Received: 18 April 2021 | Accepted: 29 June 2022

DOI: 10.1111/mafi.12361

ORIGINAL ARTICLE

MATHEMATICAL FINANCE An International Journal of Mathematics, Statistics and Financial Economics

WILEY

The American put with finite-time maturity and stochastic interest rate

Cheng Cai[1] | **Tiziano De Angelis**[2,3] | **Jan Palczewski**[1]

[1]School of Mathematics, University of Leeds, Leeds, UK

[2]School of Management and Economics, Department of ESOMAS, University of Turin, Turin, Italy

[3]Collegio Carlo Alberto, Turin, Italy

Correspondence
Jan Palczewski, School of Mathematics, University of Leeds, Woodhouse Lane, LS2 9JT Leeds, UK.
Email: J.Palczewski@leeds.ac.uk

Funding information
Engineering and Physical Sciences Research Council, Grant/Award Number: EP/R021201/1; China Scholarship Council

Abstract

In this paper, we study pricing of American put options on a nondividend-paying stock in the Black and Scholes market with a stochastic interest rate and finite-time maturity. We prove that the option value is a C^1 function of the initial time, interest rate, and stock price. By means of Itô calculus, we rigorously derive the option value's early exercise premium formula and the associated hedging portfolio. We prove the existence of an optimal exercise boundary splitting the state space into *continuation* and *stopping* region. The boundary has a parametrization as a jointly continuous function of time and stock price, and it is the unique solution to an integral equation, which we compute numerically. Our results hold for a large class of interest rate models including CIR and Vasicek models. We show a numerical study of the option price and the optimal exercise boundary for Vasicek model.

KEYWORDS

American put option, free boundary problems, integral equations, stochastic interest rate

1 | INTRODUCTION

Pricing of American options is a classical problem in mathematical finance, which has attracted continuous attention since the initial work of McKean Jr (1965). Its study has also become a benchmark for methodological developments of optimal stopping theory and the associated free boundary problems. In this paper, we contribute to this strand of research by studying the American put option on a Black and Scholes market with a stochastic interest rate and finite-time maturity. The stock price and the interest rate are driven by (possibly) correlated Brownian motions and we make minimal assumptions about the dynamics of the interest rate under the pricing measure: the coefficients are time independent and Lipschitz continuous. CIR model, which does not satisfy these conditions, is also included in our analysis.

It is well known (Bensoussan, 1984; Karatzas, 1988) that the American put option price is given by the *value function* of a related optimal stopping problem. In our model, this optimal stopping problem is three-dimensional with two-dimensional diffusive dynamics (stock price and interest rate) and time. The stopping set, that is, the set of points (t, r, x) in which it is optimal to exercise the option, is separated from the continuation set, where it is optimal to hold (or sell) the option, by a single surface called the *stopping boundary*. The value function is a classical solution to a PDE in the interior of the continuation set, that is, it is twice continuously differentiable in (r, x) and once continuously differentiable in t, whereas it coincides with the put payoff in the stopping set.

One of our technical contributions is to establish by means of probabilistic methods that the value function is globally once continuously differentiable in all variables. Then, the continuity of the gradient of the value function permits the application of a generalization of Itô's formula (due to Cai and De Angelis (2021)) and a rigorous derivation of a hedging portfolio. The hedging portfolio invests in three instruments: the money market (savings) account, the zero-coupon bond with maturity equal to the maturity of the option, and the stock. We show that the usual Delta hedging strategy is optimal: the positions in the bond and the stock are given by relevant partial derivatives of the value function. As a further consequence of the generalized Itô's formula, we also derive the decomposition of the American option price as the sum of the price of a European put option with the same maturity and the same exercise price, and an *early exercise premium*. This is known in the literature as the *early exercise premium formula*, which corresponds to Doob's decomposition of supermartingales into a martingale and a nonincreasing process (applied here to the Snell envelope of the optimal stopping problem).

Our second contribution concerns the continuity properties of the stopping boundary in our model, which have not been established in the literature. We are able to demonstrate that the stopping boundary, when parametrized as a function of (t, x), is continuous. Apart from being of interest in its own right, this enables a characterization of the stopping boundary as the unique continuous solution of an integral equation arising from the early exercise premium decomposition. When a stopping boundary is known, efficient numerical methods are at disposal for computation of the option price. One can use Monte Carlo methods based on the early exercise decomposition or classical PDE methods for Cauchy problems (in contrast to the original problem with a free boundary).

American option pricing with stochastic interest rates has already attracted a lot of attention in the literature, mainly focusing on approximations and numerical methods. Lattice (tree) based methods are employed by Appolloni et al. (2015) to price options in Black and Scholes model with CIR interest rate dynamics and by Battauz and Rotondi (2022) in a model with Vasicek interest rates. Geske and Johnson (1984)'s approximation of discretely exercised American options prices is adapted by Ho et al. (1997) and Chung (2000) to a class of stochastic interest rate models that lead to log-normally distributed bond prices. An alternative approximation is provided by Menkveld and Vorst (2000). A framework for option pricing with Heath, Jarrow, Morton's bond market model (Heath et al., 1992) is developed by Amin and Jarrow (1992) with a binomial-tree-based implementation of pricing of foreign exchange options performed in Amin and Bodurtha Jr (1995).

Detemple and Tian (2002) study the pricing of American options in a general diffusive model with a d-dimensional Brownian motion. They formulate assumptions under which there is a single exercise surface but without proving its continuity. In a Black and Scholes market model with Vasicek interest rates, they show that this exercise boundary solves an integral equation of the same form as in this paper. The uniqueness of solutions to this integral equation is not discussed and their numerical method for computing the solution is different to ours.

Hedging underlies the success of mathematical finance in derivatives markets. A rigorous theory that links hedging of American options with solutions of optimal stopping problems was initiated by Bensoussan (1984) using PDE methods and extended by Karatzas (1988) to more general models and payoffs thanks to the martingale theory of optimal stopping. A hedging strategy for an American option consists of an investment portfolio and a nondecreasing cumulative consumption process, which increases only when the state-time process is in the stopping set. In the Black and Scholes model with constant interest rate, the classical Delta hedge is known to replicate the option (Karatzas & Shreve, 1998b, Thm. 7.9, Ch. 2). This paper seems to be the first to rigorously derive the hedging strategy for American put options on a market with a stochastic interest rate. This is accomplished thanks to the C^1-regularity of the value function that we are able to prove and which did not appear in previous works.

A characterization of an optimal stopping boundary as solution to a (system of) integral equations has been known since the earliest works (see Van Moerbeke (1976)). In more recent works (Carr et al., 1992; Jacka, 1991; Kim, 1990; Myneni, 1992), the stopping boundary for the classical Black and Scholes market with constant interest rate is shown to be the unique solution to an uncountable *system* of integral equations arising from the early exercise premium decomposition of the option price. A break-through came with the work of Peskir (2005) where he shows that the stopping boundary is the unique continuous solution of a single integral equation. His key observation is that the integral equation only needs to be satisfied for stock prices at the boundary while earlier results required that it does so for all stock prices at and below the boundary. Peskir (2005)'s integral equation opens doors to side-stepping the computation of the value function in the process of determining the optimal exercise strategy; see numerical methods designed in Little et al. (2000); Kim et al. (2013). Our paper extends Peskir (2005)'s results to the market with a stochastic interest rate and the optimal boundary being a two-dimensional surface. It is also the continuity of the boundary that allows us to establish the uniqueness of solutions to the integral equation. A closely related paper that furthermore motivated our numerical approach is Detemple et al. (2018) where the authors solve an integral equation for Black and Scholes market with stochastic volatility.

The regularity of the value function in one-dimensional optimal stopping problems is often phrased as *smooth-fit* and plays a major role in determining explicit solutions. In a Black and Scholes model with constant interest rate, smooth-fit for American options with finite-time maturity is understood as continuous differentiability of the value function with respect to the stock price, for each fixed value of the time variable (see Jacka (1991) and subsequent works). That is a "directional" derivative and continuity is only considered with respect to one variable. Sobolev space regularity is studied in Jaillet et al. (1990) for American options on multiple assets and *deterministic*, time-dependent discount rate under the assumption of uniform ellipticity of the associated second-order differential operator. By Sobolev embedding, it is possible to determine continuous differentiability of the value function with respect to the initial values of all the assets but not with respect to time. Continuous differentiability with respect to time and stock price for the value of the American put with finite-time maturity and constant interest rate is obtained in De Angelis and Peskir (2020) along with other complementary findings about continuous differentiability of the value function for a large class of optimal stopping problems. In this paper, we refine the arguments from De Angelis and Peskir (2020) and remove global integrability conditions that may not hold in our set-up.

The early exercise premium formula for American options was studied in great generality, in non-Markovian problems beyond the setting of the American put option by Rutkowski (1994) with methods from martingale theory. The nature of the methods employed in Rutkowski (1994) to derive his main results is such that the emphasis is removed from the optimal boundary, which in fact only appears in specific examples (see Section 3 of that paper) as a time-dependent function. Here instead we derive the early exercise premium formula starting from the analysis of the optimal boundary (and its regularity) as a function of time and one stochastic factor from our two-factor model.

Some of the ideas in this paper find wider applicability in optimal stopping theory. The generalization of Itô's formula that we use to find the hedging portfolio has natural applications to other optimal stopping problems as discussed extensively in Section 3 of Cai and De Angelis (2021). Localization of the arguments from De Angelis and Peskir (2020) to prove continuous differentiability of the value function does not rely much on the specific structure of our problem and suggests a general recipe to address the issue. Finally, our ideas for the continuity of the optimal stopping surface have been expanded upon to cover more general settings in Cai et al. (2021).

The paper is structured as follows. Section 2 introduces the market model, main assumptions and notation. The main contributions are discussed in Section 3 while their proofs are delayed until after Section 4. A numerical study with interest rates following Vasicek model is presented in Section 4 along with a sensitivity analysis. Monotonicity and Lipschitz continuity of the value function is proved in Section 5. Existence of the stopping surface and its regularity (in the sense of diffusions) are shown in Section 6. In Section 7, we prove that the value function is continuously differentiable on the whole domain. Auxiliary estimates needed for admissibility of the hedging strategy are provided in Section 9. Four appendices contain further details.

2 | PROBLEM FORMULATION

Let $(\Omega, \mathcal{F}, \mathsf{P})$ be a complete probability space carrying two correlated Brownian motions $(B_t)_{t\geq 0}$ and $(W_t)_{t\geq 0}$ with $\mathsf{E}(W_t B_t) = \rho\, t$ for all $t \geq 0$ and a fixed $\rho \in (-1,1)$ (here $\mathsf{E}(\,\cdot\,)$ is the expectation under P). We denote by $(\mathcal{F}_t)_{t\geq 0}$ the filtration generated by (B, W) augmented with the P-null sets. On this probability space, we consider a financial market with one risky asset $(X_t)_{t\geq 0}$ and a bond. The asset and the risk-free (short) rate $(r_t)_{t\geq 0}$ take values, respectively, in intervals $\mathbb{R}_+ := (0, \infty)$ and $\mathcal{I} \subseteq \mathbb{R}$, and follow the dynamics

$$dX_t = r_t X_t dt + \sigma X_t dB_t, \qquad X_0 = x, \tag{1}$$

$$dr_t = \alpha(r_t)dt + \beta(r_t)dW_t, \qquad r_0 = r, \tag{2}$$

with $\alpha, \beta : \mathcal{I} \to \mathbb{R}$ specified below. The probability measure P is a risk neutral measure for this market. We denote by $T > 0$ a fixed finite trading horizon.

Throughout the paper, we assume $\sigma > 0$ and $\mathcal{I} = (\underline{r}, \overline{r})$ (with $\mathcal{I}$ possibly unbounded). The right boundary $\overline{r}$ is unattainable in a finite time (it is a natural or entrance-not-exit boundary). The left boundary $\underline{r}$ is either unattainable or reflecting. It will become clear later that the exact behavior of the interest rate process at this boundary is irrelevant for the majority of results and their proofs. For the dynamics of the interest rate, our benchmark example is the CIR model, but, with a relatively small additional effort, our results cover other stochastic interest rate models, for example, Vasicek model. Therefore, we make the following standing assumption:

Assumption 2.1. The coefficients α and β in Equation (2) meet one of the conditions below:

(i) *(CIR model)* For $\kappa, \theta, \gamma > 0$, we have $\alpha(r) = \kappa(\theta - r)$ and $\beta(r) = \gamma\sqrt{r}$.
(ii) α and β are globally Lipschitz and continuously differentiable on bounded subsets of $\mathcal{I}$ with $\beta(r) > 0$ for all $r \in \mathcal{I}$, and $\overline{r} > 0 \geq \underline{r}$. For any compact set $\mathcal{K} \subset \mathcal{I}$, and any $p \in [1, p']$ for some $p' > 2$ and $T > 0$, there is $C_1 > 0$ (depending on T, p, and $\mathcal{K}$) such that

$$\sup_{r \in \mathcal{K}} \mathsf{E}\left[\sup_{0 \leq s \leq T} e^{-p\int_0^s r_u du} \middle| r_0 = r \right] \leq C_1. \tag{3}$$

The assumption that $\overline{r} > 0$ cannot be relaxed without trivializing the pricing problem. A strictly positive lower boundary $\underline{r}$ could, however, be of interest. For the clarity of presentation, it is omitted but it can be studied with similar methods as those developed in this paper.

The above assumptions are sufficient to guarantee that Equation (2) admits a unique strong solution defined on $\mathcal{I}$. In the case of CIR model, we also have $\kappa\theta > 0$, which implies that the spot rate is non-negative (but not necessarily strictly positive), see, for example, (Jeanblanc et al., 2009, Sec. 6.3.1), so the left boundary $\underline{r} = 0$ is reflecting (also nonattainable if $\kappa\theta > \sigma^2/2$). Hence, the bound (3) is satisfied with the constant $C_1 = 1$. The linear growth of α and β in Equation (2) guarantees that for each $p \geq 2$ there is $C_2 > 0$ only depending on T and p, such that (Krylov, 1980, Ch. 2, Sec. 5, Thm. 9)

$$\mathsf{E}\left[\sup_{0 \leq s \leq T} |r_s|^p \middle| r_0 = r \right] \leq C_2(1 + |r|^p), \qquad \text{for } r \in \mathcal{I}. \tag{4}$$

Under Assumption 2.1, the solution of Equation (1) may be expressed as

$$X_t = x \exp\left(\sigma B_t + \int_0^t \left(r_s - \frac{\sigma^2}{2} \right) ds \right), \qquad \text{for } t \geq 0, \tag{5}$$

so that X depends on both initial values r and x. On the contrary, the dynamics of the interest rate does not depend on the initial asset value. The coupling between the processes $(r_t)_{t\geq 0}$ and $(X_t)_{t\geq 0}$ stems from formula (5) and the correlation between the Brownian motions. To keep track of the dependence of the processes on their initial values, in what follows we often use the notation $(r_t^r, X_t^{r,x})_{t\geq 0}$ for the process started at $r_0^r = r$ and $X_0^{r,x} = x$. Also, we may sometimes use the notation $\mathsf{P}_{t,r,x}(\,\cdot\,) = \mathsf{P}(\,\cdot\,| r_t = r, X_t = x)$, $\mathsf{P}_{r,x} = \mathsf{P}_{0,r,x}$, and $\mathsf{P}_r(\,\cdot\,) = \mathsf{P}(\,\cdot\,| r_0 = r)$.

According to the classical theory (Karatzas and Shreve, 1998b, Ch. 2, Thm. 5.8), the rational price of an American put option with maturity time T, strike price $K > 0$, written on the asset X

and evaluated at time $t \in [0,T]$ is given by

$$p_t = \operatorname*{ess\,sup}_{t\le\tau\le T} \mathsf{E}\left[e^{-\int_t^\tau r_s ds}(K-X_\tau)^+ \middle| \mathcal{F}_t\right], \quad 0 \le t \le T,$$

where the essential supremum is over $(\mathcal{F}_t)$-stopping times in $[t,T]$ and the function $(\cdot)^+$ denotes the positive part. In our Markovian set-up, $p_t = v(t, r_t, X_t)$ for a Borel-measurable function $v : [0,T] \times [\bar{r}, \underline{r}] \times \mathbb{R}_+$ (see Shiryaev (2008, Ch. 3)). Using that the process $(r_t, X_t)_{t\ge 0}$ is time-homogeneous and strong Markov, we can express v as

$$v(t,r,x) = \sup_{0\le\tau\le T-t} \mathsf{E}_{r,x}\left[e^{-\int_0^\tau r_s ds}(K-X_\tau)^+\right], \tag{6}$$

where $r \in \mathcal{I}$ and $x \in \mathbb{R}_+$ are, respectively, the values of the spot rate and of the asset at time t. The above is an optimal stopping problem with Markovian structure and a three-dimensional state space.

Since the process

$$t \mapsto e^{-\int_0^t r_s ds}(K-X_t)^+, \tag{7}$$

is non-negative and continuous, and thanks to the integrability condition (3), we can rely on standard optimal stopping theory (see, e.g., Appendix D in Karatzas and Shreve (1998b)) to conclude that the smallest optimal stopping time for Equation (6) is $\mathsf{P}_{r,x}$-a.s. given by

$$\tau_* := \inf\{s \ge 0 \,:\, v(t+s, r_s, X_s) = (K - X_s)^+\}, \tag{8}$$

where we note that $\tau_* \le T - t$ since $v(T,r,x) = (K-x)^+$. Clearly $\tau_* = \tau_*(t,r,x)$ depends on the initial value (t,r,x) of the three-dimensional state process $(t+s, r_s, X_s)_{s\ge 0}$.

The form (8) of τ^* gives rise to the so-called continuation set C and its complement, the stopping set $\mathcal{D}$, that is

$$C := \{(t,r,x) \in [0,T] \times \mathcal{I} \times \mathbb{R}_+ \,:\, v(t,r,x) > (K-x)^+\}, \tag{9}$$

$$\mathcal{D} := \{(t,r,x) \in [0,T] \times \mathcal{I} \times \mathbb{R}_+ \,:\, v(t,r,x) = (K-x)^+\}. \tag{10}$$

Upon observing the spot rate and the asset value, at each time, the option holder must decide whether to hold the option or to exercise it. She should *wait* (possibly trading the option on the market) if $(t, r_t, X_t) \in C$ since the option value is strictly larger than the payoff of immediate exercise. On the contrary, if $(t, r_t, X_t) \in \mathcal{D}$, the option should be immediately *exercised*. Notice that

$$\{T\} \times \mathcal{I} \times \mathbb{R}_+ \subseteq \mathcal{D}.$$

Remark 2.2. Setting

$$D_s := \exp\left(-\int_0^s r_u du\right), \quad V_s := v(t+s, r_s, X_s) \quad \text{and} \quad Y_s := D_s V_s,$$

(i.e., Y is the discounted option value process), we have that (Karatzas and Shreve, 1998b, Appendix D)

$$(Y_s)_{s\in[0,T-t]} \text{ is a right-continuous } \mathsf{P}_{r,x}\text{-supermartingale,} \tag{11}$$

$$(Y_{s\wedge\tau_*})_{s\in[0,T-t]} \text{ is a right-continuous } \mathsf{P}_{r,x}\text{-martingale.} \tag{12}$$

We will soon show (Proposition 5.1) that v is a continuous function, so that Y is a continuous process.

Notation. We set

$$\mathcal{O} := [0,T) \times \mathcal{I} \times \mathbb{R}_+, \tag{13}$$

and denote by $\partial\mathcal{C}$ the boundary of $\mathcal{C}$ in $\mathcal{O}$, that is, $\partial\mathcal{C} := (\overline{\mathcal{C}} \cap \mathcal{O}) \setminus \mathcal{C}$.

For future frequent use, we denote by $\mathcal{L}$ the infinitesimal generator of $(r_t, X_t)_{t\geq 0}$, which, for any $f \in C^2(\mathcal{I} \times \mathbb{R})$ reads

$$\mathcal{L}f := \frac{\sigma^2 x^2}{2} f_{xx} + \frac{\beta^2(r)}{2} f_{rr} + \rho\sigma x\beta(r) f_{rx} + rx f_x + \alpha(r) f_r, \tag{14}$$

where f_r, f_x and f_{rr}, f_{rx}, f_{xx} denote, respectively, the first- and second-order partial derivatives of f.

3 | MAIN RESULTS

In this section, we provide the main results of the paper. In Sections 3.1 and 3.2, under the sole Assumption 2.1, we establish continuous differentiability of the value function $v(t,r,x)$ (jointly in all variables), along with its monotonicity in (t,r,x) and convexity in x. We also prove the existence and monotonicity of an optimal exercise boundary and present two possible parametrizations of it. Then, in Sections 3.3–3.6, under a mild additional assumption on α and β (Assumption 3.6), we derive continuity of the optimal exercise boundary (as a function of two variables) and an integral equation that uniquely determines it (also under Assumption 3.8). Finally, we obtain the early exercise premium formula for the option price and the hedging portfolio that replicates the option's payoff at all times.

3.1 | Optimal stopping boundary

In the classical Black–Scholes model with constant interest rate, the stopping set is determined by a boundary: it is optimal to exercise the option the first time when the stock price drops below this boundary. A similar characterization of the stopping region $\mathcal{D}$ can be derived in our model with the difference that the stopping boundary is a surface. To this end, we research monotonicity properties of the value function.

Proposition 3.1. *The value function v is finite for all $(t,r,x) \in \mathcal{O}$ and it satisfies the following conditions:*

(i) $t \mapsto v(t,r,x)$ *is nonincreasing for all* $(r,x) \in \mathcal{I} \times \mathbb{R}_+$,
(ii) $r \mapsto v(t,r,x)$ *is nonincreasing for all* $(t,x) \in [0,T] \times \mathbb{R}_+$,
(iii) $x \mapsto v(t,r,x)$ *is convex and nonincreasing for all* $(t,r) \in [0,T] \times \mathcal{I}$.

Proof. See Section 5. □

The monotonicity in t and x and the convexity in x are the same as in the classical Black–Scholes model and the proof is very similar. The dependence on r has financial explanation: larger interest rate implies stronger discounting of future cashflows and, hence, lower present value.

Remark 3.2. In the case $T = +\infty$ (perpetual option), the discounted payoff process (7) is still uniformly integrable and continuous. This implies that, letting v_∞ denote the value of the perpetual option, the stopping time

$$\tau_\infty = \inf\{t \geq 0 \ : \ v_\infty(r_t, X_t) = (K - X_t)^+\},$$

is optimal by standard theory and Equations (11)–(12) continue to hold in this setting (see, e.g., Shiryaev (2008, Ch. 3, Thm. 3)). In particular, it can be shown that $r \mapsto v_\infty(r,x)$ is nonincreasing and $x \mapsto v_\infty(r,x)$ is convex and nonincreasing.

From the general optimal stopping theory, we expect that the value function v be continuous. Indeed, this fact is proved from first principles in our Proposition 5.1 in Section 5 (without relying on the form of the stopping set). The continuity of v means that the continuation set $\mathcal{C}$ is open and the stopping set $\mathcal{D}$ is closed. In view of the monotonicity properties established in Proposition 3.1, we can show that there is a surface splitting $\mathcal{C}$ and $\mathcal{D}$.

In models with constant interest rate, an optimal boundary is often defined as function of time, which provides a threshold for the process (X_t). A parametrization of the stopping surface as a function $b(t,r)$ of time and interest rate is also available in our setting. For the sake of mathematical tractability, we prefer to work with the parametrization $c(t,x)$ in terms of time and stock price. Due to technical reasons that will become clearer in Section 6, we are able to prove the continuity of $(t,x) \mapsto c(t,x)$ jointly in both variables (t,x), but not the joint continuity of b in (t,r). However, b is more convenient for numerical computations in Section 4 as it admits values in a bounded interval $[0,K]$. The connection between b and c is established in Proposition 3.4.

Proposition 3.3. *There exists a function* $c(t,x)$ *on* $[0,T] \times [0,\infty]$, *such that*

$$\mathcal{D} = \{(t,r,x) \in \mathcal{O} \ : \ r \geq c(t,x)\} \cup (\{T\} \times \mathcal{I} \times \mathbb{R}_+), \tag{15}$$

$$\mathcal{C} = \{(t,r,x) \in \mathcal{O} \ : \ r < c(t,x)\}. \tag{16}$$

The function $c(t,x)$ *has following properties:*

(i) *For any* $(t_0,x_0) \in [0,T) \times \mathbb{R}_+$, *the mapping* $t \mapsto c(t,x_0)$ *is right-continuous and nonincreasing and the mapping* $x \mapsto c(t_0,x)$ *is left-continuous and nondecreasing.*
(ii) $c(t,x) = \bar{r}$ *for* $(t,x) \in [0,T) \times [K,\infty)$.
(iii) $c(t,x) \geq 0$ *for* $(t,x) \in [0,T) \times \mathbb{R}_+$, *and* $\lim_{x \downarrow 0} c(t,x) = 0$ *for* $t \in [0,T)$.

Proof. See Section 6. □

Notice that (ii) and (iii) above imply that it is never optimal to exercise the option out of the money or if the interest rate is negative. This is in line with classical financial wisdom.

The following proposition whose simple proof is omitted gives details of the reparametrization of the stopping boundary as a function $b(t, r)$ of time and interest rate.

Proposition 3.4. *Define*

$$b(t,r) := \inf\{x \in \mathbb{R}_+ : c(t,x) > r\}, \qquad (t,r) \in [0,T) \times \mathcal{I}.$$

The mappings $t \mapsto b(t, r_0)$ and $r \mapsto b(t_0, r)$ are right-continuous and nondecreasing for any $(t_0, r_0) \in [0,T) \times \mathcal{I}$. For any $t \in [0,T)$, we have $K > b(t,r) > 0$ when $r > 0$, and $b(t,r) = 0$ when $r < 0$. Furthermore,

$$\mathcal{D} = \{(t,r,x) \in \mathcal{O} : x \leq b(t,r)\} \cup (\{T\} \times \mathcal{I} \times \mathbb{R}_+),$$

$$\mathcal{C} = \{(t,r,x) \in \mathcal{O} : x > b(t,r)\}.$$

3.2 | Smoothness of the value function

It is well-known that v satisfies (in the classical sense)

$$\begin{aligned} v_t(t,r,x) + (\mathcal{L} - r)v(t,r,x) = 0, &\qquad (t,r,x) \in \mathcal{C}, \\ v(t,r,x) = (K - x)^+, &\qquad (t,r,x) \in \mathcal{D}, \end{aligned} \tag{17}$$

where $\mathcal{L}$ is the generator of (r, X) defined in Equation (14). Hence, standard arguments assert that v is $C^{1,2}$ in $\mathcal{C} \cap \mathrm{int}(\mathcal{D})$. Classical optimal stopping theory identifies the boundary of the set $\mathcal{C}$ by imposing the so-called smooth-fit condition. In the American put problem with constant interest rate, this corresponds to proving that $x \mapsto v_x^\circ(t,x)$ is continuous for each $t \in [0,T)$ fixed, with v° denoting the value function associated to the option price. In our setting, we prove a stronger result and show continuous differentiability of v across the stopping boundary $\partial\mathcal{C}$, that is, the global continuity of the gradient of v (as a function of all variables) in $\mathcal{O}$. We use ideas similar to those in De Angelis and Peskir (2020) but we must refine arguments therein and use estimates with "local" nature since we are not able to directly check their assumptions. In particular, global differentiability of the flow $r \mapsto (r_s^r, X_s^{r,x})$ and related integrability conditions (see Equations (4.4)–(4.7) and Theorem 10 in De Angelis and Peskir (2020)) are not easily verifiable when, for example, the interest rate follows the CIR dynamics.

Theorem 3.5. *We have $v \in C^1(\mathcal{O})$.*

Proof. See Section 7. □

It is worth noticing that the proof of the above result combines a number of steps that may be of independent interest. In particular, we prove local Lipschitz continuity of v (Proposition 5.1) and the regularity of the stopping boundary in the sense of diffusions. The latter gives the continuity of optimal stopping times τ_* as functions of the initial state, which plays a crucial role in the proof of the theorem.

3.3 | Continuity of the stopping boundary and Dynkin's formula

Preliminary right/left-continuity properties of the stopping boundary $(t,x) \mapsto c(t,x)$ illustrated above follow from its monotonicity and the closedness of the stopping set $\mathcal{D}$ (see Proposition 3.3). However, thanks to the C^1 regularity of the value function v, we can also prove joint continuity of the stopping boundary in both variables. For this, we require local Hölder continuity of the derivatives of the coefficients in the dynamics of the short rate r.

Assumption 3.6. The functions α and β in Equation (2) have first- and second-order derivatives, respectively, Hölder continuous on any compact subset of $\mathcal{I}$.

Note that this assumption is satisfied by CIR model. It strengthens Assumption 2.1(ii) by requiring that the derivatives are not only locally continuous but also locally Hölder continuous. This technical requirement is satisfied by many popular short rate models. The joint continuity of optimal stopping boundaries depending on multiple variables has not been proved with probabilistic techniques before, so the next result is of independent mathematical interest.

Proposition 3.7. *Under Assumption 3.6, the function* $c : [0,T) \times \mathbb{R}_+ \to [0,\infty)$ *is continuous.*

Proof. See Section 8. □

Summarizing, we have $v \in C^1(\mathcal{O}) \cap C^{1,2}(\mathcal{C}) \cap C^{1,2}(\mathcal{D})$, and the optimal stopping boundary c is continuous. This is not sufficient to apply the change of variable formula developed in Peskir (2007), which is often used in optimal stopping literature to establish Itô's formula for the value function. Indeed, since Peskir (2007) deals with functions that are not necessarily C^1, it requires that $t \mapsto c(t, X_t)$ be a semi-martingale, so that the local time on the stopping boundary is well-defined. While we were unable to prove it for our optimal boundary, we can instead take advantage of the continuous differentiability of our value function and use a generalization of Itô's formula from Cai and De Angelis (2021), which only requires the monotonicity of the boundary. Notice that, interestingly, we need not control the second-order spatial derivatives near $\partial\mathcal{C}$ in order to apply results from Cai and De Angelis (2021). We do, however, need to ensure that both boundary points of the set $\mathcal{I}$ are nonattainable, because we have not proven that the derivatives $v_t(t,\underline{r},x)$, $v_r(t,\underline{r},x)$ and $v_x(t,\underline{r},x)$, understood as the limit as $r \to \underline{r}$, are well-defined.

Assumption 3.8. The lower boundary point $\underline{r}$ is nonattainable by the process (r_t). In particular, under Assumptions 2.1(i) we require $k\theta > \sigma^2/2$.

Proposition 3.9. *Under Assumption 3.8, for any* $(t,r,x) \in \mathcal{O}$ *and any stopping time* $\tau \in [0, T-t]$, *the value function satisfies the following Dynkin's formula:*

$$v(t,r,x) = \mathsf{E}_{r,x}\left[\int_0^\tau e^{-\int_0^u r_v dv} K r_u 1_{\{r_u > c(t+u, X_u)\}} du + e^{-\int_0^\tau r_v dv} v(t+\tau, r_\tau, X_\tau)\right]. \tag{18}$$

Proof. See Section 8. □

In the proof of the above proposition, we show that the discounted value function satisfies for any stopping time $\tau \in [0, T-t]$

$$\begin{aligned} & e^{-\int_0^\tau r_v dv} v(t+\tau, r_\tau, X_\tau) \\ & = v(t,r,x) - \int_0^\tau e^{-\int_0^s r_v dv} K r_s 1_{\{r_s > c(t+s, X_s)\}} ds + \int_0^\tau e^{-\int_0^s r_v dv} \sigma X_s v_x(t+s, r_s, X_s) dB_s \\ & \quad + \int_0^\tau e^{-\int_0^s r_v dv} \beta(r_s) v_r(t+s, r_s, X_s) dW_s. \end{aligned} \tag{19}$$

This representation will play a fundamental role in deriving a hedging strategy for the American put option in Section 3.6.

Remark 3.10. By arguments in Appendix B, in particular Remark B.1, the distribution of (r_u, X_u), for any $u > 0$, is absolutely continuous with respect to Lebesgue measure on any compact set. Hence, for any $(t, r, x) \in \mathcal{O}$,

$$\mathsf{E}_{r,x}\left[\int_0^{T-t} \left|1_{\{(t+u, r_u, X_u) \in \mathcal{D}\}} - 1_{\{r_u > c(t+u, X_u)\}}\right| du\right] = 0. \tag{20}$$

When we apply this result to Equation (18), we obtain

$$v(t,r,x) = \mathsf{E}_{r,x}\left[\int_0^\tau e^{-\int_0^u r_v dv} K r_u 1_{\{(t+u, r_u, X_u) \in \mathcal{D}\}} du + e^{-\int_0^\tau r_v dv} v(t+\tau, r_\tau, X_\tau)\right].$$

3.4 | Early exercise premium

Inserting $\tau = T - t$ in Equation (18), we obtain a decomposition of the American option price into a sum of the European option price v_e and an *early exercise premium* v_p (see Rutkowski (1994) for a derivation of this formula only using general martingale theory):

$$v(t,r,x) = v_p(t,r,x;T,b) + v_e(t,r,x;T), \tag{21}$$

where

$$\begin{aligned} v_e(t,r,x;T) &= \mathsf{E}_{r,x}\left[e^{-\int_0^{T-t} r_v dv}(K - X_{T-t})^+\right], \\ v_p(t,r,x;T,b) &= \mathsf{E}_{r,x}\left[\int_0^{T-t} e^{-\int_0^u r_v dv} K r_u 1_{\{r_u > c(t+u, X_u)\}} du\right] \\ &= \mathsf{E}_{r,x}\left[\int_0^{T-t} e^{-\int_0^u r_v dv} K r_u 1_{\{X_u < b(t+u, r_u)\}} du\right]. \end{aligned} \tag{22}$$

The last equality follows from $r > c(t,x) \Leftrightarrow x < b(t,r)$ by construction of b as the generalized inverse of c. By Remark 3.10, the early exercise premium also reads

$$v_p(t,r,x;T,b) = \mathsf{E}_{r,x}\left[\int_0^{T-t} e^{-\int_0^u r_v dv} K r_u 1_{\{(t+u, r_u, X_u) \in \mathcal{D}\}} du\right].$$

3.5 | Integral equation for the stopping boundary

Proposition 3.9 provides a characterization of the optimal stopping boundary $c(t,x)$. Indeed, for any $(t,x) \in [0,T) \times \mathbb{R}_+$ such that $c(t,x) \in \mathcal{I}$, inserting $\tau = T-t$ and $r = c(t,x)$ in Equation (18) yields an integral equation for c:

$$(K-x)^+ = \mathsf{E}_{c(t,x),x}\left[\int_0^{T-t} e^{-\int_0^u r_v dv} K r_u 1_{\{r_u > c(t+u,X_u)\}} du + e^{-\int_0^{T-t} r_v dv}(K - X_{T-t})^+\right]. \tag{23}$$

The condition that $c(t,x) \in \mathcal{I}$ is necessary as c can take values $\underline{r}$ and $\bar{r}$, which do not belong to the state space $\mathcal{I}$, and the interest rate process r may not be started from there. Notice also that $c(t,x) \notin \mathcal{I}$ when $x \geq K$ so the left-hand side of Equation (23) can be replaced by $(K-x)$. In line with well-known results for American options with constant interest rate (Peskir, 2005), it also turns out that c is the unique solution of the integral equation.

Proposition 3.11. *Under Assumptions 3.6 and 3.8, the function c is the unique function* $\phi : [0,T) \times \mathbb{R}_+ \to [0,\bar{r}]$ *such that:*

(1) ϕ *is continuous, nondecreasing in x and nonincreasing in t, with* $\phi(t,x) = \bar{r}$ *for* $x \geq K$,
(2) ϕ *satisfies Equation (23) (with c therein replaced by* ϕ*) for all* $(t,x) \in [0,T) \times \mathbb{R}_+$ *for which* $\phi(t,x) \in \mathcal{I}$, *and* $\{(t,x) \in [0,T) \times \mathbb{R}_+ : \phi(t,x) \in \mathcal{I}\} \neq \emptyset$.

The integral equation (23) has an analogue for the function $b(t,r)$ from Proposition 3.4. Indeed, for $b(t,r) > 0$, taking $x = b(t,r)$ and $\tau = T-t$ in Proposition 3.9 and using $v(t,r,b(t,r)) = K - b(t,r)$ we see that b solves the integral equation:

$$\begin{aligned} K - b(t,r) = \mathsf{E}_{r,b(t,r)}&\left[\int_0^{T-t} e^{-\int_0^u r_v dv} K r_u 1_{\{X_u < b(t+u,r_u)\}} du\right] \\ &+ \mathsf{E}_{r,b(t,r)}\left[e^{-\int_0^{T-t} r_v dv}(K - X_{T-t})^+\right], \end{aligned} \tag{24}$$

where we use $\{X_u < b(t+u,r_u)\} = \{r_u > c(t+u,X_u)\}$, which follows from $x > b(t,r) \Leftrightarrow r < c(t,x)$ by construction of b as the generalized inverse of c. This parametrization of the integral equation extends the one obtained in the classical American put problem with constant interest rate to our two-factor set-up. Once again, we can prove uniqueness of the solution to the integral equation but without requiring continuity of b, which is a nonstandard result for this type of equations.

Corollary 3.12. *Under the assumptions of Proposition 3.11, the function b is the unique function* $\psi : [0,T) \times \mathcal{I} \to [0,K)$ *such that:*

(1) $t \mapsto \psi(t,r)$ *and* $r \mapsto \psi(t,r)$ *are right-continuous and nondecreasing,*
(2) the generalized left-continuous inverse $\phi(t,x) := \inf\{r \in \mathcal{I} : \psi(t,r) \geq x\}$ *is continuous in* (t,x), *nondecreasing in x and nonincreasing in t,*
(3) ψ *satisfies Equation (24) with (b therein replaced by* ψ*) for all* $(t,r) \in [0,T) \times \mathcal{I}$ *such that* $\psi(t,r) > 0$, *and* $\{(t,x) \in [0,T) \times \mathcal{I} : \psi(t,r) > 0\} \neq \emptyset$.

Notice that $\phi(t,x) = \bar{r}$ for $x \geq K$ follows immediately from $\psi(t,r) < K$.

Integral equations (23) and (24) offer a method to compute the optimal stopping boundary without using the value function v. We will demonstrate it in Section 4 where we design a numerical method for solving such integral equations. Knowing the stopping boundary b, the decomposition (21) can be employed to obtain an efficient numerical estimate of the option value. This offers an alternative to numerical solution of the variational inequality for the value function v, and, subsequently, extraction of the optimal exercise boundary.

3.6 | Hedging portfolio

Thanks to the change of variable formula (19), we are also able to rigorously construct a hedging portfolio that (super)replicates the option payoff at all times. This is based on the classical delta-hedging ideas in the Black and Scholes model but its rigorous mathematical derivation requires smoothness of the option price function, which was not previously established in the literature.

Consider a market comprising three instruments: the money market account $M_t := e^{\int_0^t r_u du}$, the risky stock with the dynamics (1), and a zero-coupon bond with maturity T. We will construct a hedging portfolio for the American option on this market. We remark that the zero-coupon bond can be replaced by any other financial instrument whose dynamics depends on the Brownian motion W driving the interest rate, see Karatzas (1988).

The risk-neutral price of the zero-coupon bond at time $t \in [0,T]$ is given by

$$P(t,r) := \mathsf{E}_r\left[e^{-\int_0^{T-t} r_u du}\right], \quad P(T,r) = 1. \tag{25}$$

By standard arguments based on pathwise continuity of the flow $(t,r) \mapsto r_t^r(\omega)$, one can easily show that P is continuous on $[0,T] \times \mathcal{I}$. Then, under Assumption 2.1, the classical PDE theory (Friedman, 1964, Thm. 9, Ch. 4, Sec. 3) guarantees that P is the unique classical solution of the boundary value problem

$$\begin{aligned} &(\partial_t + \mathcal{L}_r - r)u(t,r) = 0, && (t,r) \in [0,T] \times (a,b), \\ &u(t,r) = P(t,r), && t \in [0,T),\ r \in \{a,b\} \\ &u(T,r) = 1, && r \in [a,b], \end{aligned}$$

where $\mathcal{L}_r = \alpha(r)\partial_r + \beta(r)^2/2\partial_{rr}$ and $(a,b) \subset \mathcal{I}$ is an arbitrary bounded interval. In particular, by arbitrariness of (a,b), we have $P \in C^{1,2}([0,T) \times \mathcal{I})$ and

$$(\partial_t + \mathcal{L}_r - r)P(t,r) = 0, \qquad (t,r) \in [0,T] \times \mathcal{I}.$$

Then, using Itô's formula, the discounted bond price dynamics reads

$$de^{-\int_0^s r_u du} P(s,r_s) = P_r(s,r_s)\beta(r_s)dW_s. \tag{26}$$

Denote by $\phi^{(1)}, \phi^{(2)}, \phi^{(3)}$ the holdings in the stock, the bond and the money market account, respectively. Let C be a nondecreasing continuous process starting from 0 modeling consumption.

The value of a self-financing portfolio starting at time 0 from $v(0,r,x)$ is

$$\Pi_s = v(0,r,x) + \int_0^s \phi_u^{(1)} dX_u + \int_0^s \phi_u^{(2)} dP(u,r_u) + \int_0^s \phi_u^{(3)} dM_u - C_s, \quad s \in [0,T]. \tag{27}$$

The portfolio is *admissible* if all integrals above are semimartingales. Taking the money-market account as a numeráire, we obtain from equations (27) and

$$\Pi_s := \phi_s^{(1)} X_s + \phi_s^{(2)} P(s,R_s) + \phi_s^{(3)} M_s, \quad s \in [0,T], \tag{28}$$

that the dynamics of the discounted portfolio value reads

$$\begin{aligned} de^{-\int_0^s r_u du}\Pi_s &= \phi_s^{(1)} de^{-\int_0^s r_u du} X_s + \phi_s^{(2)} de^{-\int_0^s r_u du} P(s,r_s) - e^{-\int_0^s r_u du} dC_s \\ &= e^{-\int_0^s r_u du}\phi_s^{(1)}\sigma X_s dB_s + e^{-\int_0^s r_u du}\phi_s^{(2)}\beta(r_s)P_r(s,r_s)dW_s - e^{-\int_0^s r_u du}dC_s. \end{aligned} \tag{29}$$

This means that a self-financing portfolio is uniquely determined by the processes $\phi^{(1)}, \phi^{(2)}$, and C.

Comparing Equation (29) with (19), a candidate for the hedging strategy is given by

$$\phi_s^{(1)} = v_x(s,r_s,X_s), \quad \phi_s^{(2)} = \frac{v_r(s,r_s,X_s)}{P_r(s,r_s)}, \quad C_s = \int_0^s K r_u 1_{\{r_u > c(u,X_u)\}} du. \tag{30}$$

We can indeed prove that such portfolio strategy is admissible and replicates the option's payoff.

Proposition 3.13. *Under Assumption 3.8, the portfolio $(\phi^{(1)}, \phi^{(2)}, C)$ is admissible and replicates the payoff of the American put option.*

Proof. See Section 9. □

From Equations (29) and (30), one can immediately see that if the option holder exercises the option according to the optimal rule (8), no consumption is available to the seller.

4 | NUMERICAL ANALYSIS

In the numerical analysis, we assume that the interest rate r follows Vasicek model. In particular, this means that $\mathcal{I} = \mathbb{R}$ and

$$dr_t = \kappa(\theta - r_t)dt + \beta dW_t, \tag{31}$$

whose explicit solution is given by

$$r_s = r_t e^{-(s-t)\kappa} + \theta(1 - e^{-(s-t)\kappa}) + \beta e^{-s\kappa}\int_t^s e^{\kappa u} dW_u, \quad s \geq t \geq 0. \tag{32}$$

We first derive a numerical method for computing the optimal stopping boundary using the integral equation from Equation (24). Once the boundary is obtained, we use it to also compute the value function via Equation (21). Section 4.2 contains an analysis of the effect of parameters on the stopping boundary and the value function.

4.1 | Computational approach

With an abuse of notation, we denote by $P(t,T) = P(t,r,T)$ the time-t price of a zero-coupon bond with maturity T (c.f. (25)); the dependence on the initial state r is indicated in the subscript of the expectation operator. By Proposition 3.4, we have $b(t,r) = 0$ for $r < 0$, that is, it is never optimal to stop for negative values of the interest rate. To compute $b(t,r)$ for $r \geq 0$, recall the integral equation (24): for $(t,r) \in [0,T) \times \mathcal{I}$ such that $b(t,r) > 0$, we have

$$K - b(t,r) = v_p(t,r,b(t,r);T,b) + v_e(t,r,b(t,r);T), \tag{33}$$

where v_e and v_p are stated in Equation (22). With the last parameter b of v_p, we emphasize the dependence on the function b:

$$v_p(t,r,x;T,b) = \mathsf{E}_{r,x}\left[\int_0^{T-t} e^{-\int_0^u r_v dv} K r_u 1_{\{X_u < b(t+u,r_u)\}} du\right].$$

In the numerical scheme below, we evaluate v_p for consecutive approximations of b.

In Appendix D, we derive the following formulas for v_e and v_p using well-known properties of the joint law of (r_t, X_t):

$$v_e(t,r,x;T) = P(t,T)K\mathcal{N}(d_1) - x\mathcal{N}(d_2), \tag{34}$$

$$v_p(t,r,x;T,b) = \int_t^T KP(t,u)\left[\int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{y^2}{2}} \left(q(t,u) + y\sqrt{\gamma_2(t,u)}\right)\mathcal{N}(\phi(t,u,y;b))dy\right] du, \tag{35}$$

where $\mathcal{N}(\cdot)$ is the cumulative distribution function of the standard normal distribution. An explicit formula for $P(t,T)$ is given by Equation (D.3) and the other auxiliary quantities used above are stated in Equation (D.1).

Equation (33) defines the boundary b as a fixed point of a nonlinear mapping. To compute it, we follow an iterative scheme motivated by Detemple et al. (2018). We fix $-\infty < r_{min} < r_{max} < \infty$ and discretise the variables (t,r) as follows:

$$\{(t_i, r_j) \in [t,T] \times [r_{min}, r_{max}]\}, \quad i = 1, \dots, M, j = 1, \dots, N.$$

We specify an initial approximation $b^{(0)}$ of the boundary:

$$b^{(0)}(t_i, r_j) = K, \quad \forall\, i, j.$$

For each $n \geq 1$, we compute the boundary $b^{(n)}$ at points $(t_i, r_j)_{i,j}$ by solving the algebraic equation:

$$K - b^{(n)}(t_i,r_j) - v_e\big(t_i, r_j, b^{(n)}(t_i,r_j);T\big) = v_p\big(t_i, r_j, b^{(n-1)}(t_i,r_j);T, b^{(n-1)}\big). \tag{36}$$

The right-hand side, which is difficult to compute, is independent of $b^{(n)}$, while the left-hand side is known in an explicit form. We stop iterations when, for a predetermined $\varepsilon > 0$

$$\max_{i,j} |b^{(n-1)}(t_i,r_j) - b^{(n)}(t_i,r_j)| < \varepsilon.$$

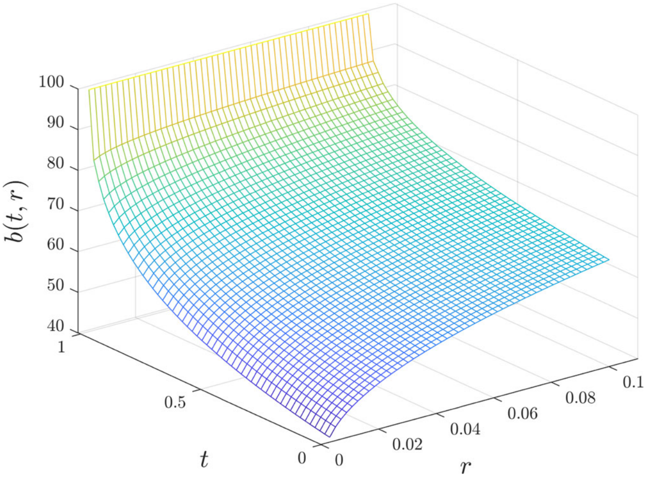


FIGURE 1 Stopping boundary surface $b(t, r)$ [Color figure can be viewed at wileyonlinelibrary.com]

The numerical evaluation of $v_p(t_i, r_j, b^{(n-1)}(t_i, r_j); T, b^{(n-1)})$ requires that the boundary $b^{(n-1)}$ be known for all points (t, r) in the state space while we compute it only on the grid (t_i, r_j). We, therefore, use Matlab interpolation function with the Modified Akima cubic Hermite polynomials ("makima") interpolation method. Integrals are computed using Matlab functions employing standard quadrature methods.

It should be remarked that the stopping boundary b may have a singularity (jump) at $r = 0$, which corresponds to a horizontal part of the parametrization c of the stopping surface: a jump occurs when $c^{-1}(\{0\}) \neq [0, T) \times \{0\}$. Furthermore, $b(T-, r) := \lim_{t \uparrow T} b(t, r)$ satisfies $b(T-, r) = 0$ for $r < 0$ and $b(T-, r) \geq b(0, r) > 0$ for $r > 0$, see Proposition 3.4. This hints at a potential numerical difficulty around $r = 0$, particularly for times t close to maturity.

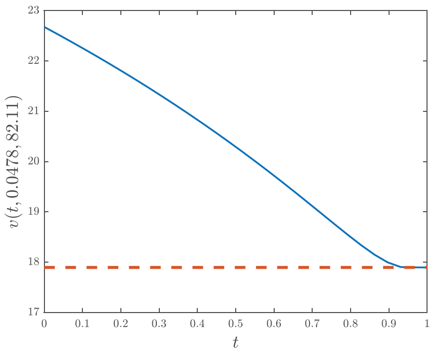
4.2 | Sensitivity analysis

Unless stated otherwise, numerical results are presented for the parameter values

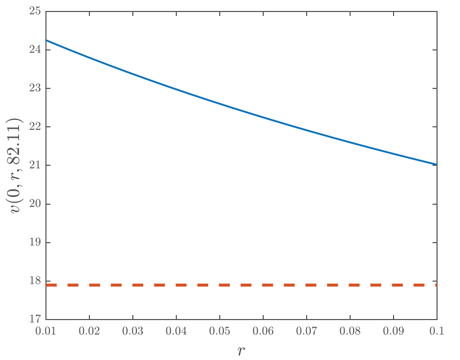
$$T = 1, K = 100, \sigma = 0.4, \kappa = 0.3, \theta = 0.05, \beta = 0.01, \rho = 0.5, \tag{37}$$

and the convergence criterion with $\varepsilon = 0.01$. The magnitude of κ, θ, and β is based on empirical findings reported in the literature, c.f. (Hull, 2009, Chapter 31) and Fergusson and Platen (2015). Although main currencies have recently enjoyed much lower interest rates, our choice of θ means that the effects of random interest rate and its parameters on the market dynamics and optimal stopping boundary are more pronounced and graphs more transparent.

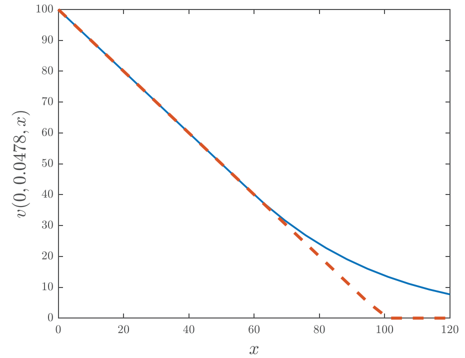
Figure 1 plots the stopping boundary $b(t, r)$ using parameters (37). The optimal stopping boundary increases as t tends to the maturity T and as the interest rate r grows (c.f. Proposition 3.4). This behavior is consistent with the one of the optimal exercise boundary for the American put option in the Black–Scholes model with a constant interest rate (Peskir, 2005). Figure 2 illustrates the value function $v(t, r, x)$ via sections in directions of t, r, and x rooted at the point $(0, 0.0478, 82.11)$, which illustrates the findings of Proposition 3.1. In Panel (Figure 2a), the value decreases to the value of the immediate exercise as the option is purchased deep in the money.



(a) $t \mapsto v(t, 0.0478, 82.11)$



(b) $r \mapsto v(0, r, 82.11)$



(c) $x \mapsto v(0, 0.0478, x)$

FIGURE 2 Sections of the value function $v(t, r, x)$ through the point $(0, 0.0478, 82.11)$. The dashed line displays the payoff $(K - x)^+$ [Color figure can be viewed at wileyonlinelibrary.com]

Effects of the interest rate. The option price is significantly affected by the initial interest rate (Figure 2b) because the maturity of the option is long (1 year). The effect depends on the mean-reversion coefficient κ and it increases when the mean reversion parameter decreases. Indeed, this tendency is clearly visible in Figure 3. A large mean-reversion speed ($\kappa = 1$) means that the interest rate is quickly pulled towards $\theta = 0.05$, diminishing the effect of the initial value. Taking expectation on both sides of Equation (32) gives that the expected interest rate at the maturity $T = 1$ is

$$\mathsf{E}_r[r_1] = re^{-\kappa} + \theta(1 - e^{-\kappa}),$$

which, for $\kappa = 1$, means $\mathsf{E}_r[r_1] \approx 0.36\,r + 0.74\,\theta$. On the contrary, we obtain $\mathsf{E}_r[r_1] \approx 0.90\,r + 0.10\,\theta$ for $\kappa = 0.1$ and so the effect of the initial interest rate on the stopping boundary (Figure 3a) and the value function (Figure 3b) is more pronounced. The optimal strategy for $\kappa = 0.1$ prescribes to be more patient compared to larger values of κ when the interest rate is near 0 and act faster when the interest rate is close to 1. Indeed, with a slow mean-reversion, the interest rate stays close to the current value for longer, so the observed behavior of the stopping boundary and the of value function is akin to that observed by a model with a constant interest rate (Broadie & Detemple, 1996; Peskir, 2005).

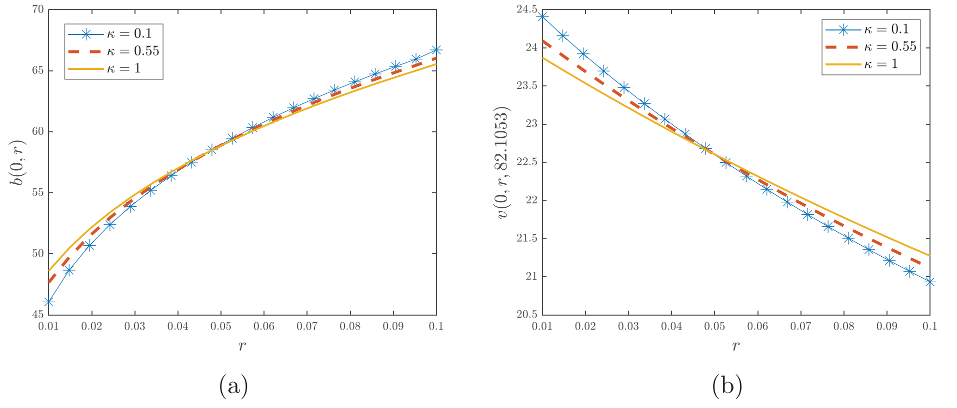


FIGURE 3 The r-sections of the stopping boundary (left panel) and the value function (right panel) for the mean-reversion parameter $\kappa \in \{0.1, 0.55, 1\}$ [Color figure can be viewed at wileyonlinelibrary.com]

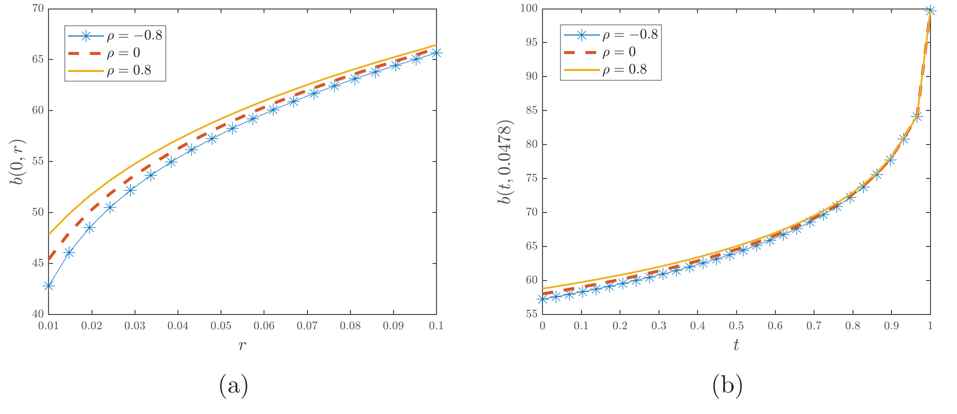


FIGURE 4 The r and t-sections of the stopping boundary for the correlation coefficient $\rho \in \{-0.8, 0, 0.8\}$ [Color figure can be viewed at wileyonlinelibrary.com]

Effects of the correlation coefficient. The sensitivity of the stopping boundary with respect to the correlation coefficient ρ between Brownian motions driving the stock price and the interest rate is displayed in Figure 4; the value function behaves accordingly and it is not displayed. High positive correlation $\rho = 0.8$ implies that the interest rate and the stock price tend to move together. The increase in the interest rate pushes the stock price up and vice versa, resulting in a more unstable environment and an earlier optimal stopping. On the contrary, a strong negative correlation sees the stock price and the interest rate dampening the effect of each other's moves: an increase in the stock price brings a drop in the interest rate, therefore, making longer waiting (lower stopping boundary) more desirable due to effect on the drift of the stock price as well as on the discount factor. Naturally, this effect diminishes the closer one gets to the maturity of the option, see Figure 4b.

Effects of the volatility of stock and interest rate. The effect of the diffusion coefficient of the spot rate β on the stopping boundary and on the value function is negligible. We compared results for $\beta \in \{0.005, 0.01, 0.015\}$, the range of values observed in empirical literature mentioned above. We

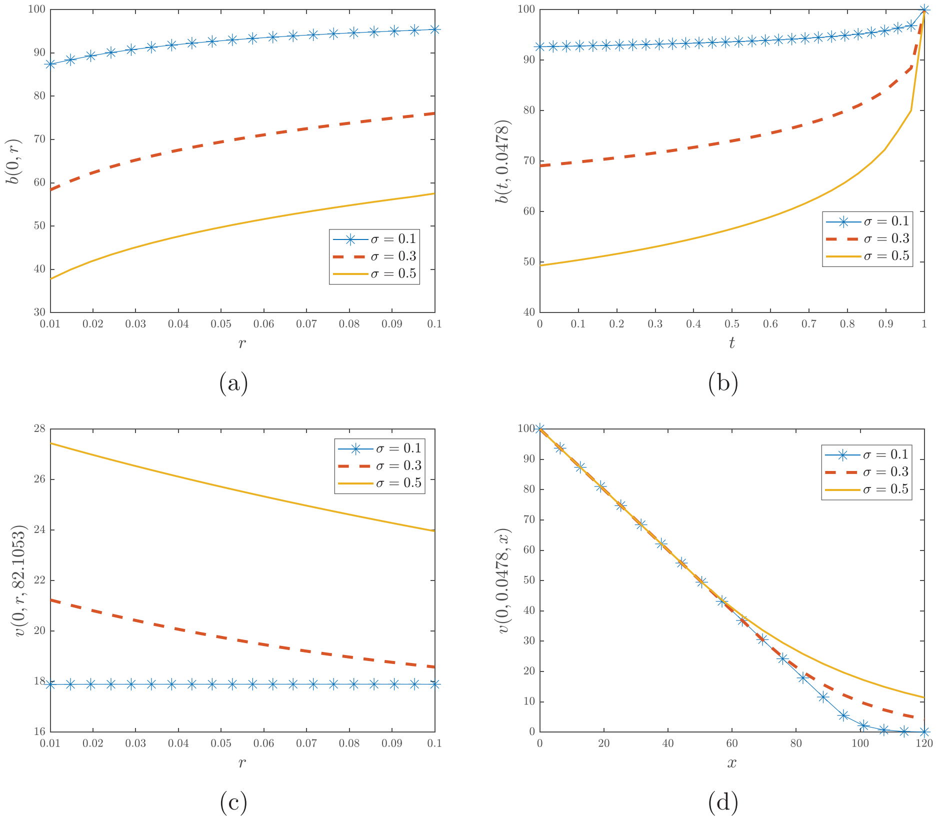


FIGURE 5 Effect of the volatility of the stock price σ. Panels (a) and (b) display the r and t-sections of the stopping boundary $b(t, r)$ and panels (c) and (d) show the r and x-sections of the value function v for $\sigma \in \{0.1, 0.3, 0.5\}$ [Color figure can be viewed at wileyonlinelibrary.com]

noticed variations in the value function of less than 0.1% and in the stopping boundary of less than 1%.

In line with the financial intuition, the value of American put option is increasing in σ, see Figures 5c and 5d. When $\sigma = 0.1$, the optimal stopping boundary is close to the exercise price K (Figure 5a), so the option is immediately exercised for the initial stock price $x = 82.1053$ presented on Panel (c), hence the flat graph. For other values of σ, the exercise boundary is below the initial stock price and the effect of the interest rate is clearly visible. The structure of results in Figure 5 is, as expected, in line with the findings for the American put option in the Black–Scholes model with constant interest rate (Broadie & Detemple, 1996; Peskir, 2005).

The remaining sections of the paper contain technical details and proofs.

5 | MONOTONICITY AND LIPSCHITZ CONTINUITY OF THE OPTION VALUE

In this section, we establish some initial regularity properties of the option value. We start with key monotonicity results and then prove Lipschitz continuity of the value function.

Proof of Proposition 3.1. Finiteness of v follows by Equation (3) and boundedness of the put payoff. Monotonicity in (i) is also a trivial consequence of the fact that the discounted put payoff is independent of time. For (ii) we argue as follows: since $r \mapsto r_t^r$ is increasing P-a.s. for all $t \in [0,T]$ (by uniqueness of the trajectories) we get, for any $\varepsilon > 0$

$$\begin{aligned} v(t, r+\varepsilon, x) &= \sup_{0\le\tau\le T-t} \mathsf{E}\left[\left(Ke^{-\int_0^\tau r_t^{r+\varepsilon}dt} - xe^{\sigma B_\tau - \frac{\sigma^2}{2}\tau}\right)^+\right] \\ &\le \sup_{0\le\tau\le T-t} \mathsf{E}\left[\left(Ke^{-\int_0^\tau r_t^{r}dt} - xe^{\sigma B_\tau - \frac{\sigma^2}{2}\tau}\right)^+\right] = v(t,r,x), \end{aligned}$$

where we took the discounting inside the positive part and used (5).

Finally, monotonicity in (iii) is a simple consequence of monotonicity of Equation (5) with respect to x and the fact that $x \mapsto (K-x)^+$ is decreasing. Convexity also follows by standard arguments: fix $\lambda \in (0,1)$, take x and y in $\mathbb{R}_+$ and denote $x_\lambda := \lambda x + (1-\lambda)y$. By the convexity of the put payoff, using that $X^{r,x_\lambda} = \lambda X^{r,x} + (1-\lambda)X^{r,y}$ and that $\sup(f+g) \le \sup f + \sup g$, it is not hard to verify that $v(t,r,x_\lambda) \le \lambda v(t,r,x) + (1-\lambda)v(t,r,y)$. □

Proposition 5.1 (Lipschitz continuity). *For any compact $\mathcal{K} \subset \mathcal{O}$, there exists a constant $L_\mathcal{K} > 0$ such that*

$$|v(t_1,r_1,x_1) - v(t_2,r_2,x_2)| \le L_\mathcal{K}(|t_1-t_2| + |r_1-r_2| + |x_1-x_2|), \tag{38}$$

for all (t_1,r_1,x_1) and (t_2,r_2,x_2) in $\mathcal{K}$.

Proof of Proposition 5.1. We look separately at Lipschitz continuity in the three variables. Arguments for r and x are quite standard while the main argument for the Lipschitz continuity in t goes back to (Jaillet et al., 1990, Thm. 3.6). However, in our framework, the interest rate is random and the coefficients of the underlying process are state-dependent, which results in some additional difficulties.

Continuity in x. Fix $(t,r) \in [0,T) \times \mathcal{I}$ and take $x_1 \le x_2$ in $\mathbb{R}_+$. Let $\tau_1 := \tau_*(t,r,x_1)$ and note that it is admissible for $v(t,r,x_2)$. Using Proposition 3.1(iii), the explicit expression for $X^{r,x}$ in Equation (5) and the Lipschitz property of the put payoff, we get

$$\begin{aligned} 0 \le v(t,r,x_1) - v(t,r,x_2) &\le \mathsf{E}\left[e^{-\int_0^{\tau_1} r_s^r ds}\left((K - X^{r,x_1})^+ - (K - X^{r,x_2})^+\right)\right] \\ &\le \mathsf{E}\left[e^{\sigma B_{\tau_1} - \frac{\sigma^2}{2}\tau_1}\right](x_2 - x_1) = (x_2 - x_1), \end{aligned}$$

where in the last equality we used Doob's optional sampling theorem.

Continuity in r. Fix $(t,x) \in [0,T) \times \mathbb{R}_+$ and take $r_1 \le r_2$ in $\mathcal{I}$ such that $(t,r_1,x) \in \mathcal{K}$. Denote, for simplicity, $r^1 := r^{r_1}$ and $r^2 := r^{r_2}$ and notice that $r_t^2 \ge r_t^1$ for all $t \ge 0$ P-a.s. Set $\tau_1 := \tau_*(t,r_1,x)$.

From Proposition 3.1(ii) and simple estimates we obtain

$$
\begin{aligned}
0 \leq v(t,r_1,x)-v(t,r_2,x) &\leq K\mathsf{E}\left[e^{-\int_0^{\tau_1} r_s^1 ds}-e^{-\int_0^{\tau_1} r_s^2 ds}\right]\\
&= K\mathsf{E}\left[e^{-\int_0^{\tau_1} r_s^1 ds}\left(1-e^{-\int_0^{\tau_1}(r_s^2-r_s^1)ds}\right)\right]\\
&\leq K\mathsf{E}\left[e^{-\int_0^{\tau_1} r_s^1 ds}\int_0^{\tau_1}(r_s^2-r_s^1)ds\right].
\end{aligned}
\tag{39}
$$

To complete the proof, we consider separately cases (i) and (ii) in Assumption 2.1. Let us start with (i): using that $r_t^1 \geq 0$ for $t \geq 0$, and the explicit form of the SDE in the CIR model, we get

$$
\begin{aligned}
\mathsf{E}\left[e^{-\int_0^{\tau_1} r_s^1 ds}\int_0^{\tau_1}(r_s^2-r_s^1)ds\right] &\leq \int_0^{T-t}\mathsf{E}\left[r_s^2-r_s^1\right]ds\\
&= \int_0^{T-t}\mathsf{E}\left[(r_2-r_1)+\int_0^s \kappa(r_u^1-r_u^2)du\right]ds \leq (T-t)(r_2-r_1),
\end{aligned}
$$

where we have used the integral equation for (r_t) and that $r_t^2 \geq r_t^1$.

If Assumption 2.1(ii) holds instead, we apply Hölder inequality:

$$
\begin{aligned}
\mathsf{E}\left[e^{-\int_0^{\tau_1} r_s^1 ds}\int_0^{\tau_1}(r_s^2-r_s^1)ds\right] &\leq \left(\mathsf{E}\left[e^{-2\int_0^{\tau_1} r_s^1 ds}\right]\right)^{\frac{1}{2}}\left(\mathsf{E}\left[\left(\int_0^{T-t}(r_s^2-r_s^1)ds\right)^2\right]\right)^{\frac{1}{2}}\\
&\leq C_1^{1/2}\left((T-t)\int_0^{T-t}\mathsf{E}\left[(r_s^2-r_s^1)^2\right]ds\right)^{\frac{1}{2}},
\end{aligned}
\tag{40}
$$

where $C_1 > 0$ is the constant from Equation (3), which depends on $\mathcal{K}$. To conclude, it is sufficient to use moment estimates for SDEs (Krylov, 1980, Ch. 2, Sec. 5, Thm. 9) which guarantee that

$$
\mathsf{E}\left[\sup_{0\leq s\leq T}(r_s^2-r_s^1)^2\right] \leq c'(r_2-r_1)^2, \tag{41}
$$

for some $c' > 0$ only depending on T and the coefficients in Equation (2).

Continuity in t. For $t \in [0,T)$, define $r_u^{T-t} := r_{u(T-t)}$ and $X_u^{T-t} := X_{u(T-t)}$ for $u \in [0,1]$. The couple $(r_u^{T-t}, X_u^{T-t})_{u\in[0,1]}$ is a strong solution to (see, e.g., Bass (1998, Ch. 1, Prop. 8.6))

$$
\begin{aligned}
dX_u^{T-t} &= (T-t)r_u^{T-t}X_u^{T-t}du + \sigma X_u^{T-t}d\widetilde{B}_u, & X_0^{T-t} &= x,\\
dr_u^{T-t} &= (T-t)\alpha(r_u^{T-t})du + \beta(r_u^{T-t})d\widetilde{W}_u, & r_0^{T-t} &= r,
\end{aligned}
$$

where $(\widetilde{B}_u, \widetilde{W}_u)_{u\in[0,1]} := (B_{u(T-t)}, W_{u(T-t)})_{u\in[0,1]}$. Using these processes, we can rewrite Equation (6) as

$$
v(t,r,x) = \sup_{0\leq\theta\leq 1}\mathsf{E}_{r,x}\left[\exp\left\{-(T-t)\int_0^{\theta} r_u^{T-t}du\right\}\left(K-X_\theta^{T-t}\right)^+\right], \tag{42}
$$

where for any $(\mathcal{F}_s)_{s\geq 0}$-stopping time τ in $[0,T-t]$, the random variable $\theta := \tau/(T-t)$ is an $(\mathcal{F}_{u(T-t)})_{u\in[0,1]}$-stopping time. Since the process $(B_{u(T-t)}, W_{u(T-t)})_{u\in[0,1]}$ is identical in law to

$(\sqrt{T-t}B_u, \sqrt{T-t}W_u)_{u\in[0,1]}$, with a slight abuse of notation, we can identify $(r_u^{T-t}, X_u^{T-t})_{u\in[0,1]}$ with the unique strong solution of

$$dX_u^{T-t} = (T-t)r_u^{T-t}X_u^{T-t}du + \sqrt{T-t}\sigma X_u^{T-t}dB_u, \qquad X_0^{T-t} = x, \tag{43}$$

$$dr_u^{T-t} = (T-t)\alpha(r_u^{T-t})du + \sqrt{T-t}\beta(r_u^{T-t})dW_u, \qquad r_0^{T-t} = r, \tag{44}$$

and take stopping times $\theta \in [0,1]$ in Equation (42) with respect to the filtration $(\mathcal{F}_t)$ generated by (B, W). In what follows, we denote by $\theta_* = \theta_*(t, r, x)$ an optimal stopping time for Equation (42).

Fix now $0 \le t_1 < t_2 < T$ and set $r^1 := r^{T-t_1}$, $r^2 := r^{T-t_2}$. Let $\theta_1 := \theta_*(t_1, r, x)$ and for $i = 1, 2$ denote also

$$R_u^i = (T-t_i)\int_0^u r_s^i ds \quad \text{and} \quad \hat{X}_u^{T-t} = \exp\left(\sqrt{T-t}\,\sigma B_u - (T-t)\frac{\sigma^2}{2}u\right),$$

so that $X_u^{T-t_i} = xe^{-R_u^i}\hat{X}_u^{T-t_i}$. We remark that θ_1 is also admissible for the problem in Equation (42) and the underlying dynamics (43)–(44) with $t = t_2$, because it is an $(\mathcal{F}_s)_{s\ge 0}$-stopping time in [0,1]. Indeed the advantage of Equation (42) with Equations (43)–(44) is that the class of admissible stopping times no longer depends on the initial time t.

Recalling Proposition 3.1(i) and using Lipschitz continuity of $x \mapsto (x)^+$ we have

$$\begin{aligned} 0 \ge v(t_2, r, x) - v(t_1, r, x) &\ge -\mathsf{E}_r\left[\left|\left(Ke^{-R^2_{\theta_1}} - x\hat{X}_{\theta_1}^{T-t_2}\right)^+ - \left(Ke^{-R^1_{\theta_1}} - x\hat{X}_{\theta_1}^{T-t_1}\right)^+\right|\right] \\ &\ge -K\mathsf{E}_r\left[\left|e^{-R^2_{\theta_1}} - e^{-R^1_{\theta_1}}\right|\right] - x\mathsf{E}\left[\left|\hat{X}_{\theta_1}^{T-t_1} - \hat{X}_{\theta_1}^{T-t_2}\right|\right]. \end{aligned} \tag{45}$$

Let us consider the second term on the right-hand side of Equation (45). By the fundamental theorem of calculus and the explicit formula for $\hat{X}^{T-t}$

$$\begin{aligned} \mathsf{E}\left[\left|\hat{X}_{\theta_1}^{T-t_1} - \hat{X}_{\theta_1}^{T-t_2}\right|\right] &= \mathsf{E}\left[\left|\int_{t_1}^{t_2} \hat{X}_{\theta_1}^{T-t}\left(\frac{\sigma^2}{2}\theta_1 - \frac{1}{2\sqrt{T-t}}\sigma B_{\theta_1}\right)dt\right|\right] \\ &\le \int_{t_1}^{t_2} \mathsf{E}\left[\left|\hat{X}_{\theta_1}^{T-t}\left(\frac{\sigma^2}{2}\theta_1 - \frac{1}{2\sqrt{T-t}}\sigma B_{\theta_1}\right)\right|\right]dt. \end{aligned} \tag{46}$$

For $t \in (t_1, t_2)$, define a measure $\tilde{\mathsf{P}}$ by $\frac{d\tilde{\mathsf{P}}}{d\mathsf{P}} := \hat{X}_1^{T-t}$. Then $\tilde{B}_s = B_s - \sigma s\sqrt{T-t}$ is a Brownian motion under $\tilde{\mathsf{P}}$ and

$$\begin{aligned} \mathsf{E}\left[\left|\hat{X}_{\theta_1}^{T-t}\left(\frac{\sigma^2}{2}\theta_1 - \frac{1}{2\sqrt{T-t}}\sigma B_{\theta_1}\right)\right|\right] &= \tilde{\mathsf{E}}\left[\left|\frac{\theta_1}{2}\sigma^2 - \frac{\sigma}{2\sqrt{T-t}}(\tilde{B}_{\theta_1} + \sqrt{T-t}\sigma\theta_1)\right|\right] \\ = \tilde{\mathsf{E}}\left[\left|\frac{\sigma}{2\sqrt{T-t}}\tilde{B}_{\theta_1}\right|\right] \le \left(\tilde{\mathsf{E}}\left[\frac{\sigma^2\tilde{B}^2_{\theta_1}}{4(T-t)}\right]\right)^{1/2} &\le \frac{\sigma}{2\sqrt{T-t}} \le \frac{\sigma}{2\sqrt{T-t_2}} =: c_1, \end{aligned}$$

where we applied Hölder inequality and used that $\theta_1 \leq 1$. Inserting the above estimate into Equation (46) gives

$$\mathsf{E}\left[\left|\hat{X}^{T-t_1}_{\theta_1} - \hat{X}^{T-t_2}_{\theta_1}\right|\right] \leq c_1(t_2 - t_1). \tag{47}$$

Next, we address the first term on the right-hand side of Equation (45). This is performed separately in cases (i) and (ii) of Assumption 2.1. We start by considering case (ii), that is, α and β in Equation (44) are Lipschitz continuous. Fundamental theorem of calculus and Hölder inequality give

$$\begin{aligned}
&\mathsf{E}_r\left[\left|e^{-R^1_{\theta_1}} - e^{-R^2_{\theta_1}}\right|\right] \\
&\leq \mathsf{E}_r\left[\max_{i=1,2}\left\{e^{-(T-t_i)\int_0^{\theta_1} r^i_u du}\right\}\left|(T-t_1)\int_0^{\theta_1} r^1_u du - (T-t_2)\int_0^{\theta_1} r^2_u du\right|\right] \\
&\leq \mathsf{E}_r\left[\max_{i=1,2}\left\{e^{-(T-t_i)\int_0^{\theta_1} r^i_u du}\right\}\left((t_2-t_1)\left|\int_0^{\theta_1} r^1_u du\right| + (T-t_2)\left|\int_0^{\theta_1}(r^2_u - r^1_u)du\right|\right)\right] \\
&\leq 2c_2\left[(t_2-t_1)\left(\mathsf{E}_r\left[\sup_{0\leq t\leq 1}\left(r^1_t\right)^2\right]\right)^{\frac{1}{2}} + (T-t_2)\left(\mathsf{E}_r\left[\int_0^1 (r^2_u - r^1_u)^2 du\right]\right)^{\frac{1}{2}}\right],
\end{aligned} \tag{48}$$

where, using Equation (3),

$$c_2 := \sup_{(t,r,x)\in\mathcal{K}}\left(\mathsf{E}_r\left[\sup_{0\leq s\leq 1} e^{-2(T-t)\int_0^s r^{T-t}_u du}\right]\right)^{\frac{1}{2}} < \infty.$$

Thanks to Equation (4), $c_3 := \sup_{(t,r,x)\in\mathcal{K}}(\mathsf{E}_r[\sup_{0\leq s\leq 1}(r^{T-t}_s)^2])^{\frac{1}{2}} < \infty$, so it remains to estimate the last term of Equation (48). By (Krylov, 1980, Ch. 2, Sec. 5, Thm. 9), there is a constant c_4 depending only on $\mathcal{K}$ and the Lipschitz constant for α and β in Equation (44) such that

$$\begin{aligned}
&\mathsf{E}_r\left[\sup_{0\leq t\leq 1}\left(r^1_t - r^2_t\right)^2\right] \\
&\leq c_4\,\mathsf{E}_r\left[\int_0^1\left(|(T-t_1)\alpha(r^1_u) - (T-t_2)\alpha(r^1_u)|^2 + |\sqrt{T-t_1}\beta(r^1_u) - \sqrt{T-t_2}\beta(r^1_u)|^2\right)du\right] \\
&\leq c_4(t_2-t_1)^2\,\mathsf{E}_r\left[\int_0^1 |\alpha(r^1_u)|^2 du\right] + c_4\,(t_2-t_1)\,\mathsf{E}_r\left[\int_0^1 |\beta(r^1_u)|^2 du\right],
\end{aligned}$$

where for the second inequality we used that $\sqrt{T-t_1} - \sqrt{T-t_2} \leq \sqrt{t_2 - t_1}$. Notice that by Equation (4) and the linear growth of α and β

$$c_5 := \sup_{(r,t,x)\in\mathcal{K}}\mathsf{E}_r\left[\int_0^1 |\alpha(r^{T-t}_u)|^2 + |\beta(r^{T-t}_u)|^2 du\right] < \infty.$$

Inserting the above estimates into Equation (48), we conclude that there is a constant c_6 such that for any $(t_1, r, x), (t_2, r, x) \in \mathcal{K}$

$$\mathsf{E}_r\left[\left|e^{-R^1_{\theta_1}} - e^{-R^2_{\theta_1}}\right|\right] \leq c_6|t_2 - t_1|.$$

This and Equation (47) feed into Equation (45) so that

$$0 \geq v(t_2, r, x) - v(t_1, r, x) \geq -c|t_2 - t_1|, \tag{49}$$

for a suitable $c > 0$ that depends on $\mathcal{K}$.

Finally, we must estimate the first term on the right-hand side of Equation (45) under the assumption that $(r_t)_{t\geq 0}$ follows the CIR dynamics (Assumption 2.1(i)). Let $\hat{r}^i_u := r^i_u/(T-t_i)$ for $u \in [0,1]$ and $i = 1,2$. The dynamics for $\hat{r}^i$ reads

$$d\hat{r}^i_u = \kappa\big(\alpha - (T-t_i)\hat{r}^i_u\big)du + \beta\sqrt{\hat{r}^i_u}dW_u, \quad u \in [0,1]. \tag{50}$$

Since $\kappa(\alpha - (T-t_1)\hat{r}) < \kappa(\alpha - (T-t_2)\hat{r})$ for $\hat{r} \geq 0$, and $\hat{r}^1_0 = r/(T-t_1) \leq r/(T-t_2) = \hat{r}^2_0$, comparison results for SDEs (Karatzas & Shreve, 1998a, Prop. 5.2.18) imply

$$\hat{r}^1_u \leq \hat{r}^2_u \quad \text{for all } u \in [0,1], \text{ P-a.s.} \tag{51}$$

Using the integral version of Equation (50) and the martingale property of the stochastic integral, we obtain

$$\begin{aligned} \mathsf{E}_r\big[\hat{r}^2_u - \hat{r}^1_u\big] &= r\Big(\tfrac{1}{T-t_2} - \tfrac{1}{T-t_1}\Big) + \mathsf{E}_r\Big[\int_0^u \big((T-t_1)\hat{r}^1_s - (T-t_2)\hat{r}^2_s\big)ds\Big] \\ &\leq r\tfrac{t_2-t_1}{(T-t_1)(T-t_2)} + (t_2-t_1)\int_0^1 \mathsf{E}_r\big[\hat{r}^1_s\big]ds + (T-t_2)\int_0^u \mathsf{E}_r\big[\hat{r}^1_s - \hat{r}^2_s\big]ds. \end{aligned}$$

Due to Equation (51), the last term is nonpositive, so

$$0 \leq \mathsf{E}_r\big[\hat{r}^2_u - \hat{r}^1_u\big] \leq (t_2-t_1)\left(\frac{r}{(T-t_1)(T-t_2)} + q_1\right) \quad \text{for all } u \in [0,1] \tag{52}$$

where

$$q_1 := \sup_{(t,r,x)\in\mathcal{K}} \frac{1}{T-t}\int_0^1 \mathsf{E}_r\big[r^{T-t}_u\big]du < \infty.$$

We use the inequalities (51)–(52) and the property that $\hat{r}^i_u \geq 0$, for $i = 1,2$, to obtain the following estimates

$$\begin{aligned} \mathsf{E}_r\left[\left|e^{-R^1_{\theta_1}} - e^{-R^2_{\theta_1}}\right|\right] &= \mathsf{E}_r\left[\left|e^{-(T-t_1)^2\int_0^{\theta_1}\hat{r}^1_u du} - e^{-(T-t_2)^2\int_0^{\theta_1}\hat{r}^2_u du}\right|\right] \\ &\leq \mathsf{E}_r\left[\left|e^{-(T-t_1)^2\int_0^{\theta_1}\hat{r}^1_u du} - e^{-(T-t_2)^2\int_0^{\theta_1}\hat{r}^1_u du}\right|\right] \\ &\quad + \mathsf{E}_r\left[\left|e^{-(T-t_2)^2\int_0^{\theta_1}\hat{r}^1_u du} - e^{-(T-t_2)^2\int_0^{\theta_1}\hat{r}^2_u du}\right|\right] \\ &\leq q_1\big((T-t_1)^2 - (T-t_2)^2\big) + (T-t_2)^2\int_0^1 \mathsf{E}_r\big[\hat{r}^2_u - \hat{r}^1_u\big]du \\ &\leq (t_2-t_1)\left(2Tq_1 + r\frac{T-t_2}{T-t_1} + q_1(T-t_2)^2\right) \leq c_7(t_2-t_1), \end{aligned} \tag{53}$$

where the constant $c_7 > 0$ depends only on $\mathcal{K}$ but not on a specific choice of t_1, t_2, r, x. Hence, as in the case of Assumption 2.1(ii), we obtain Equation (49). □

6 | PROPERTIES OF THE FREE BOUNDARY

This section is devoted to establishing the existence of an optimal stopping boundary (free boundary) and some of its main properties. In particular, we show the so-called "regularity" of the stopping boundary in the sense of diffusion theory which, together with the monotonicity, is instrumental in our proof of global C^1 regularity of the value function v.

Proof of Proposition 3.3. The payoff does not depend on (r_t) and v is nonincreasing in r by Proposition 3.1. Therefore, if $(t, r_1, x) \in \mathcal{D}$, then $(t, r_2, x) \in \mathcal{D}$ for for any $r_2 > r_1$. This allows us to represent the stopping region $\mathcal{D}$ via Equation (15) with

$$c(t,x) := \inf\{r \in \mathcal{I} : v(t,r,x) = (K-x)^+\}, \tag{54}$$

with the convention that $\inf \emptyset = \bar{r}$. It is convenient to prove (ii) first.

(ii) Fix $(t,r,x) \in [0,T) \times \mathcal{I} \times [K,\infty)$. If we show that $\mathsf{P}_{r,x}(X_\varepsilon < K) > 0$ for some $\varepsilon \in (0, T-t]$, then $v(t,r,x) > 0 = (K-x)^+$. This means that $(t,r,x) \in \mathcal{C}$ and $c(t,x) = \bar{r}$. Recall that $\rho \in (-1,1)$ is the correlation coefficient between the Brownian motions B and W driving the SDEs for X and r, respectively. Then, we can write $B_t = \rho W_t + \sqrt{1-\rho^2} B^0_t$ for some other Brownian motion B^0 independent of W. Letting $(\mathcal{F}^W_t)_{t \ge 0}$ be the filtration generated by W, using the explicit form of the dynamics of X we have

$$\begin{aligned} \mathsf{P}_{r,x}(X_\varepsilon < K) &= \mathsf{E}_{r,x}\left[\mathsf{P}_{r,x}(X_\varepsilon < K | \mathcal{F}^W_\varepsilon)\right] \\ &= \mathsf{E}_{r,x}\left[\mathsf{P}_r\left(\exp\left(\sigma\sqrt{1-\rho^2}B^0_\varepsilon\right) < (K/x)\exp\left(-\sigma\rho W_\varepsilon - \int_0^\varepsilon r_t dt + \frac{\sigma^2}{2}\varepsilon\right)\Big|\mathcal{F}^W_\varepsilon\right)\right] \\ &= \mathsf{E}_{r,x}\left[\Psi_x\left(\sigma\rho W_\varepsilon + \int_0^\varepsilon r_t dt - \frac{\sigma^2}{2}\varepsilon\right)\right], \end{aligned} \tag{55}$$

where

$$\Psi_x(z) := \mathsf{P}\Big(\exp\left(\sigma\sqrt{1-\rho^2}B^0_\varepsilon\right) < (K/x)e^{-z}\Big),$$

and the final equality above holds by the independence of B^0_ε from $\mathcal{F}^W_\varepsilon$ and the fact that $(W_\varepsilon, \int_0^\varepsilon r_t dt)$ is $\mathcal{F}^W_\varepsilon$-measurable. Since $\rho \in (-1,1)$, then $\Psi_x(z) > 0$ for any $z \in \mathbb{R}$ and we conclude that $\mathsf{P}_{r,x}(X_\varepsilon < K) > 0$.

(i) By the monotonicity of v in t, we have $(t_1, r, x) \in \mathcal{D} \Rightarrow (t_2, r, x) \in \mathcal{D}$ for any $t_2 > t_1$, hence $c(t,x)$ is nonincreasing in t.

Fix $0 \le x_1 < x_2 < K$ and let $\tau_1 := \tau_*(t,r,x_1)$ be optimal for $v(t,r,x_1)$. Then, using that $X^{r,x_1} \le X^{r,x_2}$ and recalling Equation (5), we obtain

$$\begin{aligned} v(t,r,x_2) - v(t,r,x_1) &\ge \mathsf{E}\left[e^{-\int_0^{\tau_1} r_s ds}\left(\left(K - X^{r,x_2}_{\tau_1}\right)^+ - \left(K - X^{r,x_1}_{\tau_1}\right)^+\right)\right] \\ &\ge \mathsf{E}\left[e^{-\int_0^{\tau_1} r_s ds}\left(X^{r,x_1}_{\tau_1} - X^{r,x_2}_{\tau_1}\right)\right] \\ &= x_1 - x_2 = (K-x_2)^+ - (K-x_1)^+. \end{aligned}$$

Therefore, if $(t,r,x_1) \in \mathcal{C}$ then $(t,r,x_2) \in \mathcal{C}$, which implies that $c(t,x)$ is nondecreasing in x.

Fix arbitrary $(t,x) \in [0,T) \times \mathbb{R}_+$, let $t_n \downarrow t_0$ as $n \to \infty$, then $c(t_n, x) \uparrow c(t_0+, x)$ as $n \to \infty$, where the limit exists by the monotonicity of $t \mapsto c(t,x)$. Since $(t_n, c(t_n,x), x) \in \mathcal{D}$, then also

$(t_0, c(t_0+, x), x) \in \mathcal{D}$ by the closedness of $\mathcal{D}$, hence $c(t_0+, x) \geq c(t_0, r)$, which implies $c(t_0+, r) = c(t_0, r)$. Taking $x_n \uparrow x_0$, a similar argument yields $c(t, x_0-) = c(t, x_0)$.

(iii) Under the CIR model, the positivity follows by the definition of $c(t, x)$. Only under Assumption 2.1 (ii), a proof is required. Assume that there exists $(t_0, \hat{x}) \in [0, T) \times (0, K)$ such that $c(t_0, \hat{x}) < 0$. Let $0 > r_2 > r_0 > r_1 > c(t_0, \hat{x})$ and $0 < x_0 < \hat{x}$. Define a stopping time

$$\tau_1 = \inf\{s \geq 0 : (s, r_s, X_s) \notin [0, T - t_0) \times (r_1, r_2) \times (0, \hat{x})\}.$$

By the monotonicity of $c(t, x)$, we have $(t_0, r_0, x_0) \in \mathcal{D}$. Hence, τ_1 is suboptimal and

$$K - x_0 = v(t_0, r_0, x_0) \geq \mathsf{E}_{r_0, x_0}\left[e^{-\int_0^{\tau_1} r_s ds}\left(K - X_{\tau_1}\right)^+\right] \geq K\mathsf{E}_{r_0, x_0}\left[e^{-\int_0^{\tau_1} r_s ds}\right] - x_0, \tag{56}$$

where the last inequality follows from the optional sampling theorem and the fact that $(K - X_{\tau_1})^+ \geq K - X_{\tau_1}$. Since $\mathsf{P}_{x_0, r_0}(\tau_1 > 0) = 1$ and $r_s(\omega) < r_2 < 0$ for $s \in [0, \tau_1(\omega))$, we obtain

$$K\mathsf{E}_{r_0, x_0}\left[e^{-\int_0^{\tau_1} r_s ds}\right] - x_0 > K - x_0,$$

which, in conjunction with Equation (56), leads to a contradiction.

Finally, we show that $c(t, 0+) := \lim_{x \downarrow 0} c(t, x) = 0$ for any $t \in [0, T)$. Assume $c(t, 0+) \geq \delta > 0$ for some $t \in [0, T)$. By the monotonicity of $c(t, x)$ and the openness of $\mathcal{C}$, there is $\hat{t} \in [t, T)$ such that

$$[0, \hat{t}) \times (r_1, r_2) \times (0, \infty) \subset \mathcal{C},$$

where $0 < r_1 < r_2 < \delta$. Fix $0 \leq t_0 < \hat{t}$ and $r_0 \in (r_1, r_2)$. Take an arbitrary $x_0 > 0$. Let

$$\tau_2 = \inf\{s \geq 0 : (s, r_s) \notin [0, \hat{t} - t_0) \times (r_1, r_2)\}.$$

By construction $\mathsf{P}_{r_0, x_0}((t_0 + s, r_s, X_s) \in \mathcal{C} \text{ for } s \leq \tau_2) = 1$, so $\tau_2 \leq \tau_*(t_0, r_0, x_0)$ P_{r_0, x_0}-a.s. By the martingale property of the value function, we obtain

$$\begin{aligned} K - x_0 < v(t_0, r_0, x_0) &= \mathsf{E}_{r_0, x_0}\left[e^{-\int_0^{\tau_2} r_s ds} v\left(t_0 + \tau_2, r_{\tau_2}, X_{\tau_2}\right)\right] \\ &\leq K\mathsf{E}_{r_0, x_0}[e^{-r_1 \tau_2}] = K\mathsf{E}_{r_0}[e^{-r_1 \tau_2}]. \end{aligned} \tag{57}$$

A contradiction is obtained by taking the limit $x_0 \downarrow 0$, since $\mathsf{E}_{r_0}[e^{-r_1 \tau_2}]$ is independent of X and strictly smaller than 1. □

An important consequence of Proposition 3.4 is that for $\varepsilon \in (0, x)$

$$(t, r, x) \in \mathcal{D} \Rightarrow (t + \varepsilon, r, x), (t, r + \varepsilon, x), (t, r, x - \varepsilon) \in \mathcal{D}.$$

We immediately see that $\partial\mathcal{C}$ enjoys the so-called *cone property* (Karatzas and Shreve, 1998a, Def. 4.2.18). Indeed, for any $(t_0, r_0, x_0) \in \partial\mathcal{C}$, there is an orthant $\widehat{C}_0$ with vertex in (t_0, r_0, x_0) (hence a cone with aperture $\pi/4$) that satisfies $\widehat{C}_0 \cap \mathcal{O} \subseteq \mathcal{D}$. This will be used to establish regularity of the boundary $\partial\mathcal{C}$ in the sense of diffusions, which, has important consequences for the smoothness of our value function v, as we shall see below.

To this end, we introduce the hitting time to $\mathcal{D}$, denoted $\sigma_{\mathcal{D}}$, and the entry time to the interior of $\mathcal{D}$, denoted $\mathring{\sigma}_{\mathcal{D}}$. That is, for $(t, r, x) \in \mathcal{O}$, we set $\mathsf{P}_{r,x}$-a.s.

$$\begin{aligned}\sigma_{\mathcal{D}} &:= \inf\{s > 0 \ : \ (t+s, r_s, X_s) \in \mathcal{D}\},\\ \mathring{\sigma}_{\mathcal{D}} &:= \inf\{s \geq 0 \ : \ (t+s, r_s, X_s) \in \text{int}(\mathcal{D})\} \wedge (T-t).\end{aligned} \tag{58}$$

Both $\sigma_{\mathcal{D}}$ and $\mathring{\sigma}_{\mathcal{D}}$ are stopping times with respect to the filtration $(\mathcal{F}_t)_{t\geq 0}$. We will often write $\sigma_{\mathcal{D}}(t, r, x)$ and $\mathring{\sigma}_{\mathcal{D}}(t, r, x)$ to indicate the starting point of the process.

Proposition 6.1 (Regularity of the boundary). *For $(t_0, r_0, x_0) \in \partial\mathcal{C}$, we have*

$$\mathsf{P}_{t_0,r_0,x_0}(\sigma_{\mathcal{D}} > 0) = \mathsf{P}_{t_0,r_0,x_0}(\mathring{\sigma}_{\mathcal{D}} > 0) = 0. \tag{59}$$

The proof can be found in Appendix B. It rests on Gaussian bounds for the transition density of a diffusion and ideas from the proof of well-known analogous results for multidimensional Brownian motion, see for example, Karatzas and Shreve (1998a, Thm. 4.2.19). It is also worth recalling that $\partial\mathcal{C}$ is the boundary of $\mathcal{C}$ in $\mathcal{O}$, so that it excludes $\{T\} \times \mathcal{I} \times \mathbb{R}_+$.

7 | CONTINUOUS DIFFERENTIABILITY OF THE OPTION VALUE

We start by establishing the following continuity properties of processes r and X.

Lemma 7.1. *Let $(r_n, x_n)_{n\geq 1}$ be a sequence converging to $(r, x) \in \mathcal{I} \times \mathbb{R}_+$ as $n \to \infty$. Then*

$$\lim_{n\to+\infty} \sup_{0\leq t\leq T} \left| r_t^{r_n} - r_t^r \right| = 0, \qquad \mathsf{P}\text{-}a.s. \tag{60}$$

$$\lim_{n\to+\infty} \sup_{0\leq t\leq T} \left| X_t^{r_n,x_n} - X_t^{r,x} \right| = 0, \qquad \mathsf{P}\text{-}a.s. \tag{61}$$

Proof of Lemma 7.1. Assume first that $(r_n)_{n\geq 1}$ is a monotone sequence. Define $f_t^n := r_t^{r_n} - r_t^r$. Then for a.e. $\omega \in \Omega$, $t \mapsto f_t^n(\omega)$ is continuous and $f_t^n(\omega)$ converges to 0 monotonically as $n \to \infty$ for all $t \in [0, T]$. Hence the convergence is uniform on $[0, T]$ thanks to Dini's theorem and Equation (60) holds.

For an arbitrary sequence, $(r_n)_{n\geq 1}$ define monotone sequences $\bar{r}_n = \sup_{k\geq n} r_k$ and $\underline{r}_n = \inf_{k\geq n} r_k$. Since $r_t^{\underline{r}_n} - r_t^r \leq r_t^{r_n} - r_t^r \leq r_t^{\bar{r}_n} - r_t^r$, we have

$$0 \leq \sup_{0\leq t\leq T} \left| r_t^{r_n} - r_t^r \right| \leq \sup_{0\leq t\leq T} \left| r_t^{\underline{r}_n} - r_t^r \right| + \sup_{0\leq t\leq T} \left| r_t^{\bar{r}_n} - r_t^r \right|.$$

By virtue of the first part of the proof, the terms on the right-hand side converge to 0 as $n \to \infty$, which proves Equation (60). The verification of Equation (61) is easy using the representation formula (5) for X and Equation (60). □

Lemma 7.2. *Let $(t_n, r_n, x_n)_{n\geq 1}$ be a sequence in $\mathcal{C}$ converging to $(t, r, x) \in \overline{\mathcal{C}} \cap \mathcal{O}$ as $n \to \infty$. Then*

$$\lim_{n\to\infty} \tau_*(t_n, r_n, x_n) = \tau_*(t, r, x), \qquad \mathsf{P}\text{-}a.s.$$

Proof of Lemma 7.2. The proof relies on known facts from the theory of Markov processes, which we summarize in Appendix A for the reader's convenience, combined with Proposition 6.1. Proposition 6.1 and Lemma 7.1 imply that Assumptions A.1 and A.2 are satisfied for $\mathcal{K} = \mathcal{D} \cap \mathcal{O}$. It is also immediate to see that $\sigma_{\mathcal{D}} = \sigma_{\mathcal{K}}$ P-a.s. with $\sigma_{\mathcal{K}}$ defined in Equation (A.1).

The continuity of trajectories of (r, X) means that the process cannot jump instantaneously to the stopping set $\mathcal{D}$ when starting from $\mathcal{C}$, so $\mathsf{P}_{\hat{t},\hat{r},\hat{x}}(\tau_* = \sigma_{\mathcal{D}}) = 1$ for any $(\hat{t}, \hat{r}, \hat{x}) \in \mathcal{C}$. When $(\hat{t}, \hat{r}, \hat{x}) \in \partial\mathcal{C}$, by construction we have $\tau_*(\hat{t}, \hat{r}, \hat{x}) = 0$, P-a.s., and, using Proposition 6.1, $\sigma_{\mathcal{D}}(\hat{t}, \hat{r}, \hat{x}) = 0$, P-a.s. Recalling that $\overline{\mathcal{C}} \cap \mathcal{O} = \mathcal{C} \cup \partial\mathcal{C}$, the claim then follows from Proposition A.6. □

Next, we provide gradient estimates based on probabilistic arguments.

Proposition 7.3. *Let $\mathcal{K} \subset \mathcal{O}$ be a compact set with nonempty interior. There is $L = L(\mathcal{K}) > 0$ such that for any $(t, r, x) \in (int(\mathcal{K}) \setminus \partial\mathcal{C})$ we have*

$$v_x(t,r,x) = -\mathsf{E}_{t,r,x}\left[1_{\{X_{\tau_*} \le K\}} e^{\sigma B_{\tau_*} - \frac{\sigma^2}{2}\tau_*}\right], \tag{62}$$

$$0 \ge v_t(t,r,x) \ge -L\, \mathsf{E}_{t,r,x}\left[e^{-\int_0^{\tau_{\mathcal{K}}} r_s^r ds} 1_{\{\tau_{\mathcal{K}} \le \tau_*\}}\right], \tag{63}$$

where $\tau_{\mathcal{K}} := \inf\{s \ge 0 \,:\, (t+s, r_s, X_s) \notin int(\mathcal{K})\}$.

Remark 7.4. Later on, we obtain also a bound on the derivative v_r of the value function with respect to the interest rate. We present it separately in Equation (78) because, due to the square root appearing in the diffusion coefficient of the CIR dynamics, we need to use local approximations of the stochastic dynamics of $(r_t)_{t\ge0}$. That procedure does not lead to a neat expression as in Equations (62) and (63).

Proof of Proposition 7.3. Fix $(t, r, x) \in (\text{int}(\mathcal{K}) \setminus \partial\mathcal{C})$. Recall that $\mathcal{D} \subset [0, T] \times \mathcal{I} \times [0, K]$. If $(t, r, x) \in \text{int}(\mathcal{D})$ then Equation (62) follows easily from $v(t, r, x) = K - x$ and $v_t(t, r, x) = 0$. Assume $(t, r, x) \in \mathcal{C}$ and notice that $\tau_* = \sigma_{\mathcal{D}}$, $\mathsf{P}_{t,r,x}$-a.s. We split the proof into two parts.

(*Proof of Equation* (62)) For all sufficiently small $\varepsilon > 0$, we have $(t, r, x + \varepsilon) \in \mathcal{C}$. From now on, consider such ε. To simplify notation let $\sigma_{\mathcal{D}} := \sigma_{\mathcal{D}}(t, r, x)$. Using that $\sigma_{\mathcal{D}}$ is admissible and suboptimal for $v(t, r, x + \varepsilon)$ we get

$$\begin{aligned}
&v(t,r,x+\varepsilon) - v(t,r,x)\\
&\ge \mathsf{E}\left[e^{-\int_0^{\sigma_{\mathcal{D}}} r_s^r ds}\left(\left(K-(x+\varepsilon)X^{r,1}_{\sigma_{\mathcal{D}}}\right)^+ - \left(K - xX^{r,1}_{\sigma_{\mathcal{D}}}\right)^+\right)\right]\\
&\ge \mathsf{E}\left[e^{-\int_0^{\sigma_{\mathcal{D}}} r_s^r ds} 1_{\{X^{r,x}_{\sigma_{\mathcal{D}}} \le K\}}\left(xX^{r,1}_{\sigma_{\mathcal{D}}} - (x+\varepsilon)X^{r,1}_{\sigma_{\mathcal{D}}}\right)\right] = -\varepsilon \mathsf{E}\left[1_{\{X^{r,x}_{\sigma_{\mathcal{D}}} \le K\}} e^{\sigma B_{\sigma_{\mathcal{D}}} - \frac{\sigma^2}{2}\sigma_{\mathcal{D}}}\right].
\end{aligned}$$

Dividing the above expression by ε and taking limits as $\varepsilon \to 0$ we get

$$v_x(t,r,x) = \lim_{\varepsilon\to0}\frac{1}{\varepsilon}(v(t,r,x+\varepsilon) - v(t,r,x)) \ge -\mathsf{E}\left[1_{\{X^{r,x}_{\sigma_{\mathcal{D}}} \le K\}} e^{\sigma B_{\sigma_{\mathcal{D}}} - \frac{\sigma^2}{2}\sigma_{\mathcal{D}}}\right]. \tag{64}$$

For the reverse inequality, we use that $\sigma_{\mathcal{D}}$ is admissible and suboptimal for $v(t,r,x-\varepsilon)$:

$$v(t,r,x)-v(t,r,x-\varepsilon) \le \mathsf{E}\left[e^{-\int_0^{\sigma_D} r_s^r ds}\left(\left(K - xX_{\sigma_D}^{r,1}\right)^+ - \left(K-(x-\varepsilon)X_{\sigma_D}^{r,1}\right)^+\right)\right]$$

$$\le -\varepsilon\,\mathsf{E}\left[1_{\{X_{\sigma_D}^{r,x-\varepsilon}\le K\}} e^{\sigma B_{\sigma_D}-\frac{\sigma^2}{2}\sigma_D}\right] \le -\varepsilon\,\mathsf{E}\left[1_{\{X_{\sigma_D}^{r,x}\le K\}} e^{\sigma B_{\sigma_D}-\frac{\sigma^2}{2}\sigma_D}\right],$$

where in the last inequality, we used that $X_s^{r,x-\varepsilon} < X_s^{r,x}$, $s \ge 0$. Divide the above expression by ε and take limits as $\varepsilon \to 0$:

$$v_x(t,r,x) = \lim_{\varepsilon\to 0}\frac{1}{\varepsilon}(v(t,r,x)-v(t,r,x-\varepsilon)) \le -\mathsf{E}\left[1_{\{X_{\sigma_D}^{r,x}\le K\}} e^{\sigma B_{\sigma_D}-\frac{\sigma^2}{2}\sigma_D}\right]. \tag{65}$$

Now Equations (64) and (65) imply Equation (62).

(*Proof of* (63)). The upper bound $v_t(t,r,x) \le 0$ follows from the monotonicity of v in t (Proposition 3.1). For all sufficiently small $\varepsilon > 0$, we have $(t+\varepsilon, r, x) \in \mathcal{K}\cap C$ and $\tau_{\mathcal{K}} := \tau_{\mathcal{K}}(t,r,x) \le T-t-\varepsilon$. From now on, consider such ε. Denote $\sigma_{\mathcal{D}} := \sigma_{\mathcal{D}}(t,r,x)$. Thanks to the choice of ε, the stopping time $\eta := \sigma_{\mathcal{D}} \wedge \tau_{\mathcal{K}}$ is admissible for $v(t+\varepsilon,r,x)$. Using the (super)martingale property of v (see Equations (11)–(12)), we get

$$\begin{aligned}
&v(t+\varepsilon,r,x)-v(t,r,x)\\
&\ge \mathsf{E}\left[e^{-\int_0^{\eta} r_s^r ds}\left(v(t+\varepsilon+\eta, r_\eta^r, X_\eta^{r,x}) - v(t+\eta, r_\eta^r, X_\eta^{r,x})\right)\right] \qquad (66)\\
&= \mathsf{E}\left[e^{-\int_0^{\tau_{\mathcal{K}}} r_s^r ds}\left(v(t+\varepsilon+\tau_{\mathcal{K}}, r_{\tau_{\mathcal{K}}}^r, X_{\tau_{\mathcal{K}}}^{r,x}) - v(t+\tau_{\mathcal{K}}, r_{\tau_{\mathcal{K}}}^r, X_{\tau_{\mathcal{K}}}^{r,x})\right)1_{\{\tau_{\mathcal{K}}<\sigma_{\mathcal{D}}\}}\right],
\end{aligned}$$

where the equality follows from $v(t+\varepsilon+\sigma_{\mathcal{D}}, r_{\sigma_{\mathcal{D}}}^r, X_{\sigma_{\mathcal{D}}}^{r,x}) = v(t+\sigma_{\mathcal{D}}, r_{\sigma_{\mathcal{D}}}^r, X_{\sigma_{\mathcal{D}}}^{r,x}) = K - X_{\sigma_{\mathcal{D}}}^{r,x}$ on $\{\tau_{\mathcal{K}} \ge \sigma_{\mathcal{D}}\}$ since $t \mapsto b(t,r)$ is nondecreasing (Proposition 3.4). Let $\mathcal{K}^\delta = \{(t+s,r,x) : (t,r,x) \in \mathcal{K}$ and $s \in [0,\delta]\}$. Fix a sufficiently small $\delta > 0$ so that this set is contained in $\mathcal{O}$ and set L equal to the Lipschitz constant for v on $\mathcal{K}^\delta$ (c.f. Proposition 5.1). Since $(t+\tau_{\mathcal{K}}, r_{\tau_{\mathcal{K}}}^r, X_{\tau_{\mathcal{K}}}^{r,x}) \in \partial\mathcal{K}$, we have $(t+\varepsilon+\tau_{\mathcal{K}}, r_{\tau_{\mathcal{K}}}^r, X_{\tau_{\mathcal{K}}}^{r,x}) \in \mathcal{K}^\delta$ for any $\varepsilon < \delta$. Using the Lipschitz continuity of v, we bound Equation (66) from below by

$$-\varepsilon L\,\mathsf{E}\left[e^{-\int_0^{\tau_{\mathcal{K}}} r_s^r ds} 1_{\{\tau_{\mathcal{K}}<\sigma_{\mathcal{D}}\}}\right].$$

Dividing by ε and taking the limit $\varepsilon \to 0$ completes the proof of Equation (63). □

We are now ready to prove that the value function is globally continuously differentiable on $\mathcal{O}$.

Proof of Theorem 3.5. It suffices to show that the value function has continuous partial derivatives across the stopping boundary, that is

$$\lim_{n\to\infty} v_t(t_n,r_n,x_n) = \lim_{n\to\infty} v_r(t_n,r_n,x_n) = 0, \tag{67}$$

$$\lim_{n\to\infty} v_x(t_n,r_n,x_n) = -1, \tag{68}$$

for any sequence (t_n,r_n,x_n) in C converging to $(t_0,r_0,x_0) \in \partial C$ as $n \to \infty$. Fix such a sequence and denote $\tau_n = \tau_*(t_n,r_n,x_n)$.

Convergence of v_x. Note that $\mathsf{P}_{t_n,r_n,x_n}(X_{\tau_n} = K,\, \tau_n < T - t_n) = 0$ (Proposition 3.3) and $\mathsf{P}_{t_n,r_n,x_n}(X_{\tau_n} = K,\, \tau_n = T - t_n) \le \mathsf{P}_{t_n,r_n,x_n}(X_{T-t_n} = K) = 0$ (the final equality can be shown by arguments as in (55)). From Proposition 7.3, we therefore have

$$v_x(t_n, r_n, x_n) = -\mathsf{E}\left[1_{\{X_{\tau_n}^{r_n,x_n} < K\}} e^{\sigma B_{\tau_n} - \frac{\sigma^2}{2}\tau_n}\right].$$

From Lemma 7.2, we obtain $\lim_{n\to\infty} \tau_n = 0$ P-a.s. We know from $(t_0, r_0, x_0) \in \partial\mathcal{C}$ that $x_0 < K$. Lemma 7.1 and the continuity of trajectories of (r, X) imply the convergence $1_{\{X_{\tau_n}^{r_n,x_n} < K\}} \to 1_{\{x_0 < K\}} = 1$ as $n \to \infty$. An application of the dominated convergence theorem completes the proof of Equation (68).

Convergence of v_t. Let $\mathcal{K}$ be a closed ball centered on (t_0, r_0, x_0) and contained in $\mathcal{O}$. With no loss of generality (by discarding a finite number of initial elements of the sequence) we assume that $(t_n, r_n, x_n) \in int(\mathcal{K})$ for all $n \ge 1$. Let

$$\tau_{\mathcal{K}}^n := \inf\{s \ge 0 : (t_n + s, r_s^{r_n}, X_s^{r_n,x_n}) \notin \mathcal{K}\}, \quad n \ge 0$$

and notice, in particular, that $\mathsf{P}(\tau_{\mathcal{K}}^0 > 0) = 1$. The boundary $\partial\mathcal{K}$ is regular for $\mathcal{O} \setminus \mathcal{K}$ and (t, r, X) by the same reasoning as in the proof of Proposition 6.1. Repeating arguments from the proof of Lemma 7.2 shows that $\tau_{\mathcal{K}}^n \to \tau_{\mathcal{K}}^0$, P-a.s. Fix $\varepsilon \in (0,1)$. Since $\mathsf{P}(\tau_{\mathcal{K}}^0 > 0) = 1$, there exists $\delta > 0$ such that $\mathsf{P}(\tau_{\mathcal{K}}^0 > \delta) \ge 1 - \varepsilon$. From inequality (63), we get

$$\begin{aligned}
0 \ge v_t(t_n, r_n, x_n) &\ge -L\,\mathsf{E}\left[e^{-\int_0^{\tau_{\mathcal{K}}^n} r_s^{r_n} ds} 1_{\{\tau_{\mathcal{K}}^n \le \tau_n\}}\right] \\
&= -L\,\mathsf{E}\left[e^{-\int_0^{\tau_{\mathcal{K}}^n} r_s^{r_n} ds}\left(1_{\{\tau_{\mathcal{K}}^n \le \tau_n\}\cap\{\tau_{\mathcal{K}}^n \ge \delta\}} + 1_{\{\tau_{\mathcal{K}}^n \le \tau_n\}\cap\{\tau_{\mathcal{K}}^n < \delta\}}\right)\right] \\
&\ge -L\,\mathsf{E}\left[e^{-\int_0^{\tau_{\mathcal{K}}^n} r_s^{r_n} ds}\left(1_{\{\tau_n \ge \delta\}} + 1_{\{\tau_{\mathcal{K}}^n < \delta\}}\right)\right].
\end{aligned} \tag{69}$$

Using that $|r_{t\wedge\tau_{\mathcal{K}}^n}|$ is bounded by some constant $r_{\mathcal{K}}$ for every n, we have

$$0 \ge v_t(t_n, r_n, x_n) \ge -Le^{r_{\mathcal{K}}T}\left(\mathsf{P}(\tau_n \ge \delta) + \mathsf{P}\left(\tau_{\mathcal{K}}^n < \delta\right)\right). \tag{70}$$

Lemma 7.2 guarantees that $\tau_n \to 0$ P-a.s., so the first term converges to 0 as $n \to \infty$ by the dominated convergence theorem. Fatou's lemma gives a bound for the second term:

$$\limsup_{n\to\infty} \mathsf{P}\left(\tau_{\mathcal{K}}^n < \delta\right) \le \mathsf{E}\left[\limsup_{n\to\infty} 1_{\{\tau_{\mathcal{K}}^n < \delta\}}\right] \le \mathsf{E}\left[1_{\{\tau_{\mathcal{K}}^0 \le \delta\}}\right] \le \varepsilon,$$

where we used that $\limsup_n 1_{A_n} = 1_{\limsup_n A_n}$ and the convergence of the stopping times. We obtain the convergence of v_t in Equation (67) by sending $\varepsilon \to 0$.

Convergence of v_r. Consider a sequence $(t_n, r_n, x_n) \in \mathcal{C}$ converging to $(t_0, r_0, x_0) \in \partial\mathcal{C}$. Since $\partial\mathcal{C}$ is the boundary of $\mathcal{C}$ in $\mathcal{O}$, without loss of generality, we can assume that

$$\{(r_n, x_n)\} \subset \mathrm{int}(\mathcal{K}) \quad \text{with} \quad \mathcal{K} := [r_a, r_b] \times [x_a, x_b] \subset (\underline{r}, \overline{r}) \times \mathbb{R}_+.$$

Denote $\mathcal{K}^T := [t_a, t_b] \times \mathcal{K}_0$, where $t_a = \inf_n t_n \ge 0$ and $t_b = \sup_n t_n < T$.

We know that $v_r \le 0$ on $\mathcal{C}$ (Proposition 3.1). We will now develop a lower bound for v_r on $\mathcal{C} \cap \mathcal{K}^T$, which will allow us to show that $v_r(t_n, r_n, x_n) \to 0$ as $n \to \infty$. Let $\tilde{\mathcal{K}} \subset (\underline{r}, \overline{r}) \times \mathbb{R}_+$ be compact and such that $\mathcal{K} \subset \mathrm{int}(\tilde{\mathcal{K}})$. Denote $\tilde{\mathcal{K}}^T = [t_a, \tilde{t}_b] \times \tilde{\mathcal{K}}$ for some $\tilde{t}_b \in (t_b, T)$. For $(t, r, x) \in \mathcal{C} \cap \mathcal{K}^T$, we define

$$\tau_{\tilde{\mathcal{K}}}(t, r, x) := \inf\{s \ge 0 : (r_s^r, X_s^{r,x}) \notin \tilde{\mathcal{K}}\} \wedge (T - t).$$

By the monotonicity of $r \mapsto r_s^r$ and the explicit expression (5) for $X^{r,x}$, we have, for all $(r, x) \in \mathcal{K}$,

$$r_s^{r_a} \le r_s^r \le r_s^{r_b}, \quad \text{and} \quad X_s^{r_a,x} \le X_s^{r,x} \le X_s^{r_b,x}, \quad \text{P-a.s.}$$

from which it is not hard to verify that $\tau_{\tilde{\mathcal{K}}}(t, r, x) \ge \hat{\tau}_{\mathcal{K}} > 0$, P-a.s., for all $(t, r, x) \in \mathcal{C} \cap \mathcal{K}^T$, where

$$\hat{\tau}_{\mathcal{K}} := \tau_{\tilde{\mathcal{K}}}(t_b, r_a, x_a) \wedge \tau_{\tilde{\mathcal{K}}}(t_b, r_a, x_b) \wedge \tau_{\tilde{\mathcal{K}}}(t_b, r_b, x_a) \wedge \tau_{\tilde{\mathcal{K}}}(t_b, r_b, x_b).$$

Take $(t, r, x) \in \mathcal{C} \cap \mathrm{int}(\mathcal{K}^T)$. There is $\bar{\varepsilon} > 0$ such that $(t, r + \varepsilon, x) \in \mathcal{C} \cap \mathcal{K}^T$ for all $\varepsilon \in (0, \bar{\varepsilon}]$. Denote by τ_* the optimal stopping time for (t, r, x). For any $\varepsilon \in (0, \bar{\varepsilon}]$, we apply the (super)martingale properties of the value function (11)–(12) with the stopping time $\tau_* \wedge \hat{\tau}_{\mathcal{K}}$:

$$\begin{aligned}
0 &\ge v(t, r + \varepsilon, x) - v(t, r, x) \\
&\ge \mathsf{E}\Big[e^{-\int_0^{\tau_* \wedge \hat{\tau}_{\mathcal{K}}} r_s^{r+\varepsilon} ds} v\Big(t + (\tau_* \wedge \hat{\tau}_{\mathcal{K}}), r^{r+\varepsilon}_{\tau_* \wedge \hat{\tau}_{\mathcal{K}}}, X^{r+\varepsilon,x}_{\tau_* \wedge \hat{\tau}_{\mathcal{K}}}\Big) \\
&\qquad - e^{-\int_0^{\tau_* \wedge \hat{\tau}_{\mathcal{K}}} r_s^{r} ds} v\Big(t + (\tau_* \wedge \hat{\tau}_{\mathcal{K}}), r^{r}_{\tau_* \wedge \hat{\tau}_{\mathcal{K}}}, X^{r,x}_{\tau_* \wedge \hat{\tau}_{\mathcal{K}}}\Big)\Big] \\
&\ge \mathsf{E}\Big[1_{\{\hat{\tau}_{\mathcal{K}} \le \tau_*\}}\Big(e^{-\int_0^{\hat{\tau}_{\mathcal{K}}} r_s^{r+\varepsilon} ds} v(t + \hat{\tau}_{\mathcal{K}}, r^{r+\varepsilon}_{\hat{\tau}_{\mathcal{K}}}, X^{r+\varepsilon,x}_{\hat{\tau}_{\mathcal{K}}}) - e^{-\int_0^{\hat{\tau}_{\mathcal{K}}} r_s^{r} ds} v(t + \hat{\tau}_{\mathcal{K}}, r^{r}_{\hat{\tau}_{\mathcal{K}}}, X^{r,x}_{\hat{\tau}_{\mathcal{K}}})\Big)\Big] \\
&\quad + \mathsf{E}\Big[1_{\{\hat{\tau}_{\mathcal{K}} > \tau_*\}}\Big(e^{-\int_0^{\tau_*} r_s^{r+\varepsilon} ds}(K - X^{r+\varepsilon,x}_{\tau_*})^+ - e^{-\int_0^{\tau_*} r_s^{r} ds}(K - X^{r,x}_{\tau_*})^+\Big)\Big] \\
&=: E_1 + E_2,
\end{aligned} \tag{71}$$

where for the final inequality, we used that $v(t + \tau_*, r^r_{\tau_*}, X^{r,x}_{\tau_*}) = (K - X^{r,x}_{\tau_*})^+$, P-a.s. Recalling that $r_s^{r+\varepsilon} \ge r_s^r$ and v is non-negative, we have

$$\begin{aligned}
E_1 &= \mathsf{E}\Big[1_{\{\hat{\tau}_{\mathcal{K}} \le \tau_*\}} e^{-\int_0^{\hat{\tau}_{\mathcal{K}}} r_s^{r+\varepsilon} ds}\Big(v(t + \hat{\tau}_{\mathcal{K}}, r^{r+\varepsilon}_{\hat{\tau}_{\mathcal{K}}}, X^{r+\varepsilon,x}_{\hat{\tau}_{\mathcal{K}}}) - v(t + \hat{\tau}_{\mathcal{K}}, r^{r}_{\hat{\tau}_{\mathcal{K}}}, X^{r,x}_{\hat{\tau}_{\mathcal{K}}})\Big)\Big] \\
&\quad - \mathsf{E}\Big[1_{\{\hat{\tau}_{\mathcal{K}} \le \tau_*\}}\Big(e^{-\int_0^{\hat{\tau}_{\mathcal{K}}} r_s^{r} ds} - e^{-\int_0^{\hat{\tau}_{\mathcal{K}}} r_s^{r+\varepsilon} ds}\Big) v(t + \hat{\tau}_{\mathcal{K}}, r^{r}_{\hat{\tau}_{\mathcal{K}}}, X^{r,x}_{\hat{\tau}_{\mathcal{K}}})\Big] \\
&\ge -L\,\mathsf{E}\Big[1_{\{\hat{\tau}_{\mathcal{K}} \le \tau_*\}} e^{-\int_0^{\hat{\tau}_{\mathcal{K}}} r_s^{r+\varepsilon} ds}\Big(|r^{r+\varepsilon}_{\hat{\tau}_{\mathcal{K}}} - r^{r}_{\hat{\tau}_{\mathcal{K}}}| + |X^{r+\varepsilon,x}_{\hat{\tau}_{\mathcal{K}}} - X^{r,x}_{\hat{\tau}_{\mathcal{K}}}|\Big)\Big] \\
&\quad - K\mathsf{E}\Big[1_{\{\hat{\tau}_{\mathcal{K}} \le \tau_*\}}\Big(e^{-\int_0^{\hat{\tau}_{\mathcal{K}}} r_s^{r} ds} - e^{-\int_0^{\hat{\tau}_{\mathcal{K}}} r_s^{r+\varepsilon} ds}\Big)\Big],
\end{aligned} \tag{72}$$

where the second inequality comes from the local Lipschitz property of the value function ($L > 0$ is the constant from Proposition 5.1), and the function v is bounded by the strike price K from above.

We shall now use the differentiability of the diffusion flow (r_s^r) with respect to the parameter r in the sense of (Krylov, 1980, Ch. 2, Sec. 8, Thm. 6). Apart from other assumptions, this requires that the coefficients be globally Lipschitz. As we only consider (r, X) in a compact set $\mathcal{K}$, we construct a two-dimensional diffusion $(\widetilde{r}, \widetilde{X})$, whose coefficients coincide with the coefficients of (r, X) on $\mathcal{K}$, are globally Lipschitz, continuously differentiable and with a polynomial growth. The process $(\widetilde{r}_s, \widetilde{X}_s)$ is indistinguishable from (r_s, X_s) on $\{s \le \widehat{\tau}_{\mathcal{K}}\}$, that is, on the set where it is of interest for the estimation of E_1 and E_2, so for the sake of readability, we will write (r, X) in the estimates below (we use an analogous construction in Appendix B, where full details are available).

By (Krylov, 1980, Ch. 2, Sec. 8, Thm. 6), there is a measurable in (s, ω) process $(y_s^r(\omega))_{s\ge 0}$, depending on r, such that for any $q \ge 1$

$$\lim_{\varepsilon\downarrow 0}\left\| \sup_{s\in[0,T]}\left|\frac{r_s^{r+\varepsilon}-r_s^r}{\varepsilon}-y_s^r\right|\right\|_q = 0 \quad\text{and}\quad \lim_{\varepsilon\downarrow 0}\left\|\frac{r_\cdot^{r+\varepsilon}-r_\cdot^r}{\varepsilon}-y_\cdot^r\right\|_q^* = 0, \tag{73}$$

where $\|Z\|_q = (\mathsf{E}[|Z|^q])^{1/q}$ and $\|Y_\cdot\|_q^* = (\mathsf{E}[\int_0^T |Y_s|^q ds])^{1/q}$.

Fix $\frac{1}{p}+\frac{1}{q}+\frac{1}{w}=1$ for some $p \in (1,2]$. Recalling that $r_s^{r+\varepsilon} \ge r_s^r$ and using Hölder inequality yields

$$\begin{aligned}
&\frac{1}{\varepsilon}\mathsf{E}\left[1_{\{\widehat{\tau}_{\mathcal{K}}\le\tau_*\}}e^{-\int_0^{\widehat{\tau}_{\mathcal{K}}} r_s^{r+\varepsilon}ds}|r_{\widehat{\tau}_{\mathcal{K}}}^{r+\varepsilon}-r_{\widehat{\tau}_{\mathcal{K}}}^r|\right]\\
&\le \mathsf{E}\left[1_{\{\widehat{\tau}_{\mathcal{K}}\le\tau_*\}}e^{-\int_0^{\widehat{\tau}_{\mathcal{K}}} r_s^{r+\varepsilon}ds}\left(\left|\frac{1}{\varepsilon}\left(r_{\widehat{\tau}_{\mathcal{K}}}^{r+\varepsilon}-r_{\widehat{\tau}_{\mathcal{K}}}^r\right)-y_{\widehat{\tau}_{\mathcal{K}}}^r\right|+|y_{\widehat{\tau}_{\mathcal{K}}}^r|\right)\right]\\
&\le C_1^{1/p}\mathsf{P}(\widehat{\tau}_{\mathcal{K}}\le\tau_*)^{\frac{1}{w}}\left(\left\|\frac{1}{\varepsilon}\left(r_{\widehat{\tau}_{\mathcal{K}}}^{r+\varepsilon}-r_{\widehat{\tau}_{\mathcal{K}}}^r\right)-y_{\widehat{\tau}_{\mathcal{K}}}^r\right\|_q+\|y_{\widehat{\tau}_{\mathcal{K}}}^r\|_q\right)\xrightarrow[\varepsilon\downarrow 0]{} C_1^{1/p}\mathsf{P}(\widehat{\tau}_{\mathcal{K}}\le\tau_*)^{\frac{1}{w}}\|y_{\widehat{\tau}_{\mathcal{K}}}^r\|_q,
\end{aligned}\tag{74}$$

where we used the estimate (3) in the last inequality and Equation (73) to obtain the convergence.

To bound the last term on the right-hand side of Equation (72), we observe that

$$\mathsf{E}\left[1_{\{\widehat{\tau}_{\mathcal{K}}\le\tau_*\}}\left(e^{-\int_0^{\widehat{\tau}_{\mathcal{K}}} r_s^r ds}-e^{-\int_0^{\widehat{\tau}_{\mathcal{K}}} r_s^{r+\varepsilon}ds}\right)\right]\le \mathsf{E}\left[1_{\{\widehat{\tau}_{\mathcal{K}}\le\tau_*\}}e^{-\int_0^{\widehat{\tau}_{\mathcal{K}}} r_s^r ds}\int_0^{\widehat{\tau}_{\mathcal{K}}}(r_s^{r+\varepsilon}-r_s^r)ds\right].$$

We then apply Hölder inequality and the second limit in Equation (73):

$$\begin{aligned}
&\frac{1}{\varepsilon}\mathsf{E}\left[1_{\{\widehat{\tau}_{\mathcal{K}}\le\tau_*\}}e^{-\int_0^{\widehat{\tau}_{\mathcal{K}}} r_s^r ds}\int_0^{\widehat{\tau}_{\mathcal{K}}}(r_s^{r+\varepsilon}-r_s^r)ds\right]\\
&\le \mathsf{E}\left[1_{\{\widehat{\tau}_{\mathcal{K}}\le\tau_*\}}e^{-\int_0^{\widehat{\tau}_{\mathcal{K}}} r_s^r ds}\int_0^{\widehat{\tau}_{\mathcal{K}}}\left|\frac{1}{\varepsilon}(r_s^{r+\varepsilon}-r_s^r)-y_s^r\right|+|y_s^r|ds\right]\\
&\le C_1^{1/p}\,\mathsf{P}(\widehat{\tau}_{\mathcal{K}}\le\tau_*)^{\frac{1}{w}}\left(\left\|\frac{1}{\varepsilon}(r_\cdot^{r+\varepsilon}-r_\cdot^r)-y_\cdot^r\right\|_q^*+\|y_\cdot^r\|_q^*\right)\xrightarrow[\varepsilon\downarrow 0]{} C_1^{1/p}\,\mathsf{P}(\widehat{\tau}_{\mathcal{K}}\le\tau_*)^{\frac{1}{w}}\,\|y_\cdot^r\|_q^*.
\end{aligned}\tag{75}$$

By the explicit formula (5), we have $X_t^{r,x} = e^{\int_0^t r_s^r ds}\hat{X}_t^x$, where $\hat{X}_t^x := xe^{\sigma B_t - \frac{1}{2}\sigma^2 t}$, and

$$0 \le X_t^{r+\varepsilon,x} - X_t^{r,x} \le e^{\int_0^t r_s^{r+\varepsilon}ds}\hat{X}_t^x\int_0^t (r_s^{r+\varepsilon}-r_s^r)ds.$$

We proceed similarly as in Equation (75) to obtain

$$
\begin{aligned}
&\frac{1}{\varepsilon}\mathsf{E}\left[1_{\{\widehat{\tau}_{\mathcal{K}}\le\tau_*\}}e^{-\int_0^{\widehat{\tau}_{\mathcal{K}}} r_s^{r+\varepsilon}ds}|X^{r+\varepsilon,x}_{\widehat{\tau}_{\mathcal{K}}}-X^{r,x}_{\widehat{\tau}_{\mathcal{K}}}|\right]\\
&\le \|\hat{X}^x_{\widehat{\tau}_{\mathcal{K}}}\|_p\ \mathsf{P}(\widehat{\tau}_{\mathcal{K}}\le\tau_*)^{\frac{1}{w}}\left(\left\|\frac{1}{\varepsilon}(r^{r+\varepsilon}_\cdot-r^r_\cdot)-y^r_\cdot\right\|_q^*+\|y^r_\cdot\|_q^*\right) \qquad (76)\\
&\underset{\varepsilon\downarrow 0}{\longrightarrow} \|\hat{X}^x_{\widehat{\tau}_{\mathcal{K}}}\|_p\ \mathsf{P}(\widehat{\tau}_{\mathcal{K}}\le\tau_*)^{\frac{1}{w}}\ \|y^r_\cdot\|_q^*.
\end{aligned}
$$

Similar arguments as above enable us to derive a lower bound for E_2:

$$
\begin{aligned}
\frac{1}{\varepsilon}E_2 &= \frac{1}{\varepsilon}\mathsf{E}\left[1_{\{\widehat{\tau}_{\mathcal{K}}>\tau_*\}}\left(\left(Ke^{-\int_0^{\tau_*} r_s^{r+\varepsilon}ds}-\hat{X}^x_{\tau_*}\right)^+-\left(Ke^{-\int_0^{\tau_*} r_s^{r}ds}-\hat{X}^x_{\tau_*}\right)^+\right)\right]\\
&\ge -\frac{1}{\varepsilon}K\,\mathsf{E}\left[1_{\{\widehat{\tau}_{\mathcal{K}}>\tau_*\}}\left(e^{-\int_0^{\tau_*} r_s^{r}ds}-e^{-\int_0^{\tau_*} r_s^{r+\varepsilon}ds}\right)\right]\\
&\ge -\frac{1}{\varepsilon}K\,\mathsf{E}\left[1_{\{\widehat{\tau}_{\mathcal{K}}>\tau_*\}}e^{-\int_0^{\tau_*} r_s^{r}ds}\int_0^{\tau_*}(r_s^{r+\varepsilon}-r_s^r)ds\right] \qquad (77)\\
&\ge -KC_1^{1/p}\ \mathsf{P}(\widehat{\tau}_{\mathcal{K}}>\tau_*)^{\frac{1}{w}}\left(\left\|\frac{1}{\varepsilon}(r^{r+\varepsilon}_\cdot-r^r_\cdot)-y^r_\cdot\right\|_q^*+\left(\mathsf{E}\left[\int_0^{\tau_*}|y_s^r|^q ds\right]\right)^{1/q}\right)\\
&\underset{\varepsilon\downarrow 0}{\longrightarrow} -KC_1^{1/p}\mathsf{P}(\widehat{\tau}_{\mathcal{K}}>\tau_*)^{\frac{1}{w}}\ \left(\mathsf{E}\left[\int_0^{\tau_*}|y_s^r|^q ds\right]\right)^{1/q},
\end{aligned}
$$

where in the first inequality, we used the Lipschitz property of $z\mapsto(z-\hat{X}^x_t(\omega))^+$ for any $\omega\in\Omega$.

Combining Equations (74)–(77) gives a lower bound for v_r on $C\cap\mathcal{K}^T$:

$$
\begin{aligned}
0 &\ge v_r(t,r,x)\\
&\ge -L\,\mathsf{P}(\widehat{\tau}_{\mathcal{K}}\le\tau_*)^{\frac{1}{w}}\ \left(C_1^{1/p}\|y^r_{\widehat{\tau}_{\mathcal{K}}}\|_q+\|\hat{X}^x_{\widehat{\tau}_{\mathcal{K}}}\|_p\ \|y^r_\cdot\|_q^*\right)-K\,\mathsf{P}(\widehat{\tau}_{\mathcal{K}}\le\tau_*)^{\frac{1}{w}}\ C_1^{1/p}\|y^r_\cdot\|_q^* \qquad (78)\\
&\quad -KC_1^{1/p}\mathsf{P}(\widehat{\tau}_{\mathcal{K}}>\tau_*)^{\frac{1}{w}}\ \left(\mathsf{E}\left[\int_0^{\tau_*}|y_s^r|^q ds\right]\right)^{1/q}.
\end{aligned}
$$

By (Krylov, 1980, Ch. 2, Sec. 8, Thm. 8) and standard diffusion estimates (Krylov, 1980, Ch. 2, Sec. 5, Cor. 10), the norms of y^r and $\hat{X}^x$ above are bounded uniformly for $(t,r,x)\in\mathcal{K}^T\cap C$ (recall that $\tau_*=\tau_*(t,r,x)$). Now take $(t,r,x)=(t_n,r_n,x_n)$ in Equation (78). Since $\hat{\tau}_{\mathcal{K}}>0$ P-a.s. and $\lim_{n\to\infty}\tau_*(t_n,r_n,x_n)=0$ P-a.s. by Lemma 7.2, the dominated convergence theorem gives that the first two terms of Equation (78) tend to zero as $n\to\infty$ due to $\mathsf{P}(\widehat{\tau}_{\mathcal{K}}\le\tau_*(t_n,r_n,x_n))\to 0$ and the last term converges to zero because

$$
\lim_{n\to 0}\mathsf{E}\left[\int_0^{\tau_*(t_n,r_n,x_n)}|y_s^{r_n}|^q ds\right]=0,
$$

and the mapping $r\mapsto y^r$ is continuous in the norm $\|\cdot\|_q^*$, see (Krylov 1980, Ch. 2, Sec. 8, Thm. 6). This concludes the proof. □

8 | CONTINUITY OF THE STOPPING BOUNDARY AND THE INTEGRAL EQUATION

Proof of Proposition 3.7. Since $c(t,x) = \bar{r}$ on $[0,T)\times[K,\infty)$, it remains to prove the continuity at $(t_0,x_0)\in(0,T)\times(0,K]$. It is known from Proposition 3.3 that $t\mapsto c(t,x_0)$ is nonincreasing and right-continuous at t_0, and $x\mapsto c(t_0,x)$ is nondecreasing and left-continuous at x_0.

We first show that $x\mapsto c(t_0,x)$ is right continuous at x_0. It is obvious for $x_0=K$ since $c(t_0,x)=\bar{r}$ for $x\geq K$. We proceed with an argument for $x_0<K$. Assume, by contradiction, that $c(t_0,x_0+) > c(t_0,x_0)$, so there exist r_1,r_2 such that $c(t_0,x_0+)>r_2>r_1>c(t_0,x_0)$. Let $R := (r_1,r_2)\times(x_0,x_1)$ for some $x_1\in(x_0,K)$ and $R_0 := (r_1,r_2)\times\{x_0\}$. From the monotonicity of $c(t,x)$, we have $\{t_0\}\times R\subset \mathcal{C}$ and $\{t_0\}\times R_0\subset\mathcal{D}$. Let u be a function defined on $\bar{R}$ and satisfying

$$\begin{aligned}(\mathcal{L}-r)u(r,x) &= -v_t(t_0,r,x), \qquad (r,x)\in R,\\ u(r,x) &= v(t_0,r,x), \qquad (r,x)\in\partial R.\end{aligned} \tag{79}$$

Thanks to (Friedman, 1964, Thm. 10, p. 72) we know that $(r,x)\mapsto v_t(t_0,r,x)$ is C^1 on R with Hölder continuous derivatives. Since the coefficients of Equation (14) have Hölder continuous first derivatives, there is a unique classical solution $u(r,x)$ of the above PDE (which is of elliptic type) and $u\in C^3(R)\cap C(\bar{R})$ (Friedman, 1964, Thm. 19 and 20, p. 87). From Equation (17), the function $(r,x)\mapsto v(t_0,r,x)$ satisfies Equation (79), so, by uniqueness, $u=v$ on $\bar{R}$ and $u\in C^1(\bar{R})$ by Theorem 3.5.

We differentiate the PDE in Equation (79) with respect to r and obtain

$$\frac{1}{2}\sigma^2x^2u_{rxx}(r,x) = -\mathcal{L}_1u_r(r,x) - \mathcal{L}_2u_x(r,x) - xu_x(r,x) - v_{tr}(t_0,r,x) + u(r,x), \quad (r,x)\in R, \tag{80}$$

where

$$\begin{aligned}\mathcal{L}_1 f &:= \frac{1}{2}\beta^2(r)f_{rr} + \big(\beta(r)\beta'(r)+\alpha(r)\big)f_r + \big(\alpha'(r)-r\big)f\\ \mathcal{L}_2 f &:= \rho\sigma\beta(r)xf_{rr} + \big(\rho\sigma\beta'(r)+rx\big)f_r.\end{aligned}$$

Let ϕ be a C^∞ function with compact support on (r_1,r_2) such that $\int_{r_1}^{r_2}\phi(r)dr = 1$ and for $x\in(x_0,x_1)$ define

$$F_\phi(x) = -\int_{r_1}^{r_2} u_{xx}(r,x)\phi'(r)dr.$$

Multiply Equation (80) by $\frac{2}{\sigma^2x^2}\phi(r)$ and integrate over (r_1,r_2):

$$\begin{aligned}\int_{r_1}^{r_2} u_{rxx}(x,r)\phi(r)dr = &-\int_{r_1}^{r_2}\frac{2}{\sigma^2x^2}\phi(r)[\mathcal{L}_1u_r(r,x)+\mathcal{L}_2u_x(r,x)]dr - \int_{r_1}^{r_2}\frac{2}{\sigma^2x}\phi(r)u_x(r,x)dr\\ &-\int_{r_1}^{r_2}\frac{2}{\sigma^2x^2}\phi(r)v_{tr}(t_0,r,x)dr + \int_{r_1}^{r_2}\frac{2}{\sigma^2x^2}\phi(r)u(r,x)dr.\end{aligned}$$

Intergration by parts gives

$$\begin{aligned} F_\phi(x) = &- \int_{r_1}^{r_2} \frac{2}{\sigma^2 x^2}\left[u_r(r,x)\mathcal{L}_1^*\phi(r)dr + u_x(r,x)\mathcal{L}_2^*\phi(r)\right]dr - \int_{r_1}^{r_2} \frac{2}{\sigma^2 x}\phi(r)u_x(r,x)dr \\ &+ \int_{r_1}^{r_2} \frac{2}{\sigma^2 x^2}\phi'(r)v_t(t_0,r,x)dr + \int_{r_1}^{r_2} \frac{2}{\sigma^2 x^2}\phi(r)u(r,x)dr, \end{aligned} \tag{81}$$

where $\mathcal{L}_1^*$ and $\mathcal{L}_2^*$ are adjoint operators to $\mathcal{L}_1$ and $\mathcal{L}_2$, respectively. The expression above involves only u and its first derivatives, which are continuous by Theorem 3.5. We take the limit $x \to x_0$ in Equation (81) and notice that $u_r(r,x_0) = v_r(t_0,r,x_0) = v_t(t_0,r,x_0) = 0$, $u_x(r,x_0) = v_x(t_0,r,x_0) = -1$ and $u(r,x_0) = K - x_0$. Thus,

$$\lim_{x\downarrow x_0} F_\phi(x) = \int_{r_1}^{r_2} \frac{2}{\sigma^2 x_0}\phi(r)dr + \int_{r_1}^{r_2} \frac{2}{\sigma^2 x_0^2}\phi(r)(K-x_0)dr = \frac{2K}{\sigma^2 x_0^2} > 0,$$

where we also use that $\int_{r_1}^{r_2} \mathcal{L}_2^*\phi(r)dr = 0$. Since $x \mapsto F_\phi(x)$ is continuous on (x_0, x_1) and $\lim_{x\downarrow x_0} F_\phi(x) > 0$, we have $F_\phi(x) > 0$ on $(x_0, x_0+\varepsilon)$ for any sufficiently small $\varepsilon > 0$. Using additionally that u is $C^1(\overline{R})$, we perform the following integration

$$\begin{aligned} 0 < \int_{x_0}^{x_0+\varepsilon}\int_{x_0}^{y} F_\phi(x)dxdy &= -\int_{r_1}^{r_2}\int_{x_0}^{x_0+\varepsilon}\int_{x_0}^{y} u_{xx}(r,x)dxdy\,\phi'(r)dr \\ &= -\int_{r_1}^{r_2}\int_{x_0}^{x_0+\varepsilon} (u_x(r,y)+1)dy\,\phi'(r)dr \\ &= -\int_{r_1}^{r_2} (u(r,x_0+\varepsilon) - (K-x_0) + \varepsilon)\phi'(r)dr \\ &= \int_{r_1}^{r_2} u_r(r,x_0+\varepsilon)\phi(r)dr, \end{aligned}$$

where we have used Fubini's theorem in the first equality, $u_x(r,x_0) = -1$ in the second equality, $u(r,x_0) = K - x_0$ in the third equality, and the integration by parts in the last equality. As the above inequality holds for an arbitrary smooth function ϕ with a compact support in (r_1,r_2), we must have $u_r(r,x_0+\varepsilon) = v_r(t_0,r,x_0+\varepsilon) > 0$ almost everywhere on (r_1,r_2). This contradicts that $r \mapsto v(t_0,r,x_0+\varepsilon)$ is a nonincreasing function (see Proposition 3.1), hence $x \mapsto c(t,x)$ is continuous.

We turn our attention to the left-continuity of $t \mapsto c(t,x_0)$ at t_0 (the right-continuity has already been established in Proposition 3.3). Assume, by contradiction, that the left-continuity fails at t_0. Since $t \mapsto c(t,x_0)$ is nonincreasing, there exist r_1, r_2 such that $c(t_0-,x_0) > r_2 > r_1 > c(t_0,x_0)$. By the continuity of $x \mapsto c(t_0,x)$ at x_0 and the monotonicity of $c(t,x)$, there is $x_1 \in (x_0,K)$ such that $r_1 > c(t_0,x_1) \geq c(t_0,x_0)$. Hence, for any sequence $t_n \uparrow t_0$, we have

$$c(t_n,x_1) \geq c(t_n,x_0) \geq c(t_0-,x_0) > r_2 > r_1 > c(t_0,x_1) \geq c(t_0,x_0),$$

so that

$$R := (t_1, t_0) \times (r_1, r_2) \times (x_0, x_1) \subset \mathcal{C},$$
$$R_{t_0} := \{t_0\} \times (r_1, r_2) \times (x_0, x_1) \subset \mathcal{D}.$$

Consider a PDE

$$\begin{aligned} w_t(t, r, x) + (\mathcal{L} - r)w(t, r, x) = 0, \quad & (t, r, x) \in R, \\ w(t, r, x) = v(t, r, x), \quad & (t, r, x) \in \partial_p R, \end{aligned} \tag{82}$$

where $\partial_p R$ denotes the parabolic boundary of R. By (Friedman 1964, Thm. 6, p. 65), Equation (82) admits a unique classical solution w, which coincides with v on $\overline{R}$. This also implies that $w \in C^1(\overline{R})$ by Theorem 3.5.

Let ϕ_1 be a C^∞ function with compact support in (x_0, x_1) and ϕ_2 be a C^∞ function with compact support in (r_1, r_2) such that $\int_{x_0}^{x_1} \phi_1(x)dx = \int_{r_1}^{r_2} \phi_2(r)dr = 1$. Fixing $t = t_n \in (t_1, t_0)$ from the sequence $t_n \uparrow t_0$, we multiply Equation (82) by $\phi_1(x)\phi_2(r)$ and integrate over $(r_1, r_2) \times (x_0, x_1)$:

$$\int_{r_1}^{r_2} \int_{x_0}^{x_1} \phi_1(x)\phi_2(r)\{w_t(t_n, r, x) + (\mathcal{L} - r)w(t_n, r, x)\}dxdr = 0.$$

Integration by parts gives

$$\int_{r_1}^{r_2} \int_{x_0}^{x_1} \phi_1(x)\phi_2(r)w_t(t_n, r, x)dxdr + \int_{r_1}^{r_2} \int_{x_0}^{x_1} w(t_n, r, x)(\mathcal{L}^* - r)\phi_1(x)\phi_2(r)dxdr = 0, \tag{83}$$

where $\mathcal{L}^*$ is the adjoint operator for $\mathcal{L}$. When $n \to \infty$, the first integral vanishes since $w \in C^1(\overline{R_t})$ and $w_t = v_t = 0$ on R_{t_0}. By the dominated convergence theorem, Equation (83) reads

$$\begin{aligned} 0 &= \int_{r_1}^{r_2} \int_{x_0}^{x_1} w(t_0, r, x)(\mathcal{L}^* - r)\phi_1(x)\phi_2(r)dxdr = \int_{r_1}^{r_2} \int_{x_0}^{x_1} \phi_1(x)\phi_2(r)(\mathcal{L} - r)(K - x)dxdr \\ &= \int_{r_1}^{r_2} \int_{x_0}^{x_1} \phi_1(x)\phi_2(r)(-rK)dxdr = \int_{r_1}^{r_2} \phi_2(r)(-rK)dr \end{aligned}$$

where we integrate by parts and use that $v(t, r, x) = (K - x)$ on R_{t_0} for the second equality. We obtain a contradiction because the last integral is strictly negative.

Having established the continuity in t and x separately, the monotonicity of c allows us to conclude the continuity of $(t, x) \mapsto c(t, x)$ at (t_0, x_0) (see, e.g., Kruse and Deely (1969)). ☐

Proof of Proposition 3.9. Let $\mathcal{K}_n$ be an increasing sequence of compact subsets of $\mathcal{O}$ such that $\cup_{n \in \mathbb{N}} \mathcal{K}_n = \mathcal{O}$ and define $\tau_n = \inf\{t \in [0, T - t] : (t + s, r_s, X_s) \notin \mathcal{K}_n\} \wedge (T - t - \frac{1}{n})$ for n large enough so that $\frac{1}{n} \leq T - t$. We apply a version of Itô formula from (Cai and De Angelis, 2021, Thm. 2.1), which we state in Appendix C for the reader's convenience. We delay the verification of the assumptions required until the end of the proof. Using that

$$\begin{aligned} (\partial_t + \mathcal{L} - r)v(t, r, x) = 0, \quad & r < c(t, x), \\ (\partial_t + \mathcal{L} - r)v(t, r, x) = (\partial_t + \mathcal{L} - r)(K - x) = -rK, \quad & r > c(t, x), \end{aligned} \tag{84}$$

we obtain that the dynamics of the discounted value function on $[0, \tau_n]$ is given by

$$
\begin{aligned}
&e^{-\int_0^{s\wedge\tau_n} r_v dv} v(t + s \wedge \tau_n, r_{s\wedge\tau_n}, X_{s\wedge\tau_n}) \qquad (85)\\
&= v(t, r, x) - \int_0^{s\wedge\tau_n} e^{-\int_0^u r_v dv} K r_u 1_{\{r_u > c(t+u, X_u)\}} du + \int_0^{s\wedge\tau_n} e^{-\int_0^u r_v dv} \sigma X_u v_x(t + u, r_u, X_u) dB_u\\
&\quad + \int_0^{s\wedge\tau_n} e^{-\int_0^u r_v dv} \beta(r_u) v_r(t + u, r_u, X_u) dW_u.
\end{aligned}
$$

Taking expectations and applying the optional sampling theorem, we arrive at

$$
v(t, r, x) = \mathsf{E}_{r,x}\left[\int_0^{\tau\wedge\tau_n} e^{-\int_0^u r_v dv} K r_u 1_{\{r_u > c(t+u, X_u)\}} du + e^{-\int_0^{\tau\wedge\tau_n} r_v dv} v(t + (\tau \wedge \tau_n), r_{\tau\wedge\tau_n}, X_{\tau\wedge\tau_n})\right]. \qquad (86)
$$

Using Equations (3) and (4), Hölder inequality implies

$$
\mathsf{E}_{r,x}\left[\int_0^{T-t} e^{-\int_0^u r_v dv} |K r_u| du\right] < \infty.
$$

The majorant for the second term of Equation (86) follows from Assumption 2.1 (details can be found in the proof of (90) in Lemma 9.1). The dominated convergence theorem proves Equation (18), since $\tau_n \uparrow T - t$ upon recalling that the boundary of $\mathcal{I} \times \mathbb{R}_+$ is assumed nonattainable by the process (r_t, X_t).

It remains to verify assumptions of (Cai and De Angelis 2021, Thm. 2.1) as stated in Appendix C. Identifying $X_t^1 = r_t$ and $X_t^2 = X_t$, we have

$$
\beta^{1,1}(t, r, x) = \beta^2(r), \quad \beta^{1,2}(t, r, x) = \beta^{2,1}(t, r, x) = \sigma\rho\beta(r)x, \quad \beta^{2,2}(t, r, x) = \sigma^2 x^2.
$$

By Assumption 2.1, $\beta^{i,j}$ is Lipschitz for $i, j = 1, 2$ on every compact set in $\mathcal{O}$. Indeed, it can be directly verified for the CIR process. In case (ii) of Assumption 2.1, we use Lipschitz continuity of β. The marginal distribution of the process (r_t, X_t) has density with respect to the Lebesgue measure (see Remark B.1, which makes use of Assumption 3.8), so $(t, r_t, X_t) \notin \partial\mathcal{C}$, $\mathsf{P}_{r,x}$-a.s. for any $t > 0$. Setting $\mathcal{C} = \mathcal{E}$, this verifies the first assumption in the theorem in Appendix C. For the second assumption, using Equation (84), we have

$$
\frac{1}{2} L(t, r, x) = -r x v_x(t, r, x) - \alpha(r) v_r(t, r, x) - v_t(t, r, x) + r v(t, r, x) - 1_{\{(t,r,x)\in\mathcal{D}\}}\, rK.
$$

Since $v \in C^1(\mathcal{O})$ and the function $\alpha(r)$ is continuous (see Assumption 2.1), L is continuous and bounded on $\mathcal{K}_n \setminus \partial\mathcal{C}$. We finally have that the third assumption in the theorem holds by Proposition 3.3. □

Proof of Proposition 3.11. The proof follows ideas originally developed in Peskir (2005). Assume there exists another continuous function $\tilde{c}$ that satisfies conditions (1) and (2) in the statement of this proposition. Define a function

$$\tilde{v}(t,r,x) = \mathsf{E}_{r,x}\left[\int_0^{T-t} e^{-\int_0^u r_v dv} K r_u 1_{\{r_u > \tilde{c}(t+u,X_u)\}} du + e^{-\int_0^{T-t} r_v dv}(K - X_{T-t})^+\right], \quad (t,r,x) \in \mathcal{O},$$

$$\tilde{v}(T,r,x) = (K-x)^+, \quad (r,x) \in \mathcal{I} \times \mathbb{R}_+.$$

It is not difficult to prove that $\tilde{v}$ is continuous by the continuity of $\tilde{c}$ and of the flow $(s,r,x) \mapsto (r_s^r, X_s^{r,x})$. By the Markov property of (r,X), one can also check that

$$\tilde{V}_s := \int_0^s e^{-\int_0^u r_v dv} K r_u 1_{\{r_u > \tilde{c}(t+u,X_u)\}} du + e^{-\int_0^s r_v dv} \tilde{v}(t+s, r_s, X_s), \qquad s \in [0, T-t],$$

is a continuous $\mathsf{P}_{r,x}$-martingale. Hence, for any $(t,r,x) \in \mathcal{O}$ and any stopping time $\tau \le T-t$, the optional sampling theorem yields

$$\tilde{v}(t,r,x) = \mathsf{E}_{r,x}[\tilde{V}_\tau] = \mathsf{E}_{r,x}\left[\int_0^\tau e^{-\int_0^u r_v dv} K r_u 1_{\{r_u > \tilde{c}(t+u,X_u)\}} du + e^{-\int_0^\tau r_v dv} \tilde{v}(t+\tau, r_\tau, X_\tau)\right], \tag{87}$$

which is analogous to the formula for v in Equation (18).

For an easier exposition of the arguments of proof we proceed in steps. In the first four steps, we show the equality $\tilde{c}(t,x) = c(t,x)$ for all $(t,x) \in [0,T) \times \mathbb{R}_+$ such that $\tilde{c}(t,x) \in \mathcal{I}$. Then, in the final step, we use monotonicity and continuity of $\tilde{c}$ and c to extend the equality to all $(t,x) \in [0,T) \times \mathbb{R}_+$.

Step 1. We first show that $\tilde{v}(t,r,x) = (K-x)^+$ for any $(t,r,x) \in \mathcal{O}$ such that $r \ge \tilde{c}(t,x)$. Fix $(\hat{t},\hat{r},\hat{x}) \in \mathcal{O}$ such that $\hat{r} > \tilde{c}(\hat{t},\hat{x})$ (the claim for $\hat{r} = \tilde{c}(\hat{t},\hat{x})$ follows by the continuity of $\tilde{v}$). Define a stopping time

$$\tau_1 := \inf\{s \ge 0 : r_s^{\hat{r}} \le \tilde{c}(\hat{t}+s, X_s^{\hat{r},\hat{x}})\} \wedge (T-\hat{t}).$$

By the continuity of $s \mapsto \tilde{c}(\hat{t}+s, X_s)$ and $s \mapsto r_s$, and the fact that $\underline{r}$ and $\bar{r}$ are unattainable by (r_s), we have $\tilde{c}(\hat{t}+\tau_1, X_{\tau_1}) \in \mathcal{I}$ on $\{\tau_1 < T-\hat{t}\}$. By assumption $\tilde{v}(t, \tilde{c}(t,x), x) = (K-x)^+$ and, consequently, $\tilde{v}(\hat{t}+\tau_1, \tilde{c}(\hat{t}+\tau_1, X_{\tau_1}), X_{\tau_1}) = (K-X_{\tau_1})^+$ since $\tilde{v}(T,r,x) = (K-x)^+$. In combination with Equation (87), this yields

$$\tilde{v}(\hat{t},\hat{r},\hat{x}) = \mathsf{E}_{\hat{r},\hat{x}}\left[\int_0^{\tau_1} e^{-\int_0^u r_v dv} K r_u du + e^{-\int_0^{\tau_1} r_v dv}(K-X_{\tau_1})^+\right], \tag{88}$$

where we use that $r_u > c(\hat{t}+u, X_u)$ on $\{u < \tau_1\}$. Applying Tanaka's formula to $(r,x) \mapsto (K-x)^+$ and taking expectation, we get

$$\begin{aligned}(K-\hat{x})^+ &= \mathsf{E}_{\hat{r},\hat{x}}\left[\int_0^{\tau_1} e^{-\int_0^u r_v dv} K r_u 1_{\{X_u < K\}} du + e^{-\int_0^{\tau_1} r_v dv}(K-X_{\tau_1})^+ + \frac{1}{2}\int_0^{\tau_1} e^{-\int_0^u r_v dv} dL_u^K(X)\right] \\ &= \mathsf{E}_{\hat{r},\hat{x}}\left[\int_0^{\tau_1} e^{-\int_0^u r_v dv} K r_u du + e^{-\int_0^{\tau_1} r_v dv}(K-X_{\tau_1})^+\right],\end{aligned}$$

where $L^K(X)$ is the local time of the process X at K. The local time $L^K(X)$ is null until τ_1 since $r_u > \tilde{c}(t+u, X_u) \Rightarrow X_u < K$, recalling that $\tilde{c}(t,x) = \bar{r}$ when $x \ge K$. Compare the right-hand side of the above expression to Equation (88) to conclude that $\tilde{v}(\hat{t},\hat{r},\hat{x}) = (K-\hat{x})^+$.

Step 2. The next step is to show that $\tilde{v} \le v$ for $(t,r,x) \in \mathcal{O}$. Since we have already proved $\tilde{v}(t,r,x) = (K-x)^+ \le v(t,r,x)$ when $r \ge \tilde{c}(t,x)$, we take $(\hat{t},\hat{r},\hat{x}) \in \mathcal{O}$ such that $\hat{r} < \tilde{c}(\hat{t},\hat{x})$. Define a stopping time

$$\tau_2 := \inf\{s \geq 0 : r_s^{\hat{r}} \geq \tilde{c}(\hat{t}+s, X_s^{\hat{r},\hat{x}})\} \wedge (T-\hat{t}).$$

Since $r_u < \tilde{c}(\hat{t}+u, X_u)$ on $\{u < \tau_2\}$, we obtain from Equation (87)

$$\tilde{v}(\hat{t},\hat{r},\hat{x}) = \mathsf{E}_{\hat{r},\hat{x}}\left[e^{-\int_0^{\tau_2} r_v dv}(K - X_{\tau_2})^+\right] \leq v(\hat{t},\hat{r},\hat{x}),$$

where the first equality is by $\tilde{v}(\hat{t}+\tau_2, \tilde{c}(\hat{t}+\tau_2, X_{\tau_2}), X_{\tau_2}) = (K - X_{\tau_2})^+$ and the final inequality holds by the definition of v.

Step 3. Now we show that $\tilde{c}(t,x) \leq c(t,x)$ for any $(t,x) \in [0,T) \times (0,K)$ such that $\tilde{c}(t,x) \in \mathcal{I}$ (it is immediate for $(t,x) \in [0,T) \times [K,\infty)$ as $\tilde{c}(t,x) = c(t,x) = \overline{r}$). Arguing by contradiction, assume that there exists $(\hat{t},\hat{x}) \in [0,T) \times (0,K)$ such that $\mathcal{I} \ni \tilde{c}(\hat{t},\hat{x}) > c(\hat{t},\hat{x})$. Let $\hat{r} > \tilde{c}(\hat{t},\hat{x})$, and define

$$\tau_3 := \inf\{s \geq 0 : r_s^{\hat{r}} \leq c(\hat{t}+s, X_s^{\hat{r},\hat{x}})\} \wedge (T-\hat{t}).$$

By Equations (18) and (87), we have

$$v(\hat{t},\hat{r},\hat{x}) = \mathsf{E}_{\hat{r},\hat{x}}\left[\int_0^{\tau_3} e^{-\int_0^u r_v dv} K r_u 1_{\{r_u > c(\hat{t}+u,X_u)\}} du + e^{-\int_0^{\tau_3} r_v dv} v(\hat{t}+\tau_3, r_{\tau_3}, X_{\tau_3})\right],$$

$$\tilde{v}(\hat{t},\hat{r},\hat{x}) = \mathsf{E}_{\hat{r},\hat{x}}\left[\int_0^{\tau_3} e^{-\int_0^u r_v dv} K r_u 1_{\{r_u > \tilde{c}(\hat{t}+u,X_u)\}} du + e^{-\int_0^{\tau_3} r_v dv} \tilde{v}(\hat{t}+\tau_3, r_{\tau_3}, X_{\tau_3})\right].$$

Since $\tilde{v}(\hat{t},\hat{r},\hat{x}) = (K-\hat{x})^+ = v(\hat{t},\hat{r},\hat{x})$, $r_u > c(\hat{t}+u, X_u)$ on $\{u < \tau_3\}$, and $\tilde{v} \leq v$, the above two equations imply that

$$\mathsf{E}_{\hat{r},\hat{x}}\left[\int_0^{\tau_3} e^{-\int_0^u r_v dv} K r_u 1_{\{r_u > \tilde{c}(\hat{t}+u,X_u)\}} du\right] \geq \mathsf{E}_{\hat{r},\hat{x}}\left[\int_0^{\tau_3} e^{-\int_0^u r_v dv} K r_u du\right].$$

Since the function c is non-negative, $r_u \geq 0$ on $\{u < \tau_3\}$, which allows us to conclude that

$$\mathsf{E}_{\hat{r},\hat{x}}\left[\int_0^{\tau_3} 1_{\{r_u \leq \tilde{c}(\hat{t}+u,X_u)\}} du\right] = 0. \tag{89}$$

The dynamics of (r,X) is nondegenerate on $\mathcal{I} \times \mathbb{R}_+$, so the density of (r_u, X_u) has a full support (on $\mathcal{I} \times \mathbb{R}_+$) for $u > 0$ (this can be inferred by classical Gaussian bounds as those we use in Equation (B.4) in Appendix). Hence, by the continuity of $\tilde{c}$ and c, for a sufficiently small $\varepsilon > 0$,

$$\mathsf{P}_{\hat{r},\hat{x}}\big(c(\hat{t}+u, X_u) < r_u < \tilde{c}(\hat{t}+u, X_u) \text{ for some } u \in (0,\varepsilon)\big) > 0.$$

Paired with the continuity of trajectories of (r,X), it contradicts Equation (89).

Step 4. Next, we prove $\tilde{c} = c$ at all points such that $\tilde{c} \in \mathcal{I}$. Arguing by contradiction, assume $\tilde{c}(\hat{t},\hat{x}) < c(\hat{t},\hat{x})$ for some $(\hat{t},\hat{x}) \in [0,T) \times (0,K)$ such that $\tilde{c}(\hat{t},\hat{x}) \in \mathcal{I}$. Let $\hat{r} \in (\tilde{c}(\hat{t},\hat{x}), c(\hat{t},\hat{x}))$ and define

$$\tau_4 := \inf\{s \geq 0 : r_s^{\hat{r}} \geq c(\hat{t}+s, X_s^{\hat{r},\hat{x}})\} \wedge (T-\hat{t}).$$

By Equations (18) and (87), we have

$$v(\hat{t},\hat{r},\hat{x}) = \mathsf{E}_{\hat{r},\hat{x}}\left[e^{-\int_0^{\tau_4} r_v dv} v(\hat{t}+\tau_4, r_{\tau_4}, X_{\tau_4})\right],$$

$$\tilde{v}(\hat{t},\hat{r},\hat{x}) = \mathsf{E}_{\hat{r},\hat{x}}\left[\int_0^{\tau_4} e^{-\int_0^u r_v dv} K r_u 1_{\{r_u > \tilde{c}(\hat{t}+u,X_u)\}} du + e^{-\int_0^{\tau_4} r_v dv}\tilde{v}(\hat{t}+\tau_4, r_{\tau_4}, X_{\tau_4})\right],$$

where in the first expression, we used that $1_{\{r_u > c(\hat{t}+u,X_u)\}} = 0$ on $\{u < \tau_4\}$. Since $\tilde{c}(t,x) \le c(t,x)$ for $(t,x) \in [0,T) \times (0,K)$, we have $\tilde{v}(\hat{t}+\tau_4, r_{\tau_4}, X_{\tau_4}) = (K - X_{\tau_4})^+ = v(\hat{t}+\tau_4, r_{\tau_4}, X_{\tau_4})$ by step 1. Then recalling that $\tilde{v} \le v$ and comparing the two equations above give us

$$\mathsf{E}_{\hat{r},\hat{x}}\left[\int_0^{\tau_4} e^{-\int_0^u r_v dv} K r_u 1_{\{r_u > \tilde{c}(\hat{t}+u,X_u)\}} du\right] \le 0.$$

This is a contradiction since by the continuity of (r, X) and $\tilde{c}$ there is a random variable $\eta > 0$ such that

$$r_u(\omega) > \tilde{c}(\hat{t}+u, X_u(\omega)) \quad \text{for all } u \in [0, \eta(\omega)).$$

Step 5. Here we show that $\tilde{c} = c$ on $[0,T) \times \mathbb{R}_+$. Let (t_n, x_n) be a sequence such that $\tilde{c}(t_n,x_n) \in \mathcal{I}$ and $(t_n,x_n) \to (t_0,x_0)$ with $\tilde{c}(t_0,x_0) = \bar{r}$ (respectively, $\tilde{c}(t_0,x_0) = 0$). Since $\tilde{c}(t_n,x_n) = c(t_n,x_n)$ for all n's, by the four steps above, by continuity, we also get $c(t_0,x_0) = \tilde{c}(t_0,x_0) = \bar{r}$ (respectively, $c(t_0,x_0) = \tilde{c}(t_0,x_0) = 0$). Then, by the monotonicity of both c and $\tilde{c}$, we get $c(t,x) = \tilde{c}(t,x)$ for all $(t,x) \in [0,t_0] \times [x_0,\infty)$ (respectively, $(t,x) \in [t_0,T] \times [0,x_0]$). This implies, in particular, that

$$\{(t,x) : \tilde{c}(t,x) \in \mathcal{I}\} = \{(t,x) : c(t,x) \in \mathcal{I}\},$$

which concludes the proof. □

Proof of Corollary 3.12. We can repeat the same arguments as in the proof of Proposition 3.11, always using $x > b(t,r) \iff r < c(t,x)$ to fall back into the exact set-up of steps 1–4 therein. □

9 | HEDGING STRATEGY

We start this section with an auxiliary lemma whose assertions are used to show admissibility of the hedging strategy. Estimate (90) is also used in the proof of Proposition 3.9.

Lemma 9.1. *For any compact set $\mathcal{K} \subset \mathcal{I} \times \mathbb{R}_+$, and $p \in [1,2]$, we have*

$$\sup_{(r,x)\in\mathcal{K}} \sup_{t\in[0,T]} \mathsf{E}_{r,x}\left[\sup_{0\le s\le T-t} e^{-\int_0^s r_u du} v(t+s, r_s, X_s)\right] < \infty, \tag{90}$$

$$\sup_{(r,x)\in\mathcal{K}} \sup_{s\in[0,T]} \mathsf{E}_{r,x}\left[e^{-p\int_0^s r_u du} |v_x(t+s, r_s, X_s)|^p X_s^p\right] < \infty, \tag{91}$$

$$\sup_{(r,x)\in\mathcal{K}} \sup_{s\in[0,T]} \mathsf{E}_{r,x}\left[e^{-2\int_0^s r_u du} (v_r(t+s, r_s, X_s))^2 \beta^2(r_s)\right] < \infty. \tag{92}$$

Proof. From Equation (6), we obtain an upper bound for the function v:

$$v(t,r,x) \le K\mathsf{E}_r\left[\sup_{0\le s\le T-t} e^{-\int_0^s r_u du}\right]. \tag{93}$$

Using this bound, we have $P_{r,x}$-a.s.

$$\begin{aligned} e^{-\int_0^s r_u du} v(t+s, r_s, X_s) &\le e^{-\int_0^s r_u du} K \mathsf{E}_{r_s}\left[\sup_{0\le u\le T-t-s} e^{-\int_0^u r_v dv} \right] \\ &= e^{-\int_0^s r_u du} K \mathsf{E}_{r}\left[\sup_{s\le u\le T-t} e^{-\int_s^u r_v dv} \middle| \mathcal{F}_s \right] \\ &\le K \mathsf{E}_{r}\left[\sup_{0\le u\le T-t} e^{-\int_0^u r_v dv} \middle| \mathcal{F}_s \right] \le K \mathsf{E}_{r}\left[\sup_{0\le u\le T} e^{-\int_0^u r_v dv} \middle| \mathcal{F}_s \right], \end{aligned}$$

where in the second equality, we employ the Markov property of r. By Doob's maximal inequality applied to the martingale $Y_s = \mathsf{E}_r[\sup_{0\le u\le T} e^{-\int_0^u r_v dv} | \mathcal{F}_s]$, we conclude

$$\begin{aligned} &\sup_{(r,x)\in\mathcal{K}} \sup_{t\in[0,T]} \mathsf{E}_{r,x}\left[\sup_{0\le s\le T-t} e^{-\int_0^s r_u du} v(t+s, r_s, X_s) \right] \\ &\le \sup_{(r,x)\in\mathcal{K}} \sup_{t\in[0,T]} K\, \mathsf{E}_r\left[\sup_{0\le s\le T-t} Y_s \right] \le \sup_{(r,x)\in\mathcal{K}} K\, \mathsf{E}_r\left[\sup_{0\le s\le T} Y_s \right] \\ &\le \sup_{(r,x)\in\mathcal{K}} 2K \left(\mathsf{E}_r[Y_T^2] \right)^{1/2} = \sup_{(r,x)\in\mathcal{K}} 2K \left(\mathsf{E}_r\left[\sup_{0\le u\le T} e^{-2\int_0^u r_v dv} \right] \right)^{1/2} \le 2K(C_1)^{1/2}, \end{aligned}$$

where C_1 is the constant from Equation (3). This proves (i).

We now address Equation (91). We have

$$e^{-p\int_0^s r_u du} \left| v_x(t+s, r_s, X_s) \right|^p X_s^p = \left| v_x(t+s, r_s, X_s) \right|^p x^p e^{p\sigma B_s - \frac{p}{2}\sigma^2 s} \le x^p e^{p\sigma B_s - \frac{p}{2}\sigma^2 s},$$

where we use $-1 \le v_x \le 0$ in the last inequality, which follows from Equation (62). From here, Equation (91) is immediate.

It remains to prove Equation (92). First, we consider the case of Assumption 2.1(ii). From Equations (39), (40), and (41), we deduce

$$(v_r(t,r,x))^2 \le c_1\, \mathsf{E}_r\left[\sup_{0\le s\le T-t} e^{-2\int_0^s r_u du} \right], \tag{94}$$

for some constant $c_1 > 0$ depending only on T and the coefficients of Equation (2) (notice in particular that the expected value in the right-hand side above comes from the constant C_1 in Equation (40)). Hence

$$\begin{aligned} e^{-2\int_0^s r_u du} (v_r(t+s, r_s, X_s))^2 \beta^2(r_s) &\le e^{-2\int_0^s r_u du} c_1\, \mathsf{E}_{r_s}\left[\sup_{0\le u\le T-t-s} e^{-2\int_0^u r_v dv} \right] \beta^2(r_s) \\ &\le c_1\, \mathsf{E}_r\left[\sup_{0\le u\le T-t} e^{-2\int_0^u r_v dv} \middle| \mathcal{F}_s \right] \beta^2(r_s), \end{aligned}$$

where the last inequality is by the same argument as in the proof of Equation (90). We take expectation of both sides and apply Hölder inequality with $q = p'/2$ ($p' > 2$ is defined in Assumption 2.1) and $q' = q/(q-1)$

$$\begin{aligned}\mathsf{E}_r\left[e^{-2\int_0^s r_u du}(v_r(t+s,r_s,X_s))^2\beta^2(r_s)\right] &\le c_1\left(\mathsf{E}_r\left[\mathsf{E}_r\left[\sup_{0\le u\le T-t} e^{-2\int_0^u r_v dv}\Big|\mathcal{F}_s\right]^q\right]\right)^{1/q}\left(\mathsf{E}_r[\beta^{2q'}(r_s)]\right)^{1/q'}\\ &\le c_1\left(\mathsf{E}_r\left[\sup_{0\le u\le T-t} e^{-p'\int_0^u r_v dv}\right]\right)^{1/q}\left(\mathsf{E}_r[\beta^{2q'}(r_s)]\right)^{1/q'}\\ &\le c_1 C_1^{1/q}\left(\mathsf{E}_r[\beta^{2q'}(r_s)]\right)^{1/q'},\end{aligned}$$

where the second inequality follows from Jensen's inequality and C_1 is the constant from Equation (3). Let L be the Lipschitz constant for β. Then, using triangle inequality for norms

$$\begin{aligned}\left(\mathsf{E}_r[(\beta(r_s))^{2q'}]\right)^{1/q'} \le \left(\mathsf{E}_r\left[\left|\beta(0)+L|r_s|\right|^{2q'}\right]\right)^{1/q'} &= \left(\left(\mathsf{E}_r\left[\left|\beta(0)+L|r_s|\right|^{2q'}\right]\right)^{1/2q'}\right)^2\\ \le \left(\beta(0)+L\left(\mathsf{E}_r\left[|r_s|^{2q'}\right]\right)^{1/2q'}\right)^2 &\le \left(\beta(0)+L\left(C_2(1+|r|^{2q'}\right)^{1/2q'}\right)^2,\end{aligned}$$

where the last inequality follows from Equation (4) and $2q' = p' \ge 2$. Combining the above estimates proves Equation (92).

We address the case when r follows the CIR dynamics. From the non-negativity of the process r and from Equation (94), we obtain that $(v_r(t,r,x))^2 \le c_1$ for any $(t,r,x) \in \mathcal{O}$. Hence, we write

$$\mathsf{E}_{r,x}\left[e^{-2\int_0^s r_u du}(v_r(t+s,r_s,X_s))^2\beta^2(r_s)\right] \le c_1\gamma^2\mathsf{E}_r[|r_s|],$$

where we used the explicit form of β. It remains to recall Equation (4) to conclude Equation (92). □

Proof of Proposition 3.13. The admissibility condition can be equivalently written as

$$\int_0^T e^{-2\int_0^s r_u du}\left(\phi_s^{(1)}\sigma X_s\right)^2 ds + \int_0^T e^{-2\int_0^s r_u du}\left(\phi_s^{(2)}\beta(r_s)P_r(s,r_s)\right)^2 ds < \infty, \qquad \mathsf{P}_{r,x}\text{-a.s.} \tag{95}$$

Estimates in Lemma 9.1 imply

$$\mathsf{E}_{r,x}\left[\int_0^T e^{-2\int_0^s r_u du}\left(\phi_s^{(1)}\sigma X_s\right)^2 ds + \int_0^T e^{-2\int_0^s r_u du}\left(\phi_s^{(2)}\beta(r_s)P_r(s,r_s)\right)^2 ds\right] < \infty,$$

which is a stronger condition than Equation (95). The fact that the portfolio replicates the option follows from the construction and Equation (30). □

ACKNOWLEDGMENTS

C. Cai gratefully acknowledges support by China Scholarship Council. T. De Angelis gratefully acknowledges support by EPSRC Grant EP/R021201/1.

DATA AVAILABILITY STATEMENT

Data sharing not applicable to this article as no datasets were generated or analyzed during the current study.

ORCID

Tiziano De Angelis https://orcid.org/0000-0002-0164-7936
Jan Palczewski https://orcid.org/0000-0003-0235-8746

REFERENCES

Amin, K. I., & Bodurtha Jr, J. N. (1995). Discrete-time valuation of American options with stochastic interest rates. *The Review of Financial Studies*, *8*(1), 193–234.

Amin, K. I., & Jarrow, R. A. (1992). Pricing options on risky assets in a stochastic interest rate economy. *Mathematical Finance*, *2*(4), 217–237.

Appolloni, E., Caramellino, L., & Zanette, A. (2015). A robust tree method for pricing American options with the Cox–Ingersoll–Ross interest rate model. *IMA Journal of Management Mathematics*, *26*(4), 377–401.

Bass, R. F. (1998). *Diffusions and elliptic operators*. Springer Science & Business Media.

Battauz, A., & Rotondi, F. (2022). American options and stochastic interest rates. *Computational Management Science*. https://doi.org/10.1007/s10287-022-00427-x

Bensoussan, A. (1984). On the theory of option pricing. *Acta Applicandae Mathematica*, *2*(2), 139–158.

Broadie, M., & Detemple, J. (1996). American option valuation: New bounds, approximations, and a comparison of existing methods. *The Review of Financial Studies*, *9*(4), 1211–1250.

Cai, C., & De Angelis, T. (2021). A change of variable formula with applications to multi-dimensional optimal stopping problems. *arXiv:2104.05835*.

Cai, C., De Angelis, T., & Palczewski, J. (2021). On the continuity of optimal stopping surfaces for jump-diffusions. *arXiv:2104.08502*.

Carr, P., Jarrow, R., & Myneni, R. (1992). Alternative characterizations of American put options. *Mathematical Finance*, *2*(2), 87–106.

Chung, S.-L. (2000). American option valuation under stochastic interest rates. *Review of Derivatives Research*, *3*(3), 283–307.

De Angelis, T., & Peskir, G. (2020). Global C^1 regularity of the value function in optimal stopping problems. *Annals of Applied Probability*, *30*(3), 1007–1031.

Detemple, J., Kitapbayev, Y., & Zhang, L. (2018). *American option pricing under stochastic volatility models via Picard iterations* [Working paper].

Detemple, J., & Tian, W. (2002). The valuation of American options for a class of diffusion processes. *Management Science*, *48*(7), 917–937.

Fergusson, K., & Platen, E. (2015). Application of maximum likelihood estimation to stochastic short rate models. *Annals of Financial Economics*, *10*(02), 1550009.

Friedman, A. (1964). *Partial differential equations of parabolic type*. Prentice-Hall.

Geske, R., & Johnson, H. E. (1984). The American put option valued analytically. *The Journal of Finance*, *39*(5), 1511–1524.

Heath, D., Jarrow, R., & Morton, A. (1992). Bond pricing and the term structure of interest rates: A new methodology for contingent claims valuation. *Econometrica*, *60*, 77–105.

Ho, T. S., Stapleton, R. C., & Subrahmanyam, M. G. (1997). The valuation of American options with stochastic interest rates: A generalization of the Geske–Johnson technique. *The Journal of Finance*, *52*(2), 827–840.

Hull, J. (2009). *Options, futures and other derivatives*. Prentice Hall.

Jacka, S. (1991). Optimal stopping and the American put. *Mathematical Finance*, *1*(2), 1–14.

Jaillet, P., Lamberton, D., & Lapeyre, B. (1990). Variational inequalities and the pricing of American options. *Acta Applicandae Mathematica*, *21*(3), 263–289.

Jeanblanc, M., Yor, M., & Chesney, M. (2009). *Mathematical methods for financial markets*. Springer Science & Business Media.

Karatzas, I. (1988). On the pricing of American options. *Applied Mathematics & Optimization*, *17*, 37–60.

Karatzas, I., & Shreve, S. E. (1998a). *Brownian motion and stochastic calculus*. Springer.

Karatzas, I., & Shreve, S. E. (1998b). *Methods of mathematical finance* (Vol. 39). Springer.

Kim, I. J. (1990). The analytic valuation of American options. *The Review of Financial Studies*, *3*(4), 547–572.

Kim, I. J., Jang, B.-G., & Kim, K. T. (2013). A simple iterative method for the valuation of American options. *Quantitative Finance*, *13*(6), 885–895.

Kruse, R., & Deely, J. (1969). Joint continuity of monotonic functions. *The American Mathematical Monthly*, *76*(1), 74–76.

Krylov, N. (1980). *Controlled diffusion processes*. Springer.

Little, T., Pant, V., & Hou, C. (2000). A new integral representation of the early exercise boundary for American put options. *Journal of Computational Finance*, *3*, 73–96.

McKean Jr, H. P. (1965). A free boundary problem for the heat equation arising from a problem of mathematical economics. *Industrial Management Review*, *6*, 32–39.

Menkveld, A. J., & Vorst, T. (2000). A pricing model for American options with Gaussian interest rates. *Annals of Operations Research*, *100*(1-4), 211–226.

Myneni, R. (1992). The pricing of the American option. *The Annals of Applied Probability*, *2*(1), 1–23.

Peskir, G. (2005). On the American option problem. *Mathematical Finance*, *15*(1), 169–181.

Peskir, G. (2007). A change-of-variable formula with local time on surfaces. In *Séminaire de probabilités XL* (pp. 70–96). Springer.

Rutkowski, M. (1994). The early exercise premium representation of foreign market American options. *Mathematical Finance*, *4*(4), 313–325.

Shiryaev, A. N. (2008). *Optimal stopping rules. Stochastic modelling and applied probability* (Vol. 8). Springer-Verlag. Translated from the 1976 Russian second edition by A. B. Aries, Reprint of the 1978 translation.

Van Moerbeke, P. (1976). On optimal stopping and free boundary problems. *Archive for Rational Mechanics and Analysis*, *60*(2), 101–148.

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Cai, C., De Angelis, T., & Palczewski, J. (2022). The American put with finite-time maturity and stochastic interest rate. *Mathematical Finance*, *32,* 1170–1213. https://doi.org/10.1111/mafi.12361