ORIGINAL ARTICLE

Some results on optimal stopping under phase-type distributed implementation delay

Jukka Lempa[1]

Abstract

We study optimal stopping of strong Markov processes under random implementation delay. By random implementation delay we mean the following: the payoff is not realised immediately when the process is stopped but rather after a random waiting period. The distribution of the random waiting period is assumed to be phase-type. We prove first a general result on the solvability of the problem. Then we study the case of Coxian distribution both in general and with scalar diffusion dynamics in more detail. The study is concluded with two explicit examples.

Keywords Implementation delay · Time to build · Optimal stopping · Strong Markov process · Phase-type distributions · Markov chain · Diffusion · Resolvent operator

1 Introduction

Optimal stopping problems are widely used in economic and financial applications. One of the most notable application is the real options approach to investment planning, for a textbook treatment of the topic, see, e.g., Dixit and Pindyck (1994). Since the late 1970's, this approach has been used to study broad variety of economic planning problems, such as corporate strategy (Myers 1977; Ross 1978; Dixit and Pindyck 1995; Panayi and Trigeorgis 1998; Alvarez and Stenbacka 2001), land and real estate development (Grenadier 1996; Titman 1985; Williams 1993; Capozza and Li 1994), and natural resources investment (Brennan and Schwartz 1985; Paddock et al. 1988; Schwartz 1997). In many of these applications, the option mechanism in the real investment opportunity is the same as in financial options and can be studied by modeling it as an optimal stopping problem. The task is then to determine an investment rule which maximizes the expected present value of the total investment revenue. In the classical treatment of these problems, see, e.g., McDonald and Siegel (1986),

✉ Jukka Lempa
  jumile@utu.fi

[1] Department of Mathematics and Statistics, University of Turku, 20014 Turku, Finland

Dixit and Pindyck (1994), it is assumed that once the investment decision is made, the project starts to deliver cash flows immediately. In practice, this is often not the case as most capital projects take a significant time to complete, Bar-Ilan and Strange (1996), Kydland and Prescott (1982). This completion period is usually called time-to-build, implementation delay or delivery lag.

The effect of a implementation delay to optimal policy is broadly studied in the literature. The most straightforward way to model the implementation delay is to assume that it is simply a constant, see Aïd et al. (2015), Bar-Ilan and Strange (1996), Bar-Ilan and Sulem (1995). The paper (Bar-Ilan and Strange 1996) is concerned with a classical timing problem of irreversible investment problem in the style of Dixit (1989), whereas Aïd et al. (2015) analyses capacity expansion subject to time to build under the objective of social surplus maximisation. The paper (Bar-Ilan and Sulem 1995) analyzes a continuous-time inventory system with fixed delivery lag. In Alvarez and Keppo (2002) the delivery lag is modelled as a function of the state variable at the time of the investment. Here, a lump sum is paid when the investment is made and the investment revenue accrues once the delivery lag depending on the value of the state variable at the time of the investment has elapsed. Investment timing subject to inter-temporally controllable investment rate is studied in Majd and Pindyck (1987). More precisely, the investor is appraising the real option with respect to the uncertain future revenue and chooses whether to engage the investment. It is assumed that the investment rate is uniformly bounded, this causes the time to build. Now, the time to build is a random but endogenous in the sense that it is a result of the implemented investment policy. The effect of time to build for a levered firm is studied in Sarkar and Zhang (2015). It is assumed in this paper, in the style of Margsiri et al. (2008), that a fixed proportion of the investment cost is payed at time when the investment is engaged and the rest is paid once the implementation period is elapsed. The length of the period is determined by the revenue process, it is the time that it takes the process to reach the the level that is a fixed percentage higher than the state at which the investment is made. The implementation period is random and endogenous also in this case.

The implementation delay can also be modelled as an exogenous random variable. This approach is taken in Lempa (2012b), where the implementation delay is a random variable independent of the state process and exponentially distributed. This allows the usage of resolvent formalism which can be applied in a general Markov setting. The current study is an extension of Lempa (2012b). Indeed, we assume that the implementation delay is again independent but has now a general phase-type distribution. Phase-type distributions have been broadly applied in different fields such as survival analysis (Aalen 1995), healthcare systems modelling (Fackrell 2009), insurance applications (Bladt 2005), queuing theory (Breure and Baum 2005), and population genetics (Hobolth et al. 2018). These distributions are a class of matrix exponential distributions that have a Markovian realisation: they can be identified as the absorption times of certain continuous time Markov chains. We use this connection in our analysis as follows: once the state process is stopped, the exogenous continuous time Markov chain is initiated and the payoff paid out at the time of absorption of this chain depends on the value of the state process at that time. This analysis, which is new to our best knowledge, generalises significantly the results of Lempa (2012b).

Phase-type distributions offer a flexible and convenient model for random time lags. Indeed, it is well known, see, e.g., Breure and Baum (2005), that phase-type distributions are dense in the class of probability distributions on $\mathbb{R}_+$. Furthermore, there exists well-developed methodology for estimating the parameters of phase-type distributions, see, e.g., He (2014). As we will see, the Markov realisation of the phase-type distribution allows us to work the Markov theory of diffusion to derive closed form characterisations of the optimal solution. All of these aspects of the paper at hand add to the applicability of the results.

The remainder of the paper is organised as follows. In Sect. 2 we present the optimal stopping problem. In Sect. 3 our main result on the solvability of the problem is presented. We study the case of Coxian distribution in more detail in Sect. 4. The Coxian distribution case is solved in Sect. 5 with scalar diffusion dynamics. Section 6 wraps up the study with two explicit examples.

2 The optimal stopping problem

In this section we present the optimal stopping problem. Let $(\Omega, \mathcal{F}, \mathbb{F}, \mathbf{P})$ be a complete filtered probability space satisfying the usual conditions, where $\mathbb{F} = \{\mathcal{F}_t\}_{t \geq 0}$, see Borodin and Salminen (2015), p. 2. We assume that the underlying X is a strong Markov process defined on $(\Omega, \mathcal{F}, \mathbb{F}, \mathbf{P})$ and taking values in $E \subseteq \mathbf{R}^d$ for some $d \geq 1$ with the initial state $x \in E$. We take $E = (a_1, b_1) \times \cdots \times (a_d, b_d)$, where $-\infty \leq a_i < b_i \leq \infty$ for all $i = 1, \ldots, d$. As usual, we augment the state space E with a topologically isolated element Δ if the process X is non-conservative. Then the process X can be made conservative on the augmented state space $E^\Delta := E \cup \{\Delta\}$, see Borodin and Salminen (2015), p. 4. In what follows, we drop the superscript Δ from the notation. By convention, we augment the definition of functions g on E with $g(\Delta) = 0$.

Denote as $\mathbf{P}_x$ the probability measure $\mathbf{P}$ conditioned on the initial state x and as $\mathbf{E}_x$ the expectation with respect to $\mathbf{P}_x$. The process X is assumed to evolve under $\mathbf{P}_x$ and the sample paths are assumed to be right-continuous and left-continuous over stopping times meaning the following: if the sequence of stopping times $\tau_n \uparrow \tau$, then $X_{\tau_n} \to X_\tau$ $\mathbf{P}_x$-almost surely as $n \to \infty$. There is a well-established theory of standard optimal stopping for this class of processes, see Peskir and Shiryaev (2006).

For $r > 0$, we denote by L_1^r the class of real valued measurable functions f on E satisfying the integrability condition $\mathbf{E}_x \left[\int_0^\infty e^{-rt} |f(X_t)| \, dt \right] < \infty$ for all $x \in E$. For a function $f \in L_1^r$, the *resolvent* $R_r f : E \to \mathbf{R}$ is defined as $(R_r f)(x) = \mathbf{E}_x \left[\int_0^\infty e^{-rs} f(X_s) ds \right]$ for all $x \in E$. Denote p repeated applications of R_r to the function g as $(R_r^{(p)} g)$.

Under this probabilistic setting, we study the following optimal stopping problem. We want to find a stopping time τ^* which maximizes the expected discounted value of the payoff $\tau \mapsto g(X_{\tau + \zeta})$. Here, the function g is the payoff function and the variable ζ is a phase-type distributed random time - specific assumptions on g are made later. We define next phase-type distributions. Phase-type distributions are particular cases of matrix-exponential distributions, see, e.g., He (2014), which admit a Markovian

representation. To be more precise, let Y be a continuous time Markov chain defined on $(\Omega, \mathcal{F}, \mathbf{P}; \mathbf{F})$ and taking values values on the set $(0, 1, 2, \ldots, p)$. The states $1, \ldots, p$ are transient and the state 0 is absorbing. Then Y has an intensity matrix of the form

$$\Lambda = \begin{pmatrix} 0 & \mathbf{0} \\ \mathbf{t} & \mathbf{T} \end{pmatrix},$$

where $\mathbf{T}$ is a p-dimensional real square matrix (the *subgenerator* of Y), $\mathbf{t}$ is a p-dimensional column vector and $\mathbf{0}$ is a p-dimensional row vector of zeros. Since the intensities of rows must sum to zero, we find that $\mathbf{t} = -\mathbf{Te}$, where $\mathbf{e}$ is a column vector of 1's. Let $\boldsymbol{\pi} = (\pi_1, \ldots, \pi_p)$ denote the initial distribution of Y over the transient states only. Then we say that the time of absorption $\zeta = \inf\{t \geq 0 : X_t = 0\}$ has a phase-type distribution and write $\zeta \sim \mathrm{PH}(\boldsymbol{\pi}, \mathbf{T})$.

Denote as $\mathbf{P}_{x,\boldsymbol{\pi}}$ the probability measure $\mathbf{P}$ conditioned on the initial state x and initial distribution $\boldsymbol{\pi}$. The expectation with respect to $\mathbf{P}_{x,\boldsymbol{\pi}}$ in denoted as $\mathbf{E}_{x,\boldsymbol{\pi}}$. Now, the optimal stopping problem can be expressed as follows:

$$V(x) = \sup_{\tau} \mathbf{E}_{x,\boldsymbol{\pi}} \left[e^{-r(\tau+\zeta)} g(X_{\tau+\zeta}) \mathbf{1}_{\{\tau < \infty\}} \right], \tag{2.1}$$

where τ varies over $\mathbf{F}$-stopping times and r is the discount rate. We denote an optimal stopping time as τ^*. Probabilistically, the problem (2.1) can be interpreted as follows. At the initial time $t = 0$, we choose choose a stopping rule described by the stopping time τ. When τ is realized, the Markov chain Y is initiated from the distribution $\boldsymbol{\pi}$ and the payoff is realized when Y is absorbed. The payoff is thus uncertain and we can regard Y as an additional source of noise driving the payoff.

3 The main result

In this section, we prove our main result on the solvability of the optimal stopping problem (2.1). We make the following assumption on the structure of the matrix $\mathbf{T}$.

Assumption 3.1 The eigenvalues of subgenerator $\mathbf{T}$ are real and strictly negative

The Assumption 3.1 covers a variety of interesting cases of $\mathbf{T}$. For instance, the matrix $\mathbf{T}$ can be a triangular matrix with strictly negative diagonal entries. Thus it covers, for example, the following distributions of ζ:

1. Exponential and mixtures of exponentials with mutually distinct rates λ_i,
2. Hyperexponential and hypoexponential distribution,
3. Coxian distribution,
4. Erlang-distribution,

see, e.g., Stewart (2009). We prove first an auxiliary result.

Lemma 3.2 *Let* $g \in L_1^r$ *and* $m = 0, 1, 2, \ldots$. *Then*

$$\mathbf{E}_x \left[\int_0^\infty e^{-rs} s^m g(X_s) ds \right] = m! (R_r^{(m+1)} g)(x).$$

Proof We establish the case $m = 2$, the claim follows by induction. By changing the order of integration and invoking Chapman–Kolmogorov, we obtain

$$
\int_0^\infty e^{-rs} s \int_E g(y) p(x, y; s) ds
$$
$$
= \int_0^\infty e^{-rs} \int_0^s du \int_E g(y) p(x, y; s) ds
$$
$$
= \int_0^\infty \int_u^\infty e^{-rs} \int_E g(y) p(x, y; s) ds du
$$
$$
= \int_0^\infty \int_0^\infty e^{-r(v+u)} \int_E g(y) p(x, y; v+u) dv du
$$
$$
= \int_0^\infty \int_0^\infty e^{-r(v+u)} \int_E g(y) \int_E p(x, dz; u) p(z, dy; v) dv du
$$
$$
= \int_0^\infty e^{-ru} \int_E p(x, dz; u) \left(\int_0^\infty e^{-rv} \int_E g(y) p(z, dy; v) dv \right) du,
$$

which concludes the proof. $\qquad\square$

The next theorem is the main result of this section.

Theorem 3.3 *Let Assumption* 3.1 *hold. In addition, assume that the payoff* $g : E \to \mathbf{R}$ *is in* L_r^1, *satisfies the condition* $S^+ := \{x \; : \; g(x) > 0\} \neq \emptyset$ *and the process* X *reaches a point* $y_x \in S^+$ *with positive probability for all initial states* x. *Furthermore, assume that there exist an* r-*harmonic function* $h : E \to \mathbf{R}_+$ *such that the function* $x \mapsto \frac{g(x)}{h(x)}$ *is bounded. Then the value function* V *exists and can be identified as the least* r-*excessive majorant of the function* $\mathfrak{g}_\pi : x \mapsto \sum_{i,j,k} \pi_i (R_{r+\mu_j} g)(x) \alpha_{jk}^i t_k$ *for some coefficients* α, *where* $\boldsymbol{\pi} = (\pi_1, \dots, \pi_p)$, $\mathbf{t} = (t_1, \dots, t_p)$, *and the elements* $-\mu_j < 0$ *are the eigenvalues of* $\mathbf{T}$. *Furthermore, an optimal stopping time* τ_λ^* *exists and can be expressed as* $\tau_\lambda^* = \inf\{t \geq 0 \; : \; X_t \in \Gamma_\lambda^*\}$ *where* $\Gamma_\lambda^* = \{x \; : \; V(x) = \sum_{i,j,k} \pi_i (R_{r+\mu_j} g)(x) \alpha_{jk}^i t_k\}$.

Proof It is known from Bladt (2005), that the density f of ζ reads as $f(s) = \boldsymbol{\pi} e^{\mathbf{T}s} \mathbf{t}$, where $\mathbf{t} = -\mathbf{T}\mathbf{e}$. Denote the eigenvalues $\mathbf{T}$ as $-\mu_1, \dots, -\mu_k$, $k \leq p$, with corresponding multiplicities $a_1, \dots, a_k$, where $a_1 + \cdots + a_k = p$. There exists a linear change of coordinates $\mathbf{S}$ such that

$$
\mathbf{T} = \mathbf{S} \underbrace{\begin{pmatrix} \mathbf{A}_1 & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \mathbf{A}_2 & \cdots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \cdots & \cdots & \mathbf{A}_k \end{pmatrix}}_{:=\mathbf{A}} \mathbf{S}^{-1},
$$

see Teschl (2012), Ch. 3. Here, the diagonal blocks

$$\mathbf{A}_i = \begin{pmatrix} \mu_i & 1 & 0 & \cdots & 0 \\ 0 & \mu_i & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & \ddots & \mu_i & 1 \\ 0 & 0 & 0 & \cdots & \mu_i \end{pmatrix} = \mu_i \mathbf{I} + \mathbf{N}_i$$

are $a_i \times a_i$-matrices, and the matrix $\mathbf{N}_i$ satisfies $\mathbf{N}_i^{a_i} = \mathbf{0}$, for $i = 1, \ldots, k$. Furthermore, the matrix exponent $e^{\mathbf{T}s} = \mathbf{S}e^{\mathbf{A}s}\mathbf{S}^{-1}$. We readily verify that

$$e^{\mathbf{A}s} = \begin{pmatrix} e^{\mathbf{A}_1} & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & e^{\mathbf{A}_2} & \cdots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \cdots & \cdots & e^{\mathbf{A}_k} \end{pmatrix}, \text{ where } e^{\mathbf{A}_i} = e^{-\mu_i s} \sum_{m=0}^{a_i-1} \frac{\mathbf{N}_m^k s^m}{m!}.$$

Therefore the entries of the matrix $e^{\mathbf{T}s}$ are linear combinations of the functions $s \mapsto e^{-\mu_j s} s^m$, $j = 1, \ldots, k$ and $m = 0, \ldots, a_j - 1$, that is, $\left(e^{\mathbf{T}s}\right)_{in} = \sum_{j=1}^{k} \sum_{m=0}^{a_j-1} s^m + e^{-\mu_j s} \alpha_{jmn}^i$. Finally, the density f can be expressed as

$$f(s) = \sum_{i,n=1}^{p} \sum_{j=1}^{k} \sum_{m=0}^{a_j-1} \pi_i e^{-\mu_j s} s^m \alpha_{jmn}^i t_n. \tag{3.1}$$

Since X is strong Markov, the identity

$$\mathbf{E}_x \left[e^{-r(\tau+s)} g(X_{\tau+s}) \right] = \mathbf{E}_x \left[e^{-r\tau} \mathbf{E}_{X_\tau} \left[e^{-rs} g(X_s) \right] \right],$$

holds for all $s \geq 0$. By integrating this expression over the positive reals with respect to s with weight f given by (3.1) we obtain

$$\mathbf{E}_{x,\pi} \left[e^{-r(\tau+\zeta)} g(X_{\tau+\zeta}) \right]$$
$$= \mathbf{E}_x \left[e^{-r\tau} \mathbf{E}_{X_\tau,\pi} \left[\int_0^\infty e^{-rs} g(X_s) f(s) ds \right] \right]$$
$$= \mathbf{E}_x \left[e^{-r\tau} \sum_{i,n=1}^{p} \pi_i \sum_{j=1}^{k} \sum_{m=0}^{a_j-1} \mathbf{E}_{X_\tau} \left[\int_0^\infty s^m e^{-(r+\mu_j)s} g(X_s) ds \right] \alpha_{jmn}^i t_n \right].$$

By Lemma 3.2

$$\mathbf{E}_x \left[\int_0^\infty s^m e^{-(r+\mu_j)s} g(X_s) ds \right] = m! (R_{r+\mu_j}^{(m+1)} g)(x).$$

By lumping the factorial coefficients together with constants α and still denoting the resulting constants by α, we obtain

$$\mathbf{E}_{x,\pi}\left[e^{-r(\tau+\zeta)}g(X_{\tau+\zeta})\right] = \mathbf{E}_x\left[e^{-r\tau}\sum_{i,1=1}^{p}\pi_i\sum_{j=1}^{k}\sum_{m=1}^{a_j}(R_{r+\mu_j}^{(m)}g)(X_\tau)\alpha_{jmn}^i t_n\right].$$

Since the numbers μ_j are strictly negative, we can rewrite the resolvent $(R_{r+\mu_j}^{(m)}g)$ as

$$(R_{r+\mu_j}^{(m)}g)(x) = \mathbf{E}_x\left[\int_0^\infty e^{-\mu_j t}\frac{(R_{r+\mu_j}^{(m-1)}g)(X_t)}{h(X_t)}e^{-rt}\frac{h(X_t)}{h(x)}dt\right]h(x)$$

$$= \left(R_{\mu_j}^h\frac{(R_{r+\mu_j}^{(m-1)}g)}{h}\right)(x)h(x),$$

where R^h is the resolvent of the Doob's h-transform X^h, see Borodin and Salminen (2015), p. 34. Using this, we rewrite our optimal stopping problem as

$$V(x) = \sup_\tau \mathbf{E}_x^h\left[\sum_{i,n=1}^{p}\pi_i\sum_{j=1}^{k}\sum_{m=0}^{a_j-1}\frac{(R_{r+\mu_j}^{(m)}g)(X_\tau)}{h(X_\tau)}\alpha_{jmn}^i t_n\mathbf{1}_{\{\tau<\infty\}}\right]h(x)$$

$$= \sup_\tau \mathbf{E}_x^h\left[\sum_{i,n=1}^{p}\pi_i\sum_{j=1}^{k}\sum_{m=0}^{a_j-1}\left(R_{\mu_j}^h\frac{(R_{r+\mu_j}^{(m-1)}g)}{h}\right)(X_\tau)\alpha_{jk}^i t_k\mathbf{1}_{\{\tau<\infty\}}\right]h(x).$$

Since $-\mu_j < 0$ for all j and the family $(R_\lambda^h)_{\lambda>0}$ is a contraction resolvent, we find that

$$\left\|\left(R_{\mu_j}^h\frac{(R_{r+\mu_j}^{(m-1)}g)}{h}\right)\right\|_u \leq \frac{1}{\mu_j}\left\|\frac{(R_{r+\mu_j}^{(m-1)}g)}{h}\right\|_u$$

$$= \frac{1}{\mu_j}\left\|\left(R_{\mu_j}^h\frac{(R_{r+\mu_j}^{(m-2)}g)}{h}\right)\right\|_u \leq \cdots \leq \frac{1}{\mu_j^m}\left\|\frac{g}{h}\right\|_u < \infty.$$

where $\|\cdot\|_u$ denotes the sup-norm, i.e., $\|f\|_u = \sup_{x\in E}|f(x)|$. Thus the payoff $x \mapsto \sum_{i,n=1}^{p}\pi_i\left(R_{\mu_j}^h\frac{(R_{r+\mu_j}^{(m-1)}g)}{h}\right)(x)\alpha_{jk}^i t_k$ is uniformly bounded and, in particular, continuous. The claim follows now from Peskir and Shiryaev (2006), Thrm. I.2.7. □

Theorem 3.3 gives a weak set of conditions under which the optimal stopping problem (2.1) has a well-defined solution and an optimal stopping time exists. These conditions essentially mean that the optimal stopping region is not empty and that the

payoff function g cannot grow too fast, even though it can be unbounded. Note that by Theorem 3.3, we can rewrite the optimal stopping problem (2.1) as

$$V_{\boldsymbol{\pi}}(x) = \sup_{\tau} \mathbf{E}_x\left[e^{-r\tau}\mathfrak{g}_{\boldsymbol{\pi}}(X_\tau)\mathbf{1}_{\{\tau<\infty\}}\right]. \tag{3.2}$$

This is the form of the problem we will study in what follows. We point out that the payoff function $\mathfrak{g}_{\boldsymbol{\pi}}$ can be expressed as

$$\mathfrak{g}_{\boldsymbol{\pi}}(x) = \mathbf{E}_x\left[e^{-r\zeta}g(X_\zeta)\right], \tag{3.3}$$

where $\zeta \sim \mathrm{PH}(\boldsymbol{\pi}, \mathbf{T})$.

4 Case study I: Coxian distribution

The purpose of this section is to analyze the optimal stopping problem (3.2) in more detail when the absorption time has a Coxian distribution. More precisely, we compute a representation for the payoff function (3.3) that lends itself to explicit computations in the next section. To this end, assume that the number of transient states of Y is p. The process Y is then started from the state 1, therefore the initial distribution $\boldsymbol{\pi} = (1, 0, \ldots, 0)$. The subgenerator of Y is then written as

$$\mathbf{T} = \begin{pmatrix} -(\lambda_1 + \lambda_{12}) & \lambda_{12} & 0 & \cdots & 0 \\ 0 & -(\lambda_2 + \lambda_{23}) & \lambda_{23} & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & -\lambda_p \end{pmatrix}. \tag{4.1}$$

The quantities λ_i (resp. $\lambda_{i,i+1}$) are the absorption intensities (resp. transition intensities from state i to $i+1$). The time of absorption ζ has a now Coxian distribution, see, e.g., Bladt (2005). We remark that $\mathbf{t} = (\lambda_1, \ldots, \lambda_p)^{\mathsf{T}}$. Furthermore, the eigenvalues $-\mu_i = -(\lambda_i + \lambda_{i,i+1})$ for $i = 1, \ldots, p-1$, and $-\mu_p = -\lambda_p$.

We study the function

$$\mathfrak{g}_{\boldsymbol{\pi}} : x \mapsto \sum_{j,k}(R_{r+\mu_j}g)(x)\alpha_{jk}\lambda_k,$$

where the coefficients α_{jk} are implicitly defined in the proof of Theorem 3.3. To determine these constants, we first find that the matrix $\mathbf{S}$ of the eigenvectors of $\mathbf{T}$ reads as

$$S = \begin{pmatrix} 1 & \frac{\lambda_{12}}{\mu_1-\mu_2} & \cdots & \prod_{i=1}^{p-2}\frac{\lambda_{i,i+1}}{\mu_i-\mu_{p-1}} & \prod_{i=1}^{p-1}\frac{\lambda_{i,i+1}}{\mu_i-\mu_p} \\ 0 & 1 & \cdots & \prod_{i=2}^{p-2}\frac{\lambda_{i,i+1}}{\mu_i-\mu_{p-1}} & \prod_{i=2}^{p-1}\frac{\lambda_{i,i+1}}{\mu_i-\mu_p} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & \frac{\lambda_{p-1,p}}{\mu_{p-1}-\mu_p} \\ 0 & 0 & \cdots & 0 & 1 \end{pmatrix},$$

where the ith column is the eigenvector corresponding to the eigenvalue $-\mu_i$. Thus

$$Se^{Ds} = \begin{pmatrix} e^{-\mu_1 s} & \frac{\lambda_{12}}{\mu_1-\mu_2}e^{-\mu_2 s} & \cdots & \prod_{i=1}^{p-2}\frac{\lambda_{i,i+1}}{\mu_i-\mu_{p-1}}e^{-\mu_{p-1}s} & \prod_{i=1}^{p-1}\frac{\lambda_{i,i+1}}{\mu_i-\mu_p}e^{-\mu_p s} \\ 0 & e^{-\mu_2 s} & \cdots & \prod_{i=2}^{p-2}\frac{\lambda_{i,i+1}}{\mu_i-\mu_{p-1}}e^{-\mu_{p-1}s} & \prod_{i=2}^{p-1}\frac{\lambda_{i,i+1}}{\mu_i-\mu_p}e^{-\mu_p s} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & e^{-\mu_{p-1}s} & \frac{\lambda_{p-1,p}}{\mu_{p-1}-\mu_p}e^{-\mu_p s} \\ 0 & 0 & \cdots & 0 & e^{-\mu_p s} \end{pmatrix},$$

where $D = \operatorname{diag}(-\mu_1,\ldots,-\mu_p)$. Since S is upper unitriangular, the inverse S^{-1} is also upper unitriangular. More specifically, we readily verify that

$$S^{-1} = \begin{pmatrix} 1 & \frac{-\lambda_{12}}{\mu_1-\mu_2} & \cdots & \prod_{i=1}^{p-2}\frac{-\lambda_{i,i+1}}{\mu_1-\mu_{i+1}} & \prod_{i=1}^{p-1}\frac{-\lambda_{i,i+1}}{\mu_1-\mu_{i+1}} \\ 0 & 1 & \cdots & \prod_{i=2}^{p-2}\frac{-\lambda_{i,i+1}}{\mu_2-\mu_{i+1}} & \prod_{i=2}^{p-1}\frac{-\lambda_{i,i+1}}{\mu_2-\mu_{i+1}} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & \frac{-\lambda_{p-1,p}}{\mu_{p-1}-\mu_p} \\ 0 & 0 & \cdots & 0 & 1 \end{pmatrix}.$$

Therefore the matrix exponential

$$e^{Ts} = Se^{Ds}S^{-1} = \begin{pmatrix} e^{-\mu_1 s} & b_{12}(s) & \cdots & b_{1,p-1}(s) & b_{1p}(s) \\ 0 & e^{-\mu_2 s} & \cdots & b_{2,p-1}(s) & b_{2p}(s) \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & e^{-\mu_{p-1}s} & b_{p-1,p}(s) \\ 0 & 0 & \cdots & 0 & e^{-\mu_p s} \end{pmatrix},$$

where the element

$$b_{mn}(s) = \prod_{i=m}^{n-1}\lambda_{i,i+1}\sum_{j=m}^{n}\frac{e^{-\mu_j s}}{\prod_{\substack{k=m \\ k\neq j}}^{n}(\mu_k-\mu_j)}, \quad m < n,$$

for $m = 2, \ldots, p$. By substituting this to the definition of the density f, we find

$$f(s) = \pi e^{\mathbf{T}s}\mathbf{t} = e^{-\mu_1 s}\lambda_1 + \sum_{j=2}^{p} b_{1j}(s)\lambda_j$$

$$= e^{-\mu_1 s}\lambda_1 + \sum_{k=2}^{p}\left(\prod_{i=1}^{k-1}\lambda_{i,i+1}\sum_{j=1}^{k}\frac{e^{-\mu_j s}}{\prod_{\substack{l=1\\l\neq j}}^{k}(\mu_l - \mu_j)}\right)\lambda_k.$$

From this expression, we obtain the coefficients α:

$$\alpha_{jk} = \prod_{i=1}^{k-1}\lambda_{i,i+1}\left(\prod_{\substack{l=1\\l\neq j}}^{k}(\mu_l - \mu_j)\right)^{-1}, \quad j = 1, \ldots, p, \ k \geq \max(j, 2),$$

$$\alpha_{11} = 1. \tag{4.2}$$

Having the coefficients α at our disposal, we proceed with the derivation of the payoff function $\mathfrak{g}_\pi$. Let $\mathfrak{A} = (\alpha_{jk})$ and denote the vector of resolvents $(R_{r+\mu_j}g)$, $j = 1, \ldots, p$, as $\mathbf{r}$. Then the analysis above implies that

$$\mathfrak{g}_\pi(x) = \mathbf{r}^{\mathsf{T}}\mathfrak{A}\mathbf{t}(x)$$

$$= \lambda_1(R_{r+\mu_1}g)(x) + \sum_{k=2}^{p}\left(\prod_{i=1}^{k-1}\lambda_{i,i+1}\sum_{j=1}^{k}\frac{(R_{r+\mu_j}g)(x)}{\prod_{\substack{l=1\\l\neq j}}^{k}(\mu_l - \mu_j)}\right)\lambda_k. \tag{4.3}$$

for all $x \in E$. We use the following lemma; this result can be regarded as a generalization of the resolvent equation.

Lemma 4.1 *For each $k = 2, \ldots, p$, the following holds:*

$$\sum_{j=1}^{k}\frac{(R_{r+\mu_j}g)(x)}{\prod_{\substack{l=1\\l\neq j}}^{k}(\mu_l - \mu_j)} = (R_{r+\mu_1}R_{r+\mu_2}\cdots R_{r+\mu_k}g)(x),$$

for all $x \in E$.

Proof For all $j = 1, \ldots, k$, let $U_j \sim \mathrm{Exp}(\mu_j)$. Let $S_k = \sum_{j=1}^{k}U_j$. Since the elements μ_j are distinct, we find by using the formula for the distribution of S_k from Ross (2010), p. 309 that

$$\mathbf{E}_x \left[e^{-r S_k} g(X_{S_k}) \right] = \mathbf{E}_x \left[\int_0^\infty e^{-rt} g(X_t) \prod_{i=1}^k \mu_i \sum_{j=1}^k \frac{e^{-\mu_j t}}{\prod_{\substack{l=1 \\ l \neq j}}^k (\mu_l - \mu_j)} dt \right]$$

$$= \prod_{i=1}^k \mu_i \sum_{j=1}^k \frac{(R_{r+\mu_j} g)(x)}{\prod_{\substack{l=1 \\ l \neq j}}^k (\mu_l - \mu_j)}.$$

On the other hand, since X is strong Markov, we find that

$$\mathbf{E}_x \left[e^{-r S_k} g(X_{S_k}) \right] = \mathbf{E}_x \left[e^{-r S_{k-1}} \mathbf{E}_{X_{S_{k-1}}} \left[e^{-r U_k} g(X_{U_k}) \right] \right]$$

$$= \mathbf{E}_x \left[e^{-r S_{k-1}} \mu_k (R_{r+\mu_k} g)(X_{S_{k-1}}) \right]$$

$$= \mathbf{E}_x \left[e^{-r S_{k-2}} \mu_k \mu_{k-1} (R_{r+\mu_{k-1}} R_{r+\mu_k} g)(X_{S_{k-2}}) \right]$$

$$= \ldots$$

$$= \prod_{i=1}^k \mu_i (R_{r+\mu_1} R_{r+\mu_2} \cdots R_{r+\mu_k} g)(x),$$

proving the claim. □

We can now rewrite the expression (4.3) using Lemma 4.1 to obtain the following result.

Proposition 4.2 *Assume that subgenerator* $\mathbf{T}$ *of* Y *is given by* (4.1). *Then*

$$\mathfrak{g}_\pi(x) = \lambda_1 (R_{r+\mu_1} g)(x) + \sum_{k=2}^p \prod_{i=1}^{k-1} \lambda_{i,i+1} \lambda_k (R_{r+\mu_1} R_{r+\mu_2} \cdots R_{r+\mu_k} g)(x), \quad (4.4)$$

for all $x \in E$.

The expression (4.4) has a natural interpretation. Denote the time of absorption from state i as A_i and time of transition from state i to state $i+1$ as C_i. For each k, define the random time $T_k = \sum_{i=1}^{k-1} C_i + A_k$ on the set where the process Y is absorbed from state k - for $k = 1$, obviously $T_1 = A_1$. Then

$$\mathfrak{g}_\pi(x) = \frac{\lambda_1}{\mu_1} \mu_1 (R_{r+\mu_1} g)(x)$$

$$+ \sum_{k=2}^p \prod_{i=1}^{k-1} \frac{\lambda_{i,i+1}}{\mu_i} \frac{\lambda_k}{\mu_k} (\mu_1 R_{r+\mu_1} \mu_2 R_{r+\mu_2} \cdots \mu_k R_{r+\mu_k} g)(x)$$

$$= \sum_{k=1}^p \mathbf{E}_x \left[e^{-r T_k} g(X_{T_k}) \right] \prod_{i=1}^{k-1} \mathbf{P}(C_i < A_i) \mathbf{P}(C_k > A_k).$$

This representation of the value is in line with the law of total probability. Indeed, the payoff is obtained as the sum over all paths of the process Y where the payoff corresponding to each realization is expected present value of the variable $g(X)$ sampled at the time of absorption T_k determined by this particular realization. The times T_k are sums of mutually independent exponentially distributed random times conditional on a particular realization.

5 Case study II: Coxian distribution with scalar diffusion dynamics

In this section, we elaborate the results of the previous section in the case where the process X follows a scalar diffusion. More precisely, we assume that the state process X evolves on $\mathbf{R}_+$ and follows the regular linear diffusion given as the weakly unique solution of the Itô equation $dX_t = \mu(X_t)dt + \sigma(X_t)dW_t$, $X_0 = x$. Here, W is a Wiener process on $(\Omega, \mathcal{F}, \mathbb{F}, \mathbf{P})$ and the real valued functions μ and $\sigma > 0$ are assumed to be continuous. Using the terminology of Borodin and Salminen (2015), the boundaries 0 and ∞ are either natural, entrance-not-exit, exit-not-entrance or non-singular. In the case a boundary is non-singular, it is assumed to be either killing or instantaneously reflecting, see Borodin and Salminen (2015), pp. 18–20. As usually, we denote as $\mathcal{A} = \frac{1}{2}\sigma^2(x)\frac{d^2}{dx^2} + \mu(x)\frac{d}{dx}$ the second order linear differential operator associated to X. Furthermore, we denote as, respectively, $\psi_r > 0$ and $\varphi_r > 0$ the increasing and decreasing solution of the ODE $\mathcal{A}u = ru$, where $r > 0$, defined on the domain of the characteristic operator of X. By posing appropriate boundary conditions depending on the boundary classification of the diffusion X, the functions ψ_r and φ_r are defined uniquely up to a multiplicative constant and can be identified as the minimal r-excessive functions, see (Borodin and Salminen 2015, pp. 18 – 20). Finally, we define the speed measure m and the scale function S of X via the formulæ $m'(x) = \frac{2}{\sigma^2(x)}e^{B(x)}$ and $S'(x) = e^{-B(x)}$ for all $x \in \mathbf{R}_+$, where $B(x) := \int^x \frac{2\mu(y)}{\sigma^2(y)}dy$, see Borodin and Salminen (2015), p. 17.

We know from the literature that for a given $f \in L_1^r$ the resolvent $R_r f$ can be expressed as

$$
(R_r f)(x) = B_r^{-1}\varphi_r(x)\int_0^x \psi_r(y)f(y)m'(y)dy
$$
$$
+ B_r^{-1}\psi_r(x)\int_x^\infty \varphi_r(y)f(y)m'(y)dy, \tag{5.1}
$$

for all $x \in \mathbf{R}_+$, where $B_r = \frac{\psi_r'(x)}{S'(x)}\varphi_r(x) - \frac{\varphi_r'(x)}{S'(x)}\psi_r(x)$ denotes the Wronskian determinant, see Borodin and Salminen (2015), p. 19.

We consider the optimal stopping problem

$$
V_{\boldsymbol{\pi}}(x) = \sup_\tau \mathbf{E}_x\left[e^{-r\tau}\mathfrak{g}_{\boldsymbol{\pi}}(X_\tau)\mathbf{1}_{\{\tau<\infty\}}\right]. \tag{5.2}
$$

where the payoff

$$\mathfrak{g}_\pi(x) = \frac{\lambda_1}{\mu_1} \mu_1 (R_{r+\mu_1} g)(x)$$

$$+ \sum_{k=2}^{p} \prod_{i=1}^{k-1} \frac{\lambda_{i,i+1}}{\mu_i} \frac{\lambda_k}{\mu_k} (\mu_1 R_{r+\mu_1} \mu_2 R_{r+\mu_2} \cdots \mu_k R_{r+\mu_k} g)(x). \tag{5.3}$$

The following proposition is the main result of this section.

Proposition 5.1 *Let the assumptions of Theorem 3.3 be met. In addition, assume that*

(A) *the function g is stochastically C^2, that is, continuous and twice continuously differentiable outside of a countable set D that has no accumulation points,*
(B) *the function $\frac{g(x)}{\psi_r(x)}$ attains a maximum at an interior point $\hat{x}$,*
(C) *the function $(\mathcal{A} - r)g$ is non-increasing,*
(D) *the limiting condition $(\mathcal{A}-r)(R_{r+\mu}g)(0+) := \lim_{x\to 0+}(\mathcal{A}-r)(R_{r+\mu}g)(x) > 0$ holds for all $\mu > \min\{\mu_i\}$, where μ_i are the eigenvalues of the subgenerator $\mathbf{T}$.*

Then there is a unique x^* which maximizes the function $x \mapsto \frac{\mathfrak{g}_\pi(x)}{\psi_r(x)}$. The state x^* is the optimal stopping threshold for the optimal stopping problem (5.2) and the value function V_π can be expressed as

$$V_\pi(x) = \begin{cases} \mathfrak{g}_\pi(x), & x \geq x^*, \\ \frac{\mathfrak{g}_\pi(x^*)}{\psi_r(x^*)} \psi_r(x), & x \leq x^*. \end{cases} \tag{5.4}$$

Remark 5.2 By assumption (A), we have

$$\frac{d}{dx}\left(\frac{g(x)}{\psi_r(x)}\right) = \frac{S'(x)}{\psi_r^2(x)} \int_0^x \psi_r(z)(\mathcal{A}-r)g(z)m'(z)dz.$$

Thus, by assumption (B), a threshold $x_0 < \hat{x}$ such that $(\mathcal{A}-r)g(x) \gtreqless 0$, when $x \lesseqgtr x_0$.

Example 5.3 We present a set of suffients conditions for the assumptions (A)–(D) of Proposition 5.1. Assume that the function g is

(1) non-negative and non-decreasing with $g(0+) = 0$,
(2) piecewise linear with a finite number corner points.

These assumptions cover various option-like payoffs such as $g(x) = (x - K)^+$. Now the assumption (A) is clearly satisfied. The function ψ_r is known to be increasing. If it is furthermore assumed to convex, as it is for instance for GBM (see Alvarez (2003) for general conditions for convexity of ψ_r), then the assumption (B) is also satisfied. With regard to assumption (C), we find by assuming sufficient regularity that

$$l(x) = (\mathcal{A} - r)g(x) = \mu(x)g'(x) - rg(x) = \mu(x)c_k - rg(x),$$

for some constant $c_k > 0$. Thus $l'(x) = c_k(\mu'(x) - r) < 0$ whenever $\mu'(x) < r$ for all x, that is, the growth rate of the drift function must be uniformly bounded by the

rate of discounting. This condition is not very severe and is satisfied for reasonable parameter configurations by GBM and various mean-reverting diffusions such as Cox–Ingersoll–Ross and Verhulst–Pearl processes. Finally, since $(\mathcal{A} - r)(R_{r+\mu}g)(x) = \mu(R_{r+\mu}g)(x) - g(x)$, we find that the condition of assumption (D) is satisfied in this case.

An even simpler payoff structure $g(x) = x - K$ is also covered by (A)-(D). In comparison to (1) and (2), $g(0+)$ is now negative and we cannot argue as above for (D) to hold. Assume now, that the function μ is non-negative at origin. Then $(\mathcal{A} - (r + \mu))g(0+) = \mu(0+) - (r + \mu)g(0+) > 0$ for all $\mu > 0$ and, consequently, the assumption (D) holds.

Remark 5.4 Proposition 5.1 gives sufficient conditions for the existence of a one-sided optimal stopping rule. Analogous conditions that would result in two-sided rules could most likely be provided. Indeed, as Remark 5.2 points out, the function g is r-subharmonic for on $(0, x_0)$. On the other hand, if we assume that g is r-subharmonic on an interval (a, b), where $0 < a < b < \infty$, then we would be likely to work out a set of assumptions such the resulting optimal continuation region is (z^*, y^*), where $0 < z^* < y^* < \infty$. These assumptions would most likely include boundedness and monotonicity assumptions of the functions $\frac{g}{\psi_r}$ and $\frac{g}{\varphi_r}$, see Lempa (2010). However, this generalization is beyond the scope of this paper.

Lemma 5.5 *Let the assumptions of Proposition 5.1 hold. Then, for all $\mu > 0$, the the function $(\mathcal{A} - r)(R_{r+\mu}g)$ is decreasing.*

Proof Since $(R_{r+\mu}(\mathcal{A} - (r + \mu))g) = -g = (\mathcal{A} - (r + \mu))(R_{r+\mu}g))$, we find that $(\mathcal{A} - r)(R_{r+\mu}g) = (R_{r+\mu}(\mathcal{A} - r)g)$. Now, ordinary differentiation yields

$$
B_{r+\mu}(R_{r+\mu}(\mathcal{A} - r)g)'(x) = \varphi'_{r+\mu}(x) \int_0^x \psi_{r+\mu}(z)(\mathcal{A} - r)g(z)m'(z)dz
$$
$$
+ \psi'_{r+\mu}(x) \int_x^\infty \varphi_{r+\mu}(z)(\mathcal{A} - r)g(z)m'(z)dz. \quad (5.5)
$$

We observe from (5.5) that this derivative is negative at x_0. Let $x < x_0$. Then we can rewrite (5.5) as

$$
B_{r+\mu}(R_{r+\mu}(\mathcal{A} - r)g)'(x)
$$
$$
= \varphi'_{r+\mu}(x) \int_0^x \psi_{r+\mu}(z)(\mathcal{A} - r)g(z)m'(z)dz + \psi'_{r+\mu}(x)
$$
$$
\left(\int_x^{x_0} \varphi_{r+\mu}(z)(\mathcal{A} - r)g(z)m'(z)dz + \int_{x_0}^\infty \varphi_{r+\mu}(z)(\mathcal{A} - r)g(z)m'(z)dz \right).
$$

The assumed boundary classification of 0 implies that

$$\varphi'_{r+\mu}(x) \int_0^x \psi_{r+\mu}(z)(\mathcal{A} - r)g(z)m'(z)dz$$

$$+ \psi'_{r+\mu}(x) \int_x^{x_0} \psi_{r+\mu}(z)(\mathcal{A} - r)g(z)m'(z)dz$$

$$< (\mathcal{A} - r)g(x) \left(\varphi'_{r+\mu}(x) \int_0^x \psi_{r+\mu}(z)m'(z)dz + \psi'_{r+\mu}(x) \int_x^{x_0} \varphi_{r+\mu}(z)m'(z)dz \right)$$

$$= \frac{(\mathcal{A} - r)g(x)}{r + \mu} \left(\psi'_{r+\mu}(x) \frac{\varphi'_{r+\mu}(x_0)}{S'(x_0)} - \varphi'_{r+\mu}(x) \lim_{x \to 0+} \frac{\psi'_{r+\mu}(x)}{S'(x)} \right) < 0,$$

proving that the derivative (5.5) is negative on $(0, x_0)$. The interval (x_0, ∞) is treated similarly. □

Lemma 5.6 *Let the assumptions of Proposition* 5.1 *hold. Then, for all $\mu > 0$, the function $x \mapsto \frac{(R_{r+\mu}g)(x)}{\psi_r(x)}$ is decreasing for all $x \geq \hat{x}$.*

Proof Following Remark 5.2, we find by changing the order of integration that

$$\frac{\psi_r^2(x)B_{r+\mu}}{S'(x)} \frac{d}{dx} \left(\frac{(R_{r+\mu}g)(x)}{\psi_r(x)} \right)$$

$$= B_{r+\mu} \int_0^x \psi_r(z)(\mathcal{A} - r)(R_{r+\mu}g)(z)m'(z)dz$$

$$= \int_0^x \psi_r(y)\varphi_{r+\mu}(y) \int_0^y \psi_{r+\mu}(z)(\mathcal{A} - r)g(z)m'(z)dym'(y)dy$$

$$+ \int_0^x \psi_r(y)\psi_{r+\mu}(y) \int_y^x \varphi_{r+\mu}(z)(\mathcal{A} - r)g(z)m'(z)dym'(y)dy$$

$$= \int_0^x \left(\psi_{r+\mu}(z) \int_z^x \varphi_{r+\mu}(y)\psi_r(y)m'(y)dy \right)(\mathcal{A} - r)g(z)m'(z)dz$$

$$+ \int_x^\infty \varphi_{r+\mu}(z) \int_0^x \psi_{r+\mu}(y)\psi_r(y)m'(y)dy(\mathcal{A} - r)g(z)m'(z)dz. \qquad (5.6)$$

By the virtue of Lemma 5.5, it is enough to show that the differential (5.6) is negative at $\hat{x}$. To this end, define the function $\check{\psi}_r^x(y) = \psi_r(y)\mathbf{1}_{\{y \leq x\}}$. Since the upper boundary ∞ is natural for the diffusion X, it follows from Lempa (2012a), Lemma 2.1 and the fact that the function ψ_r is non-negative that $\mu(R_{r+\mu}\check{\psi}_r^x)(y) \leq \mu(R_{r+\mu}\psi_r)(y) = \psi_r(y)$, for all $y \in (0, \infty)$. By evaluating the right hand side at the point $\hat{x}$, we then find that

$$\int_0^{\hat{x}} \left(\psi_{r+\mu}(z) \int_z^{\hat{x}} \varphi_{r+\mu}(y)\psi_r(y)m'(y)dy \right)(\mathcal{A} - r)g(z)m'(z)dz$$

$$+ \int_{\hat{x}}^\infty \varphi_{r+\mu}(z) \int_0^{\hat{x}} \psi_{r+\mu}(y)\psi_r(y)m'(y)dy(\mathcal{A} - r)g(z)m'(z)dz$$

$$< B_{r+\mu} \int_0^{\hat{x}} \psi_r(z)(\mathcal{A} - r)g(z)m'(z)dz$$

$$+ \int_{\hat{x}}^\infty \varphi_{r+\mu}(z) \int_0^{\hat{x}} \psi_{r+\mu}(y)\psi_r(y)m'(y)dy(\mathcal{A} - r)g(z)m'(z)dz < 0.$$

This proves the claim.   □

Proof of Proposition 5.1 Let $k \in \{2, \ldots, p\}$ and consider the function

$$x \mapsto \prod_{i=1}^{k-1} \frac{\lambda_{i,i+1}}{\mu_i} \frac{\lambda_k}{\mu_k} (\mu_1 R_{r+\mu_1} \mu_2 R_{r+\mu_2} \cdots \mu_k R_{r+\mu_k} g)(x).$$

We start by analyzing the properties of $\mu_k(R_{r+\mu_k}g)$ and show that it satisties the same assumptions as g. Under the assumption (A), the function $\mu_k(R_{r+\mu_k}g)$ is obviously stochastically C^2. Lemmas (5.5) and (5.6) coupled with the assumption (D) imply that $\mu_k(R_{r+\mu_k}g)$ satisfies also the assumptions (B) and (C). To see that the condition (D) is also satisfied, assume that parameters $\eta_1, \eta_2 > \min\{\mu_i\}$. Without loss of generality, we cas assume that $\eta_1 > \eta_2$. Then the resolvent equation implies that

$$(\mathcal{A} - r)(R_{r+\eta_1} R_{r+\eta_2} g)(0+)$$
$$= \frac{1}{\eta_1 - \eta_2} \left((R_{r+\eta_2}(\mathcal{A} - r)g(0+) - (R_{r+\eta_1}(\mathcal{A} - r)g(0+) \right) > 0.$$

We can now use the same procedure iteratively through the entire sum in (5.3) to conclude that the function $\mathfrak{g}_\pi$ satisfies also the same assumptions as g. The claim follows now from Alvarez (2001), Thrm. 3.   □

Remark 5.7 We observe from the proof of Proposition 5.1, in particular, from Lemma 5.6, that the optimal stopping threshold x^* is dominated by the state $\hat{x}$. This state is the optimal stopping threshold for the same problem in absence of the implementation delay. In other words, we observe that the introduction of the exercise lag accelerates the optimal exercise of the option to stop.

Remark 5.8 We observe from the proof of Proposition 5.1 also that we can modify the function (5.3) and allow for different payoffs for absorption from different phases. To elaborate, say that we have a collection of functions $(g_k)_{k=1}^p$ which all satisfy the assumptions of Proposition 5.1. Then we modify (5.3) such that if the absorption of Y occurs from the state i, the resulting payoff is given by the function g_i, that is,

$$\mathfrak{g}_\pi(x) = \frac{\lambda_1}{\mu_1} \mu_1 (R_{r+\mu_1} g_1)(x)$$

$$+ \sum_{k=2}^p \prod_{i=1}^{k-1} \frac{\lambda_{i,i+1}}{\mu_i} \frac{\lambda_k}{\mu_k} (\mu_1 R_{r+\mu_1} \mu_2 R_{r+\mu_2} \cdots \mu_k R_{r+\mu_k} g_k)(x).$$

Then we can do a similar analysis as in the proof of Proposition 5.1 to conclude that the optimal stopping threshold is given the global maximum of the function $x \mapsto \frac{g_\pi(x)}{\psi_r(x)}$.

To see why this is interesting, consider the following example. In real option applications, the investor often pays a lump sum cost K either when the investment option is exercised or when project is completed. The basic form of the exercise payoff is $x \mapsto \mathbf{E}_x\left[e^{-r\zeta} g(X_\zeta)\right]$, where the function g can, for instance, be $x \mapsto x - K$. Now the lump sum is paid at the completion time. To illustrate how to shift this payment to the start of the project, consider the following. We assume for brevity that the lag variable ζ has a two-phase Cox distribution. Furthermore, denote the constant function $x \mapsto 1$ as $\mathbb{1}$. Then

$$
\begin{aligned}
&\mathbf{E}_x\left[e^{-r\zeta} g(X_\zeta)\right] - K \\
&= \frac{\lambda_1}{\mu_1} \mu_1 (R_{r+\mu_1} g)(x) + \frac{\lambda_{12}}{\mu_1} \mu_1 (R_{r+\mu_1} \lambda_2 R_{r+\lambda_2} g)(x) - K \\
&= \frac{\lambda_1}{\mu_1} \left(\mu_1 (R_{r+\mu_1} g)(x) - K \right) + \frac{\lambda_{12}}{\mu_1} \left(\mu_1 (R_{r+\mu_1} \lambda_2 R_{r+\lambda_2} g)(x) - K \right) \\
&= \frac{\lambda_1}{\mu_1} \left(\mu_1 (R_{r+\mu_1} g)(x) - \mu_1 \left(R_{r+\mu_1} ((r + \mu_1) - \mathcal{A}) \frac{K}{\mu_1} \mathbb{1} \right)(x) \right) \\
&\quad + \frac{\lambda_{12}}{\mu_1} \left(\mu_1 (R_{r+\mu_1} \lambda_2 R_{r+\lambda_2} g)(x) - \mu_1 \left(R_{r+\mu_1} ((r + \mu_1) - \mathcal{A}) \frac{K}{\mu_1} \mathbb{1} \right)(x) \right) \\
&= \frac{\lambda_1}{\mu_1} \left(\mu_1 \left(R_{r+\mu_1} \left(g - \frac{K(r + \mu_1)}{\mu_1} \mathbb{1} \right) \right)(x) \right) \\
&\quad + \frac{\lambda_{12}}{\mu_1} \left(\mu_1 \left(R_{r+\mu_1} \left(\lambda_2 (R_{r+\lambda_2} g) - \frac{K(r + \mu_1)}{\mu_1} \mathbb{1} \right) \right)(x) \right) \\
&= \frac{\lambda_1}{\mu_1} \left(\mu_1 \left(R_{r+\mu_1} \left(g - \frac{K(r + \mu_1)}{\mu_1} \mathbb{1} \right) \right)(x) \right) \\
&\quad + \frac{\lambda_{12}}{\mu_1} \left(\mu_1 \left(R_{r+\mu_1} \lambda_2 \left(R_{r+\lambda_2} \left(g - \frac{K(r + \mu_1)(r + \lambda_2)}{\mu_1 \lambda_2} \mathbb{1} \right) \right) \right)(x) \right), \quad (5.7)
\end{aligned}
$$

for a sufficiently nice function g. Define the functions

$$
g_1(x) = g(x) - \frac{K(r + \mu_1)}{\mu_1},
$$
$$
g_2(x) = g(x) - \frac{K(r + \mu_1)(r + \lambda_2)}{\mu_1 \lambda_2}.
$$

If the functions $x \mapsto \frac{g_i(x)}{\psi_r(x)}$ satisfy the assumptions of Proposition 5.1, the conclusion of 5.1 holds for the payoff (5.7). This payoff corresponds to the case where the lump sum is paid at time when the project is initiated.

6 Two examples

The purpose of this section is to illustrate the previous results with explicit examples. We assume throughout the section that the process Y has the subgenerator

$$\mathbf{T} = \begin{pmatrix} -(\lambda_1 + \lambda_{12}) & \lambda_{12} \\ 0 & -\lambda_2 \end{pmatrix}.$$

This implies that the time of absorption ζ has a two-phase Coxian distribution.

6.1 Geometric Brownian motion

Assume that the process X is given by a geometric Brownian motion, that is, the solution of the stochastic differential equation $dX_t = \mu X_t dt + \sigma X_t dW_t$. Here, W is a Wiener process and the constants μ and $\sigma > 0$ satisfy the conditions $\mu < r$ and $\mu - \frac{1}{2}\sigma^2 > 0$. Then the optimal stopping time is almost surely finite. The scale density S' reads as $S'(x) = x^{-\frac{2\mu}{\sigma^2}}$ and the speed density m' as $m'(x) = \frac{2}{\sigma^2 x^2} x^{\frac{2\mu}{\sigma^2}}$. The differential operator $\mathcal{A} = \frac{1}{2}\sigma^2 x^2 \frac{d^2}{dx^2} + \mu x \frac{d}{dx}$ and the functions $\psi.$ and $\varphi.$ can be written as $\psi_{r+\lambda}(x) = x^{\beta_{r+\lambda}}$, $\varphi_{r+\lambda}(x) = x^{\alpha_{r+\lambda}}$, where the constants

$$\begin{cases} \beta_{r+\lambda} = \left(\frac{1}{2} - \frac{\mu}{\sigma^2}\right) + \sqrt{\left(\frac{1}{2} - \frac{\mu}{\sigma^2}\right)^2 + \frac{2(r+\lambda)}{\sigma^2}} > 1, \\ \alpha_{r+\lambda} = \left(\frac{1}{2} - \frac{\mu}{\sigma^2}\right) - \sqrt{\left(\frac{1}{2} - \frac{\mu}{\sigma^2}\right)^2 + \frac{2(r+\lambda)}{\sigma^2}} < 0. \end{cases}$$

It is a simple computation to show that the Wronskian $B_{r+\lambda} = 2 \sqrt{\left(\frac{1}{2} - \frac{\mu}{\sigma^2}\right)^2 + \frac{2(r+\lambda)}{\sigma^2}}$.

6.1.1 A problem with an explicit solution

We follow Remark 5.8 and consider the exercise payoff $\mathfrak{g}_\pi : x \mapsto \mathbf{E}_x\left[e^{-r\zeta}X_\zeta\right] - K$, where $K > 0$. Using the identity (5.7), we obtain

$$\begin{aligned}
\mathbf{E}_x &\left[e^{-r\zeta}X_\zeta\right] - K \\
&= \frac{\lambda_1}{\mu_1}\left(\mu_1\left(R_{r+\mu_1}\left(\mathrm{id} - \frac{K(r+\mu_1)}{\mu_1}\mathbb{1}\right)\right)(x)\right) \\
&\quad + \frac{\lambda_{12}}{\mu_1}\left(\mu_1\left(R_{r+\mu_1}\lambda_2\left(R_{r+\lambda_2}\left(\mathrm{id} - \frac{K(r+\mu_1)(r+\lambda_2)}{\mu_1\lambda_2}\mathbb{1}\right)\right)\right)(x)\right),
\end{aligned}$$

where id is the identity function. Let

$$g_1(x) = x - \frac{K(r + \mu_1)}{\mu_1},$$

$$g_2(x) = x - \frac{K(r + \mu_1)(r + \lambda_2)}{\mu_1 \lambda_2}.$$

We readily verify that the functions g_i satisfy the assumptions of Proposition 5.1. Next, we study the function $x \mapsto \frac{g_\pi(x)}{\psi_r(x)}$. Since

$$\mathbf{E}_x \left[e^{-r\zeta} X_\zeta \right] = \frac{\lambda_1}{\mu_1} \mu_1 (R_{r+\mu_1} \text{ id})(x) + \frac{\lambda_{12}}{\mu_1} \mu_1 (R_{r+\mu_1} \lambda_2 R_{r+\lambda_2} \text{ id})(x),$$

we start by computing the required resolvents. First, we find that

$$\mu_1(R_{r+\mu_1} \text{ id})(x) = \frac{2\mu_1}{B_{r+\mu_1} \sigma^2} \left(x^{\alpha_{r+\mu_1}} \int_0^x y^{-\alpha_{r+\mu_1}} dy + x^{\beta_{r+\mu_1}} \int_x^\infty y^{-\beta_{r+\mu_1}} dy \right)$$

$$= \frac{2\mu_1 (\beta_{r+\mu_1} - \alpha_{r+\mu_1})x}{B_{r+\mu_1} \sigma^2 (1 - \alpha_{r+\mu_1})(\beta_{r+\mu_1} - 1)} = \frac{\mu_1 x}{r + \mu_1 - \mu}$$

This yields

$$\mu_1(R_{r+\mu_1} \lambda_2 R_{r+\lambda_2} \text{ id})(x) = \frac{\mu_1 \lambda_2 x}{(r + \mu_1 - \mu)(r + \lambda_2 - \mu)}.$$

We can then write the expectation

$$\mathbf{E}_x \left[e^{-r\zeta} X_\zeta \right] = \underbrace{\frac{1}{r - \mu + \mu_1} \left(\lambda_1 + \frac{\lambda_{12} \lambda_2}{r - \mu + \lambda_2} \right)}_{:=C} x, \qquad (6.1)$$

and, consequently, express the exercise payoff as $g_\pi(x) = Cx - K$. The solution to this optimal stopping problem is well known, see, e.g., McKean (1965). The optimal stopping threshold, which can be identified as the global maximum of the function $x \mapsto \frac{g_\pi(x)}{\psi_r(x)}$, is the level $x^* = \frac{K\beta_r}{C(\beta_r - 1)}$. The value function is

$$V_\pi(x) = \begin{cases} Cx - K, & x \geq x^*, \\ \left(\frac{C}{\beta_r} \right)^{\beta_r} \left(\frac{\beta_r - 1}{K} \right)^{\beta_r - 1} x^{\beta_r}, & x \leq x^*. \end{cases}$$

6.1.2 Sensitivity analysis

We study the sensitivities of the trigger threshold x^* with respect to the jump intensities of the process Y. We first observe that

$$\frac{dx^*}{d\lambda} = -\frac{K\beta_r}{(\beta_r - 1)C^2}\frac{dC}{d\lambda},$$

here, λ denotes any of the rate variables in the coefficient C. For brevity, denote $\delta = r - \mu$. Then elementary computation yields

$$\frac{dC}{d\lambda_1} = \frac{\delta(\delta + \lambda_2)(\delta + \lambda_{12} + \lambda_2)}{(\delta + \lambda_1 + \lambda_{12})^2(\delta + \lambda_2)^2} > 0,$$

$$\frac{dC}{d\lambda_{12}} = \frac{\delta(\delta + \lambda_2)(\lambda_2 - \lambda_1)}{(\delta + \lambda_1 + \lambda_{12})^2(\delta + \lambda_2)^2} \lesseqqgtr 0,$$

$$\frac{dC}{d\lambda_2} = \frac{\delta\lambda_{12}(\delta + \lambda_{11} + \lambda_2)}{(\delta + \lambda_1 + \lambda_{12})^2(\delta + \lambda_2)^2} > 0.$$

These results are natural. Indeed, if we increase either the rate λ_1 or λ_2, we increase the net absorption rate in the process Y. This decreases the mean absorption time which, in turn accelerates the optimal exercise. The sensitivity with respect λ_{12} depends on relation of λ_1 and λ_2. Say that $\lambda_2 > \lambda_1$. Then, when we increase the rate λ_{12}, we are again increasing the net absorption rate of Y. Increased λ_{12} means that it becomes likely that Y jumps from state 1 to 2 and since $\lambda_2 > \lambda_1$, Y is more likely to be absorbed from state 2 than 1. In the complementary case $\lambda_1 > \lambda_2$, we reason similarly that increased λ_{12} decreased the net absorption rate.

6.1.3 On the effect of increased uncertainty

Conventional wisdom in options theory says that increased risk (measured by the volatility σ) postpones the optimal exercise of option. Next we study whether we can draw some similar conclusion with respect to the risk emerging from the random exercise lag. We do this as follows: by moving along the level curve

$$\bar{\zeta} = \mathbf{E}[\zeta] = \frac{\lambda_{12} + \lambda_2}{(\lambda_1 + \lambda_{12})\lambda_2} \tag{6.2}$$

of the expected absorption time, we study whether an increase in variance of the absorption time, that is, in

$$\text{Var}(\zeta) = \frac{\lambda_2^2 + 2\lambda_{12}\lambda_1 + \lambda_{12}^2}{(\lambda_1 + \lambda_{12})^2\lambda_2^2} \tag{6.3}$$

leads to an increase in the value of C defined in (6.1). First, we solve for λ_1 in (6.2) and obtain

$$\lambda_1 = \frac{\lambda_{12} + \lambda_2 - \lambda_{12}\lambda_2\bar{\zeta}}{\lambda_2\bar{\zeta}}. \tag{6.4}$$

Then substitution to the expression (6.3) and simplification yields

$$\text{Var}(\zeta) = \bar{\zeta}\frac{\bar{\zeta}\lambda_2(\lambda_2 - \lambda_{12}) + 2\lambda_{12}}{\lambda_2(\lambda_2 + \lambda_{12})}.$$

It is a matter of elementary differentiation to see that the gradient

$$\nabla\,\text{Var}(\zeta) = \left(\frac{2\bar{\zeta}(1 - \bar{\zeta}\lambda_2)}{(\lambda_{12} + \lambda_2)^2}, -\frac{2\bar{\zeta}\lambda_{12}(\lambda_{12} + \lambda_2(2 - \bar{\zeta}\lambda_2))}{\lambda_2^2(\lambda_{12} + \lambda_2)^2}\right) := (a_1, a_2). \tag{6.5}$$

Next we do the same analysis for the coefficient C. After substituting (6.4) to the definition of C, a round of simplification and differentiation yields

$$\nabla C = \left(\frac{\delta^2\bar{\zeta}\lambda_2(1 - \bar{\zeta}\lambda_2)}{(\delta + \lambda_2)(\lambda_{12} + \lambda_2(1 + \delta\bar{\zeta}))^2}, -\frac{\delta^2\bar{\zeta}\lambda_{12}(\delta + \lambda_{12} + \lambda_2(2 - \bar{\zeta}\lambda_2))}{(\delta + \lambda_2)(\lambda_{12} + \lambda_2(1 + \delta\bar{\zeta}))^2}\right) \tag{6.6}$$
$$:= (b_1, b_2),$$

where $\delta = r - \mu > 0$. We identify now the directions to which the variance is increasing. Assume that $\lambda_2 > \bar{\zeta}^{-1}$, the complementary case is studied similarly. Then the coefficient $a_1 < 0$ and, consequently, the directional derivative of the variance to the direction $\bar{u} = (u_1, u_2)$, that is,

$$\nabla\,\text{Var}(\zeta) \cdot \mathbf{u} > 0 \quad \text{when} \quad u_1 < -\frac{a_2}{a_1}u_2 = \frac{\lambda_{12}(\lambda_{12} + \lambda_2(2 - \bar{\zeta}\lambda_2))}{\lambda_2^2(1 - \bar{\zeta}\lambda_2)}u_2. \tag{6.7}$$

Assume now that $u_1 < \frac{\lambda_{12}(\lambda_{12}+\lambda_2(2-\bar{\zeta}\lambda_2))}{\lambda_2^2(1-\bar{\zeta}\lambda_2)}u_2$ and that the coefficient

$$\frac{\lambda_{12}(\lambda_{12} + \lambda_2(2 - \bar{\zeta}\lambda_2))}{\lambda_2^2(1 - \bar{\zeta}\lambda_2)} < 0, \tag{6.8}$$

the case complementary to (6.8) is studied similarly. By studying the sign of the numerator and combining the result with the condition $\lambda_2 > \bar{\zeta}^{-1}$, we find that if

$$\lambda_2 \in \left(\frac{1}{\bar{\zeta}}, \frac{1 + \sqrt{1 + \lambda_{12}\bar{\zeta}}}{\bar{\zeta}}\right), \tag{6.9}$$

then (6.8) holds.

The next task is to look at the directional derivatives of the coefficient C. We notice that under the assumption (6.9), the coefficient b_1 in (6.6) is negative. Thus

$$\nabla C \cdot \mathbf{u} > 0 \text{ when } u_1 < -\frac{b_2}{b_1} u_2 = \frac{\lambda_{12}(\delta + \lambda_{12} + \lambda_2(2 - \bar{\zeta}\lambda_2))}{\lambda_2(\delta + \lambda_2)(1 - \bar{\zeta}\lambda_2)} u_2. \tag{6.10}$$

Define the function f as

$$f(\delta) = \frac{\lambda_{12}(\delta + \lambda_{12} + \lambda_2(2 - \bar{\zeta}\lambda_2))}{\lambda_2(\delta + \lambda_2)(1 - \bar{\zeta}\lambda_2)}$$

on the positive real numbers. Elementary differentiation yields

$$f'(\delta) = -\frac{\lambda_{12}(\lambda_{12} + \lambda_2(1 - \bar{\zeta}\lambda_2))}{\lambda_2(\delta + \lambda_2)^2(1 - \bar{\zeta}\lambda_2)}.$$

We obtain by a simple computation

$$f'(\delta) \gtreqless 0 \text{ when } \lambda_2 \lesseqgtr \frac{\frac{1}{2} + \sqrt{\frac{1}{4} + \lambda_{12}\bar{\zeta}}}{\bar{\zeta}} < \frac{1 + \sqrt{1 + \lambda_{12}\bar{\zeta}}}{\bar{\zeta}}. \tag{6.11}$$

Using the function f, we observe that the condition (6.8) can be written as $f(0) < 0$ and that the sufficient condition in (6.7) can be written as $u_1 < f(0)u_2$. Furthermore, the sufficient condition in (6.10) can be written as $u_1 < f(\delta)u_2$. Assume that $u_2 < 0$, thus $f(0)u_2 > 0$. Using the condition (6.11), we find that if

$$\lambda_2 \in \left(\frac{1}{\bar{\zeta}}, \frac{\frac{1}{2} + \sqrt{\frac{1}{4} + \lambda_{12}\bar{\zeta}}}{\bar{\zeta}} \right), \tag{6.12}$$

then $0 < f(0)u_2 < f(\delta)u_2$. In other words, if the condition (6.12) is satisfied and the variance is increasing in the direction in a direction where u_2 is negative, then the coefficient C and, consequently, the optimal stopping threshold x^* is also increasing in that same direction. By reasoning similarly, we obtain the same conclusion in case $u_2 > 0$ when

$$\lambda_2 \in \left(\frac{\frac{1}{2} + \sqrt{\frac{1}{4} + \lambda_{12}\bar{\zeta}}}{\bar{\zeta}}, \frac{1 + \sqrt{1 + \lambda_{12}\bar{\zeta}}}{\bar{\zeta}} \right). \tag{6.13}$$

Summarizing, we have the following partial result: the conditions (6.12) and (6.13) identify, in their corresponding cases, parts of the level curve (6.2) where increased exercise lag risk (measured in terms of the variance of the exercise lag) postpones the optimal exercise.

6.2 Square-root diffusion

Let X be a square root diffusion known as the Cox–Ingersoll–Ross process, for the following properties of X, we refer to Campolieti and Makarov (2012), Sect. 5.2. The infinitesimal generator of X reads as

$$\mathcal{A} = (a - bx)\frac{d}{dx} + \frac{1}{2}\sigma^2 x \frac{d^2}{dx^2}.$$

Introduce the parameters

$$\mu = \frac{2a}{\sigma^2} - 1, \ \kappa = \frac{2b}{\sigma^2}.$$

The scale function and speed measure are now given by $S'(x) = x^{-\mu-1}e^{\kappa x}$ and $m'(x) = \frac{2}{\sigma^2}x^{\mu}e^{-\kappa x}$ for all $x \in (0, \infty)$. The upper boundary ∞ is natural. We assume that $\mu > 0$, then the lower boundary 0 is entrance-not-exit.

For an arbitrary $\rho > 0$, the functions ψ_ρ and φ_ρ can be represented as

$$\psi_\rho(x) = M\left(\frac{\rho}{b}, \frac{2a}{\sigma^2}, \frac{2b}{\sigma^2}x\right), \ \varphi_\rho(x) = U\left(\frac{\rho}{b}, \frac{2a}{\sigma^2}, \frac{2b}{\sigma^2}x\right)$$

where M and U are, respectively, the confluent hypergeometric functions of first and second type, for the definitions of M and U, for the properties of these functions, see Borodin and Salminen (2015), p. 647. Finally, the Wronskian determinant can then be expressed as $B_\rho = \kappa^{-\mu}\frac{\Gamma(\mu+1)}{\Gamma(\frac{\rho}{b})}$.

We consider the exercise payoff $\mathfrak{g}_\pi : x \mapsto \mathbf{E}_x\left[e^{-r\zeta}\sqrt{X_\zeta}\right] - K$, where $K > 0$. Denote the square root function as h. Then we write using the resolvent equation
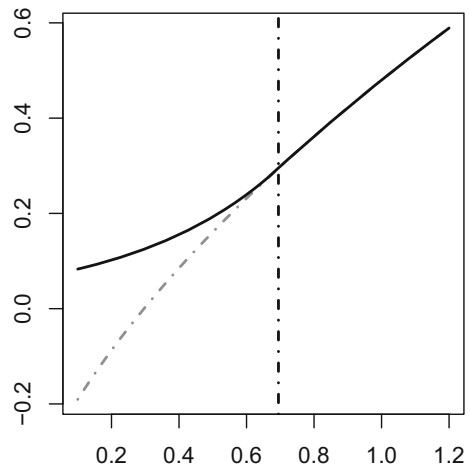
$$\mathbf{E}_x\left[e^{-r\zeta}\sqrt{X_\zeta}\right] = \frac{\lambda_1}{\mu_1}\mu_1(R_{r+\mu_1}h)(x) + \frac{\lambda_{12}}{\mu_1}\mu_1(R_{r+\mu_1}\lambda_2 R_{r+\lambda_2}h)(x)$$

$$= \left(\lambda_1 + \frac{\lambda_{12}\lambda_2}{\lambda_2 - \mu_1}\right)(R_{r+\mu_1}h)(x) - \frac{\lambda_{12}\lambda_2}{\lambda_2 - \mu_1}(R_{r+\lambda_2}h)(x).$$

We cannot find explicit expression for these resolvents, therefore we resort to numerical solution. To this end, R-packages `fAsianOptions` and `numDeriv` were employed to handle the Kummer functions numerically and to numerically determine the zero (x^*, that is) of the derivative of $\frac{\mathfrak{g}_\pi(x)}{\psi_r(x)}$, respectively. Using these, the value function were obtained numerically from Proposition 5.1.

In Fig. 1 we illustrate the solution of the problem. The solid black curve represents the value function and the grey dashed line represents the exercise payoff $x \mapsto \mathbf{E}_x\left[e^{-r\zeta}\sqrt{X_\zeta}\right] - K$ under the parameter configuration $a = 0.03$, $b = 0.05$, $\sigma = 0.2$, $r = 0.06$, $K = 1$, $\lambda_1 = 0.1$, $\lambda_{12} = 0.2$, and $\lambda_2 = 0.1$. The black dashed line indicates the position of the optimal exercise threshold $x^* = 1.701597$. As the general theory suggests, the figure indicates that the value function is smooth over the optimal exercise boundary x^*.

Fig. 1 *Square root diffusion* the solid black curve represents the value function and the grey dashed line represents the exercise payoff

$x \mapsto \mathbf{E}_x \left[e^{-r\zeta} \sqrt{X_\zeta} \right] - K$. The black dashed line indicates the position of the optimal exercise threshold $x^* = 1.701597$

References

Aalen OO (1995) Phase-type distributions in survival analysis. Scand J Stat 22:447–463

Aïd R, Federico S, Pham H, Villeneuve B (2015) Explicit investment rules with time-to-build and uncertainty. J Econ Dyn Control 51:240–256

Alvarez LHRE (2001) Reward functionals, salvage values, and optimal stopping. Math Methods Oper Res 54(2):315–337

Alvarez LHRE, Keppo J (2002) The impact of delivery lags on irreversible investment under uncertainty. Eur J Oper Res 136:173–180

Alvarez LHRE, Stenbacka R (2001) Adoption of uncertain multi-stage technology projects: a real options approach. J Math Econ 35(1):71–97

Alvarez LHRE (2003) On the properties of r-excessive mappings for a class of diffusions. Ann Appl Probab 13(4):1517–1533

Bar-Ilan A, Strange WC (1996) Investment lags. Am Econ Rev 86:610–622

Bar-Ilan A, Sulem A (1995) Explicit solution of inventory problems with delivery lags. Math Oper Res 20(3):709–720

Bladt M (2005) A review of phase-type distributions and their use in risk theory. ASTIN Bull 35(1):145–161

Borodin A, Salminen P (2015) Handbook on Brownian motion—facts and formulæ. Birkhäuser, Basel

Brennan MJ, Schwartz ES (1985) Evaluating natural resource investments. J Bus 58(2):135–157

Breure L, Baum D (2005) Introduction to queueing theory and matrix-analytic methods. Springer, Berlin

Campolieti G, Makarov RN (2012) On properties of analytically solvable families of local volatility diffusion models. Math Financ 22(3):488–518

Capozza D, Li Y (1994) The intensity and timing of investment: the case of land. Am Econ Rev 84(4):889–904

Dixit A (1989) Entry and exit decisions under uncertainty. J Polit Econ 97(3):620–638

Dixit A, Pindyck RS (1994) Investment under uncertainty. Princeton University Press, Princeton

Dixit A, Pindyck RS (1995) The options approach to capital management. Harv Bus Rev 77(3):105–115

Fackrell M (2009) Modelling healthcare systems with phase-type distributions. Health Care Manag Sci 12:11–26

Grenadier SR (1996) The strategic exercise of options: development cascades and overbuilding in real estate markets. J Financ 51:1653–1679

He Q-M (2014) Fundamentals of matrix-analytic methods. Springer, Berlin

Hobolth A, Siri-Jegousse A, Bladt M (2018) Phase-type distributions in population genetics. arXiv:1806.01416 [q-bio.PE]

Kydland FE, Prescott EC (1982) Time to build and aggregate fluctuations. Econometrica 50(6):1345–1370

Lempa J (2010) A note on optimal stopping of diffusions with a two-sided optimal rule. Oper Res Lett 38:11–16

Lempa J (2012) Optimal stopping with information constraint. Appl Math Optim 66(2):147–173

Lempa J (2012) Optimal stopping with random exercise lag. Math Methods Oper Res 75(3):273–286

Majd S, Pindyck RS (1987) Time to build, option value and investment decisions. J Financ Econ 18:7–27

Margsiri W, Mello AS, Ruckes ME (2008) A dynamic analysis of growth via acquisition. Rev Financ 12:635–671

Myers SC (1977) Determinants of corporate borrowing. J Financ Econ 5(2):147–175

McDonald R, Siegel D (1986) The value of waiting to invest. Q J Econ 104(1):707–727

McKean HP Jr (1965) Appendix: a free boundary problem for the heat equation arising from a problem of mathematical economics. Ind Manag Rev 6:32–39

Paddock JL, Siegel D, Smith JL (1988) Option valuation of claims on real assets: the case of offshore petroleum leases. Q J Econ 103(1):479–508

Panayi S, Trigeorgis L (1998) Multi-stage real options: the cases of information technology infrastructure and international bank expansion. Q Rev Econ Financ 38:675–692

Peskir G, Shiryaev A (2006) Optimal stopping and free boundary problems. Birkhäuser, Basel

Ross SA (1978) A simple approach to the valuation of risky income streams. J Bus 51(3):453–475

Ross SM (2010) Introduction to probability models, 10th edn. Academic Press, Cambridge

Sarkar S, Zhang C (2015) Investment policy with time-to-build. J Bank Financ 55:142–156

Schwartz ES (1997) The stochastic behavior of commodity prices: implications for valuation and hedging. J Financ 52(3):923–973

Stewart WJ (2009) Probability, Markov chains, queues, and simulation. Princeton University Press, Princeton

Teschl G (2012) Ordinary differential equations and dynamical systems. American Mathematical Society, Providence

Titman S (1985) Urban land prices under uncertainty. Am Econ Rev 75(3):505–514

Williams JT (1993) Equilibrium and options in real assets. Rev Financ Stud 6:825–850