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

(Article begins on next page)

06 February 2025

# THE AMERICAN PUT WITH FINITE-TIME MATURITY
# AND STOCHASTIC INTEREST RATE

CHENG CAI, TIZIANO DE ANGELIS, AND JAN PALCZEWSKI

ABSTRACT. In this paper we study pricing of American put options on 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 parametrisation 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.

## 1. INTRODUCTION

Pricing of American options is a classical problem in mathematical finance which has attracted continuous attention since the initial work of McKean in 1965 [36]. 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 [8, 28] 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 3-dimensional with 2-dimensional diffusive dynamics (stock price and interest rate) and time. The stopping set, i.e., the set of points $(t, r, x)$ for which is it 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, i.e., 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 generalisation of Itô's formula (due to [11]) 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 generalised 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 non-increasing 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 parametrised as a function of $(t, x)$, is continuous. Apart from being of interest in its own right, this enables a characterisation 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 focussing on approximations and numerical methods. Lattice (tree) based methods are employed by Appolloni, Caramellino and Zanette [3] to price options in Black and Scholes model with CIR interest rate dynamics and by Battauz and Rotondi [6] in a model with Vasicek interest rates. Geske and Johnson's ([21]) approximation of discretely exercised American options prices is adapted by Ho, Stapleton and Subrahmanyam [23] and Chung [13] 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 [37]. A framework for option pricing with Heath, Jarrow, Morton's [22] bond market model is developed by Amin and Jarrow [2] with a binomial-tree-based implementation of pricing of foreign exchange options performed in Amin and Bodurtha [1].

Detemple and Tian [16] 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 [8] using PDE methods and extended by Karatzas [28] 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 non-decreasing 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 [30, 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 characterisation of an optimal stopping boundary as solution to a (system of) integral equations has been known since the earliest works (see Van Moerbeke [44]). In more recent works [31, 25, 12, 38] 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 [39] where he shows that the stopping boundary is the unique continuous solution of a single integral equation. The key observation in [39] 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's [39] 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 [35, 32]. Our paper extends Peskir's [39] 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 [15] 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 [25] and subsequent works). That is a 'directional' derivative and continuity is only considered with respect to one variable. Sobolev space regularity is studied in [26] 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 [14] 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 [14] removing 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 [42] with methods from martingale theory. The nature of the methods employed in [42] to derive their main results is such that the emphasis is removed from the optimal boundary, which in fact only appears in specific examples ([42, Sec. 3]) 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.

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. Three 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 $\mathsf{P}$). We denote by $(\mathcal{F}_t)_{t \geq 0}$ the filtration generated by $(B, W)$ augmented with the $\mathsf{P}$-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

$$(2.1) \qquad dX_t = r_t X_t dt + \sigma X_t dB_t, \qquad X_0 = x,$$

$$(2.2) \qquad dr_t = \alpha(r_t) dt + \beta(r_t) dW_t, \qquad r_0 = r,$$

with $\alpha, \beta : \mathcal{I} \to \mathbb{R}$ specified below. The probability measure $\mathsf{P}$ 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 behaviour 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, e.g., Vasicek model. Therefore, we make the following standing assumption:

**Assumption 2.1.** *The coefficients $\alpha$ and $\beta$ in (2.2) 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) *$\alpha$ and $\beta$ are globally Lipschitz and continuously differentiable on bounded subsets of $\mathcal{I}$ with $\beta(r) > 0$ for all $r \in \mathcal{I}$, and $\bar{r} > 0 \geq \underline{r}$. 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*

$$(2.3) \qquad \sup_{r \in \mathcal{K}} \mathsf{E}\left[ \sup_{0 \leq s \leq T} e^{-p \int_0^s r_u du} \Big| r_0 = r \right] \leq C_1.$$

The assumption that $\bar{r} > 0$ cannot be relaxed without trivialising 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 (2.2) 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 e.g. [27, Sec. 6.3.1], so the left boundary $\underline{r} = 0$ is reflecting (also non-attainable if $\kappa\theta > \sigma^2/2$). Hence, the bound (2.3) is satisfied with the constant $C_1 = 1$. The linear growth of $\alpha$ and $\beta$ in (2.2) guarantees that for each $p \geq 2$ there is $C_2 > 0$ only depending on $T$ and $p$, such that [34, Thm. 2.5.9]

$$(2.4) \qquad \mathsf{E}\big[ \sup_{0 \leq s \leq T} |r_s|^p \,\big|\, r_0 = r \big] \leq C_2(1 + |r|^p), \qquad \text{for } r \in \mathcal{I}.$$

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

$$(2.5) \qquad X_t = x \exp\left( \sigma B_t + \int_0^t \left( r_s - \tfrac{\sigma^2}{2} \right) ds \right), \qquad \text{for } t \geq 0,$$

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 (2.5) 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 $\mathsf{P}_r(\cdot) = \mathsf{P}(\cdot \,|\, r_0 = r)$.

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

$$(2.6) \qquad v(t, r, x) = \sup_{0 \leq \tau \leq T-t} \mathsf{E}_{r,x}\left[ e^{-\int_0^\tau r_t dt} \left( K - X_\tau \right)^+ \right],$$

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 function $(\,\cdot\,)^+$ denotes the positive part and the optimisation runs over all $(\mathcal{F}_t)_{t \geq 0}$-stopping times bounded by $T - t$. The above is an optimal stopping problem with Markovian structure and a 3-dimensional state space.

Since the process

$$(2.7) \qquad t \mapsto e^{-\int_0^t r_s ds} \left( K - X_t \right)^+$$

is non-negative and continuous, and thanks to the integrability condition (2.3), we can rely on standard optimal stopping theory (see, e.g., [30, Appendix D]) to conclude that the smallest optimal

stopping time for (2.6) is $\mathsf{P}_{r,x}$-a.s. given by

$$(2.8) \qquad \tau_* := \inf\{s \geq 0 \,:\, v(t+s, r_s, X_s) = (K - X_s)^+\},$$

where we note that $\tau_* \leq T - t$ since $v(T, r, x) = (K - x)^+$. Clearly $\tau_* = \tau_*(t, r, x)$ depends on the initial value $(t, r, x)$ of the 3-dimensional state process $(t + s, r_s, X_s)_{s \geq 0}$.

The form (2.8) of $\tau^*$ gives rise to the so-called continuation set $\mathcal{C}$ and its complement, the stopping set $\mathcal{D}$, that is

$$(2.9) \qquad \mathcal{C} := \{(t, r, x) \in [0, T] \times \mathcal{I} \times \mathbb{R}_+ \,:\, v(t, r, x) > (K - x)^+\},$$

$$(2.10) \qquad \mathcal{D} := \{(t, r, x) \in [0, T] \times \mathcal{I} \times \mathbb{R}_+ \,:\, v(t, r, x) = (K - x)^+\}.$$

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 $(t, r_t, X_t) \in \mathcal{C}$ 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*

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

*(i.e., $Y$ is the discounted option value process), we have that [30, Appendix D]*

$$(2.11) \qquad (Y_s)_{s \in [0, T-t]} \text{ is a right-continuous } \mathsf{P}_{r,x}\text{-supermartingale},$$

$$(2.12) \qquad (Y_{s \wedge \tau_*})_{s \in [0, T-t]} \text{ is a right-continuous } \mathsf{P}_{r,x}\text{-martingale}.$$

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

**Notation.** We set

$$(2.13) \qquad \mathcal{O} := [0, T) \times \mathcal{I} \times \mathbb{R}_+,$$

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

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

$$(2.14) \qquad \mathcal{L}f := \frac{\sigma^2 x^2}{2} f_{xx} + \frac{\beta^2(r)}{2} f_{rr} + \rho \sigma x \beta(r) f_{rx} + rx f_x + \alpha(r) f_r,$$

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 parametrisations of it. Then, in Sections 3.3—3.6, under a mild additional assumption on $\alpha$ and $\beta$ (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 characterisation 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 non increasing for all $(r, x) \in \mathcal{I} \times \mathbb{R}_+$,*
(ii) $r \mapsto v(t, r, x)$ *is non increasing for all $(t, x) \in [0, T] \times \mathbb{R}_+$,*
(iii) $x \mapsto v(t, r, x)$ *is convex and non increasing for all $(t, r) \in [0, T] \times \mathcal{I}$.*

*Proof.* See Section 5. □

The monotonicity in $t$ and $x$ and the convexity in $x$ is 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 (2.7) is still uniformly integrable and continuous. This implies that, letting $v_\infty$ 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 (2.11)–(2.12) continue to hold in this setting (see, e.g., [43, Ch. 3, Thm. 3]).*

*Moreover, the proof of Proposition 3.1 can be repeated step by step, upon also noticing that*

$$\widetilde{X}_t := e^{-\int_0^t r_s ds} X_t, \qquad t \geq 0$$

*is a continuous martingale with $\lim_{t\to\infty} \widetilde{X}_t = 0$, $\mathsf{P}$-a.s. and therefore $\mathsf{E}_{r,x}[\widetilde{X}_\tau] = x$ (cf. [29, Thm. 1.3.22]). So $r \mapsto v_\infty(r, x)$ is non-increasing and $x \mapsto v_\infty(r, x)$ is convex and non-increasing.*

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 $\mathcal{D}$.

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

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

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

*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 non-increasing and the mapping $x \mapsto c(t_0, x)$ is left-continuous and non-decreasing.*
(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.

In models with constant interest rate, an optimal boundary is often defined as function of time which provides a threshold for the process $(X_t)$. An analogous representation of the continuation and stopping sets is valid in our model. The following proposition whose simple proof is omitted gives details of the reparametrisation of the stopping boundary.

**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 non-decreasing 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,*

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

The parametrisation of the stopping boundary via the function $c(t,x)$ will usually be more convenient. In particular, due to technical reasons that will become clearer in Section 6, we will be 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 $[0,K]$.

3.2. **Smoothness of the value function.** It is well-known that $v$ satisfies (in the classical sense)

$$(3.3) \quad \begin{aligned} v_t(t,r,x) + (\mathcal{L} - r)v(t,r,x) &= 0, & (t,r,x) \in \mathcal{C}, \\ v(t,r,x) &= (K-x)^+, & (t,r,x) \in \mathcal{D}, \end{aligned}$$

where $\mathcal{L}$ is the generator of $(r,X)$ defined in (2.14). Hence, standard arguments assert that $v$ is $C^{1,2,2}$ in $\mathcal{C} \cap \text{int}(\mathcal{D})$. 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 $v^\circ$ 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 $\partial\mathcal{C}$, i.e., the global continuity of the gradient of $v$ (as a function of all variables) in $\mathcal{O}$. We use ideas similar to those in [14] but we must refine arguments therein and use estimates with 'local' nature since we are not able to directly check the assumptions required in [14].

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

*Proof.* See Section 7. $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\square$

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 $\tau_*$ 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 $\alpha$ and $\beta$ in (2.2) 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. $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\square$

Summarising, 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 [40] which is often used in optimal stopping literature to establish Itô's formula for the value function. Indeed, since [40] 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 generalisation of Itô's formula from [11] 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 [11]. We do however need to ensure that both boundary points of the set $\mathcal{I}$ are non-attainable, 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 non-attainable 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:*

$$(3.4) \qquad v(t, r, x) = \mathsf{E}_{r,x} \left[ \int_0^\tau e^{-\int_0^u r_v dv} K r_u 1_{\{r_u > c(t+u, X_u)\}} du + e^{-\int_0^\tau r_v dv} v(t+\tau, r_\tau, X_\tau) \right].$$

*Proof.* See Section 8. $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\square$

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

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

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

3.4. **Early exercise premium.** Inserting $\tau = T - t$ in (3.4), 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 [42] for a derivation of this formula only using general martingale theory):

$$(3.6) \qquad\qquad\qquad v(t, r, x) = v_p(t, r, x; T, b) + v_e(t, r, x; T),$$

where

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

The last equality follows from $r > c(t, x) \Leftrightarrow x < b(t, r)$ by construction of $b$ as the generalised inverse of $c$.

3.5. **Integral equation for the stopping boundary.** Proposition 3.9 provides a characterisation 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 (3.4) yields an integral equation for $c$:

$$(3.8) \quad (K-x)^+ = \mathsf{E}_{c(t,x),x}\left[\int_0^{T-t} e^{-\int_0^u r_v dv} K r_u 1_{\{r_u > c(t+u, X_u)\}} du + e^{-\int_0^{T-t} r_v dv}(K - X_{T-t})^+\right].$$

The condition that $c(t,x) \in \mathcal{I}$ is necessary as $c$ can take values $\underline{r}$ and $\overline{r}$ which do not belong do 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 (3.8) can be replaced by $(K - x)$. In line with well-known results for American options with constant interest rate [39], it also turns out that $c$ is the unique solution of the integral equation.

**Proposition 3.10.** *Under Assumptions 3.6 and 3.8, the function $c$ is the unique function $\phi$ : $[0,T) \times \mathbb{R}_+ \to [0,\overline{r}]$ such that:*
  *(1) is continuous, non-decreasing in $x$ and non-increasing in $t$, with $\phi(t,x) = \overline{r}$ for $x \geq K$,*
  *(2) $\phi$ satisfies (3.8) (with $c$ therein replaced by $\phi$) for all $(t,x) \in [0,T) \times \mathbb{R}_+$ for which $\phi(t,x) \in \mathcal{I}$.*

The integral equation (3.8) 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:

$$(3.9) \quad \begin{aligned} K - b(t,r) &= \mathsf{E}_{r,b(t,r)}\left[\int_0^{T-t} e^{-\int_0^u r_v dv} K r_u 1_{\{X_u < b(t+u, r_u)\}} du\right] \\ &\quad + \mathsf{E}_{r,b(t,r)}\left[e^{-\int_0^{T-t} r_v dv}(X_{T-t} - K)^+\right], \end{aligned}$$

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 generalised inverse of $c$. This parametrisation 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 non-standard result for this type of equations.

**Corollary 3.11.** *Under the assumptions of Proposition 3.10, 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 non-decreasing,*
  *(2) the generalised left-continuous inverse $\phi(t,x) := \inf\{r \in \mathcal{I} : \psi(t,r) \geq x\}$ is continuous in $(t,x)$, non-decreasing in $x$ and non-increasing in $t$,*
  *(3) $\psi$ satisfies (3.9) with (b therein replaced by $\psi$) for all $(t,r) \in [0,T) \times \mathcal{I}$ such that $\psi(t,r) > 0$.*

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

Integral equations (3.8) and (3.9) 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 (3.6) 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 (3.5) 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 (2.1), 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 [28].

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

$$(3.10) \qquad P(t, r) := \mathsf{E}_r \left[ e^{-\int_0^{T-t} r_u du} \right], \quad P(T, r) = 1.$$

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 [20, 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

$$(3.11) \qquad de^{-\int_0^s r_u du} P(s, r_s) = P_r(s, r_s)\beta(r_s)dW_s.$$

Denote by $\phi^{(1)}, \phi^{(2)}, \phi^{(3)}$ the holdings in the stock, the bond and the money account, respectively. Let $C$ be a non-decreasing continuous process starting from 0 modelling consumption. The value of a self-financing portfolio starting at time 0 from $v(0, r, x)$ is

$$(3.12) \qquad \Pi_s = v(0, r, x) + \int_0^s \phi_s^{(1)} dX_s + \int_0^s \phi_s^{(2)} dP(s, r_s) + \int_0^s \phi_s^{(3)} dM_s - C_s, \quad s \in [0, T].$$

The portfolio is *admissible* if all integrals above are semimartingales. Taking the money-market account as a numeráire, we obtain from equations (3.12) and

$$(3.13) \qquad \Pi_s := \phi_s^{(1)} X_s + \phi_s^{(2)} P(s, R_s) + \phi_s^{(3)} M_s, \quad s \in [0, T],$$

that the dynamics of the discounted portfolio value reads

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

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

Comparing (3.14) with (3.5), a candidate for the hedging strategy is given by

$$(3.15) \qquad \phi_s^{(1)} = v_x(s, r_s, X_s), \quad \phi_s^{(2)} = \frac{v_r(s, r_s, X_s)}{P_r(s, r_s)}, \quad C_s = \int_0^s K r_u 1_{\{r_u > c(u, X_u)\}} du.$$

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

**Proposition 3.12.** *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. $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\quad$ $\square$

## 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

$$(4.1) \qquad dr_t = \kappa(\theta - r_t)dt + \beta dW_t,$$

whose explicit solution is given by

$$(4.2) \qquad r_s = r_t e^{-(s-t)\kappa} + \theta(1 - e^{-(s-t)\kappa}) + \beta e^{-s\kappa} \int_t^s e^{\kappa u} dW_u, \quad s \geq t \geq 0.$$

We first derive a numerical method for computing the optimal stopping boundary using the integral equation from (3.9). Once the boundary is obtained, we use it to also compute the value function via (3.6). 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. (3.10)); the dependence on the initial state $r$ is indicated in the subscript of the expectation operator. Recall the integral equation (3.9) for the boundary $b$: for $(t,r) \in [0,T) \times \mathcal{I}$ such that $b(t,r) > 0$, we have

$$(4.3) \qquad K - b(t,r) = v_p(t,r,b(t,r);T,b) + v_e(t,r,b(t,r);T),$$

where $v_e$ and $v_p$ are stated in (3.7). With the last parameter $b$ of $v_p$, we emphasise 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 C, we derive the following formulas for $v_e$ and $v_p$ using well-known properties of the joint law of $(r_t, X_t)$:

$$(4.4) \qquad v_e(t,r,x;T) = P(t,T)K\mathcal{N}(d_1) - x\mathcal{N}(d_2),$$

$$(4.5) \qquad v_p(t,r,x;T,b) = \int_t^T KP(t,u)\left[\int_{-\infty}^\infty \frac{1}{\sqrt{2\pi}} e^{-\frac{y^2}{2}}\left(q(t,u) + y\sqrt{\gamma_2(t,u)}\right)\mathcal{N}\big(\phi(t,u,y;b)\big) dy\right] du,$$

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

Equation (4.3) defines the boundary $b$ as a fixed point of a non-linear mapping. To compute it, we follow an iterative scheme motivated by [15]. We fix $-\infty < r_{min} < r_{max} < \infty$ and discretise the variables $(t,r)$ as follows:

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

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:

$$(4.6) \qquad K - b^{(n)}(t_i, r_j) - v_e\left(t_i, r_j, b^{(n)}(t_i, r_j); T\right) = v_p\left(t_i, r_j, b^{(n-1)}(t_i, r_j); T, b^{(n-1)}\right).$$

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 pre-determined $\varepsilon > 0$,

$$\max_{i,j} |b^{(n-1)}(t_i, r_j) - b^{(n)}(t_i, r_j)| < \varepsilon.$$
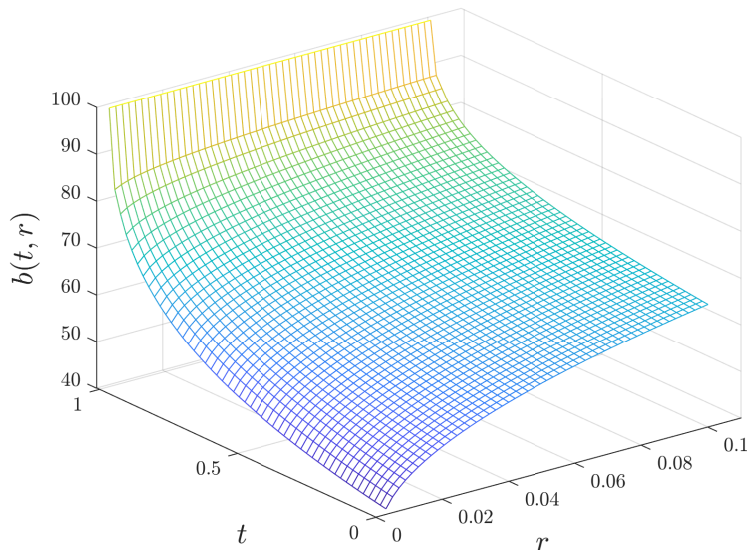
FIGURE 1. Stopping boundary surface $b(t, r)$.

The numerical evaluation of $v_p\left(t_i, r_j, b^{(n-1)}(t_i, r_j); T, b^{(n-1)}\right)$ 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 parametrisation $c$ 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

$$(4.7) \qquad T = 1, K = 100, \sigma = 0.4, \kappa = 0.3, \theta = 0.05, \beta = 0.01, \rho = 0.5,$$

and the convergence criterion with $\varepsilon = 0.01$. The magnitude of $\kappa, \theta$ and $\beta$ is based on empirical findings reported in the literature, c.f. [24, Chapter 31] and [19]. Although main currencies have recently enjoyed much lower interest rates, our choice of $\theta$ 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 (4.7). 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 behaviour 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 [39]. 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 (a), the value decreases to the value of the immediate exercise as the option is purchased deep in the money.
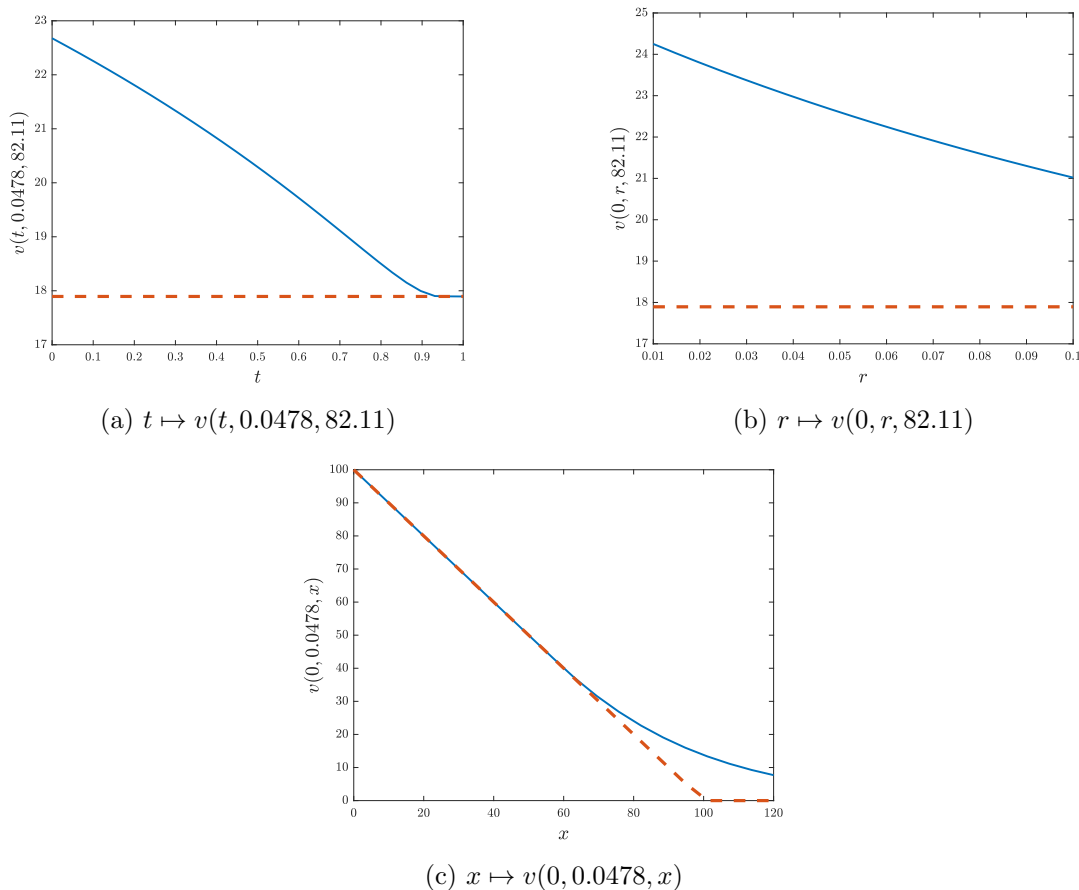
(a) $t \mapsto v(t, 0.0478, 82.11)$



(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 $(K - x)^+$.

*Effects of the interest rate.* The option price is significantly affected by the initial interest rate (Panel (b)) because the maturity of the option is long (1 year). The effect depends on the mean-reversion coefficient $\kappa$ 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 (4.2) 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 $\kappa$ 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 behaviour of the stopping boundary and the of value function is akin to that observed by a model with a constant interest rate [10, 39].

*Effects of the correlation coefficient.* The sensitivity of the stopping boundary with respect to the correlation coefficient $\rho$ 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
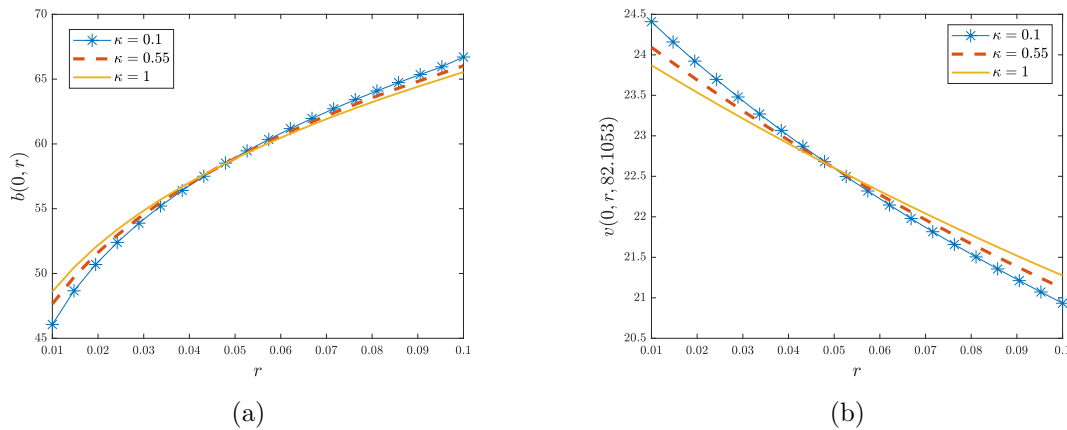
FIGURE 3. The $r$-sections of the stopping boundary (left panel) and the value function (right panel) for the mean-reversion parameter $\kappa \in \{0.1, 0.55, 1\}$.
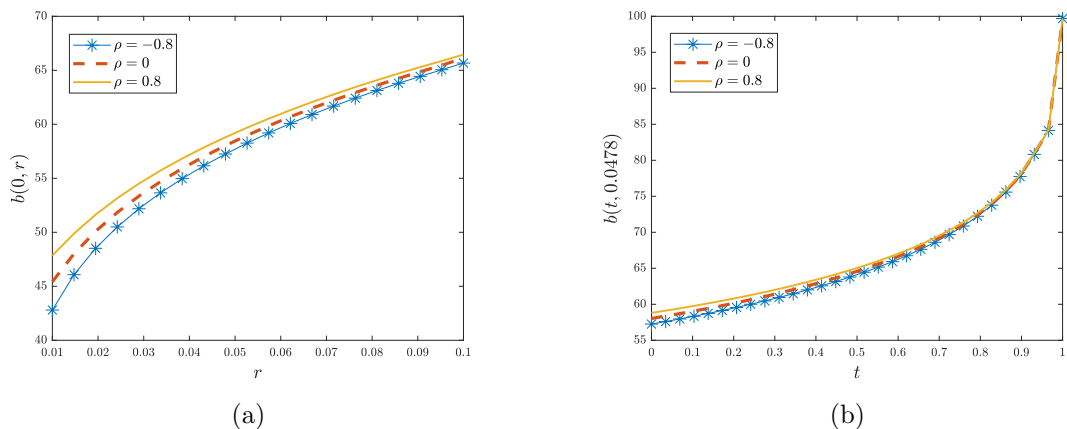


FIGURE 4. The $r$ and $t$-sections of the stopping boundary for the correlation coefficient $\rho \in \{-0.8, 0, 0.8\}$.

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 $\beta$ 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 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 $\sigma$, see Figure 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 $\sigma$, 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
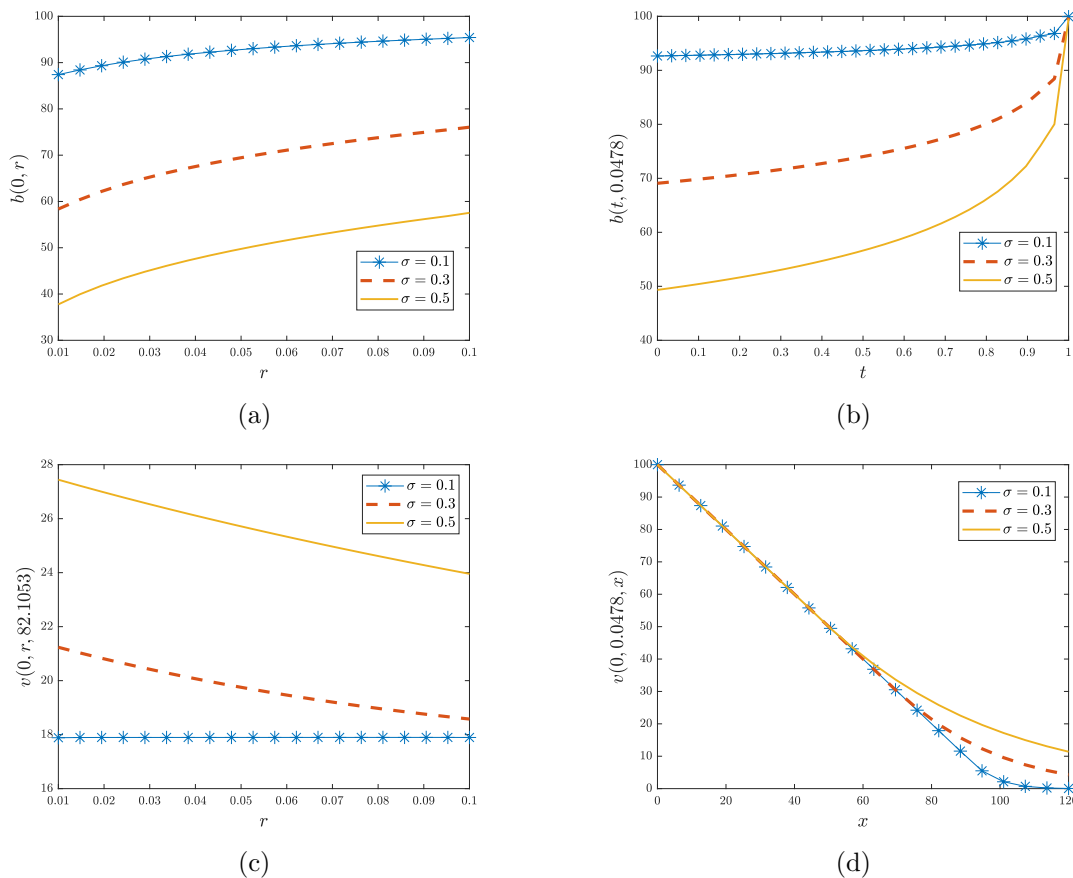
FIGURE 5. Effect of the volatility of the stock price $\sigma$. Panels (a) and (b) display the $r$ and $t$-sections of the stopping boundary $b(t, r)$ and Panels (c) and (d) show the $r$ and $x$-sections of the value function $v$ for $\sigma \in \{0.1, 0.3, 0.5\}$.

expected, in line with the findings for the American Put option in the Black-Scholes model with constant interest rate [10, 39].

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 (2.3) 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 $\varepsilon > 0$

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

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

where we took the discounting inside the positive part and used (2.5).

Finally, monotonicity in $(iii)$ is a simple consequence of monotonicity of $(2.5)$ 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 $v(t, r, x_\lambda) \le \lambda v(t, r, x) + (1 - \lambda)v(t, r, y)$. $\qquad\square$

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

$$(5.1) \qquad |v(t_1, r_1, x_1) - v(t_2, r_2, x_2)| \le L_\mathcal{K}\big(|t_1 - t_2| + |r_1 - r_2| + |x_1 - x_2|\big)$$

*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 [26, 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 $(2.5)$ and the Lipschitz property of the put payoff, we get

$$0 \le v(t, r, x_1) - v(t, r, x_2) \le \mathsf{E}\left[e^{-\int_0^{\tau_1} r_s^r ds}\Big((K - X^{r,x_1})^+ - (K - X^{r,x_2})^+\Big)\right]$$

$$\le \mathsf{E}\left[e^{\sigma B_{\tau_1} - \frac{\sigma^2}{2}\tau_1}\right](x_2 - x_1) = (x_2 - x_1),$$

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

$$(5.2) \qquad \begin{aligned} 0 &\le v(t, r_1, x) - v(t, r_2, x) \le K\mathsf{E}\left[e^{-\int_0^{\tau_1} r_s^1 ds} - e^{-\int_0^{\tau_1} r_s^2 ds}\right] = K\mathsf{E}\left[e^{-\int_0^{\tau_1} r_s^1 ds}\Big(1 - e^{-\int_0^{\tau_1}(r_s^2 - r_s^1)ds}\Big)\right] \\ &\le K\mathsf{E}\left[e^{-\int_0^{\tau_1} r_s^1 ds}\int_0^{\tau_1}(r_s^2 - r_s^1)ds\right]. \end{aligned}$$

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

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

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

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

$$(5.3) \qquad \begin{aligned} \mathsf{E}\left[e^{-\int_0^{\tau_1} r_s^1 ds}\int_0^{\tau_1}(r_s^2 - r_s^1)ds\right] &\le \left(\mathsf{E}\left[e^{-2\int_0^{\tau_1} r_s^1 ds}\right]\right)^{\frac{1}{2}}\left(\mathsf{E}\left[\Big(\int_0^{T-t}(r_s^2 - r_s^1)ds\Big)^2\right]\right)^{\frac{1}{2}} \\ &\le C_1^{1/2}\left((T - t)\int_0^{T-t} \mathsf{E}\left[(r_s^2 - r_s^1)^2\right] ds\right)^{\frac{1}{2}}, \end{aligned}$$

where $C_1 > 0$ is the constant from $(2.3)$ which depends on $\mathcal{K}$. To conclude it is sufficient to use moment estimates for SDEs [34, Thm. 2.5.9] which guarantee that

$$(5.4) \qquad \mathsf{E}\left[\sup_{0 \le s \le T}(r_s^2 - r_s^1)^2\right] \le c'(r_2 - r_1)^2$$

for some $c' > 0$ only depending on $T$ and the coefficients in (2.2).

*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., [5, Ch. 1, Prop. 8.6])

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

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 (2.6) as

$$(5.5) \qquad v(t, r, x) = \sup_{0 \le \theta \le 1} \mathsf{E}_{r,x}\left[\exp\left\{-(T-t)\int_0^\theta r_u^{T-t}du\right\}\left(K - X_\theta^{T-t}\right)^+\right],$$

where for any $(\mathcal{F}_s)_{s \ge 0}$-stopping time $\tau$ 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

$$(5.6) \qquad dX_u^{T-t} = (T-t)r_u^{T-t}X_u^{T-t}du + \sqrt{T-t}\sigma X_u^{T-t}dB_u, \qquad X_0^{T-t} = x,$$
$$(5.7) \qquad dr_u^{T-t} = (T-t)\alpha(r_u^{T-t})du + \sqrt{T-t}\beta(r_u^{T-t})dW_u, \qquad r_0^{T-t} = r,$$

and take stopping times $\theta \in [0, 1]$ in (5.5) 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 (5.5).

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 $\theta_1$ is also admissible for the problem in (5.5) and the underlying dynamics (5.6)–(5.7) with $t = t_2$, because it is an $(\mathcal{F}_s)_{s \ge 0}$-stopping time in $[0, 1]$. Indeed the advantage of (5.5) with (5.6)–(5.7) 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

$$(5.8) \qquad \begin{aligned} 0 \ge v(t_2, r, x) - v(t_1, r, x) &\ge -\mathsf{E}_r\left[\left|\left(Ke^{-R_{\theta_1}^2} - x\hat{X}_{\theta_1}^{T-t_2}\right)^+ - \left(Ke^{-R_{\theta_1}^1} - x\hat{X}_{\theta_1}^{T-t_1}\right)^+\right|\right] \\ &\ge -K\mathsf{E}_r\left[\left|e^{-R_{\theta_1}^2} - e^{-R_{\theta_1}^1}\right|\right] - x\mathsf{E}\left[|\hat{X}_{\theta_1}^{T-t_1} - \hat{X}_{\theta_1}^{T-t_2}|\right]. \end{aligned}$$

Let us consider the second term on the right hand side of (5.8). By the fundamental theorem of calculus and the explicit formula for $\hat{X}^{T-t}$

$$\mathsf{E}[|\hat{X}_{\theta_1}^{T-t_1} - \hat{X}_{\theta_1}^{T-t_2}|] = \mathsf{E}\left[\left|\int_{t_1}^{t_2} \hat{X}_{\theta_1}^{T-t}\left(\frac{\sigma^2}{2}\theta_1 - \frac{1}{2\sqrt{T-t}}\sigma B_{\theta_1}\right)dt\right|\right]$$
$$(5.9) \qquad\qquad\qquad \le \int_{t_1}^{t_2} \mathsf{E}\left[\left|\hat{X}_{\theta_1}^{T-t}\left(\frac{\sigma^2}{2}\theta_1 - \frac{1}{2\sqrt{T-t}}\sigma B_{\theta_1}\right)\right|\right]dt.$$

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

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

where we applied Hölder inequality and used that $\theta_1 \leq 1$. Inserting the above estimate into (5.9) gives

$$\text{(5.10)} \qquad \mathsf{E}\big[\big|\hat{X}_{\theta_1}^{T-t_1} - \hat{X}_{\theta_1}^{T-t_2}\big|\big] \leq c_1(t_2 - t_1).$$

Next we address the first term on the right hand side of (5.8). This is performed separately in cases (i) and (ii) of Assumption 2.1. We start by considering case (ii), i.e., $\alpha$ and $\beta$ in (5.7) are Lipschitz continuous. Fundamental theorem of calculus and Hölder inequality give

$$
\begin{aligned}
&\mathsf{E}_r\left[\left|e^{-R_{\theta_1}^1} - e^{-R_{\theta_1}^2}\right|\right] \\
&\leq \mathsf{E}_r\left[\max_{i=1,2}\left\{e^{-(T-t_i)\int_0^{\theta_1} r_u^i du}\right\}\left|(T-t_1)\int_0^{\theta_1} r_u^1 du - (T-t_2)\int_0^{\theta_1} r_u^2 du\right|\right] \\
&\leq \mathsf{E}_r\left[\max_{i=1,2}\left\{e^{-(T-t_i)\int_0^{\theta_1} r_u^i du}\right\}\left((t_2-t_1)\left|\int_0^{\theta_1} r_u^1 du\right| + (T-t_2)\left|\int_0^{\theta_1}(r_u^2 - r_u^1)du\right|\right)\right] \\
&\leq 2c_2\left[(t_2-t_1)\left(\mathsf{E}_r\big[\sup_{0\leq t\leq 1}(r_t^1)^2\big]\right)^{\frac{1}{2}} + (T-t_2)\left(\mathsf{E}_r\left[\int_0^1 (r_u^2 - r_u^1)^2 du\right]\right)^{\frac{1}{2}}\right],
\end{aligned}
$$

(5.11)

where, using (2.3),

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

Thanks to (2.4), $c_3 := \sup_{(t,r,x)\in\mathcal{K}}\left(\mathsf{E}_r\big[\sup_{0\leq s\leq 1}(r_s^{T-t})^2\big]\right)^{\frac{1}{2}} < \infty$, so it remains to estimate the last term of (5.11). By [34, Thm. 2.5.9] there is a constant $c_4$ depending only on $\mathcal{K}$ and the Lipschitz constant for $\alpha$ and $\beta$ in (5.7) such that

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

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 (2.4) and the linear growth of $\alpha$ and $\beta$

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

Inserting the above estimates into (5.11) we conclude that there is a constant $c_6$ such that for any $(t_1,r,x),(t_2,r,x)\in\mathcal{K}$

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

This and (5.10) feed into (5.8) so that

$$\text{(5.12)} \qquad 0 \geq v(t_2,r,x) - v(t_1,r,x) \geq -c|t_2 - t_1|$$

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

Finally, we must estimate the first term on the right hand side of (5.8) under the assumption that $(r_t)_{t\geq 0}$ follows the CIR dynamics (Assumption 2.1(i)). Let $\hat{r}_u^i := r_u^i/(T-t_i)$ for $u \in [0,1]$ and $i = 1,2$. The dynamics for $\hat{r}^i$ reads

$$\text{(5.13)} \qquad d\hat{r}_u^i = \kappa\big(\alpha - (T-t_i)\hat{r}_u^i\big)du + \beta\sqrt{\hat{r}_u^i}\,dW_u, \quad u \in [0,1].$$

Since $\kappa(\alpha - (T - t_1)\hat{r}) < \kappa(\alpha - (T - t_2)\hat{r})$ for $\hat{r} \geq 0$, and $\hat{r}_0^1 = r/(T - t_1) \leq r/(T - t_2) = \hat{r}_0^2$, comparison results for SDEs [29, Prop. 5.2.18] imply

$$(5.14) \qquad \hat{r}_u^1 \leq \hat{r}_u^2 \quad \text{for all } u \in [0, 1], \text{ P-a.s.}$$

Using the integral version of (5.13) and the martingale property of the stochastic integral, we obtain

$$\mathsf{E}_r\left[\hat{r}_u^2 - \hat{r}_u^1\right] = r\left(\frac{1}{T - t_2} - \frac{1}{T - t_1}\right) + \mathsf{E}_r\left[\int_0^u \left((T - t_1)\hat{r}_s^1 - (T - t_2)\hat{r}_s^2\right) ds\right]$$

$$\leq r\frac{t_2 - t_1}{(T - t_1)(T - t_2)} + (t_2 - t_1)\int_0^1 \mathsf{E}_r\left[\hat{r}_s^1\right] ds + (T - t_2)\int_0^u \mathsf{E}_r\left[\hat{r}_s^1 - \hat{r}_s^2\right] ds.$$

Due to (5.14), the last term is non-positive, so

$$(5.15) \qquad 0 \leq \mathsf{E}_r\left[\hat{r}_u^2 - \hat{r}_u^1\right] \leq (t_2 - t_1)\left(\frac{r}{(T - t_1)(T - t_2)} + q_1\right) \quad \text{for all } u \in [0, 1]$$

where

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

We use the inequalities (5.14)–(5.15) and the property that $\hat{r}_u^i \geq 0$, for $i = 1, 2$, to obtain the following estimates

$$\mathsf{E}_r\left[\left|e^{-R_{\theta_1}^1} - e^{-R_{\theta_1}^2}\right|\right] = \mathsf{E}_r\left[\left|e^{-(T-t_1)^2\int_0^{\theta_1}\hat{r}_u^1 du} - e^{-(T-t_2)^2\int_0^{\theta_1}\hat{r}_u^2 du}\right|\right]$$

$$\leq \mathsf{E}_r\left[\left|e^{-(T-t_1)^2\int_0^{\theta_1}\hat{r}_u^1 du} - e^{-(T-t_2)^2\int_0^{\theta_1}\hat{r}_u^1 du}\right|\right]$$

$$(5.16) \qquad \qquad + \mathsf{E}_r\left[\left|e^{-(T-t_2)^2\int_0^{\theta_1}\hat{r}_u^1 du} - e^{-(T-t_2)^2\int_0^{\theta_1}\hat{r}_u^2 du}\right|\right]$$

$$\leq q_1\left((T - t_1)^2 - (T - t_2)^2\right) + (T - t_2)^2\int_0^1 \mathsf{E}_r\left[\hat{r}_u^2 - \hat{r}_u^1\right] du$$

$$\leq (t_2 - t_1)\left(2Tq_1 + r\frac{T - t_2}{T - t_1} + q_1(T - t_2)^2\right) \leq c_7(t_2 - t_1),$$

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 (5.12). $\qquad \square$

## 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 non-increasing 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 (3.1) with

$$(6.1) \qquad c(t, x) := \inf\{r \in \mathcal{I} : v(t, r, x) = (K - x)^+\},$$

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 $B_t = \rho W_t + \sqrt{1 - \rho^2}B_t^0$ for some other Brownian motion $B^0$

independent of $W$. Letting $(\mathcal{F}_t^W)_{t\geq 0}$ be the filtration generated by $W$, using the explicit form of the dynamics of $X$ we have

$$\mathsf{P}_{r,x}(X_\varepsilon < K) = \mathsf{E}_{r,x}\Big[\mathsf{P}_{r,x}(X_\varepsilon < K|\mathcal{F}_\varepsilon^W)\Big]$$

$$(6.2) \qquad = \mathsf{E}_{r,x}\Big[\mathsf{P}_r\Big(\exp\big(\sigma\sqrt{1-\rho^2}B_\varepsilon^0\big) < (K/x)\exp\Big(-\sigma\rho W_\varepsilon - \int_0^\varepsilon r_t dt + \tfrac{\sigma^2}{2}\varepsilon\Big)\Big|\mathcal{F}_\varepsilon^W\Big)\Big]$$

$$= \mathsf{E}_{r,x}\Big[\Psi_x\Big(\sigma\rho W_\varepsilon + \int_0^\varepsilon r_t dt - \tfrac{\sigma^2}{2}\varepsilon\Big)\Big],$$

where

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

and the final equality above holds by the independence of $B_\varepsilon^0$ from $\mathcal{F}_\varepsilon^W$ and the fact that $(W_\varepsilon, \int_0^\varepsilon r_t dt)$ is $\mathcal{F}_\varepsilon^W$-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 $(t_1,r,x) \in \mathcal{D} \implies (t_2,r,x) \in \mathcal{D}$ for any $t_2 > t_1$, hence $c(t,x)$ is non-increasing in $t$.

Fix $0 \leq x_1 < x_2 < K$ and let $\tau_1 := \tau_*(t,r,x_1)$ be optimal for $v(t,r,x_1)$. Then, using that $X^{r,x_1} \leq X^{r,x_2}$ and recalling (2.5), we obtain

$$v(t,r,x_2) - v(t,r,x_1) \geq \mathsf{E}\Big[e^{-\int_0^{\tau_1} r_s ds}\Big(\big(K - X_{\tau_1}^{r,x_2}\big)^+ - \big(K - X_{\tau_1}^{r,x_1}\big)^+\Big)\Big]$$

$$\geq \mathsf{E}\Big[e^{-\int_0^{\tau_1} r_s ds}\big(X_{\tau_1}^{r,x_1} - X_{\tau_1}^{r,x_2}\big)\Big]$$

$$= x_1 - x_2 = (K - x_2)^+ - (K - x_1)^+.$$

Therefore, if $(t,r,x_1) \in \mathcal{C}$ then $(t,r,x_2) \in \mathcal{C}$, which implies that $c(t,x)$ is non-decreasing 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, $\tau_1$ is sub-optimal and

$$(6.3) \qquad K - x_0 = v(t_0, r_0, x_0) \geq \mathsf{E}_{r_0,x_0}\Big[e^{-\int_0^{\tau_1} r_s ds}(K - X_{\tau_1})^+\Big] \geq K\mathsf{E}_{r_0,x_0}\Big[e^{-\int_0^{\tau_1} r_s ds}\Big] - x_0,$$

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

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

which, in conjunction with (6.3), 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 $\mathsf{P}_{r_0,x_0}\big((t_0+s, r_s, X_s) \in \mathcal{C} \text{ for } s \le \tau_2\big) = 1$, so $\tau_2 \le \tau_*(t_0, r_0, x_0)$ $\mathsf{P}_{r_0,x_0}$-a.s. By the martingale property of the value function we obtain

$$
\begin{aligned}
(6.4) \qquad K - x_0 < v(t_0, r_0, x_0) &= \mathsf{E}_{r_0,x_0}\left[ e^{-\int_0^{\tau_2} r_s ds} v\left(t_0 + \tau_2, r_{\tau_3}, X_{\tau_2}\right) \right] \\
&\le K \mathsf{E}_{r_0,x_0}\left[ e^{-r_1 \tau_2} \right] = K \mathsf{E}_{r_0}\left[ e^{-r_1 \tau_2} \right].
\end{aligned}
$$

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

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

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

We immediately see that $\partial \mathcal{C}$ enjoys the so-called *cone property* [29, 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.

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

Both $\sigma_\mathcal{D}$ and $\mathring{\sigma}_\mathcal{D}$ are stopping times with respect to the filtration $(\mathcal{F}_t)_{t \ge 0}$. 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*

$$(6.6) \qquad \mathsf{P}_{t_0,r_0,x_0}(\sigma_\mathcal{D} > 0) = \mathsf{P}_{t_0,r_0,x_0}(\mathring{\sigma}_\mathcal{D} > 0) = 0.$$

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 multi-dimensional Brownian motion, see e.g. [29, 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 $(r_n, x_n)_{n \ge 1}$ be a sequence converging to $(r, x) \in \mathcal{I} \times \mathbb{R}_+$ as $n \to \infty$. Then*

$$(7.1) \qquad \lim_{n \to +\infty} \sup_{0 \le t \le T} |r_t^{r_n} - r_t^r| = 0, \qquad \mathsf{P}\text{-a.s.}$$

$$(7.2) \qquad \lim_{n \to +\infty} \sup_{0 \le t \le T} |X_t^{r_n,x_n} - X_t^{r,x}| = 0, \qquad \mathsf{P}\text{-a.s.}$$

*Proof of Lemma 7.1.* Assume first that $(r_n)_{n \ge 1}$ 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 (7.1) holds.

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

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

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 (7.1). The verification of (7.2) is easy using the representation formula (2.5) for $X$ and (7.1). $\qquad\square$

**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 summarise 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 (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.   $\square$

Next we provide gradient estimates based on probabilistic arguments.

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

$$(7.3) \qquad v_x(t, r, x) = -\mathsf{E}_{t,r,x}\left[1_{\{X_{\tau_*} \leq K\}} e^{\sigma B_{\tau_*} - \frac{\sigma^2}{2}\tau_*}\right],$$

$$(7.4) \qquad 0 \geq v_t(t, r, x) \geq -L\,\mathsf{E}_{t,r,x}\left[e^{-\int_0^{\tau_{\mathcal{K}}} r_s^r ds} 1_{\{\tau_{\mathcal{K}} \leq \tau_*\}}\right],$$

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

*Proof of Proposition 7.3.* Fix $(t, r, x) \in (\mathrm{int}(\mathcal{K}) \setminus \partial\mathcal{C})$. Recall that $\mathcal{D} \subset [0, T] \times \mathcal{I} \times [0, K]$. If $(t, r, x) \in \mathrm{int}(\mathcal{D})$ then (7.3) 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* (7.3)) For all sufficiently small $\varepsilon > 0$ we have $(t, r, x + \varepsilon) \in \mathcal{C}$. From now on, consider such $\varepsilon$. To simplify notation let $\sigma_{\mathcal{D}} := \sigma_{\mathcal{D}}(t, r, x)$. Using that $\sigma_{\mathcal{D}}$ is admissible and sub-optimal for $v(t, r, x + \varepsilon)$ we get

$$v(t, r, x + \varepsilon) - v(t, r, x)$$
$$\geq \mathsf{E}\left[e^{-\int_0^{\sigma_{\mathcal{D}}} r_s^r ds}\left(\left(K - (x + \varepsilon)X_{\sigma_{\mathcal{D}}}^{r,1}\right)^+ - \left(K - xX_{\sigma_{\mathcal{D}}}^{r,1}\right)^+\right)\right]$$
$$\geq \mathsf{E}\left[e^{-\int_0^{\sigma_{\mathcal{D}}} r_s^r ds} 1_{\{X_{\sigma_{\mathcal{D}}}^{r,x} \leq K\}}\left(xX_{\sigma_{\mathcal{D}}}^{r,1} - (x + \varepsilon)X_{\sigma_{\mathcal{D}}}^{r,1}\right)\right] = -\varepsilon\mathsf{E}\left[1_{\{X_{\sigma_{\mathcal{D}}}^{r,x} \leq K\}} e^{\sigma B_{\sigma_{\mathcal{D}}} - \frac{\sigma^2}{2}\sigma_{\mathcal{D}}}\right].$$

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

$$(7.5) \qquad v_x(t, r, x) = \lim_{\varepsilon \to 0} \frac{1}{\varepsilon}\left(v(t, r, x + \varepsilon) - v(t, r, x)\right) \geq -\mathsf{E}\left[1_{\{X_{\sigma_{\mathcal{D}}}^{r,x} \leq K\}} e^{\sigma B_{\sigma_{\mathcal{D}}} - \frac{\sigma^2}{2}\sigma_{\mathcal{D}}}\right].$$

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

$$v(t, r, x) - v(t, r, x - \varepsilon) \leq \mathsf{E}\left[e^{-\int_0^{\sigma_{\mathcal{D}}} r_s^r ds}\left(\left(K - xX_{\sigma_{\mathcal{D}}}^{r,1}\right)^+ - \left(K - (x - \varepsilon)X_{\sigma_{\mathcal{D}}}^{r,1}\right)^+\right)\right]$$
$$\leq -\varepsilon\,\mathsf{E}\left[1_{\{X_{\sigma_{\mathcal{D}}}^{r,x-\varepsilon} \leq K\}} e^{\sigma B_{\sigma_{\mathcal{D}}} - \frac{\sigma^2}{2}\sigma_{\mathcal{D}}}\right] \leq -\varepsilon\,\mathsf{E}\left[1_{\{X_{\sigma_{\mathcal{D}}}^{r,x} \leq K\}} e^{\sigma B_{\sigma_{\mathcal{D}}} - \frac{\sigma^2}{2}\sigma_{\mathcal{D}}}\right],$$

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

$$(7.6) \qquad v_x(t, r, x) = \lim_{\varepsilon \to 0} \frac{1}{\varepsilon}\left(v(t, r, x) - v(t, r, x - \varepsilon)\right) \leq -\mathsf{E}\left[1_{\{X_{\sigma_{\mathcal{D}}}^{r,x} \leq K\}} e^{\sigma B_{\sigma_{\mathcal{D}}} - \frac{\sigma^2}{2}\sigma_{\mathcal{D}}}\right].$$

Now (7.5) and (7.6) imply (7.3).

(*Proof of* (7.4)) The upper bound $v_t(t, r, x) \leq 0$ follows from the monotonicity of $v$ in $t$ (Proposition 3.1). For all sufficiently small $\varepsilon > 0$ we have $(t + \varepsilon, r, x) \in \mathcal{K} \cap \mathcal{C}$ and $\tau_{\mathcal{K}} := \tau_{\mathcal{K}}(t, r, x) \leq T - t - \varepsilon$.

From now on, consider such $\varepsilon$. Denote $\sigma_{\mathcal{D}} := \sigma_{\mathcal{D}}(t, r, x)$. Thanks to the choice of $\varepsilon$, 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 (2.11)–(2.12)) we get

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

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 $\{\tau_{\mathcal{K}} \geq \sigma_{\mathcal{D}}\}$ since $t \mapsto b(t, r)$ is non-decreasing (Proposition 3.4). Let $\mathcal{K}^\delta = \{(t + s, r, x) : (t, r, x) \in \mathcal{K} \text{ and } s \in [0, \delta]\}$. 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 (7.7) 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 $\varepsilon$ and taking the limit $\varepsilon \to 0$ completes the proof of (7.4). $\qquad\square$

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

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

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

for any sequence $(t_n, r_n, x_n)$ in $\mathcal{C}$ converging to $(t_0, r_0, x_0) \in \partial\mathcal{C}$ 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 $\mathsf{P}_{t_n, r_n, x_n}(X_{\tau_n} = K, \tau_n = T - t_n) \leq \mathsf{P}_{t_n, r_n, x_n}(X_{T-t_n} = K) = 0$ (the final equality can be shown by arguments as in (6.2)). 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 $\lim_{n \to \infty} \tau_n = 0$ $\mathsf{P}$-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 (7.9).

*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 $n \geq 1$. Let

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

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 $\tau_{\mathcal{K}}^n \to \tau_{\mathcal{K}}^0$, $\mathsf{P}$-a.s. Fix $\varepsilon \in (0, 1)$. Since $\mathsf{P}(\tau_{\mathcal{K}}^0 > 0) = 1$, there exists $\delta > 0$ such

that $\mathsf{P}(\tau_{\mathcal{K}}^0 > \delta) \geq 1 - \varepsilon$. From inequality (7.4), we get

$$
0 \geq v_t(t_n, r_n, x_n) \geq -L\, \mathsf{E}\left[e^{-\int_0^{\tau_{\mathcal{K}}^n} r_s^{r_n}\, ds} 1_{\{\tau_{\mathcal{K}}^n \leq \tau_n\}}\right]
$$

$$
(7.10) \qquad = -L\, \mathsf{E}\left[e^{-\int_0^{\tau_{\mathcal{K}}^n} r_s^{r_n}\, ds}\left(1_{\{\tau_{\mathcal{K}}^n \leq \tau_n\} \cap \{\tau_{\mathcal{K}}^n \geq \delta\}} + 1_{\{\tau_{\mathcal{K}}^n \leq \tau_n\} \cap \{\tau_{\mathcal{K}}^n < \delta\}}\right)\right]
$$

$$
\geq -L\, \mathsf{E}\left[e^{-\int_0^{\tau_{\mathcal{K}}^n} r_s^{r_n}\, ds}\left(1_{\{\tau_n \geq \delta\}} + 1_{\{\tau_{\mathcal{K}}^n < \delta\}}\right)\right].
$$

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

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

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:

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

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 (7.8) 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

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

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

We know that $v_r \leq 0$ 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 $n \to \infty$. For $(t, r, x) \in \mathcal{C} \cap \mathcal{K}^T$ we define

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

By the monotonicity of $r \mapsto r_s^r$ and the explicit expression (2.5) for $X^{r,x}$ we have, for all $(r, x) \in \mathcal{K}$,

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

from which it is not hard to verify that for all $(t, r, x) \in \mathcal{C} \cap \mathcal{K}^T$

$$
\tau_{\mathcal{K}}(t, r, x) \geq \widehat{\tau}_{\mathcal{K}} := \tau_{\mathcal{K}}(t_b, r_a, x_a) \wedge \tau_{\mathcal{K}}(t_b, r_a, x_b) \wedge \tau_{\mathcal{K}}(t_b, r_b, x_a) \wedge \tau_{\mathcal{K}}(t_b, r_b, x_b), \quad \text{P-a.s.}
$$

Take $(t, r, x) \in \mathcal{C} \cap \mathcal{K}^T$. There is $\overline{\varepsilon} > 0$ such that $(t, r + \varepsilon, x) \in \mathcal{C} \cap \mathcal{K}^T$ for all $\varepsilon \in (0, \overline{\varepsilon}]$. Denote by $\tau_*$ the optimal stopping time for $(t, r, x)$. For any $\varepsilon \in (0, \overline{\varepsilon}]$, we apply the (super)martingale properties of the value function (2.11)-(2.12) with the stopping time $\tau_* \wedge \widehat{\tau}_{\mathcal{K}}$:

$$
0 \geq v(t, r + \varepsilon, x) - v(t, r, x)
$$

$$
\geq \mathsf{E}\left[e^{-\int_0^{\tau_* \wedge \widehat{\tau}_{\mathcal{K}}} r_s^{r+\varepsilon}\, ds} v\left(t + (\tau_* \wedge \widehat{\tau}_{\mathcal{K}}), r_{\tau_* \wedge \widehat{\tau}_{\mathcal{K}}}^{r+\varepsilon}, X_{\tau_* \wedge \widehat{\tau}_{\mathcal{K}}}^{r+\varepsilon, x}\right)\right.
$$

$$
(7.12) \qquad \left. - e^{-\int_0^{\tau_* \wedge \widehat{\tau}_{\mathcal{K}}} r_s^{r}\, ds} v\left(t + (\tau_* \wedge \widehat{\tau}_{\mathcal{K}}), r_{\tau_* \wedge \widehat{\tau}_{\mathcal{K}}}^{r}, X_{\tau_* \wedge \widehat{\tau}_{\mathcal{K}}}^{r, x}\right)\right]
$$

$$
\geq \mathsf{E}\left[1_{\{\widehat{\tau}_{\mathcal{K}} \leq \tau_*\}}\left(e^{-\int_0^{\widehat{\tau}_{\mathcal{K}}} r_s^{r+\varepsilon}\, ds} v(t + \widehat{\tau}_{\mathcal{K}}, r_{\widehat{\tau}_{\mathcal{K}}}^{r+\varepsilon}, X_{\widehat{\tau}_{\mathcal{K}}}^{r+\varepsilon, x}) - e^{-\int_0^{\widehat{\tau}_{\mathcal{K}}} r_s^{r}\, ds} v(t + \widehat{\tau}_{\mathcal{K}}, r_{\widehat{\tau}_{\mathcal{K}}}^{r}, X_{\widehat{\tau}_{\mathcal{K}}}^{r, x})\right)\right]
$$

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

$$
=: E_1 + E_2,
$$

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

$$
\begin{aligned}
E_1 &= \mathsf{E}\left[1_{\{\widehat{\tau}_{\mathcal{K}} \leq \tau_*\}} e^{-\int_0^{\widehat{\tau}_{\mathcal{K}}} r^{r+\varepsilon}_s ds} \left(v(t + \widehat{\tau}_{\mathcal{K}}, r^{r+\varepsilon}_{\widehat{\tau}_{\mathcal{K}}}, X^{r+\varepsilon,x}_{\widehat{\tau}_{\mathcal{K}}}) - v(t + \widehat{\tau}_{\mathcal{K}}, r^r_{\widehat{\tau}_{\mathcal{K}}}, X^{r,x}_{\widehat{\tau}_{\mathcal{K}}})\right)\right] \\
&\quad - \mathsf{E}\left[1_{\{\widehat{\tau}_{\mathcal{K}} \leq \tau_*\}} \left(e^{-\int_0^{\widehat{\tau}_{\mathcal{K}}} r^r_s ds} - e^{-\int_0^{\widehat{\tau}_{\mathcal{K}}} r^{r+\varepsilon}_s ds}\right) v(t + \widehat{\tau}_{\mathcal{K}}, r^r_{\widehat{\tau}_{\mathcal{K}}}, X^{r,x}_{\widehat{\tau}_{\mathcal{K}}})\right] \\
&\geq -L\, \mathsf{E}\left[1_{\{\widehat{\tau}_{\mathcal{K}} \leq \tau_*\}} e^{-\int_0^{\widehat{\tau}_{\mathcal{K}}} r^{r+\varepsilon}_s ds} \left(|r^{r+\varepsilon}_{\widehat{\tau}_{\mathcal{K}}} - r^r_{\widehat{\tau}_{\mathcal{K}}}| + |X^{r+\varepsilon,x}_{\widehat{\tau}_{\mathcal{K}}} - X^{r,x}_{\widehat{\tau}_{\mathcal{K}}}|\right)\right] \\
&\quad - K\mathsf{E}\left[1_{\{\widehat{\tau}_{\mathcal{K}} \leq \tau_*\}} \left(e^{-\int_0^{\widehat{\tau}_{\mathcal{K}}} r^r_s ds} - e^{-\int_0^{\widehat{\tau}_{\mathcal{K}}} r^{r+\varepsilon}_s ds}\right)\right],
\end{aligned}
\tag{7.13}
$$

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 $(r^r_s)$ with respect to the parameter $r$ in the sense of [34, Thm. 2.8.6]. Apart from other assumptions, this requires that the coefficients are 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 $\{s \leq \widehat{\tau}_{\mathcal{K}}\}$, i.e., 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 [34, Thm. 2.8.6], there is a measurable in $(s, \omega)$ process $(y^r_s(\omega))_{s \geq 0}$, depending on $r$, such that for any $q \geq 1$

$$
\lim_{\varepsilon \downarrow 0} \left\| \sup_{s \in [0,T]} \left|\frac{r^{r+\varepsilon}_s - r^r_s}{\varepsilon} - y^r_s\right| \right\|_q = 0 \quad \text{and} \quad \lim_{\varepsilon \downarrow 0} \left\|\frac{r^{r+\varepsilon}_{\cdot} - r^r_{\cdot}}{\varepsilon} - y^r_{\cdot}\right\|^*_q = 0,
\tag{7.14}
$$

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

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

$$
\begin{aligned}
&\frac{1}{\varepsilon} \mathsf{E}\left[1_{\{\widehat{\tau}_{\mathcal{K}} \leq \tau_*\}} e^{-\int_0^{\widehat{\tau}_{\mathcal{K}}} r^{r+\varepsilon}_s ds} |r^{r+\varepsilon}_{\widehat{\tau}_{\mathcal{K}}} - r^r_{\widehat{\tau}_{\mathcal{K}}}|\right] \\
&\leq \mathsf{E}\left[1_{\{\widehat{\tau}_{\mathcal{K}} \leq \tau_*\}} e^{-\int_0^{\widehat{\tau}_{\mathcal{K}}} r^{r+\varepsilon}_s ds} \left(\left|\frac{1}{\varepsilon}(r^{r+\varepsilon}_{\widehat{\tau}_{\mathcal{K}}} - r^r_{\widehat{\tau}_{\mathcal{K}}}) - y^r_{\widehat{\tau}_{\mathcal{K}}}\right| + |y^r_{\widehat{\tau}_{\mathcal{K}}}|\right)\right] \\
&\leq C_1^{1/p} \mathsf{P}(\widehat{\tau}_{\mathcal{K}} \leq \tau_*)^{\frac{1}{w}} \left(\left\|\frac{1}{\varepsilon}(r^{r+\varepsilon}_{\widehat{\tau}_{\mathcal{K}}} - r^r_{\widehat{\tau}_{\mathcal{K}}}) - y^r_{\widehat{\tau}_{\mathcal{K}}}\right\|_q + \|y^r_{\widehat{\tau}_{\mathcal{K}}}\|_q\right) \xrightarrow[\varepsilon \downarrow 0]{} C_1^{1/p} \mathsf{P}(\widehat{\tau}_{\mathcal{K}} \leq \tau_*)^{\frac{1}{w}} \|y^r_{\widehat{\tau}_{\mathcal{K}}}\|_q,
\end{aligned}
\tag{7.15}
$$

where we used the estimate (2.3) in the last inequality and (7.14) to obtain the convergence.

To bound the last term on the right hand side of (7.13), we observe that

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

We then apply Hölder inequality and the second limit in (7.14):

$$
\begin{aligned}
&\frac{1}{\varepsilon} \mathsf{E}\left[1_{\{\widehat{\tau}_{\mathcal{K}} \leq \tau_*\}} e^{-\int_0^{\widehat{\tau}_{\mathcal{K}}} r^r_s ds} \int_0^{\widehat{\tau}_{\mathcal{K}}} (r^{r+\varepsilon}_s - r^r_s) ds\right] \\
&\leq \mathsf{E}\left[1_{\{\widehat{\tau}_{\mathcal{K}} \leq \tau_*\}} e^{-\int_0^{\widehat{\tau}_{\mathcal{K}}} r^r_s ds} \int_0^{\widehat{\tau}_{\mathcal{K}}} \left|\frac{1}{\varepsilon}(r^{r+\varepsilon}_s - r^r_s) - y^r_s\right| + |y^r_s| ds\right] \\
&\leq C_1^{1/p} \mathsf{P}(\widehat{\tau}_{\mathcal{K}} \leq \tau_*)^{\frac{1}{w}} \left(\left\|\frac{1}{\varepsilon}(r^{r+\varepsilon}_{\cdot} - r^r_{\cdot}) - y^r_{\cdot}\right\|^*_q + \|y^r_{\cdot}\|^*_q\right) \xrightarrow[\varepsilon \downarrow 0]{} C_1^{1/p} \mathsf{P}(\widehat{\tau}_{\mathcal{K}} \leq \tau_*)^{\frac{1}{w}} \|y^r_{\cdot}\|^*_q.
\end{aligned}
\tag{7.16}
$$

By the explicit formula (2.5), we have $X_t^{r,x} = e^{\int_0^t r_s^r ds} \hat{X}_t^x$, where $\hat{X}_t^x := x e^{\sigma B_t - \frac{1}{2}\sigma^2 t}$, 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 (7.16) to obtain

(7.17)
$$\frac{1}{\varepsilon} \mathsf{E}\left[ 1_{\{\hat{\tau}_{\mathcal{K}} \le \tau_*\}} e^{-\int_0^{\hat{\tau}_{\mathcal{K}}} r_s^{r+\varepsilon} ds} |X_{\hat{\tau}_{\mathcal{K}}}^{r+\varepsilon,x} - X_{\hat{\tau}_{\mathcal{K}}}^{r,x}| \right]$$
$$\le \|\hat{X}_{\hat{\tau}_{\mathcal{K}}}^x\|_p \, \mathsf{P}(\hat{\tau}_{\mathcal{K}} \le \tau_*)^{\frac{1}{w}} \left( \left\| \frac{1}{\varepsilon}(r_\cdot^{r+\varepsilon} - r_\cdot^r) - y_\cdot^r \right\|_q^* + \|y_\cdot^r\|_q^* \right)$$
$$\xrightarrow[\varepsilon\downarrow 0]{} \|\hat{X}_{\hat{\tau}_{\mathcal{K}}}^x\|_p \, \mathsf{P}(\hat{\tau}_{\mathcal{K}} \le \tau_*)^{\frac{1}{w}} \, \|y_\cdot^r\|_q^*.$$

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

(7.18)
$$\frac{1}{\varepsilon} E_2 = \frac{1}{\varepsilon} \mathsf{E}\left[ 1_{\{\hat{\tau}_{\mathcal{K}} > \tau_*\}} \left( \left( Ke^{-\int_0^{\tau_*} r_s^{r+\varepsilon} ds} - \hat{X}_{\tau_*}^x \right)^+ - \left( Ke^{-\int_0^{\tau_*} r_s^r ds} - \hat{X}_{\tau_*}^x \right)^+ \right) \right]$$
$$\ge -\frac{1}{\varepsilon} K \mathsf{E}\left[ 1_{\{\hat{\tau}_{\mathcal{K}} > \tau_*\}} \left( e^{-\int_0^{\tau_*} r_s^r ds} - e^{-\int_0^{\tau_*} r_s^{r+\varepsilon} ds} \right) \right]$$
$$\ge -\frac{1}{\varepsilon} K \mathsf{E}\left[ 1_{\{\hat{\tau}_{\mathcal{K}} > \tau_*\}} e^{-\int_0^{\tau_*} r_s^r ds} \int_0^{\tau_*} (r_s^{r+\varepsilon} - r_s^r) ds \right]$$
$$\ge -K C_1^{1/p} \, \mathsf{P}(\hat{\tau}_{\mathcal{K}} > \tau_*)^{\frac{1}{w}} \left( \left\| \frac{1}{\varepsilon}(r_\cdot^{r+\varepsilon} - r_\cdot^r) - y_\cdot^r \right\|_q^* + \left( \mathsf{E}\left[ \int_0^{\tau_*} |y_s^r|^q ds \right] \right)^{1/q} \right)$$
$$\xrightarrow[\varepsilon\downarrow 0]{} -K C_1^{1/p} \mathsf{P}(\hat{\tau}_{\mathcal{K}} > \tau_*)^{\frac{1}{w}} \left( \mathsf{E}\left[ \int_0^{\tau_*} |y_s^r|^q ds \right] \right)^{1/q},$$

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

Combining (7.15)–(7.18) gives a lower bound for $v_r$ on $\mathcal{C} \cap \mathcal{K}^T$:

(7.19)
$$0 \ge v_r(t,r,x)$$
$$\ge -L \mathsf{P}(\hat{\tau}_{\mathcal{K}} \le \tau_*)^{\frac{1}{w}} \left( C_1^{1/p} \|y_{\hat{\tau}_{\mathcal{K}}}^r\|_q + \|\hat{X}_{\hat{\tau}_{\mathcal{K}}}^x\|_p \|y_\cdot^r\|_q^* \right) - K \mathsf{P}(\hat{\tau}_{\mathcal{K}} \le \tau_*)^{\frac{1}{w}} C_1^{1/p} \|y_\cdot^r\|_q^*$$
$$- K C_1^{1/p} \mathsf{P}(\hat{\tau}_{\mathcal{K}} > \tau_*)^{\frac{1}{w}} \left( \mathsf{E}\left[ \int_0^{\tau_*} |y_s^r|^q ds \right] \right)^{1/q}.$$

By [34, Thm. 2.8.8] and standard diffusion estimates [34, Cor. 2.5.10] the norms of $y^r$ and $\hat{X}^x$ above are bounded uniformly for $(t,r,x) \in \mathcal{K}^T \cap \mathcal{C}$ (recall that $\tau_* = \tau_*(t,r,x)$). Now take $(t,r,x) = (t_n, r_n, x_n)$ in (7.19). 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 (7.19) tend to zero as $n \to \infty$ due to $\mathsf{P}(\hat{\tau}_{\mathcal{K}} \le \tau_*(t_n, r_n, x_n)) \to 0$ 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 [34, Thm. 2.8.6]. This concludes the proof. $\square$

## 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 non-increasing and right-continuous at $t_0$, and $x \mapsto c(t_0, x)$ is non-decreasing 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 $x \ge K$. 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 $\overline{R}$ and satisfying

$$(8.1) \qquad \begin{aligned} (\mathcal{L} - r)u(r, x) &= -v_t(t_0, r, x), & (r, x) \in R, \\ u(r, x) &= v(t_0, r, x), & (r, x) \in \partial R. \end{aligned}$$

Thanks to [20, Theorem 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 (2.14) have Hölder continuous first derivatives, there is a unique classical solution $u(r, x)$ of the above PDE (which is of elliptic type) and $u \in C^3(R) \cap C(\overline{R})$ [20, Theorems 19 and 20, p. 87]. From (3.3), the function $(r, x) \mapsto v(t_0, r, x)$ satisfies (8.1), so, by uniqueness, $u = v$ on $\overline{R}$ and $u \in C^1(\overline{R})$ by Theorem 3.5.

We differentiate the PDE in (8.1) with respect to $r$ and obtain

$$(8.2) \quad \frac{1}{2}\sigma^2 x^2 u_{rxx}(r, x) = -\mathcal{L}_1 u_r(r, x) - \mathcal{L}_2 u_x(r, x) - x u_x(r, x) - v_{tr}(t_0, x, r) + u(x, r), \quad (r, x) \in R,$$

where

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

Let $\phi$ be a $C^\infty$ 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 (8.2) by $\frac{2}{\sigma^2 x^2}\phi(r)$ and integrate over $(r_1, r_2)$:

$$\begin{aligned} \int_{r_1}^{r_2} u_{rxx}(x, r)\phi(r)dr = &-\int_{r_1}^{r_2} \frac{2}{\sigma^2 x^2}\phi(r)\left[\mathcal{L}_1 u_r(r, x) + \mathcal{L}_2 u_x(r, x)\right]dr - \int_{r_1}^{r_2} \frac{2}{\sigma^2 x}\phi(r)u_x(r, x)dr \\ &-\int_{r_1}^{r_2} \frac{2}{\sigma^2 x^2}\phi(r)v_{tr}(t_0, r, x)dr + \int_{r_1}^{r_2} \frac{2}{\sigma^2 x^2}\phi(r)u(r, x)dr. \end{aligned}$$

Intergration by parts gives

$$(8.3) \qquad \begin{aligned} F_\phi(x) = &-\int_{r_1}^{r_2} \frac{2}{\sigma^2 x^2}\left[u_r(r, x)\mathcal{L}_1^*\phi(r)dr + u_x(r, x)\mathcal{L}_2^*\phi(r)\right]dr - \int_{r_1}^{r_2} \frac{2}{\sigma^2 x}\phi(r)u_x(r, x)dr \\ &+\int_{r_1}^{r_2} \frac{2}{\sigma^2 x^2}\phi'(r)v_t(t_0, r, x)dr + \int_{r_1}^{r_2} \frac{2}{\sigma^2 x^2}\phi(r)u(r, x)dr, \end{aligned}$$

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 (8.3) 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

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

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 $\phi$ 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 non-increasing 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 non-increasing, 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

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

Consider a PDE

$$
\text{(8.4)} \qquad
\begin{aligned}
w_t(t,r,x) + (\mathcal{L} - r) w(t,r,x) &= 0, && (t,r,x) \in R, \\
w(t,r,x) &= v(t,r,x), && (t,r,x) \in \partial_p R,
\end{aligned}
$$

where $\partial_p R$ denotes the parabolic boundary of $R$. By [20, Theorem 6, p. 65], Equation (8.4) 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 $\phi_1$ be a $C^\infty$ function with compact support in $(x_0, x_1)$ and $\phi_2$ be a $C^\infty$ 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 (8.4) by $\phi_1(x)\phi_2(r)$ and integrate over $(r_1, r_2) \times (x_0, x_1)$:

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

Integration by parts gives

$$
\text{(8.5)} \qquad \int_{r_1}^{r_2} \int_{x_0}^{x_1} \phi_1(x)\phi_2(r) w_t(t_n, r, x) dx dr + \int_{r_1}^{r_2} \int_{x_0}^{x_1} w(t_n, r, x)(\mathcal{L}^* - r)\phi_1(x)\phi_2(r) dx dr = 0,
$$

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, (8.5) reads

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

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., [33]). □

*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 [11, Theorem 2.1]. We delay the verification of the assumptions required until the end of the proof. Using that

$$(8.6) \quad \begin{aligned} (\partial_t + \mathcal{L} - r)v(t, r, x) &= 0, & r < c(t, x), \\ (\partial_t + \mathcal{L} - r)v(t, r, x) &= (\partial_t + \mathcal{L} - r)(K - x) = -rK, & r > c(t, x), \end{aligned}$$

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

$$(8.7) \quad \begin{aligned} &e^{-\int_0^{s \wedge \tau_n} r_v dv} v(t + s \wedge \tau_n, r_{s \wedge \tau_n}, X_{s \wedge \tau_n}) \\ &= v(t, r, x) - \int_0^{s \wedge \tau_n} e^{-\int_0^u r_v dv} K r_u 1_{\{r_u > c(t+u, X_u)\}} du + \int_0^{s \wedge \tau_n} e^{-\int_0^u r_v dv} \sigma X_u v_x(t+u, r_u, X_u) dB_u \\ &+ \int_0^{s \wedge \tau_n} e^{-\int_0^u r_v dv} \beta(r_u) v_r(t+u, r_u, X_u) dW_u. \end{aligned}$$

Taking expectations and applying the optional sampling theorem we arrive at

$$(8.8)$$
$$v(t, r, x) = \mathsf{E}_{r,x}\left[\int_0^{\tau \wedge \tau_n} e^{-\int_0^u r_v dv} K r_u 1_{\{r_u > c(t+u, X_u)\}} du + e^{-\int_0^{\tau \wedge \tau_n} r_v dv} v(t+(\tau \wedge \tau_n), r_{\tau \wedge \tau_n}, X_{\tau \wedge \tau_n})\right].$$

Using (2.3) and (2.4), 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 (8.8) follows from Assumption 2.1 (details can be found in the proof of (9.1) in Lemma 9.1). The dominated convergence theorem proves (3.4), since $\tau_n \uparrow T - t$ upon recalling that the boundary of $\mathcal{I} \times \mathbb{R}_+$ is assumed non-attainable by the process $(r_t, X_t)$.

It remains to verify assumptions of [11, Theorem 2.1]. 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 $\beta$. The marginal distribution of the process $(r_t, X_t)$ has density with respect to the Lebesgue measure (Remark B.1), so $(t, r_t, X_t) \notin \partial \mathcal{C}$, $\mathsf{P}_{r,x}$-a.s. for any $t > 0$. This verifies Assumption A.1 of [11, Thm. 2.1]. In the notation of Assumption A.2 in [11, Thm. 2.1], using (8.6), we have

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

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 Assumption A.3 in [11, Thm. 2.1] holds by Proposition 3.3. □

*Proof of Proposition 3.10.* The proof follows ideas originally developed in [39]. Assume there exists another continuous function $\tilde{c}$ that satisfies conditions (1) and (2) in the statement of this

proposition. Define a function

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

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

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

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

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

$$(8.9) \quad \tilde{v}(t, r, x) = \mathsf{E}_{r,x}\big[\tilde{V}_\tau\big] = \mathsf{E}_{r,x}\left[\int_0^\tau e^{-\int_0^u r_v dv} K r_u \mathbf{1}_{\{r_u > \tilde{c}(t+u, X_u)\}} du + e^{-\int_0^\tau r_v dv}\tilde{v}(t+\tau, r_\tau, X_\tau)\right],$$

which is analogous to the formula for $v$ in (3.4).

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 $r \geq \tilde{c}(t, x)$. 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

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

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 (8.9), this yields

$$(8.10) \qquad \tilde{v}(\hat{t}, \hat{r}, \hat{x}) = \mathsf{E}_{\hat{r}, \hat{x}}\left[\int_0^{\tau_1} e^{-\int_0^u r_v dv} K r_u du + e^{-\int_0^{\tau_1} r_v dv}(K - X_{\tau_1})^+\right],$$

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

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

where $L^K(X)$ is the local time of the process $X$ at $K$. The local time $L^K(X)$ is null until $\tau_1$ since $r_u > \tilde{c}(t+u, X_u) \implies X_u < K$, recalling that $\tilde{c}(t, x) = \bar{r}$ when $x \geq K$. Compare the right-hand side of the above expression to (8.10) to conclude that $\tilde{v}(\hat{t}, \hat{r}, \hat{x}) = (K - \hat{x})^+$.

*Step 2.* The next step is to show that $\tilde{v} \leq v$ for $(t, r, x) \in \mathcal{O}$. Since we have already proved $\tilde{v}(t, r, x) = (K - x)^+ \leq v(t, r, x)$ when $r \geq \tilde{c}(t, x)$, 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 (8.9)

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

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 $\tilde{c}(t, x) = c(t, x) = \bar{r}$). 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 (3.4) and (8.9), 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

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

As the function $c$ is non-negative, $r_u \geq 0$ on $\{u < \tau_3\}$ and we conclude that

$$(8.11) \qquad \mathsf{E}_{\hat{r}, \hat{x}} \left[ \int_0^{\tau_3} 1_{\{r_u \leq \tilde{c}(\hat{t}+u, X_u)\}} du \right] = 0.$$

The dynamics of $(r, X)$ is non-degenerate 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 (B.4) in Appendix). Hence, by the continuity of $\tilde{c}$ and $c$, for a sufficiently small $\varepsilon > 0$,

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

Paired with the continuity of trajectories of $(r, X)$, it contradicts (8.11).

*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 (3.4) and (8.9), 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 $\tilde{c}(t, x) \leq c(t, x)$ 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 $\tilde{v} \leq v$ and comparing the two equations above give us

$$\mathsf{E}_{\hat{r}, \hat{x}} \left[ \int_0^{\tau_3} e^{-\int_0^u r_v dv} K r_u 1_{\{r_u > \tilde{c}(\hat{t}+u, X_u)\}} du \right] \leq 0.$$

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 3.11.* We can repeat the same arguments as in the proof of Proposition 3.10, 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 (9.1) 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*

$$(9.1) \qquad \sup_{(r,x)\in\mathcal{K}} \sup_{t\in[0,T]} \mathsf{E}_{r,x}\left[\sup_{0\leq s\leq T-t} e^{-\int_0^s r_u du} v(t+s, r_s, X_s)\right] < \infty,$$

$$(9.2) \qquad \sup_{(r,x)\in\mathcal{K}} \sup_{s\in[0,T]} \mathsf{E}_{r,x}\left[e^{-p\int_0^s r_u du}\left|v_x(t+s, r_s, X_s)\right|^p X_s^p\right] < \infty,$$

$$(9.3) \qquad \sup_{(r,x)\in\mathcal{K}} \sup_{s\in[0,T]} \mathsf{E}_{r,x}\left[e^{-2\int_0^s r_u du}\left(v_r(t+s, r_s, X_s)\right)^2 \beta^2(r_s)\right] < \infty.$$

*Proof.* From (2.6) we obtain an upper bound for the function $v$:

$$(9.4) \qquad v(t,r,x) \leq K\mathsf{E}_r\left[\sup_{0\leq s\leq T-t} e^{-\int_0^s r_u du}\right].$$

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

$$e^{-\int_0^s r_u du} v(t+s, r_s, X_s) \leq e^{-\int_0^s r_u du} K\mathsf{E}_{r_s}\left[\sup_{0\leq u\leq T-t-s} e^{-\int_0^u r_v dv}\right]$$

$$= e^{-\int_0^s r_u du} K\mathsf{E}_r\left[\sup_{s\leq u\leq T-t} e^{-\int_s^u r_v dv}\Big|\mathcal{F}_s\right]$$

$$\leq K\mathsf{E}_r\left[\sup_{0\leq u\leq T-t} e^{-\int_0^u r_v dv}\Big|\mathcal{F}_s\right] \leq K\mathsf{E}_r\left[\sup_{0\leq u\leq T} e^{-\int_0^u r_v dv}\Big|\mathcal{F}_s\right],$$

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

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

$$\leq \sup_{(r,x)\in\mathcal{K}} \sup_{t\in[0,T]} K\,\mathsf{E}_r\left[\sup_{0\leq s\leq T-t} Y_s\right] \leq \sup_{(r,x)\in\mathcal{K}} K\,\mathsf{E}_r\left[\sup_{0\leq s\leq T} Y_s\right]$$

$$\leq \sup_{(r,x)\in\mathcal{K}} 2K\left(\mathsf{E}_r[Y_T^2]\right)^{1/2} = \sup_{(r,x)\in\mathcal{K}} 2K\left(\mathsf{E}_r\left[\sup_{0\leq u\leq T} e^{-2\int_0^u r_v dv}\right]\right)^{1/2} \leq 2K(C_1)^{1/2},$$

where $C_1$ is the constant from (2.3). This proves (i).
We now address (9.2). We have

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

where we use $-1 \leq v_x \leq 0$ in the last inequality, which follows from (7.3). From here, (9.2) is immediate.

It remains to prove (9.3). First we consider the case of Assumption 2.1(ii). From (5.2), (5.3) and (5.4), we deduce

$$(9.5) \qquad \left(v_r(t,r,x)\right)^2 \leq c_1\mathsf{E}_r\left[\sup_{0\leq s\leq T-t} e^{-2\int_0^s r_u du}\right]$$

for some constant $c_1 > 0$ depending only on $T$ and the coefficients of (2.2) (notice in particular that the expected value in the right-hand side above comes from the constant $C_1$ in (5.3)). Hence

$$e^{-2\int_0^s r_u du}\big(v_r(t+s,r_s,X_s)\big)^2 \beta^2(r_s) \leq e^{-2\int_0^s r_u du} c_1 \, \mathsf{E}_{r_s}\Big[\sup_{0 \leq u \leq T-t-s} e^{-2\int_0^u r_v dv}\Big]\beta^2(r_s)$$

$$\leq c_1 \, \mathsf{E}_r\Big[\sup_{0 \leq u \leq T-t} e^{-2\int_0^u r_v dv}\Big|\mathcal{F}_s\Big]\beta^2(r_s),$$

where the last inequality is by the same argument as in the proof of (9.1). We take expectation of both sides and apply Hölder inequality with $q = p'/2$ ($p' > 2$ is defined in Assumption 2.1) and $q' = q/(q-1)$

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

$$\leq c_1 \left(\mathsf{E}_r\Big[\sup_{0 \leq u \leq T-t} e^{-p'\int_0^u r_v dv}\Big]\right)^{1/q} \big(\mathsf{E}_r[\beta^{2q'}(r_s)]\big)^{1/q'}$$

$$\leq c_1 \, C_1^{1/q}\big(\mathsf{E}_r[\beta^{2q'}(r_s)]\big)^{1/q'},$$

where the second inequality follows from Jensen's inequality and $C_1$ is the constant from (2.3). Let $L$ be the Lipschitz constant for $\beta$. Then, using triangle inequality for norms,

$$\Big(\mathsf{E}_r[(\beta(r_s))^{2q'}]\Big)^{1/q'} \leq \Big(\mathsf{E}_r\big[|\beta(0)+L|r_s||^{2q'}\big]\Big)^{1/q'} = \Big(\Big(\mathsf{E}_r\big[|\beta(0)+L|r_s||^{2q'}\big]\Big)^{1/2q'}\Big)^2$$

$$\leq \Big(\beta(0)+L\big(\mathsf{E}_r\big[|r_s|^{2q'}\big]\big)^{1/2q'}\Big)^2 \leq \Big(\beta(0)+L\big(C_2(1+|r|^{2q'})^{1/2q'}\Big)^2,$$

where the last inequality follows from (2.4) and $2q' = p' \geq 2$. Combining the above estimates proves (9.3).

We address the case when $r$ follows the CIR dynamics. From the non-negativity of the process $r$ and from (9.5) we obtain that $\big(v_r(t,r,x)\big)^2 \leq c_1$ for any $(t,r,x) \in \mathcal{O}$. Hence, we write

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

where we used the explicit form of $\beta$. It remains to recall (2.4) to conclude (9.3). $\qquad\square$

*Proof of Proposition 3.12.* The admissibility condition can be equivalently written as

$$(9.6) \qquad \int_0^T e^{-2\int_0^s r_u du}\big(\phi_s^{(1)}\sigma X_s\big)^2 ds + \int_0^T e^{-2\int_0^s r_u du}\big(\phi_s^{(2)}\beta(r_s)P_r(s,r_s)\big)^2 ds < \infty, \qquad \mathsf{P}_{r,x}\text{-a.s.}$$

Estimates in Lemma 9.1 imply

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

which is a stronger condition than (9.6). The fact that the portfolio replicates the option follows from the construction and (3.15). $\qquad\square$

# Appendices

## Appendix A. Convergence of stopping times

In this appendix we provide a self-contained exposition around the continuity of hitting/entry times to certain domains in $\mathcal{O}$. Similar results can be found in many textbooks, e.g., [17, p. 32-40], [9, Chapter 1].

Throughout this section, $\mathcal{K}$ denotes a subset of $\mathcal{O}$ that is closed in $\mathcal{O}$, i.e., $\overline{\mathcal{K}} \cap \mathcal{O} = \mathcal{K}$. Introduce the *hitting* time to $\mathcal{K}$, denoted $\sigma_{\mathcal{K}}$, and the *entry* time to the interior of $\mathcal{K}$, denoted $\mathring{\sigma}_{\mathcal{K}}$. That is, for $(t, r, x) \in [0, T] \times \mathcal{I} \times \mathbb{R}_+$, we set $\mathsf{P}_{r,x}$-a.s.

$$
(A.1) \qquad
\begin{aligned}
\sigma_{\mathcal{K}} &:= \inf\{s > 0 : (t + s, r_s, X_s) \in \mathcal{K}\} \wedge (T - t), \\
\mathring{\sigma}_{\mathcal{K}} &:= \inf\{s \geq 0 : (t + s, r_s, X_s) \in \mathrm{int}(\mathcal{K})\} \wedge (T - t).
\end{aligned}
$$

The above definition differs from (6.5) since $\{T\} \times \mathcal{I} \times \mathbb{R}_+ \subset \mathcal{D}$ guaranteeing that $\sigma_{\mathcal{D}} \leq T - t$. Both $\sigma_{\mathcal{K}}$ and $\mathring{\sigma}_{\mathcal{K}}$ are stopping times with respect to the augmentation of the filtration generated by the Brownian motions $(B, W)$. It is immediate to see that

$$
(A.2) \qquad\qquad\qquad \sigma_{\mathcal{K}} \leq \mathring{\sigma}_{\mathcal{K}}, \qquad \mathsf{P}\text{-a.s.}
$$

We will often write $\sigma_{\mathcal{K}}(t, r, x)$ and $\mathring{\sigma}_{\mathcal{K}}(t, r, x)$ to indicate the starting point of the process.

Denote by $\mathcal{A}$ the complement of $\mathcal{K}$ in $\mathcal{O}$: $\mathcal{A} := \mathcal{K}^c \cap \mathcal{O}$ (which is an open set). When we write $\partial \mathcal{A}$ we mean the boundary of $\mathcal{A}$ in $[0, T] \times \mathcal{I} \times \mathbb{R}_+$. We have two assumptions

**Assumption A.1** (Regularity). *For $(t_0, r_0, x_0) \in \partial \mathcal{A}$, we have*

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

**Assumption A.2.** *For any sequence $(r_n, x_n)_{n \geq 1}$ converging to $(r, x) \in \mathcal{I} \times \mathbb{R}_+$ as $n \to \infty$, it holds that*

$$
(A.3) \qquad \lim_{n \to +\infty} \sup_{0 \leq t \leq T} \left( |r_t^{r_n} - r_t^r| + |X_t^{r_n, x_n} - X_t^{r,x}| \right) = 0, \qquad \mathsf{P}\text{-a.s.}
$$

Note that Assumption A.1 necessitates that $\mathcal{K} = \overline{\mathrm{int}(\mathcal{K})} \cap \mathcal{O}$. It also enables establishing a connection between $\sigma_{\mathcal{K}}$ and $\mathring{\sigma}_{\mathcal{K}}$.

**Lemma A.3.** *Under Assumption A.1, $\mathsf{P}_{t,r,x}(\sigma_{\mathcal{K}} = \mathring{\sigma}_{\mathcal{K}}) = 1$ for all $(t, r, x) \in \mathcal{O}$.*

*Proof.* The equality is trivial for $(t, r, x) \in \mathrm{int}(\mathcal{K})$. Take $(t, r, x)$ in its complement, i.e., in $\overline{\mathcal{A}} \cap \mathcal{O}$. Since (A.2) holds, we only need to show that $\mathsf{P}_{t,r,x}(\sigma_{\mathcal{K}} < \mathring{\sigma}_{\mathcal{K}}) = 0$. Let us argue by contradiction and assume that $\mathsf{P}_{t,r,x}(\sigma_{\mathcal{K}} < \mathring{\sigma}_{\mathcal{K}}) > 0$. There exists $\delta > 0$ such that $\mathsf{P}_{t,r,x}(\mathring{\sigma}_{\mathcal{K}} \geq \sigma_{\mathcal{K}} + \delta) > 0$. Denoting by $\theta_t$ the shift operator, by the strong Markov property we get

$$
\begin{aligned}
\mathsf{P}_{t,r,x}(\mathring{\sigma}_{\mathcal{K}} \geq \sigma_{\mathcal{K}} + \delta) &= \mathsf{E}_{t,r,x}\left[ \mathsf{E}_{t,r,x}\left( 1_{\{\sigma_{\mathcal{K}} + \mathring{\sigma}_{\mathcal{K}} \circ \theta_{\sigma_{\mathcal{K}}} \geq \sigma_{\mathcal{K}} + \delta\}} \Big| \mathcal{F}_{\sigma_{\mathcal{K}}} \right) \right] \\
&= \mathsf{E}_{t,r,x}\left[ \mathsf{P}_{t + \sigma_{\mathcal{K}}, r_{\sigma_{\mathcal{K}}}, X_{\sigma_{\mathcal{K}}}}(\mathring{\sigma}_{\mathcal{K}} \geq \delta) \right] = 0,
\end{aligned}
$$

where the last equality follows by observing that $(t + \sigma_{\mathcal{K}}, r_{\sigma_{\mathcal{K}}}, X_{\sigma_{\mathcal{K}}}) \in \partial \mathcal{A}$, $\mathsf{P}_{t,r,x}$-a.s. and Assumption A.1. $\qquad\square$

**Lemma A.4.** *Let $\mathcal{O} \ni (t_n, r_n, x_n) \to (t, r, x) \in \mathcal{O}$ as $n \to \infty$. Under Assumption A.2, it holds $\mathsf{P}$-a.s.*

$$
\limsup_{n \to \infty} \mathring{\sigma}_{\mathcal{K}}(t_n, r_n, x_n) \leq \mathring{\sigma}_{\mathcal{K}}(t, r, x).
$$

*Proof.* For simplicity, we denote $\mathring{\sigma}_n := \mathring{\sigma}_{\mathcal{K}}(t_n, r_n, x_n)$ and $\mathring{\sigma}_{\mathcal{K}} := \mathring{\sigma}_{\mathcal{K}}(t, r, x)$. For $\mathsf{P}$-a.e. $\omega \in \Omega$ we have by (A.3)

$$
(A.4) \qquad (t_n + s, r_s^{r_n}, X_s^{r_n, x_n})(\omega) \to (t + s, r_s^r, X_s^{r,x})(\omega), \qquad s \in [0, T - t].
$$

Fix $\omega \in \Omega$ in the set of P-full measure for which the above holds. If $\mathring{\sigma}_{\mathcal{K}}(\omega) = T - t$ then the result is obvious because $\mathring{\sigma}_n(\omega) \leq T - t_n$. Assume $\mathring{\sigma}_{\mathcal{K}}(\omega) < T - t$. Take any $\delta < T - t$ such that $\mathring{\sigma}_{\mathcal{K}}(\omega) < \delta$. By the continuity of paths and the openness of $\mathrm{int}(\mathcal{K})$, there is $\delta' \in (\mathring{\sigma}_{\mathcal{K}}(\omega), \delta)$ such that $(t + \delta', r_{\delta'}^r, X_{\delta'}^{r,x})(\omega) \in \mathrm{int}(\mathcal{K})$. From (A.4) and the openness of $\mathrm{int}(\mathcal{K})$, $(t_n + \delta', r_{\delta'}^{r_n}, X_{\delta'}^{r_n,x_n})(\omega) \in \mathrm{int}(\mathcal{K})$ for all sufficiently large $n$, so $\limsup_{n\to\infty} \mathring{\sigma}_n(\omega) \leq \delta'$. As the above argument holds for any $\delta > \mathring{\sigma}_{\mathcal{K}}(\omega)$ and for a.e. $\omega \in \Omega$, we obtain the claim. $\qquad\square$

**Lemma A.5.** *Let $(t_n, r_n, x_n)_{n\geq 1}$ be a sequence converging to $(t, r, x) \in \mathcal{O}$ as $n \to \infty$. Then, under Assumptions A.1 and A.2,*

$$\text{(A.5)} \qquad \liminf_{n\to\infty} \sigma_{\mathcal{K}}(t_n, r_n, x_n) \geq \sigma_{\mathcal{K}}(t, r, x), \qquad \text{P-}a.s.$$

*Proof.* For $y, z \in \mathcal{O}$ we denote by $\mathrm{d}(y, z)$ their Euclidean distance and by $\mathrm{d}(y, \partial\mathcal{A}) = \inf\{\mathrm{d}(y, z), z \in \partial\mathcal{A}\}$. Denote $\sigma_n := \sigma_{\mathcal{K}}(t_n, r_n, x_n)$ and $\sigma_{\mathcal{K}} := \sigma_{\mathcal{K}}(t, r, x)$. To simplify notation we also set

$$\zeta_s := (t + s, r_s^r, X_s^{r,x}) \quad \text{and} \quad \zeta_s^n := (t_n + s, r_s^{r_n}, X_s^{r_n,x_n}).$$

Fix $\omega \in \Omega$ in a set of full P-measure on which trajectories of $\zeta$ are continuous and the limit (A.3) holds. If $\sigma_{\mathcal{K}}(\omega) = 0$ then (A.5) holds trivially. Otherwise, we must have $(t, r, x) \in \mathcal{A}$ by Assumption (A.1), so $\mathrm{d}(\zeta_0(\omega), \partial\mathcal{A}) > 0$ since $\mathcal{A}$ is open. For $0 < \varepsilon < \mathrm{d}(\zeta_0(\omega), \partial\mathcal{A})$, define

$$\delta_\varepsilon = \inf\{s \in [0, T - t] \,:\, \mathrm{d}(\zeta_s(\omega), \partial\mathcal{A}) \leq \varepsilon\}.$$

Using the triangle inequality we get $\mathrm{d}(\zeta_s^n(\omega), \partial\mathcal{A})) + \mathrm{d}(\zeta_s^n(\omega), \zeta_s(\omega)) > \varepsilon$ for all $s \in [0, \delta_\varepsilon]$. Thanks to Assumption A.2, $\mathrm{d}(\zeta_s^n(\omega), \zeta_s(\omega)) < \varepsilon/2$ for all $s \in [0, \delta_\varepsilon]$ and all sufficiently large $n$, so $\mathrm{d}(\zeta_s^n(\omega), \partial\mathcal{A})) \geq \varepsilon/2$ for all $s \in [0, \delta_\varepsilon]$ and all sufficiently large $n$. This gives

$$\liminf_{n\to\infty} \sigma_n(\omega) \geq \delta_\varepsilon.$$

Using the continuity of $t \mapsto \zeta_t(\omega)$ and the fact that $\zeta_{\sigma_{\mathcal{K}}}(\omega) \in \partial\mathcal{A}$, we have $\lim_{\varepsilon\to 0} \delta_\varepsilon = \sigma_{\mathcal{K}}(\omega)$. As this holds for a.e. $\omega \in \Omega$, the proof of (A.5) is complete. $\qquad\square$

Lemma A.3, A.4 and A.5 imply

**Proposition A.6.** *Let $(t_n, r_n, x_n)_{n\geq 1}$ be a sequence converging to $(t, r, x) \in \mathcal{O}$ as $n \to \infty$. Then*

$$\text{(A.6)} \qquad \begin{aligned} \lim_{n\to\infty} \mathring{\sigma}_{\mathcal{K}}(t_n, r_n, x_n) &= \mathring{\sigma}_{\mathcal{K}}(t, r, x), \qquad \text{P-}a.s. \\ \lim_{n\to\infty} \sigma_{\mathcal{K}}(t_n, r_n, x_n) &= \sigma_{\mathcal{K}}(t, r, x), \qquad \text{P-}a.s. \end{aligned}$$

## Appendix B. Proof of Proposition 6.1

Fix $(t_0, r_0, x_0) \in \partial\mathcal{C}$ and define $\mathcal{R} := [r_0, \overline{r}] \times [0, x_0]$, where we also recall that $\mathcal{I} = (\underline{r}, \overline{r})$. Since $t \mapsto c(t, x)$ is non-increasing, it is immediate to see that $[t_0, T] \times \mathcal{R} \subseteq \mathcal{D}$. Recalling the notation introduced in (A.1), set

$$\mathring{\sigma}_{\mathcal{R}}(r_0, x_0) := \inf\{s \geq 0 \,:\, (r_s^{r_0}, X_s^{r_0,x_0}) \in \mathrm{int}(\mathcal{R})\}.$$

We have $\mathring{\sigma}_{\mathcal{R}}(r_0, x_0) \geq \mathring{\sigma}_{\mathcal{D}}(t_0, r_0, x_0)$, P-a.s., and $\mathring{\sigma}_{\mathcal{R}}(r_0, x_0) \geq \sigma_{\mathcal{D}}(t_0, r_0, x_0)$ by the continuity of the process $(r, X)$. From now on we omit in the notation the dependence on $(t_0, r_0, x_0)$ since the initial point is fixed throughout the proof.

Take a compact ball $\mathcal{K} \subset \mathcal{I} \times \mathbb{R}_+$ centred at $(r_0, x_0)$. Let $\Sigma(r, x)$ denote the matrix of the diffusion coefficient for (2.1)–(2.2), i.e.

$$\Sigma(r, x) := \frac{1}{2} \begin{pmatrix} \sigma^2 x^2 & \rho\sigma x\beta(r) \\ \rho\sigma x\beta(r) & \beta^2(r) \end{pmatrix}.$$

Since the correlation coefficient $\rho \in (-1, 1)$, there is $\gamma = \gamma_{\mathcal{K}} > 0$ such that

$$\text{(B.1)} \qquad \frac{1}{\gamma}\|z\|^2 \leq \langle \Sigma(r, x)z, z \rangle \leq \gamma\|z\|^2, \qquad z \in \mathbb{R}^2 \setminus \{0\}, \ (r, x) \in \mathcal{K},$$

where $\langle\cdot,\cdot\rangle$ denotes the scalar product in $\mathbb{R}^2$ and $\|\cdot\|$ the corresponding norm.

Define a new process $(\widetilde{r}, \widetilde{X})$ with the dynamics defined on $\mathbb{R}^2$

(B.2) $$d\,\widetilde{X}_t = \mu_{\mathcal{K}}(\widetilde{r}_t, \widetilde{X}_t)dt + \sigma_{\mathcal{K}}(\widetilde{X}_t)dB_t, \qquad \widetilde{X}_0 = x_0$$

(B.3) $$d\,\widetilde{r}_t = \alpha_{\mathcal{K}}(\widetilde{r}_t)dt + \beta_{\mathcal{K}}(\widetilde{r}_t)dW_t, \qquad \widetilde{r}_0 = r_0$$

such that the coefficients coincide with the coefficients of (2.1)-(2.2) on $\mathcal{K}$, are Lipschitz continuous on $\mathbb{R}^2$ and satisfy the uniform ellipticity condition (B.1) with $\gamma$ on $\mathbb{R}^2$. Denoting $\tau_{\mathcal{K}} := \inf\{t \geq 0 : (r_t, X_t) \notin \mathrm{int}(\mathcal{K})\}$ and $\widetilde{\tau}_{\mathcal{K}} := \inf\{t \geq 0 : (\widetilde{r}_t, \widetilde{X}_t) \notin \mathrm{int}(\mathcal{K})\}$, by the uniqueness of solutions for SDEs we get indistinguishable stopped paths:

$$(r_{t\wedge\tau_{\mathcal{K}}}, X_{t\wedge\tau_{\mathcal{K}}})_{t\geq0} = (\widetilde{r}_{t\wedge\widetilde{\tau}_{\mathcal{K}}}, \widetilde{X}_{t\wedge\widetilde{\tau}_{\mathcal{K}}})_{t\geq0} \quad \mathsf{P}_{r_0,x_0}\text{-a.s.}$$

The uniform ellipticity condition (B.1) on $\mathbb{R}^2$ implies that the process $(\widetilde{r}, \widetilde{X})$ admits a transition density $\widetilde{p}(t, (r,x), (r',x'))$ which satisfies the following Gaussian bound (see, e.g., [4, 18]): there exists $m > 0$ and $\Lambda > 0$ such that

(B.4) $$\widetilde{p}(t, (r,x), (r',x')) \geq m\,t^{-1}\exp\left(-\Lambda\frac{(r'-r)^2 + (x'-x)^2}{t}\right).$$

Let $\mathcal{R}''$ be a closed cone with vertex $(r_0, x_0)$ and non-empty interior contained in $\big[(r_0, \infty) \times (-\infty, x_0)\big] \cup \{r_0, x_0\}$. Put $\mathcal{R}' = \mathcal{R}'' \cap (\mathcal{I} \times \mathbb{R}_+)$. Denote by $\mathring{\sigma}'_{\mathcal{R}}$ the entry time of $(r, X)$ to $\mathrm{int}(\mathcal{R}')$ and by $\mathring{\sigma}''_{\mathcal{R}}$ the entry time of $(\widetilde{r}, \widetilde{X})$ to $\mathrm{int}(\mathcal{R}'')$. The next estimate relies on analogous results for multi-dimensional Brownian motion ([29, Thm. 4.2.9]); in particular the second inequality below follows from (B.4):

$$\mathsf{P}_{r_0,x_0}(\mathring{\sigma}''_{\mathcal{R}} \leq t) \geq \mathsf{P}_{r_0,x_0}((\widetilde{r}_t, \widetilde{X}_t) \in \mathcal{R}'') \geq \frac{m}{t}\int_{\mathcal{R}''}\exp\left(-\Lambda\frac{(r-r_0)^2 + (x-x_0)^2}{t}\right)dr\,dx.$$

We change variables to $y := (r - r_0)/\sqrt{t}$ and $z := (x - x_0)/\sqrt{t}$ and use that $\mathcal{R}''$ is invariant under this transformation to obtain

(B.5) $$\mathsf{P}_{r_0,x_0}(\mathring{\sigma}''_{\mathcal{R}} \leq t) \geq 2\pi m\int_{\mathcal{R}''}\frac{1}{2\pi}e^{-\Lambda(y^2+z^2)}dy\,dz =: q > 0.$$

For any $t > 0$, since $\mathcal{R}' \subset \mathcal{R}$ we have

$$\mathsf{P}_{r_0,x_0}(\mathring{\sigma}_{\mathcal{R}} \leq t) \geq \mathsf{P}_{r_0,x_0}(\mathring{\sigma}'_{\mathcal{R}} \leq t) \geq \mathsf{P}_{r_0,x_0}(\mathring{\sigma}'_{\mathcal{R}} \leq t, \tau_{\mathcal{K}} > t)$$
$$= \mathsf{P}_{r_0,x_0}(\mathring{\sigma}''_{\mathcal{R}} \leq t, \widetilde{\tau}_{\mathcal{K}} > t) \geq \mathsf{P}_{r_0,x_0}(\mathring{\sigma}''_{\mathcal{R}} \leq t) - \mathsf{P}_{r_0,x_0}(\widetilde{\tau}_{\mathcal{K}} \leq t) \geq q - \mathsf{P}_{r_0,x_0}(\widetilde{\tau}_{\mathcal{K}} \leq t),$$

where the last inequality is by (B.5). As $t \downarrow 0$, we have $\mathsf{P}_{r_0,x_0}(\widetilde{\tau}_{\mathcal{K}} \leq t) \to 0$ and $\mathsf{P}_{r_0,x_0}(\mathring{\sigma}_{\mathcal{R}} \leq t) \to \mathsf{P}_{r_0,x_0}(\mathring{\sigma}_{\mathcal{R}} = 0)$, which implies that $\mathsf{P}_{r_0,x_0}(\mathring{\sigma}_{\mathcal{R}} = 0) \geq q > 0$. By the Blumenthal $0-1$ law [29, Thm. 2.7.17] we obtain $\mathsf{P}_{r_0,x_0}(\mathring{\sigma}_{\mathcal{R}} = 0) = 1$. Recalling that $\mathsf{P}_{t_0,r_0,x_0}(\mathring{\sigma}_{\mathcal{R}} \geq \mathring{\sigma}_{\mathcal{D}}) = 1$, and $\mathsf{P}_{t_0,r_0,x_0}(\mathring{\sigma}_{\mathcal{D}} \geq \sigma_{\mathcal{D}}) = 1$, we conclude $\mathsf{P}_{t_0,r_0,x_0}(\mathring{\sigma}_{\mathcal{D}} = 0) = \mathsf{P}_{t_0,r_0,x_0}(\sigma_{\mathcal{D}} = 0) = 1$. $\qquad\square$

**Remark B.1.** *It is worth noticing that the arguments above show the existence of the transition density of the process $(\widetilde{r}, \widetilde{X})$ for any compact set $\mathcal{K} \subset \mathcal{I} \times \mathbb{R}_+$ such that $\mathcal{K} = \overline{\mathrm{int}(\mathcal{K})}$. This implies that for each $t \in [0, T]$ also the law of $(r_t, X_t)$ is absolutely continuous with respect to the Lebesgue measure on $\mathcal{I}\times\mathbb{R}_+$, when the boundary of $\mathcal{I}\times\mathbb{R}_+$ is non-attainable by $(r_t, X_t)$. Indeed, let $N \subset \mathcal{I}\times\mathbb{R}_+$ be such that $\lambda(N) = 0$, with $\lambda$ denoting the Lebesgue measure on $\mathbb{R}^2$. Let $\mathcal{K} \subset \mathcal{I}\times\mathbb{R}_+$ be a compact set such that $\mathcal{K} = \overline{\mathrm{int}(\mathcal{K})}$. Then by the same construction as above*

(B.6) 
$$\mathsf{P}_{r_0,x_0}\big((r_t, X_t) \in N\big) = \mathsf{P}_{r_0,x_0}\big((r_t, X_t) \in N, t \leq \tau_{\mathcal{K}}\big) + \mathsf{P}_{r_0,x_0}\big((r_t, X_t) \in N, t > \tau_{\mathcal{K}}\big)$$
$$= \mathsf{P}_{r_0,x_0}\big((\widetilde{r}_t, \widetilde{X}_t) \in N, t \leq \widetilde{\tau}_{\mathcal{K}}\big) + \mathsf{P}_{r_0,x_0}\big((r_t, X_t) \in N, t > \tau_{\mathcal{K}}\big)$$
$$\leq \mathsf{P}_{r_0,x_0}\big((\widetilde{r}_t, \widetilde{X}_t) \in N\big) + \mathsf{P}_{r_0,x_0}\big(\tau_{\mathcal{K}} < t\big) = \mathsf{P}_{r_0,x_0}\big(\tau_{\mathcal{K}} < t\big),$$

*where the final equality uses that the transition law of $(\widetilde{r}, \widetilde{X})$ is absolutely continuous with respect to $\lambda$. Now, letting $\mathcal{K} \uparrow \mathcal{I} \times \mathbb{R}_+$, using that $0$ and $+\infty$ are not attainable by $X$ and $\underline{r}$ and $\overline{r}$ are not attainable by $r$, we can make $\mathsf{P}_{r_0, x_0}(\tau_{\mathcal{K}} < t)$ arbitrarily small, which proves the claim.*

## Appendix C. Derivation of (4.4)-(4.5)

We start by introducing the following notation (we suppress dependence on $r$ and $x$ for the sake of simplicity):

(C.1)
$$d_1(t,T) := \frac{\log\left(K\frac{P(t,T)}{x}\right) + \frac{\gamma_1(t,T)}{2}}{\sqrt{\gamma_1(t,T)}}, \qquad d_2(t,T) := d_1(t,T) - \sqrt{\gamma_1(t,T)},$$

$$q(t,u) := e^{-(u-t)\kappa}r + \theta(1 - e^{-(u-t)\kappa}) - \frac{\beta^2}{2}g(\kappa, u-t)^2,$$

$$\phi(t,u,y;b) := \frac{\log\left(\frac{P(t,u)}{x}b(u, q(t,u) + y\sqrt{\gamma_2(t,u)})\right) + \frac{\gamma_1(t,u)}{2} - \sqrt{\gamma_1(t,u)}\tilde{\rho}(t,u)y}{\sqrt{(1 - \tilde{\rho}(t,u)^2)\gamma_1(t,u)}},$$

$$\mu(t,u) := rg(\kappa, u-t) + \theta\left(u - t - g(\kappa, u-t)\right),$$

$$\gamma_1(t,u) := (u-t)\sigma^2 + \frac{2\rho\sigma\beta}{\kappa}\left(u - t - g(\kappa, u-t)\right) + \frac{\beta^2}{\kappa}\left(u - t - 2g(\kappa, u-t) + g(2\kappa, u-t)\right),$$

$$\gamma_2(t,u) := \beta^2 g(2\kappa, u-t),$$

$$\tilde{\rho}(t,u) := \frac{\rho\sigma\beta g(\kappa, u-t) + \frac{\beta^2}{2}g(\kappa, u-t)^2}{\sqrt{\gamma_1(t,u)\gamma_2(t,u)}},$$

$$g(a,u) := \frac{1 - e^{-au}}{a}.$$

**Lemma C.1.** *For a measurable bounded function $\varphi : (0,\infty) \times \mathbb{R} \to \mathbb{R}$ and $s \geq t$, the function*
$$u(t,s,r,x) = \mathsf{E}_{r,x}\left[e^{-\int_0^{s-t} r_u du}\varphi(X_{s-t}, r_{s-t})\right]$$

*has an explicit representation*

(C.2) $$u(t,s,r,x) = e^{-\mu(t,s)+\frac{1}{2}\beta^2\int_t^s g(\kappa, s-u)^2 du}\int_{\mathbb{R}^2}\varphi\left(xe^{L(t,s)+\sqrt{\gamma_1(t,s)}z}, \; q(t,s) + y\sqrt{\gamma_2(t,s)}\right)$$
$$e^{-\frac{1}{2(1-\tilde{\rho}(t,s)^2)}(z^2+y^2-2\tilde{\rho}(t,s)zy)}\frac{1}{2\pi\sqrt{1-\tilde{\rho}(t,s)^2}}dzdy,$$

*where*
$$L(t,s) = \mu(t,s) - \frac{\sigma^2}{2}(s-t) - \int_t^s\left(\beta^2 g(\kappa, s-u)^2 + \rho\sigma g(\kappa, s-u)\right)du.$$

We defer the proof until the end of this section and apply the above lemma to derive formulae for $P(t,T)$, $v_e$ and $v_p$. Taking $\varphi \equiv 1$ in (C.2), we have

(C.3) $$P(t,s) = \mathsf{E}_r\left[e^{-\int_0^{s-t} r_s ds}\right] = e^{-\mu(t,s)+\frac{1}{2}\beta^2\int_t^s g(\kappa, s-u)^2 du} = e^{-\mu(t,s)+\frac{\beta^2}{2\kappa^2}(s-t-2g(\kappa, s-t)+g(2\kappa, s-t))}.$$

This also implies that

$$e^{L(t,s)} = e^{\mu(t,s) - \frac{1}{2}\int_t^s \beta^2 g(\kappa, s-u)^2 du} \; e^{-\frac{\sigma^2}{2}(s-t) - \frac{1}{2}\int_t^s \beta^2 g(\kappa, s-u)^2 du - \int_t^s \rho\sigma g(\kappa, s-u)du} = \frac{1}{P(t,s)}e^{-\frac{\gamma_1(t,s)}{2}}.$$

Insert this and (C.3) into (C.2) and let
$$z = \tilde{\rho}(t,s)y + \sqrt{1-\tilde{\rho}(t,s)^2}\hat{z}.$$

This transforms the integral in (C.2) into an integral of two independent Gaussian variables

(C.4)
$$u(t,s,r,x) = P(t,s) \int_{\mathbb{R}^2} \varphi \left( \frac{x}{P(t,s)} e^{-\frac{\gamma_1(t,s)}{2} + \sqrt{\gamma_1(t,s)}\left(\tilde{\rho}(t,s)y + \sqrt{1-\tilde{\rho}(t,s)^2}\hat{z}\right)}, \; q(t,s) + \sqrt{\gamma_2(t,s)}y \right)$$
$$e^{-\frac{1}{2}(z^2+y^2)}\frac{1}{2\pi}d\hat{z}dy.$$

Now, letting $\varphi(x,r) = (K-x)^+$ and $\varphi(x,r) = Kr1_{\{x<b(s,r)\}}$, we obtain (4.4)-(4.5).

*Proof of Lemma C.1.* The proof follows the lines of similar computations in the literature, see [7] and [16]. Using the explicit expression of $r$ in (4.2) and stochastic Fubini's theorem [41, Theorem IV.64], we compute

(C.5)
$$\int_t^s r_u du = g(\kappa, s-t)r_t + \theta\left(s - t - g(\kappa, s-t)\right) + \beta \int_t^s g(\kappa, s-u)dW_u$$
$$= \mu(t,s) + \beta \int_t^s g(\kappa, s-u)dW_u.$$

Define $Z_t$ implicitly by
$$dB_t = \rho dW_t + \sqrt{1-\rho^2}dZ_t.$$

Then $Z$ is a Brownian motion that is independent of $W$. Using the explicit expression of $X$ and (C.5), we write $u$ as
(C.6)
$$u(t,s,r,x) = \mathsf{E}_r\left[e^{-\mu(t,s)-\beta\int_t^s g(\kappa,s-u)dW_u}\varphi\left(x\exp\left\{\mu(t,s) - \frac{\sigma^2}{2}(s-t)\right.\right.\right.$$
$$\left.\left.\left. + \beta\int_0^{s-t} g(\kappa,s-t-u)dW_u + \int_0^{s-t}\sigma(\rho dW_u + \sqrt{1-\rho^2}dZ_u)\right\}, r_{s-t}\right)\right].$$

Define a new measure $\widetilde{\mathsf{P}}$ by the Radon-Nikodym density
$$\frac{d\widetilde{\mathsf{P}}}{d\mathsf{P}} := e^{-\beta\int_0^{s-t} g(\kappa,s-t-u)dW_u - \frac{1}{2}\int_0^{s-t}\beta^2 g(\kappa,s-t-u)^2 du}.$$

The process $\widetilde{W}$ given by
$$d\widetilde{W}_t = dW_t + \beta g(\kappa, s-t)dt$$

is a Brownian motion under $\widetilde{\mathsf{P}}$. We write the explicit formula (C.5) for $r$ in terms of $\widetilde{W}$:
$$r_s = r_t e^{-(s-t)\kappa} + \theta(1 - e^{-(s-t)\kappa}) - \frac{\beta^2}{2}g(\kappa, s-t)^2 + \beta\int_t^s e^{-(s-u)\kappa}d\widetilde{W}_u.$$

Denoting by $\widetilde{\mathsf{E}}$ the expectation under $\widetilde{P}$, we obtain from (C.6)

(C.7)
$$u(t,s,r,x) = e^{-\mu(t,s)+\frac{1}{2}\beta^2\int_t^s g(\kappa,s-u)^2 du}\widetilde{\mathsf{E}}\left[\varphi\left(xe^{L(t,s)+A(t,s)}, \; q(t,s) + Y(t,s)\right)\right],$$

where
$$A(t,s) := \int_0^{s-t}(\rho\sigma + \beta g(\kappa, s-t-u))d\widetilde{W}_u + \int_0^{s-t}\sigma\sqrt{1-\rho^2}dZ_u,$$
$$Y(t,s) := \beta\int_0^{s-t} e^{-(s-t-u)\kappa}d\widetilde{W}_u.$$

For each fixed $t < s$, the random vector $(A, Y)$ is multivariate Gaussian under $\widetilde{\mathsf{P}}$ with zero mean, and variance and covariance given by

$$Var_{\widetilde{\mathsf{P}}}(A(t,s)) = \gamma_1(t,s), \qquad Var_{\widetilde{\mathsf{P}}}(Y(t,s)) = \gamma_2(t,s),$$
$$Cov_{\widetilde{\mathsf{P}}}(A(t,s), Y(t,T)) = \tilde{\rho}(t,s)\sqrt{\gamma_1(t,s)\gamma_2(t,s)}.$$

Hence, we have an explicit integral representation of (C.7)
(C.8)
$$u(t,s,r,x) = = e^{-\mu(t,s)+\frac{1}{2}\beta^2 \int_t^s g(\kappa, s-u)^2 du} \int_{\mathbb{R}^2} \varphi\Big(xe^{L(t,s)+z}, \quad q(t,s)+y\Big)$$
$$\times e^{-\frac{1}{2(1-\tilde{\rho}(t,s)^2)}\left(\frac{z^2}{\gamma_1(t,s)}+\frac{y^2}{\gamma_2(t,s)}-\frac{2\tilde{\rho}(t,s)zy}{\sqrt{\gamma_1(t,s)\gamma_2(t,s)}}\right)} \frac{1}{2\pi\sqrt{\gamma_1(t,s)\gamma_2(t,s)}\sqrt{1-\tilde{\rho}(t,s)^2}} dz dy.$$

A change of variable yields (C.2). $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad$ □

## References

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

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

[3] E. Appolloni, L. Caramellino, and A. Zanette. 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, 2015.

[4] D. G. Aronson. Bounds for the fundamental solution of a parabolic equation. *Bulletin of the American Mathematical Society*, 73(6):890–896, 1967.

[5] R. F. Bass. *Diffusions and elliptic operators.* Springer Science & Business Media, 1998.

[6] A. Battauz and F. Rotondi. American options and stochastic interest rates. *Working Paper*, 2019.

[7] M. Baxter and A. Rennie. *Financial calculus: an introduction to derivative pricing.* Cambridge University Press, 1996.

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

[9] R. Blumenthal and R. Getoor. *Markov Processes and Potential Theory.* Academic Press, New York, 1968.

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

[11] C. Cai and T. De Angelis. A change of variable formula with applications to multi-dimensional optimal stopping problems. *arXiv:2104.05835*, 2021.

[12] P. Carr, R. Jarrow, and R. Myneni. Alternative characterizations of American put options. *Mathematical Finance*, 2(2):87–106, 1992.

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

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

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

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

[17] E. Dynkin. *Markov Processes*, volume 2. Academic Press, New York, 1965.

[18] E. Fabes and D. Stroock. A new proof of Mosers parabolic Harnack inequality using the old ideas of Nash. In *Analysis and Continuum Mechanics*, pages 459–470. Springer, 1989.

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

[20] A. Friedman. *Partial Differential Equations of Parabolic Type*. Prentice-Hall, 1964.

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

[22] D. Heath, R. Jarrow, and A. Morton. Bond pricing and the term structure of interest rates: A new methodology for contingent claims valuation. *Econometrica*, pages 77–105, 1992.

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

[24] J. Hull. *Options, futures and other derivatives*. Upper Saddle River, NJ: Prentice Hall, 2009.

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

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

[27] M. Jeanblanc, M. Yor, and M. Chesney. *Mathematical methods for financial markets*. Springer Science & Business Media, 2009.

[28] I. Karatzas. On the pricing of American options. *Applied Mathematics & Optimization*, 17:37–60, 1988.

[29] I. Karatzas and S. E. Shreve. *Brownian Motion and Stochastic Calculus*. Springer, 1998.

[30] I. Karatzas and S. E. Shreve. *Methods of mathematical finance*, volume 39. Springer, 1998.

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

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

[33] R. Kruse and J. Deely. Joint continuity of monotonic functions. *The American Mathematical Monthly*, 76(1):74–76, 1969.

[34] N. Krylov. *Controlled Diffusion Processes*. Springer, 1980.

[35] T. Little, V. Pant, and C. Hou. A new integral representation of the early exercise boundary for American put options. *Journal of Computational Finance*, 3:73–96, 2000.

[36] H. P. McKean Jr. A free boundary problem for the heat equation arising from a problem of mathermatical economics. *Industrial Management Review*, 6:32–39, 1965.

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

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

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

[40] G. Peskir. A change-of-variable formula with local time on surfaces. In *Séminaire de probabilités XL*, pages 70–96. Springer, 2007.

[41] P. E. Protter. *Stochastic integration and differential equations*. Springer, 2nd edition, 2005.

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

[43] A. N. Shiryaev. *Optimal stopping rules*, volume 8 of *Stochastic Modelling and Applied Probability*. Springer-Verlag, Berlin, 2008. Translated from the 1976 Russian second edition by A. B. Aries, Reprint of the 1978 translation.

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

C. Cai: School of Mathematics, University of Leeds, Woodhouse Lane, LS2 9JT Leeds, UK.
*Email address*: mmcca@leeds.ac.uk

T. De Angelis: School of Management and Economics, Dept. ESOMAS, University of Turin, C.so Unione Sovietica 218bis, 10134, Turin, ITALY.
*Email address*: tiziano.deangelis@unito.it

J. Palczewski: School of Mathematics, University of Leeds, Woodhouse Lane, LS2 9JT Leeds, UK.
*Email address*: j.palczewski@leeds.ac.uk