 0, b, c < 0, a + b + c = 1 \) and \( a_1 = a_N = a, c_1 = b_N = 1 - a \)—this is actually the type of matrix to which we apply Proposition A.1; see (3.8). One can easily check that both (A.4) and the boundary requirements in (A.5) hold, so Proposition A.1 can be applied. Moreover, estimates (A.6) and (A.7) can be rewritten as follows: for \( k = 2, \ldots, N - 1 \),

\[
|A^{-1} v_{N+1-k}| \leq \left(\frac{|b|}{\gamma^*}\right)^{N-k} \quad \text{with} \quad \gamma^* = \min \left( a - \frac{|c(1-a)|}{a}, \frac{a + \sqrt{a^2 - 4|bc|}}{2} \right),
\]

\[
|A^{-1} v_N| \leq \left(\frac{|c|}{\gamma^*}\right)^{N-k} \quad \text{with} \quad \gamma^* = \min \left( a - \frac{|b(1-a)|}{a}, \frac{a + \sqrt{a^2 - 4|bc|}}{2} \right).
\]
In fact, as for the second inequality, (A.7) gives $|(A^{-1}v)_k| \leq x_N(|c|/\gamma^*)^{N-k}$, where $x_N = 1/\gamma_N$ and $\gamma_N$ is defined in (A.8), together with $\gamma_1, \ldots, \gamma_{N-1}$. Since $\gamma_{N-1} > |c|$, (A.8) gives

$$\gamma_N = a - \frac{|(1 - a)c|}{\gamma_{N-1}} \geq a - |1 - a|.$$ 

But $1 - a = b + c \leq 0$, so $\gamma_N \geq 1$. This implies $x_N \leq 1$ and then $|(A^{-1}v)_k| \leq (|c|/\gamma^*)^{N-k}$. Similarly, one has $\tilde{\gamma}_N \geq 1$, and the first inequality holds as well.

**Remark A.3** Suppose now that $A$ has the form (A.3) with $a > 0$, $b, c < 0$, $a + b + c = 1$ and $a_1 = a_N = 1$, $c_1 = b_N = 0$—this is the matrix in (3.27), that is the matrix one has to deal with when the boundary conditions are of a time-independent Dirichlet type. Here, (A.4) and (A.5) both hold, so Proposition A.1 can be applied. Moreover, one immediately gets $\gamma_N = \tilde{\gamma}_N = 1$, $\tilde{\gamma}_2 = \gamma_2 = a < z_+$, so that $\min(\tilde{\gamma}_2, z_+) = \min(\gamma_2, z_+) = a$. Therefore, for $k = 2, \ldots, N - 1$,

$$|(A^{-1}v_1)_{N+1-k}| \leq \left(\frac{|b|}{a}\right)^{N-k} \text{ and } |(A^{-1}v_N)_k| \leq \left(\frac{|c|}{a}\right)^{N-k}.$$ 

**References**

**Alfonsi, A.** (2010) High order discretization schemes for the CIR process: application to affine term structure and Heston models. *Math. Comput.*, 79, 209–237.

**Amin, K. & Khanna, A.** (1994) Convergence of American option values from discrete-to continuous-time financial models. *Math. Finance*, 4, 289–304.

**Appolloni, E., Caramellino, L. & Zanette, A.** (2014) A robust tree method for pricing American options with CIR stochastic interest rate. *IMA J. Manage. Math.* (to appear), ArXiv:1305.0479. doi:10.1093/imaman/dpt030.

**Berman, A. & Plemmons, R. J.** (1994) *Nonnegative Matrices in the Mathematical Sciences*. Philadelphia, PA: Society for Industrial and Applied Mathematics (SIAM).

**Billingsley, P.** (1968) *Convergence of Probability Measures*. New York: Wiley.

**Carr, P. & Madan, D.** (1999) Option valuation using the Fast Fourier transform. *J. Comput. Finance*, 3, 463–520.

**Chiarella, C., Kang, B. & Meyer, G. H.** (2012) The evaluation of barrier option prices under stochastic volatility. *Comput. Math. Appl.*, 64, 2034–2048.

**Clarke, N. & Parrott, K.** (1999) Multigrid for American option pricing with stochastic volatility. *Appl. Math. Finance*, 6, 177–195.

**Cox, J. C., Ingersoll, J. & Ross, S.** (1985) A theory of the term structure of interest rates. *Econometrica*, 53, 385–407.

**D’Halluin, V., Forsyth, P. A. & Labahn, G.** (2005) A semi-Lagrangian approach for American Asian options under jump-diffusion. *Siam J. Sci. Comput.*, 27, 315–345.

**Ethier, S. N. & Kurtz, T.** (1986) *Markov Processes: Characterization and Convergence*. New York: John Wiley.

**Fang, F. & Oosterlee, C. W.** (2011) A Fourier-based valuation method for Bermudan and barrier options under Heston’s model. *Siam J. Fin. Math.*, 2, 439–463.

**Florescu, I. & Viens, F.** (2005) A binomial tree approach to stochastic volatility driven model of the stock price. *An. Univ. Craiova Ser. Mat. Inform.*, 32, 126–142.

**Florescu, I. & Viens, F.** (2008) Stochastic volatility: option pricing using a multinomial recombining tree. *Appl. Math. Finance J.*, 15, 151–181.

**Guan, L. K. & Xiaoqiang, G.** (2000) Pricing American options with stochastic volatility: evidence from S&P futures options. *J. Futures Markets*, 20, 625–659.

**Heston, S. L.** (1993) A closed-form solution for options with stochastic volatility with applications to bond and currency options. *Rev. Financ. Stud.*, 6, 327–343.
Hilliard, J. E. & Schwartz, A. (1996) Binomial option pricing under stochastic volatility and correlated state variables. *J. Derivatives*, 23–39.

Hull, J. & White, A. (1987) The pricing of options on assets with stochastic volatilities. *J. Finance*, 42, 281–300.

Ikonen, S. & Toivanen, J. (2009) Operator splitting methods for pricing American options under stochastic volatility. *Numer. Math.*, 113, 299–324.

in’t Hout, K. J. & Foulon, S. (2010) ADI finite difference schemes for option pricing in the Heston model with correlation. *Int. J. Numer. Anal. Mod.*, 7, 303–320.

Lamberton, D. & Pagès, G. (1990) Sur l’approximation des réduites. *Ann. Inst. H. Poincaré, Probab. Statist.*, 26, 331–355.

Leisen, D. P. J. (2000) Stock evolution under stochastic volatility: a discrete approach. *J. Derivatives*, 8, 8–27.

Longstaff, F. A. & Schwartz, E. S. (2001) Valuing American options by simulations: a simple least squares approach. *Rev. Financ. Stud.*, 14, 113–148.

Nelson, D. B. & Ramaswamy, K. (1990) Simple binomial processes as diffusion approximations in financial models. *Rev. Financ. Stud.*, 3, 393–430.

Oosterlee, C. W. (2003) On multigrid for linear complimentarity problems with application to American-style options. *Electronic Trans. Numer. Anal.*, 15, 165–185.

PREMIA. An Option Pricer. [http://www.premia.fr](http://www.premia.fr).

Stein, E. & Stein, J. (1991) Stock price distributions with stochastic volatility: an analytic approach. *Rev. Financ. Stud.*, 4, 727–752.

Strikwerda, J. C. (2004) *Finite Difference Schemes and Partial Differential Equations*. Philadelphia, PA: Society for Industrial and Applied Mathematics (SIAM).

Stroock, D. W. & Varadhan, S. R. S. (1979) *Multidimensional Diffusion Processes*. Berlin: Springer.

Thomas, L. H. (1949) Elliptic Problems in Linear Differential Equations over a Network. *Watson Scientific Computing Laboratory Report*. New York: Columbia University.

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

Zvan, R., Forsyth, P. & Vetzal, K. (1998) A penalty method for American options with stochastic volatility. *J. Comput. Appl. Math.*, 92, 199–218.