

# A robust tree method for pricing American options with CIR stochastic interest rate

ELISA APPOLLONI

MEMOTEF

Università di Roma *La Sapienza*

`elisa.appolloni@uniroma1.it`

LUCIA CARAMELLINO

Dipartimento di Matematica

Università di Roma *Tor Vergata*

`caramell@mat.uniroma2.it`

ANTONINO ZANETTE

Dipartimento di Scienze Economiche e Statistiche

Università di Udine

`antonino.zanette@uniud.it`

## Abstract

We propose a robust and stable lattice method which permits to obtain very accurate American option prices under the CIR stochastic interest rate without any numerical restriction on its parameters. Numerical results show the reliability and the accuracy of the proposed method.

*Keywords:* American options; tree methods; CIR process; stochastic interest rate;  
*2000 MSC:* 91G60, 60H30, 65C20.

*Corresponding author:* Antonino Zanette, Dipartimento di Scienze Economiche e Statistiche, Università di Udine, via Tomadini 30/A, I-33100 Udine, Italy.

## Premia 22

### 1 Introduction

The scenarios that have been prevailing on the financial markets in the last decade, as well as the financial and insurance applications to long-lived contracts such as equity-linked policies (see Brennan and Schwartz [4]), suggest that the equity models need to take into account for stochastic and positive interest rates. We consider therefore the problem of pricing American options when the equity asset follows a

log-normal type diffusion process and the interest rate risk  $r$  is described by the Cox, Ingersoll and Ross [7] (CIR hereafter) model:

$$dr(t) = \kappa(\theta - r(t))dt + \sigma_r \sqrt{r(t)}dZ_r(t), \quad r(0) = r_0 > 0.$$

In this framework, we consider tree methods for pricing American options.

The existing literature for pricing American options on lattice bivariate models is based mainly on the technique introduced by Nelson and Ramaswamy [13], who show that all 1-dimensional diffusion processes can be approximated with a computationally simple binomial process by transforming the original process into a diffusion with unit variance. In particular, Wei [17] proposes a multidimensional lattice using two transformations: the first one maps both the processes (interest rate and equity asset) into two diffusions with unit variance and the second one brings to an auxiliary process which is orthogonal to the transformed interest rate process. This allows one to build up easily a recombining bivariate lattice to guarantee the tractability of the model from a computational point of view. Hilliard, Schwartz and Tucker [9] propose different transformations, allowing one to treat with a constant variance and to eliminate the correlation between processes, and then proceed similarly to Wei. Nevertheless, numerical experiments show that the methods by Wei and by Hilliard Schwartz and Tucker are not stable and robust from a practical point of view when  $\sigma_r$  is large.

We propose here a new lattice approach which permits accurate and efficient option pricing numerical results without any restriction on the volatility parameter  $\sigma_r$ . This procedure is structurally different from the previous ones: even if the tree structure is again built by using suitable transformations, the transition probabilities are independent of these maps and are set starting from the original (and not transformed) diffusion processes. As an immediate consequence, our approximating tree strongly takes into account the structure of the covariance between the spot rate and the asset price. Comparisons for European and American options pricing problems show that the proposed method is numerically very precise, efficient, stable and robust. Moreover, the numerical efficiency of our method turns out to be particularly important when applied to the more elaborated numerical algorithm in [5], where an hybrid tree-finite difference approach is set in order to numerically compute American options in the Heston model. In fact, the tree component for the hybrid model is the one set for handling the volatility process, and turns out to be the tree which is presented here for the interest rate process.

The paper is organized as follows. In Section 2, we describe the bivariate model for the continuous-time joint evolution of the equity and of the spot interest rate processes. In Section 3 we recall the algorithms introduced by Wei [17] and by Hilliard, Schwartz and Tucker [9]. Section 4 is devoted to the description of our proposed tree algorithm, whose convergence is proved in Section 5. Numerical results and comparisons related to the evaluation tree-algorithms are the contents of Section 6.

## 2 The bivariate continuous model

We are concerned in a geometric Brownian motion describing the evolution of the equity value with drift driven by a square root process. So, we consider, under the risk-neutral probability measure, the following dynamics for the equity value

$$\frac{dS(t)}{S(t)} = r(t)dt + \sigma_S dZ_S(t), \quad S(0) = S_0 > 0, \quad (1)$$

where  $r$  is the short interest rate process,  $\sigma_S$  is the constant stock price volatility and  $Z_S$  is a standard Brownian motion. The risk-neutralized process for the short rate is described as in the CIR model, that is

$$dr(t) = \kappa(\theta - r(t))dt + \sigma_r \sqrt{r(t)} dZ_r(t), \quad r(0) = r_0 > 0, \quad (2)$$

where  $\kappa$  is a constant representing the reversion speed,  $\theta$  is the long term reversion target,  $\sigma_r > 0$  is constant and  $Z_r$  is a standard Brownian motion. When dealing with the CIR model, the Feller condition should be cited: if  $r_0 > 0$  and if  $2\kappa\theta \geq \sigma_r^2$  then a.s. the process  $r$  never hits 0.

Finally, the Brownian noises  $Z_S$  and  $Z_r$  are supposed to be correlated and we let  $\rho$  denote the correlation:

$$d\langle Z_S, Z_r \rangle(t) = \rho dt.$$

## 3 The Wei and Hilliard-Schwartz-Tucker procedures

Starting from the model in Section 2, we describe here the bivariate lattices proposed by Wei [17] and by Hilliard, Schwartz and Tucker [9], following the Nelson and Ramaswamy [13] technique. We concentrate our attention mainly to the Wei procedure. Furthermore, we briefly remind the Hilliard-Schwartz-Tucker transformations.

Actually, in Wei [17] the underlying dynamics for the short rate is given by the Vasicek model. The natural extension to the CIR process can be found in Costabile, Gaudenzi, Massabò and Zanette [6], and works as follows. First, both processes  $S$  and  $r$  are transformed into two diffusions with unit variance and then, an auxiliary process orthogonal to the transformed interest rate process is taken into account. This allows one to build up easily a recombining bivariate lattice to guarantee the tractability of the model from a computational point of view.

The first step for the construction of a recombining bivariate tree is to transform processes (1) and (2) into diffusions with unit variance. This is done by introducing the variables

$$X = (\log S)/\sigma_S \quad \text{and} \quad R = 2\sqrt{r}/\sigma_r$$

respectively. The dynamics of  $X$  and  $R$  may be easily derived applying Ito's Lemma. Hence,

$$\begin{aligned} dX(t) &= \mu_X(R(t))dt + dZ_S(t), & X(0) &= (\log S_0)/\sigma_S, \\ dR(t) &= \mu_R(R(t))dt + dZ_r(t), & R(0) &= 2\sqrt{r_0}/\sigma_r \end{aligned} \quad (3)$$

where

$$\mu_X(R) = \frac{\sigma_r^2 R^2/4 - \sigma_S^2/2}{\sigma_S} \quad \text{and} \quad \mu_R(R) = \frac{\kappa(4\theta - R^2\sigma_r^2) - \sigma_r^2}{2R\sigma_r^2}. \quad (4)$$

Both the transformed processes  $X$  and  $R$  have unit variance, hence they may be discretized independently using recombining binomial trees. Consider now the following process:

$$Y(t) = \frac{X(t) - \rho R(t)}{\sqrt{1 - \rho^2}}.$$

Standard calculations give that the processes  $Y(t)$  and  $R(t)$  have null covariance. In fact, the dynamics of the diffusion process  $Y(t)$  has the form

$$dY(t) = \mu_Y(R(t))dt + dZ_Y(t), \quad Y(0) = Y_0 = \frac{1}{\sqrt{1 - \rho^2}}[(\log S_0)/\sigma_S - 2\rho\sqrt{r_0}/\sigma_r], \quad (5)$$

where

$$\mu_Y(R) = \frac{\mu_X(R) - \rho\mu_R(R)}{\sqrt{1 - \rho^2}}. \quad (6)$$

and  $Z_Y$  is the standard Brownian motion given by

$$Z_Y = \frac{Z_S - \rho Z_r}{\sqrt{1 - \rho^2}},$$

so that it is orthogonal to  $Z_r$ . The underlying asset  $S$  and the spot interest rate  $r$  can be obtained from the state variables  $Y$  and  $R$  using the relations:

$$r(t) = \begin{cases} \frac{R^2(t)\sigma_r^2}{4} & \text{if } R(t) > 0 \\ 0 & \text{otherwise} \end{cases} \quad S(t) = \exp[\sigma_S(\sqrt{1 - \rho^2}Y(t) + \rho R(t))]. \quad (7)$$

In order to construct the discrete approximations of the processes  $R$  and  $Y$ , one divides the time to maturity  $T$  into  $N$  intervals of length  $h = T/N$  and, since  $R$  and  $Y$  are both processes with unit variance, one sets the size of each up-step and each down-step equal to  $\sqrt{h}$  and  $-\sqrt{h}$  respectively. As usual, a binomial tree is considered to describe the evolution of the discrete approximating processes. We label  $(0,0)$  the starting node where the  $R$ -process has value  $R(0)$ . After  $i$  time steps ( $i = 0, \dots, N$ ),  $R$  may be located at one of the nodes  $(i, k)$  ( $k = 0, \dots, i$ ) corresponding to the values

$$R_{i,k} = R_0 + (2k - i)\sqrt{h}. \quad (8)$$

Analogously, for the discrete process approximating  $Y$ , after  $i$  time steps  $Y$  may be located at one of the nodes  $(i, j)$  ( $j = 0, \dots, i$ ) corresponding to the values

$$Y_{i,j} = Y_0 + (2j - i)\sqrt{h}. \quad (9)$$

Transition probabilities have to be specified to ensure the matching of the local drift and of the local variance between the discrete and continuous model of  $(Y, R)$ . This will guarantee that the 2-dimensional discretized process weakly converges to the corresponding 2-dimensional continuous-time process given by (3) and (5) (see Nelson and Ramaswamy [13] for a detailed description in the one-dimensional setting). To do this one has to take into account that in some regions of the tree it may happen that multiple jumps are needed to satisfy properly the matching conditions. Hence, starting from  $R_{i,k}$  at time  $ih$ , the process  $R$  may jump at time  $(i+1)h$  to the value  $R_{i+1,k_d}$  or  $R_{i+1,k_u}$ , with  $k_d$  and  $k_u$  defined as

$$k_d = \begin{cases} 0 & \text{if } R_{i,k} + (\mu_R)_{i,k}h < R_{i+1,0} \\ i & \text{if } R_{i,k} + (\mu_R)_{i,k}h > R_{i+1,i+1} \\ \text{the largest index } k^* \in [0, i] \text{ s.t.} & \\ R_{i,k} + (\mu_R)_{i,k}h \geq R_{i+1,k^*} & \text{otherwise} \end{cases}$$

and

$$k_u = k_d + 1,$$

in which  $(\mu_R)_{i,k} = \mu_R(R_{i,k})$ . It is clear that in the case  $R_{i,k} = 0$  the drift  $\mu_R$  cannot be evaluated. So, in practice one chooses the number  $N$  of the monitoring instants in such a way that the lattice for  $R$  never hits 0. We briefly recall here the construction of the discrete approximation for  $R$  as described in Remark 1 in [6]: setting

$$k_0 = \text{int}\left(\frac{R_0}{\sqrt{h}}\right) \quad \text{and} \quad \gamma = \min\left(|R_0 - k_0\sqrt{h}|, |R_0 - (k_0 + 1)\sqrt{h}|\right)$$

(that is,  $\gamma$  is the minimum, in absolute value, of the lattice points for  $R$ ), the number of steps  $N$  giving  $h = T/N$  is chosen such that  $\gamma$  is not too small ( $\gamma \geq 10^{-6}$ ). Hereafter,  $\text{int}(x)$  denotes the integer part of  $x \in \mathbb{R}$ , that is, for  $x \geq 0$  then  $\text{int}(x)$  is the largest integer not exceeding  $x$  and for  $x < 0$  we set  $\text{int}(x) = -\text{int}(-x)$ .

One easily gets

$$k_d = k + \text{int}\left(\frac{(\mu_R)_{i,k}\sqrt{h} + 1}{2}\right).$$

Now, the probability that the  $R$ -process, located at the node  $R_{i,k}$ , reaches  $R_{i+1,k_u}$ , is well defined by setting

$$p_{i,k} = 0 \vee \frac{(\mu_R)_{i,k}h + R_{i,k} - R_{i+1,k_d}}{R_{i+1,k_u} - R_{i+1,k_d}} \wedge 1.$$

Obviously, the probability to reach  $R_{i+1,k_d}$  is  $1 - p_{i,k}$ .

Concerning the  $Y$ -process, one has to take into account also the behavior of  $R$  because here  $\mu_Y = \mu_Y(R)$ , see (6). So, let us assume that the pair is located at  $(Y_{i,j}, R_{i,k})$  at the time-step  $i$ . Then, for the jumps in the  $Y$ -direction, one sets

$$j_d = \begin{cases} 0 & \text{if } Y_{i,j} + (\mu_Y)_{i,k}h < Y_{i+1,0} \\ i & \text{if } Y_{i,j} + (\mu_Y)_{i,k}h > Y_{i+1,i+1} \\ \text{the largest index } j^* \in [0, i] \text{ s.t.} & \\ Y_{i,j} + (\mu_Y)_{i,k}h \geq Y_{i+1,j^*} & \text{otherwise} \end{cases}$$

and  $j_u = j_d + 1$ , in which  $(\mu_Y)_{i,k} = \mu_Y(R_{i,k})$ . Again, one has

$$j_d = j + \text{int}\left(\frac{(\mu_Y)_{i,k}\sqrt{h} + 1}{2}\right).$$

Furthermore, the transition probability that the  $Y$ -process jumps to  $Y_{i+1,j_u}$  is defined by setting

$$\hat{p}_{i,j,k} = 0 \vee \frac{(\mu_Y)_{i,k}h + Y_{i,j} - Y_{i+1,j_d}}{Y_{i+1,j_u} - Y_{i+1,j_d}} \wedge 1$$

and  $1 - \hat{p}_{i,j,k}$  is the probability of the down-jump.

We are now ready to describe the discrete approximation scheme for the joint evolution of the processes  $Y$  and  $R$  by considering a bivariate tree. To this purpose, one takes a recombining structure for the tree by merging the two univariate binomial trees for the state variables  $R$  and  $Y$ . At each time step  $i$  ( $i = 0, \dots, N$ ), the tree have  $(i+1)^2$  nodes that we label  $(i, j, k)$  corresponding to the values  $R_{i,k}$  and  $Y_{i,j}$  ( $k, j = 0, \dots, i$ ).

Starting from the node  $(i, j, k)$ , in consideration of possible multiple jumps and taking into account the tree structure, the process may reach one of the following four nodes:

$$\begin{aligned} (i+1, j_u, k_u), & \quad \text{with probability } q_{i,j_u,k_u}, \\ (i+1, j_u, k_d), & \quad \text{with probability } q_{i,j_u,k_d}, \\ (i+1, j_d, k_u), & \quad \text{with probability } q_{i,j_d,k_u}, \\ (i+1, j_d, k_d), & \quad \text{with probability } q_{i,j_d,k_d}, \end{aligned}$$

where  $j_u, j_d, k_u, k_d$  are related to multiple jumps on the tree in the  $Y$  and  $R$  direction respectively, and  $q_{i,j_u,k_u}, q_{i,j_u,k_d}, q_{i,j_d,k_u}, q_{i,j_d,k_d}$  are the associated transition probabilities. Such probabilities can be computed, due to the orthogonality of the noises driving the two processes, as follows

$$\begin{aligned} q_{i,j_u,k_u} &= p_{i,k}\hat{p}_{i,j,k}, & q_{i,j_u,k_d} &= (1 - p_{i,k})\hat{p}_{i,j,k} \\ q_{i,j_d,k_u} &= p_{i,k}(1 - \hat{p}_{i,j,k}), & q_{i,j_d,k_d} &= (1 - p_{i,k})(1 - \hat{p}_{i,j,k}). \end{aligned} \tag{10}$$

The tree for the joint evolution of the processes  $r$  and  $S$  is derived simply by applying the maps in (7) to the discrete scheme just defined. Therefore, each node  $(i, j, k)$  of the bivariate tree for  $(S, r)$  corresponds to the values

$$S_{i,j,k} = \exp[\sigma_S(\sqrt{1 - \rho^2}Y_{i,j} + \rho R_{i,k})]. \tag{11}$$

and

$$r_{i,k} = \begin{cases} \frac{R_{i,k}^2 \sigma_r^2}{4} & \text{if } R_{i,k} > 0 \\ 0 & \text{otherwise.} \end{cases} \tag{12}$$

The resulting successor values are easily identified with  $r_{i+1,k_u}, r_{i+1,k_d}$  and  $S_{i+1,j_u,k_u}, S_{i+1,j_u,k_d}, S_{i+1,j_d,k_u}, S_{i+1,j_d,k_d}$ , and the transition probabilities are kept those defined in (10).

Once the approximating Markov chain structure is built, the price at time 0 of an American Put option with maturity  $T$  and strike  $K$  is computed by the following backward dynamic programming equations:

$$\begin{cases} v_{N,j,k} = (K - S_{N,j,k})_+ \\ v_{i,j,k} = \max \left( (K - S_{i,j,k})_+, e^{-r_{i,k}h} \left[ q_{i,j_u,k_u} v_{i+1,j_u,k_u} + q_{i,j_u,k_d} v_{i+1,j_u,k_d} + \right. \right. \\ \left. \left. + q_{i,j_d,k_u} v_{i+1,j_d,k_u} + q_{i,j_d,k_d} v_{i+1,j_d,k_d} \right] \right), \end{cases}$$

where  $v_{i,j,k}$ ,  $i = 0, \dots, n$  and  $j, k = 0, \dots, i$ , provides the American option price at every node  $(i, j, k)$  of the tree structure.

Hilliard, Schwartz and Tucker [9] propose a similar procedure, which can be briefly recalled as follows. The first transformation (ensuring constant variance) is similar to the one proposed by Wei:

$$X = (\log S)/\sigma_S, \quad R = 2\sqrt{r}.$$

The second transformation (removing the correlation between processes) is given by

$$X_1 = \sigma_r X + R, \quad X_2 = \sigma_r X - R.$$

Then a recombining bivariate tree is constructed by means of the discrete approximations for  $X_1$  and  $X_2$ . The multiple jumps procedure and the transition probabilities are set by using a procedure which is analogous to the one just described.

Notice that in the procedures by Wei and by Hilliard, Schwartz and Tucker, the structure of the transition probabilities is of a product type, since it comes from the non-correlation between the processes they construct ( $R$  and  $Y$  for Wei,  $X_1$  and  $X_2$  for Hilliard, Schwartz and Tucker). Moreover, as already discussed in Tian [16] and stressed in [6], the convergence in law of the tree for  $R$  holds under the condition  $4\kappa\theta \geq \sigma_r^2$ , and this is a limit for these procedures. These points, that is the product-structure for the transition probabilities and the above requirement for the parameters of the CIR process, make the main difference with the method we propose in this paper (see also next Remark 1).

## 4 The robust tree algorithm

In this section we introduce a new procedure for the construction of a lattice for the description of the approximating evolution of the pair  $(S, r)$  as in (1)-(2). Here, we often use the drift coefficients appearing in (1) and (2), so we recall that they are given by

$$\mu_S(S) = rS \quad \text{and} \quad \mu_r(r) = \kappa(\theta - r)$$

respectively.

Consider first the interest rate process  $r$ . Following a remark of Tian (see [16] at page 100, or also [15]), the right implementation of the Nelson-Ramaswamy algorithm for the CIR process consists in setting the probabilistic structure directly on  $r$ , and not through  $R$ . Thus, we first construct a tree for  $r$  by transforming the computationally simple lattice for  $R$ , by means of the map in (12) : at time step  $i = 0, 1, \dots, N$ , for  $k = 0, \dots, i$  we set

$$r_{ik} = \begin{cases} \frac{R_{i,k}^2 \sigma_r^2}{4} & \text{if } R_{i,k} > 0 \\ 0 & \text{otherwise,} \end{cases}$$

$R_{i,k}$  being defined in (8). Secondly, the transition probabilities are specified directly on  $r$ : the matching of the local drift between the discrete and the continuous model is set by using the original interest rate CIR process  $r$ . According to Nelson and Ramaswamy [13], at the node  $(i, k)$ , we set  $(\mu_r)_{i,k} = \mu_r(r_{i,k})$  and we define

$$k_d(i, k) = \max\{k^* : 0 \leq k^* \leq k \text{ and } r_{i,k} + (\mu_r)_{i,k}h \geq r_{i+1,k^*}\}, \quad (13)$$

$$k_u(i, k) = \min\{k^* : k+1 \leq k^* \leq i+1 \text{ and } r_{i,k} + (\mu_r)_{i,k}h \leq r_{i+1,k^*}\} \quad (14)$$

with the understanding  $k_d(i, k) = 0$  if  $\{k^* : 0 \leq k^* \leq k \text{ and } r_{i,k} + (\mu_r)_{i,k}h \geq r_{i+1,k^*}\} = \emptyset$  and  $k_u(i, k) = i+1$  if  $\{k^* : k+1 \leq k^* \leq i+1 \text{ and } r_{i,k} + (\mu_r)_{i,k}h \leq r_{i+1,k^*}\} = \emptyset$ . The transition probabilities are now defined as follows: starting from  $(i, k)$ , the probability that the process jumps to  $(i+1, k_u(i, k))$  is set as

$$p_{i,k} = 0 \vee \frac{(\mu_r)_{i,k}h + r_{i,k} - r_{i+1,k_d(i,k)}}{r_{i+1,k_u(i,k)} - r_{i+1,k_d(i,k)}} \wedge 1. \quad (15)$$

And of course, the jump to  $(i+1, k_d(i, k))$  happens with probability  $1 - p_{i,k}$ . Concerning the lattice for  $S$ , we proceed as follows. We first consider a computationally simple tree-structure  $(U_{i,j})_{i,j}$  defined as

$$U_{i,j} = U_0 + (2j - i)\sqrt{h}, \quad U_0 = \frac{1}{\sigma_S} \log S_0, \quad i = 0, \dots, N, \quad j = 0, \dots, i. \quad (16)$$

Then, we apply the (new) transformation

$$S = e^{\sigma_S U} \quad (17)$$

and we get a tree  $(S_{i,j})_{i,j}$  for  $S$ :  $S_{i,j} = e^{\sigma_S U_{i,j}}$ . By combining the two lattices, we obtain a 2-dimensional tree

$$(S_{i,j}, r_{i,k}), \quad i = 0, \dots, N, \quad j, k = 0, \dots, i.$$

It is worth to say that the transformation (17) does not seem to be natural in order to describe the evolution of the pair. Nevertheless, this is not important. In fact, at this stage, we only need to set up the state-space of the Markov chain that we want to approximate the continuous time process  $(S, r)$ , and in fact as  $h \rightarrow 0$  one really gets that the discrete state-space converges to  $\mathbb{R}_+^2$ . What is important now is to define the transition probabilities in order to link the tree with the diffusion pair  $(S, r)$ .



We first set the probabilistic behavior of the jumps for  $S$  at the time-step  $i + 1$  given the position  $(S_{i,j}, r_{i,k})$  by matching w.r.t. the drift  $\mu_S$  of  $S$ . So, we set  $(\mu_S)_{i,j,k} = \mu_S(S_{i,j}, r_{i,k})$  and we define

$$j_d(i, j, k) = \max\{j^* : 0 \leq j^* \leq j \text{ and } S_{i,j} + (\mu_S)_{i,j,k}h \geq S_{i+1,j^*}\} \quad (18)$$

$$j_u(i, j, k) = \min\{j^* : j + 1 \leq j^* \leq i + 1 \text{ and } S_{i,j} + (\mu_S)_{i,j,k}h \leq S_{i+1,j^*}\} \quad (19)$$

with the usual understanding  $j_d(i, j, k) = 0$  if  $\{j^* : 0 \leq j^* \leq j \text{ and } S_{i,j} + (\mu_S)_{i,j,k}h \geq S_{i+1,j^*}\} = \emptyset$  and  $j_u(i, j, k) = i + 1$  if  $\{j^* : j + 1 \leq j^* \leq i + 1 \text{ and } S_{i,j} + (\mu_S)_{i,j,k}h \leq S_{i+1,j^*}\} = \emptyset$ . Now, starting from  $(i, j, k)$ , the probability of an up-jump of the tree for  $S$  is set as

$$\hat{p}_{i,j,k} = 0 \vee \frac{(\mu_S)_{i,j,k}h + S_{i,j} - S_{i+1,j_d(i,j,k)}}{S_{i+1,j_u(i,j,k)} - S_{i+1,j_d(i,j,k)}} \wedge 1. \quad (20)$$

The down-jump obviously may happen with probability  $1 - \hat{p}_{i,j,k}$ .

We must finally set the covariance structure for the joint evolution of the processes  $S$  and  $r$  on the bivariate tree. Starting from the node  $(i, j, k)$ , by considering possible multiple jumps and by taking into account the tree structure, the process may reach one of the four nodes:  $(i + 1, j_u, k_u)$  with probability  $q_{i,j_u,k_u}$ ,  $(i + 1, j_u, k_d)$  with probability  $q_{i,j_u,k_d}$ ,  $(i + 1, j_d, k_u)$  with probability  $q_{i,j_d,k_u}$ ,  $(i + 1, j_d, k_d)$  with probability  $q_{i,j_d,k_d}$ . The transition probabilities  $q_{i,j_u,k_u}, q_{i,j_u,k_d}, q_{i,j_d,k_u}, q_{i,j_d,k_d}$  are computed by matching (at the first order in  $h$ ) the conditional mean and the conditional covariance between the continuous and the discrete processes of  $S$  and  $r$ . The matching conditions lead to solving the following system:

$$\begin{cases} q_{i,j_u,k_u} + q_{i,j_u,k_d} = \hat{p}_{i,j,k} \\ q_{i,j_u,k_u} + q_{i,j_d,k_u} = p_{i,k} \\ q_{i,j_u,k_u} + q_{i,j_d,k_u} + q_{i,j_u,k_d} + q_{i,j_d,k_d} = 1 \\ m_{i,j_u,k_u}q_{i,j_u,k_u} + m_{i,j_u,k_d}q_{i,j_u,k_d} + m_{i,j_d,k_u}q_{i,j_d,k_u} + m_{i,j_d,k_d}q_{i,j_d,k_d} = \rho\sigma_r\sqrt{r_{i,k}}\sigma_S S_{i,j}h \end{cases} \quad (21)$$

where  $\hat{p}_{i,j,k}$  and  $p_{i,k}$  are given in (15) and (20) respectively and

$$\begin{aligned} m_{i,j_u,k_u} &= (S_{i+1,j_u} - S_{i,j})(r_{i+1,k_u} - r_{i,k}), & m_{i,j_u,k_d} &= (S_{i+1,j_u} - S_{i,j})(r_{i+1,k_d} - r_{i,k}), \\ m_{i,j_d,k_u} &= (S_{i+1,j_d} - S_{i,j})(r_{i+1,k_u} - r_{i,k}), & m_{i,j_d,k_d} &= (S_{i+1,j_d} - S_{i,j})(r_{i+1,k_d} - r_{i,k}). \end{aligned} \quad (22)$$

This completely concludes the description of the approximating process for the pair  $(S, r)$ , and we study in Section 5 the convergence to the continuous time process. Now, the backward induction procedure (3) can be applied, by using the probabilities solving the system (21). In Section 6 we give numerical results from the use of this procedure and we compare them with the ones provided by the methods of Wei and of Hilliard, Schwartz and Tucker.

**Remark 1.** *Our procedure uses directly both the evolution of the process  $r$  and the correlation between  $S$  and  $r$ . Therefore, the transformed process  $R$  is exploited only to construct the state-space of the Markov chain approximating  $r$ , and not to build up the transition probabilities. This is a structural difference with what done in Wei [17], in Costabile et al. [6] and in Hilliard, Schwartz and Tucker [9].*

Indeed, their methods are strongly based on the use of a bivariate diffusion process whose components are driven by uncorrelated noises, leading to the definition of the transition probabilities by means of a product. And this can be done only by handling the process  $R$ .

## 5 Convergence results

We consider here the theoretical problem of the convergence of our tree approximation. Specifically, in Section 5.1 we first study the convergence in law of our bivariate tree model to the diffusion model given by (1) and (2), whereas in Section 5.2 we discuss the convergence of the binomial American price to the true one.

### 5.1 The convergence in law

Let us briefly recall the lattice structure for the process  $r$ : for a fixed node  $(i, k)$ ,  $i = 0, 1, \dots, N$ ,  $k = 0, 1, \dots, i + 1$ , one has

$$r_{i,k} = \frac{R_{i,k}^2 \sigma_r^2}{4} \mathbf{1}_{R_{i,k} > 0} \quad \text{and} \quad R_{i,k} = R_0 + (2k - i)\sqrt{h}, \quad (23)$$

with  $R_0 = \frac{2}{\sigma_r} \sqrt{r_0}$ . We recall that  $k_d(i, k) = \max K_d(i, k)$  and  $k_u(i, k) = \min K_u(i, k)$ , where

$$K_d(i, k) = \{k^* : 0 \leq k^* \leq k \text{ and } r_{i,k} + (\mu_r)_{i,k} h \geq r_{i+1,k^*}\}, \quad (24)$$

$$K_u(i, k) = \{k^* : k + 1 \leq k^* \leq i + 1 \text{ and } r_{i,k} + (\mu_r)_{i,k} h \leq r_{i+1,k^*}\} \quad (25)$$

with the understanding  $k_d(i, k) = 0$  if  $K_d(i, k) = \emptyset$  and  $k_u(i, k) = i + 1$  if  $K_u(i, k) = \emptyset$ . The probability that the process jumps to  $(i + 1, k_u(i, k))$  is set as

$$p_{i,k} = 0 \vee \frac{(\mu_r)_{i,k} h + r_{i,k} - r_{i+1,k_d(i,k)}}{r_{i+1,k_u(i,k)} - r_{i+1,k_d(i,k)}} \wedge 1 \quad (26)$$

and of course, the jump to  $(i + 1, k_d(i, k))$  happens with probability  $1 - p_{i,k}$ . The behavior of the up and down jumps is given in the following

**Lemma 2.** *Let  $\theta_* < \theta$  and  $\theta^* > \theta$  be such that*

$$0 < \theta_* < \frac{(\theta \wedge r_0)}{2} \quad \text{and} \quad \theta^* > 2(\theta \vee r_0).$$

*Then there exists a positive constant  $h_1 = h_1(\theta_*, \theta^*, \kappa, \theta, \sigma_r) < 1$  such that for every  $h < h_1$  the following statements hold.*

- (i) *If  $0 \leq r_{i,k} < \theta_* \sqrt{h}$  then  $k_d(i, k) = k$  and  $k_u(i, k) \in \{k + 1, \dots, i + 1\}$ . Moreover, there exists a positive constant  $C_* > 0$  such that*

$$|r_{i+1,k_d(i,k)} - r_{i,k}| \leq C_* h^{3/4} \quad \text{and} \quad |r_{i+1,k_u(i,k)} - r_{i,k}| \leq C_* h^{3/4}. \quad (27)$$

(ii) If  $\theta_*\sqrt{h} \leq r_{i,k} \leq \theta^*/\sqrt{h}$  then  $k_d(i, k) = k$  and  $k_u(i, k) = k + 1$ . Moreover, one has

$$r_{i+1,k_d(i,k)} - r_{i,k} = -\sigma_r \sqrt{r_{i,k}h} + \frac{\sigma_r^2}{4} h \quad \text{and} \quad r_{i+1,k_u(i,k)} - r_{i,k} = \sigma_r \sqrt{r_{i,k}h} + \frac{\sigma_r^2}{4} h. \quad (28)$$

**Proof.** We start by considering  $h_1 \in (0, 1)$  and in what follows, we will “calibrate” the value of  $h_1$ .

(i) Let us first notice that since  $r_{i,k} \leq \theta$  one has  $(\mu_r)_{i,k} \geq 0$  and then for every  $k^* \leq k$  one has  $r_{i+1,k^*} \leq r_{i,k} \leq r_{i,k} + (\mu_r)_{i,k}h$ , which gives  $k_d(i, k) = k$ . Let us prove (27) for  $k_d(i, k)$ . If  $r_{i,k} = 0$  then  $r_{i+1,k} = 0$  as well, and (27) trivially holds. If instead  $0 < r_{i,k} < \theta_*\sqrt{h}$ , then one can have both  $r_{i+1,k} = 0$  and  $r_{i+1,k} > 0$ . In the first case, it must be

$$0 \geq R_{i+1,k} = R_{i,k} - \sqrt{h},$$

so that one actually has  $0 < R_{i,k} < \sqrt{h}$  and then  $r_{i,k} \leq \sigma_r^2 h/4$ . Therefore,  $|r_{i+1,k_d} - r_{i,k}| = r_{i,k} \leq \sigma_r^2 h/4$ , and (27) holds. Consider now the second case, that is  $0 < r_{i,k} < \theta_*\sqrt{h}$  and  $r_{i+1,k} > 0$ . Then it must be

$$r_{i+1,k} = \frac{\sigma_r^2}{4} R_{i+1,k}^2 = \frac{\sigma_r^2}{4} (R_{i,k} - \sqrt{h})^2 = r_{i,k} - \sigma_r \sqrt{r_{i,k}h} + \frac{\sigma_r^2}{4} h$$

so that

$$|r_{i+1,k_d(i,k)} - r_{i,k}| = \left| -\sigma_r \sqrt{r_{i,k}h} + \frac{\sigma_r^2}{4} h \right| \leq \sigma_r \sqrt{r_{i,k}h} + \frac{\sigma_r^2}{4} h \leq \left( \sigma_r \sqrt{\theta_*} + \frac{\sigma_r^2}{4} \right) h^{3/4}$$

and (27) again holds.

Let us now discuss the up-jump. We notice that

$$r_{i+1,i+1} - r_{i,k} - (\mu_r)_{i,k}h \geq r_0 - \theta_*\sqrt{h} - \kappa\theta h \geq r_0 - \theta_* - \kappa\theta h.$$

So, by taking  $h_1 < (r_0 - \theta_*)/(\kappa\theta)$  we get  $K_u(i, k) \neq \emptyset$ , and we can proceed by looking for the smallest integer  $k^* \geq k + 1$  such that  $k^* \leq i + 1$  and the following statements hold:

$$r_{i+1,k^*} \geq r_{i,k} + (\mu_r)_{i,k}h \quad \text{and} \quad r_{i+1,k^*-1} < r_{i,k} + (\mu_r)_{i,k}h. \quad (29)$$

Notice that in particular the first condition gives  $r_{i+1,k^*} > 0$ . Assume first the case  $r_{i+1,k^*-1} = 0$ , that is  $R_{i+1,k^*} > 0$  and  $R_{i+1,k^*-1} \leq 0$ . Then,

$$0 < R_{i+1,k^*} = R_{i+1,k^*-1} + 2\sqrt{h} \leq 2\sqrt{h},$$

so that  $r_{i+1,k^*} \leq \sigma_r^2 h$ . Since  $r_{i+1,k^*} \geq r_{i,k}$ , one gets  $|r_{i+1,k^*} - r_{i,k}| \leq r_{i+1,k^*} \leq \sigma_r^2 h$ , and this proves (27). So, it remains to study the case  $r_{i+1,k^*-1} > 0$ . Here, we have

$$\begin{aligned} r_{i+1,k^*} &= \frac{\sigma_r^2}{4} R_{i+1,k^*}^2 = \frac{\sigma_r^2}{4} (R_{i+1,k^*-1} + 2\sqrt{h})^2 \\ &= r_{i+1,k^*-1} + 2\sigma_r \sqrt{r_{i+1,k^*-1}h} + \sigma_r^2 h. \end{aligned}$$

Now, by using (29) we get

$$\begin{aligned} 0 \leq r_{i+1,k^*} - r_{i,k} &= r_{i+1,k^*-1} - r_{i,k} + 2\sigma_r \sqrt{r_{i+1,k^*-1}h} + \sigma_r^2 h \\ &\leq (\mu_r)_{i,k} h + 2\sigma_r \sqrt{(r_{i,k} + (\mu_r)_{i,k} h)h} + \sigma_r^2 h \end{aligned}$$

and by recalling that  $r_{i,k} \leq \theta_* \sqrt{h}$ , we get (27) also in this last case.

(ii) Here, we split our reasonings in two different cases: (ii.a)  $\theta_* \sqrt{h} \leq r_{i,k} \leq \theta$  and (ii.b)  $\theta \leq r_{i,k} \leq \theta^*/\sqrt{h}$ .

Assume (ii.a). As for the down-jump, here we have  $(\mu_r)_{i,k} \geq 0$ . Since for every  $k^* \leq k$  one has  $r_{i+1,k^*} \leq r_{i,k}$ , we can immediately conclude that  $k_d(i,k) = k$ . Concerning the up-jump, we first notice that for  $k^* \geq k+1$  one has  $R_{i+1,k^*} = R_{i,k} + (2(k^* - k) - 1)\sqrt{h} > 0$ , so that using the square relation with  $R$  we can write

$$r_{i+1,k^*} - r_{i,k} = (2(k^* - k) - 1)\sigma_r \sqrt{r_{i,k}h} + \frac{(2(k^* - k) - 1)^2}{4} \sigma_r^2 h.$$

Therefore, we must look for the smallest integer  $k^* \geq k+1$  such that the above r.h.s. is larger or equal to  $(\mu_r)_{i,k} h$ , and this reduces to require that

$$2(k^* - k) - 1 \geq 2 \frac{\sqrt{r_{i,k} + (\mu_r)_{i,k} h} - \sqrt{r_{i,k}}}{\sigma_r \sqrt{h}}.$$

So, we prove that we can choose  $h_1$  such that for every  $h < h_1$  the above inequality holds for  $k^* = k+1$ , that is

$$1 \geq 2 \frac{\sqrt{r_{i,k} + (\mu_r)_{i,k} h} - \sqrt{r_{i,k}}}{\sigma_r \sqrt{h}}.$$

In fact, by recalling that  $0 \leq (\mu_r)_{i,k} = \kappa(\theta - r_{i,k}) \leq \kappa\theta$  and  $r_{i,k} \geq \theta_* \sqrt{h}$ , we can write

$$\begin{aligned} 2 \frac{\sqrt{r_{i,k} + (\mu_r)_{i,k} h} - \sqrt{r_{i,k}}}{\sigma_r \sqrt{h}} &= \frac{2(\mu_r)_{i,k} \sqrt{h}}{\sigma_r (\sqrt{r_{i,k} + (\mu_r)_{i,k} h} + \sqrt{r_{i,k}})} \\ &\leq \frac{\kappa\theta \sqrt{h}}{\sigma_r \sqrt{r_{i,k}}} \leq \frac{\kappa\theta}{\sigma_r \sqrt{\theta_*}} h^{1/4} \end{aligned}$$

and the statement actually holds for  $h < h_1 < (\frac{\sigma_r \sqrt{\theta_*}}{\kappa\theta})^4$ .

(ii.b) This case can be treated similarly to the previous one, so we omit the proof.  $\square$

**Remark 3.** As a consequence of the proof of Lemma 2, for  $h < h_1$  we get that  $r_{i+1,k_u(i,k)} - r_{i+1,k_d(i,k)} > 0$ ,  $(\mu_r)_{i,k} h + r_{i,k} - r_{i+1,k_d(i,k)} \geq 0$  and  $r_{i+1,k_u(i,k)} - r_{i,k} -$

$(\mu_r)_{i,k}h \geq 0$ . Therefore, we can actually drop the  $0 \vee$  and  $\wedge 1$  appearing in (26) and for  $h < h_1$  we can directly write

$$p_{i,k} = \frac{(\mu_r)_{i,k}h + r_{i,k} - r_{i+1,k_d(i,k)}}{r_{i+1,k_u(i,k)} - r_{i+1,k_d(i,k)}}. \quad (30)$$

We also notice that whenever  $r_{i,k} = 0$  one has  $r_{i+1,k_d(i,k)} = 0$  as well and we have seen that  $(\mu_r)_{i,k}h \leq r_{i+1,k_u} = r_{i+1,k_u-1} + 2\sigma_r\sqrt{r_{i+1,k_u-1}h} + \sigma_r^2h$ . So, since  $r_{i+1,k_u-1} < r_{i,k} + (\mu_r)_{i,k}h$ , we get

$$\kappa\theta h \leq r_{i+1,k_u} \leq (\sqrt{\kappa\theta} + \sigma_r)^2h$$

and by substituting in (26) we obtain

$$p_{i,k} = \frac{\kappa\theta h}{r_{i+1,k_u(i,k)}} \in \left[ \frac{1}{(1 + \sigma_r/\sqrt{\kappa\theta})^2}, 1 \right]. \quad (31)$$

So, we state a lower bound for the up-jump probability, which is close to 1 when the volatility parameter  $\sigma_r$  is close to 0.

The behavior of the up and down jumps for the process  $S$  are much easier to study. So, we recall that for a fixed node  $(i, j)$ ,  $i = 0, 1, \dots, N$ ,  $j = 0, 1, \dots, i+1$ , the lattice structure on  $S$  is given by

$$S_{i,j} = e^{\sigma_S U_{i,j}} \quad \text{and} \quad U_{i,j} = U_0 + (2j - i)\sqrt{h}, \quad (32)$$

with  $U_0 = \frac{1}{\sigma_S} \log S_0$ . Here, the up and down jumps depend also on the position of the process  $r$ : for a fixed  $k = 1, \dots, i+1$ ,  $j_d(i, j, k) = \max J_d(i, j, k)$  and  $j_u(i, j, k) = \min J_u(i, j, k)$ , where

$$J_d(i, j, k) = \{j^* : 0 \leq j^* \leq j \text{ and } S_{i,j} + (\mu_S)_{i,j,k}h \geq S_{i+1,j^*}\}, \quad (33)$$

$$J_u(i, j, k) = \{j^* : j+1 \leq j^* \leq i+1 \text{ and } S_{i,j} + (\mu_S)_{i,j,k}h \leq S_{i+1,j^*}\} \quad (34)$$

again with the understanding  $j_d(i, j, k) = 0$  if  $J_d(i, j, k) = \emptyset$  and  $j_u(i, j, k) = i+1$  if  $J_u(i, j, k) = \emptyset$ . The probability that the process jumps to  $(i+1, j_u(i, j, k))$  is set as

$$\hat{p}_{i,j,k} = 0 \vee \frac{(\mu_S)_{i,j,k}h + S_{i,j} - S_{i+1,j_d(i,j,k)}}{S_{i+1,j_u(i,j,k)} - S_{i+1,j_d(i,j,k)}} \wedge 1 \quad (35)$$

and of course, the jump to  $(i+1, j_d(i, j, k))$  happens with probability  $1 - \hat{p}_{i,j,k}$ . The behavior of the up and down jumps is given in the following

**Lemma 4.** *Let  $r_* > 0$  be fixed. Then there exists  $h_2 = h_2(r_*, \sigma_S) < 1$  such that for every  $h < h_2$  and  $(i, k)$  such that  $r_{i,k} \in [0, r_*]$  one has*

$$j_d(i, j, k) = j \quad \text{and} \quad j_u(i, j, k) = j+1.$$

As a consequence, for  $h < h_2$  and for every  $(i, k)$  such that  $r_{i,k} \in [0, r_*]$  one has

$$S_{i+1,j_u(i,j,k)} - S_{i,j} = S_{i,j}(e^{\sigma_S\sqrt{h}} - 1) \quad \text{and} \quad S_{i+1,j_d(i,j,k)} - S_{i,j} = S_{i,j}(e^{-\sigma_S\sqrt{h}} - 1). \quad (36)$$

**Proof.** First of all, notice that for  $j^* \leq j$  one has  $U_{i+1,j^*} \leq U_{i,j}$ , so that  $S_{i+1,j^*} - S_{i,j} \leq 0 \leq (\mu_S)_{i,j,k}h$ . This gives  $j_d(i, j, k) = j$ , for every  $i, j, k$  and  $h$ , so that

$$S_{i+1,j_d(i,j,k)} - S_{i,j} = S_{i+1,j} - S_{i,j} = S_{i,j}(e^{-\sigma_S\sqrt{h}} - 1).$$

Concerning the up-jump, it is sufficient to prove that for every  $h$  sufficiently small one has

$$S_{i+1,j+1} - S_{i,j} \geq (\mu_S)_{i,j,k}h,$$

that is equivalent to

$$S_{i,j}(e^{\sigma_S\sqrt{h}} - 1 - r_{i,k}h) \geq 0.$$

Since  $e^x - 1 \geq x$  for  $x > 0$ , for  $r_{i,k} \leq r_*$  we can write

$$S_{i,j}(e^{\sigma_S\sqrt{h}} - 1 - r_{i,k}h) \geq S_{i,j}(\sigma_S\sqrt{h} - r_*h) = S_{i,j}\sqrt{h}(\sigma_S - r_*\sqrt{h}),$$

and the last term is positive for  $h < h_2 < (\frac{\sigma_S}{r_*})^2$ . Finally,

$$S_{i+1,j_u(i,j,k)} - S_{i,j} = S_{i+1,j+1} - S_{i,j} = S_{i,j}(e^{\sigma_S\sqrt{h}} - 1).$$

□

**Remark 5.** We can state a remark similar to Remark 3: in Lemma 4 we actually proved that for  $h < h_2$  we have  $S_{i+1,j_u(i,j,k)} - S_{i+1,j_d(i,j,k)} > 0$ ,  $(\mu_S)_{i,j,k}h + S_{i,k} - S_{i+1,j_d(i,j,k)} \geq 0$  and  $S_{i+1,j_u(i,j,k)} - (\mu_S)_{i,j,k}h - S_{i,k} \geq 0$ . Therefore, as  $h < h_2$ , the up-jump probability in (35) can be rewritten as

$$\hat{p}_{i,j,k} = \frac{(\mu_S)_{i,j,k}h + S_{i,j} - S_{i+1,j_d(i,j,k)}}{S_{i+1,j_u(i,j,k)} - S_{i+1,j_d(i,j,k)}}. \quad (37)$$

We are now ready to set up and discuss the transition matrix of the bivariate Markov chain. We recall that, starting from the node  $(i, j, k)$ , we have used the following notations:  $q_{i,j_u,k_u}$ ,  $q_{i,j_u,k_d}$ ,  $q_{i,j_d,k_u}$ , and  $q_{i,j_d,k_d}$  stand for the probability to reach  $(i + 1, j_u, k_u)$ ,  $(i + 1, j_u, k_d)$ ,  $(i + 1, j_d, k_u)$  and  $(i + 1, j_d, k_d)$  respectively.

**Proposition 6.** Let  $r_* > 0$  and  $S_* > 0$  be fixed and set  $A_* = \{(i, j, k) : r_{i,k} \leq r_*, S_{i,j} \leq S_*\}$ . Let  $\theta_*$  be as in Lemma 2 and  $(i, j, k) \in A_*$ . We set:

i) if  $(i, j, k) \in A_*$  and  $r_{i,k} \leq \theta_*\sqrt{h}$  then

$$\begin{aligned} q_{i,j_u,k_u} &= \hat{p}_{i,j_u,k_u}p_{i,k_u}, & q_{i,j_u,k_d} &= \hat{p}_{i,j_u,k_u}(1 - p_{i,k_u}), \\ q_{i,j_d,k_u} &= (1 - \hat{p}_{i,j_u,k_u})p_{i,k_u}, & q_{i,j_d,k_d} &= (1 - \hat{p}_{i,j_u,k_u})(1 - p_{i,k_u}); \end{aligned}$$

ii) if  $(i, j, k) \in A_*$  and  $r_{i,k} \geq \theta_*\sqrt{h}$  then  $q_{i,j_u,k_u}$ ,  $q_{i,j_u,k_d}$ ,  $q_{i,j_d,k_u}$  and  $q_{i,j_d,k_d}$  are set as the solutions of the linear system (21).

Then there exists  $h_3 < 1$  and a positive constant  $C$  such that for every  $h < h_3$  and  $(i, j, k) \in A_*$  the above probabilities are actually well defined.

**Proof.** We fix the node  $(S_{i,j}, r_{i,k})$ , with  $(i, j, k) \in A_*$ .

i) If  $r_{i,k} = 0$ , the defined transition probabilities solve system (21). And if  $r_{i,k}$  is small enough then the transition probabilities are supposed to be close to the ones in 0. So, what is said in i) is that for  $r_{i,k}$  positive but small, we just consider the behavior in 0.

ii) We assume here that  $r_{i,k} \geq \theta_* \sqrt{h}$ . If  $M$  denotes the  $4 \times 4$  matrix underlying the linear system (21), then straightforward computations give

$$\det M = -(S_{i+1,j_u(i,j,k)} - S_{i+1,j_d(i,j,k)})(r_{i+1,k_u(i,k)} - r_{i+1,k_d(i,k)})$$

which is non null because, by construction, both factors are positive. So, a unique solution  $x \in \mathbb{R}^4$  really exists for every  $h$ . One has only to check that this actually gives a probability distribution, and this reduces to check that all entries of  $x$  are non negative. We set the solution as follows:

$$\begin{aligned} x_1 &= \hat{p}_{i,j_u,k_u} p_{i,k_u} (1 + g_{i,j,k}(\rho)), & x_2 &= \hat{p}_{i,j_u,k_u} (1 - p_{i,k_u}) (1 - g_{i,j,k}(\rho)), \\ x_3 &= (1 - \hat{p}_{i,j_u,k_u}) p_{i,k_u} (1 - g_{i,j,k}(\rho)), & x_4 &= (1 - \hat{p}_{i,j_u,k_u}) (1 - p_{i,k_u}) (1 + g_{i,j,k}(\rho)) \end{aligned}$$

(just to be clear, we have implicitly given to the four transition nodes the following ordering:  $(i+1, j_u, k_u)$ ,  $(i+1, j_u, k_d)$ ,  $(i+1, j_d, k_u)$ ,  $(i+1, j_d, k_d)$ ). It is clear that  $x$  solves the first three equations. As for the fourth one, we take  $h < h_1 \wedge h_2$ ,  $h_1$  and  $h_2$  given in Lemma 2 and Lemma 4 respectively, so that we can use both (30) and (37). And easy computations give

$$\begin{aligned} g_{i,j,k}(\rho) &= \frac{\rho \sigma_r \sigma_S \sqrt{r_{i,k}} S_{i,j} h - (\mu_r)_{i,k} (\mu_S)_{i,j,k} h^2}{\left( \Delta_u S_{i,j,k} \hat{p}_{i,j,k} - \Delta_d S_{i,j,k} (1 - \hat{p}_{i,j,k}) \right) \left( \Delta_u r_{i,k} p_{i,k} - \Delta_d r_{i,k} (1 - p_{i,k}) \right)} \\ &= \frac{\sigma_r \sigma_S \sqrt{r_{i,k}} S_{i,j} h}{\left( \Delta_u S_{i,j,k} \hat{p}_{i,j,k} - \Delta_d S_{i,j,k} (1 - \hat{p}_{i,j,k}) \right) \left( \Delta_u r_{i,k} p_{i,k} - \Delta_d r_{i,k} (1 - p_{i,k}) \right)} \times \\ &\quad \times \left( \rho - \frac{(\mu_r)_{i,k} (\mu_S)_{i,j,k} h^2}{\sigma_r \sigma_S \sqrt{r_{i,k}} S_{i,j} h} \right), \end{aligned}$$

in which we have set  $\Delta_u S_{i,j,k} = S_{i+1,j_u(i,j,k)} - S_{i,j}$ ,  $\Delta_d S_{i,j,k} = S_{i+1,j_d(i,j,k)} - S_{i,j}$ ,  $\Delta_u r_{i,k} = r_{i+1,k_u(i,k)} - r_{i,k}$  and  $\Delta_d r_{i,k} = r_{i+1,k_d(i,k)} - r_{i,k}$ . So, we only need to show that for small values of  $h$  one gets

$$\sup_{(i,j,k) \in A_* \text{ and } r_{i,k} \geq \theta_* \sqrt{h}} |g(\rho)| < 1.$$

We write

$$g_{i,j,k}(\rho) = \frac{1}{\alpha_{i,j,k}} \left( \rho - \frac{(\mu_r)_{i,k} (\mu_S)_{i,j,k} h^2}{\sigma_r \sigma_S \sqrt{r_{i,k}} S_{i,j} h} \right).$$

with

$$\alpha_{i,j,k} = \frac{\left( \Delta_u S_{i,j,k} \hat{p}_{i,j,k} - \Delta_d S_{i,j,k} (1 - \hat{p}_{i,j,k}) \right) \left( \Delta_u r_{i,k} p_{i,k} - \Delta_d r_{i,k} (1 - p_{i,k}) \right)}{\sigma_r \sigma_S \sqrt{r_{i,k}} S_{i,j} h}$$

We use (28), (36) and we write, for  $h$  small,

$$\begin{aligned}\Delta_u S_{i,j,k} &= S_{i,j} e^{\sigma_S \sqrt{h}} (1 - e^{-\sigma_S \sqrt{h}}) \geq S_{i,j} (1 - e^{-\sigma_S \sqrt{h}}) = -\Delta_d S_{i,j,k} \\ \Delta_u r_{i,k} &= \sigma_r \sqrt{r_{i,k} h} + \frac{\sigma_r^2}{4} h \geq \sigma_r \sqrt{r_{i,k} h} - \frac{\sigma_r^2}{4} h = -\Delta_d r_{i,k}\end{aligned}$$

so that

$$\alpha_{i,j,k} \geq \frac{S_{i,j} (1 - e^{-\sigma_S \sqrt{h}}) (\sigma_r \sqrt{r_{i,k} h} - \frac{\sigma_r^2}{4} h)}{\sigma_r \sigma_S \sqrt{r_{i,k}} S_{i,j} h} = \frac{(1 - e^{-\sigma_S \sqrt{h}}) (\sigma_r \sqrt{r_{i,k} h} - \frac{\sigma_r^2}{4} h)}{\sigma_r \sigma_S \sqrt{r_{i,k}} h}.$$

By recalling that  $1 - e^{-\sigma_S \sqrt{h}} \geq \sigma_S \sqrt{h} - \frac{\sigma_S^2}{2} h e^{\sigma_S}$ , we get

$$\alpha_{i,j,k} \geq \left(1 - \frac{\sigma_S e^{\sigma_S}}{2} \sqrt{h}\right) \left(1 - \frac{\sigma_r}{4} \sqrt{\frac{h}{r_{i,k}}}\right) \geq \left(1 - \frac{\sigma_S e^{\sigma_S}}{2} h^{\frac{1}{2}}\right) \left(1 - \frac{\sigma_r}{4\sqrt{\theta_*}} h^{\frac{1}{4}}\right)$$

in which we have used that  $r_{i,k} \geq \theta_* \sqrt{h}$ . Setting  $c_* = \frac{\sigma_r}{4\sqrt{\theta_*}} \wedge \frac{\sigma_S e^{\sigma_S}}{2}$ , we get

$$\alpha_{i,j,k} \geq \left(1 - c_* h^{\frac{1}{4}}\right)^2$$

Therefore, we obtain

$$\begin{aligned}|g_{i,j,k}(\rho)| &\leq \frac{1}{(1 - c_* h^{\frac{1}{4}})^2} \left| \rho - \frac{(\mu_r)_{i,k} (\mu_S)_{i,j,k} h^2}{\sigma_r \sigma_S \sqrt{r_{i,k}} S_{i,j} h} \right| \\ &\leq \frac{1}{(1 - c_* h^{\frac{1}{4}})^2} \left( |\rho| + \frac{(\mu_r)_{i,k} (\mu_S)_{i,j,k} h^2}{\sigma_r \sigma_S \sqrt{r_{i,k}} S_{i,j} h} \right) \\ &\leq \frac{1}{(1 - c_* h^{\frac{1}{4}})^2} \left( |\rho| + \frac{\kappa(\theta + r_*)}{\sigma_r \sqrt{\theta_*}} h^{1/4} + \frac{r_*}{\sigma_S} h^{1/2} \right)\end{aligned}$$

and since  $|\rho| < 1$ , the last r.h.s. can be set less than 1 for  $h$  small enough.  $\square$

We can now state the main result. In order to do this, we set  $(S_i^h, r_i^h)_{i=0,1,\dots,N}$  the Markov chain running on the bivariate lattice structure, that is:

- $S_0^h = S_0$  and  $r_0^h = r_0$ ;
- at time  $ih$ , the state-space for the pair  $(S_i^h, r_i^h)$  is given by  $\{(S_{i,j}, r_{i,k}) : j, k = 0, 1, \dots, i\}$ ;
- from time  $ih$  to time  $(i+1)h$  the transition law on  $\mathbb{R}^2$  is given by

$$\begin{aligned}\Pi_h(r_{i,k}, S_{i,j}; dx) &= q_{i,j_u,k_u} \delta_{\{(S_{i+1,j_u}, r_{i+1,k_u})\}}(dx) + q_{i,j_u,k_d} \delta_{\{(S_{i+1,j_u}, r_{i+1,k_d})\}}(dx) \\ &\quad + q_{i,j_d,k_u} \delta_{\{(S_{i+1,j_d}, r_{i+1,k_u})\}}(dx) + q_{i,j_d,k_d} \delta_{\{(S_{i+1,j_d}, r_{i+1,k_d})\}}(dx),\end{aligned}$$

where  $\delta_{\{a\}}$  denotes here the Dirac mass in  $a \in \mathbb{R}^2$  and the above probabilities are given in Proposition 6.



We set now  $(\bar{S}_t^h, \bar{r}_t^h)_{t \in [0, T]}$  as the continuous-time process defined through the piecewise constant and càdlàg interpolation (in time) of the chain: for  $t \in [ih, (i+1)h]$ ,

$$\bar{S}_t^h = S_i^h \quad \text{and} \quad \bar{r}_t^h = r_i^h$$

We are now ready to state the convergence in law of the Markov chain  $(r_i^h, S_i^h)_{i=0, \dots, N}$  to the diffusion process of our interest. To this purpose, we set  $D([0, T]; \mathbb{R}^2)$  the set of the càdlàg functions on  $[0, T]$  taking values in  $\mathbb{R}^2$ , endowed with the Skorohod topology (see e.g. Billingsley [3]).

**Theorem 7.** *As  $h \rightarrow 0$ , the family of Markov processes  $(\bar{S}^h, \bar{r}^h)_h$  converges in law on the space  $D([0, T]; \mathbb{R}^2)$  to the diffusion process  $(S, r)$  solution to the equations (1) and (2).*

**Proof.** The proof is standard, see e.g. Nelson and Ramaswamy [13] or also classical books such as Billingsley [3], Ethier and Kurtz [8] or Stroock and Varadhan [14].

To simplify the notations, let us set

$$\begin{aligned} \mathcal{M}_{i,j,k}^S(\ell) &= \mathbb{E}\left((S_{i+1}^h - S_i^h)^\ell \mid (S_i^h, r_i^h) = (S_{i,j}, r_{i,k})\right), \quad \ell = 1, 2, 4, \\ \mathcal{M}_{i,j,k}^r(\ell) &= \mathbb{E}\left((r_{i+1}^h - r_i^h)^\ell \mid (S_i^h, r_i^h) = (S_{i,j}, r_{i,k})\right), \quad \ell = 1, 2, 4, \\ \mathcal{M}_{i,j,k}^{S,r} &= \mathbb{E}\left((S_{i+1}^h - S_i^h)(r_{i+1}^h - r_i^h) \mid (S_i^h, r_i^h) = (S_{i,j}, r_{i,k})\right). \end{aligned}$$

It is clear that  $\mathcal{M}_{i,j,k}^S(\ell)$  is the local moment of order  $\ell$  at time  $ih$  related to  $S$ , as well as  $\mathcal{M}_{i,j,k}^r(\ell)$  is similar but related to the component  $r$ , and  $\mathcal{M}_{i,j,k}^{S,r}$  is the local cross-moment of the two components at the generic time step  $i$ .

So, the proof of the theorem relies in checking that for  $r_* > 0$  and  $S_* > 0$  fixed, setting  $A_* = \{(i, j, k) : r_{i,k} \leq r_*, S_{i,j} \leq S_*\}$ , then the following properties *i*), *ii*) and *iii*) hold:

*i*) (convergence of the local drift)

$$\begin{aligned} \lim_{h \rightarrow 0} \sup_{(i,j,k) \in A_*} \frac{1}{h} \left| \mathcal{M}_{i,j,k}^S(1) - (\mu_S)_{i,j,k} h \right| &= 0, \\ \lim_{h \rightarrow 0} \sup_{(i,j,k) \in A_*} \frac{1}{h} \left| \mathcal{M}_{i,j,k}^r(1) - (\mu_r)_{i,j,k} h \right| &= 0; \end{aligned}$$

*ii*) (convergence of the local diffusion coefficient)

$$\begin{aligned} \lim_{h \rightarrow 0} \sup_{(i,j,k) \in A_*} \frac{1}{h} \left| \mathcal{M}_{i,j,k}^S(2) - \sigma_S^2 S_{i,j}^2 h \right| &= 0, \\ \lim_{h \rightarrow 0} \sup_{(i,j,k) \in A_*} \frac{1}{h} \left| \mathcal{M}_{i,j,k}^r(2) - \sigma_r^2 r_{i,k}^2 h \right| &= 0 \\ \lim_{h \rightarrow 0} \sup_{(i,j,k) \in A_*} \frac{1}{h} \left| \mathcal{M}_{i,j,k}^{S,r} - \rho \sigma_r \sigma_S S_{i,j} \sqrt{r_{i,k}} h \right| &= 0; \end{aligned}$$

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

$$\lim_{h \rightarrow 0} \sup_{(i,j,k) \in A_*} \frac{1}{h} \mathcal{M}_{i,j,k}^S(4) = 0,$$

$$\lim_{h \rightarrow 0} \sup_{(i,j,k) \in A_*} \frac{1}{h} \mathcal{M}_{i,j,k}^r(4) = 0.$$

So, we consider  $h$  such that  $h < \min(h_1, h_2, h_3)$ ,  $h_1$ ,  $h_2$  and  $h_3$  being given in Lemma 2, Lemma 4 and Proposition 6 respectively, so that we can use both Remark 3 and Remark 5.

*Proof of i).* By using (30) and (37), we immediately get that  $\mathcal{M}_{i,j,k}^S(1) = (\mu_S)_{i,j,k}h$  and  $\mathcal{M}_{i,j,k}^r(1) = (\mu_r)_{i,k}h$ , and *i)* holds.

*Proof of ii).* As for the cross-moment, this is really immediate when  $r_{i,k} \geq \theta_*\sqrt{h}$ : the transition probabilities solve system (21), whose last equation is actually  $\mathcal{M}_{i,j,k}^{S,r} = \rho\sigma_r\sigma_S S_{i,j}\sqrt{r_{i,k}}h$ . Consider now the case  $r_{i,k} < \theta_*\sqrt{h}$ : here the transition probabilities are of the product-type, so that  $\mathcal{M}_{i,j,k}^{S,r} = (\mu_r)_{i,k}(\mu_S)_{i,j,k}h^2$ . Therefore,

$$\begin{aligned} |\mathcal{M}_{i,j,k}^{S,r} - \rho\sigma_r\sigma_S S_{i,j}\sqrt{r_{i,k}}h| &= (\mu_r)_{i,k}(\mu_S)_{i,j,k}h^2 + \sigma_r\sigma_S S_{i,j}\sqrt{r_{i,k}}h \\ &\leq \kappa(\theta + r_*)r_*S_*h^2 + \sigma_r\sigma_S S_*\sqrt{\theta_*}h^{5/4} \\ &\leq Ch^{5/4}. \end{aligned}$$

As for the second order moment, again by using (30) and (37), it follows that

$$\begin{aligned} \mathcal{M}_{i,j,k}^S(2) &= (S_{i+1,j_u} + S_{i+1,j_d} - 2S_{i,j})(\mu_S)_{i,j,k}h + (S_{i+1,j_u} - S_{i,j})(S_{i,j} - S_{i+1,j_d}) \\ \mathcal{M}_{i,k}^r(2) &= (r_{i+1,k_u} + r_{i+1,k_d} - 2r_{i,k})(\mu_r)_{i,k}h + (r_{i+1,k_u} - r_{i,k})(r_{i,k} - r_{i+1,k_d}). \end{aligned}$$

So, by using (36), we get

$$\mathcal{M}_{i,j,k}^S(2) = S_{i,j}^2 \left( 2r_{i,k} (\cosh(\sigma_S \sqrt{h}) - 1)h + e^{-\sigma_S \sqrt{h}} (e^{\sigma_S \sqrt{h}} - 1)^2 \right).$$

Therefore, for  $(i, j, k) \in A_*$  we get

$$|\mathcal{M}_{i,j,k}^S(2) - \sigma_S^2 S_{i,j}^2 h| \leq 2S_*^2 r_* (\cosh(\sigma_S \sqrt{h}) - 1)h + S_*^2 \left| e^{-\sigma_S \sqrt{h}} (e^{\sigma_S \sqrt{h}} - 1)^2 - \sigma_S^2 h \right|$$

and the statement holds. Concerning the second moment on  $r$ , we first study the case  $r_{i,k} \leq \theta_*\sqrt{h}$ ,  $\theta_*$  as in Lemma 2. Then by using (27) we have

$$|\mathcal{M}_{i,k}^r(2) - \sigma_r^2 r_{i,k}h| \leq 2C_*\kappa\theta h^{3/4+1} + C_*^2 h^{3/2} \leq Ch^{3/2}.$$

Consider now the case  $\theta_*\sqrt{h} \leq r_{i,k} \leq r_*$ . Then, for  $h \leq (\theta^*/r_*)^2$  then  $r_{i,k} \leq \theta^*/\sqrt{h}$ ,  $\theta^*$  as in Lemma 2. So, we use (28) and we obtain

$$|\mathcal{M}_{i,k}^r(2) - \sigma_r^2 r_{i,k}h| \leq \kappa(\theta + r_*) \frac{\sigma_r^2}{4} h^2 + \frac{\sigma_r^4}{16} h^2$$

and the statement holds as well.

*Proof of iii).* First, straightforward computations give

$$\begin{aligned}\mathcal{M}_{i,j,k}^S(4) &= \left( (\mu_S)_{i,j,k}h + S_{i,j} - S_{i+1,j_d} \right) (S_{i+1,j_u} + S_{i+1,j_d} - 2S_{i,j}) \times \\ &\quad \times \left( (S_{i+1,j_u} - S_{i,j})^2 + (S_{i+1,j_d} - S_{i,j})^2 \right) + (S_{i+1,j_d} - S_{i,j})^4 \\ \mathcal{M}_{i,k}^r(4) &= \left( (\mu_r)_{i,k}h + r_{i,k} - r_{i+1,k_d} \right) (r_{i+1,k_u} + r_{i+1,k_d} - 2r_{i,k}) \times \\ &\quad \times \left( (r_{i+1,k_u} - r_{i,k})^2 + (r_{i+1,k_d} - r_{i,k})^2 \right) + (r_{i+1,k_d} - r_{i,k})^4.\end{aligned}$$

As for the first quantity, we immediately get

$$\begin{aligned}\mathcal{M}_{i,j,k}^S(4) &\leq 4S_*^4(r_*h + 1 - e^{-\sigma_S\sqrt{h}}) \left( \cosh(\sigma_S\sqrt{h}) - 1 \right) 2(e^{\sigma_S\sqrt{h}} - 1)^2 + \\ &\quad + S_*^4(e^{-\sigma_S\sqrt{h}} - 1)^4 \leq Ch^2\end{aligned}$$

Concerning the 4th moment for  $r$ , first notice that if  $r_{i,k} \leq \theta_*\sqrt{h}$  then (27) gives  $\mathcal{M}_{i,k}^r(4) \leq C_*h^3$ . If instead  $\theta_*\sqrt{h} \leq r_{i,k} \leq r_*$  then from (28) one gets

$$\begin{aligned}\mathcal{M}_{i,k}^r(4) &\leq \left( \kappa(\theta + r_*)h + \sigma_r\sqrt{r_*h} + \frac{\sigma_r^2}{4}h \right) \sigma_r^2h \left( \sigma_r\sqrt{r_*h} + \frac{\sigma_r^2}{4}h \right)^2 + \left( \sigma_r\sqrt{r_*h} + \frac{\sigma_r^2}{4}h \right)^4 \\ &\leq Ch^2.\end{aligned}$$

So, the proof is completed.  $\square$

We finally stress that, in order to obtain the weak convergence, we do not need to require the Feller condition ( $2\kappa\theta \geq \sigma_r^2$ ) or the “convergence condition” ( $4\kappa\theta \geq \sigma_r^2$ ), the latter being necessary to get the tree approximation of the process  $R$ . Here, we only request that  $\kappa > 0$  and  $\theta > 0$ , and these are natural conditions in a financial setting.

## 5.2 Convergence of the American prices

Theorem 7 can be used in order to discuss the convergence of the binomial price to the true one for American style options. Let  $f(t, x) : [0, T] \times C([0, T]) \rightarrow [0, +\infty)$  denote a (nice) payoff function, so that the price in the continuous-time model and in the discrete model are given by

$$\sup_{\tau \in \mathcal{G}_{0,T}} \mathbb{E} \left( e^{-\int_0^\tau r_s ds} f(\tau, S) \right) \quad \text{and} \quad \sup_{\sigma \in \mathcal{G}_{0,T}^h} \mathbb{E} \left( e^{-\int_0^\sigma \bar{r}_s^h ds} f(\sigma, \bar{S}^h) \right)$$

respectively, where  $\mathcal{G}_{0,T}$  and  $\mathcal{G}_{0,T}^h$  denote the stopping times in  $[0, T]$  w.r.t. the filtration  $\mathcal{F}_t = \sigma((S_s, r_s) : s \leq t)$  and  $\mathcal{F}_t^h = \sigma((\bar{S}_s, \bar{r}_s^h) : s \leq t)$  respectively. Let us say that, in principle, we should use different notations for the expectations above, that is we should write  $\mathbb{E}_{\mathbb{P}}$  and  $\mathbb{E}_{\mathbb{P}^h}$  related to the measures  $\mathbb{P}$  and  $\mathbb{P}^h$  of the probability space where the processes  $(S, r)$  and  $(\bar{S}^h, \bar{r}^h)$  are defined, but for the sake of simplicity we do not stress these notations.

Consider the following assumptions:

- (H1)  $f$  is a continuous function (in the product topology) and for every  $x, y \in C([0, T])$  such that  $x_s = y_s$  for every  $s \in [0, t]$  then  $f(t, x) = f(t, y)$ ;
- (H2) there exists  $\delta > 1$  and  $h_* > 0$  such that  $\sup_{h < h_*} \mathbb{E} \left( \sup_{t \leq T} |e^{-\int_0^t \bar{r}_s^h ds} f(t, \bar{S}^h)|^\delta \right) < \infty$ .

Then under (H1) and (H2), Theorem 7 allows one to approximate the price in the continuous-time model by means of the price computed in the discrete model. In fact, by Theorem 7, the binomial process weakly converges to the bivariate diffusion. In the European case, this is enough to ensure the convergence of the price for continuous and bounded payoffs. In the American case, this is a much more difficult problem, due to the presence of optimal stopping times. However, under (H1) and (H2), this statement is proved in Amin and Khanna [2] (but see also Lamberton and Pagès [12] or the discussion and the references quoted in the notes by Lamberton [11]). We stress that the boundedness condition is weakened by requiring the uniform integrability property in (H2).

As an immediate consequence, American put options can be numerically evaluated by means of the algorithm described in Section 4. But other options can be considered, for example when  $f$  is continuous and fulfills the following polynomial-growth condition: there exist  $C > 0$  and  $\gamma > 1$  such that

$$\sup_{t \leq T} |f(t, x)| \leq C \left( 1 + \sup_{t \leq T} |x_t|^\gamma \right). \quad (38)$$

In fact, by developing the arguments used in [2] Section 5 (namely, martingale properties and Doob's inequality), in next Proposition 8 we prove that for every  $p > 1$  there exists  $h_* < 1$  such that

$$\sup_{h < h_*} \mathbb{E} \left( \sup_{t \leq T} e^{-p \int_0^t \bar{r}_s^h ds} (\bar{S}_t^h)^p \right) < \infty. \quad (39)$$

Therefore, if (38) holds then assumption (H2) is satisfied, and if (H1) holds as well the convergence of the binomial American price will follow.

So, we conclude this section by proving that (39) really holds.

**Proposition 8.** *For every  $p > 1$  there exist  $h_* < 1$ , a positive constant  $C_{p,T}$  depending on  $p, T$  and a universal constant  $c > 0$  such that*

$$\sup_{h < h_*} \mathbb{E} \left( \sup_{t \leq T} e^{-p \int_0^t \bar{r}_s^h ds} (\bar{S}_t^h)^p \right) \leq C_{p,T} S_0^p \exp \left( c \frac{p(p-1)}{2} \sigma_S^2 T \right).$$

**Proof.** We use here the notation  $\mathbb{E}_i$  to denote the conditional expectation given the  $\sigma$ -algebra  $\mathcal{F}_i^h = \sigma((S_k^h, r_k^h) : k \leq i)$ ,  $i \leq N$ . Moreover, we notice that

$$\sup_{t \leq T} e^{-p \int_0^t \bar{r}_s^h ds} (\bar{S}_t^h)^p = \sup_{i \leq N} e^{-p \sum_{k=0}^{i-1} r_k^h h} (S_i^h)^p,$$

with the understanding  $\sum_{k=0}^{-1}(\cdot) = 0$ , so we work with the quantity in the r.h.s. above.

For  $i = 0, 1, \dots, N$ , we set  $\tilde{S}_i^h = \exp(-\sum_{k=0}^{i-1} r_k^h h) S_i^h$ , so that

$$\tilde{S}_{i+1}^h = \tilde{S}_i^h \times e^{-r_i^h h} \xi_{i+1}^h \quad \text{with} \quad \xi_{i+1}^h = \frac{S_{i+1}^h}{S_i^h}.$$

We recall that, by construction, the conditional law of  $S_{i+1}^h$  given  $\mathcal{F}_i^h$  is the following:

- on the set  $\{1 + r_i^h h \leq e^{\sigma_S \sqrt{h}}\}$ ,  $S_{i+1}^h$  can assume the values  $S_i^h e^{\sigma_S \sqrt{h}}$  and  $S_i^h e^{-\sigma_S \sqrt{h}}$  with probability

$$\hat{p}_i^h = \frac{1 + r_i^h h - e^{-\sigma_S \sqrt{h}}}{e^{\sigma_S \sqrt{h}} - e^{-\sigma_S \sqrt{h}}} \quad \text{and} \quad 1 - \hat{p}_i^h = \frac{e^{\sigma_S \sqrt{h}} - 1 - r_i^h h}{e^{\sigma_S \sqrt{h}} - e^{-\sigma_S \sqrt{h}}}$$

respectively; we also recall that, on this set,  $\mathbb{E}_i(S_{i+1}^h) = S_i^h(1 + r_i^h h)$ ;

- on the set  $\{1 + r_i^h h > e^{\sigma_S \sqrt{h}}\}$ ,  $S_{i+1}^h$  assumes the (unique) value  $S_i^h e^{\sigma_S \sqrt{h}}$ .

As a consequence, we get

$$e^{-r_i^h h} \mathbb{E}_i(\xi_{i+1}^h) = e^{-r_i^h h} (1 + r_i^h h) \mathbf{1}_{\{1 + r_i^h h \leq e^{\sigma_S \sqrt{h}}\}} + e^{-r_i^h h} e^{\sigma_S \sqrt{h}} \mathbf{1}_{\{1 + r_i^h h > e^{\sigma_S \sqrt{h}}\}} \leq 1$$

and

$$\left( \xi_{i+1}^h - \mathbb{E}_i(\xi_{i+1}^h) \right) \mathbf{1}_{\{1 + r_i^h h > e^{\sigma_S \sqrt{h}}\}} = 0$$

Therefore, we can write

$$\begin{aligned} \tilde{S}_{i+1}^h &= \tilde{S}_i^h e^{-r_i^h h} \mathbb{E}_i(\xi_{i+1}^h) + \tilde{S}_i^h e^{-r_i^h h} (\xi_{i+1}^h - \mathbb{E}_i(\xi_{i+1}^h)) \\ &\leq \tilde{S}_i^h + \tilde{S}_i^h e^{-r_i^h h} (\xi_{i+1}^h - (1 + r_i^h h)) \mathbf{1}_{\{1 + r_i^h h \leq e^{\sigma_S \sqrt{h}}\}}. \end{aligned}$$

So, by setting

$$Y_{i+1}^h = e^{-r_i^h h} (\xi_{i+1}^h - (1 + r_i^h h)) \mathbf{1}_{\{1 + r_i^h h \leq e^{\sigma_S \sqrt{h}}\}}$$

we get

$$\tilde{S}_{i+1}^h \leq \tilde{S}_i^h + \tilde{S}_i^h Y_{i+1}^h = \tilde{S}_i^h (1 + Y_{i+1}^h).$$

This first gives that  $1 + Y_{i+1}^h \geq 0$  for every  $i$  and secondly,  $\tilde{S}_n^h \leq S_0 M_n^h$  for every  $n = 0, 1, \dots, N$ , where

$$M_0^h = 1 \text{ and for } n = 1, \dots, N, \quad M_n^h = \prod_{k=1}^n (1 + Y_k^h).$$

Since  $\sup_{n \leq N} (\tilde{S}_n^h)^p \leq S_0^p \sup_{n \leq N} (M_n^h)^p$ , we proceed by proving that  $\sup_{n \leq N} (M_n^h)^p$  is integrable and by studying an upper estimate for its expectation which is independent of the choice of  $h$ .

We first notice that  $\mathbb{E}_n(Y_{n+1}^h) = 0$ , so that  $(M_n^h)_{n \leq N}$  is a martingale. Let us prove that it is bounded in  $L^p$  - this will allow us to apply the Doob inequality.

On the set  $\{1 + r_i^h h \leq e^{\sigma_S \sqrt{h}}\}$ , we can find  $h_1$  such that for every  $h < h_1$  we have  $|\xi_{i+1}^h - 1| \leq 2\sigma_S \sqrt{h}$  and  $r_i^h h \leq 2\sigma_S \sqrt{h}$ . This gives that

$$|Y_{i+1}^h| \leq 4\sigma_S \sqrt{h}$$

as  $h < h_1$ . By using the Taylor expansion to the function  $x \mapsto (1+x)^p$  for  $|x| \leq 4\sigma_S \sqrt{h}$ , there exists a positive (non random) constant  $c_{1,p}$  such that

$$(1 + Y_{i+1}^h)^p \leq 1 + pY_{i+1}^h + \frac{p(p-1)}{2} (Y_{i+1}^h)^2 + c_{1,p}h\sqrt{h},$$

so that

$$\mathbb{E}_i((1 + Y_{i+1}^h)^p) \leq 1 + \frac{p(p-1)}{2} \mathbb{E}_i((Y_{i+1}^h)^2) + c_{1,p}h\sqrt{h}.$$

Now, straightforward computations give

$$\mathbb{E}_i((Y_{i+1}^h)^2) = e^{-2r_i^h h} (e^{\sigma_S \sqrt{h}} - 1 - r_i^h h)(1 + r_i^h h - e^{-\sigma_S \sqrt{h}}) \mathbf{1}_{\{1+r_i^h h \leq e^{\sigma_S \sqrt{h}}\}}$$

and for  $h$  small enough we have

$$\begin{aligned} \mathbb{E}_i((Y_{i+1}^h)^2) &\leq (\sigma_S^2 h - (r_i^h h)^2 + c_2 h \sqrt{h}) \mathbf{1}_{\{1+r_i^h h \leq e^{\sigma_S \sqrt{h}}\}} \\ &\leq 5\sigma_S^2 h + c_2 h \sqrt{h} \end{aligned}$$

where  $c_2$  denotes a suitable positive constant. So, there exist  $h_* > 0$  and  $c_p > 0$  such that for every  $h < h_*$

$$\mathbb{E}_i((1 + Y_{i+1}^h)^p) \leq 1 + 5\frac{p(p-1)}{2} \sigma_S^2 h + c_p h \sqrt{h}.$$

Now we are done, because

$$\mathbb{E}((M_{n+1}^h)^p) = \mathbb{E}((M_n^h)^p \mathbb{E}_n((1 + Y_{n+1}^h)^p)) \leq \mathbb{E}((M_n^h)^p) \left(1 + 5\frac{p(p-1)}{2} \sigma_S^2 h + c_p h \sqrt{h}\right)$$

and by iteration,

$$\begin{aligned} \mathbb{E}((M_{n+1}^h)^p) &\leq \left(1 + 5\frac{p(p-1)}{2} \sigma_S^2 h + c_p h \sqrt{h}\right)^{n+1} \\ &\leq \exp\left(5\frac{p(p-1)}{2} \sigma_S^2 h(n+1) + c_p h \sqrt{h}(n+1)\right). \end{aligned}$$

In particular,

$$\sup_{n \leq N-1} \mathbb{E}((M_{n+1}^h)^p) \leq \exp\left(5\frac{p(p-1)}{2} \sigma_S^2 T + c_p T \sqrt{h}\right)$$

and this holds for every  $h < h_*$ . We can apply the Doob inequality and we get

$$\begin{aligned} \mathbb{E}\left(\sup_{n \leq N-1} (M_{n+1}^h)^p\right) &\leq \left(\frac{p}{p-1}\right)^p \exp\left(1 + 5\frac{p(p-1)}{2} \sigma_S^2 T + c_p T \sqrt{h}\right) \\ &\leq C_{p,T} \exp\left(5\frac{p(p-1)}{2} \sigma_S^2 T\right) \end{aligned}$$

and we stress the the above r.h.s. does not depend on  $h$ . The statement now follows.  $\square$

## 6 Numerical results

In this section we compare the performance of our lattice algorithm (called ACZ) with the procedures of Wei (WEI) and of Hilliard, Schwartz and Tucker (HST) for the computation of European and American options with the CIR stochastic interest rate.

In the European and American option contracts we are dealing with, we consider a set of parameters already used in Hilliard, Schwartz and Tucker [9]:  $S_0 = 100$ ,  $\sigma_S = 0.25$ ,  $r_0 = 0.06$ ,  $\theta = 0.1$ ,  $\kappa = 0.5$ ,  $\rho = -0.25$ . In order to study the numerical robustness of the algorithms, we choose different values for  $\sigma_r$ : we set  $\sigma_r = 0.08, 0.5, 1, 3$ . Let us remark that for  $\sigma_r = 0.5, 1, 3$ , the Feller condition  $2\kappa\theta \geq \sigma_r^2$  is not satisfied at all. The parameters of the option contracts are the following: the strike is  $K = 100$  and the maturity  $T$  is varying, since we set  $T = 1, 2$  years. Finally, the number of time steps  $N$  varies:  $N = 50, 100, 150, 200, 300$ .

Tables 1 and 2 report European put option prices for  $T = 1$  and  $T = 2$  respectively. The benchmark value is obtained using a Monte Carlo with very large number of Monte Carlo simulation (10 million iterations) using the accurate Alfonsi [1] discretization scheme for the CIR process with  $M = 300$  discretization time steps (this method provides a Monte Carlo weak second-order scheme for the CIR process, without any restriction on its parameters). We also provide results for American put option prices, as reported in Table 3 and Table 4 (no benchmarks are available in this case).

The numerical results show that our method provides very reliable and stable outcomes. Even if no odd results appear as  $\sigma_S$  varies, on the contrary both WEI and HST fail when  $\sigma_r$  increases. As already observed by Tian [15], [16], this follows from the procedure set up to approximate the CIR process and can be explained by looking at the behavior of the drift  $\mu_R$  (see (4)) associated to the transformed process  $R$ , which is the one used to define the transition probabilities (see Remark 1). Indeed, one can write

$$\mu_R(R) = \frac{1}{2\sigma_r^2} \left( \frac{4\kappa\theta - \sigma_r^2}{R} - \kappa\sigma_r^2 R \right).$$

We recall that the convergence is achieved when  $4\kappa\theta - \sigma_r^2 \geq 0$ . As  $\sigma_r = 0.5, 1, 3$ , one gets  $4\kappa\theta - \sigma_r^2 < 0$ , so that when  $R \downarrow 0$  one has  $\mu_R(R) \downarrow -\infty$  and  $p_{i,k} \rightarrow 0$ . This gives that, in some sense, the process  $r$  tends to be absorbed in 0. As  $\sigma_r = 0.35$  the Feller condition is not satisfied but the convergence condition holds, instead as  $\sigma_r = 0.08$  both conditions are true.

	$N$	WEI	HST	ACZ	MC Benchmark
$\sigma_r = 0.08$	50	6.599590	6.579999	6.547556	
	100	6.596199	6.586245	6.569844	(6.580864)
	150	6.594964	6.588312	6.577486	6.586622
	200	6.594367	6.589255	6.581016	(6.592380)
	300	6.593705	6.590247	6.584744	
$\sigma_r = 0.35$	50	6.525565	6.501327	6.456607	
	100	8.098863	6.485796	6.480764	(6.487199)
	150	6.506420	6.488178	6.488831	6.492901
	200	6.518722	6.453131	6.492840	(6.498603)
	300	6.506538	6.502675	6.496892	
$\sigma_r = 0.50$	50	6.522023	6.495957	6.515875	
	100	6.606752	nan	6.543780	(6.544551)
	150	6.708564	6.430882	6.551733	6.550315
	200	6.588212	6.308088	6.556571	(6.556078)
	300	6.492437	6.280427	6.560239	
$\sigma_r = 1.00$	50	7.397553	3.640714	7.054686	
	100	4.786837	nan	7.109525	(7.153304)
	150	8.046617	4.094663	7.123767	7.159471
	200	4.846171	4.082429	7.127271	(7.165637)
	300	7.340537	4.135390	7.126012	
$\sigma_r = 3.00$	50	9.582836	0.042281	8.491568	
	100	3.026463	0.074981	8.636541	(8.756826)
	150	15.966217	0.085990	8.671644	8.763625
	200	9.103675	0.081284	8.648793	(8.770423)
	300	1.604996	0.084082	8.681638	

Table 1: *European put options with  $T = 1$ ,  $S_0 = 100$ ,  $K = 100$ ,  $\sigma_S = 0.25$ ,  $r_0 = 0.06$ ,  $\theta = 0.1$ ,  $\kappa = 0.5$ ,  $\rho = -0.25$ ,  $\sigma_r$  varying.*



	$N$	WEI	HST	ACZ	MC Benchmark
$\sigma_r = 0.08$	50	7.113252	7.075978	7.044844	
	100	7.109863	7.090808	7.075287	(7.090164)
	150	7.108682	7.095834	7.085454	7.096171
	200	7.108057	7.098328	7.090541	(7.102178)
	300	7.107381	7.100858	7.095698	
$\sigma_r = 0.35$	50	19.700288	6.894164	7.150781	
	100	7.392442	7.187941	7.184543	(7.180155)
	150	7.190269	7.138343	7.196207	7.186305
	200	8.938681	7.089212	7.201783	(7.192455)
	300	7.205803	7.103088	7.207558	
$\sigma_r = 0.50$	50	7.590023	nan	7.532236	
	100	7.384775	7.151360	7.573197	(7.575268)
	150	8.099945	7.157489	7.590692	7.581702
	200	7.866820	nan	7.598707	(7.588135)
	300	8.627843	7.029057	7.600614	
$\sigma_r = 1.00$	50	10.511291	nan	9.029287	
	100	10.377910	1.513991	9.113234	(9.312621)
	150	11.204261	1.603589	9.204665	9.319903
	200	3.979039	nan	9.192707	(9.327185)
	300	12.541620	1.839346	9.206984	
$\sigma_r = 3.00$	50	0.003130	0.000216	11.801614	
	100	13.515823	0.000344	11.656249	(12.046144)
	150	12.799415	0.000594	11.820232	12.054222
	200	1.864718	0.000832	11.873238	(12.062299)
	300	24.377573	0.001110	11.919532	

Table 2: *European put options with  $T = 2$ ,  $S_0 = 100$ ,  $K = 100$ ,  $\sigma_S = 0.25$ ,  $r_0 = 0.06$ ,  $\theta = 0.1$ ,  $\kappa = 0.5$ ,  $\rho = -0.25$ ,  $\sigma_r$  varying.*

	$N$	WEI	HST	ACZ
$\sigma_r = 0.08$	50	7.459623	7.422693	7.438208
	100	7.456394	7.437777	7.445392
	150	7.455096	7.442827	7.447790
	200	7.454461	7.445104	7.448865
	300	7.453750	7.447485	7.449971
$\sigma_r = 0.35$	50	6.025684	7.539697	7.536313
	100	5.953750	7.544594	7.545435
	150	7.547122	7.547791	7.548400
	200	6.179678	7.554119	7.549846
	300	7.551032	7.550363	7.551299
$\sigma_r = 0.50$	50	7.647938	7.596269	7.650304
	100	7.698584	nan	7.662723
	150	7.751229	7.586111	7.665955
	200	7.684290	7.472087	7.668003
	300	7.620723	7.627011	7.669464
$\sigma_r = 1.00$	50	8.187062	6.511775	8.084019
	100	6.750600	nan	8.106474
	150	8.743109	6.506826	8.113036
	200	6.822003	6.501217	8.115025
	300	8.164598	6.534435	8.116760
$\sigma_r = 3.00$	50	9.699021	2.902701	8.953802
	100	5.821256	3.165914	9.009359
	150	16.069688	3.405395	9.028517
	200	9.212406	3.307452	9.025299
	300	4.740153	3.049339	9.037878

Table 3: *American put options with  $T = 1$ ,  $S_0 = 100$ ,  $K = 100$ ,  $\sigma_S = 0.25$ ,  $r_0 = 0.06$ ,  $\theta = 0.1$ ,  $\kappa = 0.5$ ,  $\rho = -0.25$ ,  $\sigma_r$  varying.*

	$N$	WEI	HST	ACZ
$\sigma_r = 0.08$	50	9.161152	9.087704	9.149096
	100	9.162292	9.125087	9.155979
	150	9.162352	9.137484	9.158033
	200	9.162318	9.143667	9.159035
	300	9.162238	9.149867	9.160028
$\sigma_r = 0.35$	50	4.459093	9.284505	9.499658
	100	4.956463	9.479192	9.507529
	150	9.494545	9.442510	9.510407
	200	4.826695	9.480988	9.511623
	300	9.505248	9.494247	9.512845
$\sigma_r = 0.50$	50	9.814438	nan	9.821675
	100	9.717844	9.555458	9.833291
	150	10.113582	9.570541	9.838509
	200	9.986565	nan	9.841491
	300	10.358344	9.576959	9.842231
$\sigma_r = 1.00$	50	11.421349	6.586697	10.790323
	100	11.365461	6.934987	10.834057
	150	11.826422	6.392858	10.868361
	200	7.790298	nan	10.863552
	300	13.160854	6.897618	10.871842
$\sigma_r = 3.00$	50	2.837551	3.291170	12.488520
	100	13.585301	2.903662	12.458445
	150	12.865822	2.672643	12.516032
	200	6.335779	3.166955	12.539396
	300	24.420598	3.405717	12.564801

Table 4: *American put options with  $T = 2$ ,  $S_0 = 100$ ,  $K = 10$ ,  $\sigma_S = 0.25$ ,  $r_0 = 0.06$ ,  $\theta = 0.1$ ,  $\kappa = 0.5$ ,  $\rho = -0.25$ ,  $\sigma_r$  varying.*

## References

- [1] Alfonsi, A. (2010): High order discretization schemes for the CIR process: application to affine term structure and Heston models. *Mathematics of Computation*, **79**, 209-237. [23](#)
- [2] Amin, K., Khanna, A. (1994): Convergence of American option values from discrete-to continuous-time financial models. *Mathematical Finance*, **4**, 289-304. [20](#)
- [3] Billingsley, P. (1968): *Convergence of probability measures*. Wiley, New York. [17](#)
- [4] Brennan, M. J., Schwartz, E. S. (1976): The pricing of equity-linked life insurance policies with an asset value guarantee. *Journal of Financial Economics*, **3**, 195-213. [1](#)
- [5] Briani, M., Caramellino, L., Zanette, A. (2013): Pricing American options in the Heston model: an hybrid tree-finite difference approach. Preprint, ArXiv:1307.7178. [2](#)
- [6] Costabile, M., Gaudenzi, M., Massabò, I., Zanette, A. (2009): Evaluating fair premiums of equity-linked policies with surrender option in a bivariate model. *Insurance: Mathematics and Economics*, **45**, 286-295. [3](#), [5](#), [7](#), [9](#)
- [7] Cox, J.C., Ingersoll, J., Ross, S. (1985): A theory of the term structure of interest rates. *Econometrica*, **53**, 385-407. [2](#)
- [8] Ethier, S.N., Kurtz, T. (1986): *Markov processes: characterization and convergence*. John Wiley & Sons, New York. [17](#)
- [9] Hilliard, J.E., Schwartz, A.L., Tucker, A.L. (1996): Bivariate binomial pricing with generalized interest rate processes. *The Journal of Financial Research*, **XIX-4**, 585-602. [2](#), [3](#), [7](#), [9](#), [23](#)
- [10] Kushner, H., Dupuis, P.G. (1992): *Numerical methods for stochastic control problems in continous time*. Springer Verlag.
- [11] Lamberton, D. (2009): *Optimal stopping and American options*. Ljubljana Summer School on Financial Mathematics, <http://www.fmf.uni-lj.si/finmath09/ShortCourseAmericanOptions.pdf>. [20](#)
- [12] Lamberton, D., Pagès G. (1990): Sur l'approximation des réduites. *Ann. Inst. H. Poincaré, Probab. et Statistiques* **26**, 331-355. [20](#)
- [13] Nelson, D. B., Ramaswamy, K. (1990): Simple binomial processes as diffusion approximations in financial models. *The Review of Financial Studies*, **3**, 393-430. [2](#), [3](#), [5](#), [8](#), [17](#)
- [14] Stroock, D. W., Varadhan, S.R.S. (1979): *Multidimensional Diffusion Processes*. Springer, Berlin. [17](#)

- [15] Tian, Y. (1992): A simplified binomial approach to the pricing of interest-rate contingent claims. *Journal of Financial Engineering*, **1**, 4-37. 8, 23
- [16] Tian, Y. (1994): A reexamination of lattice procedures for interest rate-contingent claims. *Advances in Futures and Options Research*, **7**, 87-110. 7, 8, 23
- [17] Wei, J., Z. (1996): Valuing American equity options with a stochastic interest rate: a note. *The Journal of Financial Engineering*, **2**, 195-206. 2, 3, 9