

# American Options and Stochastic Interest Rates

Anna Battauz

Department of Finance, Baffi-Carefin and IGIER,  
Bocconi University,  
Milan, Italy

Francesco Rotondi

Department of Finance,  
Bocconi University,  
Milan, Italy

May 2019

## Abstract

We study American equity options in a stochastic interest framework of Vasicek type (Vasicek (1977)). We allow for a non-zero correlation between the innovations driving the equity and the interest rates. We also allow for the interest rate to assume negative values, which is the case for some investment grade government bonds in Europe in recent years. We develop a bivariate discretization of the equity and interest rates processes by matching their original moments. The discretized processes converge in probability to their joint limiting distribution as the time step shrinks. The discretization is described by a recombining quadrinomial tree whose computational complexity is quadratic in the number of steps. We exploit our quadrinomial tree to evaluate American put options on the risky equity asset, characterizing also the optimal exercise policy. We analyze the two-dimensional free boundary, i.e. the underlying asset and the interest rate values that trigger the optimal exercise of the option. We document in the stochastic interest rates framework non-standard exercise policies associated with the double continuation region first described by Battauz et al. (2015) in constant interest rate framework.

*JEL Classification:* G13.

*Keywords:* American options, stochastic interest rates, binomial tree, quadrinomial tree.

# 1 Introduction

In an arbitrage-free financial market the role of the short-term interest rate is twofold: on one hand it represents the rate at which the equity appreciates; on the other hand it drives the locally risk-free asset and the related discount rate. Therefore, neglecting the variability of short-term interest rates may induce significant mispricings on both interest rates and equity derivatives. This issue is particularly relevant when derivatives are path-dependent and therefore sensitive to the entire path of the short-term interest rate, and not just its expected value at maturity. American equity call and put options, due to the optionality of their exercise policy, fall within this category. In fact, the holder of an American option has to timely choose when to cash in by exercising the option, balancing the effects from the discount rate and from the rate of return of the underlying asset. When both of these effects depend on a stochastic process, the valuation of the option becomes tricky. Our paper offers an intuitive and effective lattice method to compute both the price and the optimal exercise policies of American options on a risky asset with constant volatility in a stochastic interest rate framework of Vasicek type (see Vasicek (1977)). We employ the Vasicek mean-reverting model for the interest rate, because it allows the interest rate to assume mildly negative values, as the ones documented in recent years in the Eurozone<sup>1</sup>. We also allow for a correlation between the Brownian innovations of the interest rate and the risky security processes. The feasibility of negative interest rates within the Vasicek model, once a source of major criticism, has very recently become the reason of renewed interest in the model itself because of the aforementioned market circumstances. The literature on American equity options has so far focused on alternative stochastic interest rates models, such as the CIR one, based on the seminal work of Cox et al. (1985) (See<sup>2</sup> Medvedev and

---

<sup>1</sup>It is widely accepted to proxy the risk-free rate in Europe by the recently negative rates of German bonds.

<sup>2</sup>Medvedev and Scaillet (2010) introduce an analytical approach to price American options using a short-maturity asymptotic expansion. They perform a throughout numerical investigation for American call and

Scaillet (2010) and Boyarchenko and Levendorskiĭ (2013)). Our paper is, to our knowledge, the first that addresses the evaluation of American equity options in a stochastic interest rate framework of Vasicek type, allowing for the possibility of negative interest rates (see Detemple (2014) for an exhaustive review of the state of the art and Fabozzi et al. (2016) for a new recent quasi-analytic method to price and hedge American options on a lognormal asset with constant interest rate). First attempts to evaluate the impact of stochastic interest rates on American derivatives date back to Amin and Jr. (1995) and Ho et al. (1997). Nevertheless, both of them proxy American with Bermudan options with few exercise dates. Although this allows them to obtain closed form solutions for both the price of the options and their optimal exercise policy, the approximation of a continuum of exercise dates by just a couple of possible exercise dates leads to a heavy mispricing of the options and provides no accurate insight on the free boundary.

In the spirit of Cox et al. (1979), we propose a lattice-based approach to compute an American option's price, its optimal exercise policy and the related free boundary. Building on Nelson and Ramaswamy (1990), who provide a tree approximation for an univariate process, we construct a discrete joint approximation for the both the stock price and the interest rate processes. Hahn and Dyer (2008) develop a similar discretization for a correlated two-dimensional mean reverting process representing the price of two correlated commodities and they use it to evaluate the value of an oil and gas switching option. Our paper is different, as the mean reverting stochastic interest rate process enters the risk-neutral drift of our stock price, that has constant volatility and correlates with the interest rate.

---

put options with both stochastic CIR interest rates and stochastic underlying's volatility. Analogously, Boyarchenko and Levendorskiĭ (2013) consider a stochastic volatility equity and stochastic interest rates depending on two CIR factors, allowing for non-zero correlations between all the underlying processes. They provide a sophisticated iterative algorithm to price American derivatives that exploits a sequence of embedded perpetual options and their pricing results are in line with those of the Longstaff and Schwartz method and the asymptotics of Medvedev and Scaillet (2010).

In this framework, we provide a throughout investigation of American equity call and put options and their free boundaries. Our findings contribute to the literature on American options with stochastic interest rates, that usually restricts on non-negative interest rates. Detemple (2014) develops a state-of-the-art investigation of the American options' pricing problem in a stochastic non-negative interest rate framework. The free boundary is retrieved by a discretization of an integral equation for the early exercise premium decomposition. In particular, for American call options he argues that the exercise region is connected in the upward direction, namely, if immediate exercise is optimal at some underlying asset price at some fixed date  $t$ , it remains optimal at any higher asset price at date  $t$ . The property holds true in a non-negative interest rate framework. On the contrary, we show that such property may fail if interest rates may assume even only slightly negative values. In this case, we document the existence of a non standard double continuation region first described by Battauz et al. (2015) in a constant interest rate framework. In particular, a non-standard additional continuation region appears where the option is most deeply in the money. In this case, a mildly negative interest rates may lead to optimal postponement of the deeply in the money option as the holder is confident the option will still be in the money later on prefers to delay the cash-in.

The remaining of the paper is organized as follows: in Section 2 we introduce the financial framework we work in describing the stochastic processes we have and the related traded assets. We develop here the lattice-based discretization of the market, that we call *quadrinomial tree*. In Section 3 we deal with American put options describing numerical pricing results, showing the differences arising from the standard constant interest rate case and providing a graphical characterization of the optimal exercise policy and of the free boundaries. Section 4 concludes.

## 2 The market and the Quadrinomial Tree

### 2.1 The Assets in the Market

Consider a stylized financial market in a continuous time framework with investment horizon  $T > 0$ . A risky security  $S(t)$  is traded. Following the seminal work of Vasicek (1977), we assume a mean-reverting stochastic process for the prevailing short term interest rate on the market  $r(t)$ . We allow for a non zero correlation between the innovations of  $S$  and  $r$ . A market player can invest in the short-term interest rate, which is locally risk-free, through the money market account  $B(t)$ .

The dynamics of the risky security, of the short-term interest rate and of the money market account under the risk-neutral measure  $\mathbb{Q}$  are:

$$\begin{cases} \frac{dS(t)}{S(t)} = (r(t) - q)dt + \sigma_S dW_S^{\mathbb{Q}}(t) \\ dr(t) = \kappa(\theta - r(t))dt + \sigma_r dW_r^{\mathbb{Q}}(t) \\ dB(t) = r(t)B(t)dt \end{cases} \quad (1)$$

with  $\langle dW_S^{\mathbb{Q}}(t), dW_r^{\mathbb{Q}}(t) \rangle = \rho dt$  and given some initial conditions  $S(0) = S_0$ ,  $r(0) = r_0$  and  $B(0) = 1$ . Moreover:  $q$  is the deterministic constant annual dividend rate of the stock,  $\sigma_S > 0$  the volatility of the stock,  $\kappa$  the speed of mean-reversion of the short-term interest rate,  $\theta$  its long-run mean,  $\sigma_r > 0$  the volatility of the short-term interest rate and  $\rho \in [-1, 1]$  the correlation between the Brownian shocks on  $S$  and  $r$ .

System (1) can be rewritten equivalently in the following vectorial specification:

$$\begin{cases} \frac{dS(t)}{S(t)} = \mu_S dt + \nu_S \cdot dW^{\mathbb{Q}}(t) \\ dr(t) = \mu_r dt + \nu_r \cdot dW^{\mathbb{Q}}(t) \end{cases} \quad (2)$$

where  $\mu_S = (r(t) - q)$ ,  $\mu_r = \kappa(\theta - r(t))$ ,  $\nu_S = [\sigma_S \ 0]$ ,  $\nu_r = [\sigma_r \rho \ \sigma_r \sqrt{1 - \rho^2}]$ ,  $W^{\mathbb{Q}}(t) = [W_1^{\mathbb{Q}}(t) \ W_2^{\mathbb{Q}}(t)]'$  is a standard two-dimensional Brownian motion and  $\cdot$  is the matrix product.

The explicit solution to the system of SDEs in (1) is

$$\begin{cases} S(t) = S_0 \exp \left[ \int_0^t r(s) ds - \left( q + \frac{\sigma_S^2}{2} \right) t + \sigma_S W_S(t) \right] \\ r(t) = r_0 e^{-\kappa t} + \theta(1 - e^{-\kappa t}) + \sigma_r \int_0^t e^{-\kappa(t-s)} dW_r(s) \\ B(t) = \exp \left[ \int_0^t r(s) ds \right] \end{cases} \quad (3)$$

The zero-coupon bond with maturity  $T$  pays 1 at its holder at  $T$  and its price at  $t \in (0, T)$  is labelled with  $p(t, T)$ . By no arbitrage valuation, we have

$$p(t, T) = \mathbb{E}^{\mathbb{Q}} \left[ \frac{1}{B(T)} \middle| \mathcal{F}_t \right] = \mathbb{E}^{\mathbb{Q}} \left[ \exp \left[ - \int_0^T r(s) ds \right] \middle| \mathcal{F}_t \right],$$

that admits a closed formula solution as derived in Section 3.2.1 of Brigo and Mercurio (2007):

$$p(t, T) = e^{A(t, T) - B(t, T)r(t)} \quad (4)$$

where:

$$B(t, T) = \frac{1}{\kappa} \left( 1 - e^{-\kappa(T-t)} \right) \quad (5)$$

$$A(t, T) = \left( \theta - \frac{\sigma_r^2}{2\kappa^2} \right) (B(t, T) - (T - t)) - \frac{\sigma_r^2 B^2(t, T)}{4\kappa}. \quad (6)$$

In this fairly general pricing framework, the price of European put (and also of an European call) option on  $S$  can be derived in closed formulas. Exploiting a change of numéraire as described in Geman et al. (1995), it is possible to obtain a generalization of the Black-Scholes formula for the price of a put option:

**Proposition 1 (Value of the European put option)** *In the financial market specified in (1), the price at  $t \in [0, T]$  of an European put option on  $S$  with strike  $K$  is equal to*

$$\pi_E(t, S(t), r(t)) = Kp(t, T)N(-\tilde{d}_2) - S(t)e^{-q(T-t)}N(-\tilde{d}_1)$$

with<sup>3</sup>:

$$\begin{aligned}\tilde{d}_1 &= \frac{1}{\sqrt{\Sigma_{t,T}^2}} \left( \ln \frac{S(t)}{Kp(t,T)} + \frac{1}{2} \Sigma_{t,T}^2 - q(T-t) \right), \\ \tilde{d}_2 &= \tilde{d}_1 - \sqrt{\Sigma_{t,T}^2}, \\ \Sigma_{t,T}^2 &= \sigma_S^2(T-t) + 2\sigma_S\sigma_r\rho \left( \frac{-1 + e^{-\kappa(T-t)} + \kappa(T-t)}{k^2} \right) + \\ &\quad -\sigma_r^2 \left( \frac{3 + e^{-2\kappa(T-t)} - 4e^{-\kappa(T-t)} - 2\kappa(T-t)}{2k^3} \right).\end{aligned}$$

**Proof.** See Appendix B. ■

## 2.2 The Quadrinomial Tree

In their seminal work, Cox et al. (1979) show how to discretize a lognormal risky security and how to easily exploit such a binomial discretization in order to evaluate derivatives written on the primary asset. Embedding this geometric Brownian motion case into a more general class of diffusion processes, Nelson and Ramaswamy (1990) propose a one-dimensional scheme to properly define a binomial process that approximates a one-dimensional diffusion process. They do so by matching the diffusion's instantaneous drift and its variance and imposing a recombining structure to their discretized process. Can scheme can be used also for the univariate Ornstein–Uhlenbeck process.

The discretization via a tree/lattice structure of correlated processes, possibly of different kind, is more challenging. Gamba and Trigeorgis (2007) model two or more correlated geometric Brownian motion representing the price processes of different risky assets exploiting a log-transformation of the processes first and then an orthogonal decomposition of the shocks. In this way they are able to efficiently price derivatives on five correlated assets. Moving away from lognormality, Hahn and Dyer (2008) construct a quadrinomial lattice

---

<sup>3</sup>Notice that the current value of the interest rate  $r(t)$  enters the current price of the zero-coupon bond  $p(t, T)$  in  $\tilde{d}_1$  and  $\tilde{d}_2$ .

to approximate two mean-reverting processes in order to model two correlated one-factor commodity prices and evaluate derivatives on them.

We propose here a quadrinomial tree to jointly modelling a mean-reverting process for the short term interest rate as suggested first by Vasicek (1977) and the process for the risky stock's price with constant volatility and the drift that embeds the stochastic interest rate as in (1). A non constant short-term interest rates is surely more suitable from an option pricing perspective and an Ornstein–Uhlenbeck process enables us to investigate some interesting features of options when the discount rate becomes slightly negative. This situation, which was one of the major critique to the model proposed by Vasicek, actually reflects the present situation of European risk-free rates<sup>4</sup> with maturities up to 5 years.

We first show how to build the discretization of the processes  $(S(t), r(t))$  described by (1) and then we address the convergence issue of the discretization itself.

We apply Itô's Lemma to  $Y(t) := \ln(S(t))$  and we get

$$\begin{cases} dY(t) = \mu_Y dt + \sigma_S dW_S(t) \\ dr(t) = \mu_r dt + \sigma_r dW_r(t) \end{cases} \quad (7)$$

where  $\mu_Y := \left( r(t) - q - \frac{\sigma_S^2}{2} \right)$  and  $\mu_r := \kappa(\theta - r(t))$ . Again, the vectorial version of (2) is:

$$d \begin{bmatrix} Y(t) \\ r(t) \end{bmatrix} = \begin{bmatrix} \left( r(t) - q - \frac{\sigma_S^2}{2} \right) \\ \kappa(\theta - r(t)) \end{bmatrix} dt + \begin{bmatrix} \sigma_S & 0 \\ \sigma_r \rho & \sigma_r \sqrt{1 - \rho^2} \end{bmatrix} \cdot dW(t). \quad (8)$$

We refer here to the general technique of Section 11.3 of Stroock and Varadhan (1997) exploiting the very convenient notation introduced in Section 3.2.1 of Prigent (2003). For the ease of the reader we recall here their template. Consider the following bivariate SDE:

$$dX(t) = \mu(x, t)dt + \sigma(x, t) \cdot dW(t) \quad (9)$$

where  $X(t)_{t \geq 0} = (Y(t), r(t))_{t \geq 0}$ ,  $W(t)$  is a standard two-dimensional Brownian motion,  $\mu(x, t) : (\mathbb{R} \times \mathbb{R}^+) \times \mathbb{R}^+ \rightarrow \mathbb{R}^2$ ,  $\sigma(x, t) : (\mathbb{R} \times \mathbb{R}^+) \times \mathbb{R}^+ \rightarrow \mathbb{R}^{2 \times 2}$  and an initial condition

---

<sup>4</sup>it is widely accepted to proxy the risk-free rate in Europe by the rates of German bonds.



$X(0) = (x_0, r_0)$  is given.

Consider a discrete uniform partition of the time interval  $[0, T]$  like  $\{i\frac{T}{n}, i = 1, \dots, n\}$  and define  $\Delta t := \frac{T}{n}$ . For each  $n$  consider a bivariate stochastic process  $\{X_n\}$  on  $[0, T]$  which is constant between the nodes of the partition. At each node, both of the components of  $X_n$  jump up (or down) a certain distance with a certain probability. The sizes of the jumps and the probabilities are allowed to be time-dependent and state-contingent. Since after any jump each component of  $X_n$  can assume two new different values, there will be globally four possible outcomes after the jumps. For sake of clarity, fix  $n$  and consider the generic  $i$ -th step of the bivariate discrete process  $X_i = (Y_i, r_i)$ . In the following step, the process can assume only one of the following four values:

$$(Y_{i+1}, r_{i+1}) = \begin{cases} (Y_i + \Delta Y^+, r_i + \Delta r^+) & \text{with probability } q_{uu} \\ (Y_i + \Delta Y^+, r_i + \Delta r^-) & \text{with probability } q_{ud} \\ (Y_i + \Delta Y^-, r_i + \Delta r^+) & \text{with probability } q_{du} \\ (Y_i + \Delta Y^-, r_i + \Delta r^-) & \text{with probability } q_{dd} \end{cases} \quad (10)$$

where  $\Delta Y^\pm, \Delta r^\pm$  are the jumping increments and the four transition probabilities are both time-dependent and state-contingent. Figure (1) provides a graphical intuition for the bivariate binomial discretization over one step.

Globally, there are 8 parameters to pin down at each step:  $\Delta Y^\pm, \Delta r^\pm, q_{uu}, q_{ud}, q_{du}$  and  $q_{dd}$ .

In order to obtain a discretization that converges in probability to the solution of (9), we need to match the first two moments of  $Y(t)$  and  $r(t)$  as well as their cross variation. In this way we impose 5 conditions on the 8 parameters we need to determine. One more constraint has to be imposed on the four transition probabilities that need to sum up to one. Finally, we may want to impose a recombining structure to our quadrinomial tree in order to preserve tractability. Setting  $\Delta Y^- = -\Delta Y^+ := \Delta Y$  and  $\Delta r^- = -\Delta r^+ := \Delta r$  makes the number of different outcomes of our discretization grow quadratically (and non exponentially) in the number of steps. Figure (2) provides a graphical intuition of this

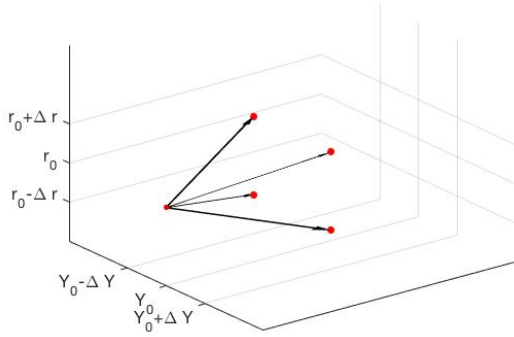


Figure 1: One step of the bivariate binomial discretization.

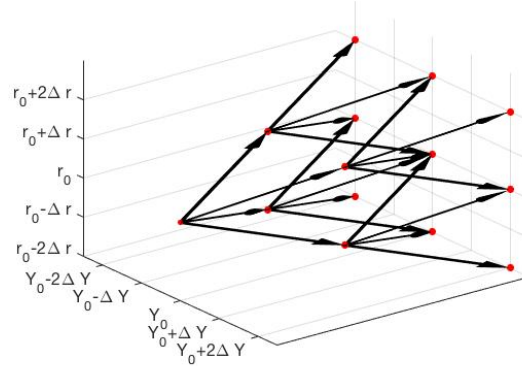


Figure 2: Two steps of the bivariate binomial discretization.

trick: starting from  $(Y_0, r_0)$ , after two steps the bivariate binomial process may assume 9 possible values, namely all the possible ordered couples of  $\{Y_0 - 2\Delta Y, Y_0, Y_0 + 2\Delta Y\}$  and of  $\{r_0 - 2\Delta r, r_0, r_0 + 2\Delta r\}$ . Thus, for a generic number of time steps  $n$ , the final possible outcomes of the discretization are  $(n+1)^2$  rather than  $2^{n+1}$ , the number of possible outcomes along a non recombining tree.

We now derive the explicit expressions of the increments and of the transition probabilities of our discretization for the bivariate SDEs (2), which, again, is just a simple instance of (9).

Matching the first two moments of  $Y(t)$  and  $S(t)$  as well as their cross-variation, neglecting the  $\Delta t$ -second order terms, imposing the proper constraint on the probabilities and imposing a recombining tree as explained above lead to the following system of 8 equations in 8

unknowns:

$$\left\{ \begin{array}{l} \mathbb{E}[\Delta Y] = (q_{uu} + q_{ud})\Delta Y^+ + (q_{du} + q_{dd})\Delta Y^- \stackrel{!}{=} \mu_Y \Delta t \\ \mathbb{E}[\Delta r] = (q_{uu} + q_{du})\Delta r^+ + (q_{ud} + q_{dd})\Delta r^- \stackrel{!}{=} \mu_r \Delta t \\ \mathbb{E}[\Delta Y^2] = (q_{uu} + q_{ud})(\Delta Y^+)^2 + (q_{du} + q_{dd})(\Delta Y^-)^2 \stackrel{!}{=} \sigma_S^2 \Delta t \\ \mathbb{E}[\Delta r^2] = (q_{uu} + q_{du})(\Delta r^+)^2 + (q_{ud} + q_{dd})(\Delta r^-)^2 \stackrel{!}{=} \sigma_r^2 \Delta t \\ \mathbb{E}[\Delta Y \Delta r] = q_{uu}\Delta Y^+ \Delta r^+ + q_{ud}\Delta Y^+ \Delta r^- + \\ \quad + q_{du}\Delta Y^- \Delta r^+ + q_{dd}\Delta Y^- \Delta r^- \stackrel{!}{=} \rho \sigma_S \sigma_r \Delta t \\ \quad \quad \quad q_{uu} + q_{ud} + q_{du} + q_{dd} \stackrel{!}{=} 1 \\ \quad \quad \quad \Delta Y^+ \stackrel{!}{=} -\Delta Y^- \\ \quad \quad \quad \Delta r^+ \stackrel{!}{=} -\Delta r^- \end{array} \right.$$

Imposing  $\Delta Y^+ > \Delta Y^-$  and  $\Delta r^+ > \Delta r^-$  we get:

$$\begin{aligned} \Delta Y^+ &= \sigma_S \sqrt{\Delta t} = -\Delta Y^- := \Delta Y \\ \Delta r^+ &= \sigma_r \sqrt{\Delta t} = -\Delta r^- := \Delta r \end{aligned} \tag{11}$$

$$\begin{aligned} q_{uu} &= \frac{\mu_Y \mu_r \Delta t + \mu_Y \Delta r + \mu_r \Delta Y + (1 + \rho) \sigma_r \sigma_S}{4\sigma_r \sigma_S} \\ q_{ud} &= \frac{-\mu_Y \mu_r \Delta t + \mu_Y \Delta r - \mu_r \Delta Y + (1 - \rho) \sigma_r \sigma_S}{4\sigma_r \sigma_S} \\ q_{du} &= \frac{-\mu_Y \mu_r \Delta t - \mu_Y \Delta r + \mu_r \Delta Y + (1 - \rho) \sigma_r \sigma_S}{4\sigma_r \sigma_S} \\ q_{dd} &= \frac{\mu_Y \mu_r \Delta t - \mu_Y \Delta r - \mu_r \Delta Y + (1 + \rho) \sigma_r \sigma_S}{4\sigma_r \sigma_S}. \end{aligned} \tag{12}$$

As noted in Nelson and Ramaswamy (1990), the four transition probabilities are not necessarily positive. In the limit, namely as  $\Delta t \rightarrow 0$ , we have  $\Delta Y, \Delta r \rightarrow 0$  and, therefore,  $q_{uu}, q_{dd} \rightarrow \frac{(1+\rho)}{4} > 0$  and  $q_{ud}, q_{du} \rightarrow \frac{(1-\rho)}{4} > 0$ . For  $\Delta t > 0$ , some of the four probabilities in (12) may become non-positive. In Appendix A we derive conditions ensuring the four probability in (12) are non-negative. Intuitively we get that  $q_{uu}, q_{du}, q_{ud}$  and  $q_{dd}$  are positive at any  $t$  as long as:

$$\underline{r} \leq r(t) \leq \bar{r}$$

where  $\underline{r}$  and  $\bar{r}$  depend on the parameters of the model but not on  $t$ . Namely, there exist two boundaries for the interest rate process outside which we need to “adjust” the transition probabilities in order to avoid the negative ones. We do so by setting equal to zero any

transition probability that becomes negative and normalizing to 1 the other ones. This ensures that our bivariate discretization covers in the limit the entire support of the bivariate vector  $(Y, r)$ .

Once we obtained a discretization of (2), we can map back the rate of return  $Y(t)$  to the level  $S(t)$  of the risky stock. In this way we have a lattice discretization of the solution of (2) and we name this discretization *quadrinomial tree*.

The following proposition shows that our quadrinomial tree converges in distribution to the solution of (2).

**Proposition 2 (Convergence of the quadrinomial tree)** *The bivariate discrete process  $(X_i)_i$  defined in (10) with the parameters in (11) and (12) converges in distribution to  $X(t) = (Y(t), r(t))$ .*

**Proof.** See Appendix B. ■

### 3 American Options

We now turn to American options. The holder of an American option can exercise it at any time before maturity  $T$ . This feature leads to an optimization problem as the holder tries to find the optimal exercise time that maximize her discounted profit. This flexibility does not come for free. Consequently, American options are usually more expensive than their European counterparts. This difference in the price is known as *early exercise premium*.

From now on, we will focus on American put option. In a continuous time framework, as described in Section 21.5 of Björk (2009), the value at  $t \leq T$  of an American put option on  $S$  with strike price  $K$  and maturity  $T$  is:

$$\begin{aligned} V(t) &= \text{ess sup}_{t \leq \tau \leq T} \mathbb{E}^{\mathbb{Q}} \left[ \frac{B(t)}{B(\tau)} (K - S(\tau))^+ \middle| \mathcal{F}_t \right] \\ &= \text{ess sup}_{t \leq \tau \leq T} \mathbb{E}^{\mathbb{Q}} \left[ e^{-\int_t^\tau r(s) ds} (K - S(\tau))^+ \middle| \mathcal{F}_t \right] \end{aligned} \quad (13)$$

where the *essential supremum* accounts for the fact that the sup is taken on an (uncountable) family of random variables defined up to zero-probability sets. In words, the value of the American put option is determined by the optimal stopping time  $\tau$  that maximizes the discounted payoff. It is well known that equation (13) admits no explicit formulation even in the standard Black and Scholes market. In our case, since  $r$  is stochastic and not independent of  $S$ , we cannot split the conditional expected value in (13) into two simpler separate ones. Nevertheless, the following Proposition shows that, as in the standard case of deterministic interest rate, the value of an American put option  $V(t)$  can be expressed as a deterministic function of time  $t$  (or, equivalently, of time to maturity  $T - t$ ) and of the current value of both the underlying asset  $S(t)$  and the short term interest rate  $r(t)$ .

**Proposition 3** (*Value of the American put option as a deterministic function*)

*In the market described by (1), the value of an American put option on  $S$  (13) is of the form:*

$$V(t) = F(t, S(t), r(t))$$

with  $F : [0, T] \times \mathbb{R}^+ \times \mathbb{R} \mapsto \mathbb{R}^+$  given by:

$$F(t, S, r) = \sup_{0 \leq \eta \leq T-t} \mathbb{E}^{\mathbb{Q}} \left[ \exp \left( - \int_0^{\eta} r(s) ds \right) \cdot \left( K - S \exp \left( \int_0^{\eta} r(s) ds - \left( q + \frac{1}{2} \sigma_S^2 \right) \eta + \sigma_S W_S(\eta) \right) \right)^+ \right]. \quad (14)$$

**Proof.** See Appendix B. ■

At each point in time the holder of the option compares the immediate payoff she would get from an immediate exercise and the discounted future value of the option. As at any  $t$  both the present and the expected future values of the option depend on the current value of the underlying asset,  $S$ , and of the interest rate,  $r$ , it is convenient to divide the plane  $(S, r)$  into two complementary regions: the *early exercise region* and the *continuation region*. Provided that the holder of the option can distinguish the two, if for a given  $t$  she

observes in the market a couple of values  $(S, r)$  that belongs to the early exercise region, the option will be exercised. If the observed couple belongs to the continuation one, the option will not be exercised at that moment.

We can now formally define the aforementioned regions. At each  $t \in [0, T]$ , the plane  $(S, r) \in \mathbb{R}^+ \times \mathbb{R}$  can be divided into:

- the continuation region  $CR(t) = \{(S, r) \in \mathbb{R}^+ \times \mathbb{R} : F(t, S, r) > (K - S)^+\}$ , the set of couples  $(S, r)$  where it is optimal to continue the option at  $t$ ;
- the early exercise region  $EEER(t) = \{(S, r) \in \mathbb{R}^+ \times \mathbb{R} : F(t, S, r) = (K - S)^+\}$ , the set of couples  $(S, r)$  where it is optimal to exercise the option at  $t$ .

When  $r = r_0$  is deterministic and strictly positive, the two regions are separated by a *free boundary*  $t \mapsto S^*(t)$ , called *critical price*. It has no explicit expression but its asymptotic approximation as the time to maturity shrinks has been derived by Evans et al. (2002) and Lamberton and Villeneuve (2003). In this case, the exercise region is downward connected, as shown in Figure (3): if it is optimal at  $t$  to exercise the option for some underlying values, it so for all the lower values of the underlying.

When, on the contrary,  $r = r_0 < 0$  and  $r_0 - q - \frac{\sigma_S^2}{2} > 0$ , Battauz et al. (2015) show that the early exercise region lies in between two continuation regions. In this case there are two free boundaries: an upper one,  $t \mapsto \bar{S}^*(t)$ , and a lower one,  $t \mapsto \underline{S}^*(t)$ . Again, they have no explicit expressions but their asymptotic behaviours have been derived by Battauz et al. (2015). In this case, the exercise region for the American put option is no more downward connected, as shown in Figure (4).

When  $r$  is constant, the intuition is the following. The value of the American put option, namely the counterpart of (14) is a deterministic function  $G(t, S) : \mathbb{R}^+ \times \mathbb{R} \rightarrow \mathbb{R}^+$  of the time and the current value of the underlying:

$$G(t, S) = \sup_{0 \leq \eta \leq T-t} \mathbb{E}^{\mathbb{Q}} \left[ e^{-r\eta} \left( K - S \exp \left( \left( r - q - \frac{\sigma_S^2}{2} \right) \eta + \sigma_S W_S(\eta) \right) \right)^+ \right]. \quad (15)$$

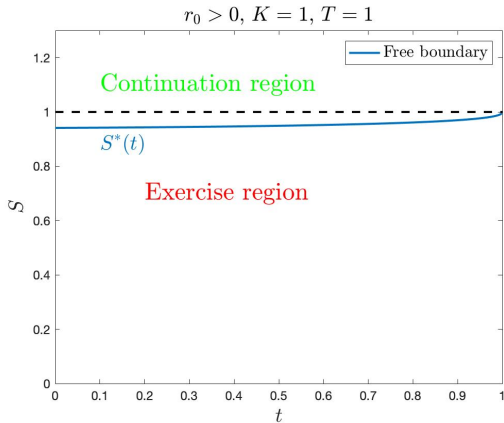


Figure 3: Free boundary of an American put option with  $r_0 > 0$ .

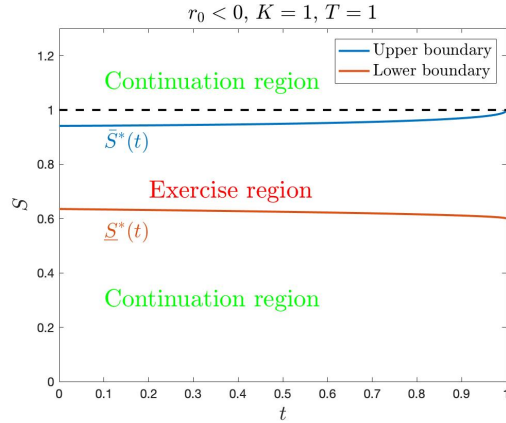


Figure 4: Free boundaries of an American put option with  $r_0 - q - \frac{\sigma_S^2}{2} < 0$  and  $r_0 < 0$ .

Denote by  $\mu := \left( r - q - \frac{\sigma_S^2}{2} \right)$  the drift of the risky stock's value. The value of the risky asset is increasing in  $\mu$  and, consequently, the value of the American put option is on average decreasing in  $\mu$ . The value of the American option is also decreasing, on average, with respect to the discount rate  $r$ . Therefore, the impact on the value of the American option of  $r$  streams from two channels: the discount factor  $e^{-r\eta}$  and the drift  $\mu$  of the risky asset. When  $r$  decreases and becomes negative, the negative discount rate will make the holder of the American option eager to wait and cash in later on to gain from the negative interest rate. If also  $\mu$  stays negative, the value of the underlying will decrease on average thus making the put option more valuable. Therefore, both of the effects will make the investor wait and exercise the option as late as possible, namely at maturity. In this case, the early exercise premium will disappear and the value of the American option will collapse into the value of its European counterpart. If, on the contrary,  $\mu$  stays positive, the value of the underlying will increase on average thus making the put option less valuable. In this case the effects of the discount factor and of the the drift are not aligned and a trade-off arises: the

investor would like to wait in order to cash in later on and gain from the negative interest rate but in this way she faces the risk of the payoff's depreciation, as the value underlying is expected to increase. In this case,  $r < 0$  and  $\mu = r - q - \frac{\sigma_S^2}{2} > 0$ , the aforementioned tradeoff translates in a non-empty exercise region surrounded by a double continuation region. In particular, a non standard continuation region appears when the option is mostly deeply in the money. Here the risk of the payoff's depreciation is very mild and the preference for the postponement due to the negative interest rate prevails.

Notice that in the latter case, as  $r < 0$  and  $\sigma_S > 0$ ,  $\mu > 0$  implies  $q < 0$ , which can be counter intuitive in the standard interpretation where  $q$  represents the continuous dividend yield of the risky asset. Nevertheless, a negative  $q$  can appear as a storage cost when dealing with options on commodities or as an "artificial drift" capturing the interplay of domestic and foreign interest rate when evaluating quanto options (see Battauz et al. (2018)).

When the interest rate is stochastic, both the discount factor  $e^{-\int_0^\eta r(s)ds}$  and the drift  $\mu := \int_0^\eta r(s)ds - (q + \frac{1}{2}\sigma_S^2)\eta$  in (14) depend on the whole path of  $r$ , from its current value up to the exercise time  $\eta$ . Therefore, at a given instant of time  $t \in [0, T]$  and for a generic couple of state variables  $(S, r)$ , the tradeoff between the discount factor and the expected drift of  $S$ , and consequently the presence of a singular or a double continuation region, depend on the all the possible future realizations of  $r$ , especially on its expected sign. The following Proposition formalizes this intuition and provides two necessary conditions for the existence of a double continuation region as the time to maturity of the option shrinks.

**Proposition 4** (*Asymptotic necessary conditions for the existence of a double continuation region*) *In the market described by (1), at any point in time  $t$  and given the current value of the interest rate  $r(t) = r$ , the following are jointly necessary conditions for the existence of the lower free boundary for the American put option for  $T - t \approx 0$ :*

[NC0] *the current level of the interest rate  $r$  is such that either  $T - t > \frac{2}{\kappa} \cdot \frac{r}{r - \theta}$  if  $r > \theta$  or*



$\kappa r < 0$  if  $r < \theta$ ;

[NC1] the dividend yield is non positive,  $q \leq 0$ ;

[NC2] for some  $S$ ,  $\pi_E(t, S, r) = (K - S)^+$ , where  $\pi_E(t, S, r)$  is the value of the European put option defined in Proposition 1.

**Proof.** See Appendix B. ■

With respect to [NC0], the case  $r < \theta$  is way much more likely: the long-run mean of the interest rate  $\theta$  is commonly assumed to be positive and above the current short-term interest rate. Furthermore, the speed of mean reversion  $\kappa$  of  $r$  is usually assumed to be positive: a negative  $\kappa$  would make the process  $r$  diverge, which is not the case according to traditional economic theory. Therefore, in most of the cases, [NC0] boils down to  $r < 0$ .

[NC1] ensures that the discounted price of the risky security is not a supermartingale. If this was the case, we show in the proof that this would lead automatically to optimal exercise at maturity only, under condition [NC0].

Under [NC0], [NC2] ensures that the price of the European put option  $\pi_E(t, S, r)$  does not dominate the immediate payoff value. If this was the case, the American put option would dominate the immediate payoff value as well, thus preventing the existence of early optimal exercise opportunities.

Condition [NC1] is poor. We empirically find the following finer condition:

[NC1'] the expected drift of  $S$  is positive, namely,  $(r - \theta)\alpha + \left(\theta - q - \frac{\sigma_S^2}{2}\right)(T - t) > 0$  where  $\alpha := \frac{1 - e^{-\kappa(T-t)}}{\kappa}$ .

Although the formal proof of the necessary conditions in Proposition 4 requires the time to maturity to be small enough, we show in the following section that actually those conditions ([NC1'] and [NC0]) correctly identify nodes on the tree in which a double continuation region appears along the whole lifetime of the option (see Figure 7).

### 3.1 Numerical examples

We now present and describe three illustrative numerical examples that show the optimal exercise strategies and the possible characterizations of the continuation region for the American put option in the market described by (1).

We exploit our quadrinomial tree to evaluate the American put option by backward induction. Once the whole quadrinomial tree, namely all the couples  $(S, r)$  and the related transition probabilities, have been generated, we start from the values of the state variables  $S$  and  $r$  at maturity  $T$ . At maturity, the American option is exercised in all the nodes in which it is in the money; the resulting payoff is the value of the American option at  $T$ . At any other generic instant  $t \in \{0, \Delta t, 2\Delta t, \dots, T - \Delta t\}$ , and for any couple  $(S(t), r(t))$ , we compute the immediate payoff  $(K - S(t))^+$  and we compare it to the continuation value of the option. The continuation value is obtained as the discounted (by the current realization of  $r(t)$ ) expected value (according the transition probabilities computed at  $(S(t), r(t))$ ) of the four values of the American option at  $t + \Delta t$  connected on the tree to the current node. From the comparison between the immediate exercise and the continuation value, we get the value of the American option in the node  $(S(t), r(t))$ . Going backward, we finally get the price of the American option at  $t = 0$ .

Proposition 2 showed that the quadrinomial tree we proposed converges in distribution to the bivariate process that solves (1), as the time step shrinks. Mulinacci and Pratelli (1998) prove that the convergence in distribution of the lattice-based approximation of the underlying state variables implies that the price of the American option evaluated according to the backward procedure described above converges to its theoretical value given by (13).

In all of the three following examples the parameters are:  $T = 1$ ,  $n = 125$ ,  $S_0 = K = 1$ ,  $\sigma_S = 0.15$ ,  $r_0 = 0$ ,  $\theta = 0.02$ ,  $\kappa = 1$ ,  $\sigma_r = 0.01$  and  $\rho = 0.05$ . The dividend yield  $q$  of the risky stock is the only parameter that varies across the examples: in the first one we set

$q = 0$ , in the second  $q = 0.02$  and  $q = -0.02$  in the last one.

For each example we:

- compute the value of the expected drift of  $S$  at  $t = 0$  over the whole investment horizon

$$\bar{\mu}|_{t=0} = \int_0^T \mathbb{E}^{\mathbb{Q}} [r(s)] d(s) - \left( q + \frac{\sigma_S^2}{2} \right) T$$

as this is one of the two key drivers of the option's exercise policy as explained below;

- compute the value at inception of the European counterpart  $\pi_E$  obtained both with the formula of Proposition 1 and along the quadrinomial tree (the values obtained in the two ways are indistinguishable);
- compute the value at inception of the American put option  $\pi_A$  along the quadrinomial tree;
- compute the price of the American put option,  $\pi_A^{r_0}$ , evaluated along the standard binomial tree of Cox et al. (1979) with a deterministic interest rate  $r = r_0 = 0^5$ . Our aim is to quantify the error that an “unsophisticated” investor would make by evaluating the American put option within a flat term structure framework rather than within a fluctuating one;
- graphically show how the lower and upper boundary (if any) look like in the space  $(t, S, r)$ . These graphs characterize the optimal exercise policy: at any  $t$ , the investor should look at the current values of  $(S(t), r(t))$ . If  $(S(t), r(t))$  is below the upper boundary and above the lower one (if any), she should exercise the option; otherwise, she is better off by postponing the exercise.

---

<sup>5</sup>we also evaluate the American put option with a deterministic interest rate equal to the expected value of  $r$  over the investment period; namely, we also set  $r = \mathbb{E}^{\mathbb{Q}} [r(T)] = 1.26\%$ . This exercise delivers qualitatively similar results. With respect to the last column of Table 1, the relative errors in this case are, respectively, 4.58%, 4.64%, 4.52%.

Numerical results are summed up in Table 1.

Figure	$q$	$\bar{\mu} _{t=0}$	$\pi_E$	$\pi_A$	$\pi_A^{r_0}$	$ \pi_A - \pi_A^{r_0} /\pi_A$
5	0%	0.139%	5.620%	5.990%	5.712%	4.87%
6	2%	2.139%	6.565%	6.973%	6.570%	6.14%
7	-2%	-1.861%	4.763%	5.240%	5.030%	4.18%

Table 1: Results from the three numerical examples.

First example:  $q = 0\%$ . If the underlying pays no dividend and its volatility is reasonably small, the expected drift of  $S$ ,  $\bar{\mu}$ , basically coincides with  $r(t)$ . This splits the domain of  $r$  in two complementary regions according to the sign of  $r$  (and of  $\bar{\mu}$ ), as shown in the bottom panel of Figure 5. In the left one where both  $r$  and  $\bar{\mu}$  are negative, the investor is willing to wait and to postpone the exercise as much as possible in order to gain from both the negative discount rate and the depreciation of  $S$ . In the right region, on the contrary, we have the standard tradeoff between a positive discount rate (that makes the investor willing to exercise the option as soon as possible) and a negative expected drift of  $S$  (that makes the investor willing to wait for a larger payoff). This generates the standard upper boundary shown in the top panels Figure 5.

The investor who believes that the term structure is flat and evaluates the American option with a constant discount rate makes a relative error almost equal to 5%.

Second example:  $q = 2\%$ . If the underlying pays (positive) dividends, the instantaneous drift of  $S$ ,  $\mu(t)$ , is equal to  $r(t)$  plus a negative quantity ( $-q - \frac{\sigma_S^2}{2} < 0$ ). As a result,  $\mu(t)$  (and  $\bar{\mu}$ ) is positive if and only if  $r$  (and its expected value) is sufficiently high. This splits the domain of  $r$  into three complementary regions, as shown in the bottom panel of Figure 6: the one in which  $r$  and  $\bar{\mu}$  are both negative, the one in which  $r$  is positive but small so that  $\bar{\mu}$  is still negative, the last one in which  $r$  and  $\bar{\mu}$  are both positive. In the first

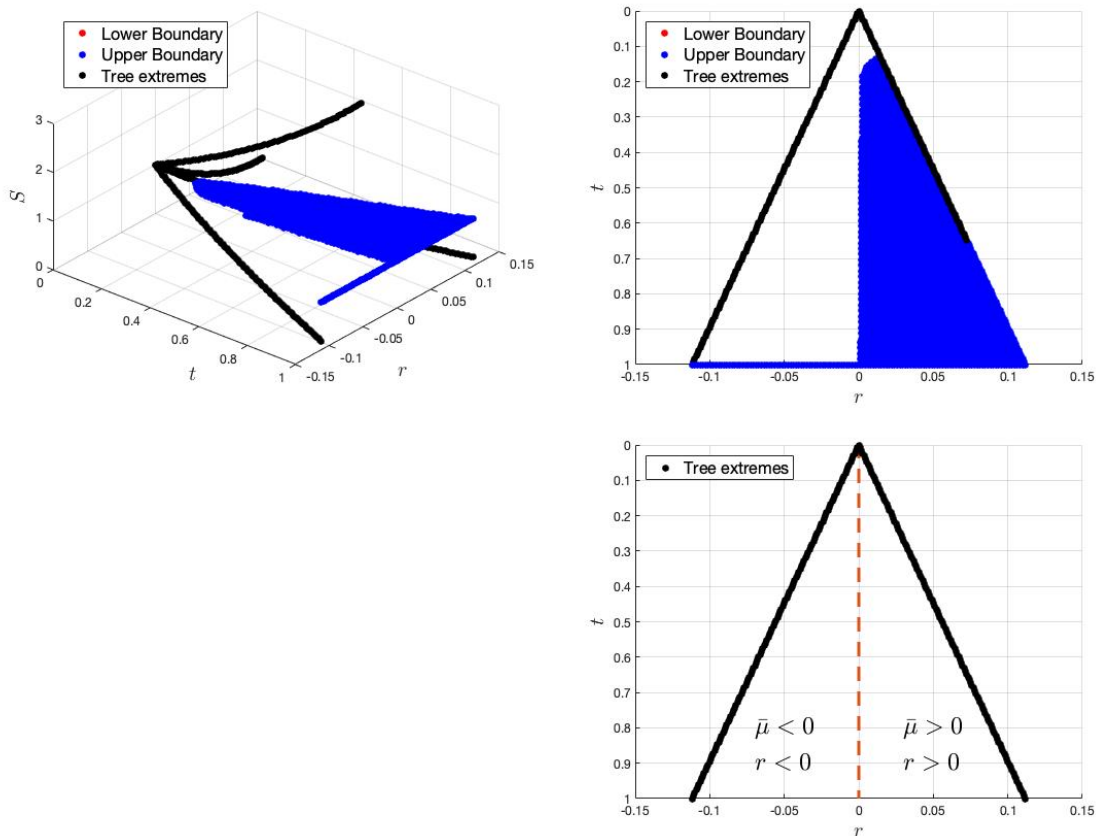


Figure 5: First example:  $q = 0\%$ .

one, the option is optimally exercised at maturity, as before. In the middle region there is a new tradeoff: the investor would like to cash in as soon as possible due to  $r > 0$  but the value of  $S$  is expected to decrease as  $\bar{\mu} < 0$ . This last condition allows for a standard upper boundary. The critical price below which the investor will exercise, though, becomes smaller as  $r$  approaches 0: as  $r$  decreases the threat of the positive discount rate weakens and, therefore, the investor would postpone the exercise unless the underlying reaches a very low level. In other words, if the discount is not that strong, the investor prefers to gain the high dividend yield keeping the asset as long as possible. In the last region, we find the standard behaviour already outlined in the first example.

The investor who believes that the term structure is flat and evaluates the American option

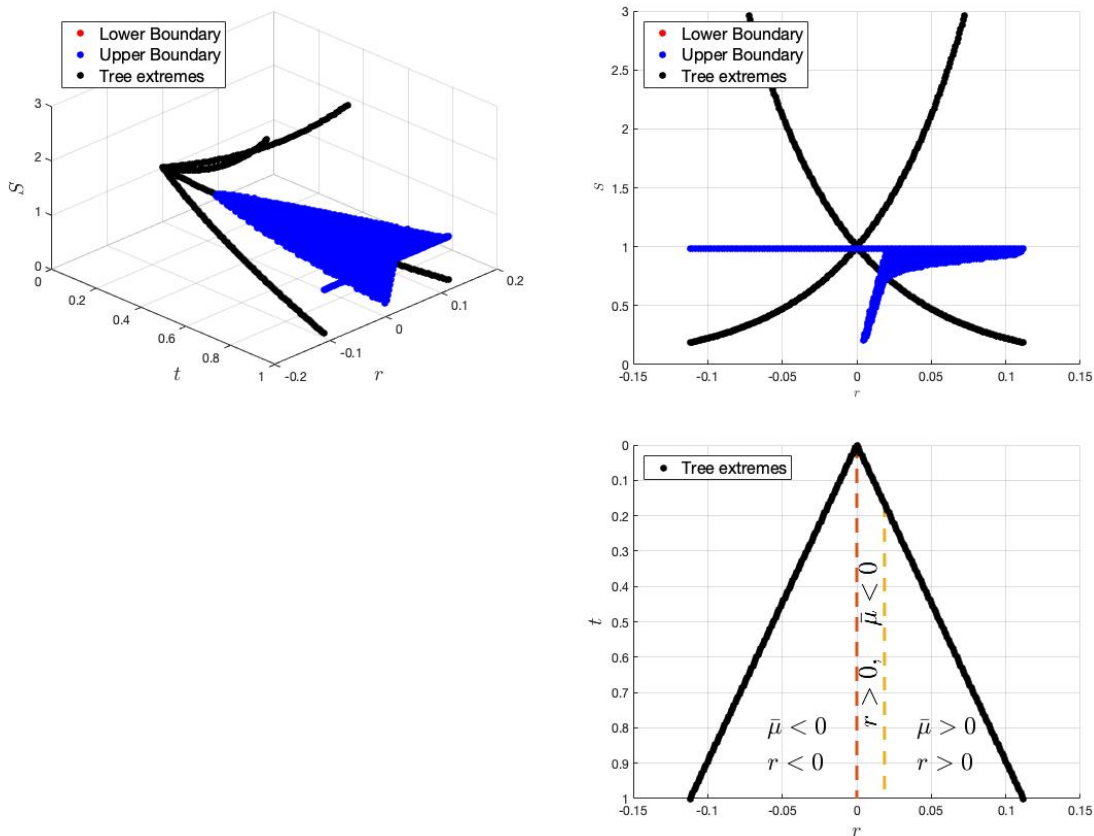


Figure 6: Second example:  $q = 2\%$ .

with a constant interest rate makes here an even higher relative error than before (6.14%).

Third example:  $q = -2\%$ . In the case of negative dividends, the instantaneous drift of  $S$ ,  $\mu(t)$ , is equal to  $r(t)$  plus a quantity which is now positive ( $-q - \frac{\sigma_S^2}{2} > 0$ ). As a result,  $\mu(t)$  (and  $\bar{\mu}$ ) may be positive also when  $r$  is mildly negative. This splits again the domain of  $r$  into three complementary regions, as shown in the middle panel of Figure 7: the one in which  $r$  and  $\bar{\mu}$  are both negative, the one in which  $r$  is negative but  $\bar{\mu}$  is positive and the last one in which  $r$  and  $\bar{\mu}$  are both positive. In the first region, the option is again optimally exercised at maturity as in the previous examples. In the middle region a double continuation region appears: this is the case in which the necessary conditions in Proposition 4 are satisfied as documented in the bottom panels of Figure 7. In the last region

where both  $r$  and  $\bar{\mu}$  are positive, we find the standard behaviour already outlined in the first two examples.

Finally, this is the case in which the investor who believes that the term structure is flat and evaluates the American option with a constant discount rate makes the largest relative error (6.14%).

## 4 Conclusions

In this paper we have studied American equity options in a stochastic correlated interest framework of Vasicek (1977) type. We have introduced a tractable lattice-based discretization of the equity and interest rate processes by means of a quadrinomial tree. Our quadrinomial tree matches the joint discretized moments of the equity and the stochastic interest rate and converges in probability to the continuous time original processes. This allowed us to employ our quadrinomial tree to characterize the two-dimensional free boundary for an American equity put option, that consists of the underlying asset and the interest rate values that trigger the optimal exercise of the option. Our results are in line with the existing literature when interest rates lie in the positive realm. In particular, the higher the dividend yield, the higher the benefits from deferring the option exercise. Moreover, in this case, the exercise region is downward connected with respect to the underlying asset value (see, for instance, Detemple (2014)). On the contrary, when interest rate are likely to assume even mildly negative values, optimal exercise policies change, depending on the tradeoff between the interest rate and the expected rate of return on the equity asset. If such expected rate of return is negative, optimal exercise occurs at maturity only as the option goes (on average) deeper in the money as time goes by and the negative interest rates make the investor willing to cash in as late as possible. If the expected rate of return on the equity asset is positive, the option is expected to move towards the out of the money re-

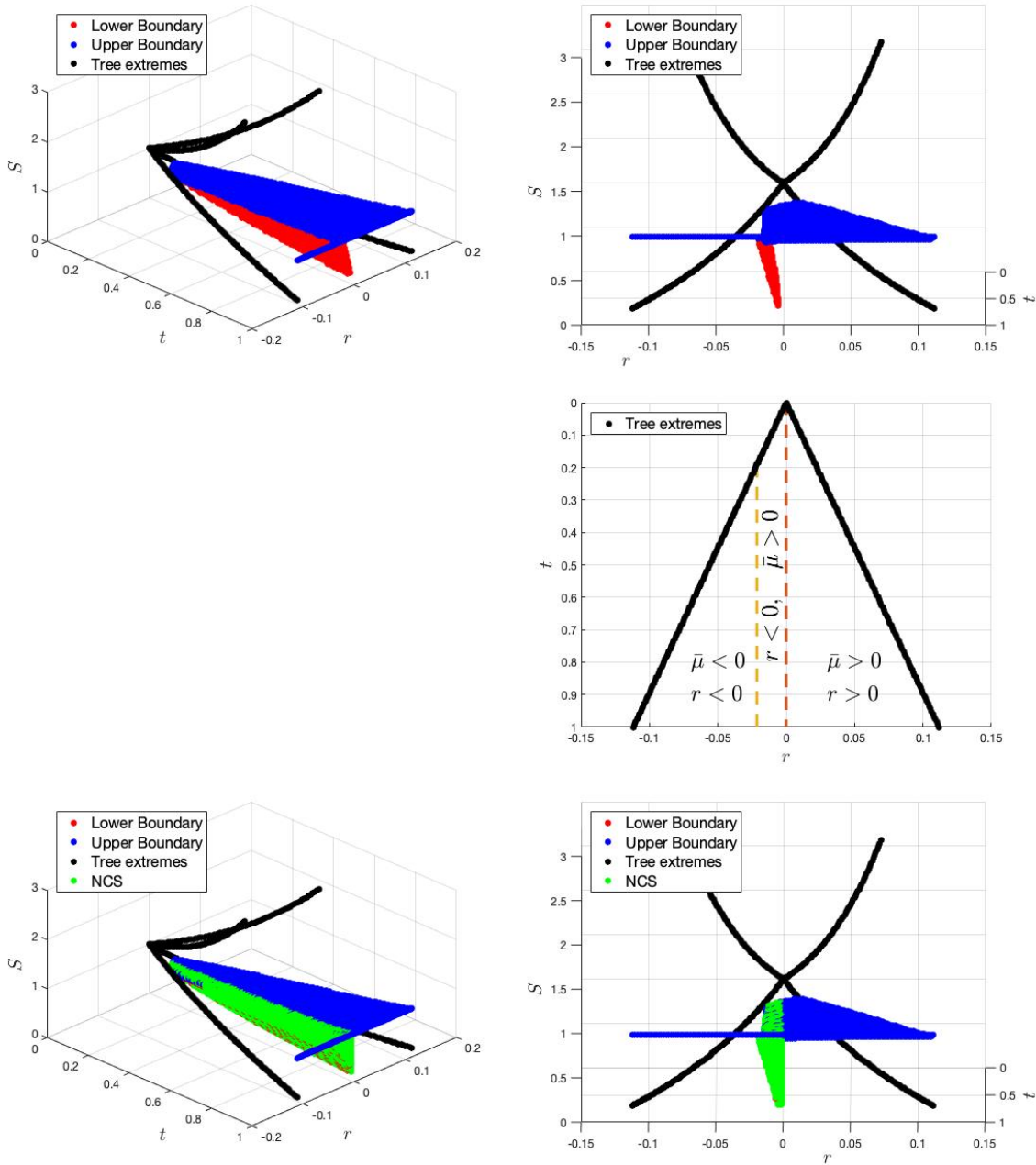


Figure 7: Third example:  $q = -2\%$ . Green points in the bottom panels show the nodes of the quadrinomial tree in which the two necessary conditions of Proposition 4 for a double continuation region hold simultaneously.



gion. This effect is compensated by the preference to postponement due to negative interest rates. The tradeoff results in a non-standard double continuation region that violates the aforementioned downward connectedness of the exercise region. Finally, we quantified the pricing error that an investor would make assuming a constant interest rate and therefore neglecting the variability (and the related risk) of the term structure.

## A Bounds of the probabilities in the Quadrinomial Tree

Recall that at each  $t$  the four probabilities of an upward/downward movement of  $r/Y$  on the tree are:

$$\begin{aligned}
 q_{uu} &= \frac{\mu_Y \mu_r \Delta t + \mu_Y \Delta r^+ + \mu_r \Delta Y^+ + (1 + \rho) \sigma_r \sigma_S}{4 \sigma_r \sigma_S} \\
 q_{ud} &= \frac{-\mu_Y \mu_r \Delta t + \mu_Y \Delta r^+ - \mu_r \Delta Y^+ + (1 - \rho) \sigma_r \sigma_S}{4 \sigma_r \sigma_S} \\
 q_{du} &= \frac{-\mu_Y \mu_r \Delta t - \mu_Y \Delta r^+ + \mu_r \Delta Y^+ + (1 - \rho) \sigma_r \sigma_S}{4 \sigma_r \sigma_S} \\
 q_{dd} &= \frac{\mu_Y \mu_r \Delta t - \mu_Y \Delta r^+ - \mu_r \Delta Y^+ + (1 + \rho) \sigma_r \sigma_S}{4 \sigma_r \sigma_S}
 \end{aligned} \tag{A1}$$

with  $\Delta r^+ = \sigma_r \sqrt{\Delta t}$ ,  $\Delta Y^+ = \sigma_S \sqrt{\Delta t}$ ,  $\mu_Y = r(t) - q - \frac{\sigma_S^2}{2}$  and  $\mu_r = \kappa(\theta - r(t))$ . From now on we light the notation writing  $r$  instead of  $r(t)$ . Nevertheless, it is crucial to remember that these probabilities are different for each node of the quadrinomial tree.

As already pointed out, the four probabilities sum up to one by construction. Unfortunately, they do not necessarily lie in  $(0,1)$ . As a first control, we investigate what happens as the length of the time step goes to zero, namely, as  $\Delta t \rightarrow 0$ . We have

$$\begin{aligned}
 \lim_{\Delta t \rightarrow 0} q_{uu} &= \lim_{\Delta t \rightarrow 0} q_{dd} = \frac{1 + \rho}{4}, \\
 \lim_{\Delta t \rightarrow 0} q_{ud} &= \lim_{\Delta t \rightarrow 0} q_{du} = \frac{1 - \rho}{4}
 \end{aligned}$$

which are all positive quantities (at least as  $\rho \in (-1,1)$ ). Therefore, the problem of having possibly negative probabilities is only due to the discretization procedure.

For instance, with  $n = 250$  steps and  $T = 1$  (that corresponds to  $\Delta t = 0.004$ ), we need to impose the positivity constraint on all the four numerators in (A1).

Imposing  $q_{uu} \geq 0$  and solving with respect to  $r$  leads to:

$$A_{uu} r^2 + B_{uu} r + C_{uu} \leq 0 \tag{A2}$$

where:

$$\begin{aligned}
A_{uu} &= \kappa \\
B_{uu} &= -\kappa \left( \theta + q + \frac{\sigma_S^2}{2} - \frac{\sigma_S}{\sqrt{\Delta t}} \right) - \frac{\sigma_r}{\sqrt{\Delta t}} \\
C_{uu} &= -\kappa \theta \left( -q - \frac{\sigma_S^2}{2} + \frac{\sigma_S}{\sqrt{\Delta t}} \right) - \frac{\sigma_r}{\sqrt{\Delta t}} \left( -q - \frac{\sigma_S^2}{2} \right) - \frac{(1+\rho)\sigma_r\sigma_S}{\Delta t}.
\end{aligned}$$

Provided that the discriminant of equation (A2) is positive, which surely holds true as  $\Delta t \rightarrow 0$ , the solution is  $r_{uu} \leq r \leq \bar{r}_{uu}$ , where, of course,

$$r_{uu} = \frac{-B_{uu} - \sqrt{B_{uu}^2 - 4A_{uu}C_{uu}}}{2A_{uu}} \quad \text{and} \quad \bar{r}_{uu} = \frac{-B_{uu} + \sqrt{B_{uu}^2 - 4A_{uu}C_{uu}}}{2A_{uu}}.$$

Similarly, we can work out all of the other probabilities.

Imposing  $q_{ud} \geq 0$  leads to:

$$A_{ud}r^2 + B_{ud}r + C_{ud} \geq 0$$

where:

$$\begin{aligned}
A_{ud} &= \kappa \\
B_{ud} &= -\kappa \left( \theta + q + \frac{\sigma_S^2}{2} - \frac{\sigma_S}{\sqrt{\Delta t}} \right) + \frac{\sigma_r}{\sqrt{\Delta t}} \\
C_{ud} &= -\kappa \theta \left( -q - \frac{\sigma_S^2}{2} + \frac{\sigma_S}{\sqrt{\Delta t}} \right) - \frac{\sigma_r}{\sqrt{\Delta t}} \left( q + \frac{\sigma_S^2}{2} \right) + \frac{(1-\rho)\sigma_r\sigma_S}{\Delta t},
\end{aligned}$$

that is solved by  $r \leq r_{ud} \cup r \geq \bar{r}_{ud}$ .

Imposing  $q_{du} \geq 0$  leads to:

$$A_{du}r^2 + B_{du}r + C_{du} \geq 0$$

where:

$$\begin{aligned}
A_{du} &= k \\
B_{du} &= -\kappa \left( \theta + q + \frac{\sigma_S^2}{2} + \frac{\sigma_S}{\sqrt{\Delta t}} \right) - \frac{\sigma_r}{\sqrt{\Delta t}} \\
C_{du} &= -\kappa \theta \left( -q - \frac{\sigma_S^2}{2} - \frac{\sigma_S}{\sqrt{\Delta t}} \right) + \frac{\sigma_r}{\sqrt{\Delta t}} \left( q + \frac{\sigma_S^2}{2} \right) + \frac{(1-\rho)\sigma_r\sigma_S}{\Delta t},
\end{aligned}$$

that is solved by  $r \leq \underline{r}_{du} \cup r \geq \bar{r}_{du}$ .

Finally, imposing  $q_{dd} \geq 0$  leads to:

$$A_{dd}r^2 + B_{dd}r + C_{dd} \leq 0$$

where:

$$\begin{aligned} A_{dd} &= \kappa \\ B_{dd} &= -\kappa \left( \theta + q + \frac{\sigma_S^2}{2} + \frac{\sigma_S}{\sqrt{\Delta t}} \right) + \frac{\sigma_r}{\sqrt{\Delta t}} \\ C_{dd} &= -\kappa \theta \left( -q - \frac{\sigma_S^2}{2} - \frac{\sigma_S}{\sqrt{\Delta t}} \right) + \frac{\sigma_r}{\sqrt{\Delta t}} \left( -q - \frac{\sigma_S^2}{2} \right) - \frac{(1 + \rho)\sigma_r\sigma_S}{\Delta t}. \end{aligned}$$

that is solved by  $\underline{r}_{dd} \leq r \leq \bar{r}_{dd}$ .

Summing up, probabilities in (A1) stay positive as long as  $r$  satisfies:

$$\begin{cases} \underline{r}_{uu} \leq r \leq \bar{r}_{uu} \\ r \leq \underline{r}_{ud} \cup r \geq \bar{r}_{ud} \\ r \leq \underline{r}_{du} \cup r \geq \bar{r}_{du} \\ \underline{r}_{dd} \leq r \leq \bar{r}_{dd} \end{cases}$$

The solution to the previous system of inequalities depends on the sign of the correlation  $\rho$ . Given the sign of  $\rho$ , the eight extremes values  $\underline{r}_{uu}, \underline{r}_{ud}, \dots, \bar{r}_{du}, \bar{r}_{dd}$  always satisfy the same chain of inequalities. Furthermore, notice that this eight values depend only on the parameters of the model and not on  $t$ .

When  $\rho \in (0, 1]$ , the only interval on which all of the inequalities hold true is  $\bar{r}_{ud} \leq r \leq \underline{r}_{du}$  as it can be conveniently seen in Figure 8.

The intuition here is that when  $r$  and  $S$  move together and the discretization of  $r$  reaches values far away from its long run mean  $\theta$ , a further movement of  $r$  away from  $\theta$  and in the opposite direction of  $S$  is extremely unlikely and, eventually, happens “with a negative probability”.

If, for example,  $r(0) = 0$ ,  $\theta = 0.02$ ,  $\sigma_r = 0.01$ ,  $\kappa = 0.7$ ,  $S(0) = 1$ ,  $\sigma_S = 0.15$ ,  $q = 0$ ,  $\rho = 0.5$ ,  $T = 1$  and  $n = 125$ , after  $m = 100$  steps, namely at  $t = m \cdot \Delta t = m \cdot \frac{T}{n} = 0.8$ ,  $r(t)$  spans the

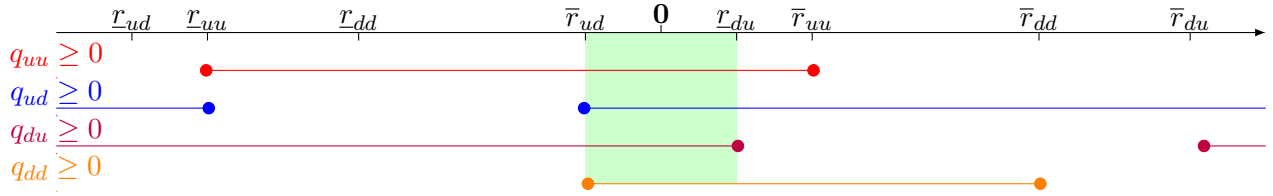


Figure 8: Graphical solution to the system of inequalities when  $\rho \in (0, 1]$ .

interval  $[-0.0885, 0.0885]$  and  $S(t)$  the interval  $[-1.3282, 1.3282]$ , both of them assuming  $m = 101$  different values. Hence, at  $t = 0.8$  there are  $101^2 = 10201$  possible nodes on tree. As an instance, at the node  $(S(t), r(t))_{t=0.8} = (0.5847, -0.0751)$  the four probabilities are:

$$q_{uu} = 0.4885$$

$$q_{ud} = -0.0143$$

$$q_{du} = 0.2780$$

$$q_{dd} = 0.2478.$$

Indeed, with the given parameters, probabilities are all positive as long as  $\bar{r}_{ud} = -0.0660 \leq r(t) \leq 0.0861 = r_{du}$ , which is not our case. As  $r(t)$  is extremely far away from its long-run mean and since  $\rho > 0$  implies that  $r$  and  $S$  are likely to move together in the same direction,  $q_{ud}$ , namely the probability that  $r$  deviates even further from its long-run mean and also against  $S$ , becomes negative. Notice that  $q_{ud} > q_{dd}$ , meaning that the force that drives  $r$  towards its long-run mean prevails on the positive correlation between the two processes. When such a scenario happens, we adjust the probabilities by setting the negative one to 0 and normalizing to 1 the others. From the example above we would then get:

$$q_{uu} = 0.4816$$

$$q_{ud} = 0$$

$$q_{du} = 0.2741$$

$$q_{dd} = 0.2443.$$

A very similar situation happens when  $\rho \in [-1, 0)$  and the four probabilities stay positive as long as  $\underline{r}_{dd} \leq r \leq \bar{r}_{uu}$ . Figure 9 shows the solution to the system of inequalities in this case. Now  $q_{uu}$  or  $q_{dd}$  might become negative. This is due to the negative correlation: as  $r$  and  $S$  are likely to move in the opposite direction, when  $r$  is far away from its long-run mean, moving even further in the same direction of  $S$  may result in a negative probability. Again, we correct for such a phenomenon with the normalization described above.

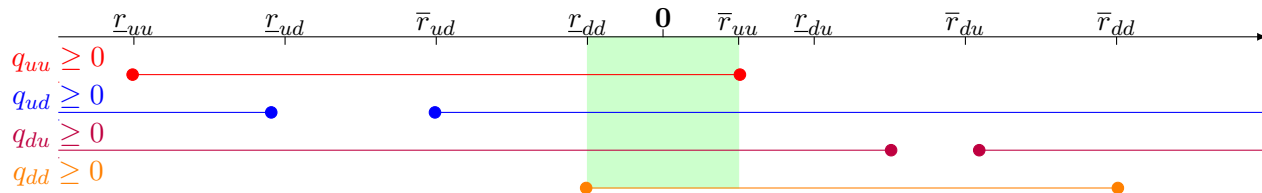


Figure 9: Graphical solution to the system of inequalities when  $\rho \in [-1, 0)$ .

For sake of completeness, we briefly discuss also the limit of zero correlation between  $r$  and  $S$ . When  $\rho = 0$ ,  $\underline{r}_{uu} = \underline{r}_{ud}$ ,  $\bar{r}_{ud} = \underline{r}_{dd}$ ,  $\bar{r}_{uu} = \underline{r}_{du}$  and  $\bar{r}_{du} = \bar{r}_{dd}$ . Hence, the two intervals we found for the two previous cases,  $\bar{r}_{ud} \leq r \leq \underline{r}_{du}$  when  $\rho \in (0, -1]$  and  $\underline{r}_{dd} \leq r \leq \bar{r}_{uu}$  when  $\rho \in [-1, 0)$ , coincide. When  $\rho = 0$ , probabilities stay positive as long as  $r$  belong to that interval.

Since the support of the discretization of  $r(t)$  is known at each  $t$ , we can retrieve the maximum  $t$  before which no normalization of the probabilities is needed.

Given the two thresholds  $\underline{r}$  and  $\bar{r}$  (where  $\underline{r} = \bar{r}_{ud}$ ,  $\bar{r} = \underline{r}_{du}$  if  $\rho > 0$  and  $\underline{r} = \underline{r}_{dd}$ ,  $\bar{r} = \bar{r}_{uu}$  if  $\rho < 0$ ) we can set  $\underline{t}$  and  $\bar{t}$  as:

$$\underline{t} := \min_{s \in \{0, \Delta t, 2\Delta t, \dots, T\}} \{r(s) \geq \underline{r}\} \quad \text{and} \quad \bar{t} := \max_{s \in \{0, \Delta t, 2\Delta t, \dots, T\}} \{r(s) \leq \bar{r}\}.$$

Given the binomial structure of the discretization, after  $m$  steps we have we have:

$$r(0) - m\Delta r^- = r(0) - m\sigma_r \Delta t \leq r(t) \leq r(0) + m\sigma_r \Delta t = r(0) + m\Delta r^+$$

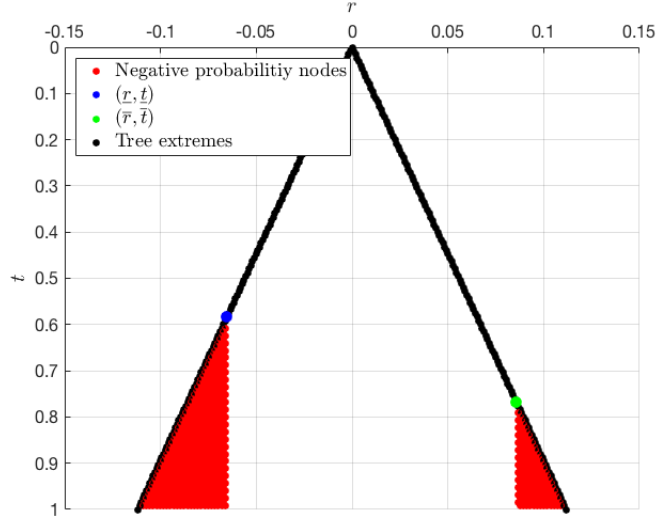


Figure 10: Section of the quadrinomial tree for  $S = 0$ . Red points indicate nodes at which one transition probability becomes negative. Parameters:  $r(0) = 0$ ,  $\theta = 0.02$ ,  $\sigma_r = 0.01$ ,  $\kappa = 0.7$ ,  $S(0) = 1$ ,  $\sigma_S = 0.15$ ,  $q = 0$ ,  $\rho = 0.5$ ,  $T = 1$ ,  $n = 125$ .

and, therefore, from

$$\begin{aligned} r(0) - \underline{m}\sigma_r\Delta t &\geq \underline{r} \\ r(0) + \overline{m}\sigma_r\Delta t &\leq \overline{r} \end{aligned} \tag{A3}$$

we can explicitly compute:

$$\begin{aligned} \underline{t} = \underline{m}\Delta t &= \frac{r(0) - \underline{r}}{\sigma_r\sqrt{\Delta t}}\Delta t = \frac{r(0) - \underline{r}}{\sigma_r}\sqrt{\Delta t} \\ \overline{t} = \overline{m}\Delta t &= \frac{\overline{r} - r(0)}{\sigma_r\sqrt{\Delta t}}\Delta t = \frac{\overline{r} - r(0)}{\sigma_r}\sqrt{\Delta t}. \end{aligned}$$

Of course, neither  $\underline{r}$ ,  $\overline{r}$  nor  $\underline{t}$ ,  $\overline{t}$  are likely to correspond to any node of the discretized process  $r(t)$  or to the discretized time line  $\{0, \Delta t, 2\Delta t, \dots, T\}$ . In this case, we set  $\underline{r}$ ,  $\overline{r}$  and  $\underline{t}$ ,  $\overline{t}$  equal to the smallest values on the grid of  $r(t)$  and  $t$  that satisfy the constraints in (A3). Going back to the numerical example above, we have that  $\underline{t} = 0.5840$  and  $\overline{t} = 0.7680$ . A section of the quadrinomial tree in this case is displayed in Figure 10.

## B Proofs of the Claims

**Proof of Proposition (1): value of the European put option.** We derive the price of the European put option at  $t = 0$ . As the payoff of the derivative depends only on the final value of  $S$ , the price at any  $t$  is obtained straightforwardly thanks to the Markovianity of  $(S, r)$  by replacing  $S_0$  by  $S(t)$ ,  $r_0$  by  $r(t)$  and  $T$  by  $T - t$ .

In the market described by (1), the risk-neutral price at  $t = 0$  of the European put option on  $S$  with strike price  $K$  and maturity  $T$  is given by:

$$\begin{aligned} \pi_{EuropeanPut}(0) &= \mathbb{E}^{\mathbb{Q}}[e^{-\int_0^T r(s)ds}(K - S(T))^+] \\ &= \mathbb{E}^{\mathbb{Q}}\left[\frac{(K - S(T))}{B(T)}\mathbf{1}_{\{K - S(T) > 0\}}\right] \\ &= K\mathbb{E}^{\mathbb{Q}}\left[\frac{\mathbf{1}_{\{K - S(T) > 0\}}}{B(T)}\right] - \mathbb{E}^{\mathbb{Q}}\left[\frac{S(T)\mathbf{1}_{\{K - S(T) > 0\}}}{B(T)}\right]. \end{aligned} \quad (\text{B1})$$

Since  $B(T)$  depends on  $r$  which is correlated with  $S$ , in order to compute the two expected values we would need to know their joint distribution under  $\mathbb{Q}$  and then evaluate a double integral. This turns out to be rather complicated. Nevertheless, we can greatly simplify the computation of the two expected values applying a change of numéraire.

We start from the first expectation in (B1). Consider the  $T$ -forward measure  $\mathbb{Q}^T$ , namely the martingale measure for the numéraire process  $p(t, T)$ . The Radon-Nikodym derivative of  $\mathbb{Q}^T$  with respect to  $\mathbb{Q}$  (whose numéraire process is the money market  $B(t) = e^{\int_0^t r(s)ds}$ ) is:

$$\frac{d\mathbb{Q}^T}{d\mathbb{Q}} = L^T(t) = \frac{p(t, T)}{B(t)p(0, T)} \text{ on } \mathcal{F}_t, 0 \leq t \leq T.$$



As  $p(T, T) = 1$  and since  $p(0, T)$  is a scalar, we get:

$$\begin{aligned}
\mathbb{E}^{\mathbb{Q}} \left[ \frac{\mathbf{1}_{\{K-S(T)>0\}}}{B(T)} \right] &= p(0, T) \mathbb{E}^{\mathbb{Q}} \left[ \frac{p(T, T)}{B(T)p(0, T)} \mathbf{1}_{\{K-S(T)>0\}} \right] \\
&= p(0, T) \mathbb{E}^{\mathbb{Q}} [L^T(T) \mathbf{1}_{\{K-S(T)>0\}}] \\
&= p(0, T) \mathbb{E}^{\mathbb{Q}^T} [\mathbf{1}_{\{K-S(T)>0\}}] \\
&= p(0, T) \mathbb{Q}^T (S(T) < K) \\
&= p(0, T) \mathbb{Q}^T \left( \frac{S(T)}{p(T, T)} < K \right).
\end{aligned}$$

By definition, under the T-forward measure  $\mathbb{Q}^T$  the discounted process of the risky asset, when accounting for the dividend,  $Z_{0,T}(t) := \frac{S(t)e^{qt}}{p(t, T)}$  is a martingale. Applying the multidimensional Itô's Lemma to  $Z_{0,T}(t)$  under  $\mathbb{Q}$  we get:

$$dZ_{0,T}(t) = (\dots) dt + (\nu_S + B(t, T)\nu_r) \cdot dW^{\mathbb{Q}}(t)$$

where,  $\nu_S = [\sigma_S \ 0]$ ,  $\nu_r = [\sigma_r \rho \ \sigma_r \sqrt{1 - \rho^2}]$  and  $W^{\mathbb{Q}}(t) = [W_1^{\mathbb{Q}}(t) \ W_2^{\mathbb{Q}}(t)]'$  is standard two-dimensional Brownian motion under  $\mathbb{Q}$ . Since the volatility process  $\sigma_{0,T}(t) := \nu_S + B(t, T)\nu_r$  is constant, we can apply a suitable change of measure from  $\mathbb{Q}$  to  $\mathbb{Q}^T$  to get rid of the deterministic drift of  $Z_{0,T}(t)$ . Under the T-forward measure we get:

$$dZ_{0,T}(t) = +\sigma_{0,T}(t) \cdot dW^{\mathbb{Q}^T}(t)$$

as we expected. The process  $Z_{0,T}(t)$  is now a geometric Brownian motion driven by a bi-dimensional Wiener process. Hence, its solution is:

$$\begin{aligned}
Z_{0,T}(t) &= Z_{0,T}(0) \exp \left\{ -\frac{1}{2} \int_0^t \|\sigma_{0,T}^2(s)\| ds + \int_0^t \sigma_{0,T}(s) \cdot dW^{\mathbb{Q}^T}(s) \right\} \\
&= \frac{S(0)}{p(0, T)} \exp \left\{ -\frac{1}{2} \int_0^t \|\sigma_{0,T}^2(s)\| ds + \int_0^t \sigma_{0,T}(s) \cdot dW^{\mathbb{Q}^T}(s) \right\}.
\end{aligned}$$

Notice that, due to Itô's Isometry,

$$-\frac{1}{2} \int_0^t \|\sigma_{0,T}^2(s)\| ds + \int_0^t \sigma_{0,T}(s) \cdot dW^{\mathbb{Q}^T}(s) \sim \mathcal{N} \left( -\frac{1}{2} \Sigma_{0,T}^2(t), \Sigma_{0,T}^2(t) \right)$$

where:

$$\begin{aligned}
\Sigma_{0,T}^2(t) &:= \int_0^t \|\sigma_{0,T}^2(s)\| ds \\
&= \int_0^t \sigma_S^2 + 2\sigma_S\sigma_r\rho B(s,T) + B(s,T)^2\sigma_r^2 ds \\
&= \sigma_S^2 t + 2\sigma_S\sigma_r\rho \left( \frac{\kappa t - e^{-\kappa T}(-1 + e^{\kappa t})}{k^2} \right) + \sigma_r^2 \left( \frac{e^{-2\kappa T}(-1 + e^{2\kappa t} + 4e^{\kappa T} - 4\kappa(t+T) + 2e^{2\kappa T}\kappa t)}{2k^3} \right).
\end{aligned}$$

When  $t = T$ , the expression above simplifies to:

$$\Sigma_{0,T}^2(T) = \sigma_S^2 T + 2\sigma_S\sigma_r\rho \left( \frac{-1 + e^{-\kappa T} + \kappa T}{k^2} \right) + \sigma_r^2 \left( -\frac{3 + e^{-2\kappa T} - 4e^{-\kappa T} - 2\kappa T}{2k^3} \right).$$

Finally, we can compute the T-forward probability that the put option closes in the money

as:

$$\begin{aligned}
\mathbb{Q}^T \left( \frac{S(T)}{p(T,T)} < K \right) &= \mathbb{Q}^T \left( \frac{S(T)e^{qT}}{p(T,T)} < Ke^{qT} \right) \\
&= \mathbb{Q}^T (Z_{0,T}(T) < Ke^{qT}) \\
&= \mathbb{Q}^T \left( \frac{S(0)}{p(0,T)} \exp \left\{ -\frac{1}{2} \int_0^T \|\sigma_{0,T}\|^2(s) ds + \int_0^T \sigma_{0,T}(s) \cdot dW^{\mathbb{Q}^T}(s) \right\} < Ke^{qT} \right) \\
&= \mathbb{Q}^T \left( \mathcal{N} \left( -\frac{1}{2} \Sigma_{0,T}^2(T), \Sigma_{0,T}^2(T) \right) < \ln \frac{p(0,T)Ke^{qT}}{S(0)} \right) \\
&= N(-\tilde{d}_2)
\end{aligned}$$

with  $\tilde{d}_2 = \frac{1}{\sqrt{\Sigma_{0,T}^2(T)}} \left( \ln \frac{S(0)}{p(0,T)K} - \frac{1}{2} \Sigma_{0,T}^2(T) - qT \right)$ . Hence,

$$\mathbb{E}^{\mathbb{Q}} \left[ \frac{\mathbf{1}_{\{K-S(T)>0\}}}{B(T)} \right] = p(0,T)N(-\tilde{d}_2).$$

We now turn to the second expected value in (B1). Consider the martingale measure  $\mathbb{Q}^S$  with numéraire process  $S(t)e^{qt}$ . The Radon-Nikodym derivative of  $\mathbb{Q}^S$  with respect to  $\mathbb{Q}$  is:

$$\frac{d\mathbb{Q}^S}{d\mathbb{Q}} = L^S(t) = \frac{S(t)e^{qt}}{S(0)B(t)} \text{ on } \mathcal{F}_t, 0 \leq t \leq T.$$

As both  $S(0)$  and  $e^{qT}$  are scalars, we have:

$$\begin{aligned}
\mathbb{E}^{\mathbb{Q}} \left[ \frac{S(T) \mathbf{1}_{\{K-S(T)>0\}}}{B(T)} \right] &= S(0) e^{-qT} \mathbb{E}^{\mathbb{Q}} \left[ \frac{S(T) e^{qT}}{S(0) B(T)} \mathbf{1}_{\{K-S(T)>0\}} \right] \\
&= S(0) e^{-qT} \mathbb{E}^{\mathbb{Q}} [L^S(T) \mathbf{1}_{\{K-S(T)>0\}}] \\
&= S(0) e^{-qT} \mathbb{E}^{\mathbb{Q}^S} [\mathbf{1}_{\{K-S(T)>0\}}] \\
&= S(0) e^{-qT} \mathbb{Q}^S(S(T) < K).
\end{aligned}$$

Under  $\mathbb{Q}^S$ , the process  $Y_{0,T}(t) := \frac{p(0,t)}{S(t)e^{qt}}$  is a martingale. Notice that  $Y_{0,T}(t) = Z_{0,T}(t)^{-1}$ .

Then, Itô's Lemma tells us immediately that:

$$dY_{0,T}(t) = (\dots) dt - (\nu_S + B(t,T)\nu_r) \cdot dW^{\mathbb{Q}}(t)$$

and after a suitable change of measure we get that under  $\mathbb{Q}^S$ :

$$dY_{0,T}(t) = -\sigma_{0,T}(t) \cdot dW^{\mathbb{Q}^S}(t).$$

As before, we get:

$$\begin{aligned}
Y_{0,T}(t) &= Y_{0,T}(0) \exp \left\{ -\frac{1}{2} \int_0^t \|\sigma_{0,T}^2(s)\| ds - \int_0^t \sigma_{0,T}(s) \cdot dW^{\mathbb{Q}^2}(s) \right\} \\
&= \frac{p(0,T)}{S(0)} \exp \left\{ -\frac{1}{2} \int_0^t \|\sigma_{0,T}^2(s)\| ds - \int_0^t \sigma_{0,T}(s) \cdot dW^{\mathbb{Q}^S}(s) \right\}.
\end{aligned}$$

and, since again  $p(T,T) = 1$ , the  $\mathbb{Q}^S$  probability that the put option closes in the money is:

$$\begin{aligned}
\mathbb{Q}^S(S(T) < K) &= \mathbb{Q}^S \left( \frac{1}{S(T)} > \frac{1}{K} \right) \\
&= \mathbb{Q}^S \left( \frac{p(T,T)}{S(T)e^{qT}} > \frac{1}{Ke^{qT}} \right) \\
&= \mathbb{Q}^S \left( Y_{0,T}(T) > \frac{1}{Ke^{qT}} \right) \\
&= \mathbb{Q}^T \left( \frac{p(0,T)}{S(0)} \exp \left\{ -\frac{1}{2} \int_0^T \|\sigma_{0,T}^2(s)\| ds - \int_0^T \sigma_{0,T}(s) \cdot dW^{\mathbb{Q}^T}(s) \right\} > \frac{1}{Ke^{qT}} \right) \\
&= \mathbb{Q}^T \left( \mathcal{N} \left( -\frac{1}{2} \Sigma_{0,T}^2(T), \Sigma_{0,T}^2(T) \right) > \ln \frac{S(0)}{p(0,T)Ke^{qT}} \right) \\
&= \mathbb{Q}^T \left( \mathcal{N}(0,1) > \frac{1}{\sqrt{\Sigma_{0,T}^2(T)}} \left( \ln \frac{S(0)}{p(0,T)K} + \frac{1}{2} \Sigma_{0,T}^2(T) - qT \right) \right) \\
&= N(-\tilde{d}_1)
\end{aligned}$$

where  $\tilde{d}_1 = \frac{1}{\sqrt{\Sigma_{0,T}^2(T)}} \left( \ln \frac{S(0)}{p(0,T)K} + \frac{1}{2}\Sigma_{0,T}^2(T) - qT \right) = \tilde{d}_2 + \sqrt{\Sigma_{0,T}^2(T)}$ . Hence,

$$\mathbb{E}^{\mathbb{Q}} \left[ \frac{S(T)\mathbf{1}_{\{K-S(T)>0\}}}{B(T)} \right] = S(0)e^{-qT}N(-\tilde{d}_1).$$

Finally, putting everything together we find the initial price of the put option:

$$\begin{aligned} \pi_{EuropeanPut}(0) &= K\mathbb{E}^{\mathbb{Q}}[p(0,T)\mathbf{1}_{\{K-S(T)>0\}}] - \mathbb{E}^{\mathbb{Q}}[p(0,T)S(T)\mathbf{1}_{\{K-S(T)>0\}}] \\ &= Kp(0,T)N(-\tilde{d}_2) - S(0)e^{-qT}N(-\tilde{d}_1). \end{aligned}$$

**Proof of Proposition (2): convergence of the quadrinomial tree.** We now need to show that the bivariate discrete process  $(X_i)_i$  defined in (10) with the parameters in (11) and (12) converges in distribution to  $X(t) = (Y(t), r(t))$  that solves (7). With the notation of the general case in (9) and exploiting the result of Section 11.3 of Stroock and Varadhan (1997), the desired result holds true if the following four conditions are met:

(A1) the functions  $\mu(x, t)$  and  $\sigma(x, t)$  are continuous and  $\sigma(x, t)$  is non negative;

(A2) with probability 1 a solution  $(X_t)_t$  to the SDE:

$$X_t = X_0 + \int_0^t \mu(X_s, s)ds + \int_0^t \sigma(X_s, s) \cdot dW(s)$$

exists for  $0 < t < +\infty$  and it is unique in law;

(A3) for all  $\delta, T > 0$

$$\begin{aligned} \lim_{n \rightarrow +\infty} \sup_{\|x\| \leq \delta, 0 \leq t \leq T} |\Delta Y^\pm| &= 0 \\ \lim_{n \rightarrow +\infty} \sup_{\|x\| \leq \delta, 0 \leq t \leq T} |\Delta r^\pm| &= 0; \end{aligned}$$

(A4) let  $X_{i,j}$  indicate the  $j$ -th entry of  $X_i$  and let  $\mathcal{F}_i = \sigma(X_1, \dots, X_i)$  be the filtration generated by the discrete bivariate process  $(X_i)$ . Define:

$$\mu_i(x, t) := \begin{bmatrix} \mu_{i,1}(x, t) \\ \mu_{i,2}(x, t) \end{bmatrix} \text{ and } \sigma_i^2(x, t) := \begin{bmatrix} \sigma_{i,1}^2(x, t) \\ \sigma_{i,2}^2(x, t) \end{bmatrix}$$

where  $\mu_{i,j}(x, t) = \frac{\mathbb{E}^{\mathbb{Q}}[X_{i+1,j} - X_{i,j} | \mathcal{F}_i]}{\frac{T}{n}}$  and  $\sigma_{i,j}^2(x, t) = \frac{\mathbb{E}^{\mathbb{Q}}[(X_{i+1,j} - X_{i,j})^2 | \mathcal{F}_i]}{\frac{T}{n}}$  for  $j = 1, 2$ . Let  $\rho_i(x, t) = \frac{\mathbb{E}^{\mathbb{Q}}[(X_{i+1,1} - X_{i,1})(X_{i+1,2} - X_{i,2}) | \mathcal{F}_i]}{\frac{T}{n}}$  and  $\rho(x, t) = \sigma_1(x, t) \cdot \sigma_2(x, t)'$  where  $\sigma_j(x, t)$  is the  $j$ -th row of  $\sigma(x, t)$ . Then, for all  $\delta, T > 0$ ,

$$\lim_{n \rightarrow +\infty} \sup_{\|x\| \leq \delta, 0 \leq t \leq T} \|\mu_i(x, t) - \mu(x, t)\| = 0$$

$$\lim_{n \rightarrow +\infty} \sup_{\|x\| \leq \delta, 0 \leq t \leq T} \|\sigma_i^2(x, t) - \sigma^2(x, t) \cdot \mathbf{I}_2\| = 0$$

$$\lim_{n \rightarrow +\infty} \sup_{\|x\| \leq \delta, 0 \leq t \leq T} |\rho_i(x, t) - \rho(x, t)| = 0$$

where  $\mathbf{I}_n$  is the column vector with all of the  $n$  entries equal to one.

For our quadrinomial tree we have  $X_t = [Y(t), \quad r(t)]'$ ,

$$\mu(X_t, t) = \begin{bmatrix} \left( r(t) - q - \frac{\sigma_S^2}{2} \right) \\ \kappa(\theta - r(t)) \end{bmatrix} \text{ and } \sigma(X_t, t) = \begin{bmatrix} \sigma_S & 0 \\ \sigma_r \rho & \sigma_r \sqrt{1 - \rho^2} \end{bmatrix}.$$

Assumption (A1) trivially holds true.

Assumption (A2) holds true if the standard conditions for the existence and the uniqueness of the solution to an SDE are met. According, e.g., to Proposition 5.1 in Björk (2009), it is sufficient to show that there exists a constant  $K$  such that the following are satisfied for all  $x_i = [y_i, \quad r_i]'$ ,  $i = 1, 2$  and  $t$ :

$$\|\mu(x_1, t) - \mu(x_2, t)\| \leq K \|x_1 - x_2\|,$$

$$\|\sigma(x_1, t) - \sigma(x_2, t)\| \leq K \|x_1 - x_2\|,$$

$$\|\mu(x_1, t)\| + \|\sigma(x_1, t)\| \leq K (1 + \|x_1\|).$$

Notice that the second and the third conditions involve the operator norm of a matrix  $A \in \mathbb{R}^n$  defined as  $\|A\| := \sup_{\|x\|=1} \{ \|A \cdot x\| : x \in \mathbb{R}^n \}$ .

As  $\|\mu(x_1, t) - \mu(x_2, t)\| = \sqrt{1 + \kappa^2} |r_1 - r_2|$  and  $(r_1 - r_2)^2 \leq \|x_1 - x_2\|^2$ , the first condition is surely satisfied for any  $K \geq \sqrt{1 + \kappa^2}$ . As  $\sigma(x_i, t)$  is actually constant and independent of  $x_i$  and  $t$ ,  $\|\sigma(x_1, t) - \sigma(x_2, t)\| = 0$  and the second condition is surely satisfied for any  $K \geq 0$ . Finally, as

$$\|\sigma(x_1, t)\| = \sigma_S^2 + \rho^2 \frac{\sigma_r^2}{2} + |\rho| \frac{\sigma_r}{2} \sqrt{4\sigma_S^2 + \sigma_r^2}$$

is constant and as

$$\|\mu(x_1, t)\| = \sqrt{\left(r_1 - q - \frac{\sigma_S^2}{2}\right)^2 + \kappa^2(\theta - r_1)^2}$$

can be bounded from above by  $\sqrt{2(1 + \kappa^2)r_1^2}$ , we have

$$\|\mu(x_1, t)\| + \|\sigma(x_1, t)\| \leq \sqrt{2(1 + \kappa^2)} \|x_1\| + \|\sigma(x_1, t)\| \leq K(1 + \|x_1\|)$$

for any  $K \geq \max\{\sqrt{2(1 + \kappa^2)}, \|\sigma(x_1, t)\|\}$ . As the three conditions hold true simultaneously for any  $K \geq \max\{\sqrt{2(1 + \kappa^2)}, \|\sigma(x_1, t)\|\}$ , assumption (A2) is satisfied.

As the increments of the bivariate discrete process  $\Delta Y^\pm = \pm \sigma_S \sqrt{\Delta t} = \pm \sigma_S \sqrt{\frac{T}{n}}$ ,  $\Delta r^\pm = \pm \sigma_r \sqrt{\Delta t} = \pm \sigma_r \sqrt{\frac{T}{n}}$  are constant and do not depend neither on  $x_i$ ,  $i = 1, 2$ , nor on  $t$ ,

$$\sup_{\|x\| \leq \delta, 0 \leq t \leq T} |\Delta Y^\pm| = |\Delta Y^\pm| = \sigma_S \sqrt{\frac{T}{n}},$$

$$\sup_{\|x\| \leq \delta, 0 \leq t \leq T} |\Delta r^\pm| = |\Delta r^\pm| = \sigma_r \sqrt{\frac{T}{n}}.$$

As both of the sups are infinitesimal with respect to  $n$ , (A3) holds true as well.

As the parameters in (11) and (12) of the bivariate discretization  $X_i = (Y_i, r_i)$  are chosen in order to match the first two moments and the cross-variation of  $X(t) = (Y(t), r(t))$ , we have  $\mu_i(x, t) = \mu(x, t)$ ,  $\sigma_i^2(x, t) = \sigma^2(x, t) \cdot \mathbf{I}_2$  and  $\rho_i(x, t) = \rho(x, t)$ . Hence, assumption (A4) is satisfied by construction.

Theorem 11.3.3 of Stroock and Varadhan (1997) allows us to conclude.

**Proof of Proposition (3): value of the American put option as a deterministic function.** Let  $\eta := \tau - t$ . Then we can rewrite the value of the American put option (13) as:

$$\tilde{V}(t) = \text{ess sup}_{0 \leq \eta \leq T-t} \mathbb{E}^{\mathbb{Q}} \left[ e^{-\int_t^{t+\eta} r(s) ds} (K - S(t+\eta))^+ \mid \mathcal{F}_t \right] \quad (\text{B2})$$

Recalling the dynamics of the risky stock conditional on  $S(t) = S$  we can further rewrite  $\tilde{V}(t)$  as:

$$\begin{aligned} \tilde{V}(t) = \text{ess sup}_{0 \leq \eta \leq T-t} \mathbb{E}^{\mathbb{Q}} & \left[ \exp \left( - \int_t^{t+\eta} r(s) ds \right) \cdot \right. \\ & \left. \cdot \left( K - S \exp \left( \int_t^{t+\eta} r(s) ds - \left( q + \frac{1}{2} \sigma_S^2 \right) \eta + \sigma_S (W_S(t+\eta) - W_S(t)) \right) \right)^+ \right] \Bigg| \mathcal{F}_t \end{aligned}$$

Therefore,  $\tilde{V}$  depends on the expectation of two random variables: the Brownian increment  $W_S(t+\eta) - W_S(t)$  and the integral  $\int_t^{t+\eta} r(s) ds$ , which appears both in the drift part of the underlying and in the discount factor. The first of the two random variables is  $\mathcal{F}_t$ -independent by definition:

$$W_S(t+\eta) - W_S(t) \perp \mathcal{F}_t,$$

and, moreover,

$$W_S(t+\eta) - W_S(t) \stackrel{\mathbb{Q}}{\sim} W_S(\eta).$$

We now show that also  $\int_t^{t+\eta} r(s) ds$  is independent of  $\mathcal{F}_t$  and that  $\int_t^{t+\eta} r(s) ds \stackrel{\mathbb{Q}}{\sim} \int_0^\eta r(s) ds$  as well. Recalling the solution to the SDE driving the short term interest rate conditional on  $r(t) = r$  we have:

$$\begin{aligned} \int_t^{t+\eta} r(s) ds &= \int_t^{t+\eta} \left[ r e^{-\kappa(s-t)} + \theta(1 - e^{-\kappa(s-t)}) + \sigma_r \int_t^s e^{-\kappa(s-y)} dW_r(y) \right] ds \\ &= -\frac{e^{-\kappa\eta}}{\kappa} (r - \theta) + \frac{r - \theta}{\kappa} + \theta\eta + \sigma_r \int_t^{t+\eta} \int_t^s e^{-\kappa(s-y)} dW_r(y) ds. \end{aligned}$$

The constant  $\alpha := -\frac{e^{-\kappa\eta}}{\kappa} (r - \theta) + \frac{r - \theta}{\kappa} + \theta\eta$  does not depend on  $t$ . Exploiting the definition of stochastic integral with  $\{t_i\}_{i=1, \dots, N}$  such that  $t_0 = t$ ,  $t_N = t + \eta$  and  $\|\{t_i\}\| \rightarrow 0$ , we get:

$$\int_t^{t+\eta} r(s) ds = \alpha + \sigma_r \int_t^{t+\eta} \sum_{i=0}^{N-1} e^{-\kappa(s-t_i)} (W_r(t_{i+1}) - W_r(t_i)) ds.$$

Since  $t_{i+1} > t_i > t$  for any  $i = 1, \dots, N-1$ ,  $W_r(t_{i+1}) - W_r(t_i) \perp \mathcal{F}_t$  by definition and for any value of  $s$ . Hence,

$$\sum_{i=0}^{N-1} e^{-\kappa(s-t_i)} (W_r(t_{i+1}) - W_r(t_i)) \perp \mathcal{F}_t \quad \forall s.$$

Since the sum is independent of  $\mathcal{F}_t$  for any  $s$ , the outer integral in  $ds$  preserves such independence. As a result,

$$\int_t^{t+\eta} r(s) ds \perp \mathcal{F}_t.$$

Furthermore, we need to show that the distribution of

$$\int_t^{t+\eta} r(s) ds$$

does not depend on  $t$ . Recalling that:

$$\int_t^{t+\eta} r(s) ds = \alpha + \sigma_r \int_t^{t+\eta} \int_t^s e^{-\kappa(s-y)} dW_r(y) ds$$

and setting  $a := s - t$  in the outer integral in  $ds$ , we get:

$$\int_t^{t+\eta} r(s) ds = \alpha + \sigma_r \int_0^\eta \int_t^{a+t} e^{-\kappa(a+t-y)} dW_r(y) da.$$

The argument of the inner stochastic integral is deterministic in  $y$  and, therefore:

$$\begin{aligned} \int_t^{a+t} e^{-\kappa(s-y)} dW_r(y) &\stackrel{\mathbb{Q}}{\sim} \mathcal{N} \left( 0, \int_t^{a+t} e^{-2\kappa(s-y)} dy \right) \\ &\stackrel{\mathbb{Q}}{\sim} \mathcal{N} \left( 0, \frac{1}{2\kappa} (1 - e^{-2\kappa a}) \right), \end{aligned}$$

which does not depend on  $t$ . Thanks to a little abuse of notation we see that:

$$\int_t^{t+\eta} r(s) ds = \alpha + \sigma_r \int_0^\eta \mathcal{N} \left( 0, \frac{1}{2\kappa} (1 - e^{-2\kappa a}) \right) da,$$

where the right-hand side of the equation does not depend on  $t$ . Hence:

$$\int_t^{t+\eta} r(s) ds \stackrel{\mathbb{Q}}{\sim} \int_0^\eta r(s) ds.$$



We now go back to the original expression (B2). Thanks to the independence of  $\mathcal{F}_t$ , the conditional expected value turns into the unconditional one:

$$\mathbb{E}^{\mathbb{Q}} \left[ e^{-\int_t^{t+\eta} r(s) ds} (K - S(t + \eta))^+ \middle| \mathcal{F}_t \right] = \mathbb{E}^{\mathbb{Q}} \left[ e^{-\int_t^{t+\eta} r(s) ds} (K - S(t + \eta))^+ \right],$$

and setting  $S(0) = S(t) = S$ ,  $r(0) = r(t) = r$ ,

$$\mathbb{E}^{\mathbb{Q}} \left[ e^{-\int_t^{t+\eta} r(s) ds} (K - S(t + \eta))^+ \right] = \mathbb{E}^{\mathbb{Q}} \left[ e^{-\int_0^\eta r(s) ds} (K - S(\eta))^+ \right].$$

Therefore, the value on an American option on  $S$  defined in (13) reduces to

$$\tilde{V} = F(t, S(t), r(t))$$

with

$$F(t, S, r) = \sup_{0 \leq \eta \leq T-t} \mathbb{E}^{\mathbb{Q}} \left[ \exp \left( -\int_0^\eta r(s) ds \right) \cdot \left( K - S \exp \left( \int_0^\eta r(s) ds - \left( q + \frac{1}{2} \sigma_S^2 \right) \eta + \sigma_S W_S(\eta) \right) \right)^+ \right].$$

where  $t$  enters only the upper bound of  $\eta$ , namely the time to maturity of the option.

**Proof of Proposition (4): asymptotic necessary conditions for the existence of a double continuation region.** When the interest rate is constant, the value at  $t$  of the American put option is  $G(t, S(t))$ , where  $G$  is a deterministic function defined in (15). As Battauz et al. (2015) show in Section 2, necessary conditions for the double continuation region to appear at a generic  $t$  are that the drift of  $S$  is positive and  $G(t, 0) > K$ .

According to Proposition (3), the value of the American put on  $S$  in the market described by (1) is of the form  $\tilde{V}(t) = F(t, S(t), r(t))$  where  $F$  is a deterministic function implicitly defined by (14). Following the same logic of Battauz et al. (2015), we need to impose analogous conditions on the drift of  $S$  and on  $F(t, S(t), r(t))$ .

The counterpart of  $G(t, 0) > K$  in Battauz et al. (2015), is here  $F(t, 0, r) > K$  that delivers

[NC0].

According to (14), we have:

$$F(t, 0, r) = K \cdot \sup_{0 \leq \eta \leq T-t} \mathbb{E}^{\mathbb{Q}} \left[ \exp \left( - \int_0^{\eta} r(s) ds \right) \right].$$

Therefore, we need to check under which condition the supremum on the right hand side of the last equation is strictly greater than 1. Namely, we have to derive conditions under which there exists a  $\eta \in [0, T-t]$  such that the supremum is strictly greater than 1. Exploiting Jensen's inequality and the uniform integrability of  $r(s)$ , we get:

$$\mathbb{E}^{\mathbb{Q}} \left[ \exp \left( - \int_0^{\eta} r(s) ds \right) \right] \geq \exp \left( - \mathbb{E}^{\mathbb{Q}} \left[ \int_0^{\eta} r(s) ds \right] \right) = \exp \left( - \int_0^{\eta} \mathbb{E}^{\mathbb{Q}} [r(s)] ds \right).$$

As before, thanks to (3), we have:

$$\mathbb{E}^{\mathbb{Q}} \left[ \exp \left( - \int_0^{\eta} r(s) ds \right) \right] \geq \exp \left( - \int_0^{\eta} r e^{-\kappa s} + \theta(1 - e^{-\kappa s}) ds \right) = \exp(r\alpha - \theta(\alpha + \eta))$$

where we set  $\alpha := \frac{e^{-\kappa\eta} - 1}{\kappa}$ . Notice that  $\alpha \leq 0$  for any  $\kappa$  and  $\eta \in [0, T-t]$ .

If  $r\alpha - \theta(\alpha + \eta) > 0$ , then  $F(t, 0, r) > K$ .

Notice that, when  $\kappa \cdot \eta \approx 0$ :

$$\alpha = \frac{e^{-\kappa\eta} - 1}{\kappa} \approx \frac{1 - \kappa\eta + \frac{\kappa^2\eta^2}{2} - 1}{\kappa} = -\eta + \frac{\kappa\eta^2}{2}.$$

Therefore, when  $\kappa \cdot \eta \approx 0$ :

$$r\alpha - \theta(\alpha + \eta) = \alpha(r - \theta) - \theta\eta \approx \left( -\eta + \frac{\kappa\eta^2}{2} \right) (r - \theta) - \theta\eta.$$

Hence, since  $\eta > 0$ ,  $r\alpha - \theta(\alpha + \eta) > 0$  holds true when  $\kappa \cdot \eta \approx 0$  if:

$$\left( -1 + \frac{\kappa\eta}{2} \right) (r - \theta) - \theta > 0. \tag{B3}$$

If  $r > \theta$ , (B3) can be rewritten as:

$$\eta > \frac{2}{\kappa} \cdot \frac{\theta}{r - \theta}.$$

Since  $\eta \in [0, T - t]$ , if

$$T - t > \frac{2}{\kappa} \cdot \frac{r}{r - \theta},$$

(B3) is surely satisfied for at least a  $\eta \in [0, T - t]$ .

If, on the contrary,  $r < \theta$ , (B3) leads to:

$$\eta < \frac{2}{\kappa} \cdot \frac{r}{r - \theta}.$$

Since  $\eta \in [0, T - t]$ , if

$$0 < \frac{2}{\kappa} \cdot \frac{r}{r - \theta},$$

(B3) is surely satisfied for at least a  $\eta \in [0, T - t]$ . Notice that the last inequality holds true if  $\kappa < 0$  and  $r > 0$  or if  $\kappa > 0$  and  $r < 0$ .

Summing up, when  $\kappa \cdot \eta \approx 0$ ,  $F(t, 0, r(t)) > K$  if:

- $T - t > \frac{2}{\kappa} \cdot \frac{r}{r - \theta}$  if  $r > \theta$ ,
- $\kappa r < 0$  if  $r < \theta$ .

Under [NC0], if [NC1], i.e.  $q > 0$ , is not satisfied, than the discounted risky security  $\tilde{S}$  is driven by

$$d\tilde{S}(t) = -qdt + \sigma_S dW_S^{\mathbb{Q}}(t),$$

and  $\tilde{S}$  is a supermartingale. Thus, for any  $t < \tau < T$ ,

$$\mathbb{E}^{\mathbb{Q}} \left[ S(\tau) e^{-\int_t^{\tau} r(s) ds} \right] \leq S(t)$$

and, by Jensen's inequality,

$$\begin{aligned} \mathbb{E}^{\mathbb{Q}} \left[ (K - S(\tau))^+ e^{-\int_t^{\tau} r(s) ds} \right] &\geq \left( K \mathbb{E}^{\mathbb{Q}} \left[ e^{-\int_t^{\tau} r(s) ds} \right] - S(t) \right)^+ \\ &\geq (K - S(t))^+, \end{aligned}$$

where the last inequalities holds under [NC0]. This shows that, under [NC0], if [NC1] is violated, early exercise is never optimal at  $t$ .

Under [NC0], if [NC2] is not satisfied, then  $\pi_A(t, S, r) \geq \pi_E(t, S, r) > (K - S)^+$ , that means that early exercise is never optimal at  $t$ .

We conclude the proof by motivating condition [NC1']. When interest rates are stochastic, the drift of  $S$  is no longer constant. Therefore, we need to impose a condition on its expected value under  $\mathbb{Q}$  up to maturity:

$$\bar{\mu} := \mathbb{E}^{\mathbb{Q}} \left[ \int_0^{T-t} r(s) ds - \left( q + \frac{1}{2} \sigma_S^2 \right) (T-t) \right].$$

Recalling the dynamics of  $r(s)$  in (3), we have:

$$\bar{\mu} = (r - \theta)\alpha + \left( \theta - q - \frac{\sigma_S^2}{2} \right) (T-t)$$

where we set  $\alpha := \frac{1 - e^{-\kappa(T-t)}}{\kappa}$ .

If  $\bar{\mu}$  is negative, the value of  $S$  is expected to decrease and, therefore, the early exercise of the American option is never optimal. If  $\bar{\mu}$  is positive, early exercise might be optimal as the value of  $S$  is expected to increase thus reducing the value of the put option. Therefore,  $\bar{\mu} > 0$  is a necessary condition for the presence of a critical price at  $t$ .

## References

- Amin, K.I. and J.N. Bodurtha Jr.**, “Discrete-time valuation of American options with stochastic interest rates,” *The Review of Financial Studies*, 1995, 8 (1), 193–234.
- Battauz, A., M. De Donno, and A. Sbuelz**, “Real options and American derivatives: the double continuation region,” *Management Science*, 2015, 61 (5), 1094–1107.
- , – , and – , “On the exercise of American quanto options,” *Working paper*, 2018.
- Björk, T.**, *Arbitrage theory in continuous time*, 3 ed., Oxford Finance, 2009.
- Boyarchenko, S. and S. Levendorskiĭ**, “American options in the Heston model with stochastic interest rate and its generalizations,” *Applied Mathematical Finance*, 2013, 20 (1), 26–49.
- Brigo, D. and F. Mercurio**, *Interest rate models-theory and practice: with smile, inflation and credit*, Springer Science & Business Media, 2007.
- Cox, J.C., J.E. Ingersoll, and S.A. Ross**, “A theory of the term structure of interest rates,” *Econometrica*, 1985, 53, 385–407.
- , **S.A. Ross**, and **M. Rubinstein**, “Option pricing: a simplified approach,” *Journal of Financial Economics*, 1979, 7 (3), 229–263.
- Detemple, J.**, “Optimal exercise for derivative securities,” *Annual Review of Financial Economics*, 2014, 6, 459–487.
- Evans, J.D., R. Kuske, and J.B. Keller**, “American options on asset with dividends near expiry,” *Mathematical Finance*, 2002, 12 (1), 219–237.
- Fabozzi, F.J., T. Paletta, S. Stanescu, and R. Tunaru**, “An improved method for pricing and hedging long dated American options,” *European Journal of Operational Research*, 2016, 254 (1), 656–666.

- Gamba, A. and L. Trigeorgis**, “An improved binomial lattice method for multi-dimensional options,” *Applied Mathematical Finance*, 2007, *14* (5), 453–475.
- Geman, H., N. El Karoui, and J.C. Rochet**, “Changes of numéraire, changes of probability measure and option pricing,” *Journal of Applied Probability*, 1995, *32* (2), 443–458.
- Hahn, W.J. and J.S. Dyer**, “Discrete time modeling of mean-reverting stochastic processes for real option valuation,” *European Journal of Operational Research*, 2008, *184* (2), 534–548.
- Ho, T.S., R.C. Stapleton, and M.G. Subrahmanyam**, “The valuation of American options with stochastic interest rates: A generalization of the Geske-Johnson technique,” *The Journal of Finance*, 1997, *52* (2), 827–840.
- Lamberton, D. and S. Villeneuve**, “Critical price near maturity for an American option on dividend-paying stock,” *Annals of Applied Probability*, 2003, *13* (1), 800–815.
- Medvedev, A. and O. Scaillet**, “Pricing American options under stochastic volatility and stochastic interest rates,” *Journal of Financial Economics*, 2010, *98* (1), 145–159.
- Mulinacci, S. and M. Pratelli**, “Functional convergence of Snell envelopes: application to American option approximations,” *Finance and Stochastics*, 1998, *2* (3), 311–327.
- Nelson, D.B. and K. Ramaswamy**, “Simple binomial processes as diffusion approximations in financial models,” *The Review of Financial Studies*, 1990, *3* (3), 393–430.
- Prigent, J.L.**, *Weak convergence of financial markets*, Springer, 2003.
- Stroock, D.W. and S.R.S. Varadhan**, *Multidimensional diffusion processes*, Springer, 1997.

**Vasicek, O.**, “An equilibrium characterization of the term structure,” *Journal of Financial Economics*, 1977, 5 (2), 177–188.