OPTIMAL STOPPING FOR PAIRS TRADING STRATEGIES

A thesis submitted to the University of Manchester
for the degree of Doctor of Philosophy
in the Faculty of Science and Engineering

2023

Zeping Bian
Department of Mathematics
School of Natural Sciences

Contents

Word count xxxxx

The University of Manchester

Zeping Bian
Doctor of Philosophy
Optimal Stopping for Pairs Trading Strategies
July 4, 2023

This PhD thesis studies the detection problem associated with pairs trading strategies. We model the difference between the prices of the two underlying stocks in pairs trading strategies as an Ornstein-Uhleneck process based on its mean-reverting property. The mean-reversion rate (and the mean-reverting level) of the underlying Ornstein-Uhleneck process is assumed to have a change at some random/unobservable time. We consider the problem of detecting the random/unobservable time as accurately as possible. The problem is formulated as a quickest detection problem. We consider the most random scenario where the unobservable/random time is (i) exponentially distributed and (ii) independent from the underlying Ornstein-Uhleneck process prior to its change in the mean-reversion rate.

We formulate the quickest detection problem as a *Lagrangian* of the probability of *false alarm* and the expected *detection delay*. By changing the probability measures we transform the quickest detection problem to a two-dimensional *Lagrange* formulated optimal stopping problem. We reduce the underlying process to its canonical form and transform the optimal stopping problem correspondingly. We also decouple the coefficients on the diffusion term of the underlying two-dimensional process by time change. The Lagrange formulated optimal stopping problem is transformed to its *Mayer* formulation through Itô's formula and the optional sampling theorem. We reduce the Lagrange formulated optimal stopping problem to a free-boundary problem where the derivatives are understood in the sense of Schwartz distribution. We then verify that the canonical infinitesimal generator satisfies Hörmander's condition which upgrades the weak solution to a strong (smooth/classic) solution.

We determine an upper bound on the rates of convergence in the *Wald-Bellman equations* depending on the gain function and the transition density function of the underlying Markov process. Making use of theories from parabolic second-order partial differential equations, we derive an upper bound on the rates of convergence when the transition density function is not known explicitly. We then introduce a technique of constructing the value functions to enable the convergence of Wald-Bellman equations to continuous-time Mayer formulated optimal stopping problems with finite horizon. The technique is first applied to the detection problem associated with pairs trading strategies, then applied to various Mayer formulated continuous-time optimal stopping problems. Numerical approximations of Wald-Bellman equations are obtained through Mathematica algorithms. The value functions are then used to generate the corresponding optimal stopping boundaries through numerical calculations in Mathematica.

Declaration

No portion of the work referred to in the thesis has been submitted in support of an application for another degree or qualification of this or any other university or other institute of learning.

Copyright Statement

Acknowledgements

I would like to thank my family for all the supports they have provided throughout my life.

My thanks to my friends who have provided many supports through the hard time, to fellow PhD students Dr Tianren Bu, Dr Yuwei Chu, Dr Min Gao, Muye Li, Dr Shi Qiu, Yingjie Wang and Dr Jingsi Xu for the valuable discussions, and to Dr Kristoffer Glover for discussions on numerical analysis.

My supervisor Prof. Goran Peskir, your clear and in-depth explanations always inspire my passion for mathematics. I appreciate your guidance not only on mathematics but also on my life path.

Chapter 1

Introduction

The *pairs trading* strategies have been popular since introduced by Morgan Stanley in 1980s (cf. [33]). The two stock prices are assumed to have stationary and mean-reverting spread around zero that means when the prices of the underlying stocks diverge, the price spread is expected to revert to zero. The trader would short the outperforming stock and long the underperforming stock to construct the portfolio. If the spread reverts to zero as expected, then the investor would gain profits from both long and short positions. The pairs trading strategies can be considered as market neutral. For instance, in a market depression, the trader would incur losses from the long position, but gain profits from the short position, which offsets the losses. In practice, however, the assumed mean-reverting property may disappear at some point which could cause large losses to the trader if the portfolio is not adjusted.

Currently, the investors manage such type of risk by using a stop-loss strategy (cf. [17], [15] and [16]) where the underlying stocks are treated based on an assumed mean-reverting model until they reach a predetermined loss level, then the investors may liquidate their stocks with a loss. Among the academic literature, the spread of the prices was first modeled as a mean-reverting process in [7]. The spread was first modeled as an Ornstein-Uhlenbeck process in [6]. [9] was the first paper to investigate in the detection problem on a change of the mean-reverting property. [9] formulated the problem as a quickest detection problem for Ornstein-Uhlenbeck processes. A similar quickest detection problems for Bessel processes was discussed in [14]. One may find applications of the quickest detection problems of Ornstein-Uhlenbeck processes beyond pairs trading strategy. For instance, an Ornstein-Uhlenbeck process with negative initial mean-reversion rate may be used to model financial bubbles (see [22] and [23]). The quickest detection problems for such Ornstein-Uhlenbeck processes may help detect the *bursting* of such financial bubbles.

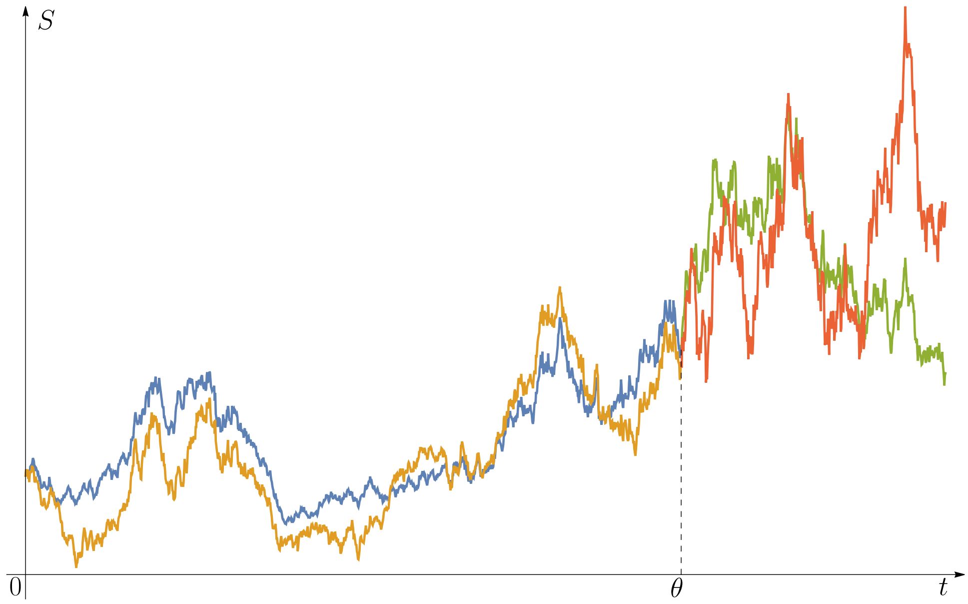

Figure 1. The stationary (Ornstein-Uhlenbeck process with positive mean-reversion rate) spread becomes recurrent (Brownian motion) at time θ.

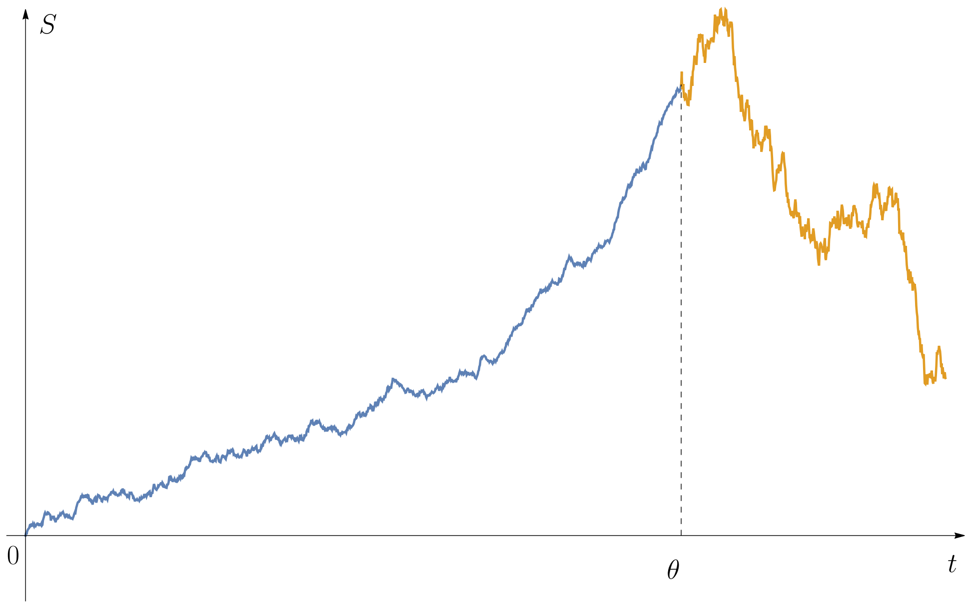

Figure 2. A transient Itô process (Ornstein-Uhlenbeck process with negative mean-reversion rate) becomes recurrent (Brownian motion) at time θ.

This PhD project studies the similar quickest detection problems for Ornstein-Uhlenbeck processes introduced in [9] with a different method namely the Wald-Bellman equations and a more general choice of the mean-reverting levels. We consider the problem in the most random situation, where we assume that the random time, when a change of the mean-reverting rate happens, is (i) exponentially distributed and (ii) independent from the previous Ornstein-Uhlenbeck process which is prior to the unobservable random time. To detect the random optimal stopping time as accurately as possible, we minimize the Lagrangian of the probability of the *false alarm* and the expected *detection delay* . The formulation of general quickest detection problems were first introduced in [28] and still being studied to date (cf. [29]). We reformulate the quickest detection problems to optimal stopping problems by changing probability measures. The optimal stopping problems are solved by Wald-Bellman equations (cf. [21, Theorem 1.7]). Numerical approximations are obtained through Mathematica algorithms. We also study the Wald-Bellman equations in the general case for n-dimensional time-homogeneous Markov processes and apply the Wald-Bellman equations to various optimal stopping problems. This thesis is organised as follows:

In Chapter 2, we formulate the financial problem as a mathematical problem. In Section 2.2, we formulate the quickest detection problem as a Lagrangian of the probability of *false alarm* and expected *detection delay* under the assumption that the observed Ornstein-Uhlenbeck process changes its initial mean-reverting rate (with a change of mean-reversion level) at a random/unobservable time. The quickest detection problem is then reformulated as an optimal stopping problem in terms of the *posterior probability distribution* process. Then we define the *likelihood ratio process* and the *posterior probability ratio* process. In Section 2.3, we transform the optimal stopping which is a function of the *posterior probability distribution* process to an optimal stopping problem in terms of the *posterior probability ratio* process by changing the probability measure. The one-dimensional problem becomes two-dimensional from where we consider the underlying stochastic process in two-dimension. In Section 2.4, we reduce the underlying two-dimensional stochastic process to its canonical form by using standard partial differential equations arguments. We also apply transformations to the canonical process to simplify the problem. The optimal stopping problems are reformulated correspondingly. In Section 2.5, we apply time change technique to the observed Ornstein-Uhlenbeck process to decouple the coefficients on the diffusion terms in the system of stochastic differential equations. We then derive the optimal stopping problem in terms of the time-changed process. In Section 2.6, we introduce the Mayer formulation for the optimal stopping problem. We have not found the Mayer formulation in the general case. The Mayer formulation is essential for the application of Wald-Bellman equations. The lack of Mayer formulation puts restrictions on the choices of mean-reversion rates.

In Chapter 3, we first derive the system of free-boundary problem which the value function satisfies where the derivatives are understood in the sense of Schwartz distribution. We then verify that the given weak solution satisfies the Hörmander's condition which implies that the value function is actually a strong (smooth/classic) solution.

In Chapter 4, we first recall the basic setting for n-dimensional Wald-Bellman equations. We then introduce an upper bound on the rates of convergence in the Wald-Bellman equations. We also introduce a technique of constructing the value functions associated with *finite-horizon continuous-time Mayer* formulated optimal stopping problems. The upper bound depends on the gain function and the transition density function of the underlying process. Making use of theories from parabolic second-order partial differential equations, we derive an upper bound on the rates of convergence when the transition density function is not known explicitly.

In Chapter 5, we apply the Wald-Bellman equations to the Mayer formulated optimal stopping problem introduced in Section 2.6. By using *Euler approximation* (cf. [13, Chapter 2.1,]) we obtain numerical approximations of the value functions and optimal stopping boundaries through Mathematica algorithms. The method of Wald-Bellman equations is applicable to the optimal stopping problems which are Mayer formulated with known stochastic differential equations for the underlying diffusion processes.

In Chapter 6, we apply the Wald-Bellman equations to a Wald's type optimal stopping problem for one-dimensional Brownian motion with finite horizon of which the analytical solution is known in the case of infinite horizon. Numerical approximations of the value functions and optimal stopping boundaries are obtained through Mathematica algorithms.

In Chapter 7, we study a *time-inhomogeneous* optimal stopping problem for one-dimensional Brownian motion with finite horizon. The problem is solved by using time-inhomogeneous Wald-Bellman equations. Numerical approximations of the value functions and optimal stopping boundaries are obtained through Mathematica algorithms.

In Chapter 8, we study the finite-horizon American put option. The problem is solved in two different ways. We first simplify the problem by "killing" the sample path of the underlying process and then transform to a time-homogeneous optimal stopping problem. The simplified problem is solved by the time-homogeneous Wald-Bellman equations. We then solve the problem by using time-inhomogeneous Wald-Bellman equations directly. Numerical approximations of the value functions and optimal stopping boundaries are obtained through Mathematica algorithm.

In Chapter 9, we study the same quickest detection problem for Bessel processes introduced in [14]. The Mayer formulated optimal stopping problem is solved by Wald-Bellman equations. Numerical approximations of the value functions and optimal stopping boundaries are obtained through Mathematica algorithms.

Sections 2.1-2.3, 2.5 and Chapter 3 are generalizations of Sections 3-5 and 6 in [9]. Parts of Chapters 4-8 form a paper [3] which is submitted to the journal "Stochastic Processes and their Applications" for consideration of publication.

Chapter 2

Formulation of the problem

In this Chapter, we formulate the optimal stopping problem for pairs trading strategies as quickest detection problems for Ornstein-Uhlenbeck processes.

2.1 Contributions

The setting of quickest detection problems in Section 2.2 is a generalization of [9, Section 3] in which the authors set the mean-reverting levels to 0. Such setting helps to obtain the symmetry property of the underlying process around 0 with some losses of generality which simplifies further analysis. In this project, we attempt to consider the problem in full generality where the mean-reverting levels are arbitrary real values.

The likelihood ratio process plays an important role to connect the underlying process with the random/unobservable time that we intend to detect. The likelihood ratio process is defined as a Radon–Nikodym derivative where the explicit expression in terms of the underlying process can be obtained. The previous literature makes use of measure theory and Bayes formula to obtain the expression. In this research, we use a different but less general argument through the Girsanov theorem. Moreover, we use a more probabilistic method to determine the explicit expression of the posterior probability ratio process in terms of the underlying process.

Since we consider the problem in full generality, the signal-to-noise ratio will always contain a constant term which is dependent on the diffusion coefficient, the mean-reversion rates and the mean-reverting levels. This constant term makes the system of stochastic differential equations for the underlying process more complicated and the analysis on the corresponding infinitesimal generator more challenging than [9, Section 5]. We have also applied time-change transformation to decouple the stochastic processes on the diffusion term in Section 2.5. The method is similar to [14, Section 7] whereas the technique is not used in [9].

We determine the Mayer formulation of the optimal stopping problem through Itô's formula and the optional sampling theorem in Section 2.6 with restriction on the mean-reverting rates. The Mayer formulation is only available when the mean-reverting is not equal to $-\lambda/2$ before the change and equal to 0 after the change. Under this restriction, we lose the generality of the problem because, when the mean-reversion rate equals 0 after the change, the problem will be simplified to the cases when the mean-reverting levels are 0 as considered in [9].

2.2 Setting

In this section we introduce the setting for the corresponding quickest detection problem. The quickest detection problem is then transferred to a *Bolza* formulated optimal stopping problem (cf. [21, Chapter III]).

1. Let the unobservable/random time θ take value 0 with probability $\pi \in [0,1]$ and follow exponential distribution with parameter $\lambda > 0$ given $\theta > 0$. Given $\theta > 0$ for any subset $\mathbb{A}$ of all positive real numbers $I\!R^+$, the probability that θ equals t is $(1-\pi)\int_{\mathbb{A}} \lambda\, e^{-\lambda t}\, dt$ for $t \in \mathbb{A}$. The setting for θ can be realised on the probability space $(\Omega, \mathcal{F}, \mathsf{P}_\pi)$ where the probability measure P_π is defined by

$$\mathsf{P}_\pi = \pi\, \mathsf{P}^0 + (1-\pi) \int_0^\infty \lambda\, e^{-\lambda t}\, \mathsf{P}^t\, dt. \tag{2.1}$$

P^t is the probability measure under which θ equals t for $t \in [0,\infty)$. The underlying Ornstein-Uhlenbeck process X changes its mean-reversion rate from $\beta_0 \in I\!R$ to $\beta_1 \in I\!R$ (and mean-reverting level $x_0 \in I\!R$ to $x_1 \in I\!R$) at time t under P^t. The unobservable/random time θ is a non-negative random variable such that $\mathsf{P}_\pi(\theta = 0) = \pi$ and $\mathsf{P}_\pi(\theta > t \mid \theta > 0) = e^{-\lambda t}$ for $t \in I\!R^+$. The probability of the underlying Ornstein-Uhlenbeck process X changes its mean reversion rate $\beta_0 \in I\!R$ to $\beta_1 \in I\!R$ (and mean-reverting level $x_0 \in I\!R$ to $x_1 \in I\!R$) at time $t > 0$ under probability measure P^t is the same as the probability of X changes its mean reversion rate given that θ equals

$t > 0$ under probability measure P_π for $t \in I\!R^+$, i.e. $\mathsf{P}^t(X \in \cdot\,) = \mathsf{P}_\pi(X \in \cdot \mid \theta = t)$ for $t > 0$, which is the probability law of the underlying Ornstein-Uhlenbeck process X. When the mean-reversion rate of the underlying Ornstein-Uhlenbeck process X does not change and remains β_0 (with mean-reverting x_0) all the time, we have $\theta = t = \infty$ and denote such probability measure by P^∞. The probability law of the Ornstein-Uhlenbeck process X is $\mathsf{P}^\infty(X \in \cdot\,) = \mathsf{P}_\pi(X \in \cdot \mid \theta = \infty)$ in this case.

2. The observed Ornstein-Uhlenbeck process X satisfies the stochastic differential equation

$$dX_t = [\mu_0(X_t) + I(t \geq \theta)\,(\mu_1(X_t) - \mu_0(X_t))]\,dt + \sigma(X_t)\,dB_t \tag{2.2}$$

where B is a standard Brownian motion under P_π and

$$\mu_0(x) = \beta_0\,(x_0 - x) \quad \& \quad \mu_1(x) = \beta_1\,(x_1 - x) \quad \& \quad \sigma(x) = \sigma > 0 \tag{2.3}$$

for $x \in I\!R$ with the assumption that θ and the standard Brownian motion B are independent under P_π for $\pi \in [0, 1]$.

3. There are eleven cases of β_0 and β_1 mentioned in [9] which are

Case 0: $\beta_0 = \beta_1 \neq 0$ (with $x_0 \neq x_1$).

Case 1: $\beta_0 > 0$ and $\beta_1 = 0$ (One may set $x_0 = x_1 = 0$ without loss of generality).

Case 2: $\beta_0 = 0$ and $\beta_1 > 0$ (One may set $x_0 = x_1 = 0$ without loss of generality).

Case 3: $\beta_0 < 0$ and $\beta_1 = 0$ (One may set $x_0 = x_1 = 0$ without loss of generality).

Case 4: $\beta_0 = 0$ and $\beta_1 < 0$ (One may set $x_0 = x_1 = 0$ without loss of generality).

Case 5: $0 < \beta_0 < \beta_1$ (with $x_0 = x_1 = 0$).

Case 6: $0 < \beta_1 < \beta_0$ (with $x_0 = x_1 = 0$).

Case 7: $\beta_1 < 0 < \beta_0$ (with $x_0 = x_1 = 0$).

Case 8: $\beta_1 < \beta_0 < 0$ (with $x_0 = x_1 = 0$).

Case 9: $\beta_0 < \beta_1 < 0$ (with $x_0 = x_1 = 0$).

Case 10: $\beta_0 < 0 < \beta_1$ (with $x_0 = x_1 = 0$).

In Cases 5 - 10 one may set $x_0 \neq x_1$ with non-zero real values of x_0 and x_1 for full generality. Such cases may bring more difficulties to the problem which we will discuss in later chapters. When both β_0 and β_1 are non-zero with non-zero mean-reverting levels x_0 and x_1, there will be a change in mean-reversion level at the random time θ. If either β_0 or β_1 is zero, then the corresponding mean-reversion level can be set to 0, and thus the other mean-reverting level can be set to zero by making an appropriate translation of the underlying Ornstein-Uhlenbeck process X. One may also set $x_0 = 0$ and $x_1 \neq 0$ in Cases 5 - 10 without loss of generality by making an appropriate translation. [9] discussed the above eleven cases when $x_0 = x_1 = 0$. We will discuss the cases when $x_0 = 0$ and $x_1 \neq 0$.

4. Based on the observed Ornstein-Uhlenbeck process X we want to find a $\mathcal{F}_t^X$-measurable stopping time τ_* that is closest to the unobservable/random time θ where $\mathcal{F}_t^X$ is the natural filtration of X defined by $\mathcal{F}_t^X = \sigma(X_s \mid 0 \leq s \leq t)$ for $t \geq 0$. The formulation of such quickest detection problem was first studied in [28], based on which the value function is defined by

$$V(\pi) = \inf_{\tau \geq 0} \left[\mathsf{P}_\pi(\tau < \theta) + c\, \mathsf{E}_\pi(\tau - \theta)^+ \right] \tag{2.4}$$

where $\pi \in [0, 1]$ is the probability that $\theta = 0$ and $c \in I\!R^+$ is given and fixed. We aim to find the optimal stopping time τ_* when the infimum in (2.4) is attained and the corresponding value of the value function V. $\mathsf{P}_\pi(\tau < \theta)$ is the probability of *false alarm* and $\mathsf{E}_\pi(\tau - \theta)^+$ is the expected *detection delay*.

5. We define the *posterior probability distribution* process $\Pi = (\Pi_t)_{t \geq 0}$ of θ given X by

$$\Pi_t = \mathsf{P}_\pi(\theta \leq t \mid \mathcal{F}_t^X) \tag{2.5}$$

for $t \geq 0$. Π can be interpreted as the best prediction of the distribution function of the unobservable/random time θ given the sample path of the observed Ornstein-Uhlenbeck process X. (2.4) can be rewritten as

$$V(\pi) = \inf_{\tau \geq 0} \mathsf{E}_\pi \left(1 - \Pi_\tau + c \int_0^\tau \Pi_t \, dt \right) \tag{2.6}$$

for $\pi \in [0, 1]$ using similar arguments in [29, Section 4.3, Theorem 7]. The *signal-to-noise ratio* is defined by

$$\rho(x) = \frac{\mu_1(x) - \mu_0(x)}{\sigma(x)} \tag{2.7}$$

for $x \in I\!R$. If we consider the case of Ornstein-Uhlenbeck process, then the *signal-to-noise ratio* is simplified to

$$\rho(x) = \frac{\mu_1(x) - \mu_0(x)}{\sigma(x)} = \frac{\beta_0 - \beta_1}{\sigma} x + \frac{\beta_1 x_1 - \beta_0 x_0}{\sigma} = \delta x + \gamma \tag{2.8}$$

where $\delta = (\beta_0 - \beta_1)/\sigma$ and $\gamma = (\beta_1\, x_1 - \beta_0\, x_0)/\sigma$ for $x \in I\!R$.

6. We denote the restrictions of the probability measures P^0 and P^∞ to $\mathcal{F}_t^X$ by P_t^0 and P_t^∞ respectively. The likelihood ratio process $L = (L_t)_{t \geq 0}$ is defined by

$$L_t = \frac{d\mathsf{P}_t^0}{d\mathsf{P}_t^\infty}. \tag{2.9}$$

By the Girsanov theorem (cf. [26, Theorem 38.5]), there exists a previsible process $c = (c_t)_{t \geq 0}$ in $I\!R$ under P^∞ such that

$$L_t = \frac{d\mathsf{P}_t^0}{d\mathsf{P}_t^\infty} = \exp\left(\int_0^t c_s\, dB_s - \frac{1}{2}\int_0^1 c_s^2\, ds\right). \tag{2.10}$$

The measure P_t^∞ is equivalent to the measure P_t^0 on $t \geq 0$ because the underlying Ornstein-Uhlenbeck process is indistinguishable when the observing time is finite. Let $B = (B_t)_{t \geq 0}$ be a standard Brownian motion under P^∞. Then by the Girsanov theorem, $\tilde{B}_t = B_t - \int_0^t c_s\, ds$ is a standard Brownian motion under P^0 for $t \geq 0$. Under measure P^0 let $\tilde{B}$ be the driving Brownian motion for X. By definition of P^0, under P^0, the underlying Ornstein-Uhlenbeck process X satisfies the stochastic differential equation

$$dX_t = \mu_1(X_t)\, dt + \sigma\, d\tilde{B}_t = [\mu_1(X_t) - \sigma c_t]\, dt + \sigma\, dB_t. \tag{2.11}$$

The equation (2.11) may be viewed under measure P^∞. Let $\hat{B} = (\hat{B}_t)_{t \geq 0}$ be the driving standard Brownian motion of X under P^∞. Then X satisfies the stochastic differential equation

$$dX_t = \mu_0(X_t)\, dt + \sigma\, d\hat{B}_t \tag{2.12}$$

under P^∞. Note that we cannot conclude that the standard Brownian motion $\hat{B}$ is equal to B at the moment. We aim to show that the drift coefficient in the latter part of (2.11) is equivalent to the drift coefficient in (2.12), i.e. $\mu_0(X_t) = \mu_1(X_t) - \sigma c_t$. (2.11) and (2.12) may be rewritten in integral form

$$X_t = X_0 + \int_0^t [\mu_1(X_s) - \sigma c_s]\, ds + \int_0^t \sigma\, dB_s \tag{2.13}$$

$$X_t = X_0 + \int_0^t \mu_0(X_s)\, ds + \int_0^t \sigma\, d\hat{B}_s. \tag{2.14}$$

Subtracting (2.13) by (2.14) and rearranging, we obtain

$$\int_0^t \mu_0(X_s) - \mu_1(X_s) + \sigma c_s \, ds = \sigma \left(B_t - \hat{B}_t \right) \tag{2.15}$$

for $t \geq 0$. The left hand side of (2.15) is a process of bounded variation. The right hand side of (2.15) is a continuous martingale. Then, it is well know that

$$\int_0^t (\mu_0(X_s) - \mu_1(X_s) + \sigma c_s) \, ds = \sigma \left(B_t - \hat{B}_t \right) = 0 \tag{2.16}$$

P^∞ almost surely. We thus have shown $B_t = \hat{B}_t$ for $t \geq 0$ and

$$\int_0^t \frac{\mu_1(X_s) - \mu_0(X_s)}{\sigma} \, ds = \int_0^t c_s \, ds \tag{2.17}$$

P^∞ almost surely.

7. Given (2.17), we aim to show $c_t = (\mu_1(X_t) - \mu_0(X_t))/\sigma$ which helps us determine L in (2.10). Define functions $f(s) = f^+(s) - f^-(s) := (\mu_1(X_s) - \mu_0(X_s))/\sigma$ and $g(s) = g^+(s) - g^-(s) := c_s$ for $s \in [0, t]$ and $t \geq 0$. Then (2.17) can be rewritten as

$$\int_0^t \left(f^+(s) + g^-(s) \right) ds = \int_0^t \left(g^+(s) + f^-(s) \right) ds. \tag{2.18}$$

We define Lebesgue measures μ_1 and μ_2 by

$$\mu_1(\mathcal{A}) = \int_{\mathcal{A}} \left(f^+(s) + g^-(s) \right) ds \tag{2.19}$$

$$\mu_2(\mathcal{A}) = \int_{\mathcal{A}} \left(g^+(s) + f^-(s) \right) ds \tag{2.20}$$

for $\mathcal{A} := \{(t_1, t_2] \mid 0 \leq t_1 < t_2 < +\infty\} \bigcup \{\emptyset\}$ with $\mu_1(\emptyset) = 0$ and $\mu_2(\emptyset) = 0$ recalling that $f^+(s) := \max(f(s), 0)$ and $f^-(s) := \max(-f(s), 0)$. It can be shown that $\mu_1(\mathcal{A}) = \mu_2(\mathcal{A}) < \infty$. Note that $\mathcal{A}$ defines a π-system and $\mathcal{B}([0, +\infty)) = \sigma(\mathcal{A})$, then by [35, Lemma 1.6(a)], we have $\mu_1 = \mu_2$ on $\mathcal{B}([0, +\infty))$. Define $\mathcal{A}_1, \mathcal{A}_2 \in \mathcal{B}([0, +\infty))$ where $\mathcal{A}_1 := \{s \mid f^+(s) + g^-(s) > g^+(s) + f^-(s)\} = \{s \mid f(s) > g(s)\}$ and $\mathcal{A}_2 := \{s \mid f^+(s) + g^-(s) < g^+(s) + f^-(s) = \{s \,|\, f(s) < g(s)\}$. Then we have

$$\begin{aligned} \int_{\mathcal{A}_1} \left(f^+(s) + g^-(s) - g^+(s) - f^-(s) \right) ds &= \int_{\mathcal{A}_1} (f(s) - g(s)) \, ds \\ &= \int_{[0,+\infty)} I_{\mathcal{A}_1}(s) \, (f - g)(s) \, ds = 0. \end{aligned} \tag{2.21}$$

Since $I_{\mathcal{A}_1}(s)(f - g)(s) \geq 0$ for $s \in [0, t]$ and $t \geq 0$, by [5, Corollary 2.3.12], we have $I_{\mathcal{A}_1}(s)(f - g)(s) = 0$ λ-*almost everywhere* where λ is the Lebesgue measure. Then we have $f(s) \le g(s)$ λ-*almost everywhere*. By analogy it can be shown that $f(s) \ge g(s)$ λ-*almost everywhere*. Combining the results we have

$$c_t = \frac{\mu_1(X_t) - \mu_0(X_t)}{\sigma} \tag{2.22}$$

for $t \ge 0$ λ-*almost everywhere*. The explicit form of L in (2.10) under P^∞ is thus given by

$$L_t = \exp\left(\int_0^t \frac{\mu_1(X_s) - \mu_0(X_s)}{\sigma}\, dB_s - \frac{1}{2}\int_0^t \left(\frac{\mu_1(X_s) - \mu_0(X_s)}{\sigma}\right)^2 ds \right). \tag{2.23}$$

By applying (2.12), the explicit form of L in (2.23) can be reformulated as

$$L_t = \exp\left(\int_0^t \frac{\mu_1(X_s) - \mu_0(X_s)}{\sigma^2(X_s)}\, dX_s - \frac{1}{2}\int_0^t \frac{\mu_1^2(X_s) - \mu_0^2(X_s)}{\sigma^2(X_s)}\, ds \right) \tag{2.24}$$

for $t \ge 0$. The above arguments can be applied to one-dimensional quickest detection problems of which the underlying process X satisfies the conditions that $[\mu_0(X_t) - \mu_1(X_t) + \sigma(X_t)\, c_t]$ is continuous and $\sigma(X_t)$ is a constant for $t \ge 0$ with the previsible process $c = (c_t)_{t\ge 0}$ taking values in $I\!R$. A more general argument for the formulation in (2.23) is performed in [18, Section 7.5] . We are not able to verify that (2.23) satisfies the Novikov's condtion nor the Kazamaki's condition (see [25]) which motivates the above arguments. For the Novikov's condition, by (2.8), we have attempted

$$\begin{aligned} \mathsf{E}^\infty\left[\exp\left(\frac{1}{2}\int_0^t \left[\frac{\mu_1(X_u) - \mu_0(X_u)}{\sigma(X_u)} \right]^2 du \right) \right] &= \mathsf{E}^\infty\left[\exp\left(\frac{1}{2}\int_0^t [\delta X_s + \gamma]^2\, ds \right) \right] \\ &\le \int_0^t \mathsf{E}^\infty \left[\exp\left(\frac{t}{2} (\delta X_s + \gamma)^2 \right) \right] \frac{ds}{t} \end{aligned} \tag{2.25}$$

where Jensen's inequality and Fubini's theorem are used in the last inequality in (2.25). X is known to be normally distributed and thus the expectation in (2.21) is only bounded for small t. For the Kazamaki's condition, since $\int_0^t (\mu_1(X_s) - \mu_0(X_s))\,/\sigma(X_s)\, dB_s$ is pure stochastic, we are not able to verify that $\exp\left(\frac{1}{2}\int_0^t (\mu_1(X_s) - \mu_0(X_s))\,/\sigma(X_s)\, dB_s \right)$ is uniformly integrable.

8. The *posterior probability ratio* process $\varPhi = (\varPhi_t)_{t\ge 0}$ of θ given X is defined by

$$\varPhi_t = \frac{\varPi_t}{1 - \varPi_t}. \tag{2.26}$$

Making use of (2.1), $\varPhi$ in (2.26) can be reformulated as follows

$$
\begin{aligned}
\varPhi_t &= \frac{\mathsf{P}_\pi(\theta \le t \mid \mathcal{F}_t^X)}{\mathsf{P}_\pi(\theta > t \mid \mathcal{F}_t^X)} \\
&= \frac{\pi \mathsf{P}^0(\theta \le t \mid \mathcal{F}_t^X)\frac{d\mathsf{P}_t^0}{d\mathsf{P}_{\pi,t}} + (1-\pi)\int_0^\infty \lambda e^{-\lambda s}\mathsf{P}^s(\theta \le t \mid \mathcal{F}_t^X)\frac{d\mathsf{P}_t^s}{d\mathsf{P}_{\pi,t}}\,ds}{\pi \mathsf{P}^0(\theta > t \mid \mathcal{F}_t^X)\frac{d\mathsf{P}_t^0}{d\mathsf{P}_{\pi,t}} + (1-\pi)\int_0^\infty \lambda e^{-\lambda s}\mathsf{P}^s(\theta > t \mid \mathcal{F}_t^X)\frac{d\mathsf{P}_t^s}{d\mathsf{P}_{\pi,t}}\,ds} \\
&= \frac{\pi \frac{d\mathsf{P}_t^0}{d\mathsf{P}_{\pi,t}} + (1-\pi)\int_0^t \lambda e^{-\lambda s}\frac{d\mathsf{P}_t^s}{d\mathsf{P}_{\pi,t}}\,ds}{(1-\pi)\int_t^\infty \lambda e^{-\lambda s}\frac{d\mathsf{P}_t^\infty}{d\mathsf{P}_{\pi,t}}\,ds} \\
&= \frac{\pi \frac{d\mathsf{P}_t^0}{d\mathsf{P}_{\pi,t}} + (1-\pi)\int_0^t \lambda e^{-\lambda s}\frac{d\mathsf{P}_t^s}{d\mathsf{P}_{\pi,t}}\,ds}{(1-\pi)e^{-\lambda t}\frac{d\mathsf{P}_t^\infty}{d\mathsf{P}_{\pi,t}}} \\
&= \frac{\pi e^{\lambda t}}{1-\pi}\frac{d\mathsf{P}_t^0}{d\mathsf{P}_t^\infty} + e^{\lambda t}\int_0^t \lambda e^{-\lambda s}\frac{d\mathsf{P}_t^s}{d\mathsf{P}_t^0}\frac{d\mathsf{P}_t^0}{d\mathsf{P}_t^\infty}\,ds \\
&= e^{\lambda t}L_t\left(\varPhi_0 + \int_0^t \lambda e^{-\lambda s}\frac{d\mathsf{P}_t^s}{d\mathsf{P}_t^0}\,ds\right). \qquad (2.27)
\end{aligned}
$$

By the similar arguments we performed for (2.23) above, under P^0,

$$
\begin{aligned}
\frac{d\mathsf{P}_t^s}{d\mathsf{P}_t^0} &= \exp\Bigg(- \int_0^t \frac{1}{\sigma(X_u)}\left[\mu_1(X_u) - \mu_0(X_u) - \mathrm{I}(u \ge s)\left(\mu_1(X_u) - \mu_0(X_u)\right)\right] dB_u \\
&\qquad - \frac{1}{2}\int_0^t \left(\frac{1}{\sigma(X_u)}\left[\mu_1(X_u) - \mu_0(X_u) - \mathrm{I}(u \ge s)\left(\mu_1(X_u) - \mu_0(X_u)\right)\right]\right)^2 du\Bigg) \\
&= \exp\left(-\int_0^s \frac{\mu_1(X_u)-\mu_0(X_u)}{\sigma(X_u)}\,dB_u - \frac{1}{2}\int_0^s \left[\frac{\mu_1(X_u)-\mu_0(X_u)}{\sigma(X_u)}\right]^2 du\right) \\
&= \exp\Bigg(- \int_0^s \frac{\mu_1(X_u)-\mu_0(X_u)}{\sigma(X_u)}\left[\frac{dX_u}{\sigma\left(X_u\right)} - \frac{\mu_1(X_u)}{\sigma(X_u)}du\right] \\
&\qquad - \frac{1}{2}\int_0^s \left[\frac{\mu_1(X_u)-\mu_0(X_u)}{\sigma(X_u)}\right]^2 du\Bigg) \\
&= \exp\left(-\int_0^s \frac{\mu_1(X_u)-\mu_0(X_u)}{\sigma^2(X_u)}\,dX_u + \frac{1}{2}\int_0^s \frac{\mu_1^2(X_u)-\mu_0^2(X_u)}{\sigma^2(X_u)}\,du\right) = \frac{1}{L_s} \qquad (2.28)
\end{aligned}
$$

for $0 \le s \le t$. Substituting the result from (2.28) into (2.27) we obtain

$$
\varPhi_t = e^{\lambda t}L_t\left(\varPhi_0 + \int_0^t \lambda e^{-\lambda s}\frac{1}{L_s}\,ds\right) \qquad (2.29)
$$

for $t \ge 0$ and $\varPhi_0 = \pi/\left(1-\pi\right)$.

9. By Itô's formula, we know that L satisfies the stochastic differential equation

$$
dL_t = \frac{\mu_1(X_t)-\mu_0(X_t)}{\sigma^2(X_t)}L_t\left[dX_t - \mu_0(X_t)\,dt\right] \qquad (2.30)
$$

with $L_0 = 1$. The stochastic differential equations that Π and Φ satisfy are given by

$$d\Pi_t = \lambda\,(1-\Pi_t)\,dt + \rho(X_t)\,\Pi_t\,(1-\Pi_t)\,d\bar{B}_t \tag{2.31}$$

$$d\Phi_t = \left[\lambda\,(1+\Phi_t) + \rho^2(X_t)\,\frac{\Phi_t^2}{1+\Phi_t}\right]dt + \rho(X_t)\,\Phi_t\,d\bar{B}_t \tag{2.32}$$

for X satisfying the stochastic differential equation

$$dX_t = [\mu_0(X_t) + \Pi_t\,(\mu_1(X_t) - \mu_0(X_t))]\,dt + \sigma(X_t)\,d\bar{B}_t \tag{2.33}$$

where $\bar{B} = (\bar{B}_t)_{t\geq 0}$ is the *innovation process* defined by

$$\bar{B}_t = \int_0^t \frac{dX_s}{\sigma(X_s)} - \int_0^t \left[\frac{\mu_0\,(X_s)}{\sigma(X_s)} + \Pi_s\,\frac{\mu_1(X_s)-\mu_0\,(X_s)}{\sigma(X_s)}\right]ds \tag{2.34}$$

for $t \geq 0$ and $x \in I\!R$. By Lévy's characterisation theorem we know $\bar{B}$ is a standard Brownian motion with respect to $(\mathcal{F}_t^X)_{t\geq 0}$ under probability measure P_π for $\pi \in [0,1]$.

2.3 Measure change

In this section, we simplify the optimal stopping problem to its *Lagrange formulation* by changing the probability measure from P_π to P^∞. The optimal stopping problem correspondingly becomes two-dimensional. Similar change of probability measures was introduced in [9, Section 4] and [14, Section 4].

1. By [14, Lemma 1] it is identified that the Radon-Nikodym derivative of the probability measure change from P_π to P^∞ is given by

$$\frac{d\mathsf{P}_{\pi,\tau}}{d\mathsf{P}^\infty_\tau} = e^{-\lambda\tau}\,\frac{1-\pi}{1-\Pi_\tau} \tag{2.35}$$

for all stopping times τ of X with $\pi \in [0,1)$ where P^∞_τ and $\mathsf{P}_{\pi,\tau}$ are the restrictions of the probability measures P^∞ and P_π to $\mathcal{F}_\tau^X$ respectively. Under P^∞ the two-dimensional stochastic process (Φ, X) satisfies the system of stochastic differential equations

$$d\Phi_t = \lambda\,(1+\Phi_t)\,dt + \rho(X_t)\,\Phi_t\,dB_t \tag{2.36}$$

$$dX_t = \mu_0(X_t)\,dt + \sigma(X_t)\,dB_t \tag{2.37}$$

for $t \geq 0$. The stochstic differential equation (2.30) is reduced to

$$dL_t = \rho(X_t)\,L_t\,dB_t. \tag{2.38}$$

The explicit expression of L in terms of X is now given by

$$L_t = \exp\left(\int_0^t \rho(X_s)\, dB_s - \frac{1}{2}\int_0^t \rho^2(X_s)\, ds\right) \tag{2.39}$$

for $t \geq 0$.

2. Under P^∞ the optimal stopping problem in (2.6) can be reformulated in terms of (Φ, X) in (2.36)+(2.37) where the starting point of Φ is $\pi/(1-\pi)$. We denote the flow by $\Phi^{\pi/(1-\pi)}$. By [14, Proposition 2] it is identified that the value function V in (2.6) satisfies the identity

$$V(\pi) = (1-\pi)\left[1 + c\hat{V}(\pi)\right] \tag{2.40}$$

where $\hat{V}$ is the equivalent optimal stopping problem defined by

$$\hat{V}(\pi) = \inf_{\tau \geq 0} \mathsf{E}^\infty \left[\int_0^\tau e^{-\lambda t}\left(\Phi_t^{\pi/(1-\pi)} - \frac{\lambda}{c}\right) dt\right] \tag{2.41}$$

for $\pi \in [0,1)$ and all stopping times τ of X.

3. Since the system of stochastic differential equations (2.36)+(2.37) has a unique weak solution, then (Φ, X) is a two-dimensional strong Markov process (cf. [26, pp. 166-173]) under P^∞. We let the two-dimensional strong Markov process (Φ, X) start at any point $(\varphi, x) \in [0,\infty)\times I\!R$. Under $\mathsf{P}^\infty_{\varphi,x}$ the optimal stopping problem (2.41) becomes

$$\hat{V}(\varphi, x) = \inf_{\tau \geq 0} \mathsf{E}^\infty_{\varphi,x} \left[\int_0^\tau e^{-\lambda t}\left(\Phi_t - \frac{\lambda}{c}\right) dt\right] \tag{2.42}$$

for $\varphi \in [0,\infty)\times I\!R$, $\mathsf{P}^\infty_{\varphi,x}((\Phi_0, X_0) = (\varphi, x)) = 1$ and all stopping times τ of (Φ, X). The value function $\hat{V}$ is non-positive since we can choose to stop at once and set $\tau = 0$. We have reduced the quickest detection problem in (2.4) to the optimal stopping problem in (2.42) for the strong Markov process (Φ, X) satisfying the system of stochastic differential equations

$$d\Phi_t = \lambda\,(1+\Phi_t)\, dt + (\delta X_t + \gamma)\,\Phi_t\, dB_t \tag{2.43}$$

$$dX_t = \beta_0\,(x_0 - X_t)\, dt + \sigma\, dB_t \tag{2.44}$$

under P^∞ with $(\varphi, x) \in [0,\infty)\times I\!R$.

2.4 Reduction to the canonical process

Under $\mathsf{P}^{\infty}_{\varphi,x}$ with $(\varphi, x) \in [0,\infty) \times I\!R$, the two-dimensional strong Markov process (Φ, X) satisfies the system of stochastic differential equations (2.43)+(2.44). Without loss of generality we set $\sigma = 1$ because we can replace the underlying Ornstein-Uhlenbeck process X by X/σ. For full generality we consider the case when $x_0 = 0$ and $x_1 \neq 0$. The system of stochastic differential equations (2.43)+(2.44) is simplified to

$$d\Phi_t = \lambda\,(1+\Phi_t)\,dt + (\delta X_t + \gamma)\,\Phi_t\,dB_t \tag{2.45}$$

$$dX_t = -\beta_0\,X_t\,dt + dB_t \tag{2.46}$$

with $(\varphi, x) \in [0,\infty) \times I\!R$, $\delta = \beta_0 - \beta_1$ and $\gamma = \beta_1\,x_1$ under $\mathsf{P}^{\infty}_{\varphi,x}$. The underlying Ornstein-Uhlenbeck process X is on the diffusion coefficient in (2.45) which makes the application of comparison theorems difficult. In this section, we reduce the system of stochastic differential equations (2.45)+(2.46) to its canonical form and derive the corresponding transformed optimal stopping problem.

1. The infinitesimal generator of (Φ, X) in (2.45)+(2.46) is given by

$$I\!L_{\Phi,X} = \lambda\,(1+\varphi)\,\partial_\varphi - \beta_0\,x\,\partial_x + (\delta x+\gamma)\,\varphi\,\partial_{\varphi x} + \frac{1}{2}\,(\delta x+\gamma)^2\,\varphi^2\,\partial_{\varphi\varphi} + \frac{1}{2}\,\partial_{xx}. \tag{2.47}$$

We denote the coefficients in (2.47) by

$$a\,(\varphi, x) = \frac{1}{2}\,(\delta x+\gamma)^2\,\varphi^2 \quad \& \quad 2\,b\,(\varphi, x) = (\delta x+\gamma)\,\varphi \quad \& \quad c\,(\varphi, x) = \frac{1}{2} \tag{2.48}$$

with $(\varphi, x) \in [0,\infty) \times I\!R$ given and fixed. The infinitesimal generator $I\!L_{\Phi,X}$ is of *parabolic* type because

$$b^2\,(\varphi, x) - a\,(\varphi, x)\,c\,(\varphi, x) = \frac{1}{4}\,(\delta x + \gamma)^2\,\varphi^2 - (\delta x + \gamma)^2\,\varphi^2\frac{1}{4} = 0. \tag{2.49}$$

Therefore, it has a unique family of characteristic curves given by

$$\frac{dx}{d\varphi} = \frac{b\,(\varphi, x)}{a\,(\varphi, x)} = \frac{1}{(\delta x+\gamma)\,\varphi}. \tag{2.50}$$

Solving the separable ordinary differential equation (2.50) we obtain

$$(\delta x + \gamma)\,dx = \frac{1}{\varphi}\,d\varphi \quad \Leftrightarrow \quad \frac{\delta}{2}x^2 + \gamma\,x = \log\varphi + u \tag{2.51}$$

for a constant $u \in I\!R$. Setting $\xi(\varphi, x) = u = \frac{\delta}{2}x^2 + \gamma\,x - \log\varphi$ and $\eta(\varphi, x) = x$, the Jacobian is given by

$$J = \frac{\partial\,(\xi, \eta)}{\partial\,(\varphi, x)} = \xi_\varphi\,\eta_x - \xi_x\,\eta_\varphi = -\frac{1}{\varphi} \notin \{0, \infty\} \tag{2.52}$$

for $\varphi \neq 0$ which indicates that the inverse function theorem is applicable. The coefficients of (2.47) in the canonical form are given by

$$\partial_\varphi = \xi_\varphi\, \partial_\xi + \eta_\varphi\, \partial_\eta = -\frac{1}{\varphi}\, \partial_\xi \tag{2.53}$$

$$\partial_x = \xi_x\, \partial_\xi + \eta_x\, \partial_\eta = (\delta x + \gamma)\, \partial_\xi + \partial_\eta \tag{2.54}$$

$$\partial_{\varphi\varphi} = \xi_{\varphi\varphi}\, \partial_\xi + \xi_\varphi^2\, \partial_{\xi\xi} + 2\, \xi_\varphi\, \eta_\varphi\, \partial_{\xi\eta} + \eta_{\varphi\varphi}\, \partial_\eta + \eta_\varphi^2\, \partial_{\eta\eta} = \frac{1}{\varphi^2}\, (\partial_\xi + \partial_{\xi\xi}) \tag{2.55}$$

$$\begin{aligned} \partial_{xx} &= \xi_{xx}\, \partial_\xi + \xi_x^2\, \partial_{\xi\xi} + 2\, \xi_x\, \eta_x\, \partial_{\xi\eta} + \eta_{xx}\, \partial_\eta + \eta_x^2\, \partial_{\eta\eta} \\ &= \delta\, \partial_\xi + (\delta x{+}\gamma)^2\, \partial_{\xi\xi} + 2\, (\delta x{+}\gamma)\, \partial_{\xi\eta} + \partial_{\eta\eta} \end{aligned} \tag{2.56}$$

$$\begin{aligned} \partial_{\varphi x} &= \xi_{\varphi x}\, \partial_\xi + \xi_\varphi\, \xi_x\, \partial_{\xi\xi} + \xi_\varphi\, \eta_x\, \partial_{\xi\eta} + \eta_{\varphi x}\, \partial_\eta + \eta_\varphi\, \eta_x\, \partial_{\eta\eta} + \eta_\varphi\, \xi_x\, \partial_{\varphi x} \\ &= -\frac{1}{\varphi}\, (\delta x + \gamma)\, \partial_{\varphi\eta} - \frac{1}{\varphi}\, \partial_{\xi\eta}. \end{aligned} \tag{2.57}$$

Substituting the above expressions into (2.47) we obtain the canonical infinitesimal generator

$$I\!L_{U,X} = \left[\frac{\delta}{2}\, (1{-}\kappa\, x^2) - \lambda\, \Big(1{+}e^{u - \frac{\delta}{2} x^2 - \gamma\, x}\Big) - \frac{\beta_1^2\, x_1}{2}\, x + \frac{\gamma^2}{2}\right] \partial_u - \beta_0\, x\, \partial_x + \frac{1}{2}\, \partial_{xx} \tag{2.58}$$

for $\kappa = \beta_0 {+} \beta_1$ and $\delta = \beta_0 {-} \beta_1$. From (2.51) we know that the process U satisfies the identity

$$U_t = \frac{\delta}{2} X_t^2 + \gamma X_t - \log \varPhi_t \tag{2.59}$$

for $t \geq 0$. The canonical process (U, X) satisfies the system of stochastic differential equations

$$dU_t = \left[\frac{\delta}{2}\, (1{-}\kappa X_t^2) - \lambda\, \Big(1{+}e^{U_t - \frac{\delta}{2} X_t^2 - \gamma X_t}\Big) - \frac{\beta_1^2 x_1}{2} X_t + \frac{\gamma^2}{2}\right] dt \tag{2.60}$$

$$dX_t = -\beta_0\, X_t\, dt + dB_t \tag{2.61}$$

under $\mathsf{P}_{u,x}^\infty$ with $\mathsf{P}_{u,x}^\infty\, ((U_0, X_0) = (u, x)) = 1$ for $(u, x) \in I\!R \times I\!R$. Note that U is a process of bounded variation. The system of stochastic differential equations (2.60)+(2.61) has a unique weak solution (cf. [26, pp. 166-173]) which indicates that the process (U, X) in (2.60)+(2.61) is a strong Markov process (cf. [26, pp. 158-163]) under $\mathsf{P}_{u,x}^\infty$ with $(u, x) \in I\!R \times I\!R$. The underlying Ornstein-Uhlenbeck process X in (2.61) has a unique strong solution given by

$$X_t^x = e^{-\beta_0 t} \left(x + \int_0^t e^{\beta_0 s}\, dB_s\right) \tag{2.62}$$

for $t \geq 0$ where the superscript $x \in I\!R$ indicates the initial point under P^∞. The unique strong solution in (2.62) defines a Markovian flow. The above arguments for reducing the strong Markov process (Φ, X) to its canonical form is analytic. Probabilistic arguments for the same reduction to *canonical process* were introduced in [14, Proposition 4].

2. The bounded variation U can be expressed in terms of the underlying Ornstein-Uhlenbeck process X. We define functions f and g by

$$f(x) = \frac{\delta}{2}\left(1 - \kappa\, x^2\right) - \frac{\beta_1^2\, x_1}{2}\, x + \frac{\gamma^2}{2} - \lambda \quad \& \quad g(x) = \lambda\, e^{-\frac{\delta}{2}\, x^2 - \frac{\gamma^2}{2}\, x} \tag{2.63}$$

for $x \in I\!R$. (2.60) can be rewritten as

$$\frac{dU_t}{dt} = f(X_t) - g(X_t)\, e^{U_t}. \tag{2.64}$$

Setting $R_t = e^{U_t}$ then (2.64) can be rewritten as

$$\frac{dR_t}{dt} = f(X_t) R_t - g(X_t) R_t^2 \tag{2.65}$$

which is a Bernoulli equation and can be solved explicitly. Setting $S_t = 1/R_t$ then (2.65) becomes

$$\frac{dS_t}{dt} = g(X_t) - f(X_t)\, S_t. \tag{2.66}$$

Using the integrating factor $I(t) := \exp\left(\int_0^t f(X_s)\, ds\right)$ for $t \geq 0$, we obtain the general solution to (2.64)

$$U_t = \int_0^t f(X_s)\, ds - \log\left[\int_0^t e^{\int_0^s f(X_r)\, dr} g(X_s)\, ds + e^{-U_0}\right]. \tag{2.67}$$

3. We define the two-dimensional process (V, Z) by

$$(V_t, Z_t) = \left(e^{-U_t}, X_t^2\right) \tag{2.68}$$

for $t \geq 0$. (2.67) can be rewritten as

$$V_t = e^{-\int_0^t f\left(\sqrt{Z_s}\right) ds}\left[\int_0^t e^{\int_0^s f\left(\sqrt{Z_r}\right) dr} g\left(\sqrt{Z_s}\right) ds + V_0\right] \tag{2.69}$$

for $t \geq 0$. With X defined in (2.61), by Itô's formula, we obtain

$$Z_t = Z_0 + \int_0^t (1 - 2\,\beta_0\, Z_s)\, ds + 2\int_0^t \sqrt{Z_s}\, d\tilde{B}_s \tag{2.70}$$

where $\tilde{B}_t := \int_0^t \text{sign}(X_s)\, dB_s$ for $t \geq 0$. By Lévy's characterisation theorem, we know that $\tilde{B}$ is a standard Brownian motion. From (2.62) we know that

$$Z_t^z = e^{-2\beta_0 t}\left(\sqrt{z} + \int_0^t e^{-\beta_0 s}\, dB_s\right)^2 \tag{2.71}$$

under P^∞. The superscript $z \in [0, \infty)$ denotes the initial point of Z. (2.71) defines a Markovian flow. From (2.60) and (2.70), we know that the process (V, Z) solves the system of stochastic differential equations

$$dV_t = \left[\left(\lambda - \frac{\delta}{2}(1 - \kappa Z_t) + \frac{\beta_1^2 x_1}{2}\sqrt{Z_t} - \frac{\gamma^2}{2}\right) V_t + \lambda e^{-\frac{\delta}{2} Z_t - \gamma\sqrt{Z_t}}\right] dt \tag{2.72}$$

$$dZ_t = (1 - 2\,\beta_0\, Z_t)\, dt + 2\sqrt{Z_t}\, d\tilde{B}_t \tag{2.73}$$

under $\mathsf{P}_{v,z}^\infty$ with $\mathsf{P}_{v,z}^\infty\left((V_0, Z_0) = (v, z)\right) = 1$ for $(v, z) \in [0, \infty) \times [0, \infty)$. The system of stochastic differential equations (2.72)+(2.73) has a unique weak solution (cf. [26, pp. 166-173]) which indicates that the process (V, Z) in (2.68) is a two-dimensional strong Markov process under $\mathsf{P}_{v,z}^\infty$. The infinitesimal generator $\mathbb{L}_{V,Z}$ is given by

$$\mathbb{L}_{V,Z} = \left[\left(\lambda - \frac{\delta}{2}(1 - \kappa z) + \frac{\beta_1^2 x_1}{2}\sqrt{z} - \frac{\gamma^2}{2}\right) v + \lambda e^{-\frac{\delta}{2} z - \gamma\sqrt{z}}\right] \partial_v + (1 - 2\,\beta_0\, z)\, \partial_z + 2\, z\, \partial_{zz}. \tag{2.74}$$

Proposition 1. *The value function* $\hat{V}$ *in* (2.42) *satisfies the identity*

$$\hat{V}(\varphi, x) = \tilde{V}\left(\varphi\, e^{-\frac{\delta}{2} x^2 - \gamma x}, x^2\right) \tag{2.75}$$

for $(\varphi, x) \in [0, \infty) \times \mathbb{R}$ *where* $\tilde{V}$ *is the equivalent value function defined by*

$$\tilde{V}(v, z) = \inf_{\tau \geq 0} \mathsf{E}_{v,z}^\infty \left[\int_0^\tau e^{-\lambda t}\left(V_t\, e^{\frac{\delta^2}{2} Z_t + \gamma\sqrt{Z_t}} - \frac{\lambda}{c}\right) dt\right] \tag{2.76}$$

with $(v, z) \in (0, \infty] \times (0, \infty]$ *for all stopping times* τ *of the strong Markov process (V, Z) solving the system of stochastic differential equations* (2.72)+(2.73) *where the explicit solutions are given in* (2.69)+(2.71).

Proof. By (2.59) and (2.68) we have $v = e^{-u} = \varphi\, e^{-\frac{\delta}{2} z - \gamma\sqrt{z}}$ since $z = x^2$. Substituting into (2.42) we obtain

$$\tilde{V}(v, z) = \inf_{\tau \geq 0} \mathsf{E}_{v,z}^\infty \left[\int_0^\tau e^{-\lambda t}\left(V_t\, e^{\frac{\delta^2}{2} Z_t + \gamma\sqrt{Z_t}} - \frac{\lambda}{c}\right) dt\right]. \tag{2.77}$$

□

This section is a generalization of [9, Section 5]. Probabilistic reduction to canonical form were introduced in [14, Section 6].

2.5 Time change

In this section we use time change technique to decouple the stochastic processes on the diffusion term and then derive the optimal stopping problem for the time-changed process.

1. We define an additive function $A = (A(t))_{t\geq 0}$ by

$$A(t) = \int_0^t (\delta X_s + \gamma)^2 \, ds \tag{2.78}$$

recalling $\delta = \beta_0 - \beta_1$ and $\gamma = \beta_1 x_1$. $A(t)$ is continuous and strictly increasing P^∞-almost surely. The inverse function of $A(t)$ defined by

$$T(t) = A^{-1}(t) = \inf\{s \geq 0 \mid A(s) > t\} \tag{2.79}$$

is also continuous and strictly increasing. Since $A(t)$ is $(\mathcal{F}_t^X)_{t\geq 0}$-measurable $T(t)$ is a stopping time of (Φ, X) with respect to $(\mathcal{F}_t^X)_{t\geq 0}$. Therefore, $T(t)$ defines a time change transformation with respect to $(\mathcal{F}_t^X)_{t\geq 0}$. The time-changed process $(\hat{\Phi}, \hat{X}) = ((\hat{\Phi}_t, \hat{X}_t))_{t\geq 0}$ is defined by

$$(\hat{\Phi}_t, \hat{X}_t) = (\Phi_{T(t)}, X_{T(t)}) \tag{2.80}$$

for $t \geq 0$ and $(\varphi, x) \in [0, \infty) \times I\!R$ under $\mathsf{P}^\infty_{\varphi,x}$. The infinitesimal generator of $(\hat{\Phi}, \hat{X})$ is given by

$$I\!L_{\hat{\Phi},\hat{X}} = \frac{1}{(\delta\, x + \gamma)^2} \, I\!L_{\Phi,X} \tag{2.81}$$

where $I\!L_{\Phi,X}$ is the infinitesimal generator of (Φ, X) defined in (2.47) above. The natural filtration generated by $(\hat{\Phi}, \hat{X})$ is equivalent to the time-changed natural filtration generated by (Φ, X), i.e. $\hat{\mathcal{F}}_t^{\Phi,X} = \mathcal{F}_{T(t)}^{\Phi,X}$ with $t \geq 0$. The explicit expression of $T(t)$ in (2.79) is given by

$$T(t) = \int_0^t \frac{1}{(\delta \hat{X}_s + \gamma)^2} \, ds \tag{2.82}$$

for $t \geq 0$.

2. By (2.45)+(2.46), the time-changed process $(\hat{\Phi}, \hat{X})$ satisfies the system of stochastic integral equations

$$\hat{\Phi}_t = \Phi_{T(t)} = \Phi_0 + \int_0^{T(t)} \lambda\,(1+\Phi_s)\,ds + \int_0^{T(t)} (\delta X_s + \gamma)\,\Phi_s\,dB_s$$
$$= \hat{\Phi}_0 + \int_0^t \lambda \frac{1+\hat{\Phi}_s}{(\delta \hat{X}_s+\gamma)^2}\,ds + \int_0^t \Big(\delta \hat{X}_s + \gamma\Big)\,\hat{\Phi}_s\,d\hat{B}_s$$
$$= \hat{\Phi}_0 + \int_0^t \lambda \frac{1+\hat{\Phi}_s}{(\delta \hat{X}_s+\gamma)^2}\,ds + \int_0^t \hat{\Phi}_s\,d\tilde{B}_s \tag{2.83}$$
$$\hat{X}_t = X_{T(t)} = X_0 + \int_0^{T(t)} -\beta_0\,X_s\,ds + \int_0^{T(t)} dB_s$$
$$= \hat{X}_0 - \int_0^t \frac{\beta_0\,\hat{X}_s}{(\delta \hat{X}_s+\gamma)^2}\,ds + \int_0^t \frac{1}{(\delta \hat{X}_s+\gamma)}\,d\tilde{B}_s \tag{2.84}$$

where the process $\tilde{B} = (\tilde{B}_t)_{t\geq 0}$ is defined by

$$\tilde{B}_t = \int_0^t \Big(\delta \hat{X}_s + \gamma\Big)\,d\hat{B}_s = \int_0^{T(t)} (\delta X_s + \gamma)\,dB_s = M_{T(t)}. \tag{2.85}$$

for $M_t = \int_0^t (\delta X_s + \gamma)\,dB_s$ with $t \geq 0$. $M = (M_t)_{t\geq 0}$ is a continuous local martingale with respect to $(\mathcal{F}_t^X)_{t\geq 0}$ which implies that $\tilde{B} = (\tilde{B}_t)_{t\geq 0}$ is a continuous local martingale with respect to $(\hat{\mathcal{F}}_t^X)_{t\geq 0}$. Further, $\langle \tilde{B}, \tilde{B}\rangle_t = \langle M, M\rangle_{T(t)} = \int_0^{T(t)} (\delta X_s+\gamma)^2\,ds = A(T(t)) = t$ for $t \geq 0$. By Lévy's characterisation theorem, $\tilde{B}$ is a standard Brownian motion with respect to $(\hat{\mathcal{F}}_t^X)_{t\geq 0}$ under $\mathsf{P}_{\varphi,x}^{\infty}$. (2.83)+(2.84) can be written as a system of stochastic differential equations

$$d\hat{\Phi}_t = \lambda \frac{1+\hat{\Phi}_t}{(\delta \hat{X}_t+\gamma)^2}\,dt + \hat{\Phi}_t\,d\tilde{B}_t \tag{2.86}$$
$$d\hat{X}_t = -\frac{\beta_0\,\hat{X}_t}{(\delta \hat{X}_t+\gamma)^2}\,dt + \frac{1}{(\delta \hat{X}_t+\gamma)}\,d\tilde{B}_t \tag{2.87}$$

with $t \geq 0$ under $\mathsf{P}_{\varphi,x}^{\infty}$ for $(\varphi, x) \in [0,\infty)\times I\!R$. Since (2.86)+(2.87) has a unique weak solution (cf. [26, pp. 166-173]), $(\hat{\Phi}, \hat{X})$ is a strong Markov process (cf. [26, pp. 158-163]) under $\mathsf{P}_{\varphi,x}^{\infty}$ for $(\varphi, x) \in [0,\infty)\times I\!R$. The time-changed process $\hat{L} = (L_{T(t)})_{t\geq 0}$ of L from (2.38)+(2.39) satisfies the stochastic differential equation

$$d\hat{L}_t = \hat{L}_t\,d\tilde{B}_t \tag{2.88}$$

where the explicit form of $\hat{L}$ is given by

$$\hat{L}_t = \exp\left(\tilde{B}_t - \frac{t}{2}\right) \tag{2.89}$$

for $t \geq 0$.

3. We now derive the time-changed version of the optimal stopping problem in (2.42).

Proposition 2. *The value function $\hat{V}$ defined in* (2.42) *satisfies the identity*

$$\hat{V}(\varphi, x) = \inf_{\sigma \geq 0} \mathsf{E}^{\infty}_{\varphi,x} \left[\int_0^{\sigma} e^{-\lambda \int_0^t (\delta \hat{X}_s + \gamma)^{-2}\, ds} \, \frac{\hat{\Phi}_t - \frac{\lambda}{c}}{(\delta \hat{X}_s + \gamma)^2} \, dt \right] \tag{2.90}$$

for $(\varphi, x) \in [0, \infty) \times I\!R$ *and all stopping times* σ *of* $(\hat{\Phi}, \hat{X})$.

Proof. $\tau = T(\sigma)$ is a stopping time of (Φ, X) if and only if $\sigma = A(\tau)$ is a stopping time of $(\hat{\Phi}, \hat{X})$. We can construct σ or τ if either τ or σ is known respectively. From (2.82) we obtain

$$\begin{aligned} \mathsf{E}^{\infty}_{\varphi,x} \left[\int_0^{\tau} e^{-\lambda t} \left(\Phi_t - \frac{\lambda}{c} \right) dt \right] &= \mathsf{E}^{\infty}_{\varphi,x} \left[\int_0^{T(\sigma)} e^{-\lambda t} \left(\Phi_t - \frac{\lambda}{c} \right) dt \right] \\ &= \mathsf{E}^{\infty}_{\varphi,x} \left[\int_0^{\sigma} e^{-\lambda T(t)} \left(\hat{\Phi}_t - \frac{\lambda}{c} \right) dT(t) \right] \\ &= \mathsf{E}^{\infty}_{\varphi,x} \left[\int_0^{\sigma} e^{-\lambda \int_0^t (\delta \hat{X}_s + \gamma)^{-2}\, ds} \, \frac{\hat{\Phi}_t - \frac{\lambda}{c}}{(\delta \hat{X}_s + \gamma)^2} \, dt \right]. \end{aligned} \tag{2.91}$$

The result follows by taking the infimum over all τ and σ on both sides. □

Similar arguments for the time change of Bessel processes were introduced in [14, Section 7].

2.6 Mayer formulation

We recall that the optimal stopping problem in (2.42) is Lagrange formulated. To derive the corresponding *Mayer* formulation (cf. [21, Section 7.5]) we need to find a function $\hat{M}(\varphi, x)$ which solves the partial differential equation

$$(I\!L_{\Phi,X} \hat{M} - \lambda \hat{M})(\varphi, x) = \varphi - \frac{\lambda}{c} \tag{2.92}$$

for $(\varphi, x) \in [0, \infty) \times I\!R$. By [9, Proposition 4] we know that if $\beta_0 \neq -\lambda/2$ and $\beta_1 = 0$, then the value function $\hat{V}$ in (2.42) can be reformualted as

$$\hat{V}(\varphi, x) = \inf_{\tau \geq 0} \mathsf{E}^{\infty}_{\varphi,x} \left[e^{-\lambda \tau} \hat{M}(\Phi_\tau, X_\tau) \right] - \hat{M}(\varphi, x) \tag{2.93}$$

with $(\varphi, x) \in [0, \infty) \times I\!R$ for all stopping times τ of (Φ, X) where $\hat{M}$ is given by

$$\hat{M}(\varphi, x) = \varphi x^2 + \left(x^2 + \frac{1}{\lambda}\right) / \left(1 + \frac{2\beta_0}{\lambda}\right) + \frac{1}{c} \tag{2.94}$$

for $(\varphi, x) \in [0, \infty) \times I\!R$. The transformation from Lagrange formulation to Mayer formulation is based on Itô's formula and the optional sampling theorem. The Mayer formulation is essential for the applications of *Wald-Bellman equations* (cf. [21, Section 1.2]). We note that $\hat{M}$ in (2.94) is not the unique solution of the partial differential equation (2.92).

Chapter 3

Free-boundary problem

In this chapter we verify that the value function $\hat{V}$ in (2.42) is a weak solution to a system of free-boundary problem where the derivatives are understood in the sense of Schwartz distribution (cf. [27]). We then verify that the weak solution satisfies Hörmander's condition (cf. [12]) which indicates that the value function is smooth and therefore a strong/classic solution. The verification of the Hörmander condition in this chapter is a generalization of the technique introduced in [9].

1. We consider the Lagrange formulated optimal stopping problem in (2.42). By analogous to Section 7 in [9], the continuation set C and stopping set D are defined by

$$C = \{ (\varphi, x) \in [0, \infty) \times I\!R \mid \hat{V}(\varphi, x) < 0 \} \tag{3.1}$$

$$D = \{ (\varphi, x) \in [0, \infty) \times I\!R \mid \hat{V}(\varphi, x) = 0 \}. \tag{3.2}$$

By Proposition 5 in [9] we conclude that the stopping set D is not empty which indicates that the incentive to continue observing is not endless. The first entry time of the process $(\varPhi, X)$ into the stopping set D defined by

$$\tau_D = \inf\{ t \geq 0 \mid (\varPhi_t, X_t) \in D \} \tag{3.3}$$

is optimal in (2.42) whenever $\mathsf{P}^{\infty}_{\varphi,x}(\tau_D < \infty) = 1$ for $(\varphi, x) \in [0, \infty) \times I\!R$. The optimal stopping boundary between C and D is defined by

$$b(x) = \inf\{ \varphi \geq 0 \mid (\varphi, x) \in D \} \tag{3.4}$$

for $x \in I\!R$ (cf. [9, Section 7]). The optimal stopping problem in (2.42) can be rewritten as

$$\hat{V}(\varphi, x) = \mathsf{E}^{\infty}_{\varphi,x}\left[\int_0^{\tau_D} e^{-\lambda t}\left(\Phi_t - \frac{\lambda}{c}\right)dt\right] \tag{3.5}$$

with $(\varphi, x) \in [0,\infty)\times I\!R$, $\lambda > 0$ and $c > 0$. By Corollary 6 in [20] the value function $\hat{V}$ solves the free-boundary problem

$$I\!L_{\Phi,X}\hat{V} - \lambda\hat{V} \overset{w}{=} -L \quad \text{in} \ \ C \tag{3.6}$$

$$\hat{V} = 0 \quad \text{on} \ \ D \ \ \text{(instaneous stopping)} \tag{3.7}$$

$$\hat{V} = 0 \quad \text{at} \ \ \partial C \tag{3.8}$$

where $L(\varphi) = \lambda/c - \varphi$ for $\varphi \in [0,\infty)$, $\partial C = \{\, (\varphi, x) \in [0,\infty)\times I\!R \mid \varphi = b(x)\,\}$ is the (optimal stopping) boundary between the sets C and D (cf. [21, Chapter 7]) and the derivatives in (3.6) are understood in the sense of Schwartz distributions, i.e. we have

$$\begin{aligned}\langle I\!L_{\Phi,X}\hat{V}, f\rangle &= \int_0^{\infty}\int_{-\infty}^{\infty}\Big(I\!L_{\Phi,X}\hat{V}(\varphi,x)\Big)\,f(\varphi,x)\,d\varphi\,dx\\ &= \int_0^{\infty}\int_{-\infty}^{\infty}\hat{V}(\varphi,x)\,I\!L^*_{\Phi,X}f(\varphi,x)\,d\varphi\,dx\\ &= \langle\hat{V}, I\!L^*_{\Phi,X}f\rangle\end{aligned} \tag{3.9}$$

for all $f \in C_c^\infty$. $C_c^\infty = C_c^\infty\left([0,\infty)\times I\!R\right)$ denotes the space of all infinitely differentiable functions from $[0,\infty)\times I\!R$ into $I\!R$ with compact supports contained in the interior of $[0,\infty)\times I\!R$ which indicates that the function f vanishes to 0 at the ending points for all $f \in C_c^\infty$. $I\!L^*_{\Phi,X}$ is the *adjoint* of the infinitesimal generator $I\!L_{\Phi,X}$ in (2.47) defined by

$$\begin{aligned}I\!L^*_{\Phi,X}f\,(\varphi,x) =& -\partial_\varphi\left(\lambda\,(1+\varphi)\,f\,(\varphi,x)\right) + \partial_x\left(\beta_0\,x\,f\,(\varphi,x)\right)\\ &+ \partial_{\varphi x}\left((\delta x+\gamma)\,\varphi\,f\,(\varphi,x)\right) + \partial_{\varphi\varphi}\left(\frac{1}{2}\,(\delta x+\gamma)^2\,\varphi^2\,f\,(\varphi,x)\right)\\ &+ \partial_{xx}\left(\frac{1}{2}\,f\,(\varphi,x)\right)\end{aligned} \tag{3.10}$$

derived through integration by parts.

2. The weak solution $\hat{V}$ in (3.5) solving the system (3.6)-(3.8) has an opportunity to upgrade into a strong (classic/smooth) if it is hypoelliptic, e.g. when $\hat{V}$ satisfies the Hörmander's condition (cf. [12]).

Proposition 3. *The value function* $\hat{V}$ *in* (3.4) *solving the system* (3.5)-(3.7) *is smooth which indicates that* $\hat{V} \in C^\infty$ *is infinitely differentiable.*

Proof. We recall the infinitesimal generator of the canonical process (U, X) in (2.58) can be written as

$$\mathbb{L}_{U,X} = a\,\partial_u - \beta_0\,x\,\partial_x + \frac{1}{2}\,\partial_{xx} \tag{3.11}$$

where the coefficient a is defined by

$$a := a(u,x) = \frac{\delta}{2}\left(1-\kappa\,x^2\right) - \lambda\left(1+e^{u-\frac{\delta}{2}x^2-\gamma\,x}\right) - \frac{\beta_1^2\,x_1}{2}\,x + \frac{\gamma^2}{2} \tag{3.12}$$

with $\kappa = \beta_0 + \beta_1$, $\delta = \beta_0 - \beta_1$, $u \in \mathbb{R}$, $x \in \mathbb{R}$, $x_1 \neq 0$, $\beta_0 \in \mathbb{R}$ and $\beta_1 \in \mathbb{R}$. We now verify that the infinitesimal generator $\mathbb{L}_{U,X}$ satisfies the Hörmander's condition (cf. [12]). The coefficients in (3.8) are smooth which belong to C^∞ on the continuation set C. The infinitesimal generator $\mathbb{L}_{U,X}$ can be written as

$$\mathbb{L}_{U,X} = D_0 + D_1^2 = D_0 + D_1\,D_1 \tag{3.13}$$

where D_0 and D_1 are first-order differential operators given by

$$D_0 = a\,\partial_u - \beta_0\,x\,\partial_x \tag{3.14}$$

$$D_1 = \frac{1}{\sqrt{2}}\,\partial_x. \tag{3.15}$$

The product $D_1\,D_1$ means the composition of first-order differential operators. The *Lie bracket* of D_0 and D_1 understood as first-order differential operators is given by

$$\begin{aligned}
[D_0, D_1] &= D_0\,D_1 - D_1\,D_0 = -[D_1, D_0] = [-D_1, D_0] = [D_1, -D_0] \\
&= \frac{a}{\sqrt{2}}\,\partial_u\,\partial_x - \frac{\beta_0\,x}{\sqrt{2}}\,\partial_x\,\partial_x - \frac{1}{\sqrt{2}}\,\partial_x\,(a\,\partial_u) + \frac{\beta_0}{\sqrt{2}}\,\partial_x\,(x\,\partial_x) \\
&= \frac{a}{\sqrt{2}}\,\partial_u\,\partial_x - \frac{\beta_0\,x}{\sqrt{2}}\,\partial_x\,\partial_x - \frac{1}{\sqrt{2}}\,((\partial_x\,a)\,\partial_u + a\,\partial_x\partial_u) \\
&\qquad + \frac{\beta_0}{\sqrt{2}}\,(\partial_x + x\,\partial_x\,\partial_x) \\
&= -\frac{1}{\sqrt{2}}\,(\partial_x\,a)\,\partial_u + \frac{\beta_0}{\sqrt{2}}\,\partial_x.
\end{aligned} \tag{3.16}$$

We refer to the Lie bracket $[D_i, D_j]$ of D_i and D_j as the commutator of step 1 for $i, j = 0, 1$. By analogy we obtain the Lie bracket of $[D_0, D_1]$ and D_1

$$\begin{aligned}
[[D_0, D_1], D_1] &= [D_0, D_1]\,D_1 - D_1\,[D_0, D_1] = -[[D_1, D_0], D_1] = [D_1, [D_1, D_0]] \\
&= \frac{1}{\sqrt{2}}\,\partial_x\left(-\frac{1}{\sqrt{2}}\,(\partial_x\,a)\,\partial_u + \frac{\beta_0}{\sqrt{2}}\,\partial_x\right) - \frac{1}{2}\,(\partial_x\,a)\,\partial_u\,\partial_x + \frac{\beta_0}{2}\,\partial_x\,\partial_x \\
&= \frac{1}{2}\left((\partial_x^2\,a)\,\partial_u + (\partial_x\,a)\,\partial_x\,\partial_u - \beta_0\,\partial_x^2\right) - \frac{1}{2}\,(\partial_x\,a)\,\partial_u\,\partial_x + \frac{\beta_0}{2}\,\partial_x^2 \\
&= \frac{1}{2}\left(\partial_x^2\,a\right)\partial_u.
\end{aligned} \tag{3.17}$$

We refer to the Lie bracket $[[D_i, D_j], D_k]$ as the commutator of step 2 for $i, j, k = 0, 1$. Inductively we obtain the commutator of step n given by

$$[D_1, [D_1, ...[D_1, [D_1, D_0]]...]] = 2^{-\frac{n}{2}} \left(\partial_x^n a\right) \partial_u \tag{3.18}$$

for $n \geq 2$. The first order differential operators D_0, D_1 and their commutator of any finite step n can be identified with the coefficients as vector fields

$$D_0 \sim (a(u, x), -\beta_0 \, x) \tag{3.19}$$

$$D_1 \sim \left(0, \frac{1}{\sqrt{2}}\right) \tag{3.20}$$

$$[D_1, D_0] \sim \left(\frac{\partial_x \, a(u, x)}{\sqrt{2}}, -\frac{\beta_0}{\sqrt{2}}\right) \tag{3.21}$$

$$[D_1, [D_1, D_0]] \sim \left(\frac{1}{2}\partial_x^2 \, a(u, x), 0\right) \tag{3.22}$$

$$\vdots$$

$$[D_1, [D_1, ...[D_1, [D_1, D_0]]...]] \sim \left(2^{-\frac{n}{2}} \partial_x^n \, a(u, x), 0\right). \tag{3.23}$$

The *Lie algebra* generated by the vector fields D_0 and D_1 is the set of linear combination of the vector fields D_0, D_1 and their commutators of any finite step which is given by

$$Lie\,(D_0, D_1) = \text{span}\{\, D_i, [D_i, D_j], [[D_i, D_j], D_k], ... \mid i, j, k, ... = 0, 1 \,\}. \tag{3.24}$$

We see that $\dim Lie\,(D_0, D_1) = 2$ and the Hörmander's condition (cf. [12]) is satisfied for $\mathbb{L}_{U,X}$ because $\partial_x^n \, a(u, x)$ are not 0 at the same time for $n \geq 0$ with $\partial_x^0 \, a(u, x) = a(u, x)$ otherwise we would have $a(u_0, x) = 0$ for $x \in \mathbb{R}$ with $u_0 \in \mathbb{R}$ given and fixed by Taylor expansion. Recalling (2.59) is a smooth diffeomorphism $u(\varphi, x) := \frac{\delta}{2} \, x^2 + \gamma \, x - \log \varphi$, combining Corollary 8 in [20], we conclude that $\hat{V} \in C^\infty$ is infinitely differentiable. $\square$

We note that $\hat{V}$ in (3.4) is a strong (classic/smooth) solution to the system (3.6)-(3.8) even if it is not the unique solution. The derivatives in (3.12) are now understood in the classic sense where

$$\mathbb{L}_{\Phi,X} \hat{V} - \lambda \hat{V} = - \, L \quad \text{in} \;\; C. \tag{3.25}$$

A general introduction to Hörmander's condition could be found in [4, Chapter 1-2].

Chapter 4

General Wald-Bellman equations

In this chapter we recall the basic setting of time-homogeneous and time-inhomogeneous Wald-Bellman equations. We then introduce an upper bound on the rates of convergence in the time-homogeneous Wald-Bellman equations in each iteration. We finally introduce a technique of constructing the value functions which enables the applications of Wald-Bellman equations to finite-horizon continuous-time Mayer formulated optimal stopping problems.

1. Consider a continuous time-homogeneous Markov process $X = (X_t)_{t\geq 0}$ defined on a filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, \mathsf{P}_x)$ and taking values in $I\!R^d$ for $d \geq 1$. We assume that the process X starts at x under P_x for $x \in I\!R^d$. It is also assumed that the mapping $x \to \mathsf{P}_x(F)$ is measurable for each $F \in \mathcal{F}$. It follows that the mapping $x \to \mathsf{E}_x(Z)$ is measurable for each random variable Z.

Given a measurable function $G : I\!R^d \to I\!R$ satisfying the condition:

$$\mathsf{E}_x\left(\sup_{0\leq t<\infty} |G(X_t)|\right) < \infty \tag{4.1}$$

for all $x \in I\!R^d$, we consider the optimal stopping problem

$$V(x) = \sup_{0\leq\tau<\infty} \mathsf{E}_x G(X_\tau). \tag{4.2}$$

The corresponding continuation set C and stopping set D are given by

$$C = \{\, x \in I\!R^d \mid V(x) > G(x) \,\} \tag{4.3}$$

$$D = \{\, x \in I\!R^d \mid V(x) = G(x) \,\}. \tag{4.4}$$

The transition operator $\mathbb{P}_t$ of X is defined by

$$\mathbb{P}_t G(x) = \mathsf{E}_x G(X_t) = \int_{I\!R^d} G(y)\, p(t;x;y)\, dy \tag{4.5}$$

for $0 \leq t < \infty$ and $y \in I\!R^d$. The transition density function p is the unique non-negative solution to the Kolmogorov backward equation

$$\frac{\partial}{\partial t} p(t;x;y) = I\!L_X(p)(t;x;y) \tag{4.6}$$

$$p(0+;x;y) = \delta_x(y) \quad (weakly) \tag{4.7}$$

where $\int_{I\!R^d} p(t;x;y)\, dy = 1$ for $t > 0$ with x and y in $I\!R^d$ and δ_x is the Dirac measure at x. $I\!L_X$ is the infinitesimal generator of X given by

$$I\!L_X f(x) = \frac{1}{2} \sum_{i=1}^{d} \sum_{k=1}^{d} a_{ik}(x) \frac{\partial^2 f(x)}{\partial x_i \partial x_k} + \sum_{i=1}^{d} b_i(x) \frac{\partial f(x)}{\partial x_i} \tag{4.8}$$

for twice continuously differentiable function $f : I\!R^d \to I\!R$ where

$$a_{ik}(x) = \sum_{j=1}^{r} \sigma_{ij}(x) \sigma_{kj}(x). \tag{4.9}$$

for a measurable function $\sigma : I\!R^d \to I\!R^{d\times r}$ with $1 \leq i \leq d$, $1 \leq j \leq r$ and $r \geq 1$. The Wald-Bellman equations (cf. [29, Chapter 3]) satisfy the identity

$$V(x) = \lim_{n\to\infty} \lim_{N\to\infty} V_n^N(x) \tag{4.10}$$

where we set

$$V_n^N(x) = \sup_{\tau \in \mathcal{T}_n^N} \mathsf{E}_x G(X_\tau) = \max\big(G(x), \mathbb{P}_{2^{-n}} V_n^{N-1}(x)\big) \tag{4.11}$$

with $V_n^0(x) = G(x)$, $n \in I\!N$, $N \in I\!N$ and $\mathcal{T}_n^N = \{0, 2^{-n}, 2\times 2^{-n}, ..., N \times 2^{-n}\}$.

2. Now consider a continuous time-inhomogeneous optimal stopping problem (cf. [21, Section 1.2] and [21, Section 2.2]). Given a measurable function $G : [0,T) \times I\!R^d \to I\!R$ satisfying the condition:

$$\mathsf{E}_{t,x}\left(\sup_{0\leq k \leq T-t} |G(t+k, X_{t+k})| \right) < \infty \tag{4.12}$$

for all $t \in [0,T)$ and $x \in I\!R^d$, we consider the optimal stopping problem

$$V^T(t,x) = \sup_{0\leq \tau \leq T-t} \mathsf{E}_{t,x} G(t+\tau, X_{t+\tau}). \tag{4.13}$$

The continuation set C and stopping set D (cf. [21, Chapter I, 1.2]) are given by

$$C = \{ (t,x) \in [0,T) \times I\!R^d \mid V^T(t,x) > G(t,x) \} \tag{4.14}$$

$$D = \{ (t,x) \in [0,T) \times I\!R^d \mid V^T(t,x) = G(t,x) \}. \tag{4.15}$$

The time-inhomogeneous Wald-Bellman equations (cf. [21, Theorem 1.9]) are defined by

$$V_n^N(t,x) = \sup_{\tau \in \mathcal{T}_n^{N-t2^n}} \mathsf{E}_{t,x} G(t+\tau, X_{t+\tau}) = \max\left(G(t,x), \mathbb{P}_{2^{-n}} V_n^N(t,x)\right) \tag{4.16}$$

for $t \in \mathcal{T}_n^{N-1}$, $x \in I\!R^d$, $n \geq 0$, $N \in I\!N$, $\mathcal{T}_n^{N-1} = \{ 0, 2^{-n}, 2 \times 2^{-n}, ..., (N-1) \times 2^{-n} \}$ and $\mathcal{T}_n^{N-t2^n} = \{ 0, 2^{-n}, 2 \times 2^{-n}, ..., (N - t2^n) \times 2^{-n} \}$ where $\mathbb{P}_{2^{-n}} V_n^N((N-1)2^{-n}, x) = \mathsf{E}_{(N-1)2^{-n},x} G(N2^{-n}, X_{N2^{-n}})$. $\mathbb{P}_{2^{-n}}$ is the transition operator defined by

$$\mathbb{P}_{2^{-n}} V_n^N(t,x) = \mathsf{E}_{t,x} V_n^N(t+2^{-n}, X_{t+2^{-n}}). \tag{4.17}$$

3. We now introduce an upper bound on the rates of convergence in the time-homogeneous Wald-Bellman equations (4.11) in each iteration

$$|V_n^N(x) - V_n^{N-1}(x)| = V_n^N(x) - V_n^{N-1}(x) \tag{4.18}$$

where $N \to V_n^N(x)$ is increasing.

Theorem 4. *The rates of convergence in the time-homogeneous Wald-Bellman equations* (4.11) *in each iteration are determined by*

$$|V_n^N(x) - V_n^{N-1}(x)| \leq \int_{I\!R^d} \left(\mathbb{P}_{2^{-n}} G(y) - G(y)\right)^+ p\left((N-1)2^{-n}; x; y\right) dy \tag{4.19}$$

with $V_n^0(x) = G(x)$, $x \in I\!R^d$, $n \geq 0$ *and* $N \in I\!N$.

Proof. Applying [29, Lemma 2.15] iteratively we obtain

$$\begin{aligned}
V_n^N(x) - V_n^{N-1}(x) &\leq \mathbb{P}_{2^{-n}} \left(V_n^{N-1}(x) - V_n^{N-2}(x)\right) \\
&\leq \mathbb{P}_{2^{-n}} \mathbb{P}_{2^{-n}} \left(V_n^{N-2}(x) - V_n^{N-3}(x)\right) \\
&= \mathbb{P}_{2 \times 2^{-n}} \left(V_n^{N-2}(x) - V_n^{N-3}(x)\right) \\
&\vdots \\
&\leq \mathbb{P}_{(N-1) \times 2^{-n}} \left(V_n^1(x) - G(x)\right) \\
&\leq \mathsf{E}_x \left[V_n^1\left(X_{(N-1)2^{-n}}\right) - G\left(X_{(N-1)2^{-n}}\right)\right]
\end{aligned}$$

$$= \int_{\mathbb{R}^d} \left[V_n^1(y) - G(y)\right] p\left((N-1)2^{-n}; x; y\right) dy$$
$$= \int_{\mathbb{R}^d} \left(\mathbb{P}_{2^{-n}} G(y) - G(y)\right)^+ p\left((N-1)2^{-n}; x; y\right) dy. \tag{4.20}$$

where $V_n^1(y) = \max\left(G(x), \mathbb{P}_{2^{-n}} G(x)\right)$. □

The upper bounds on the rates of convergence provides information on how many iterations are expected to obtain an accurate value function. We extend Theorem 4 to the rate of convergence in the time-homogeneous Wald-Bellman equations with infinite number of iterations under certain conditions.

Lemma 5. *The rates of convergence in the time-homogeneous Wald-Bellman equations* (4.11) *with infinite number of iterations are determined by*

$$|V_n(x) - V_n^{N-1}(x)| \leq \int_{\mathbb{R}^d} \frac{K_d \left(\mathbb{P}_{2^{-n}} G(y) - G(y)\right)^+}{(\epsilon 2^{-n})^{d/2}} \sum_{i=0}^{\infty} (N+i)^{-d/2} \exp\left(\frac{-2^n |y-x|^2}{4A(N+i)}\right) dy \tag{4.21}$$

if each of the derivatives of function a in (4.9) *is bounded and $a(x) \geq \epsilon I$ for some $\epsilon > 0$ where $K_d < \infty$ is a universal constant depending only on d and $A = \sup_{x \in \mathbb{R}^d} ||a(x)||_{\text{op}}$ for x and y in $\mathbb{R}^d$, $d \geq 3$, $n \geq 0$ and $N \in \mathbb{N}$.*

Proof. The upper bound on the rates of convergence in Theorem 4 can be extended to the infinite-horizon case

$$|V_n(x) - V_n^N(x)| = V_n(x) - V_n^N(x) = \lim_{k \to \infty} V_n^{N+k}(x) - V_n^N(x) \tag{4.22}$$

for $k \in \mathbb{N}$ where

$$V_n^{N+k}(x) - V_n^N(x)$$
$$= V_n^{N+k}(x) - V_n^{N+k-1}(x) + V_n^{N+k-1}(x) - \ldots - V_n^{N+1}(x) + V_n^{N+1}(x) - V_n^N(x)$$
$$= \sum_{i=0}^{k-1} V_n^{N+i+1}(x) - V_n^{N+i}(x) \tag{4.23}$$

which implies that

$$V_n(x) - V_n^N(x) = \sum_{i=0}^{\infty} V_n^{N+i+1}(x) - V_n^{N+i}(x). \tag{4.24}$$

By Theorem 4 we obtain

$$V_n(x) - V_n^N(x) \leq \sum_{i=0}^{\infty} \int_{\mathbb{R}^d} \left(\mathbb{P}_{2^{-n}} G(y) - G(y)\right)^+ p\left((N+i)2^{-n}; x; y\right) dy. \tag{4.25}$$

If each of the derivatives of function a in (4.9) is bounded and $a(x) \geq \epsilon I$ for some $\epsilon > 0$, then [31, Theorem 4.1.11] introduced an upper bound on the transition density function which leads to

$$\begin{aligned} V_n(x) - V_n^N(x) &\leq \sum_{i=0}^{\infty} \int_{I\!R^d} \frac{K_d\left(\mathbb{P}_{2^{-n}}G(y) - G(y)\right)^+}{(\epsilon(N+i)2^{-n})^{d/2}} \exp\left(\frac{-|y-x|^2}{4A(N+i)2^{-n}}\right) dy \\ &= \int_{I\!R^d} \frac{K_d\left(\mathbb{P}_{2^{-n}}G(y)-G(y)\right)^+}{(\epsilon 2^{-n})^{d/2}} \sum_{i=0}^{\infty}(N+i)^{-d/2} \exp\left(\frac{-2^n|y-x|^2}{4A(N+i)}\right) dy. \end{aligned} \tag{4.26}$$

The series in (4.26) is convergent for $d \geq 3$ and divergent otherwise. □

The upper bounds in Theorem 4 and Lemma 5 converge to 0 as N increases to ∞ when the underlying process X is a diffusion process with starting point $x \in I\!R^d$ given and fixed.

4. The identity in (4.10) shows that the discrete finite-horizon Wald-Bellman equations converge to the continuous infinite-horizon value function. However, the order of the limits cannot be interchanged because the value function $V_n^N(x)$ is not necessarily monotone in n when $N \in I\!N$ and $x \in I\!R$ are given and fixed. To overcome this restriction we define the equivalent value function $\bar{V}$ by

$$\bar{V}_n^T(x) = V_n^{T2^n}(x) = \sup_{\tau \in \mathcal{T}_n^{T2^n}} \mathsf{E}_x G(X_\tau) \tag{4.27}$$

for V_n^N defined in (4.11), $x \in I\!R^d$, $n \geq 0$, $T2^n \in I\!N$ and $\mathcal{T}_n^{T2^n} = \{0, 2^{-n}, 2\times 2^{-n}, ..., T\}$. The value function $\bar{V}_n^T(x)$ is increasing in T and n when x is given and fixed.

Proposition 6. *The value function $\bar{V}_n^T(x)$ defined in (4.27) converges to the continuous infinite-horizon value function $V(x)$ defined in (4.2) as n and T increase to ∞. $\bar{V}_n^T(x)$ satisfies the identities*

$$V(x) = \lim_{n\to\infty}\lim_{T\to\infty} \bar{V}_n^T(x) = \lim_{T\to\infty}\lim_{n\to\infty} \bar{V}_n^T(x) \tag{4.28}$$

for $x \in I\!R^d$, $n \geq 0$ and $T2^n \in I\!N$.

Proof. From (4.10) we claim that

$$V(x) = \lim_{n\to\infty}\lim_{N\to\infty} V_n^N(x) = \lim_{n\to\infty}\lim_{T\to\infty} \bar{V}_n^T(x). \tag{4.29}$$

By (4.11) and (4.27) we know that the value function $\bar{V}_n^T(x)$ is increasing in n and T because for $n \geq 0$, $N \in I\!N$, $\mathcal{T}_n^N = \{\, 0, 2^{-n}, 2 \times 2^{-n}, ..., N \times 2^{-n} \,\}$ we have $\mathcal{T}_n^N \subset \mathcal{T}_n^{N+1} \subset \mathcal{T}_{n+1}^{N+1}$ which implies that

$$V_n^N(x) \leq V_n^{N+1}(x) \leq V_{n+1}^{N+1}(x). \tag{4.30}$$

Consequently, we obtain

$$\begin{aligned}\bar{V}_n^T(x) = V_n^{T2^n}(x) &\leq V_n^{(T+1)2^n}(x) = \bar{V}_n^{T+1}(x) \\ &\leq V_{n+1}^{(T+1)2^{n+1}}(x) = \bar{V}_{n+1}^{T+1}(x)\end{aligned} \tag{4.31}$$

which leads to the identities

$$V(x) = \lim_{n\to\infty} \lim_{T\to\infty} \bar{V}_n^T(x) = \lim_{T\to\infty} \lim_{n\to\infty} \bar{V}_n^T(x) \tag{4.32}$$

with $x \in I\!R^d$, $n \geq 0$ and $T2^n \in I\!N$. □

5. The Wald-Bellman equations in (4.11) provide a constructive method to solve finite-horizon discrete-time Mayer formulated optimal stopping problems. We now show that the value function $\bar{V}_n^T$ in (4.36) converge to finite-horizon continuous-time optimal stopping problems.

Theorem 7. *The value function $\bar{V}_n^T(x)$ in* (4.27) *satisfies the identity*

$$\bar{V}^T(x) := \sup_{0\leq\tau\leq T} \mathsf{E}_x G(X_\tau) = \lim_{n\to\infty} \bar{V}_n^T(x) \tag{4.33}$$

if the gain function G is continuous and satisfies (4.1) *with $x \in I\!R^d$, $n \geq 0$ and $T2^n \in I\!N$.*

Proof. (i). By the definition of $\bar{V}_n^T(x)$ in (4.27) we have

$$\bar{V}^T(x) = \sup_{0\leq\tau\leq T} \mathsf{E}_x G(X_\tau) \geq \bar{V}_n^T(x). \tag{4.34}$$

for $n \geq 0$. Since $\bar{V}_n^T(x)$ is increasing in n we obtain

$$\bar{V}^T(x) = \sup_{0\leq\tau\leq T} \mathsf{E}_x G(X_\tau) \geq \sup_{n\geq 0} \bar{V}_n^T(x) = \lim_{n\to\infty} \bar{V}_n^T(x). \tag{4.35}$$

(ii). Let $\tau \in [0, T]$ be a stopping time of the underlying process X. Define τ_n by

$$\tau_n = k2^{-n} \wedge T = \min\big(k2^{-n}, T\big) \quad \text{if } (k-1)2^{-n} \leq \tau < k2^{-n} \tag{4.36}$$

for $k \in \{1, 2, 3, ..., T2^n + 1\}$, $T2^n \in I\!N$ and $n \geq 0$. Then τ_n is a stopping time of X and $\tau_n \downarrow \tau$ as $n \to \infty$ that $\tau = \lim_{n\to\infty} \tau_n$. τ_n is an element of the set of stopping times of X defined by $\mathcal{T}_n^{T2^n} = \{0, 2^{-n}, 2 \times 2^{-n}, ..., T\}$ which implies that

$$\mathsf{E}_x G(X_{\tau_n}) \leq \bar{V}_n^T(x) = \sup_{\tau \in \mathcal{T}_n^{T2^n}} \mathsf{E}_x G(X_\tau) \tag{4.37}$$

for $\tau_n \in \mathcal{T}_n^{T2^n}$. Since X is a continuous Markov process, by using that G is a continuous function satisfying (4.1) and dominated convergence theorem, we obtain

$$\lim_{n\to\infty} \mathsf{E}_x G(X_{\tau_n}) = \mathsf{E}_x G(X_\tau) \leq \lim_{n\to\infty} \bar{V}_n^T(x) \tag{4.38}$$

for $\tau \in [0, T]$. By taking the supremum over the stopping time $\tau \in [0, T]$ we obtain

$$\sup_{0\leq\tau<T} \mathsf{E}_x G(X_\tau) \leq \lim_{n\to\infty} \bar{V}_n^T(x). \tag{4.39}$$

Combining the above results we conclude that

$$\bar{V}^T(x) = \sup_{0\leq\tau\leq T} \mathsf{E}_x G(X_\tau) = \lim_{n\to\infty} \bar{V}_n^T(x) \tag{4.40}$$

with $x \in I\!R^d$, $n \geq 0$ and $T2^n \in I\!N$. $\square$

6. We now consider the time-inhomogeneous Wald-Bellman equations. Define the value function $\bar{V}_n^T$ by

$$\bar{V}_n^T(t, x) = V_n^{T2^n}(t, x) = \sup_{\tau \in \mathcal{T}_n^{(T-t)2^n}} \mathsf{E}_{t,x} G(t + \tau, X_{t+\tau}) \tag{4.41}$$

for V_n^N defined in (4.16), $x \in I\!R^d$, $T2^n \in I\!N$, $t \in \mathcal{T}_n^{T2^n-1}$, $\mathcal{T}_n^{T2^n-1} = \{0, 2^{-n}, 2 \times 2^{-n}, ..., T - 2^{-n}\}$ and $\mathcal{T}_n^{(T-t)2^n} = \{0, 2^{-n}, 2 \times 2^{-n}, ..., T - t\}$. The value function $\bar{V}_n^T(t, x)$ is increasing in T and n when t and x are given and fixed.

Theorem 8. *The value function $\bar{V}_n^T(t, x)$ in (4.41) satisfies the identity*

$$V^T(t, x) := \sup_{0\leq\tau\leq T-t} \mathsf{E}_{t,x} G(t + \tau, X_{t+\tau}) = \lim_{n\to\infty} \bar{V}_n^T(t, x) \tag{4.42}$$

if the gain function G is continuous and satisfies (4.12) with $x \in I\!R^d$, $n \geq 0$, $T2^n \in I\!N$ and $t \in [0, T)$.

Proof. (i). By the definition of $\bar{V}_n^T(t, x)$ in (4.41) we have

$$V^T(t, x) = \sup_{0\leq\tau\leq T-t} \mathsf{E}_{t,x} G(t + \tau, X_{t+\tau}) \geq \bar{V}_n^T(t, x) \tag{4.43}$$

for $n \geq 0$. Since $\bar{V}_n^T(t,x)$ is increasing in n we obtain

$$V^T(t,x) = \sup_{0 \leq \tau \leq T-t} \mathsf{E}_{t,x} G(t+\tau, X_{t+\tau}) \geq \sup_{n \geq 0} \bar{V}_n^T(t,x) = \lim_{n\to\infty} \bar{V}_n^T(t,x). \tag{4.44}$$

(ii). Let $\tau \in [0, T-t]$ be a stopping time of the underlying process X with $t \in [0,T)$ and $(T-t)2^n \in \mathbb{N}$. Define τ_n by

$$\tau_n = k2^{-n} \wedge T - t \quad \text{if } (k-1)2^{-n} \leq \tau < k2^{-n} \tag{4.45}$$

for $k \in \{1, 2, 3, ..., (T-t)2^n + 1\}$, $(T-t)2^n \in \mathbb{N}$, $t \in [0,T)$ and $n \geq 0$. Then τ_n is a stopping time of X and $\tau_n \downarrow \tau$ as $n \to \infty$ that $\tau = \lim_{n\to\infty} \tau_n$. τ_n is an element of the set of stopping times of X defined by $\mathcal{T}_n^{(T-t)2^n} = \{0, 2^{-n}, 2 \times 2^{-n}, ..., T-t\}$ which implies that

$$\mathsf{E}_{t,x} G(t+\tau_n, X_{t+\tau_n}) \leq \bar{V}_n^T(t,x) = \sup_{\tau \in \mathcal{T}_n^{(T-t)2^n}} \mathsf{E}_{t,x} G(t+\tau, X_{t+\tau}) \tag{4.46}$$

for $\tau_n \in \mathcal{T}_n^{(T-t)2^n}$. Since X is a continuous Markov process, by using that G is a continuous function satisfying (4.12) and dominated convergence theorem, we obtain

$$\lim_{n\to\infty} \mathsf{E}_{t,x} G(t+\tau_n, X_{t+\tau_n}) = \mathsf{E}_{t,x} G(t+\tau, X_{t+\tau}) \leq \lim_{n\to\infty} \bar{V}_n^T(t,x) \tag{4.47}$$

for $\tau \in [0, T-t]$, $(T-t)2^n \in \mathbb{N}$ and $t \in [0,T)$. By taking the supremum over $\tau \in [0, T-t]$ we obtain

$$V^T(t,x) = \sup_{0 \leq \tau \leq T-t} \mathsf{E}_{t,x} G(t+\tau, X_{t+\tau}) \leq \lim_{n\to\infty} \bar{V}_n^T(t,x). \tag{4.48}$$

Combining the above results we conclude that

$$V^T(t,x) = \sup_{0 \leq \tau \leq T-t} \mathsf{E}_{t,x} G(t+\tau, X_{t+\tau}) = \lim_{n\to\infty} \bar{V}_n^T(t,x) \tag{4.49}$$

with $x \in \mathbb{R}^d$, $n \geq 0$, $(T-t)2^n \in \mathbb{N}$ and $t \in [0,T)$. □

In the case of infinite horizon, the time-inhomogeneous problems can be treated in the same way as time-homogeneous problems by altering the remaining time (cf. [21, p. 18]). Further details to Wald-Bellman equations could be found in [21, Chapter I] and [29, Chapter 1-3].

7. The Wald-Bellman equations were first characterized implicitly in Wald's sequential analysis [34] and then stated explicitly in [1, p. 219]. They are the simplest case of

dynamic programming developed by Bellman [2]. Upper bounds on the rates of convergence in the Wald-Bellman equations were characterized implicitly first by Ray [24] and then Grigelionis and Shiryayev [10]. The same authors introduced the technique of constructing the value functions associated with infinite-horizon continuous-time Mayer formulated optimal stopping problems in [11].

Chapter 5

Quickest detection problems for Ornstein-Uhlenbeck processes

In this chapter, we apply the results obtained in Chapter 4 to the Mayer formulated optimal stopping problem introduced in Section 2.6. Numerical approximations of the Wald-Bellman equations are obtained through Mathematica algorithms.

1. The value function $\hat{V}$ in (2.93) can be rewritten as

$$\hat{V}(\varphi, x) = \tilde{V}(\varphi, x) - \hat{M}(\varphi, x) \tag{5.1}$$

where $\tilde{V}$ is the equivalent optimal stopping problem defined by

$$\tilde{V}(\varphi, x) = \inf_{\tau \geq 0} \mathsf{E}^{\infty}_{\varphi,x}\left[e^{-\lambda\tau}\hat{M}(\Phi_\tau, X_\tau)\right] \tag{5.2}$$

with $(\varphi, x) \in [0, \infty) \times I\!R$ for all stopping times τ of (Φ, X). By killing the sample path of (Φ, X) at the rate λ (cf. [21, Subsection 5.4] and [21, Subsection 6.3]), we obtain the identities

$$\mathsf{E}^{\infty}_{\varphi,x}\left[e^{-\lambda t}\hat{M}(\Phi_t, X_t)\right] = \mathsf{E}^{\infty}_{\varphi,x}\left[\hat{M}(\tilde{\Phi}_t, \tilde{X}_t)\right] \tag{5.3}$$

$$e^{-\lambda t}p(t; \varphi, x; \eta, z) = \tilde{p}(t; \varphi, x; \eta, z) \tag{5.4}$$

for $t > 0$ with (φ, x) and (η, z) in $[0, \infty) \times I\!R$ where $(\tilde{\Phi}, \tilde{X})$ is the killed process starting at $(\varphi, x) \in [0, \infty) \times I\!R$. p is the transition density function of (Φ, X) and $\tilde{p}$ is the transition density function of $(\tilde{\Phi}, \tilde{X})$ under $\mathsf{P}^{\infty}_{\varphi,x}$. The value function in (5.2) can be simplified to

$$\tilde{V}(\varphi, x) = \inf_{\tau \geq 0} \mathsf{E}^{\infty}_{\varphi,x}\left[\hat{M}(\tilde{\Phi}_\tau, \tilde{X}_\tau)\right] \tag{5.5}$$

with $(\varphi, x) \in [0,\infty) \times I\!R$ for all stopping times τ of $(\tilde{\Phi}, \tilde{X})$.

2. By the Wald-Bellman equations defined in (4.11) and Theorem 5, the value function in (5.5) satisfy the identities

$$\tilde{V}(\varphi, x) = \lim_{n\to\infty} \lim_{T\to\infty} \bar{V}_n^{T2^n}(\varphi, x) = \lim_{T\to\infty} \lim_{n\to\infty} \bar{V}_n^{T2^n}(\varphi, x) \tag{5.6}$$

where $\bar{V}_n^N$ is defined by

$$\bar{V}_n^N(\varphi, x) = \inf_{\tau\in\mathcal{T}_n^N} \mathsf{E}^{\infty}_{\varphi,x} \hat{M}(\tilde{\Phi}_\tau, \tilde{X}_\tau) = \min\Big(\hat{M}(\varphi, x), \mathbb{P}_{2^{-n}} \bar{V}_n^{N-1}(\varphi, x)\Big) \tag{5.7}$$

for $\bar{V}_n^0(\varphi, x) = \hat{M}(\varphi, x)$, $(\varphi, x) \in [0,\infty) \times I\!R$, $n \geq 0$, $T2^n \in I\!N$, $N \in I\!N$ and $\mathcal{T}_n^N = \{0, 2^{-n}, 2\times 2^{-n}, ..., N\times 2^{-n}\}$. $\mathbb{P}_{2^{-n}}$ is the transition operator of $(\tilde{\Phi}, \tilde{X})$ defined by

$$\mathbb{P}_{2^{-n}} \bar{V}_n^{N-1}(\varphi, x) = \mathsf{E}^{\infty}_{\varphi,x} \bar{V}_n^{N-1}(\tilde{\Phi}_{2^{-n}}, \tilde{X}_{2^{-n}}) = \mathsf{E}^{\infty}_{\varphi,x} e^{\lambda 2^{-n}} \bar{V}_n^{N-1}(\Phi_{2^{-n}}, X_{2^{-n}}) \tag{5.8}$$

for $(\varphi, x) \in [0,\infty) \times I\!R$.

3. The continuation set C and stopping set D (cf. [9, Section 7]) are defined by

$$C = \{\, (\varphi, x) \in [0,\infty)\times I\!R \mid \bar{V}(\varphi, x) < \hat{M}(\varphi, x) \,\} \tag{5.9}$$

$$D = \{\, (\varphi, x) \in [0,\infty)\times I\!R \mid \bar{V}(\varphi, x) = \hat{M}(\varphi, x) \,\}. \tag{5.10}$$

The first entry time of (Φ, X) into D defined by

$$\tau_D = \inf\{\, t \geq 0 \mid (\Phi_t, X_t) \in D \,\}. \tag{5.11}$$

is optimal in (5.5) (cf. [21, Theorem 1.7]). The optimal stopping boundary between C and D is defined by

$$b(x) = \inf\{\, \varphi \geq 0 \mid (\varphi, x) \in D \,\} \tag{5.12}$$

for $x \in I\!R$.

4. We recall that the transition density function p of (Φ, X) is the unique non-negative solution to the Kolmogorov backward equation

$$p_t(t; \varphi, x; \eta, z) = I\!L_{\Phi,X}(p)(t; \varphi, x; \eta, z) \tag{5.13}$$

$$p(0+; \varphi, x; \eta, z) = \delta_{\varphi,x}(\eta, z) \quad (weakly) \tag{5.14}$$

under $\mathsf{P}^{\infty}_{\varphi,x}$ satisfying $\int_0^\infty \int_{-\infty}^\infty p(t; \varphi, x; \eta, z)\, d\eta\, dz = 1$ for $t > 0$ with (φ, x) and (η, z) in $[0,\infty)\times I\!R$ (cf. [8]) where $I\!L_{\Phi,X}$ is the infinitesimal generator defined in (2.47) and $\delta_{\varphi,x}$

is the Dirac measure at (φ, x). As an alternative method to compute the expectations in the Wald-Bellman equations (5.7), we use *Euler approximation* (cf. [13, Chapter 2.1]) where

$$\begin{aligned}\Phi_{2^{-n}} &= \varphi + \int_0^{2^{-n}} \lambda(1+\Phi_s)\, ds + \int_0^{2^{-n}} \beta_0 X_s \Phi_s \, dB_s \\ &\approx \varphi + \int_0^{2^{-n}} \lambda(1+\varphi)\, ds + \int_0^{2^{-n}} \beta_0 x \varphi \, dB_s \\ &= \varphi + \lambda(1+\varphi)2^{-n} + \beta_0 x \varphi B_{2^{-n}} \end{aligned} \tag{5.15}$$

$$X_{2^{-n}} \approx x - \beta_0 x 2^{-n} + B_{2^{-n}} \tag{5.16}$$

for large $n \in I\!N$. The expectations associated with (Φ, X) can be approximated by using the law of standard Brownian motion.

It remains our interest to find an explicit general solution to the Kolmogorov backward equation (5.13)+(5.14). We suspect that the general solutions may not exist in all cases. However, we may attempt to find the explicit solutions in some special cases by adding some restrictions. For instance, we may first focus on one-dimensional cases and assume that the coefficients are smooth and locally integrable with the infinitesimal generator of the underlying Markov process satisfying Hörmander's condition (cf. [12]). In addition, instead of treating the problem through pure partial differential equations methods, we may also embed probability theories because we could follow the sample path of the underlying Markov process which provides additional information to the classical partial differential equations methods that may lead to potential further research.

5. We now introduce the algorithm of Wald-Bellman equations and the corresponding numerical analysis. The computation is conducted in Mathematica language on a computer with 64-bit Windows 10 system, CPU i5-11500 and 32 GB RAM. The idea of Wald-Bellman equations is value iteration which starts with the corresponding gain function. Consider the Wald-Bellman equations in (5.6)+(5.7), the iterations are given by

$$\bar{V}_n^0(\varphi, x) = \hat{M}(\varphi, x) \tag{5.17}$$

$$\begin{aligned}\bar{V}_n^1(\varphi, x) &= \inf_{\tau \in \mathcal{T}_n^1} \mathsf{E}^\infty_{\varphi,x} \hat{M}(\tilde{\Phi}_\tau, \tilde{X}_\tau) = \min\Big(\hat{M}(\varphi, x), \mathbb{P}_{2^{-n}} \bar{V}_n^0(\varphi, x)\Big) \\ &= \min\Big(\hat{M}(\varphi, x), \mathsf{E}^\infty_{\varphi,x} e^{\lambda 2^{-n}} \bar{V}_n^0(\Phi_{2^{-n}}, X_{2^{-n}})\Big)\end{aligned}$$

$$
\begin{aligned}
&= \min\Big(\hat{M}(\varphi,x), \int_0^\infty \int_{-\infty}^\infty e^{\lambda 2^{-n}}\, \bar{V}_n^0(\eta,z)\, p(2^{-n};\varphi,x;\eta,z)\, d\eta\, dz\Big)\\
&\approx \min\Big(\hat{M}(\varphi,x),\\
&\qquad \int_{-\infty}^\infty e^{\lambda 2^{-n}}\, \bar{V}_n^0(\varphi+\lambda(1+\varphi)2^{-n}+\beta_0 x\varphi y, x-\beta_0 x 2^{-n}+y)\, \hat{p}(2^{-n};y)\, dy\Big)
\end{aligned} \tag{5.18}
$$

$$
\begin{aligned}
\bar{V}_n^2(\varphi,x) &= \inf_{\tau\in\mathcal{T}_n^2} \mathsf{E}_{\varphi,x}^\infty \hat{M}(\tilde{\Phi}_\tau,\tilde{X}_\tau) = \min\Big(\hat{M}(\varphi,x), \mathbb{P}_{2^{-n}}\bar{V}_n^1(\varphi,x)\Big)\\
&\approx \min\Big(\hat{M}(\varphi,x),\\
&\qquad \int_{-\infty}^\infty e^{\lambda 2^{-n}} \bar{V}_n^1(\varphi+\lambda(1+\varphi)2^{-n}+\beta_0 x\varphi y, x-\beta_0 x 2^{-n}+y)\, \hat{p}(2^{-n};y) dy\Big)
\end{aligned} \tag{5.19}
$$

$$\vdots$$

$$
\begin{aligned}
\bar{V}_n^{T2^n}(\varphi,x) &= \inf_{\tau\in\mathcal{T}_n^{T2^n}} \mathsf{E}_{\varphi,x}^\infty \hat{M}(\tilde{\Phi}_\tau,\tilde{X}_\tau) = \min\Big(\hat{M}(\varphi,x), \mathbb{P}_{2^{-n}}\bar{V}_n^{T2^n-1}(\varphi,x)\Big)\\
&\approx \min\Big(\hat{M}(\varphi,x),\\
&\qquad \int_{-\infty}^\infty e^{\lambda 2^{-n}} \bar{V}_n^{T2^n-1}(\varphi+\lambda(1+\varphi)2^{-n}+\beta_0 x\varphi y, x-\beta_0 x 2^{-n}+y)\hat{p}(2^{-n};y)dy\Big)
\end{aligned} \tag{5.20}
$$

for $(\varphi,x)\in[0,\infty)\times I\!R$, $n\geq 0$, $T2^n\in I\!N$, $N\in I\!N$ and $\mathcal{T}_n^N=\{0,2^{-n},2\times 2^{-n},...,N\times 2^{-n}\}$ where $\hat{p}$ is the transition density function of standard Brownian motion defined by

$$\hat{p}(t;y) = \frac{e^{-y^2/(2t)}}{\sqrt{2\pi t}} \tag{5.21}$$

with $t>0$ and $y\in I\!R$. We first construct $\bar{V}_n^1$ based on (5.18) by using $\bar{V}_n^0$ in (5.17) and continue constructing value functions iteratively until $\bar{V}_n^{T2^n}$. Starting from $\bar{V}_n^2$, the value functions are defined as nested integrals which cannot be computed explicitly. To overcome this problem, we use interpolations functions to approximate the value functions and decouple the integrals. We now illustrate an example code in Mathematica where the idea is applicable to other similar problems in different programming languages. Firstly, the coefficients are defined by

Code 1.

```
c=1.; time=1.; phistep=phiubound/100.;
xstep=(xubound-xlbound)/100.; h=0.5^10.; w=time/h;
```

```
lambda=1.  ;beta0=1.; xubound =1.; xlbound =-1.;
phiubound=2.; philbound=0.;
lossfunction[phi_,x_]:=phi*x^2.+(x^2.+1./lambda)/(1.+(2.*beta0)/lambda)
                              +1./c;
functionlist={lossfunction};
```

where `c` is the constant $c \in I\!R^+$ in (2.4). `time` is T in (5.6). `phiubound` and `philbound` are the upper and lower bounds of the interpolation domain for φ. Similarly, `xubound` and `xlbound` are the upper and lower bounds of the interpolation domain for x. `h` is 2^{-n} in (5.7). `w` is the number of total iterations defined as $T2^n \in I\!N$ in (5.6). `lambda` is $\lambda > 0$ in (2.1). `beta0` is the mean-reversion rate $\beta_0 \in I\!R$ in (2.3). `phistep` denotes the distance between the grids for interpolating φ. `xstep` denotes the distance between the grids for interpolating x. `lossfunction` is $\hat{M}$ in (2.94). `functionlist` is a `list` of functions which stores the interpolated value functions. After the value iterations, the `functionlist` should consists of approximated value functions where `functionlist` $= \{\bar{V}_n^0, \bar{V}_n^1, \bar{V}_n^2, ..., \bar{V}_n^{T2^n}\}$. We now introduce the algorithm of iterations

Code 2.

```
Do[Subscript[valuefunctionexact,n][phi_,x_]:=
 Min[lossfunction[phi, x],
  NIntegrate[
   functionlist[[n]][lambda*(1.+phi)*h+beta0*x*phi*b+phi,-beta0*x*h+b+x]*
    Exp[-lambda*h]/Sqrt[2.*Pi*h]*Exp[-0.5/h*b^2.],{b,-0.5,0.5},
     Method->{Automatic,"SymbolicProcessing"->0},AccuracyGoal->10]];
 Subscript[table,n]=N[Flatten[
  Table[{phi,x,Subscript[valuefunctionexact,n][phi, x]},
   {phi,0,phiubound,phistep},{x,xlbound,xubound,xstep}],1]];
 Subscript[valuefunction,n]=Interpolation[Subscript[table,n]];
 functionlist=Insert[functionlist,Subscript[valuefunction,n],-1],{n,w}].
```

We recall that `NIntegrate` is a built-in function in Mathematica to compute integrals numerically. The option `Method->{Automatic,"SymbolicProcessing"->0}` is used to let Mathematica start numerical integrating without simplifying the integrals which increases the computation speed of the algorithm. As mentioned above, `functionlist` stores the generated value functions. The command `functionlist[[n]]` calls the n-*th* element from the `functionlist`. `Table` is a built-in function which generates a `List` depending on the inputs. `N` is a built-in function which converts its inputs into `Float` type data. `Flatten` is a built-in function which relax the nested levels of the `List`. `Interpolation` is a built-in function which generates an interpolation function based on the input `List`. `Insert` is a built-in function which insert an element to the required position of the `List`.

The idea of Code 2 is to first construct an approximated value function called `Subscript[valuefunctionexact,n]` based on the value function given by Wald-Bellman equations in (5.7) with Euler approximations in (5.15)+(5.16). The integrals for computing expectations are taken from -0.5 to 0.5 instead from $-\infty$ to ∞ because the algorithm is computation intensive. Shrinking the integral interval helps reduce the computation time. The marginal distribution of (φ, x) is approximated by standard Brownian motion which has light tailed distribution, thus the approximation should still remain reliable. Then we construct a `list` of data `Subscript[table,n]` which has the form $\{\{$`phi,x,Subscript[valuefunctionexact,n][phi,x]`$\}\}$ for all `phi=philbound, philbound+phistep, philbound+2*phistep, ..., phiubound` and `x=xlbound, xlbound+xstep, xlbound+2*xstep, ..., xubound`. The interpolation function `Subscript[valuefunction,n]` is generated from the `list` of data `Subscript[table,n]` where the default interpolation order in Mathematica is 3. The interpolation function `Subscript[valuefunction,n]` is added to the last position of the `functionlist`. The whole procedure is repeated $T2^n \in I\!N$ times to obtain $\bar{V}_n^{T2^n}$.

6. The generation of value functions $\bar{V}_n^N$ defined in (5.7) through Wald-Bellman equations can be imaged as a surface put under the loss function $\hat{M}$ defined in (2.94) being pulled down towards the exact the value function $\tilde{V}$ defined in (5.2). The more iterations conducted in Wald-Bellman equations, the closer the approximated value functions to $\tilde{V}$ as shown in Figure 6. We recall in Theorem 5 that the finite-horizon Wald-Bellman equations converge to the infinite-horizon optimal stopping problem as the number of iterations goes to infinity. We generate the value function $\hat{V}$ defined in (5.1) according to $\bar{V}_n^N$ as shown in Figure 7. The generation of value functions

$\hat{V}$ defined in (5.1) through Wald-Bellman equations can be similarly imaged as a surface put at 0 being pulled down. Figure 8 shows a comparison between the results generated by free-boundary problem and Wald-Bellman equations. We recall that the value function $\hat{V}$ in (2.42) is non-positive because one could always choose to stop at once. It is remarkable that the constructive solution through Wald-Bellman equations is very reliable and close to the result generated through free-boundary problem. It is also remarkable that the value functions generated by Wald-Bellman equations comply with smooth fit (cf. [21, Chapter IV, 9.1] and [9, Section 7, 8]) at the touching points between the value functions and loss function as shown in Figure 9 and 10.

7. We construct the optimal stopping boundary by the definition in (5.12). Recall that for $\hat{V}$ defined in (2.42), it is not optimal to stop when $\Phi_t < \frac{\lambda}{c}$. The definitions of continuation set C and stopping set D in (5.9)+(5.10) can be written as

$$C = \{ (\varphi, x) \in [0, \infty) \times I\!R \mid \hat{V}(\varphi, x) < 0 \} \tag{5.22}$$

$$D = \{ (\varphi, x) \in [0, \infty) \times I\!R \mid \hat{V}(\varphi, x) = 0 \}. \tag{5.23}$$

for $\hat{V}$ defined in (5.1) (cf. [9, Section 7]). By observing Figure 7 and 8, we see that the boundary $b(x)$ between C and D satisfies the identity $b(x) \geq \frac{\lambda}{c}$. The optimal stopping boundary $b(x)$ is constructed in Mathematica as follows

Code 3.

```
boundary=Table[{phi/.FindRoot[functionlist[[w+1]][phi, x]]-
 lossfunction[phi, x],{phi,lambda/c},AccuracyGoal->10,
  PrecisionGoal -> 10], x},{x, xlbound, xubound, xstep}];.
```

The built-in function `FindRoot` finds the roots of the input equations starting from a given data point. As we mentioned above, it is optimal to stop when $\varphi \geq \lambda/c$. Therefore, we choose `lambda/c` to be the starting points for `FindRoot`. The symbol `/.` converts the result generated by `FindRoot` from `Array` to `Float` data type which is necessary. The generated `boundary` contains the `List` of data points in the form $(b(x), x)$. The smoothness of optimal stopping boundary is controlled by the number of grids used for interpolation as shown in Figure 11. The more grids used, the smoother and more accurate the boundary would be. The construction of optimal stopping boundaries along with Wald-Bellman equations can be imaged as a rope being pulled towards the true boundary as shown in Figure 12. The numerical approximation converges from finite-horizon to the infinite-horizon optimal stopping boundary as the iteration continues which complies with Theorem 5.

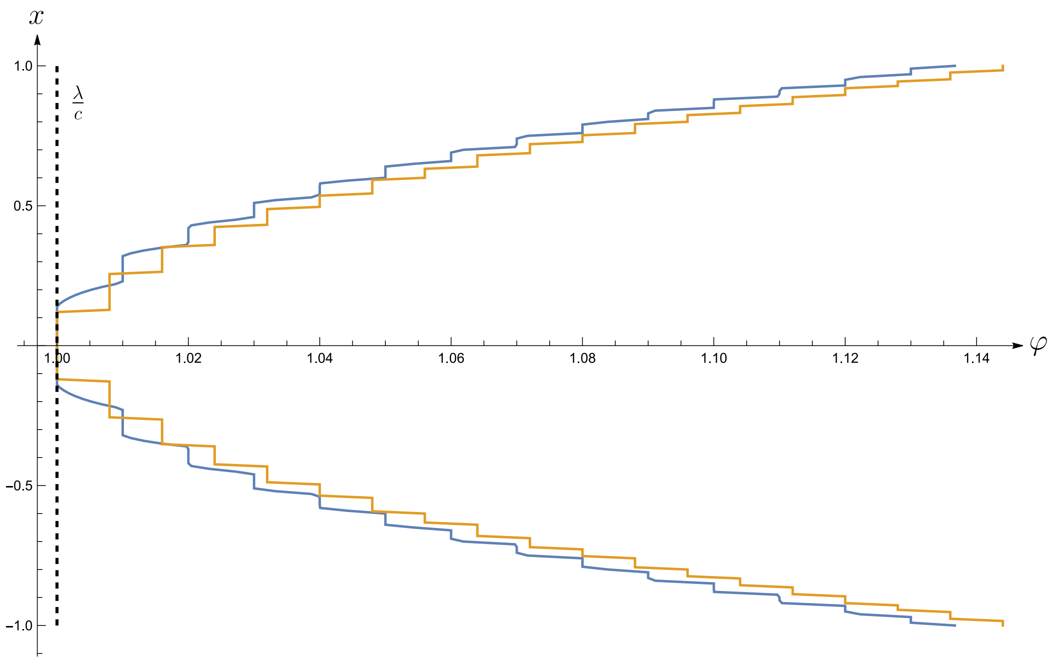

Figure 3. Numerical approximations of optimal stopping boundaries generated through free-boundary problem in [9, Section 8] (yellow) and Wald-Bellman equations in (5.6)+(5.7) (blue) with $\lambda = 1$, $\beta_0 = 1$, $T = 1$, $c = 1$ and $\sigma = 1$.

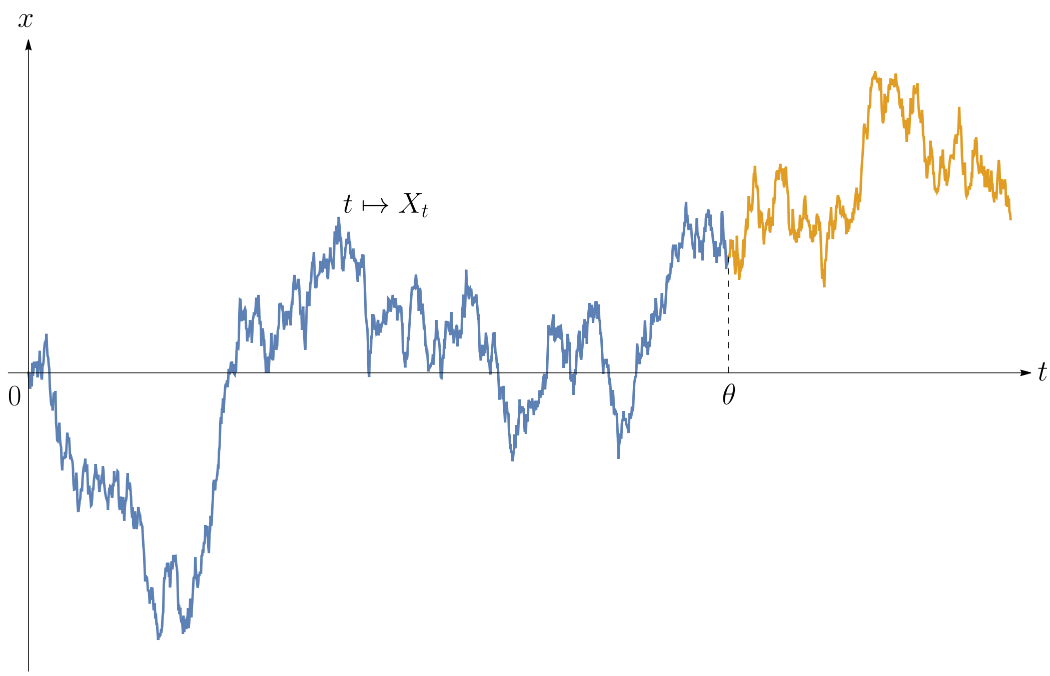

Figure 4. Simulated sample path of the Ornstein-Uhlenbeck process X with $\beta_0 = 1$ and $\sigma = 1$.

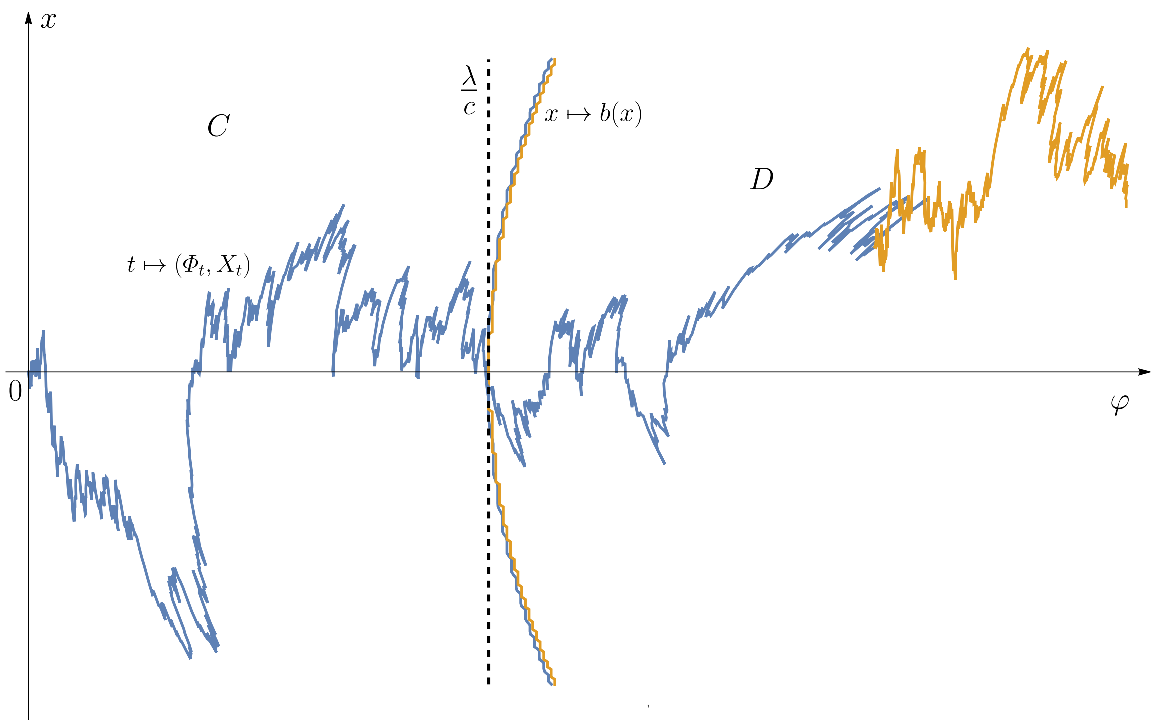

Figure 5. Kinematics of the process (Φ, X) associated with the sample path from Figure 4 and locations of the optimal stopping boundaries from Figure 3.

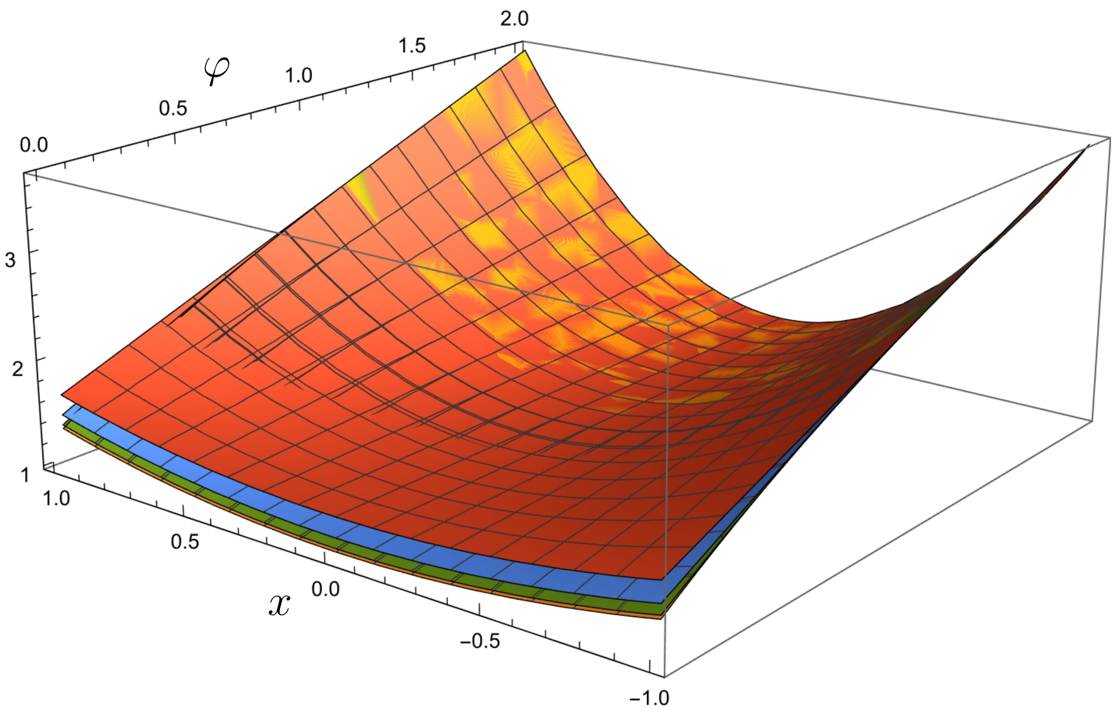

Figure 6. Numerical approximations of value functions $\bar{V}_n^N$ defined in (5.7) generated by Code 1 and 2 with 251 iterations (blue), 501 iterations (green) and 1024 iterations (yellow). $\hat{M}$ defined in (2.94) (red).

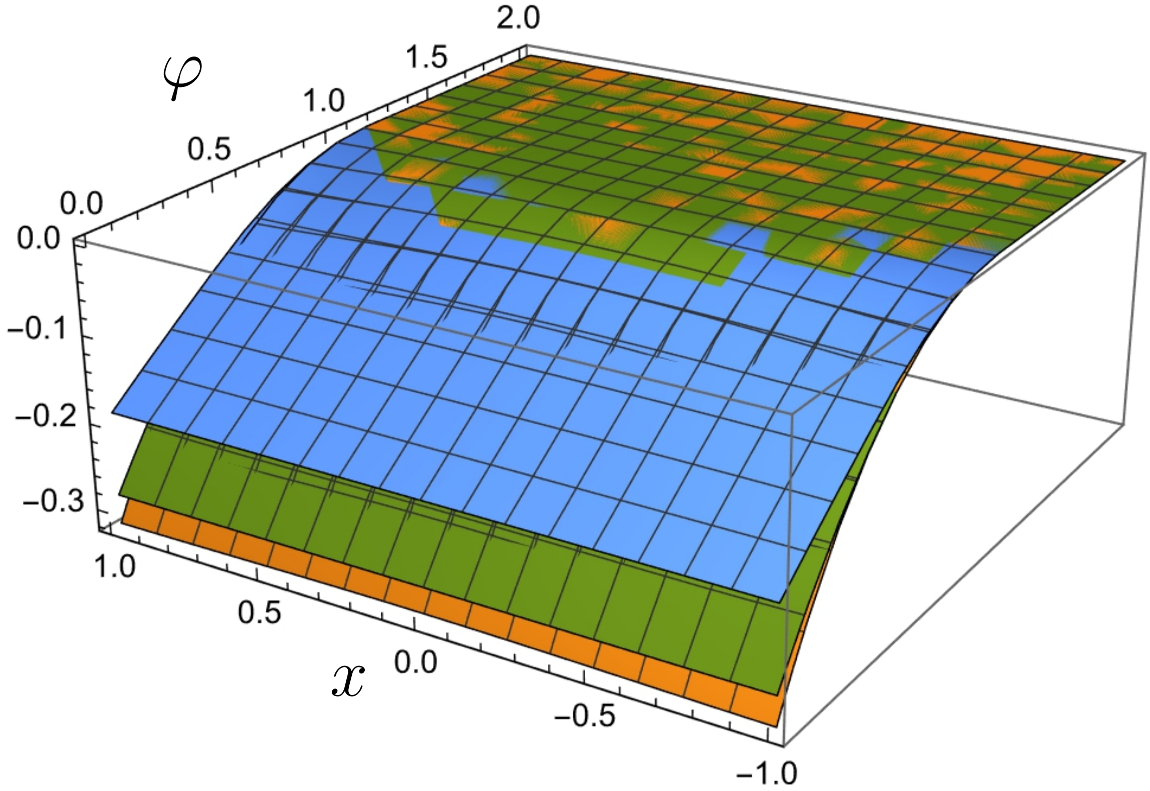

Figure 7. Numerical approximations of value functions $\hat{V}$ defined in (5.1) according to $\bar{V}_n^N$ in Figure 6 with 251 iterations (blue), 501 iterations (green) and 1024 iterations (yellow).

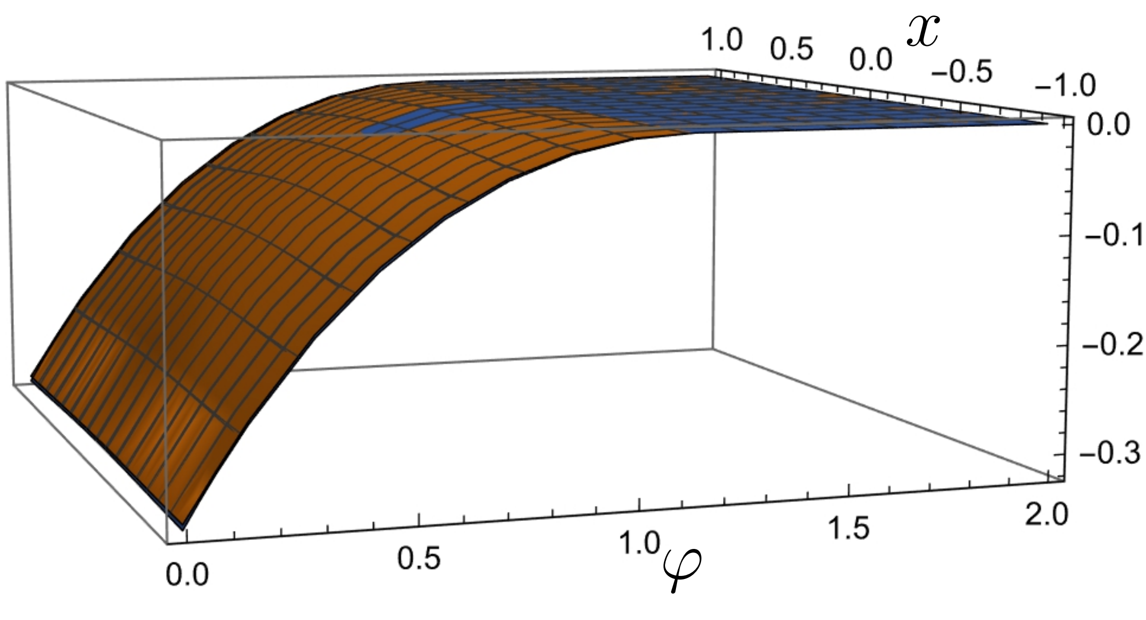

Figure 8. Numerical approximations of value functions $\hat{V}$ according to Figure 6 with 1024 iterations (blue) and value function generated through finite difference method in [9, Section 8] (yellow).

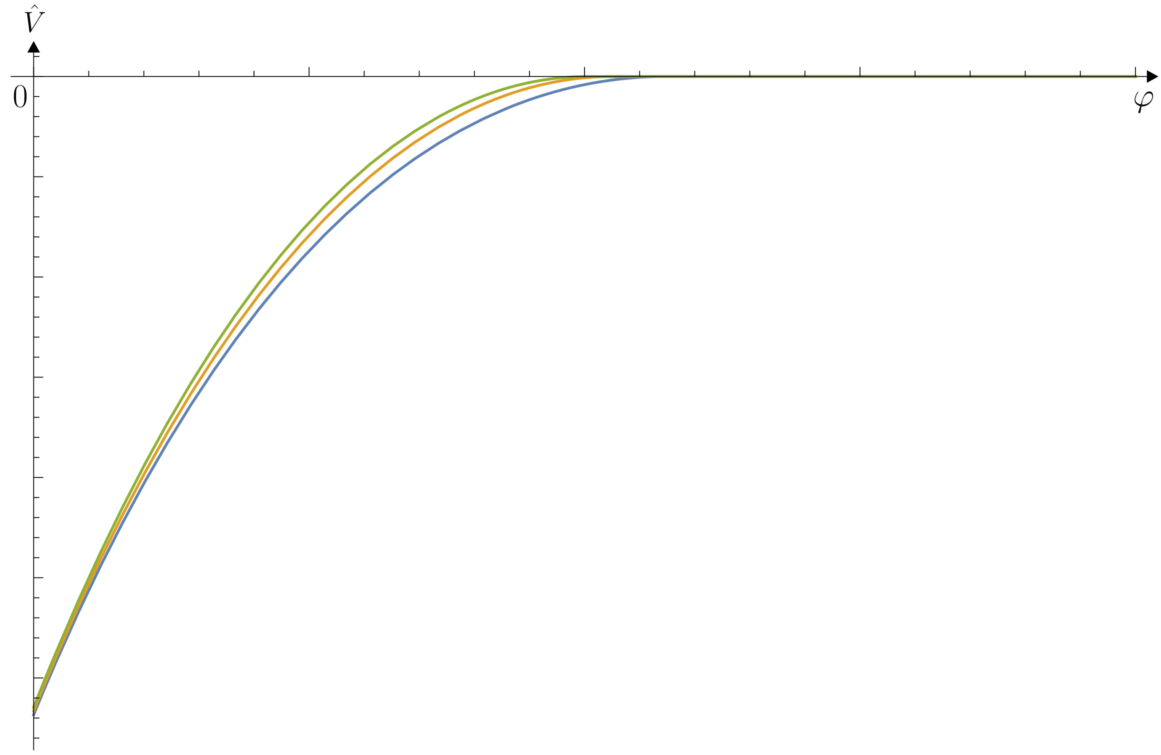

Figure 9. Numerical approximations of value functions $\hat{V}$ according to Figure 6 with 1024 iterations when $\varphi \in [0, 2]$ and $x = -1$ (blue), $x = -0.6$ (yellow) and $x = -0.2$ (green).

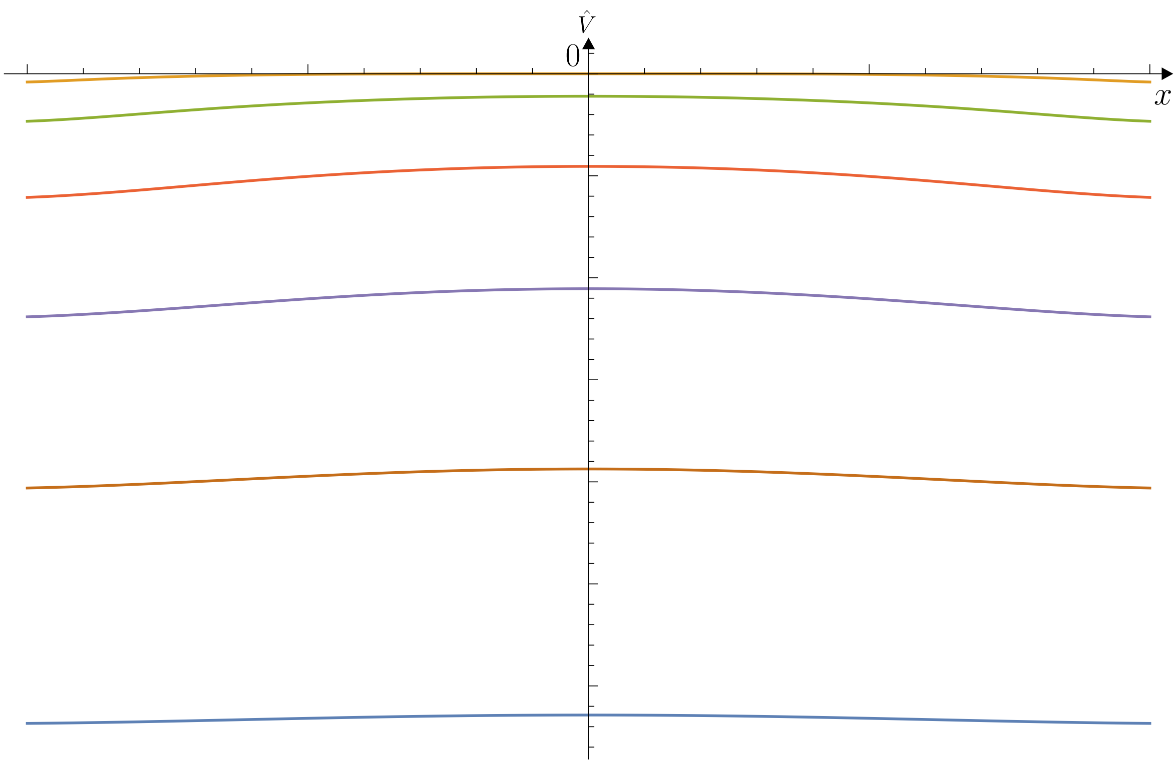

Figure 10. Numerical approximations of value functions $\hat{V}$ according to Figure 6 with 1024 iterations when $x \in [-1, 1]$ and $\varphi = 0$ (blue), $\varphi = 0.2$ (brown), $\varphi = 0.4$ (purple), $\varphi = 0.6$ (red), $\varphi = 0.8$ (green) and $\varphi = 1$ (yellow).

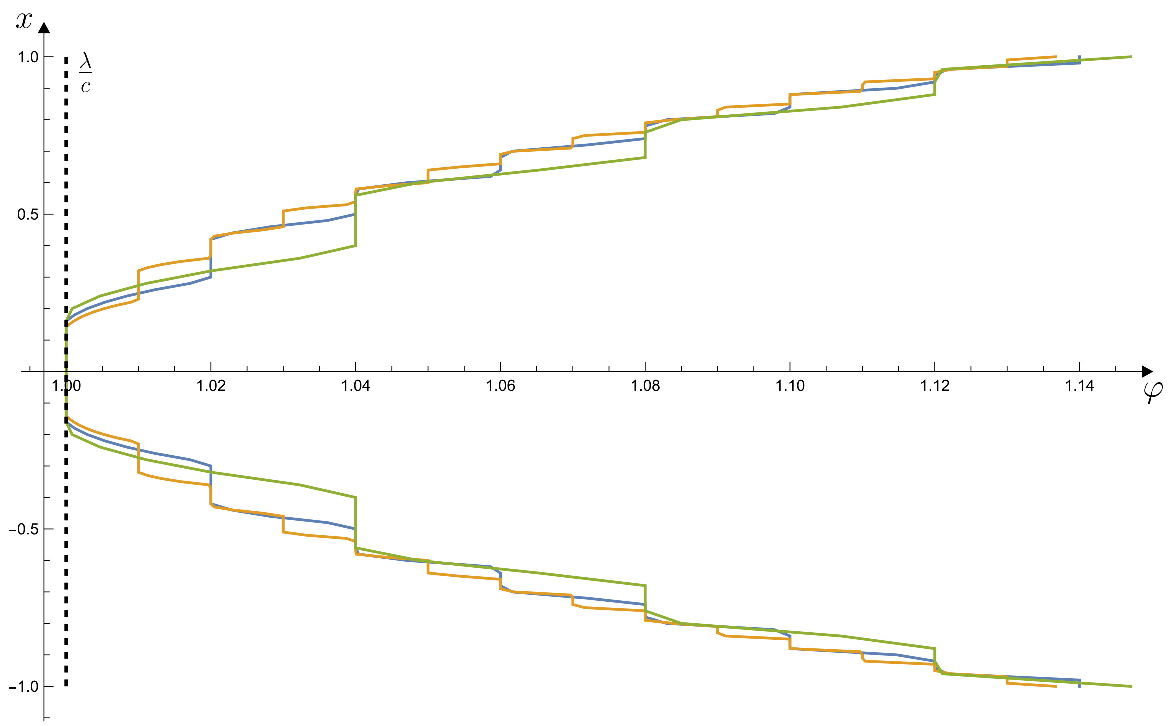

Figure 11. Numerical approximations of optimal stopping boundaries according to Code 1, 2 and 3 when the number of grids used for interpolation are 50 (green), 100 (blue) and 200 (yellow) for φ and x.

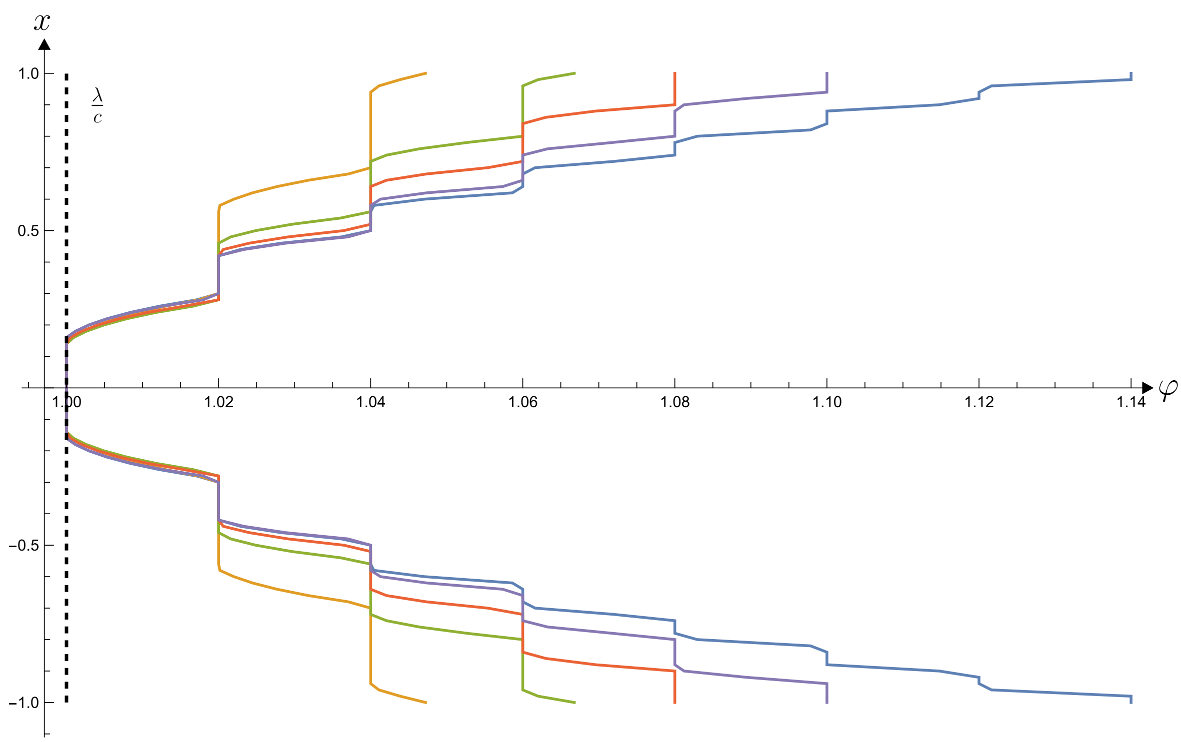

Figure 12. Numerical approximations of optimal stopping boundaries according to Code 1, 2 and 3 when the number of iterations are 10 (yellow), 20 (green), 30 (red), 50 (purple) and 1024 (blue).

Chapter 6

Applications to one-dimensional Brownian motion problem

In this Chapter we consider a Wald's type optimal stopping problem for one-dimensional Brownian motion with *finite horizon* of which the analytical solution is known in the case of *infinite horizon*.

1. We consider the optimal stopping problem given by

$$V^T(x) = \sup_{0 \le \tau \le T} \mathsf{E}_x \left(|X_\tau| - c\tau \right) \tag{6.1}$$

where τ is a stopping time of the Brownian motion $X = (X_t)_{t \ge 0}$ satisfying $X_t = x + B_t$. $B = (B_t)_{t \ge 0}$ is a standard Brownian motion starting at 0 under P_x for $T > 0$ and $c \in I\!R^+$. P_x is the probability measure under which $X_0 = x \in I\!R$.

2. We know that the process $X_t^2 - t$ is a martingale for $0 \le t \le T$. Consider a stopped version of the process $X_t^2 - t$, we obtain

$$\begin{aligned} |X^2_{t\wedge\tau} - t \wedge \tau| &\le |X^2_{t\wedge\tau}| + |t \wedge \tau| = x^2 + |2x B_{t\wedge\tau}| + |B^2_{t\wedge\tau}| + |t \wedge \tau| \\ &= x^2 + |2x| \max_{0 \le t \le \tau} |B_t| + \max_{0 \le t \le \tau} |B_t^2| + \tau. \end{aligned} \tag{6.2}$$

If τ is a bounded stopping time such that $\tau \le M < \infty$, then we have

$$|X^2_{t\wedge\tau} - t \wedge \tau| \le x^2 + |2x| \max_{0 \le t \le M} |B_t| + \max_{0 \le t \le M} |B_t^2| + M := Z. \tag{6.3}$$

We claim Z is integrable. It is enough to show that

$$|2x| \max_{0 \le t \le M} |B_t| + \max_{0 \le t \le M} |B_t^2| \tag{6.4}$$

is integrable. We have

$$|2x|\,\mathsf{E}_x\left(\max_{0\le t\le M}|B_t|\right)+\mathsf{E}_x\left(\max_{0\le t\le M}|B_t^2|\right)\le 2|x|\sqrt{\frac{\pi}{2}M}+4\mathsf{E}_x(B_M^2)$$
$$=2|x|\sqrt{\frac{\pi}{2}M}+4M<\infty \tag{6.5}$$

by applying Doob's maximal inequality. By optional sampling theorem we have

$$\mathsf{E}_x(X_\tau^2-\tau)=\mathsf{E}_x(X_0^2)=x^2 \tag{6.6}$$

which shows that

$$\mathsf{E}_x(X_\tau^2)=\mathsf{E}_x(\tau)+x^2. \tag{6.7}$$

In the general case when τ is not bounded such that $\mathsf{E}_x(\tau)<\infty$, we apply the result in (6.7) to the bounded stopping time $\tau_n := \tau\wedge n$ and conclude that:

$$\mathsf{E}_x(X_{\tau_n}^2)=\mathsf{E}_x(\tau_n)+x^2 \tag{6.8}$$

for $n\ge 1$. By applying monotone convergence theorem to the right hand side of (6.8), we have

$$\lim_{n\to\infty}\mathsf{E}_x(\tau_n)+x^2=\mathsf{E}_x\left(\lim_{n\to\infty}\tau_n\right)+x^2=\mathsf{E}_x(\tau)+x^2. \tag{6.9}$$

For the left hand side of (6.8), note that

$$|X_{\tau_n}^2|=|\,(x+B_{\tau_n})^2\,|=|x^2+2xB_{\tau_n}+B_{\tau_n}^2|\le x^2+2|x|\max_{0\le t\le\tau_n}|B_t|+\max_{0\le t\le\tau_n}B_t^2:=Z'. \tag{6.10}$$

We claim that Z' is integrable. It is enough to show that

$$2|x|\max_{0\le t\le\tau_n}|B_t|+\max_{0\le t\le\tau_n}B_t^2 \tag{6.11}$$

is integrable.

$$|2x|\,\mathsf{E}_x\left(\max_{0\le t\le\tau_n}|B_t|\right)+\mathsf{E}_x\left(\max_{0\le t\le\tau_n}B_t^2\right)\le 2C|x|\sqrt{\mathsf{E}_x(\tau_n)}+4\mathsf{E}_x(\tau_n) \tag{6.12}$$

by applying Doob's maximal inequality with the best constant C equal to $\sqrt{2}$. By applying monotone convergence theorem to both sides of (6.12) we have

$$|2x|\,\mathsf{E}_x\left(\max_{0\le t\le\tau}|B_t|\right)+\mathsf{E}_x\left(\max_{0\le t\le\tau}B_t^2\right)\le 2C|x|\sqrt{\mathsf{E}_x(\tau)}+4\mathsf{E}_x(\tau)<\infty \tag{6.13}$$

which shows that Z' is integrable. By applying dominated convergence theorem to the left hand side of (6.8) we have

$$\lim_{n\to\infty}\mathsf{E}_x(X_{\tau_n}^2)=\mathsf{E}_x\left(\lim_{n\to\infty}X_{\tau_n}^2\right)=\mathsf{E}_x(X_\tau^2). \tag{6.14}$$

Combining the results from (6.9) and (6.14) we conclude that

$$\mathsf{E}_x(X_\tau^2) = \mathsf{E}_x(\tau) + x^2 \tag{6.15}$$

for stopping time τ satisfying $\mathsf{E}_x(\tau) < \infty$. Applying the result in (6.15) to the optimal stopping problem in (6.1) we have

$$V^T(x) = \sup_{0\le\tau\le T} \mathsf{E}_x(|X_\tau| - cX_\tau^2 + cx^2) = cx^2 + \sup_{0\le\tau\le T} \mathsf{E}_x(|X_\tau| - cX_\tau^2) = cx^2 + \tilde{V}^T(x) \tag{6.16}$$

where we set

$$\tilde{V}^T(x) = \sup_{0\le\tau\le T} \mathsf{E}_x(|X_\tau| - cX_\tau^2). \tag{6.17}$$

The optimal stopping problem in (6.17) is equivalent to the optimal stopping problem in (6.1).

3. The optimal stopping problem (6.17) has analytical solution in the case of infinite horizon that is

$$\tilde{V}(x) = \sup_{0\le\tau<\infty} \mathsf{E}_x(|X_\tau| - cX_\tau^2). \tag{6.18}$$

By setting $F(a) = a - ca^2$ for $a \in I\!R$, we have

$$\tilde{V}(x) = \sup_{0\le\tau<\infty} \mathsf{E}_x(F(|X_\tau|)). \tag{6.19}$$

We could calculate the maximum of function F by setting $F'(a) = 1 - 2ca = 0$ which implies $a = \frac{1}{2c}$ and $F(\frac{1}{2c}) = \frac{1}{4c}$. If we define a stopping time cccccc

4. We apply Wald-Bellman equations to the optimal stopping problems in (6.17). By (4.11) and Theorem 7 we obtain the identities

$$\tilde{V}^T(x) = \lim_{n\to\infty} \bar{V}_n^{T2^n}(x) \tag{6.20}$$

where

$$\bar{V}_n^N(x) = \sup_{\tau\in\mathcal{T}_n^N} \mathsf{E}_x G(X_\tau) = \max\big(G(x), \mathbb{P}_{2^{-n}}\bar{V}_n^{N-1}(x)\big) \tag{6.21}$$

for $\bar{V}_n^0(x) = G(x) = |x| - cx^2$, $x \in I\!R$, $n \ge 0$, $N \in I\!N$, $T2^n \in I\!N$ and $\mathcal{T}_n^N = \{0, 2^{-n}, 2\times 2^{-n}, ..., N\times 2^{-n}\}$. $\mathbb{P}_{2^{-n}}$ is the associated transition operator of X defined in (4.5).

5. The sections of the continuation set C and stopping set D are defined by

$$C_t = \{x \in I\!R \mid \tilde{V}^{T-t}(x) > G(x)\} \tag{6.22}$$

$$D_t = \{x \in I\!R \mid \tilde{V}^{T-t}(x) = G(x)\} \tag{6.23}$$

for $t \in [0,T]$. The first entry time τ_D defined by

$$\tau_D = \inf\{ t \in [0,T] \mid X_t \in D_t \}. \tag{6.24}$$

is the optimal stopping time for the problem (6.18) (cf. [21, Theorem 1.7]). The optimal stopping boundary between C and D is defined by

$$b^+(t) = \inf\{ x \in I\!R^+ \mid x \in D_t \} \tag{6.25}$$

$$b^-(t) = \sup\{ x \in I\!R^- \mid x \in D_t \}. \tag{6.26}$$

for $t \in [0,T]$ where $b^+(t)$ denotes the positive part of the boundary and $b^-(t)$ denotes the negative part of the boundary.

6. By Theorem 4, the rates of convergence in the Wald-Bellman equations in (6.21) is given by

$$\bar{V}_n^N(x) - \bar{V}_n^{N-1}(x) \le \int_{-\infty}^{\infty} \left(\mathbb{P}_{2^{-n}} G(y) - G(y)\right)^+ p\left((N-1)2^{-n}; x; y\right) dy \tag{6.27}$$

where $\mathbb{P}_{2^{-n}}$ is the associated transition operator of X defined in (4.5) and $p\left((N-1)2^{-n}; x; y\right)$ is the transition density function of Brownian motion given by

$$p\left((N-1)2^{-n}; x; y\right) = \sqrt{\frac{2^{n-1}}{\pi(N-1)}} \exp\left(\frac{2^{n-1}(x-y)^2}{1-N}\right) \tag{6.28}$$

with $n \ge 0$, $N \in I\!N$, x and y in $I\!R$.

7. We now introduce the algorithm of Wald-Bellman equations in Mathematica and the corresponding numerical analysis. The algorithm is similar to Chapter 5, but the optimal stopping problem is now seeking for supremum. The marginal distribution of one-dimensional Brownian motion is know explicitly, thus Euler approximation is not

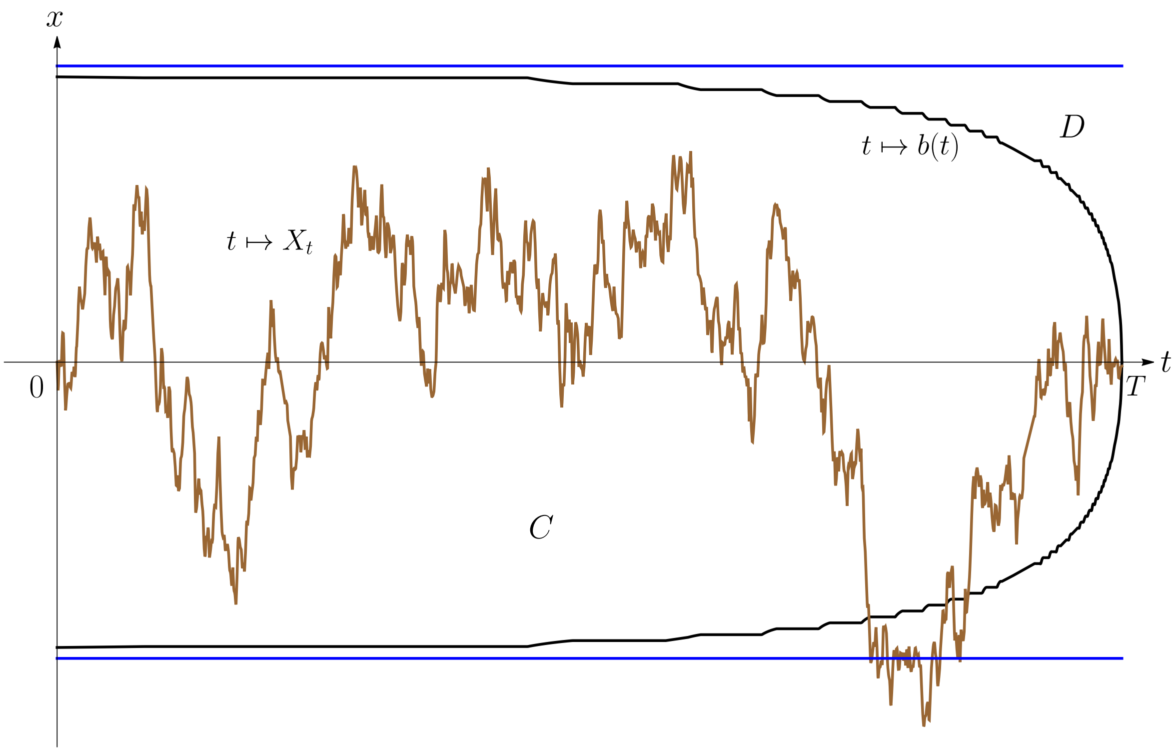

Figure 13. Simulated sample path of the Brownian motion X starting at 0. Numerical approximations of optimal stopping boundaries generated by Wald-Bellman equations for (6.20)+(6.21) with finite horizon for $T = 1$ and $c = 1$ (black). Analytical optimal stopping boundaries for (2.62) with infinite horizon (blue).

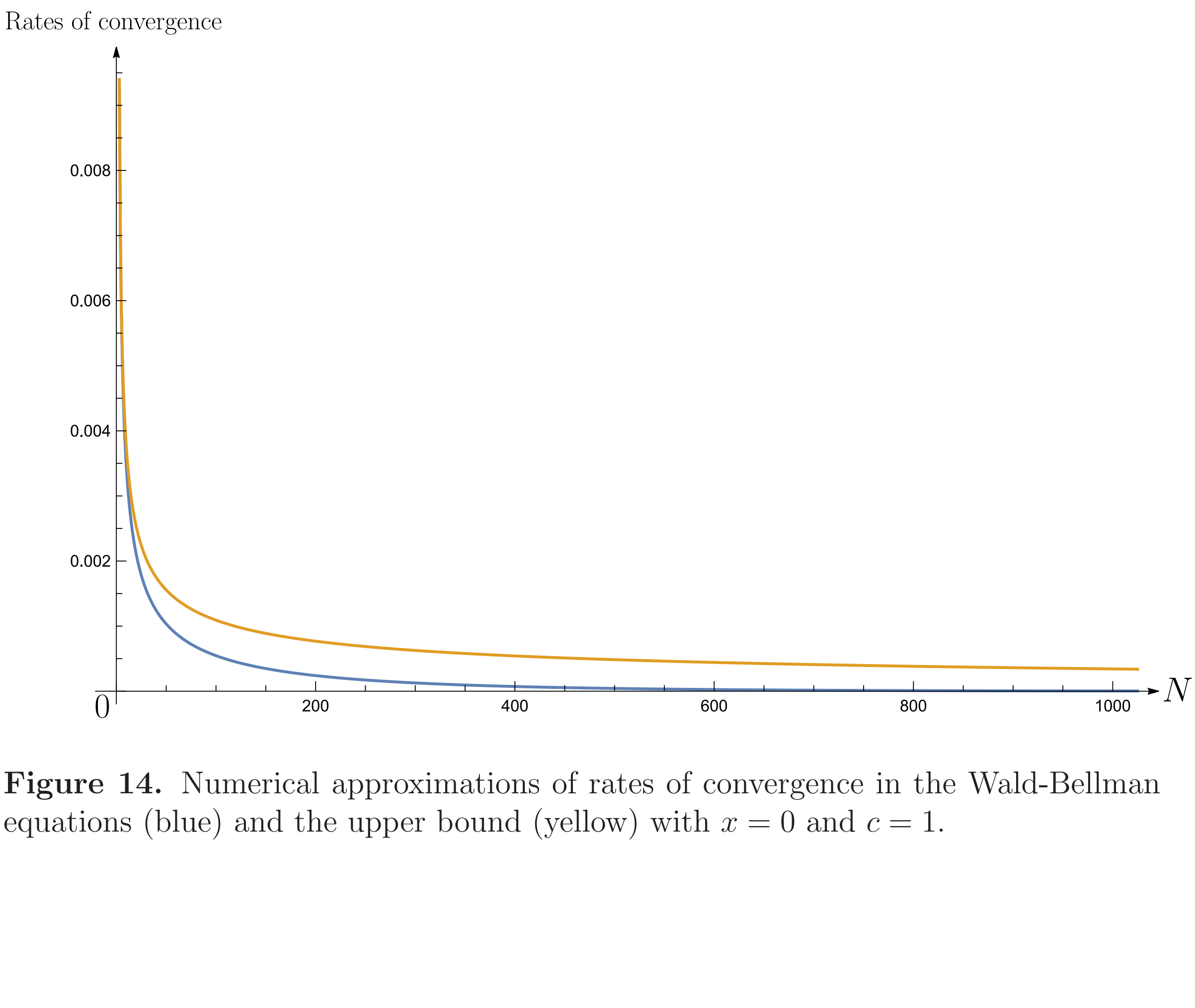

Figure 14. Numerical approximations of rates of convergence in the Wald-Bellman equations (blue) and the upper bound (yellow) with $x = 0$ and $c = 1$.

used. The iterations for Wald-Bellman equations in (6.20)+(6.21) are given by

$$\bar{V}_n^0(x) = G(x) = |x| - cx^2 \tag{6.29}$$

$$\begin{aligned}\bar{V}_n^1(x) &= \sup_{\tau \in \mathcal{T}_n^1} \mathsf{E}_x G(X_\tau) = \max\big(G(x), \mathbb{P}_{2^{-n}} \bar{V}_n^0(x)\big) \\ &= \max\big(G(x), \mathsf{E}_x \bar{V}_n^0(X_{2^{-n}})\big) \\ &= \max\left(G(x), \int_{-\infty}^{\infty} \bar{V}_n^0(y)\, p(2^{-n}; x; y)\, dy\right)\end{aligned} \tag{6.30}$$

$$\begin{aligned}\bar{V}_n^2(x) &= \sup_{\tau \in \mathcal{T}_n^2} \mathsf{E}_x G(X_\tau) = \max\big(G(x), \mathbb{P}_{2^{-n}} \bar{V}_n^1(x)\big) \\ &= \max\left(G(x), \int_{-\infty}^{\infty} \bar{V}_n^1(y)\, p(2^{-n}; x; y)\, dy\right)\end{aligned} \tag{6.31}$$

$$\vdots$$

$$\begin{aligned}\bar{V}_n^{T2^n}(x) &= \sup_{\tau \in \mathcal{T}_n^{T2^n}} \mathsf{E}_x G(X_\tau) = \max\big(G(x), \mathbb{P}_{2^{-n}} \bar{V}_n^{T2^n-1}(x)\big) \\ &= \max\left(G(x), \int_{-\infty}^{\infty} \bar{V}_n^{T2^n-1}(y)\, p(2^{-n}; x; y)\, dy\right)\end{aligned} \tag{6.32}$$

for $x \in I\!R$, $n \geq 0$, $T2^n \in I\!N$, $N \in I\!N$ and $\mathcal{T}_n^N = \{0, 2^{-n}, 2 \times 2^{-n}, ..., N \times 2^{-n}\}$ where p is the transition density function of Brownian motion staring at x under P_x defined in (6.28) with $t > 0$ and $y \in I\!R$. We now illustrate an example code in Mathematica. The coeffcients are defined by

Code 4.

```
h=0.5^8.; time=1.; w=time/h; c=1.;
gainfunction[x_]:=Abs[x]-c*Abs[x]^2; xubound= 1.;
xlbound=-1.; xsize=200.; xstep=(xubound-xlbound)/xsize;
functionlist={gainfunction}; intbound=Infinity;
```

where `h` denotes 2^{-n} for $n \geq 0$ in (6.21), `time` denotes T, `w` is the number of iterations $T2^n \in I\!N$, `c` is the constant $c \in I\!R^+$ in (6.1), `gainfunction` is G in (6.21), `xlbound` and `xubound` are the lower and upper bound for interpolation functions respectively, `xsize` is the number of grids used for interpolation in each iteration and `intbound` is used for the integration intervals to compute expectations in Wald-Bellman equations. Since the problem is one-dimensional which is less computation intensive, we compute the expectation by integrating from $-\infty$ to ∞ instead of truncating the interval in Chapter 5. The generated value functions are stored as interpolation functions in `functionlist` where `functionlist` $= \{\bar{V}_n^0, \bar{V}_n^1, \bar{V}_n^2, ..., \bar{V}_n^{T2^n}\}$. The algorithm for generating value functions is similar to Code 2 whereas the problem is now looking for a supremum instead of an infimum.

Code 5.

```
Do[Subscript[valuefunctionexact,n][x_]:=
 Max[gainfunction[x], NIntegrate[
  functionlist[[n]][b]/Sqrt[2.*Pi*h]*Exp[-0.5/h*(b - x)^2.],
   {b,-intbound,intbound},Method->{Automatic,"SymbolicProcessing"->0},
    AccuracyGoal->10]];
 Subscript[table,n]=N[Table[{x,Subscript[valuefunctionexact,n][x]},
  {x,xlbound,xubound,xstep}]];
 Subscript[valuefunction,n]=Interpolation[Subscript[table,n]];
 functionlist=Insert[functionlist,Subscript[valuefunction,n],-1],{n,w}].
```

8. The construction of value functions $\bar{V}_n^N$ defined in (6.21) through Wald-Bellman equations can be imaged as a rope put on the gain function G being pulled up towards the exact value function with infinite horizon $\tilde{V}$ defined in (6.18). The rates of convergence in Figure 14 illustrates the speed that the rope being pulled up. The more iterations conducted in Wald-Bellman equations, the closer the approximated value functions to $\tilde{V}$ as shown in Figure 15. The finite-horizon Wald-Bellman equations converge to the infinite-horizon optimal stopping problem as the number of iterations goes to infinity. Figure 16 and 17 illustrate the smooth fit of the value function at touching points with the gain function.

9. The optimal stopping boundary is constructed based on the definitions in (6.25) +(6.26). The example code in Mathematica is given as follows

Code 6.

```
boundarylistpositive=Table[{N[(w+1-k)*h],
 N[x/.FindRoot[functionlist[[k]][x]-gainfunction[x], {x, 0.000001}]]},
```

```
 {k,w + 1}];

boundarylistnegative=Table[{N[(w+1-k)*h],
 N[x/.FindRoot[functionlist[[k]][x]-gainfunction[x], {x, -0.000001}]]},
  {k,w + 1}];

boundary=Flatten[{Reverse[boundarylistpositive],boundarylistnegative},
 1];
```

We recall that `Reverse` is a built-in function which reverts the order of the `List`. Since the optimal stopping boundary is defined as two parts in (6.25)+(6.26), we generate the boundaries separately and combine them. `boundarylistpositive` is the positive part defined in (6.25) with the form $(t, b^+(t))$. `boundarylistnegativ` is the negative part defined in (6.26) with the form $(t, b^-(t))$. `boundary` is the combined `List` of data points. The construction of optimal stopping boundaries along with Wald-Bellman equations can be imaged as a rope being pulled towards the infinite-horizon boundary as shown in Figure 18. The numerical approximation through Wald-Bellman equations converges to the infinite-horizon optimal stopping boundary as the iteration continues which complies with Theorem 6. The smoothness of the optimal stopping boundary is controlled by the number of grids used for interpolation as shown in Figure 19. The more grids used, the smoother and more accurate the boundary would be.

10. Since the transition density function p is known, we can compute the upper bound on rates of convergence in the Wald-Bellman equations explicitly based on (6.27)+(6.28). We recall that the upper bounds on the rates of convergence in the Wald-Bellman equations (6.21) are given by

$$\bar{V}_n^N(x) - \bar{V}_n^{N-1}(x) \leq \int_{-\infty}^{\infty} \left[\bar{V}_n^1(y) - G(y)\right] p\left((N-1)2^{-n}; x; y\right) dy \tag{6.33}$$

for p defined in (6.28). The Mathematica code for computing the rates of convergence and the corresponding upper bounds when $x = 0$ are given as follows

Code 7.

```
ratetable=Table[{i,Abs[functionlist[[i]][0.]-functionlist[[i-1]][0.]]},
 {i,3,w + 1}];

boundfunction[m_]:=NIntegrate[(functionlist[[2]][b]-
```

```
 gainfunction[b])/Sqrt[h*2*Pi*m]*Exp[-0.5/h*(b^2)/m],
  {b,-intbound,intbound},
   Method->{Automatic, "SymbolicProcessing"->0},AccuracyGoal->10];

boundlist=Table[N[{i+2, boundfunction[i]}], {i,1,w-1}];
```

where `ratetable` is a `List` containing the rates of convergence in the Wald-Bellman equations when $x = 0$ namely $\bar{V}_n^N(0) - \bar{V}_n^{N-1}(0)$. `ratetable` has the form $((N, \bar{V}_n^N(0) - \bar{V}_n^{N-1}(0)))$. `boundfunction` is a function which computes the upper bounds in (6.29). `boundlist` is a `List` containing the upper bounds on the rates of convergence in the Wald-Bellman equations when $x = 0$ which has the form $((N, \int_{-\infty}^{\infty} \left[\bar{V}_n^1(y) - G(y)\right] p\,((N-1)2^{-n}; 0; y)\, dy))$. n is given and fixed.

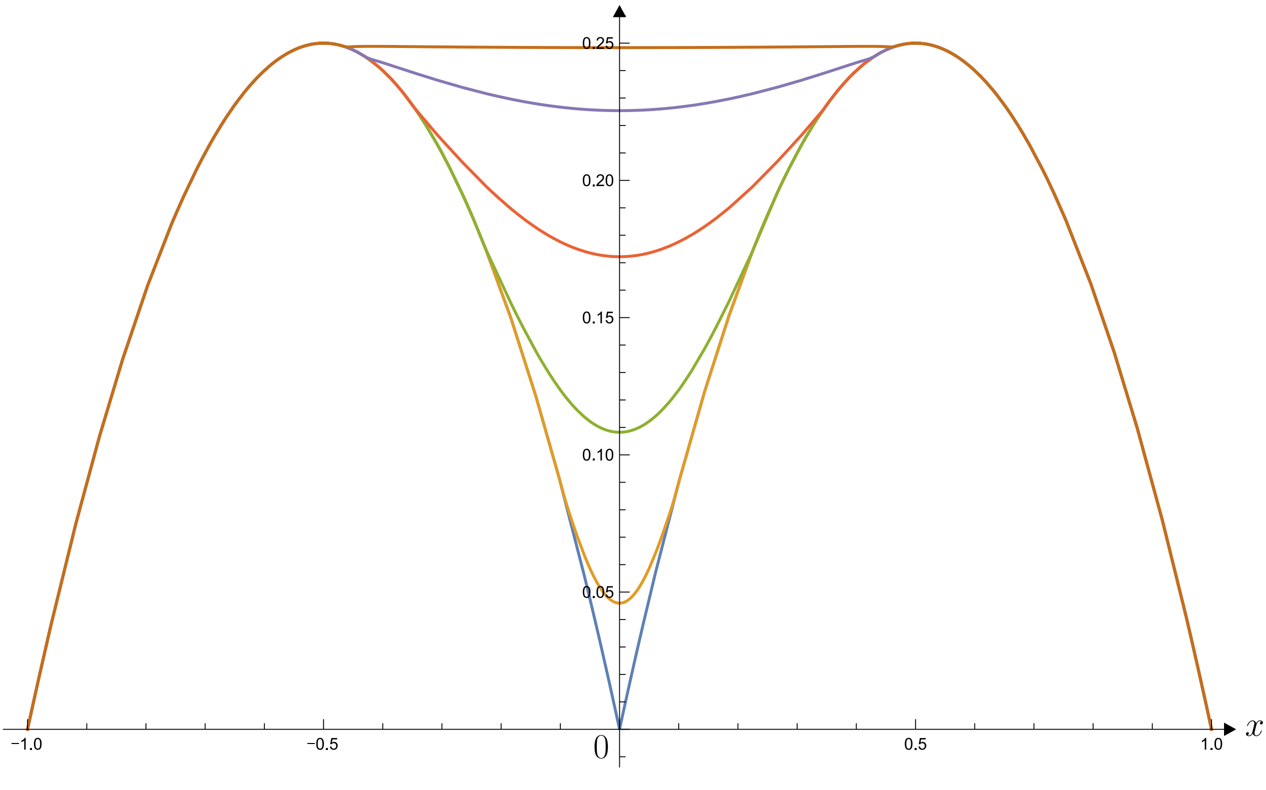

Figure 15. Numerical approximations of value functions with $c = 1$, $n = 8$ and $T = 2^{-n}$ (yellow), $T = 7 \times 2^{-n}$ (green), $T = 25 \times 2^{-n}$ (red), $T = 75 \times 2^{-n}$ (purple) and $T = 256 \times 2^{-n}$ (brown). Gain function G (blue).

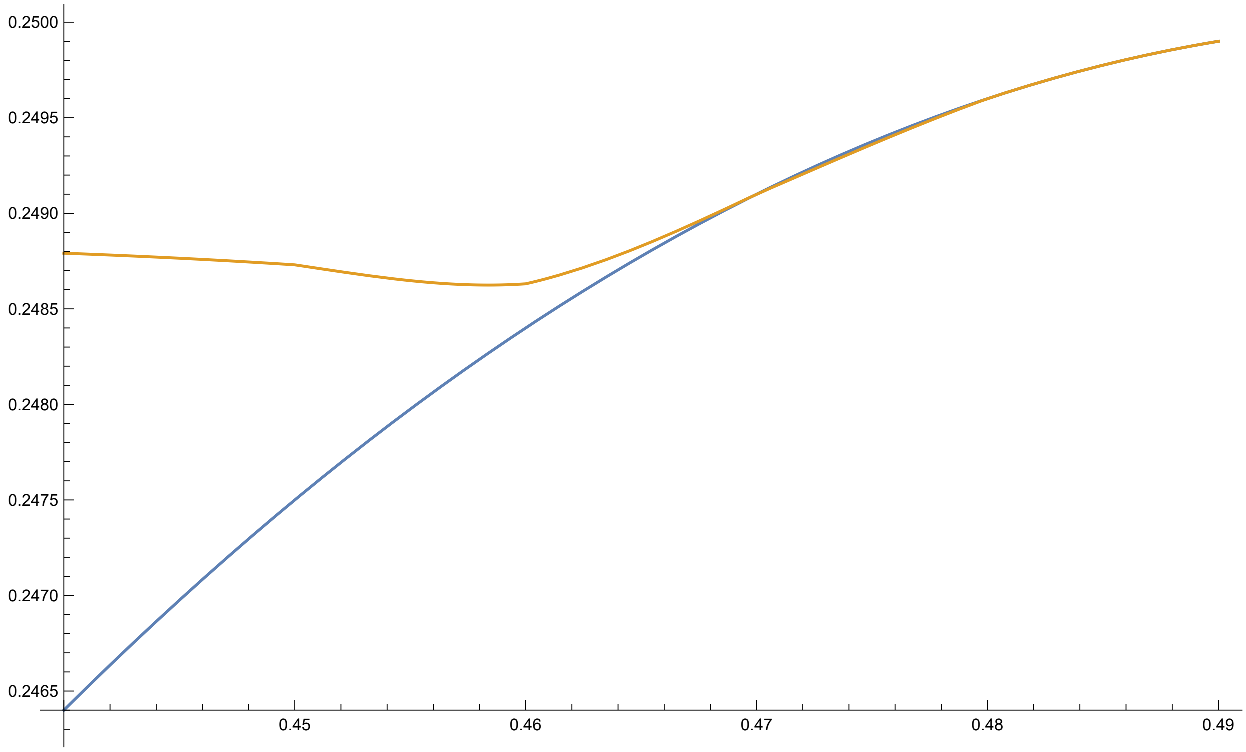

Figure 16. Right touching points of value function with 256 iterations (yellow) and gain function (blue) according to Figure 15.

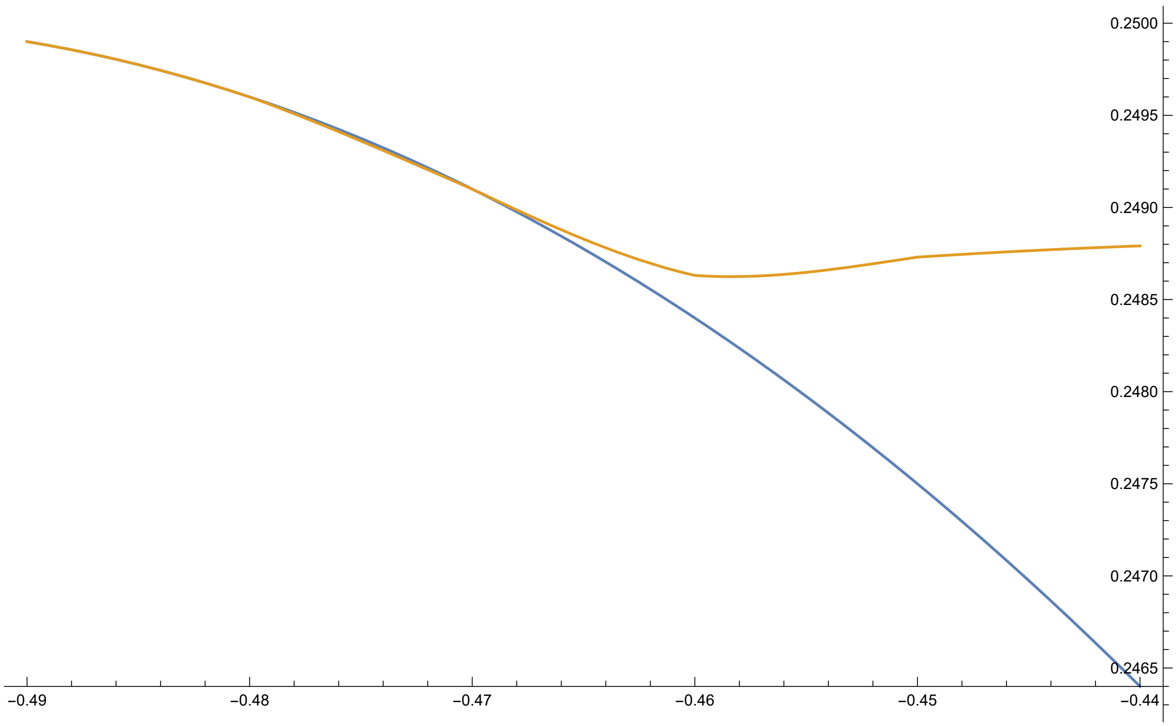

Figure 17. Left touching points of value function with 256 iterations (yellow) and gain function (blue) according to Figure 15.

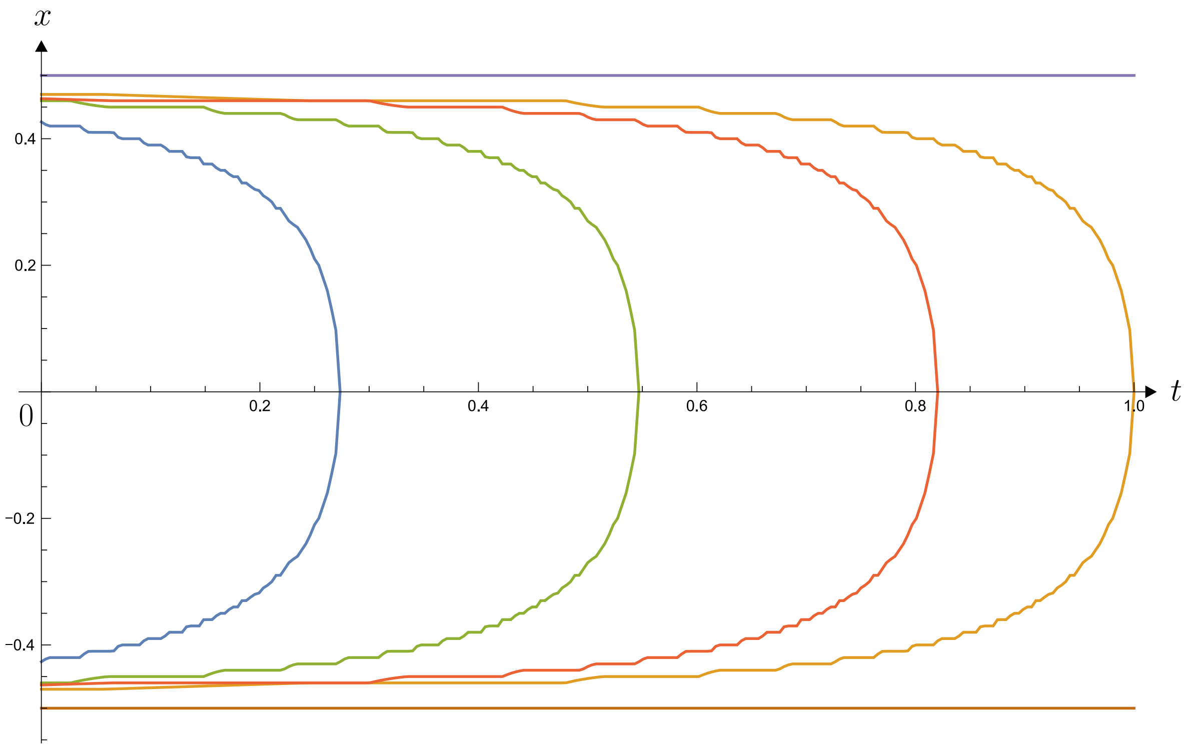

Figure 18. Numerical approximations of optimal stopping boundaries with $c = 1$, $n = 8$ and $T = 70 \times 2^{-n}$ (blue), $T = 140 \times 2^{-n}$ (green), $T = 210 \times 2^{-n}$ (red) and $T = 256 \times 2^{-n}$ (yellow). Optimal stopping boundaries with infinite horizon (purple and brown).

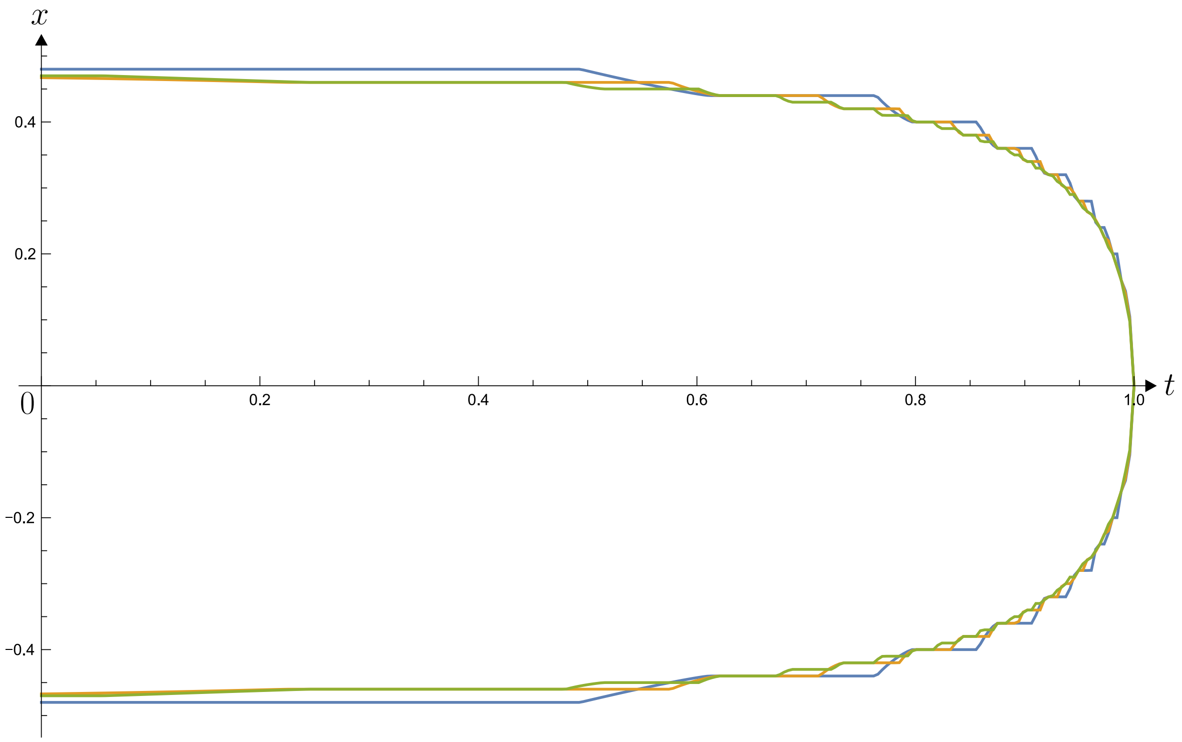

Figure 19. Numerical approximations of optimal stopping boundaries through Wald-Bellman equations with number of grids used for interpolation 50 (blue), 100 (yellow) and 200 (green).

Chapter 7

Applications to two-dimensional diffusion process problem

In this Chapter, we consider a nonlinear time-inhomogeneous optimal stopping problem for one-dimensional Brownian motion with *finite horizon*. This optimal stopping problem was introduced in [19, Example 3.2].

1. Consider the optimal stopping problem

$$V(t,x) = \sup_{0\le\tau\le 1-t} \mathsf{E}_{t,x}\left((1-t-\tau)\,X_{t+\tau}\right) \tag{7.1}$$

where τ is a stopping time of the Brownian motion $X = (X_{t+s})_{0\le s\le 1-t}$ satisfying $X_{t+s} = x + B_s$. $B = (B_t)_{t\ge 0}$ is a standard Brownian motion starting at 0 under $\mathsf{P}_{t,x}$. $\mathsf{P}_{t,x}$ is the probability measure under which $X_t = x \in I\!R$ for $t \in [0,1)$.

2. By Theorem 8 and (4.16), the Wald-Bellman equations satisfy the identities

$$V(t,x) = \lim_{n\to\infty} \bar{V}_n^{2^n}(t,x) \tag{7.2}$$

where

$$\bar{V}_n^N(t,x) = \sup_{\tau\in\mathcal{T}_n^{N-t2^n}} \mathsf{E}_{t,x} G(t+\tau, X_{t+\tau}) = \max\big(G(t,x), \mathbb{P}_{2^{-n}}\bar{V}_n^N(t,x)\big) \tag{7.3}$$

for $\mathbb{P}_{2^{-n}}\bar{V}_n^N((N-1)2^{-n},x) = \mathsf{E}_{(N-1)2^{-n},x} G(N2^{-n}, X_{N2^{-n}})$, $t \in \mathcal{T}_n^{N-1}$, $n \ge 0$, $2^n \in I\!N$, $x \in I\!R$, $N \in I\!N$, $\mathcal{T}_n^{N-1} = \{0, 2^{-n}, 2\times 2^{-n}, ..., (N-1)\times 2^{-n}\}$ and $\mathcal{T}_n^{N-t2^n} = \{0, 2^{-n}, 2\times 2^{-n}, ..., (N-t2^n)\times 2^{-n}\}$. $\mathbb{P}_{2^{-n}}V_n^N$ is the transition operator defined in (4.17). The gain function G is defined by $G(t,x) = (1-t)x$.

3. The continuation set C and stopping set D are defined by

$$C = \{\, (t, x) \in [0, 1) \times I\!R \mid V(t, x) > G(t, x) \,\} \tag{7.4}$$

$$D = \{\, (t, x) \in [0, 1) \times I\!R \mid V(t, x) = G(t, x) \,\}. \tag{7.5}$$

The first entry time of $(t, X_t)_{0 \le t < 1}$ into D defined by

$$\tau_D = \inf\{\, t \in [0, 1) \mid (t, X_t) \in D \,\} \tag{7.6}$$

is the optimal stopping time (cf. [21, Theorem 1.9]). The optimal stopping boundary between C and D is defined by

$$b(t) = \inf\{\, x \in I\!R \mid (t, x) \in D \,\} \tag{7.7}$$

for $t \in [0, 1)$.

4. We now introduce the algorithm of time-inhomogeneous Wald-Bellman equations in Mathematica and the corresponding numerical analysis. The idea of time-inhomogeneous Wald-Bellman equations starts with the corresponding gain function and continues backwardly. Consider the Wald-Bellman equations in (7.2)+(7.3), the iterations are given by

$$\bar{V}_n^{2^n}((2^n - 1)2^{-n}, x) = \sup_{\tau \in \mathcal{T}_n^1} \mathsf{E}_{(2^n-1)2^{-n}, x} G\left((2^n - 1)2^{-n} + \tau, X_{(2^n-1)2^{-n}+\tau}\right) \tag{7.8}$$

$$\begin{aligned} \bar{V}_n^{2^n}(1 - 2^{-n}, x) &= \sup_{\tau \in \mathcal{T}_n^1} \mathsf{E}_{1-2^{-n}, x} G\left(1 - 2^{-n} + \tau, X_{1-2^{-n}+\tau}\right) \\ &= \max\left(G(1 - 2^{-n}, x), \mathbb{P}_{2^{-n}} \bar{V}_n^{2^n}(1 - 2^{-n}, x)\right) \\ &= \max\left(G(1 - 2^{-n}, x), \mathsf{E}_{1-2^{-n}, x} G(1, X_1)\right) \\ &= \max\left(G(1 - 2^{-n}, x), 0\right) \end{aligned} \tag{7.9}$$

$$\begin{aligned} \bar{V}_n^{2^n}((2^n - 2)2^{-n}, x) &= \sup_{\tau \in \mathcal{T}_n^2} \mathsf{E}_{(2^n-2)2^{-n}, x} G\left((2^n - 2)2^{-n} + \tau, X_{(2^n-2)2^{-n}+\tau}\right) \\ \bar{V}_n^{2^n}(1 - 2 \times 2^{-n}, x) &= \sup_{\tau \in \mathcal{T}_n^2} \mathsf{E}_{1-2\times 2^{-n}, x} G\left(1 - 2 \times 2^{-n} + \tau, X_{1-2\times 2^{-n}+\tau}\right) \\ \bar{V}_n^{2^n}(1 - 2^{-n+1}, x) &= \max\left(G(1 - 2^{-n+1}, x), \mathbb{P}_{2^{-n}} \bar{V}_n^{2^n}(1 - 2^{-n+1}, x)\right) \\ &= \max\left(G(1 - 2^{-n+1}, x), \mathsf{E}_{1-2^{-n+1}, x} \bar{V}_n^{2^n}(1 - 2^{-n}, X_{1-2^{-n}})\right) \\ &= \max\left(G(1 - 2^{-n+1}, x), \int_{-\infty}^{\infty} \bar{V}_n^{2^n}(1 - 2^{-n}, y)\, p(2^{-n}; x; y)\, dy\right) \end{aligned} \tag{7.10}$$

$$\vdots$$

$$\begin{aligned}\bar{V}_n^{2^n}(0,x) &= \sup_{\tau\in\mathcal{T}_n^{2^n}} \mathsf{E}_{0,x}G(\tau,X_\tau) = \max\big(G(0,x),\mathbb{P}_{2^{-n}}\bar{V}_n^{2^n}(0,x)\big)\\ &= \max\big(G(0,x),\mathsf{E}_{0,x}\bar{V}_n^{2^n}(2^{-n},X_{2^{-n}})\big)\\ &= \max\left(G(0,x),\int_{-\infty}^{\infty}\bar{V}_n^{2^n}(2^{-n},y)\,p(2^{-n};x;y)\,dy\right) \qquad (7.11)\end{aligned}$$

where p is the transition density function of Brownian motion defined in (6.28) with $2^n \in I\!N$, $x \in I\!R$ and $y \in I\!R$. We first construct $\bar{V}_n^{2^n}((2^n-1)2^{-n},x)$ based on the definition in (7.3), then construct $\bar{V}_n^{2^n}((2^n-2)2^{-n},x)$ by using results in (7.9) and continue constructing value functions iteratively until $\bar{V}_n^{2^n}(0,x)$. We now illustrate an example code in Mathematica. The coefficients are defined by

Code 8.

```
h=0.5^10.; time=1.; intbound = Infinity; xubound =1.; xlbound =-2.;
xsize = 200; xstep=(xubound-xlbound)/xsize; w=time/h;
gainfunction[t_, x_]:=(1-t)*x; functionlist={ };
```

where `time` is the upper bound on time horizon in (7.1). `xubound` and `xlbound` are the upper and lower bounds of the interpolation domain for x. `h` is 2^{-n} in (7.3). `w` is the number of total iterations defined as $2^n \in I\!N$ in (7.2). `intbound` is used to control the integral interval for the computation of expectations. `xsize` is the number of grids used for interpolation. `xstep` denotes the distance between the grids for interpolating x. `gainfunction` is G in (7.3). `functionlist` is a `list` of functions which stores the interpolated value functions. After the value iterations, the `functionlist` should consists of approximated value functions where `functionlist` $=$ $\{\bar{V}_n^{2^n}((2^n-1)2^{-n},x),\bar{V}_n^{2^n}((2^n-2)2^{-n},x),\bar{V}_n^{2^n}((2^n-3)2^{-n},x),...,\bar{V}_n^{2^n}(0,x)\}$. The difference between time-homogeneous and time-inhomogeneous is that the index $N \in I\!N$, which denotes the number of iterations now, is also a coordinate of the value function. The time coordinates of the value functions in `functionlist` is not a variable which means that these value functions can be considered as one-dimensional functions in x. The gain function G is two-dimensional. Therefore, we first generate the value function in (7.8), then construct the value functions iteratively by considering the value function as one-dimensional functions. The example code in Mathematica is given by

Code 9.

```
Subscript[valuefunctionexact, 1][x_]:=
 Max[gainfunction[(w-1)*h,x],NIntegrate[
  gainfunction[w*h,b]/Sqrt[2*Pi*h]*Exp[-0.5/h*(b-x)^2],
   {b,-intbound,intbound},Method->{Automatic,"SymbolicProcessing"->0}
    AccuracyGoal->10]];

Subscript[table,1]=N[
 Table[x,Subscript[valuefunctionexact,1][x],x,xlbound,xubound,xstep]];

Subscript[valuefunction,1]=Interpolation[Subscript[table,1]];
functionlist=Insert[functionlist,Subscript[valuefunction,1],-1];.
```

The value function $\bar{V}_n^{2^n}((2^n-1)2^{-n},x)$ has now been constructed. Although the function has two coordinates, the time coordinate $(2^n-1)2^{-n}$ is given and fixed. We then construct the remaining value functions iteratively by treating them as one-dimensional functions. The example code in Mathematica is given by

Code 10.

```
AbsoluteTiming[Do[Subscript[valuefunctionexact,n][x_]:=
 Max[gainfunction[(w-n)*h,x],NIntegrate[
  functionlist[[n-1]][b]/Sqrt[2*Pi*h]*Exp[-0.5/h*(b-x)^2],
   {b,-intbound,intbound},Method->{Automatic,"SymbolicProcessing"->0},
    AccuracyGoal->10]];

 Subscript[table,n]=N[
  Table[x,Subscript[valuefunctionexact,n][x],{x,xlbound,xubound,xstep}]];

 Subscript[valuefunction,n]=Interpolation[Subscript[table,n]];

 functionlist=Insert[functionlist,Subscript[valuefunction,n],-1],
                  {n,2,w}]].
```

Since the time coordinates of the value functions in `functionlist` is fixed, we need to restore the two-dimensional value function defined in (7.2)+(7.3). The idea to construct an interpolation function based on data grids generated by `functionlist`. The example code in Mathematica is given as follows

Code 11.

```
valuetabletx=Flatten[Table[{(w-n+1)*h,x,functionlist[[n]][x]},
 {x,xlbound,xubound,xstep},{n,w}],1];

valuetx=Interpolation[valuetabletx];
```

where `valuetabletx` is a `List` which stores data pairs with the form $\{\{t, x, \bar{V}_n^{2^n}(t, x)\}\}$. `valuetx` is the interpolated value function generated from `valuetabletx`. The value function $\bar{V}$ in (7.1) is non-negative because one could always to choose stop at the terminal time as shown in Figure 21. The value functions generated by Wald-Bellman equations comply with smooth fit at the touching points between the value functions and the gain function.

5. We construct the optimal stopping boundary by the definition in (7.7). The example code in Mathematica is given as follows

Code 12.

```
boundary= Table[{(w-k)*h,N[x/.FindRoot[
 functionlist[[k]][x]-gainfunction[(w-k)*h,x], {x, 0.005},
  AccuracyGoal->6,PrecisionGoal->6]]},{k,w}];.
```

The smoothness of optimal stopping boundary is controlled by the number of grids used for interpolation as shown in Figure 22. The more grids used, the smoother the boundary would be.

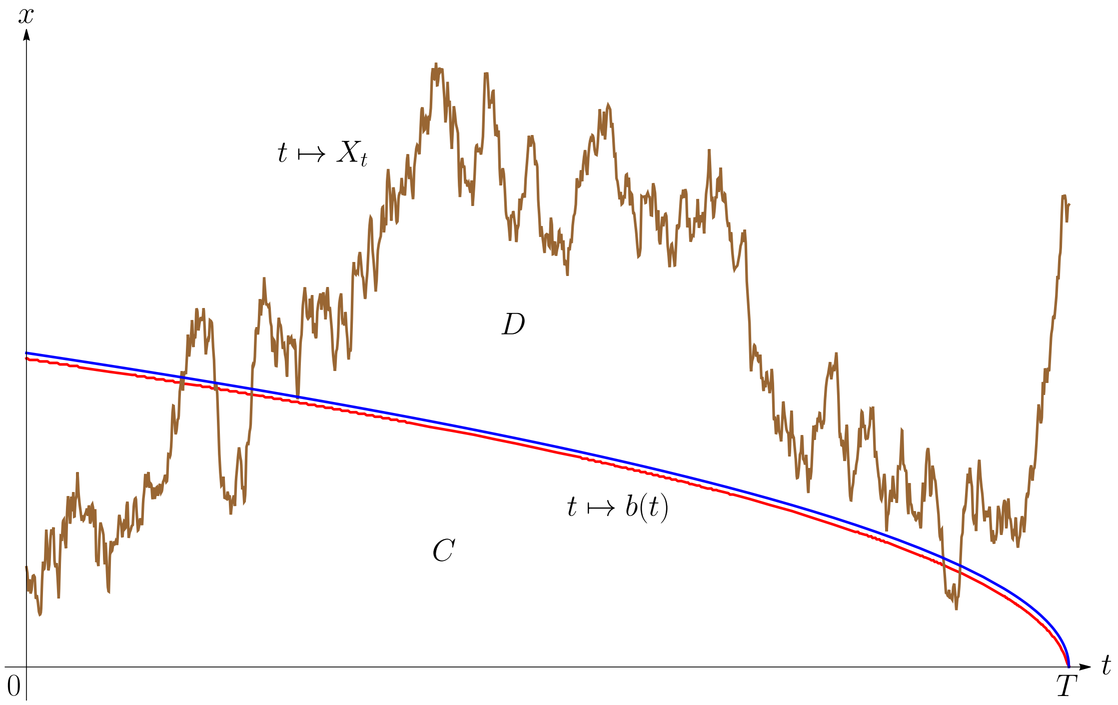

Figure 20. Simulated sample path of the Brownian motion X starting at 0.2. Numerical approximations of optimal stopping boundary generated by Wald-Bellman equations (6.20)+(6.21) with $n = 10$ and $c = 1$ (red). Analytical boundary $t \mapsto 0.63\sqrt{1-t}$ introduced in [19, Example 3.2] (blue).

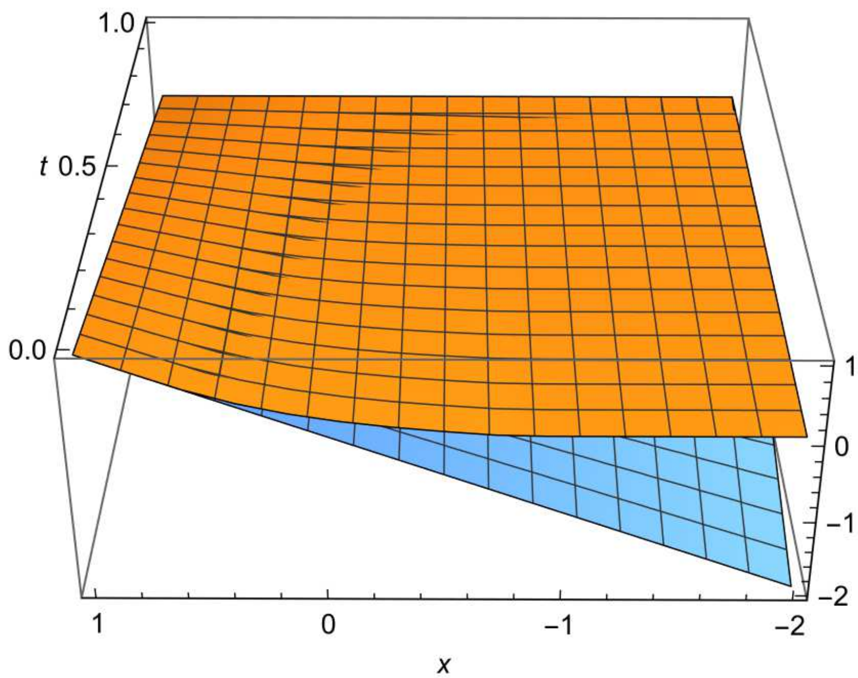

Figure 21. Numerical approximations of value function in (7.1) through Wald-Bellman equations (7.2)+(7.3) according to code 8, 9, 10 and 11 with $n = 10$ and $c = 1$ (yellow). Gain function (blue).

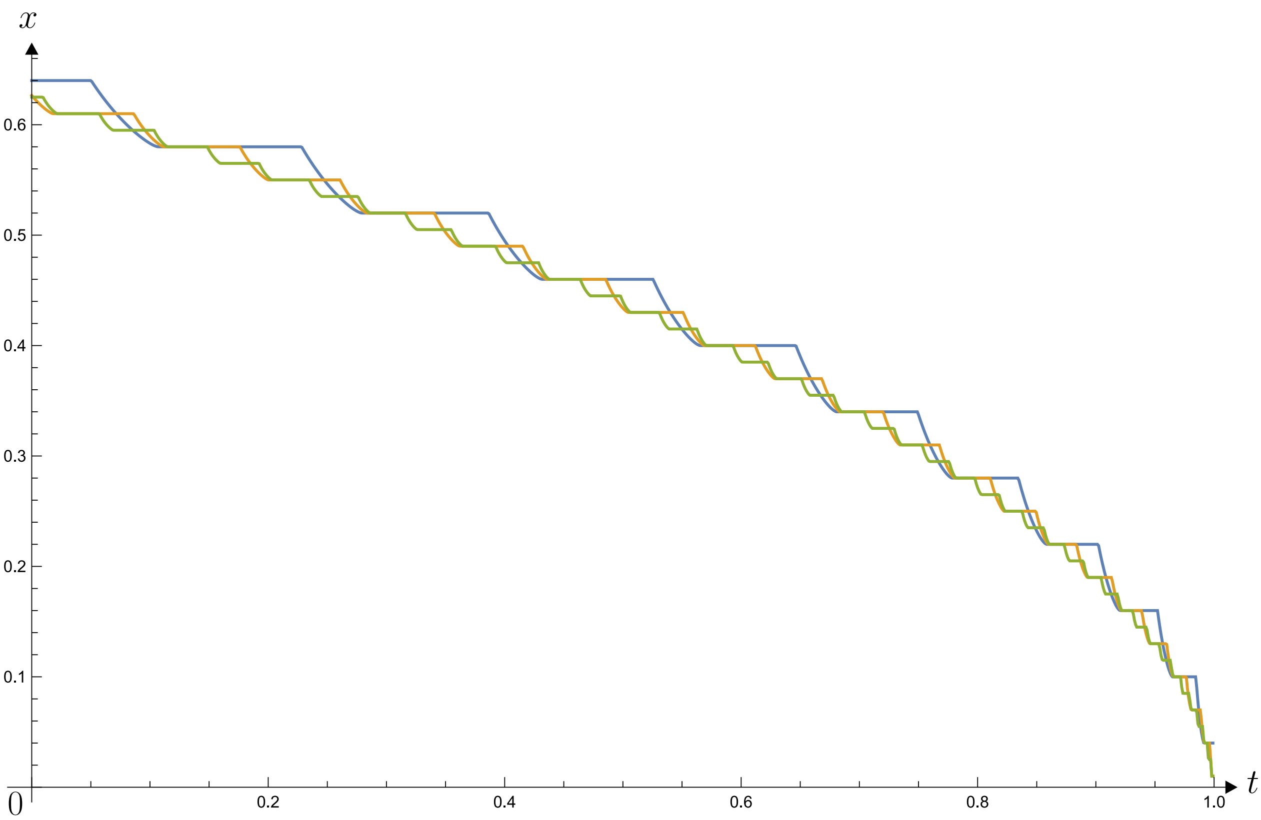

Figure 22. Numerical approximations of optimal stopping boundaries through Wald-Bellman equations according to Code 12 with the number of grids used for interpolation 50 (blue), 100 (yellow) and 200 (green).

Chapter 8

Applications to American put option

In this section, we consider the optimal stopping problem associated with the arbitrage-free price of the American put option with finite horizon (cf. [21, Chapter VII]). The optimal stopping problem will be discussed both as time-homogeneous and time-inhomogeneous problems.

1. The arbitrage-free price of the American put option with finite horizon is given by

$$V^T(t,x) = \sup_{0 \le \tau \le T-t} \mathsf{E}_{t,x}\left(e^{-r\tau}\left(K - X_{t+\tau}\right)^+\right) \tag{8.1}$$

for $t \in [0,T)$ and $X_t = x \in I\!R^+$ under $\mathsf{P}_{t,x}$ where τ is a stopping time of the geometric Brownian motion $X = (X_{t+s})_{s \ge 0}$ solving

$$dX_{t+s} = rX_{t+s}\,ds + \sigma X_{t+s}\,dB_s. \tag{8.2}$$

$B = (B_s)_{s \ge 0}$ is a standard Brownian motion starting at 0. $T > 0$ is the expiration date (maturity). $r > 0$ is the interest rate. $K > 0$ is the strike (exercise) price. $\sigma > 0$ is the volatility coefficient. The strong solution of (8.2) under $\mathsf{P}_{t,x}$ is given by

$$X_{t+s} = x\exp\left(\sigma B_s + \left(r - \sigma^2/2\right)s\right). \tag{8.3}$$

2. Since X is a time-homogeneous Markov process, (8.1) could be reformulated as a time-homogeneous optimal stopping problem (cf. [21, p. 16])

$$V^T(t,x) = \hat{V}^{T-t}(x) = \sup_{0 \le \tau \le T-t} \mathsf{E}_x\left(e^{-r\tau}\left(K - X_\tau\right)^+\right) \tag{8.4}$$

where the geometric Brownian motion $X = (X_s)_{s\geq 0}$ now satisfies

$$dX_s = rX_s\,ds + \sigma X_s\,dB_s \tag{8.5}$$

with the strong solution given by

$$X_s = x\exp\left(\sigma B_s + \left(r - \sigma^2/2\right)s\right) \tag{8.6}$$

under P_x with $X_0 = x \in I\!R^+$. The value function $\hat{V}^{T-t}(x)$ in (8.4) is equivalent to

$$\hat{V}^{T-t}(x) = \sup_{0\leq\tau\leq T-t} \mathsf{E}_x\left(e^{-r\tau}G(X_\tau)\right) \tag{8.7}$$

for gain function G defined by $G(x) = (K-x)^+$. $\hat{V}^{T-t}$ is a time-homogeneous optimal stopping problem.

3. By killing the sample path of X at rate r (cf. [21, Chapter 5.4 and Chapter 6.3]) we obtain the identities

$$\mathsf{E}_x\left[e^{-rt}G(X_t)\right] = \mathsf{E}_x\left[G(\tilde{X}_t)\right] \tag{8.8}$$

$$e^{-rt}p(t;x;y) = \tilde{p}(t;x;y) \tag{8.9}$$

where $\tilde{X}$ is the killed geometric Brownian motion starting at x, $\tilde{p}$ is the transition density function of the killed geometric Brownian motion $\tilde{X}$ and p is the transition density function of the geometric Brownian motion X under P_x. The optimal stopping problem in (8.7) is simplified to

$$\hat{V}^{T-t}(x) = \sup_{0\leq\tau\leq T-t} \mathsf{E}_x\left(G(\tilde{X}_\tau)\right) \tag{8.10}$$

for stopping times τ of $\tilde{X}$.

4. By Theorem 7 and (4.11), the Wald-Bellman equations satisfy the identities

$$\hat{V}^{T-t}(x) = \lim_{n\to\infty} \bar{V}_n^{(T-t)2^n}(x) \tag{8.11}$$

where we set

$$\bar{V}_n^N(x) = \sup_{\tau\in\mathcal{T}_n^N} \mathsf{E}_x G(\tilde{X}_\tau) = \max\left(G(x), \mathbb{P}_{2^{-n}}\bar{V}_n^{N-1}(x)\right) \tag{8.12}$$

for $\bar{V}_n^0(x) = G(x)$, $x \in I\!R^+$, $n \geq 0$, $t \in [0,T)$, $(T-t)2^n \in I\!N$, $N \in I\!N$ and $\mathcal{T}_n^N = \{0, 2^{-n}, 2\times 2^{-n}, ..., N\times 2^{-n}\}$. $\mathbb{P}_{2^{-n}}$ is the transition operator of $\tilde{X}$ defined by

$$\mathbb{P}_{2^{-n}}\bar{V}_n^{N-1}(x) = \mathsf{E}_x\bar{V}_n^{N-1}(\tilde{X}_{2^{-n}}) = \mathsf{E}_x e^{r2^{-n}}\bar{V}_n^{N-1}(X_{2^{-n}}). \tag{8.13}$$

5. The sections of the continuation set C and the stopping set D are defined by

$$C_t = \{\, x \in (0,\infty) \mid \hat{V}^{T-t}(x) > G(x) \,\} \tag{8.14}$$

$$D_t = \{\, x \in (0,\infty) \mid \hat{V}^{T-t}(x) = G(x) \,\} \tag{8.15}$$

for $t \in [0,T)$. The first entry time of the process X into D defined by

$$\tau_D = \inf\{\, t \in [0,T) \mid X_t \in D_t \,\} \tag{8.16}$$

is optimal in (8.10) (cf. [21, Theorem 1.7]). The optimal stopping boundary between C and D is defined by

$$b(t) = \sup\{\, x \in (0,\infty) \mid x \in D_t \,\} \tag{8.17}$$

for $t \in [0,T)$.

6. By Theorem 4, the rates of convergence in the Wald-Bellman equations (8.12) are given by

$$\begin{aligned}\bar{V}_n^N(x) - \bar{V}_n^{N-1}(x) &\leq \mathsf{E}_x \left(\bar{V}_n^1(\tilde{X}_{(N-1)2^{-n}}) - G(\tilde{X}_{(N-1)2^{-n}}) \right) \\ &= e^{-r(N-1)2^{-n}} \mathsf{E}_x \left(\bar{V}_n^1(X_{N2^{-n}}) - G(X_{N2^{-n}}) \right) \\ &= e^{-r(N-1)2^{-n}} \mathsf{E}_x \left(\mathbb{P}_{2^{-n}} G(X_{N2^{-n}}) - G(X_{N2^{-n}}) \right)^+ \end{aligned} \tag{8.18}$$

for $\bar{V}_n^1(x) = \max(G(x), \mathbb{P}_{2^{-n}} G(x))$, $x \in \mathbb{R}^+$, $n \geq 0$, $N \in \mathbb{N}$ and $r > 0$.

7. We now introduce the algorithm of the time-homogeneous Wald-Bellman equations in Mathematica and the corresponding numerical analysis. The algorithm is similar to Chpater 10. The iterations for Wald-Bellman equations in (8.11)+(8.12) are given by

$$\bar{V}_n^0(x) = G(x) = (K-x)^+ \tag{8.19}$$

$$\begin{aligned}\bar{V}_n^1(x) &= \sup_{\tau \in \mathcal{T}_n^1} \mathsf{E}_x G(\tilde{X}_\tau) = \max\left(G(x), \mathbb{P}_{2^{-n}} \bar{V}_n^0(x) \right) \\ &= \max\left(G(x), \mathsf{E}_x \bar{V}_n^0(\tilde{X}_{2^{-n}}) \right) \\ &= \max\left(G(x), \int_{-\infty}^{\infty} \bar{V}_n^0 \left(x \exp\left(\sigma y + (r - \sigma^2/2) 2^{-n} \right) \right) p(2^{-n}; y)\, dy \right) \end{aligned} \tag{8.20}$$

$$\begin{aligned}\bar{V}_n^2(x) &= \sup_{\tau \in \mathcal{T}_n^2} \mathsf{E}_x G(\tilde{X}_\tau) = \max\left(G(x), \mathbb{P}_{2^{-n}} \bar{V}_n^1(x) \right) \\ &= \max\left(G(x), \int_{-\infty}^{\infty} \bar{V}_n^1 \left(x \exp\left(\sigma y + (r - \sigma^2/2) 2^{-n} \right) \right) p(2^{-n}; y)\, dy \right) \end{aligned} \tag{8.21}$$

$$\vdots$$

$$\bar{V}_n^{(T-t)2^n}(x) = \sup_{\tau \in \mathcal{T}_n^{(T-t)2^n}} \mathsf{E}_x G(\tilde{X}_\tau) = \max\left(G(x), \mathbb{P}_{2^{-n}} \bar{V}_n^{(T-t)2^n-1}(x)\right)$$
$$= \max\left(G(x), \int_{-\infty}^{\infty} \bar{V}_n^{(T-t)2^n-1}\left(x \exp\left(\sigma y + (r-\sigma^2/2)2^{-n}\right)\right) p(2^{-n}; y)\, dy\right) \tag{8.22}$$

for $x \in I\!R^+$, $y \in I\!R$, $n \geq 0$, $t \in [0, T)$, $(T-t)2^n \in I\!N$, $N \in I\!N$ and $\mathcal{T}_n^N = \{0, 2^{-n}, 2 \times 2^{-n}, ..., N \times 2^{-n}\}$ where p is the transition density function of standard Brownian motion staring at 0 under P_x defined by

$$\frac{1}{\sqrt{2\pi t}} \exp\left(-\frac{y^2}{t}\right) \tag{8.23}$$

with $t > 0$ and $y \in I\!R$.

8. We illustrate an example code in Mathematica. The coefficients are defined by

Code 13.

```
h=0.5^12.; time=1.; w=time/h; r=0.1; k=1.; sigma=1.;
gainfunction[x_]:=Max[k-x,0.]; binf=k/(1.+sigma^2./(2.*r));
xubound= 2.; xlbound=0.; xsize=100.; xstep=(xubound-xlbound)/xsize;
functionlist={gainfunction}; intbound=Infinity;
Subscript[c,2]=(sigma^2./(2.*r))*(binf^(1.+(2.*r)/sigma^2.));
valuefunctionth=Piecewise[
 {{Subscript[c,2]*x^((-2.*r)/sigma^2.)),x >= binf},{k-x,0.<=x<binf}}];
```

where `h` denotes 2^{-n} for $n \geq 0$ in (8.12), `time` denotes T. `w` is the number of iterations $T2^n \in I\!N$. `r` is the interest rate. `k` is the strike price. `sigma` is the volatility. `gainfunction` is G in (8.7). `binf` is the optimal stopping boundary with infinite horizon (cf. [21, Chapter VII]). `xlbound` and `xubound` are the lower and upper bound for interpolation functions respectively. `xsize` is the number of grids used for interpolation in each iteration. `intbound` is used for the integration intervals to compute expectations in Wald-Bellman equations. `valuefunctionth` is the explicit value function with infinite horizon (cf. [21, Chapter VII]). Since the problem is one-dimensional which is less computation intensive, we compute the expectation by integrating from $-\infty$ to ∞. The generated value function are stored as interpolation functions in functionlist where functionlist $= \{\bar{V}_n^0, \bar{V}_n^1, \bar{V}_n^2, ..., \bar{V}_n^{T2^n}\}$. The value iterations in Mathematica are given as follows

Code 14.

```
Do[Subscript[valuefunctionexact,n][x_]:=
 Max[gainfunction[x],
  NIntegrate[
   Exp[-r*h]*
    functionlist[[n]][x*Exp[sigma*b+(r-(sigma^2)/2)*h]]/
     Sqrt[2*Pi*h]*Exp[-0.5/h*b^2],{b,-intbound,intbound},
     Method->{Automatic,"SymbolicProcessing"->0},AccuracyGoal->10]];
 Subscript[table,n]=N[Table[{x,Subscript[valuefunctionexact,n][x]},
  {x,xlbound,xubound,xstep}]];
 Subscript[valuefunction,n]=Interpolation[Subscript[table,n]];
 functionlist=Insert[functionlist,Subscript[valuefunction,n],-1],{n,w}].
```

We now restore the two-dimensional value function defined in (8.1) based on the identity in (8.4). The example code in Mathematica is given as follows

Code 15.

```
valuetabletx=Flatten[Table[{(w-n+1)*h,x,functionlist[[n]][x]},
 {x,xlbound,xubound,xstep},{n,w}],1];

valuetx=Interpolation[valuetabletx];
```

where valuetabletx is a List which stores data pairs with the form $\{\{t, x, \bar{V}_n^{(T-t)2^n}(t,x)\}\}$. valuetx is the interpolated value function generated from valuetabletx. The construction of value functions $\bar{V}_n^N$ defined in (8.12) through Wald-Bellman equations can be imaged as a rope put on the gain function G being pulled up towards the explicit value function with infinite horizon (cf. [21, Chapter VII]). The rates of convergence in Figure 24 illustrates the speed that the rope being pulled up. The more iterations conducted in Wald-Bellman equations, the closer the approximated value functions to the explicit value function as shown in Figure 25. The finite-horizon Wald-Bellman equations converge to the infinite-horizon optimal stopping problem as the number of iterations goes to infinity. Figure 25 and 26 illustrate the smooth fit of the value function at touching points with the gain function.

9. The optimal stopping boundary is constructed based on the definitions in (8.17). The example code in Mathematica is given as follows

Code 16.

```
boundarylist=Table[{N[(w+1-i)*h],
 N[x/.FindRoot[functionlist[[i]][x]-gainfunction[x],{x,0.999999}]]},
  {i,w + 1}];
```

The construction of optimal stopping boundaries along with Wald-Bellman equations can be imaged as a rope being pulled towards the infinite-horizon boundary as shown in Figure 27. The numerical approximation through Wald-Bellman equations converges to the infinite-horizon optimal stopping boundary as the iteration continues which complies with Proposition 6. The smoothness of optimal stopping boundary is controlled by the number of grids used for interpolation as shown in Figure 28. The more grids used, the smoother and more accurate the boundary would be.

10. We now consider (8.1) as a time-inhomogeneous optimal stopping problem. Since the gain function in (8.1) does not have time coordinate t, the value function V^T in (8.1) can be modified as

$$V^T(t,x) = e^{rt}\,\tilde{V}^T(t,x) \tag{8.24}$$

where we set

$$\tilde{V}^T(t,x) = \sup_{0\le\tau\le T-t} \mathsf{E}_{t,x}\hat{G}\left(t+\tau, X_{t+\tau}\right). \tag{8.25}$$

The gain function $\hat{G}$ is defined by $\hat{G}(t,x) = e^{-rt}\,(K-x)^+$ for $t \in [0,T)$, $x \in I\!R^+$ and $r > 0$. The problem in (8.25) is time-inhomogeneous.

11. By Theorem 8 and (4.16), the time-inhomogeneous Wald-Bellman equations satisfy the identities

$$\tilde{V}^T(t,x) = \lim_{n\to\infty} \breve{V}_n^{T2^n}(t,x) \tag{8.26}$$

where we set

$$\breve{V}_n^N(t,x) = \sup_{\tau\in\mathcal{T}_n^{N-t2^n}} \mathsf{E}_{t,x}\hat{G}(t+\tau, X_{t+\tau}) = \max\Big(\hat{G}(t,x), \mathbb{P}_{2^{-n}}\breve{V}_n^N(t,x)\Big) \tag{8.27}$$

for $t \in \mathcal{T}_n^{N-1}$, $n \geq 0$, $T2^n \in I\!N$, $N \in I\!N$, $\mathcal{T}_n^{N-1} = \{0, 2^{-n}, 2 \times 2^{-n}, ..., (N-1) \times 2^{-n}\}$ and $\mathcal{T}_n^{N-t2^n} = \{0, 2^{-n}, 2 \times 2^{-n}, ..., (N-t2^n) \times 2^{-n}\}$ where $\mathbb{P}_{2^{-n}} \breve{V}_n^N((N-1)2^{-n}, x) = \mathsf{E}_{(N-1)2^{-n},x} \hat{G}(N2^{-n}, X_{N2^{-n}})$. $\mathbb{P}_{2^{-n}}$ is the transition operator defined by

$$\mathbb{P}_{2^{-n}} \breve{V}_n^N(t, x) = \mathsf{E}_{t,x} \breve{V}_n^N(t + 2^{-n}, X_{t+2^{-n}}). \tag{8.28}$$

12. The continuation set $\hat{C}$ and stopping set $\hat{D}$ are defined by

$$\hat{C} = \{(t, x) \in [0, T) \times I\!R \mid \tilde{V}^T(t, x) > \hat{G}(t, x)\} \tag{8.29}$$

$$\hat{D} = \{(t, x) \in [0, T) \times I\!R \mid \tilde{V}^T(t, x) = \hat{G}(t, x)\}. \tag{8.30}$$

The first entry time of $(t, X_t)_{0 \leq t < T}$ into $\hat{D}$ defined by

$$\hat{\tau_D} = \inf\{t \in [0, T) \mid (t, X_t) \in \hat{D}\} \tag{8.31}$$

is the optimal stopping time in (8.25) (cf. [21, Theorem 1.9]). The optimal stopping boundary between $\hat{C}$ and $\hat{D}$ is defined by

$$\hat{b}(t) = \sup\{x \in (0, \infty) \mid (t, x) \in \hat{D}\} \tag{8.32}$$

for $t \in [0, T)$.

13. The algorithm is similar to Chapter 6. Consider the Wald-Bellman equations in (8.26)+(8.27), the iterations are given by

$$\breve{V}_n^{T2^n}((T2^n - 1)2^{-n}, x) = \sup_{\tau \in \mathcal{T}_n^1} \mathsf{E}_{(T2^n-1)2^{-n},x} \hat{G}\left((T2^n - 1)2^{-n} + \tau, X_{(T2^n-1)2^{-n}+\tau}\right) \tag{8.33}$$

$$\begin{aligned}
\breve{V}_n^{T2^n}(T - 2^{-n}, x) &= \sup_{\tau \in \mathcal{T}_n^1} \mathsf{E}_{T-2^{-n},x} \hat{G}\left(T - 2^{-n} + \tau, X_{T-2^{-n}+\tau}\right) \\
&= \max\left(\hat{G}(T - 2^{-n}, x), \mathbb{P}_{2^{-n}} \breve{V}_n^{T2^n}(T - 2^{-n}, x)\right) \\
&= \max\left(\hat{G}(T - 2^{-n}, x), \mathsf{E}_{t-2^{-n},x} \hat{G}(T, X_T)\right) \\
&= \max\Big(\hat{G}(T - 2^{-n}, x), \\
&\qquad \int_{-\infty}^{\infty} \hat{G}\left(T, x \exp\left(\sigma y + (r - \sigma^2/2)2^{-n}\right)\right) p(2^{-n}; y)\, dy\Big)
\end{aligned} \tag{8.34}$$

$$\breve{V}_n^{T2^n}((T2^n - 2)2^{-n}, x) = \sup_{\tau \in \mathcal{T}_n^2} \mathsf{E}_{(T2^n-2)2^{-n},x} \hat{G}\left((T2^n - 2)2^{-n} + \tau, X_{(T2^n-2)2^{-n}+\tau}\right)$$

$$\breve{V}_n^{T2^n}(T - 2 \times 2^{-n}, x) = \sup_{\tau \in \mathcal{T}_n^2} \mathsf{E}_{T-2\times2^{-n},x} \hat{G}\left(T - 2 \times 2^{-n} + \tau, X_{T-2\times2^{-n}+\tau}\right)$$

$$\begin{aligned}\breve{V}_n^{T2^n}(T-2^{-n+1},x) &= \max\Big(\hat{G}(T-2^{-n+1},x), \mathbb{P}_{2^{-n}}\breve{V}_n^{T2^n}(T-2^{-n+1},x)\Big)\\ &= \max\Big(\hat{G}(T-2^{-n+1},x), \mathsf{E}_{T-2^{-n+1},x}\breve{V}_n^{T2^n}(T-2^{-n},X_{T-2^{-n}})\Big)\\ &= \max\Big(\hat{G}(T-2^{-n+1},x),\\ &\qquad \int_{-\infty}^{\infty}\breve{V}_n^{T2^n}\left(T-2^{-n},x\exp\left(\sigma y+(r-\sigma^2/2)2^{-n}\right)\right)p(2^{-n};y)\,dy\Big)\end{aligned} \tag{8.35}$$

$$\vdots$$

$$\begin{aligned}\breve{V}_n^{T2^n}(0,x) = \sup_{\tau\in\mathcal{T}_n^{T2^n}}\mathsf{E}_{0,x}G(\tau,X_\tau) &= \max\Big(\hat{G}(0,x), \mathbb{P}_{2^{-n}}\breve{V}_n^{T2^n}(0,x)\Big)\\ &= \max\Big(\hat{G}(0,x), \mathsf{E}_{0,x}\breve{V}_n^{T2^n}(2^{-n},X_{2^{-n}})\Big)\\ &= \max\Big(\hat{G}(0,x),\\ &\qquad \int_{-\infty}^{\infty}\breve{V}_n^{T2^n}\left(2^{-n},x\exp\left(\sigma y+(r-\sigma^2/2)2^{-n}\right)\right)p(2^{-n};y)\,dy\Big)\end{aligned} \tag{8.36}$$

where p is the transition density function of Brownian motion defined in (6.28) with $t > 0$, $2^n \in I\!N$, $x \in I\!R$ and $y \in I\!R$. We first construct $\bar{V}_n^{2^n}((2^n-1)2^{-n},x)$ based on the definitions in (7.3), then construct $\bar{V}_n^{2^n}((2^n-2)2^{-n},x)$ by using results in (7.9) and continue constructing value functions iteratively until $\bar{V}_n^{2^n}(0,x)$. We now illustrate an example code in Mathematica. The coefficients are defined by

Code 17.

```
h=0.5^12.; time=1.; w=time/h; r=0.1; k=1.; sigma=1.;
gainfunction[x_]:=Exp[-r*t]*Max[k-x,0.]; binf=k/(1.+sigma^2./(2.*r));
xubound= 2.; xlbound=0.; xsize=100.; xstep=(xubound-xlbound)/xsize;
functionlist={gainfunction}; intbound=Infinity;
Subscript[c,2]=(sigma^2./(2.*r))*(binf^(1.+(2.*r)/sigma^2.));
valuefunctionth=Piecewise[
 {{Subscript[c,2]*x^((-2.*r)/sigma^2.)),x >= binf},{k-x,0.<=x<binf}}];
```

where the coefficients are defined similar to Code 13, but `gainfunction` is $\hat{G}$ in (8.25). `functionlist` is a `list` of functions which stores the interpolated value functions. After the value iterations, the `functionlist` should consists of approximated

value functions where `functionlist` =
$\{\breve{V}_n^{T2^n}((T2^n-1)2^{-n},x), \breve{V}_n^{T2^n}((T2^n-2)2^{-n},x), \breve{V}_n^{T2^n}((T2^n-3)2^{-n},x), ..., \breve{V}_n^{T2^n}(0,x)\}$.
We first generate the value function in (8.33), then construct the value functions iteratively by considering the value function as one-dimensional functions similarly to Code 9. The example code in Mathematica is given by

Code 18.

```
Subscript[valuefunctionexact, 1][x_]:=
 Max[gainfunction[(w-1)*h,x],NIntegrate[
  gainfunction[w*h,x*Exp[sigma*b+(r-(sigma^2)/2)*h]]/
   Sqrt[2*Pi*h]*Exp[-0.5/h*b^2],{b,-intbound,intbound},
    Method->{Automatic,"SymbolicProcessing"->0},
     AccuracyGoal->10]];

 Subscript[table,1]=N[
  Table[x,Subscript[valuefunctionexact,1][x],x,xlbound,xubound,xstep]];

 Subscript[valuefunction,1]=Interpolation[Subscript[table,1]];
 functionlist=Insert[functionlist,Subscript[valuefunction,1],-1];.
```

The value function $\breve{V}_n^{T2^n}((T2^n-1)2^{-n},x)$ has now been constructed. We then construct the remaining value functions iteratively by treating them as one-dimensional functions. The example code in Mathematica is given by

Code 19.

```
Do[Subscript[valuefunctionexact,n][x_]:=
 Max[gainfunction[(w-n)*h,x],NIntegrate[
  functionlist[[n-1]][x*Exp[sigma*b+(r-(sigma^2)/2)*h]]/
   Sqrt[2*Pi*h]*Exp[-0.5/h*b^2],{b,-intbound,intbound},
    Method->{Automatic,"SymbolicProcessing"->0},
     AccuracyGoal->10]];

 Subscript[table,n]=N[
  Table[x,Subscript[valuefunctionexact,n][x],{x,xlbound,xubound,xstep}]];

 Subscript[valuefunction,n]=Interpolation[Subscript[table,n]];
```

```
functionlist=Insert[functionlist,Subscript[valuefunction,n],-1],{n,2,w}].
```

We need to restore the two-dimensional value function defined in (7.2)+(8.25). The idea is similar to Code 11. The example code in Mathematica is given as follows

Code 20.

```
valuetabletx=Flatten[Table[{(w-n+1)*h,x,functionlist[[n]][x]},
 {x,xlbound,xubound,xstep},{n,w}],1];

valuetx=Interpolation[valuetabletx];
```

where `valuetabletx` is a `List` which stores data pairs with the form $\{\{t, x, \check{V}_n^{T2^n}(t, x)\}\}$. `valuetx` is the interpolated value function generated from `valuetabletx`. The value functions generated with time-homogeneous and time-inhomogeneous Wald-Bellman equations are identical. The value functions should in fact be strictly positive which is different from Figure 25. This is caused by the interpolating technique. The interpolated value functions could be more accurate by enlarging the interpolation interval and the computation time would be increased as well. If the value function is not of the interest, and the optimal stopping boundary is more important, then we could only focus on the touching points between the value functions and gain functions. Smooth fit at the touching points between the value functions and the gain function is persisted.

14. We construct the optimal stopping boundary by the definition in (8.32). The example code in Mathematica is given as follows

Code 21.

```
boundary= Table[{(w-k)*h,N[x/.FindRoot[
 functionlist[[k]][x]-gainfunction[(w-k)*h,x], {x, 0.005},
  AccuracyGoal->6,PrecisionGoal->6]]},{k,w}];.
```

The smoothness of optimal stopping boundary is controlled by the number of grids used for interpolation as shown in Figure 28. The more grids used, the smoother the boundary would be.

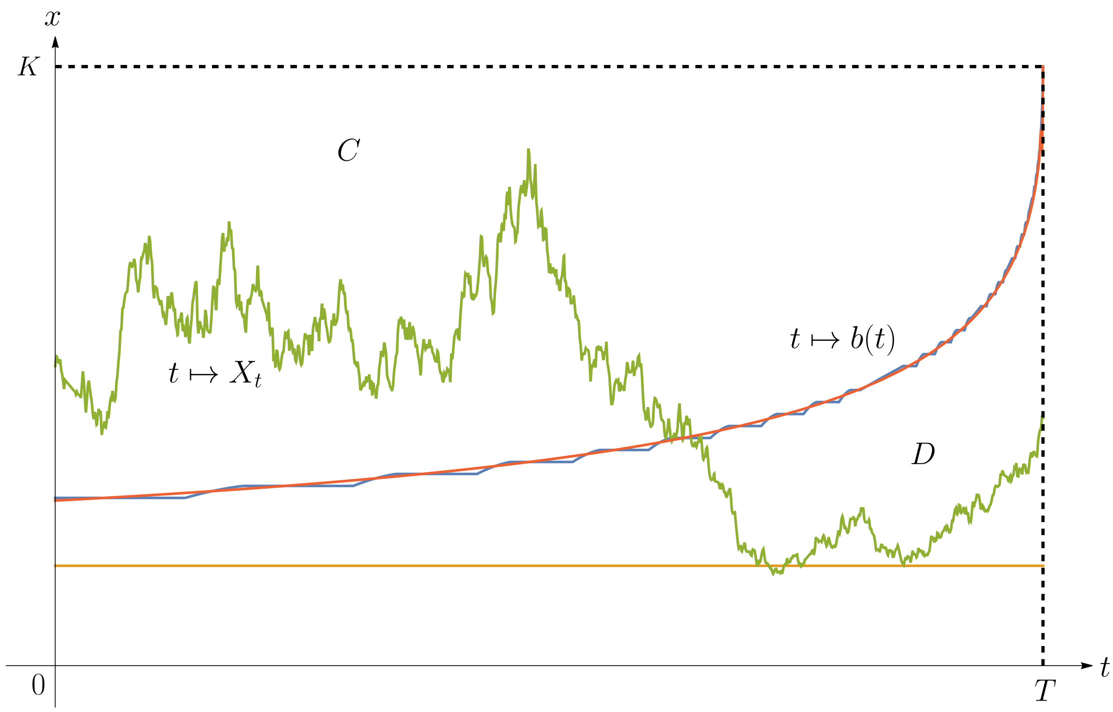

Figure 23. Simulated sample path of the geometric Brownian motion X starting at 0.5. Numerical approximations of optimal stopping boundary generated by Wald-Bellman equations in (8.11)+(8.12) and (8.26)+(8.27) with $T = 1$, $r = 0.1$, $k = 1$, $n = 12$ and $\sigma = 1$ (blue). Numerical approximations of optimal stopping boundary generated by nonlinear integral equation introduced in [21, Chapter 25.2] (red). Associated analytical boundary with infinite horizon introduced in [21, Chapter 25.1] (yellow).

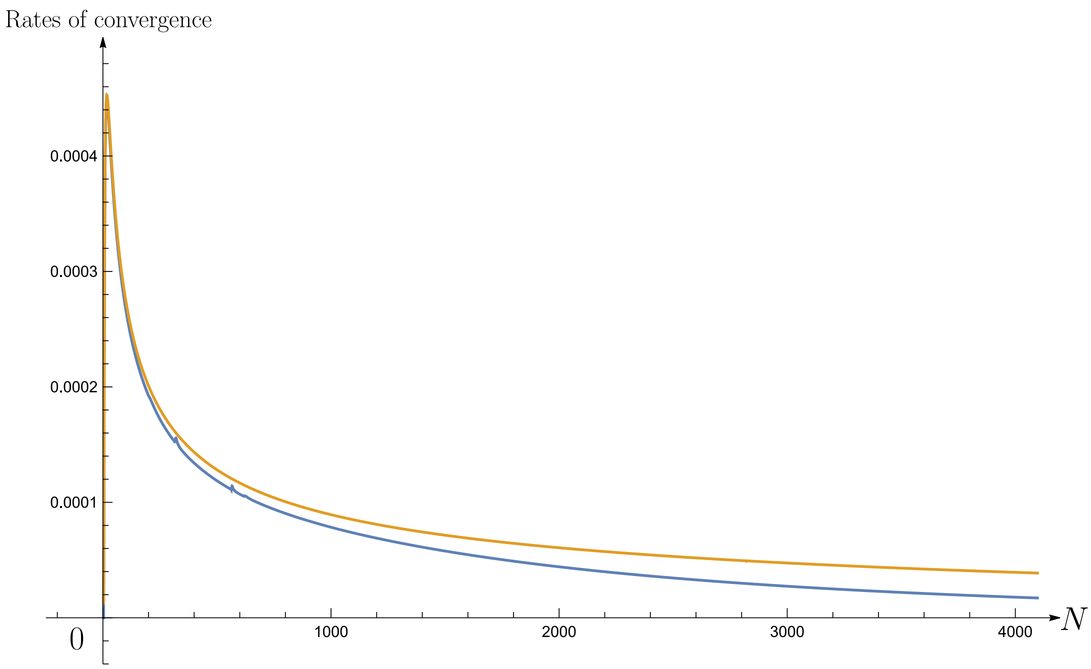

Figure 24. Numerical approximations of rates of convergence in the Wald-Bellman equations (blue) and the upper bound (yellow) in (8.18) $x = 0.94$, $T = 1$, $r = 0.1$, $k = 1$, $n = 12$ and $\sigma = 1$.

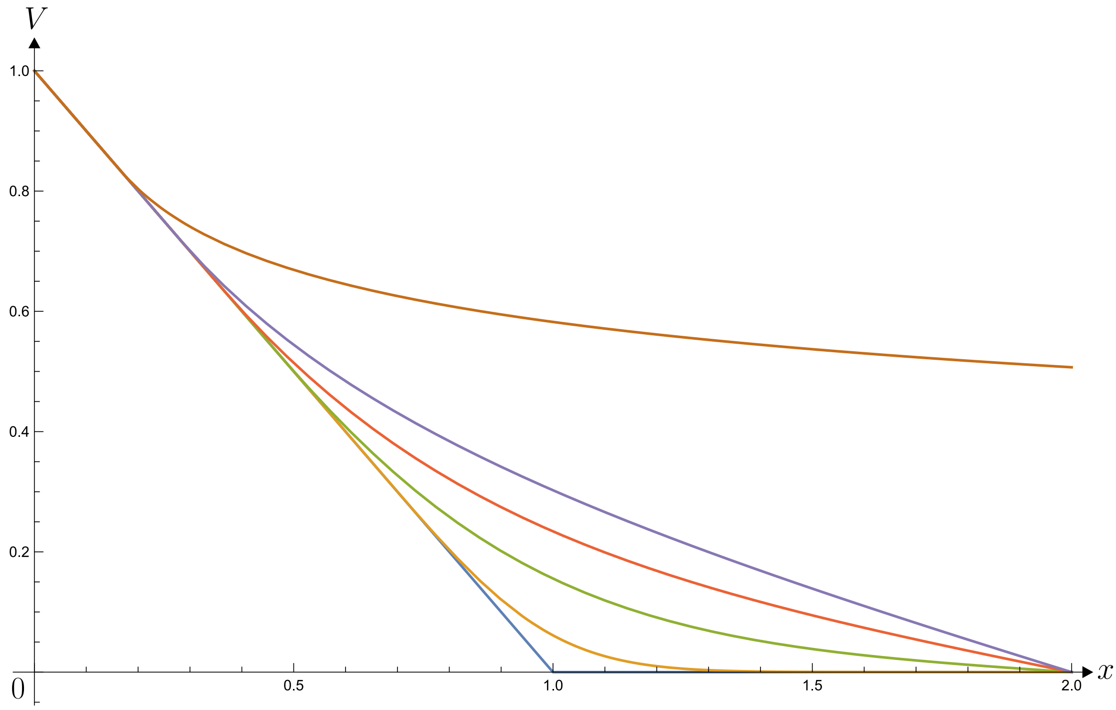

Figure 25. Numerical approximations of value functions in (8.10) through Wald-Bellman equations in (8.11)+(8.12) with $T = 1$, $r = 0.1$, $k = 1$, $\sigma = 1$, $n = 12$ and $T - t = 100 \times 2^{-n}$ (yellow), $T - t = 700 \times 2^{-n}$ (green), $T - t = 1800 \times 2^{-n}$ (red) and $T - t = 4096 \times 2^{-n}$ (purple). Gain function G in (8.10) (blue). Value function in (8.1) with infinite horizon (brown) (cf. [21, Chapter VII]).

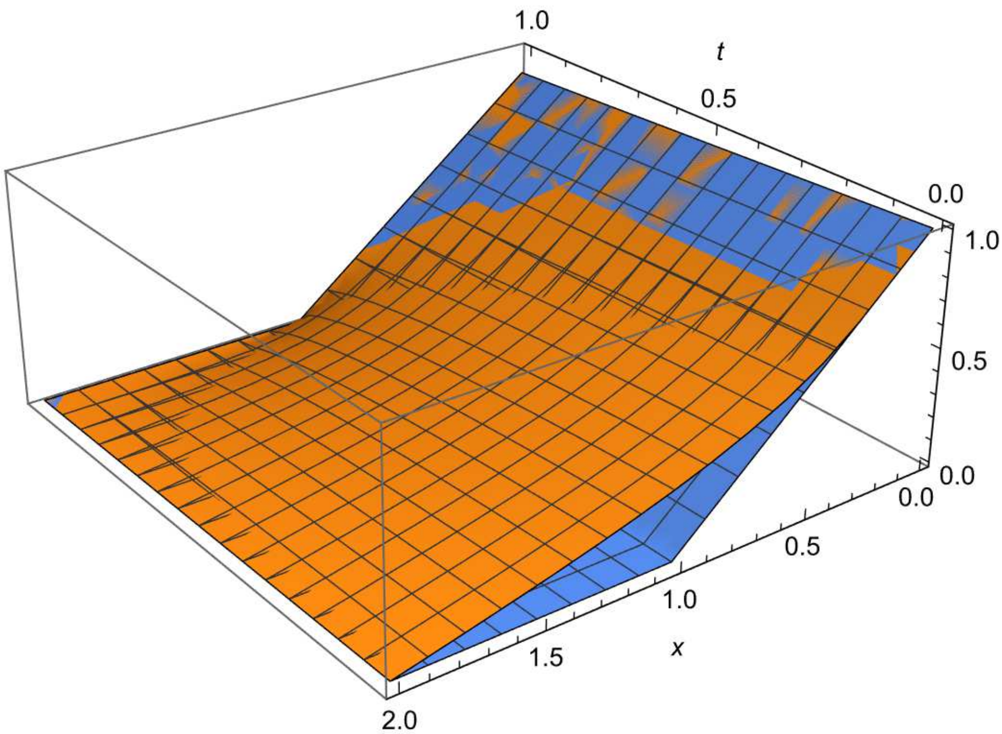

Figure 26. Numerical approximations of value function (yellow) and gain function (blue) in (8.25) according to Wald-Bellman equations in (8.26)+(8.27) with $T = 1$, $r = 0.1$, $k = 1$, $n = 12$ and $\sigma = 1$.

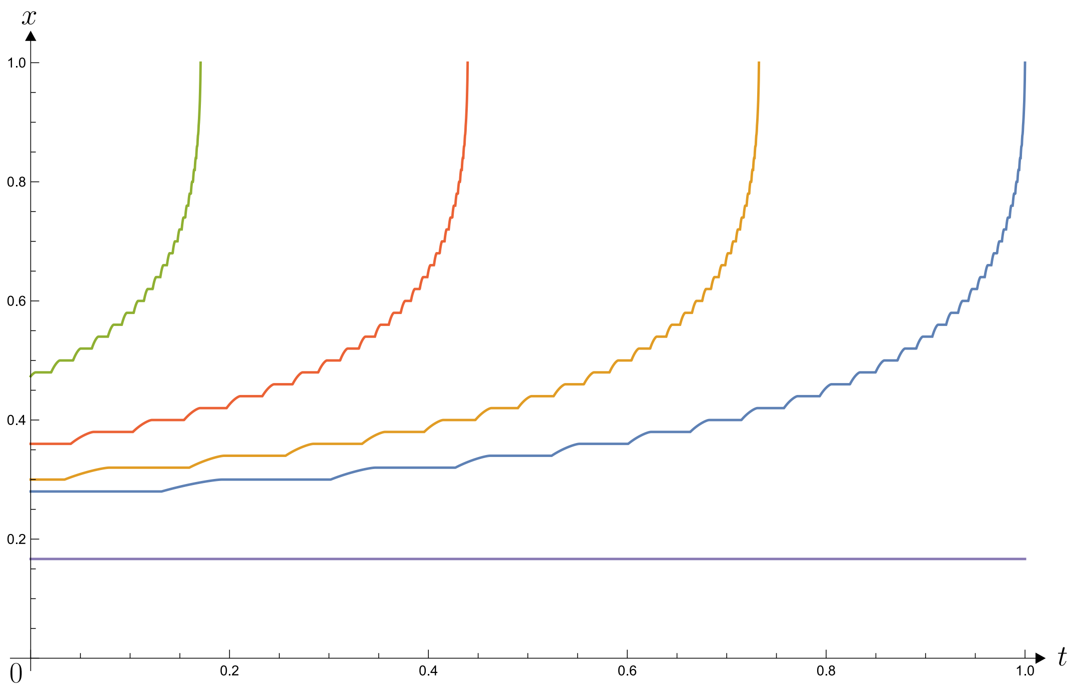

Figure 27. Numerical approximations of optimal stopping boundaries in (8.17) and (8.32) through Wald-Bellman equations in (8.11)+(8.12) and (8.26)+(8.27) with $r = 0.1$, $k = 1$, $\sigma = 1$, $n = 12$ and $T = 700 \times 2^{-n}$ (green), $T = 1800 \times 2^{-n}$ (red), $T = 3000 \times 2^{-n}$ (yellow) and $T = 4096 \times 2^{-n}$ (blue). Optimal stopping boundary with infinite horizon (purple) (cf. [21, Chapter VII]).

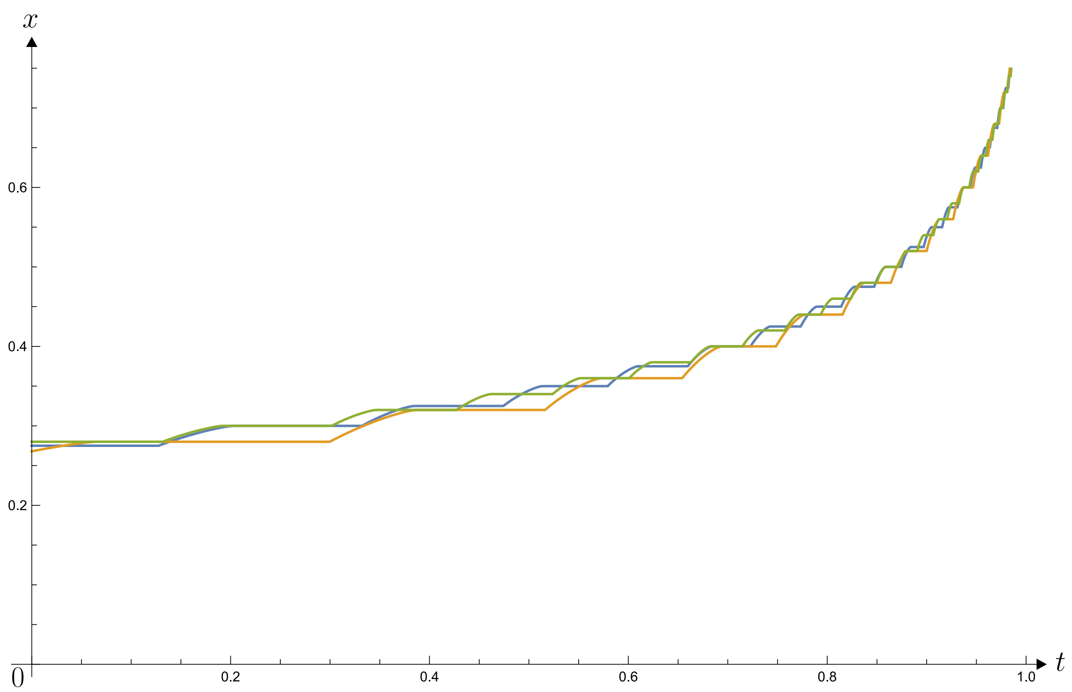

Figure 28. Numerical approximations of optimal stopping boundaries through Wald-Bellman equations according to Code 12 with the number of grids used for interpolation 50 (yellow), 80 (blue) and 100 (green).

Chapter 9

Quickest detection problems for Bessel processes

In this chapter, we consider the *Mayer* formulated optimal stopping problem which appeared in the quickest detection problems for Bessel processes in [14, Section 6]. The problem considered the movement of a particle which initially follows δ_0-dimensional Brownian motion and then after some random/unobservable time θ becomes δ_1-dimensional Brownian motion with $\delta_0 \geq 2$ and $\delta_1 > \delta_0$. It is assumed that only the distance of the particle to the origin is being observed, and the problem is to detect the time θ at which the particle changes its dimension as accurately as possible. The random/unobservable time θ is assumed to be (i) exponentially distributed and (ii) independent from the initial motion of the particle. By proper transformations introduced in [14], the quickest detection problem is formulated as a two-dimensional optimal stopping problem.

1. We consider the two-dimensional optimal stopping problem given by

$$V(\varphi, x) = \inf_{\tau \geq 0} \mathsf{E}_{\varphi,x}\left[e^{-\lambda\tau} M(\Phi_\tau, X_\tau)\right] \tag{9.1}$$

for $(\varphi, x) \in [0, \infty) \times (0, \infty)$ and $\lambda > 0$. τ is a stopping time of the two-dimensional strong Markov process (Φ, X) satisfying the system of stochastic differential equations

$$d\Phi_t = \lambda\,(1+\Phi_t)\,dt + \frac{1}{2}(\delta_1 - \delta_0)\frac{\Phi_t}{X_t}\,dB_t \tag{9.2}$$

$$dX_t = \frac{\delta_0 - 1}{2X_t}\,dt + dB_t \tag{9.3}$$

where $\delta_0 \geq 2$ is the dimension of the Brownian motion before the change and $\delta_1 > \delta_0$ is the dimension after the change with $t > 0$ and $\mathsf{P}_{\varphi,x}\left((\Phi_0, X_0) = (\varphi, x)\right) = 1$. $B =$

$(B_t)_{t\geq 0}$ is a standard Brownian motion starting at 0 under $\mathsf{P}_{\varphi,x}$. X is known to be a Bessel process. $M(\varphi,x)$ is a solution of the partial differential equation

$$(I\!\!L_{\Phi,X}M - \lambda M)(\varphi,x) = \varphi - \frac{\lambda}{c} \tag{9.4}$$

with $c > 0$ given and fixed where $I\!\!L_{\Phi,X}$ is the infinitesimal generator of the strong Markov process (Φ, X) given by

$$I\!\!L_{\Phi,X} = \lambda\,(1+\varphi)\,\partial_\varphi + \frac{\delta_0 - 1}{2x}\,\partial_x + \frac{1}{2}(\delta_1-\delta_0)\frac{\varphi}{x}\,\partial_{\varphi x} + \frac{1}{8}(\delta_1-\delta_0)^2\frac{\varphi^2}{x^2}\,\partial_{\varphi\varphi} + \frac{1}{2}\,\partial_{xx}. \tag{9.5}$$

An explicit expression of M is derived in [14, Section 5] where

$$M(\varphi,x) = \frac{1}{\delta_1}\left[(1+\varphi)x^2 + \frac{\delta_0}{\lambda} + \frac{\delta_1}{c}\right] \tag{9.6}$$

with $(\varphi,x) \in [0,\infty)\times(0,\infty)$.

2. By killing the sample path of (Φ, X) at the rate $\lambda > 0$, we obtain the identities

$$\mathsf{E}_{\varphi,x}\left[e^{-\lambda t}M(\Phi_t,X_t)\right] = \mathsf{E}_{\varphi,x}\left[M(\tilde{\Phi}_t,\tilde{X}_t)\right] \tag{9.7}$$

$$e^{-\lambda t}p(t;\varphi,x;\eta,z) = \tilde{p}(t;\varphi,x;\eta,z) \tag{9.8}$$

for $t > 0$ with (φ,x) and (η,z) in $[0,\infty)\times(0,\infty)$ where $(\tilde{\Phi},\tilde{X})$ is the killed process starting at $(\varphi,x) \in [0,\infty)\times(0,\infty)$. p is the transition density function of (Φ, X) and $\tilde{p}$ is the transition density function of $(\tilde{\Phi},\tilde{X})$ under $\mathsf{P}_{\varphi,x}$. The value function in (9.1) can be simplified to

$$V(\varphi,x) = \inf_{\tau\geq 0}\ \mathsf{E}_{\varphi,x}\left[M(\tilde{\Phi}_\tau,\tilde{X}_\tau)\right] \tag{9.9}$$

for $(\varphi,x) \in [0,\infty)\times(0,\infty)$ and stopping times τ of $(\tilde{\Phi},\tilde{X})$.

3. By the Wald-Bellman equations in (4.11) and Proposition 6, the value function in (9.9) satisfy the identities

$$V(\varphi,x) = \lim_{n\to\infty}\lim_{T\to\infty}\bar{V}_n^{T2^n}(\varphi,x) = \lim_{T\to\infty}\lim_{n\to\infty}\bar{V}_n^{T2^n}(\varphi,x) \tag{9.10}$$

where we set

$$\bar{V}_n^N(\varphi,x) = \inf_{\tau\in\mathcal{T}_n^N}\mathsf{E}_{\varphi,x}M(\tilde{\Phi}_\tau,\tilde{X}_\tau) = \min\left(M(\varphi,x), \mathbb{P}_{2^{-n}}\bar{V}_n^{N-1}(\varphi,x)\right) \tag{9.11}$$

for $\bar{V}_n^0(\varphi,x) = M(\varphi,x)$, $(\varphi,x) \in [0,\infty)\times(0,\infty)$, $n \geq 0$, $T2^n \in I\!N$, $N \in I\!N$ and $\mathcal{T}_n^N = \{0, 2^{-n}, 2\times 2^{-n}, ..., N\times 2^{-n}\}$. $\mathbb{P}_{2^{-n}}$ is the transition operator of $(\tilde{\Phi},\tilde{X})$ defined by

$$\mathbb{P}_{2^{-n}}\bar{V}_n^{N-1}(\varphi,x) = \mathsf{E}_{\varphi,x}\bar{V}_n^{N-1}(\tilde{\Phi}_{2^{-n}},\tilde{X}_{2^{-n}}) = \mathsf{E}_{\varphi,x}e^{\lambda 2^{-n}}\bar{V}_n^{N-1}(\Phi_{2^{-n}},X_{2^{-n}}) \tag{9.12}$$

for $(\varphi, x) \in [0, \infty) \times (0, \infty)$.

4. The continuation set C and stopping set D (cf. [14, Section 8]) are defined by

$$C = \{ (\varphi, x) \in [0, \infty) \times (0, \infty) \mid V(\varphi, x) < M(\varphi, x) \} \tag{9.13}$$

$$D = \{ (\varphi, x) \in [0, \infty) \times (0, \infty) \mid V(\varphi, x) = M(\varphi, x) \}. \tag{9.14}$$

The first entry time of (Φ, X) into D defined by

$$\tau_D = \inf\{ t \geq 0 \mid (\Phi_t, X_t) \in D \} \tag{9.15}$$

is optimal in (9.1) (cf. [14, Section 8]). The optimal stopping boundary between C and D is defined by

$$b(x) = \inf\{ \varphi \geq 0 \mid (\varphi, x) \in D \} \tag{9.16}$$

for $x \in \mathbb{R}^+$.

5. We recall that the transition density function p of (Φ, X) is the unique non-negative solution to the Kolmogorov backward equation

$$p_t(t; \varphi, x; \eta, z) = \mathbb{L}_{\Phi, X}(p)(t; \varphi, x; \eta, z) \tag{9.17}$$

$$p(0+; \varphi, x; \eta, z) = \delta_{\varphi, x}(\eta, z) \quad (weakly) \tag{9.18}$$

under $\mathsf{P}_{\varphi,x}$ satisfying $\int_0^\infty \int_0^\infty p(t; \varphi, x; \eta, z)\, d\eta\, dz = 1$ for $t > 0$ with (φ, x) and (η, z) in $[0, \infty) \times (0, \infty)$ (cf. [8]) where $\mathbb{L}_{\Phi, X}$ is the infinitesimal generator defined in (9.5) and $\delta_{\varphi,x}$ is the Dirac measure at (φ, x). As an alternative method to compute the expectations in the Wald-Bellman equations (9.11), we use *Euler approximation* (cf. [13, Chapter 2.1]) where

$$\begin{aligned} \Phi_{2^{-n}} &= \varphi + \int_0^{2^{-n}} \lambda(1 + \Phi_s)\, ds + \int_0^{2^{-n}} \frac{1}{2}(\delta_1 - \delta_0) \frac{\Phi_s}{X_s}\, dB_s \\ &\approx \varphi + \int_0^{2^{-n}} \lambda(1 + \varphi)\, ds + \int_0^{2^{-n}} \frac{1}{2}(\delta_1 - \delta_0) \frac{\varphi}{x}\, dB_s \\ &= \varphi + \lambda(1 + \varphi) 2^{-n} + \frac{1}{2}(\delta_1 - \delta_0) \frac{\varphi}{x} B_{2^{-n}} \end{aligned} \tag{9.19}$$

$$X_{2^{-n}} \approx x + \frac{\delta_0 - 1}{2x} 2^{-n} + B_{2^{-n}}. \tag{9.20}$$

for large $n \in \mathbb{N}$. The expectations associated with (Φ, X) can be approximated by using the law of standard Brownian motion.

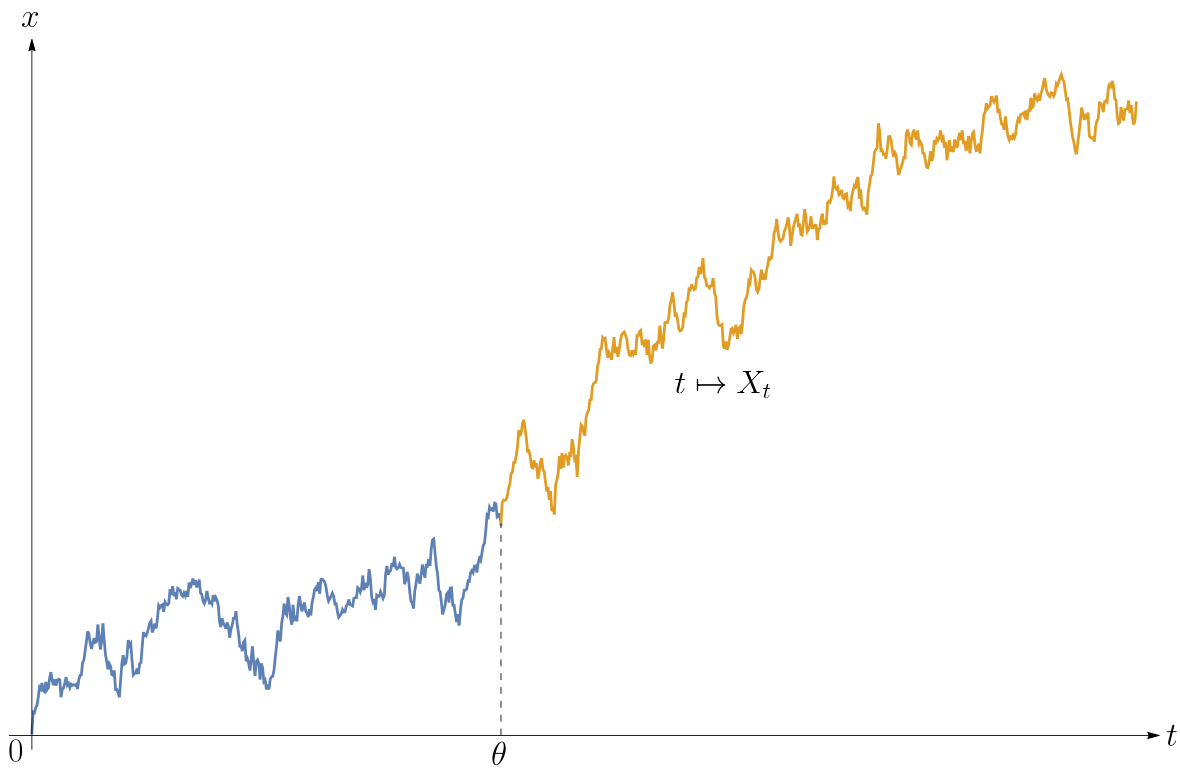

Figure 29. Simulated sample path of the Bessel process X with $\delta_0 = 2$ and $\delta_1 = 3$.

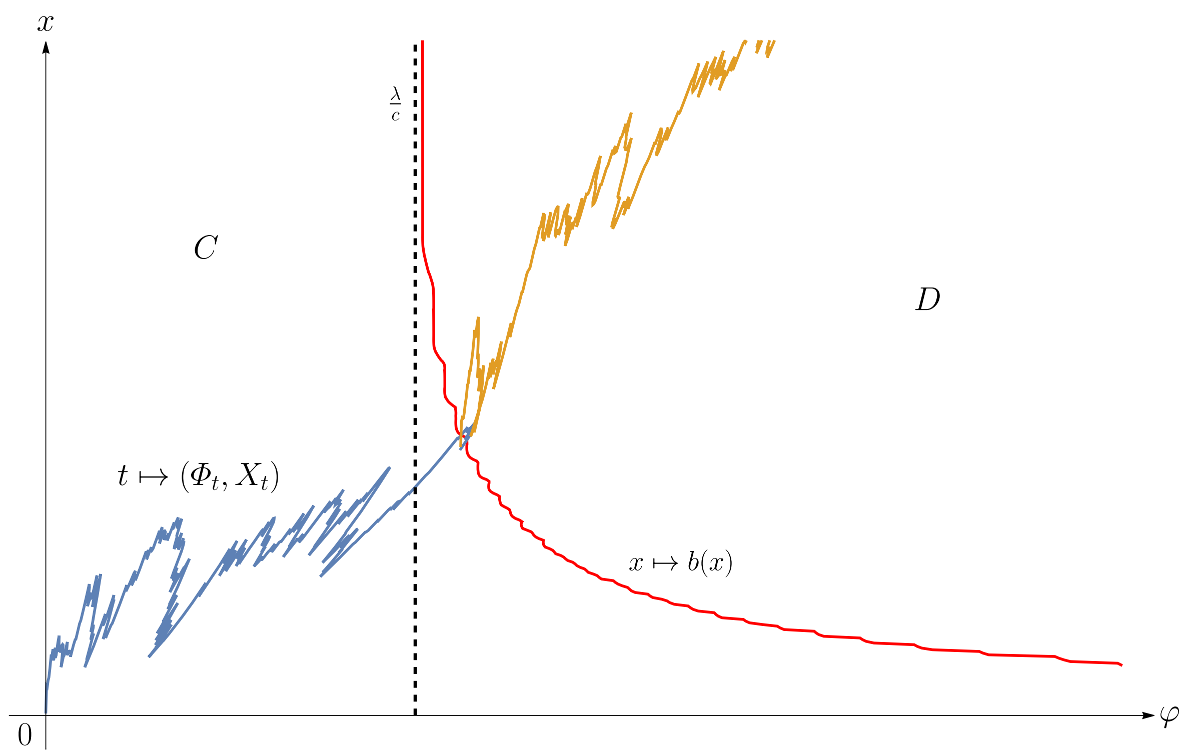

Figure 30. Kinematics of the process (Φ, X) associated with the sample path from Figure 11 and locations of the optimal stopping boundary when $\lambda = 1$ and $c = 2$.

6. We now introduce the algorithm of Wald-Bellman equations and the corresponding numerical analysis. The idea is similar to Chapter 5. Consider the Wald-Bellman equations in (9.10)+(9.11), the iterations are given by

$$
\begin{aligned}
\bar{V}_n^0(\varphi,x) &= M(\varphi,x) \qquad (9.21)\\
\bar{V}_n^1(\varphi,x) &= \inf_{\tau\in\mathcal{T}_n^1} \mathsf{E}_{\varphi,x}^{\infty} M(\tilde{\Phi}_\tau,\tilde{X}_\tau) = \min\big(M(\varphi,x), \mathbb{P}_{2^{-n}}\bar{V}_n^0(\varphi,x)\big)\\
&= \min\Big(M(\varphi,x), \mathsf{E}_{\varphi,x}^{\infty} e^{\lambda 2^{-n}}\bar{V}_n^0(\Phi_{2^{-n}},X_{2^{-n}})\Big)\\
&= \min\Big(M(\varphi,x), \int_0^\infty\int_0^\infty e^{\lambda 2^{-n}}\,\bar{V}_n^0(\eta,z)\,p(2^{-n};\varphi,x;\eta,z)\,d\eta\,dz\Big)\\
&\approx \min\Big(M(\varphi,x),\\
&\int_{-\infty}^{\infty} e^{\lambda 2^{-n}}\,\bar{V}_n^0(\varphi+\lambda(1+\varphi)2^{-n}+\frac{1}{2}(\delta_1-\delta_0)\frac{\varphi}{x}y, x+\frac{\delta_0-1}{2x}2^{-n}+y)\,\hat{p}(2^{-n};y)\,dy\Big)
\end{aligned} \qquad (9.22)
$$

$$
\begin{aligned}
\bar{V}_n^2(\varphi,x) &= \inf_{\tau\in\mathcal{T}_n^2} \mathsf{E}_{\varphi,x}^{\infty} \hat{M}(\tilde{\Phi}_\tau,\tilde{X}_\tau) = \min\Big(\hat{M}(\varphi,x), \mathbb{P}_{2^{-n}}\bar{V}_n^1(\varphi,x)\Big)\\
&\approx \min\Big(\hat{M}(\varphi,x),\\
&\int_{-\infty}^{\infty} e^{\lambda 2^{-n}}\,\bar{V}_n^1(\varphi+\lambda(1+\varphi)2^{-n}+\frac{1}{2}(\delta_1-\delta_0)\frac{\varphi}{x}y, x+\frac{\delta_0-1}{2x}2^{-n}+y)\,\hat{p}(2^{-n};y)\,dy\Big)
\end{aligned} \qquad (9.23)
$$

$$\vdots$$

$$
\begin{aligned}
\bar{V}_n^{T2^n}(\varphi,x) &= \inf_{\tau\in\mathcal{T}_n^{T2^n}} \mathsf{E}_{\varphi,x}^{\infty} M(\tilde{\Phi}_\tau,\tilde{X}_\tau) = \min\big(M(\varphi,x), \mathbb{P}_{2^{-n}}\bar{V}_n^{T2^n-1}(\varphi,x)\big)\\
&\approx \min\Big(M(\varphi,x),\\
&\int_{-\infty}^{\infty} e^{\lambda 2^{-n}}\bar{V}_n^{T2^n-1}(\varphi+\lambda(1+\varphi)2^{-n}+\frac{1}{2}(\delta_1-\delta_0)\frac{\varphi}{x}y, x+\frac{\delta_0-1}{2x}2^{-n}+y)\hat{p}(2^{-n};y)\,dy\Big)
\end{aligned} \qquad (9.24)
$$

for $(\varphi,x)\in[0,\infty)\times(0,\infty)$, $n\geq 0$, $T2^n\in I\!N$, $N\in I\!N$ and $\mathcal{T}_n^N = \{\,0,2^{-n},2\times 2^{-n},...,N\times 2^{-n}\,\}$ where $\hat{p}$ is the transition density function of standard Brownian motion defined in (5.21). We now illustrate an example code in Mathematica. The coefficients are defined by

Code 22.

```
c=2.; time=1.; phistep=phiubound/100.;
xstep=(xubound-xlbound)/100.; h=0.5^10.; w=time/h;
```

```
lambda=1.  ;delta0=2.; delta1=3.; xubound=2.; xlbound=0.1;
phiubound=1.5.; philbound=0.;
lossfunction[phi_,x_]:=((1.+phi)*x^2.+delta0/lambda+delta1/c)/delta1;
functionlist={lossfunction};
```

where `c` is the constant $c \in I\!R^+$ in (9.4). `time` is T in (9.10). `phiubound` and `philbound` are the upper and lower bounds of the interpolation domain for φ. Similarly, `xubound` and `xlbound` are the upper and lower bounds of the interpolation domain for x. `h` is 2^{-n} in (9.10). `w` is the number of total iterations defined as $T2^n \in I\!N$ in (9.10). `lambda` is $\lambda > 0$ in (9.1). `delta0` is the initial dimension $\delta_0 \geq 2$ in (9.2). `delta1` is dimension $\delta_1 > \delta_0$ after the change. `phistep` denotes the distance between the grids for interpolating φ. `xstep` denotes the distance between the grids for interpolating x. `lossfunction` is M in (9.6). `functionlist` is a `list` of functions which stores the interpolated value functions. After the value iterations, the `functionlist` should have the form `functionlist` $= \{\bar{V}_n^0, \bar{V}_n^1, \bar{V}_n^2, ..., \bar{V}_n^{T2^n}\}$. We now introduce the algorithm of value iterations

Code 23.

```
Do[Subscript[valuefunctionexact,n][phi_,x_]:=
 Min[lossfunction[phi, x],
  NIntegrate[
   functionlist[[n]][lambda*(1.+phi)*h+((delta1-delta0)/2.)*
    (phi/x)*b+phi,((delta0-1.)/(2.*x))*h+b+x]*
     Exp[-lambda*h]/Sqrt[2.*Pi*h]*Exp[-0.5/h*b^2.],
      {b,-1.,1.},Method->{Automatic,"SymbolicProcessing"->0},
       AccuracyGoal->10]];
 Subscript[table,n]=N[Flatten[
  Table[{phi,x,Subscript[valuefunctionexact,n][phi, x]},
   {phi,0,phiubound,phistep},{x,xlbound,xubound,xstep}],1]];
 Subscript[valuefunction,n]=Interpolation[Subscript[table,n]];
 functionlist=Insert[functionlist,Subscript[valuefunction,n],-1],{n,w}].
```

The integrals for computing expectations are taken from -1 to 1 instead from $-\infty$

to ∞ because the algorithm is computation intensive. Shrinking the integral interval helps reduce the computation time.

7. The constructions of value functions $\bar{V}_n^N$ defined in (9.9) through Wald-Bellman equations can be imaged as a surface placed under the loss function M in (2.94) being pulled down towards the exact the value function $\tilde{V}$ defined in (9.1) as the number of iterations tends infinity. The more iterations conducted in Wald-Bellman equations, the closer the approximated value functions to $\tilde{V}$ as shown in Figure 31. The value functions generated by Wald-Bellman equations comply with smooth fit (cf. [21, Chapter IV, 9.1] and [14]) at the touching points between the value functions and loss function as shown in Figure 32 and 33. The value function in Figure 32 is decreasing when x is small and φ is large. This is caused by the numerical approximations. We recall that x is strictly positive. The behaviour of the value function when $x \downarrow 0$ is challenging to obtain because the Euler approximations in (9.20) would tend to infinity. The singularity cannot be handled by a computer. However, although there are some biases when x is small, the optimal stopping may not be affected because the constructed value functions are accurate around the touching points with the loss function M.

8. We construct the optimal stopping boundary by the definition in (9.16). The optimal stopping boundary $b(x)$ is constructed in Mathematica as follows

Code 24.

```
boundary=Table[{phi/.FindRoot[functionlist[[w+1]][phi, x]]-
 lossfunction[phi, x],{phi,lambda/c},AccuracyGoal->3,
  PrecisionGoal -> 3], x},{x, 0.15, xubound, 0.005}];.
```

The smoothness of optimal stopping boundary is controlled by the number of grids used for interpolation as shown in Figure 34. The more grids used, the smoother and more accurate the boundary would be. The construction of optimal stopping boundaries along with Wald-Bellman equations can be imaged as a rope being pulled towards the true boundary as shown in Figure 35. The numerical approximation converges from finite-horizon to the infinite-horizon optimal stopping boundary as the iteration continues which complies with Proposition 6.

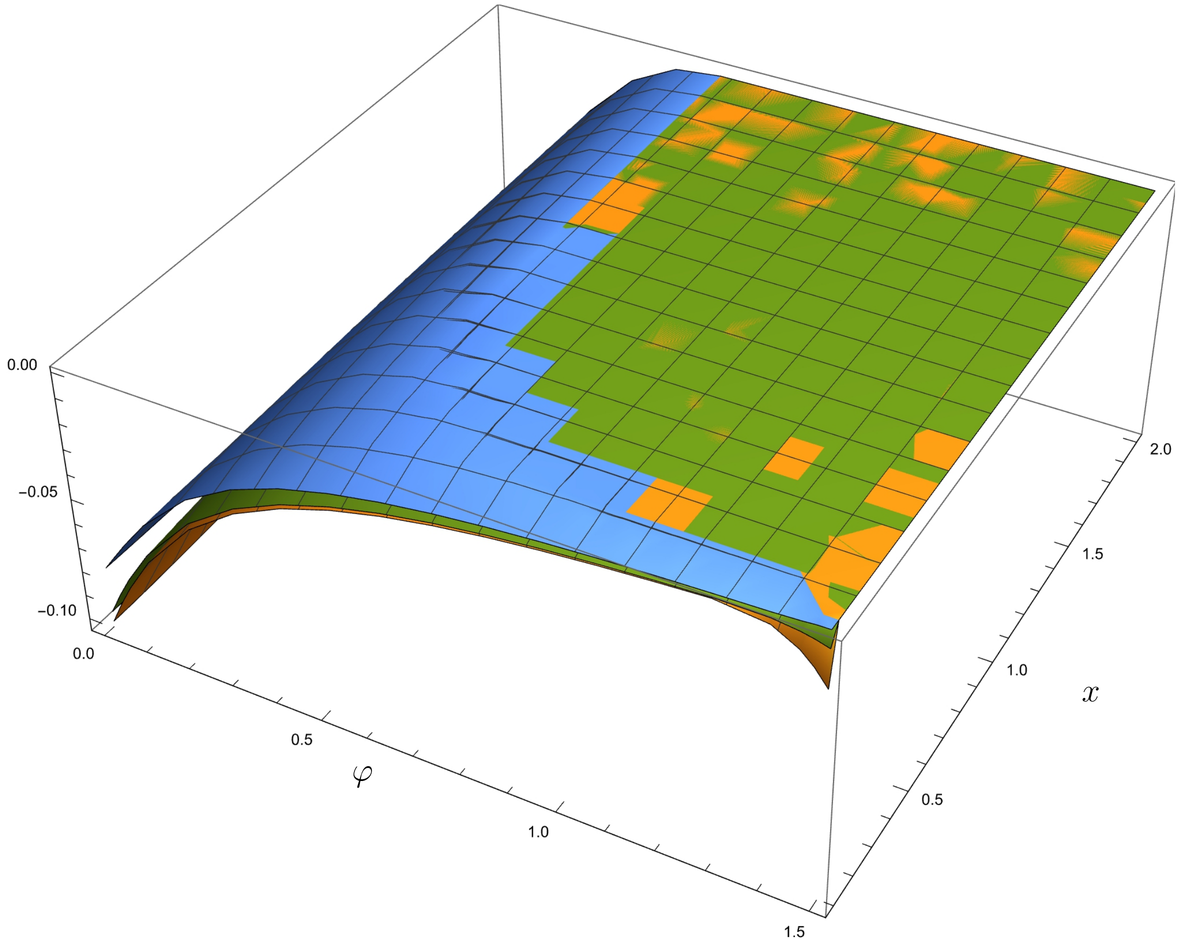

Figure 31. Numerical approximations of value functions with 251 iterations (blue), 501 iterations (green) and 1024 iterations (yellow).

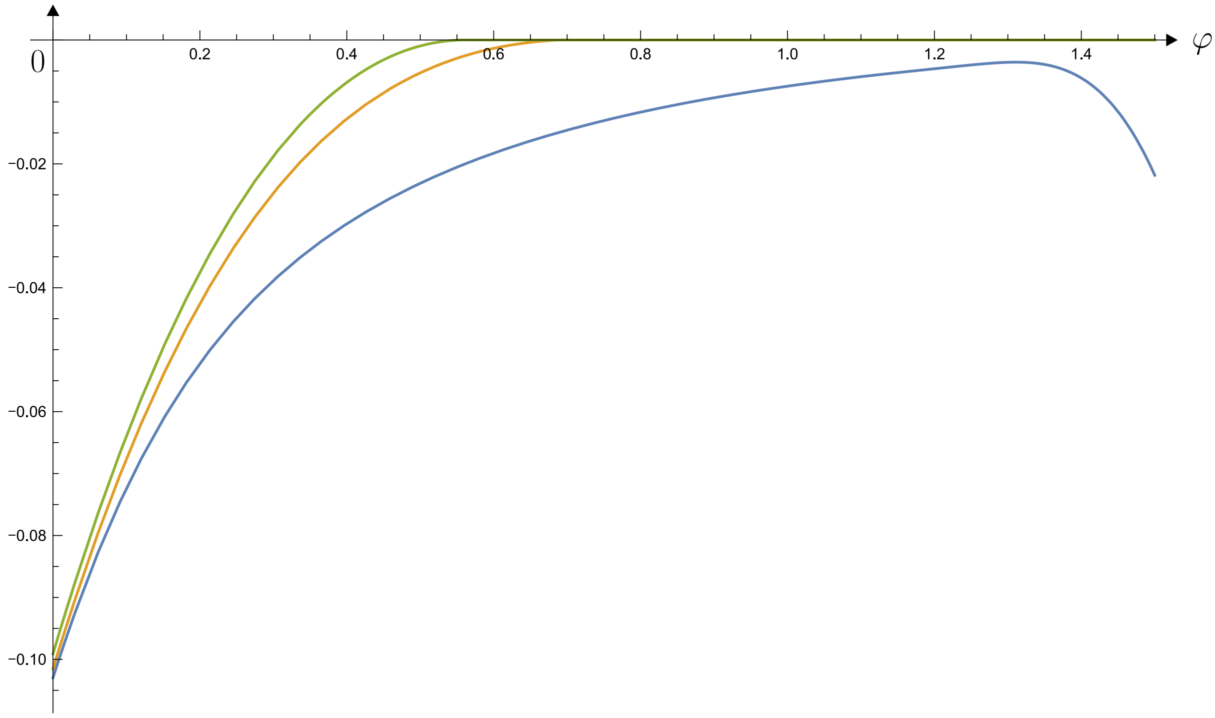

Figure 32. Numerical approximations of value functions according to Figure 31 with 1024 iterations when $\varphi \in [0, 1.5]$ and $x = 0.1$ (blue), $x = 0.38$ (yellow) and $x = 0.76$ (green).

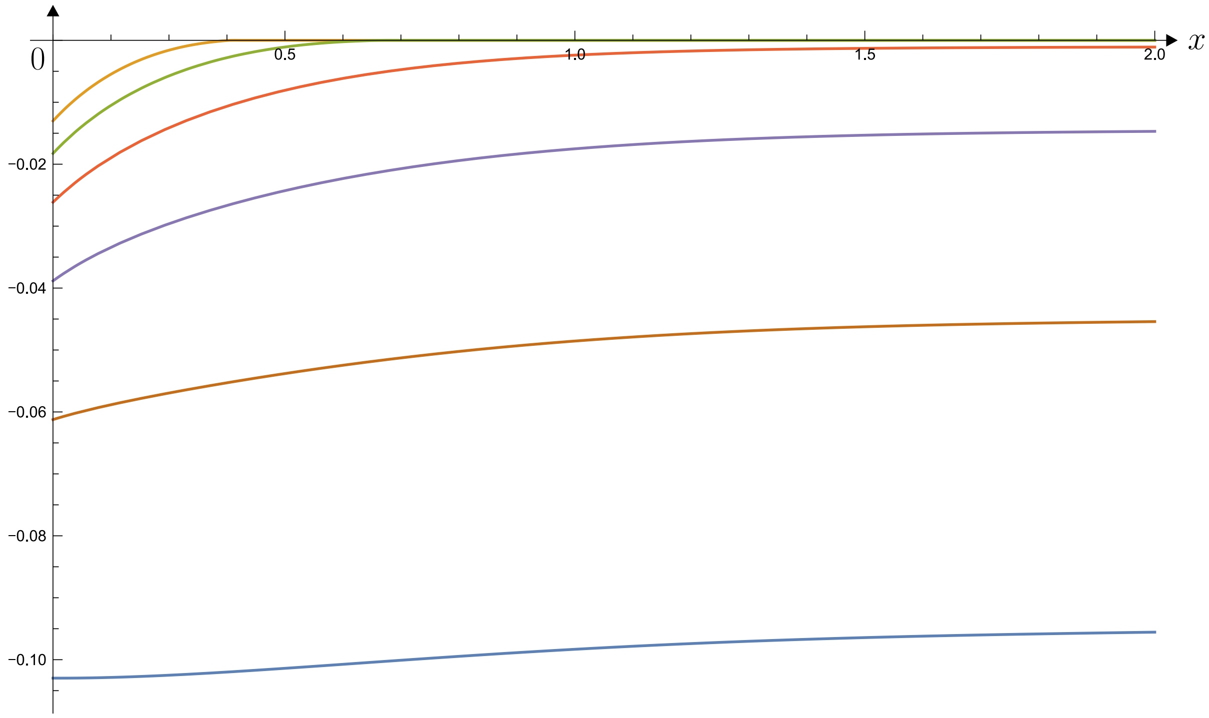

Figure 33. Numerical approximations of value functions with 1024 iterations when $x \in [0.1, 2]$ and $\varphi = 0$ (blue), $\varphi = 0.15$ (brown), $\varphi = 0.3$ (purple), $\varphi = 0.45$ (red), $\varphi = 0.6$ (green) and $\varphi = 0.75$ (yellow).

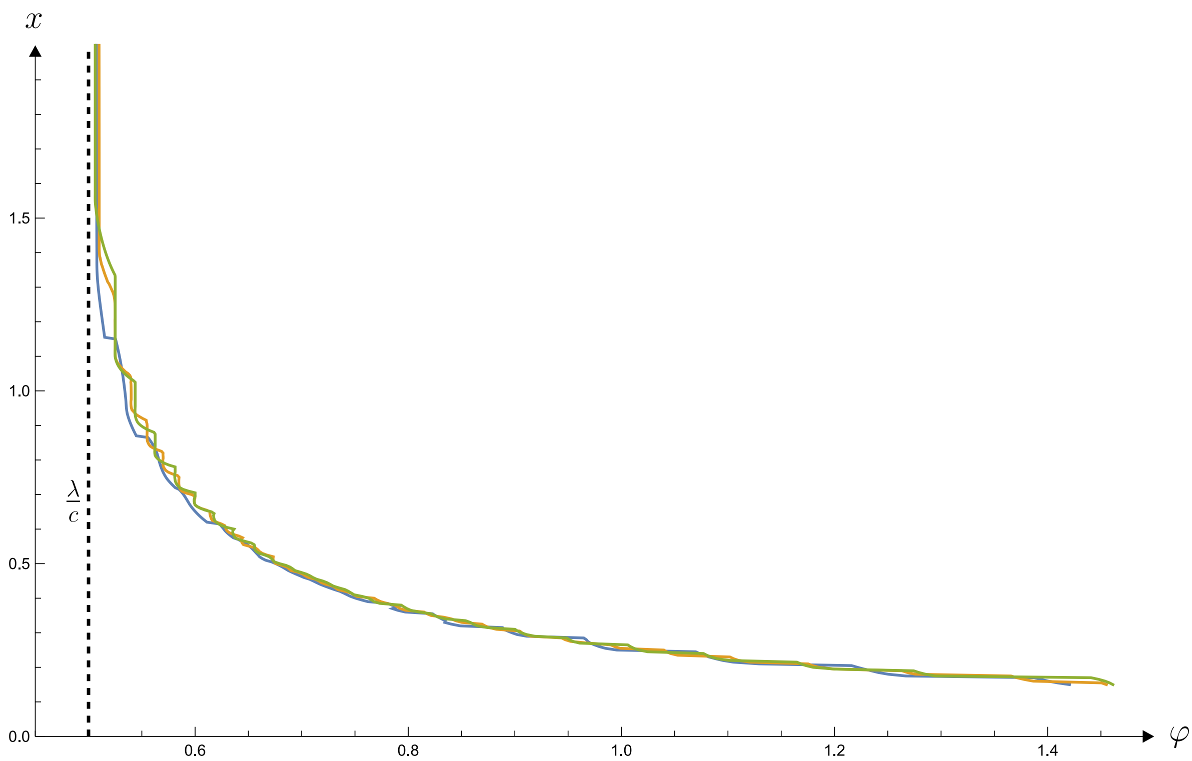

Figure 34. Numerical approximations of optimal stopping boundaries when the number of grids used for interpolation are 50 (green), 100 (blue) and 200 (yellow).

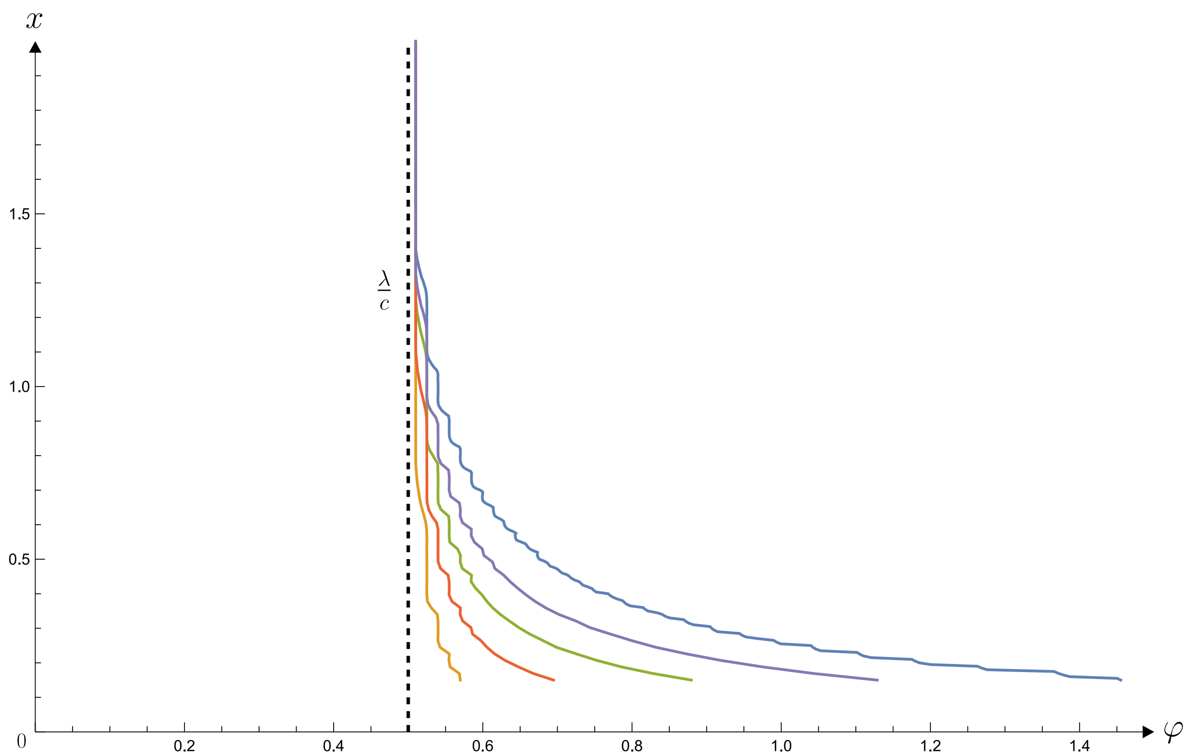

Figure 35. Numerical approximations of optimal stopping boundaries when the number of iterations are 10 (yellow), 20 (red), 30 (green), 50 (purple) and 1024 (blue).

Bibliography

[1] ARROW, K. J., BLACKWELL, D. *and* GIRSHICK, M. A. (1949). Bayes and minimax solutions of sequential decision problems. *Econometrica* 17 (213-244).

[2] BELLMAN, R. (1957). *Dynamic Programming*. Princeton Univ. Press, Princeton.

[3] BIAN, Z. (2023). Rates of convergence in the Wald-Bellman equations with applications. *Research Report* No. 2, *Probab. Statist. Group Manchester* (26 pp). (Submitted to *Stochastic Process. Appl.*)

[4] BRAMANTI, M. (2014). *An Invitation to Hypoelliptic Operators and Hörmander's Vector Fields*. SpringerBriefs in Mathematics. Springer, Cham.

[5] COHN, L. D. (2nd. ed. 2013). *Measure Theory.* Birkhäuser.

[6] EKSTRÖM, E. LINDBERG, C. *and* TYSK, J. (2011). Optimal liquidation of a pairs trade. *Advanced Mathematical Methods for Finance*. Springer (247-255).

[7] ELLIOTT, R. VAN DER HOEK, J. *and* MALCOLM, W. (2005). Pairs trading. *Quant. Finance* 5 (271-276).

[8] FELLER, W. (1952). The parabolic differential equations and the associated semigroups of transformations. *Ann. of Math.* 55 (468-519).

[9] GLOVER, K. *and* PESKIR, G. (2017). Quickest detection problems for Ornstein-Uhlenbeck processes. *Research Report* No. 11, *Probab. Statist. Group Manchester* (28 pp).

[10] GRIGELIONIS, B. I. *and* SHIRYAEV, A. N. (1965). The "truncation" criteria for the optimal stopping time in the sequential analysis. *Theory Probab. Appl.* 10 (601-613).

[11] Grigelionis, B. I. *and* Shiryaev, A. N. (1966). On Stefan's problem and optimal stopping rules for Markov processes. *Theory Probab. Appl.* 11 (541-558).

[12] Hörmander, L. (1967). Hypoelliptic second order differential equations. *Acta Math.* 119 (147-171).

[13] Iacus, S. M. (2008). *Simulation and Inference for Stochastic Differential Equations.* Springer.

[14] Johnson, P. *and* Peskir, G. (2017). Quickest detection problems for Bessel processes. *Ann. Appl. Probab.* 27 (1003-1056).

[15] Leung, T. *and* Li, X. (2015). Optimal mean reversion trading with transaction costs & stop-loss exit. *Int. J. Theor. Appl. Finance* 18 (31 pp).

[16] Leung, T. *and* Li, X. (2016). *Optimal Mean Reversion Trading: Mathematical Analysis and Practical Applications.* World Scientific.

[17] Lindberg, C. (2014). Pairs trading with opportunity cost. *J. Appl. Probab.* 51 (282-286).

[18] Lipster, R. S. *and* Shiryaev, A. N. (1977, 2nd. ed. 2001). *Statistics of Random Processes I: General Theory.* Springer.

[19] Pedersen, J. L. *and* Peskir, G. (2002). On nonlinear integral equations arising in problems of optimal stopping. *Proc. Funct. Anal.* VII (Dubrovnik 2001), *Various Publ. Ser.* 46 (159-175).

[20] Peskir, G. (2022). Weak solutions in the sense of Schwartz to Dynkin's characteristic operator equation. *Research Report* No. 1, *Probab. Statist. Group Manchester* (20 pp).

[21] Peskir, G. *and* Shiryaev, A. N. (2006). *Optimal Stopping and Free-Boundary Problems.* Lectures in Mathematics, ETH Zürich, Birkhäuser.

[22] Phillips, P. C. B. Wu, Y. *and* Yu, J. (2011). Explosive behavior in the 1990s NASDAQ: When did exuberance escalate asset values? *Int. Econ. Rev.* 52 (201-226).

[23] Phillips, P. C. B. *and* Yu, J. (2011). Dating the timeline of financial bubbles during the subprime crisis. *Quant. Econ.* 2 (455-491).

[24] Ray, S. N. (1965). Bounds on the maximum sample size of a Bayes sequential procedure. *Ann. Math. Statist.* 36(3) (859-878).

[25] Revuz, D. *and* Yor, M. (1999). *Continuous Martingales and Brownian Motion.* Springer-Verlag.

[26] Rogers, L. C. G. *and* Williams, D. (2000). *Diffusions, Markov Processes and Martingles: Itô Calculus* (Vol 2). Cambridge University Press.

[27] Schwartz, L. (1951). *Théory des Distributions.* Hermann.

[28] Shiryaev, A. N. (1961). The problem of the most rapid detection of a disturbance of a stationary regime. *Soviet Math. Dokl.* 2 (795-799).

[29] Shiryaev, A. N. (1978). *Optimal Stopping Rules.* Springer.

[30] Shiryaev, A. N. (2010). Quickest detection problems: Fifty years later. *Sequential Anal.* 29 (345-385).

[31] Stroock, D. W. (2008). *Partial Differential Equations for Probabalists.* Cambridge University Press.

[32] Uhlenbeck, G. E. *and* Ornstein, L. S. (1930). On the theory of Brownian motion. *Phys. Rev.* 36 (823-841).

[33] Vidyamurphy, G. (2004). *Pairs Trading: Quantitative Methods and Analysis.* John Wiley.

[34] Wald, A. (1976). *Sequential Analysis.* Wiley, New York; Chapman & Hall, London.

[35] Williams, D. (1991). *Probability with Martingales.* Cambridge University Press.