COMPOUND POISSON DISORDER WITH GENERAL PRIOR AND MISSPECIFIED WIENER DISORDER PROBLEM

A THESIS SUBMITTED TO

THE GRADUATE SCHOOL OF ENGINEERING AND SCIENCE

OF BILKENT UNIVERSITY

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR

THE DEGREE OF

MASTER OF SCIENCE

IN

INDUSTRIAL ENGINEERING

By

Deniz Şahin

July 2024

Compound Poisson Disorder with General Prior and Misspecified Wiener Disorder Problem
By Deniz Şahin
July 2024

We certify that we have read this thesis and that in our opinion it is fully adequate, in scope and in quality, as a thesis for the degree of Master of Science.

Savaş Dayanık(Advisor)

Semih Onur Sezer

Özlem Çavuş İyigün

Approved for the Graduate School of Engineering and Science:

Orhan Arıkan
Director of the Graduate School

ABSTRACT

COMPOUND POISSON DISORDER WITH GENERAL PRIOR AND MISSPECIFIED WIENER DISORDER PROBLEM

Deniz Şahin
M.S. in Industrial Engineering
Advisor: Savaş Dayanık
July 2024

For a system modeled with a compound Poisson or a Wiener process, let us assume that the underlying model parameters change at an unknown and unobservable time. For a compound Poisson process, these are arrival rate and mark distribution while for a Wiener process, it is the drift parameter. Suppose the decision maker knows the pre- and post-disorder process parameters, as well as the prior density of the disorder time. In this case, finding a stopping rule that optimizes a Bayesian penalty function is called the compound Poisson and Wiener disorder problem, respectively. For the compound Poisson problem, we consider a general prior distribution where the decision maker has more general knowledge about the disorder time than exponential and uniform priors which were addressed in the previous studies. For the Wiener problem, we revisit the asset selling problem with an exponential prior, where the decision maker specifies problem parameters incorrectly. In both cases, the original problems reduce to optimal stopping problems. We use time discretization and successive approximation methods for the first case and Markov chain approximation and Monte Carlo simulations for the second case. We provide the quickest detection rules and discuss various numerical examples.

Keywords: Compound Poisson disorder problem; Wiener disorder problem; quickest detection; optimal stopping; parameter misspecification.

ÖZET

GENEL DAĞILAN DÜZENSİZLİK ZAMANLI BİLEŞİK POISSON VE YANLIŞ BELİRTİLMİŞ WIENER DÜZENSİZLİK PROBLEMİ

Deniz Şahin
Endüstri Mühendisliği, Yüksek Lisans
Tez Danışmanı: Savaş Dayanık
Temmuz 2024

Bileşik Poisson veya Wiener süreçleriyle modellenen bir sistem için, model parametrelerinin bilinmeyen ve gözlenemeyen bir zamanda değiştiğini varsayalım. Bileşik Poisson süreci için bunlar varış hızı ve işaret dağılımıyken, Wiener süreci için sürüklenme parametresidir. Karar vericinin, düzensizlik öncesi ve sonrası süreç parameterlerini, aynı zamanda düzensizlik ön dağılımını bildiğini kabul edelim. Bu durumda, bir Bayes ceza fonksiyonunu en iyileyen bir durma kuralı bulmaya sırasıyla bileşik Poisson ve Wiener düzensizlik problemi adı verilir. Bileşik Poisson problemi için, daha önce çalışılmış üstel ve düzenli ön dağılımlar yerine, karar verenin düzensizlik zamanı hakkında daha çok genel bilgi sahibi olduğu genel bir ön dağılımı değerlendiriyoruz. Wiener problemi içinse karar vericinin problem parametrelerini yanlış belirttiği, üstel ön dağılımlı varlık-satış problemini yeniden inceliyoruz. Her iki durumda da orijinal problemler en iyi durma problemlerine indirgenmektedir. İlk durumda zaman ayrıklaştırma ve ardışık yakınsamalar, ikinci durumdaysa Markov zinciri yakınsaması ve Monte Carlo simülasyonları metodlarını kullanıyoruz. En hızlı tespit kurallarını veriyor ve çeşitli sayısal örnekleri tartışıyoruz.

Anahtar sözcükler: Bileşik Poisson düzensizlik problemi; Wiener düzensizlik problemi; en hızlı tespit; en iyi durma; parametre yanlış belirtmesi.

Acknowledgement

First and foremost, I would like to express my deepest gratitude to my advisor, Prof. Savaş Dayanık, for his endless support and understanding. Besides guiding me through quality research, I have learned a lot about teaching and being a good mentor as well.

I would like to thank Assoc. Prof. Semih Onur Sezer for his guidance starting from my undergraduate studies. It was my first course with him that motivated me to pursue an academic career. I am forever grateful for his support.

I want to thank my family, Hayrettin Şahin, Gülcan Şahin, and Damla Şahin Güner, for their lifelong and unconditional support. I feel like the luckiest person to have you.

Special thanks to Ecenur Oğuz, Göksu Ece Okur, and Deniz Akkaya for their unmatched company on this journey and for always being just a phone call away for an amazing evening. Cheers!

I want to thank my friends, Elif Rana Yöner, Elif Sena Işık, Mustafa Çolak, Doğuş Berk Koçak, Osman Kağan Yayla, Zeynep Göze Gürhan, Gülin Yurter, Ali Eren Demir, Yunus Emre Çakır, Ali İlhan Haliloğlu, İrem Bahtiyar and Zehranaz Dönmez for being an incredible group of people. I feel truly fortunate to see each of you every day.

A heartfelt thanks to Berçem Ceylan, Ece Ulutürk, Ekin Gün, Melis Koçer, Mervan Önen, Sıla Öztürk, Yılmazcem Özardıç, and Zeynep Hekimoğlu for always being there whenever I needed, even from another city or country. I feel so lucky to be your friend.

Finally, I want to thank Onur Yaman and Derya Şahin for making my experience in the capital city a hundred times better. I will miss these three years a lot.

Contents

1 Introduction **1**

I Compound Poisson Disorder Problem **9**

2 Model and Problem Formulation **11**

2.1 A Markov Pair 14

2.2 Truncation of C-Space 18

3 Successive Approximations **21**

3.1 Defining Successive Approximations 23

3.2 Implementation of the Method 27

4 Discretization **29**

4.1 Implementation of the Method 33

4.2 Alarm Rule 34

5 Numerical Examples **38**

5.1 Mixed Uniform Problem 38

5.2 Gamma Problem 39

5.3 Phase-Type Problem 42

5.4 An Asymptotic Problem 44

5.5 Mixed Uniform Distribution for Approximation 45

5.6 Fraud Problem 46

II Misspecified Wiener Disorder Problem **50**

6 Model and Problem Description **52**

6.1 Model 52

6.2 Problem with Correct Parameters 54

6.3 Problem with Incorrect Parameters 55

7 Problem with Correct Parameters **56**

7.1 The Dynamics of Φ under $\mathbb{P}_Q$ 56

7.2 Solution of the Original Problem 57

7.3 Solution Using Markov Chain Approximation 60

8 Misspecification **62**

8.1 The Dynamics of $\tilde{\Phi}$ under $\mathbb{P}_Q$ 62

8.2 On the Function $H(\phi, \tilde{\phi})$ 63

9 Solution of the Misspecified Problem 65

9.1 Markov Chain Approximation Method 66

9.2 Euler's Method 67

9.2.1 Method of Antithetic Variables 68

9.2.2 Method of Control Variables 69

9.3 A Sample Path Example 71

10 Numerical Examples 73

10.1 First Scenario: Misspecification of the Post-Disorder Drift 74

10.2 Second Scenario: Misspecification of the Pre- and Post-disorder Drifts 77

10.3 Third Scenario: Misspecification of the Mean Time to Disorder . . 79

11 Conclusion 83

A Detailed Calculations for the Compound Poisson Disorder Problem 90

A.1 Calculation of the Bayes Risk 90

A.2 Dynamics of Q and C 91

A.3 Proof of Lemma 2 93

A.4 Proof of Proposition 2 98

A.5 Variational Inequalities 102

B Detailed Calculations for the Wiener Disorder Problem 105

B.1 The Dynamics of ξ and Φ 105

B.2 Independence of θ and X under $\mathbb{P}_Q$ 106

B.3 Derivation of the Objective Function 107

B.4 Proof of Lemma 5 108

B.5 Proof of Lemma 6 110

List of Figures

2.1 A sample path of the process Q 15

4.1 Sample path for the disorder problem with the stopping boundary, $\alpha = 5$ 35

4.2 Sample path for the disorder problem with the stopping boundary, $\alpha = 1$ 37

5.1 Stopping boundaries for two different mixed uniform priors 39

5.2 Risk for the stochastic and the deterministic problems for three different α values 41

5.3 Stopping boundaries for two different Gamma priors 42

5.4 Stopping boundary for the phase-type prior 43

5.5 Stopping boundary for the asymptotic prior 44

5.6 Stopping boundaries for the exponential problem 45

5.7 Density plots for the fraud problem 47

5.8 Plots on the fraud problem 48

5.9 A sample path example of the fraud problem 49

9.1 Sample path for the Φ and $\tilde{\Phi}$ processes 72

10.1 The effect of the decision maker's parameter for a fixed correct parameter 75

10.2 The effect of the correct parameter when the decision maker's parameter is fixed 75

10.3 The effect of the correct and incorrect parameters for Scenario 1 . 75

10.4 1% Tolerance regions for seven different user parameters 77

10.5 The effect of the decision maker's parameters when the correct parameters are $\mu = -\nu = 0.5 \cdot 0.226$ 78

10.6 The effect of the correct parameters when the decision maker's parameters are $\tilde{\mu} = -\tilde{\nu} = 0.5 \cdot 0.226$ 78

10.7 The effect of the correct and incorrect parameters for Scenario 2 . 78

10.8 0.5% Tolerance regions for five different user parameters 79

10.9 The effect of the decision maker's parameter for a fixed correct parameter 80

10.10 The effect of the correct parameter when the decision maker's parameter is fixed 80

10.11 The effect of the correct and incorrect parameters for Scenario 3 . 80

10.12 1% Tolerance regions for five different user parameters (a) 81

10.13 1% Tolerance regions for five different user parameters (b) 81

10.141% Tolerance regions for five different user parameters 81

Chapter 1

Introduction

In real life, when the dynamics of a system undergo a change, it is quite important and sometimes crucial to quickly detect the moment of disorder to avoid unwanted consequences. Unfortunately, for many systems, it is not possible to observe or know the disorder time beforehand. Thus, the best one can do is to observe the system responses and try to find a disorder detection rule to alert the user. Poisson and compound Poisson processes are useful for modeling discrete event arrivals, possibly attached to some random marks. Wiener models, on the other hand, are useful in financial models. There exists a vast literature about the detection problems on these mathematical models.

Basevelle and Nikiforov [1], Poor and Hadijiliadis [2] and Tartakovsky, Nikiforov and Baseville [3] have investigated various change detection problems using different methods and algorithms. However, the Bayesian formulations of the problem date back to Kolmogorov's studies on the sufficiency of Bayesian statistics and Shiryaev's work on the Bayesian methods for solving detection problems [4]. Bayesian methods aim to minimize a cost function

$$\mathbb{P}\{\tau < \theta\} + \alpha \mathbb{E}(\tau - \theta)^{+}$$

of detection rule τ for random disorder time θ and consists of the false alarm frequency and the cost of expected delay, using prior information about the disorder distribution. Here α denotes the unit cost of expected delay. Because of their convenience, geometric and exponential prior disorder distributions were used widely for discrete and continuous time problems, respectively. The main idea behind solving these problems is to find an optimal stopping rule by utilizing free boundary problems. Peskir and Shiryaev [5] explained how to use Markovian and martingale methods on the solution of these problems. Furthermore, Shiryaev [6] presented extended results on Markovian methods for optimal stopping problems.

The disorder problem for a Poisson process is firstly studied by Galchuk and Rozovskii [7], under certain assumptions regarding the Poisson arrival rates and the disorder density parameter λ of the zero-weighted exponential prior. In 1976, Davis [8] relaxed these assumptions and solved the problem for a more general scenario. Finally, Peskir and Shiryaev [9] solved the Poisson disorder problem with the linear penalty completely without any of these assumptions. Bayraktar et al. [10, 11] revisited the problem for different Bayesian risk functions

$$\mathcal{R}_\tau^{(1)}(\pi) := \mathbb{P}\{\tau < \theta - \varepsilon\} + \alpha\mathbb{E}(\tau-\theta)^+, \quad \mathcal{R}_\tau^{(2)}(\pi) := \mathbb{P}\{\tau < \theta\} + \alpha\mathbb{E}(\tau-\theta)^+,$$
$$\mathcal{R}_\tau^{(3)}(\pi) := \mathbb{E}(\theta-\tau)^+ + \alpha\mathbb{E}(\tau-\theta)^+, \quad \mathcal{R}_\tau^{(4)}(\pi) := \mathbb{P}\{\tau < \theta\} + c\mathbb{E}[e^{\tilde{\alpha}(\tau-\theta)^+} - 1],$$

and considered randomness of the post-disorder arrival rate. Buonaguidi [12] worked on the disorder of Lévy processes with completely monotone jumps. Gapeev [13] worked on some special cases of Lévy processes and solved the compound Poisson disorder problem assuming exponentially distributed marks. Dayanık and Sezer [14] used the method of Gugerli [15] and Davis[16], and solved the compound Poisson disorder problem for general mark distribution completely, using a piecewise-deterministic process and the method of successive approximations. Later, their methodology to solve compound Poisson disorder problem was used for a fraud detection problem by Buonaguidi et al. [17], which will be revisited later in this thesis. To relax the exponential prior assumption, Bayraktar and Sezer [18] considered a phase-type prior for the Poisson disorder problem, which is the distribution of time until absorption in a Markov chain. Phase-type priors have the property that any positive distribution can be approximated by a phase-type distribution, see [19, 20]. On the other hand, uniform prior was firstly used for Wiener disorder problems, see Zhitlukhin and Shiryaev [21, 22] and Sokko [23]. For the compound Poisson case, the problem was revisited by Ürü et al. [24], using the same methodology with compound Poisson problem with exponential prior. The problem turns out to be a finite horizon problem when the underlying prior is uniformly distributed.

Previous studies assumed continuous observation of the processes. Herberts and Jensen [25] solved the Poisson disorder problem by considering different levels of information availability, adopting various observation schemes such as sequential observation, ex post analysis, and a mixture of the two. Moreover, Brown [26, 27] considered the case where the process is observed at discrete time points and random time intervals, respectively.

When there is no realistic pre-existing statistical model for disorder, especially for surveillance and inspection systems, non-Bayesian methods are used, see Poor and Hadjiliadis [2]. Karoui, Loisel and Salhi [28] approached the problem using Lorden's minimax method, a popular robust optimization technique where the worst-case value of the delay is considered. Later, Figueroa-López and Ólafsson [29] worked on Lévy processes and used CUSUM (cumulative sum) procedure to solve the problem for the Poisson process where the jump intensity changes.

As mentioned earlier, in Bayesian formulations, mostly exponential distribution is used as prior disorder density, since it is a convenient distribution having memoryless property and frequently used in areas concerning reliability and waiting times. Moreover, because of its memoryless property, exponential distribution is essential for one-dimensional Markov sufficient statistic. On the other hand, uniform distribution is a good alternative for modeling the disorder time in cases where there is insufficient information about when the disorder occurs, or when it is known that the disorder is equally likely to occur in each time interval. Finally, it was mentioned that phase-type distributions are used for approximating other distributions. However, approximating any distribution can result in a large Markov chain, and the curse of dimensionality prohibits the calculation of the optimal detection rule.

When we have a better knowledge about the density of the disorder time and it is neither of the cases mentioned above, these models do not work very well and require better alternatives. For instance, we may have a knowledge that no disorder will occur during a specific interval of time, or we may know that the prior has a Gamma, Weibull or Lognormal density. Moreover, mixture of uniform distributions might be useful for approximating more complex densities, for instance. For this thesis, we will be dealing with a general zero-modified density as the disorder distribution.

The compound Poisson process is denoted by

$$X_t = X_0 + \sum_{k=1}^{N_t} Y_k,\ t \geq 0,$$

where $N = \{N_t : t \geq 0\}$ denotes a Poisson process with arrival rate $\lambda_0 > 0$, and the Y_i's are i.i.d. distributed $\mathbb{R}^d$ random variables coming from distribution ν_0 independent of N. At an unknown and unobserved time θ, the parameters of the compound Poisson process change from λ_0 to λ_1, and from ν_0 to ν_1, respectively. Moreover, the change time has a zero-modified prior distribution

$$\begin{aligned} &\mathbb{P}\{\theta = 0\} = \pi, \\ &\mathbb{P}\{\theta > t\} = (1-\pi)\int_t^{\infty} f_\theta(s)\ ds, \end{aligned}$$

where $f_\theta(\cdot)$ denotes the conditional disorder time density given that disorder happens after 0. Our aim is to minimize the Bayesian risk function

$$\mathcal{R}_\tau(\pi) := \mathbb{P}\{\theta > \tau\} + \alpha\mathbb{E}[(\tau-\theta)^+] \quad \pi \in [0,1], \alpha \in \mathbb{R}_+,$$

where τ is a stopping time of the natural filtration $\mathcal{F}$, of the process X. The first term denotes the false alarm frequency and the second term denotes the cost of expected delay, where α denotes the unit delay cost. A stopping time τ is called a Bayes optimal alarm time if it attains the minimum

$$V(\pi) := \inf_{\tau\in\mathcal{F}} \mathcal{R}_\tau(\pi). \tag{1.1}$$

A further investigation on the Bayesian risk problem in (1.1) results in an optimal stopping problem of a two-dimensional Markov sufficient process one of which is the deterministic clock process. Using the Markov property of this process and the properties of the value function, the problem reduces into a free boundary problem where the state space is separated into continuation and stopping regions, by a boundary curve. The properties of the boundary curve depend on the prior density of the disorder time. Moreover, there are no restrictions on the underlying density, in the context of its support, the problem can be an infinite horizon problem (for example, when the disorder density is from the Gamma family). For these cases, we will provide required conditions for the truncation of the time axis and the optimality and convergence criteria for this updated optimization problems. To calculate the value function, method of successive approximations and time-discretization will be used, and the optimal stopping rule will be explained in detail.

The Wiener disorder problem in a Bayesian framework was firstly formulated by Shiryaev [4, 30]. Here, the drift of a Wiener process changes at unobservable and unknown time θ, which was assumed to have zero-modified exponential distribution. The purpose of the problem was to minimize a Bayesian risk function consisting of a false alarm frequency and a cost of expected delay, of the form

$$\mathcal{R}^{\pi}(\tau) := \mathbb{P}^{\pi}\{\tau < \theta\} + \alpha\mathbb{E}^{\pi}(\tau - \theta)^{+},$$

where τ denotes the alarm time and $\alpha > 0$. By means of a sufficient statistic, the problem turned out to be a one dimensional free boundary problem, which was explained in detail by Shiryaev [31]. Beibel [32] later revisited the Wiener disorder problem considering the randomness of the post-disorder drift and an exponential penalty for the delay in

$$\mathcal{R}(\tau) := \mathbb{P}\{\tau < \theta\} + \alpha\mathbb{E}[e^{\tilde{\alpha}(\tau-\theta)^{+}} - 1],$$

where $\tilde{\alpha} > 0$ [33].

Gapeev and Peskir [34] and Peskir and Shiryaev [5] considered the finite horizon version of the Wiener disorder problem. In this case, the problem turns out to be two-dimensional, requiring non-linear integral equations to be solved. Dayanik et al. [35] addressed the disorder problem where the local characteristics of multi-dimensional independent compound Poisson and Wiener processes change simultaneously. They solved sequential optimal stopping problems using a jump operator. Sezer [36] considered the case where the drift change occurs due to shocks arriving according to a Poisson process, with the drift changing at one of these arrivals. Dayanık [37] revisited the disorder problem where the system is only observable at discrete time epochs.

Studies mentioned so far assumed that the disorder has mostly a zero-modified exponential prior density. A more general problem was addressed by Zhitlukhin and Shiryaev [21], a G-model where a general density for the disorder time was considered. Sokko [23] gave the complete solution of the problem where a zero-modified uniform prior density was adopted. This model fits well where there is no prior information where the disorder will occur or it is equally likely to happen at any time on a finite interval.

Because of its continuous characteristics, the Wiener process is often used in finance, especially to model stock prices. Moreover, changes in the drift of a stock price and detecting them quickly give rise to another formulation of the disorder problem. Beibel and Lerche [38], Shiryaev and Novikov [39], Ekström and Lindberg [40] considered the detection problem which aims to find an optimal stopping time where the drift of a geometric Brownian motion S changes with the dynamics

$$\frac{dS_t}{S_t} = \begin{cases} \mu dt + \sigma dW_t, & \text{for } t \leq \theta, \\ \nu dt + \sigma dW_t, & \text{for } t > \theta \end{cases}$$

and aims to maximize its expected discounted price

$$\mathbb{E}[e^{-r\tau} S_\tau 1_{\{\tau<\infty\}}].$$

This formulation determines the optimal selling time for an investor holding a stock. These formulations used zero-modified exponential prior and later, Shiryaev and Zhitlukhin [41] visited the case where the disorder was distributed with a zero-modified uniform prior density.

When stock prices are modeled using Brownian motion, detecting anomalies and drastic changes can be framed as a change-point detection problem. Market bubbles and crises exemplify these kinds of situations. Shiryaev, Zhitlukhin and Ziemba [42] applied the model to Apple stock price bubble of 2009-2012 and internet bubble crash of 2000-2002, using a uniform prior for the disorder. Later, they worked on the land and stock bubbles in Japan circa 1990 and 2003 [43], this time using a discrete time model with geometric Gaussian random walk, where the disorder has a discrete probability density with probabilities $p_1, p_2, \ldots, p_T$. Finally, Zhitlukhin and Ziemba [44] worked on certain crashes in the US and China markets, using a discrete uniform density for the disorder time.

The drift after the disorder was assumed to be known, so far. However, this is generally not the case for the real-life scenarios. Muravlev and Shiryaev [45] considered the two-sided disorder problem where the drift after disorder is ν_1 with probability ρ_1 and ν_2 with probability $1-\rho_1$. Later, Ekström and Vaicenavicius [46] assumed that the post-disorder drift is a random variable and investigated the robustness of the problem by taking the misspecification of the process parameters into account.

The papers on crashes and bubbles addressed retrospective problems using data that had already been observed. Assuming that the drift before the disorder is already known from the data, a performance analysis was conducted for various options for the post-disorder drift. With historical stock prices, the optimal choice for the post-disorder was determined. For real-time financial models, misspecification of the problem parameters poses a problem for solving the change-point detection. Blanchet and Scalliet et al. [47, 48] worked on the performance of various trading strategies under parameter misspecification.

In this thesis, we deal with an asset selling problem where we consider a simple market model with a risky and a risk free asset. It is assumed that the dynamics of the stock change at an unknown and unobserved time θ, which has a zero-weighted exponential prior density

$$\mathbb{P}\{\theta = 0\} = \pi \quad \text{and} \quad \mathbb{P}\{\theta > t\} = (1-\pi)e^{-\lambda t}.$$

Thus, the dynamics of the stock can be modeled as

$$\frac{dR_t}{R_t} = r dt \quad \text{and} \quad \frac{dS_t}{S_t} = [\mu 1_{\{t<\theta\}} + \nu 1_{\{\theta \le t\}}]dt + \sigma dW_t.$$

Here, r denotes the risk free rate, μ and ν are drift parameters before and after the disorder. W denotes the Wiener process, and σ is the volatility parameter.

The objective function of the problem, which concerns the maximization of the present value of the stock sold at optimal time τ using linear utility, can be written as

$$\sup_{\tau} \mathbb{E}[1_{\{\tau<\infty\}} P_\tau],$$

where $P_t := e^{-rt} S_t$ is the present value of the stock price at time t, using risk-free interest rate r.

In the second part of the thesis, firstly, we will provide solution methods to this problem. While solving, we will introduce the odds-ratio process

$$\Phi_t := \frac{\mathbb{E}[1_{\{\theta\leq t\}}|\mathcal{F}_t]}{\mathbb{E}[1_{\{\theta>t\}}|\mathcal{F}_t]}.$$

We will give an alternative representation of the objective function, and it turns out to be an optimal stopping problem, with the process Φ. Moreover, we will see that there exists a ϕ^* that divides the state space into continuation and stopping regions. For the solution of the original problem, we will use Markov chain approximation method of Kushner and Dupuis [49].

Secondly, we will deal with the misspecification of the problem parameters. The misspecified problem parameters are denoted $\tilde{\pi}$, $\tilde{\lambda}$, $\tilde{\mu}$ and $\tilde{\nu}$. The decision maker solves the problem with these parameters and finds an incorrect stopping rule, thus, calculates an incorrect value function. We will present the dynamics of the process $\tilde{\Phi}$, which is the misspecified version of the process Φ. It turns out that both Φ and $\tilde{\Phi}$ are Itô processes, run by the same Wiener process. We will use this characterization to calculate incorrectly calculated value function by the help of Monte Carlo simulations.

This thesis is divided into two parts, Part I: Compound Poisson Disorder Problem (Chapters 2-5) and Part II: Misspecified Wiener Disorder Problem (Chapters 6-10).

Part I

Compound Poisson Disorder Problem

In Chapter 2, the compound Poisson disorder problem and the required modelling elements will be explained and the value function for the optimal stopping problem will be constructed. Moreover, the time-truncated version of the problem will be provided in this chapter, as well. In Chapters 3 and 4, successive approximation and time-discretization methods, and their implementation will be presented respectively, and the procedure to stop a system with disorder by alarming the user will be explained. In Chapter 5, various numerical examples with different prior densities will be discussed. Long proofs and some other mathematical results will be presented in Appendix A.

Chapter 2

Model and Problem Formulation

In this chapter, we will define our model and the auxiliary value function which turns the detection problem into an optimal stopping problem. Moreover, we will introduce the time truncated version of the value function for the densities with infinite support.

Let $(\Omega, \mathcal{F}, \mathbb{P}_0)$ be a probability space hosting a compound Poisson process X,

$$X_t = X_0 + \sum_{k=1}^{N_t} Y_k,\ t \geq 0.$$

N is the standard Poisson process with parameter λ_0, and Y_i's are i.i.d. $\mathbb{R}^d$-valued random variables with distribution ν_0, independent of N. The jump times of this Poisson process are denoted by

$$W_n \equiv \inf\{t > W_{n-1} : X_t \neq X_{t-}\},\ n \geq 1\ (W_0 \equiv 0).$$

Let $\mathcal{F}$ denote the augmentation of this process' natural filtration $\sigma(X_s, s < t)$, $t \geq 0$ with $\mathbb{P}_0$-null sets. ($\mathcal{F} = \{\mathcal{F}_t; t \geq 0\}$).

Suppose the dynamics of the compound Poisson process X change at some unknown and unobserved time θ. In particular, Poisson arrival rate changes from λ_0 to λ_1 and jump distribution changes from ν_0 to ν_1. This change time has a prior distribution with

$$\begin{aligned}
&\mathbb{P}_0\{\theta = 0\} = \pi, \\
&\mathbb{P}_0\{\theta \in dt | \theta > 0\} = f_\theta(t)dt \quad \mathbb{P}_0\{\theta \le t\} = \pi + (1-\pi)\int_0^t f_\theta(s)ds, \\
&F_\theta(t) = \mathbb{P}_0\{\theta \le t | \theta > 0\} = \int_0^t f_\theta(s)ds.
\end{aligned}$$

Here, the disorder time θ has a known conditional distribution with density $f_\theta(t)$ and cumulative distribution function $F_\theta(t)$.

We will denote the enlargement of $\mathcal{F}$ with $\sigma(\theta)$ with $\mathcal{G} = \{\mathcal{G}_t; t \ge 0\}$. Thus, $\mathcal{G}_t \equiv \mathcal{F}_t \vee \sigma(\theta)$.

Let ν_1 be absolutely continuous with ν_0. (Otherwise, the problem becomes trivial.) A new probability measure $\mathbb{P}$ is defined on $(\Omega, \vee_{s\ge 0}\mathcal{G}_s)$ by a change of measure using the Radon-Nikodym derivative of $\mathbb{P}$ with respect to $\mathbb{P}_0$ as

$$\left.\frac{d\mathbb{P}}{d\mathbb{P}_0}\right|_{\mathcal{G}_t} = Z_t \equiv 1_{\{t<\theta\}} + 1_{\{t\ge\theta\}} e^{-(\lambda_1-\lambda_0)(t-\theta)} \prod_{k=N_{\theta-}+1}^{N_t} \left[\frac{\lambda_1}{\lambda_0} h(Y_k)\right],\ t \ge 0.$$

Here $N_{\theta-}$ denotes the number of arrivals in the interval $[0,\theta)$, whereas $h(\cdot)$ denotes the Radon-Nikodym derivative of ν_1 with respect to ν_0 as

$$h(y) \equiv \left.\frac{d\nu_1}{d\nu_0}\right|_{\mathcal{B}(\mathbb{R}^d)}(y),\ y \in \mathbb{R}^d.$$

Moreover Z_t can be written as

$$Z_t \equiv 1_{\{t<\theta\}} + 1_{\{t\ge\theta\}} \frac{R_t}{R_\theta},\ t \ge 0$$

using the likelihood ratio process $R = \{R_t, t \ge 0\}$ defined as

$$R_t \equiv e^{-(\lambda_1-\lambda_0)t} \prod_{k=1}^{N_t} \left[\frac{\lambda_1}{\lambda_0} h(Y_k)\right] \text{ for } t \ge 0.$$

Observe that under the measure $\mathbb{P}$ the change occurs at θ, while change never occurs under the measure $\mathbb{P}_0$. Moreover, $Z_0 = 1$ $\mathbb{P}_0$-a.s. therefore $\mathbb{P}_0$ and $\mathbb{P}$ agree on $\mathcal{G}_0 = \sigma(\theta)$. This means that θ has the same distribution under these probability measures. Thus, the properties of X and θ remain the same within two measures. Note that we are motivated to work with measure $\mathbb{P}_0$, since it is simpler to make calculations under this measure.

We want to find a quickest detection rule for the unknown and unobserved disorder time θ. In other words, we want to find an $\mathcal{F}$-stopping time τ that minimizes the Bayesian risk function defined as

$$\mathcal{R}_\tau(\pi) \equiv \mathbb{P}\{\theta > \tau\} + \alpha\mathbb{E}[(\tau - \theta)^+] \quad \pi \in [0,1], \alpha \in \mathbb{R}_+, \tau \in \mathcal{F},$$

where $\mathbb{P}\{\theta > \tau\}$ represents the false alarm frequency and $\alpha\mathbb{E}[(\tau-\theta)^+]$ represents the cost of the expected delay. Note that α can also be considered as a penalty for the delay. Therefore, it is a measure of decision maker's tendency to wait for new information.

If τ attains the function

$$V(\pi) \equiv \inf_{\tau \in \mathcal{F}} \mathcal{R}_\tau(\pi) \quad \pi \in [0,1],$$

then it is a Bayes-optimal alarm time. The Bayes risk function and the corresponding objective function can be written as

$$\begin{aligned}
\mathcal{R}_\tau(\pi) &= \mathbb{P}\{\theta > \tau\} + \alpha\mathbb{E}[(\tau - \theta)^+] \\
&= (1-\pi) + \alpha(1-\pi)\mathbb{E}_0 \int_0^\tau \Big(\frac{1}{1-\pi}\mathbb{E}_0[Z_t 1_{\{\theta \le t\}} | \mathcal{F}_t] - \frac{1}{\alpha} f_\theta(t)\Big) dt, \\
V(\pi) &\equiv \inf_{\tau \in \mathcal{F}} \mathcal{R}_\tau(\pi) \\
&= (1-\pi) + \alpha(1-\pi)\mathbb{E}_0 \int_0^\tau \Big(\frac{1}{1-\pi}\mathbb{E}_0[Z_t 1_{\{\theta \le t\}} | \mathcal{F}_t] - \frac{1}{\alpha} f_\theta(t)\Big) dt \\
&= (1-\pi) + \alpha(1-\pi) \inf_{\tau \in \mathcal{F}} \mathbb{E}_0 \int_0^\tau R_t \Big[\frac{\pi}{1-\pi} + \int_0^t f_\theta(u) \frac{1}{R_u} du\Big] - \frac{1}{\alpha} f_\theta(t)\ dt.
\end{aligned}$$

The term $R_t \left[\frac{\pi}{1-\pi} + \int_0^t f_\theta(u) \frac{1}{R_u} du\right]$ motivates us to define a new process.

2.1 A Markov Pair

Let us define the pair process (Q, C) where

$$\begin{aligned} C_t &:= c + t, \\ Q_t &:= R_t\Big[q + \int_0^t f_\theta(C_u)\frac{1}{R_u}\, du\Big] = R_t\Big[q + \int_0^t f_\theta(c+u)\frac{1}{R_u}\, du\Big], \end{aligned} \tag{2.1}$$

with $Q_0 = q$ and $C_0 = c$. The process C can be regarded as the deterministic clock process starting from an initial value c. With $c = 0$ and $q = \frac{\pi}{1-\pi}$, Q_t coincides with $\frac{1}{1-\pi}\mathbb{E}_0[Z_t 1_{\{\theta \le t\}}|\mathcal{F}_t] = R_t\left[\frac{\pi}{1-\pi} + \int_0^t f_\theta(u)\frac{1}{R_u}du\right]$. Therefore, with this initial values, we can write the expectation in our Bayes risk, using (Q, C) as

$$\mathbb{E}_0 \int_0^\tau R_t\Big[\frac{\pi}{1-\pi} + \int_0^t f_\theta(u)\frac{1}{R_u}du\Big] - \frac{1}{\alpha}f_\theta(t)\, dt = \mathbb{E}_0\int_0^\tau \Big(Q_t - \frac{1}{\alpha}f_\theta(C_t)\Big)dt.$$

For $t, s \ge 0$, it is clear that $C_{s+t} = C_s + t$, and for Q, we have

$$\begin{aligned} Q_{s+t} &= R_{s+t}\Big[q + \int_0^{s+t} f_\theta(C_u)\frac{1}{R_u}\, du\Big] \\ &= \frac{R_{s+t}}{R_s}R_s\Big[q + \int_0^s f_\theta(C_u)\frac{1}{R_u}\, du + \int_s^{s+t} f_\theta(C_u)\frac{1}{R_u}\, du\Big] \\ &= \frac{R_{s+t}}{R_s}\Big[Q_s + \int_s^{s+t} f_\theta(C_u)\frac{1}{R_u/R_s}\, du\Big]. \end{aligned} \tag{2.2}$$

We can write $\frac{R_{s+t}}{R_s}$ as

$$\frac{R_{s+t}}{R_s} = e^{-(\lambda_1-\lambda_0)t}\prod_{k=N_s+1}^{N_{s+t}}\Big[\frac{\lambda_1}{\lambda_0}h(Y_k)\Big].$$

Note that conditioned on $\mathcal{F}_s$, $\frac{R_{s+t}}{R_s}$ has the same same law as the unconditional law of R_t. Thus, if we define $\tilde{R}_t := \frac{R_{s+t}}{R_s}$, we can write (2.2) as

$$\begin{aligned} Q_{s+t} &= \tilde{R}_t\Big[Q_s + \int_s^{s+t} f_\theta(C_u)\frac{1}{\widetilde{R}_{u-s}}\, du\Big] \\ &= \tilde{R}_t\Big[Q_s + \int_s^{s+t} f_\theta(C_s + (u-s))\frac{1}{\widetilde{R}_{u-s}}\, du\Big] \\ &= \tilde{R}_t\Big[Q_s + \int_0^t f_\theta(C_s + w)\frac{1}{\widetilde{R}_w}\, dw\Big] \\ &= \tilde{R}_t\Big[q + \int_0^t f_\theta(c+w)\frac{1}{\widetilde{R}_w}\, dw\Big]\Big|_{q=Q_s,\ c=C_s}. \end{aligned} \tag{2.3}$$

Comparing this with (2.1) we see that (Q,C) is a time-homogeneous Markov pair. Moreover, from the dynamics of (Q,C), we conclude that, between arrivals, Q follows deterministic curves. If we define

$$x(q,c,t)=\begin{cases} e^{-(\lambda_1-\lambda_0)t}\Big[q+\int_0^t e^{(\lambda_1-\lambda_0)s}f_\theta(c+s)\ ds\Big], & \text{if } \lambda_0\neq\lambda_1,\\ q+\int_0^t f_\theta(c+s)\ ds, & \text{if } \lambda_0=\lambda_1,\end{cases}$$

then for $i\geq 0$ we have

$$Q_t=\begin{cases} x(Q_{W_i},C_{W_i},t-W_i), & \text{if } W_i\leq t<W_{i+1},\\ x(Q_{W_i},C_{W_i},W_{i+1}-W_i)\frac{\lambda_1}{\lambda_0}h(Y_{i+1}), & \text{if } t=W_{i+1},\end{cases}$$
$$C_t=C_{W_i}+t-W_i,\quad \text{if } W_i\leq t\leq W_{i+1}.$$

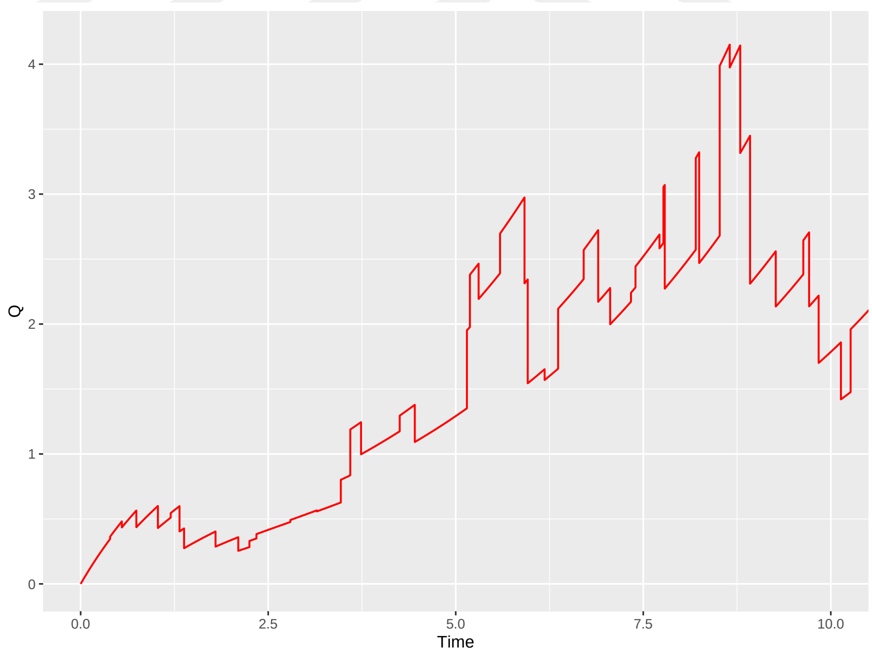


Figure 2.1: A sample path of the process Q

Figure 2.1 illustrates a sample path for the process Q. As it seen on the figure, the process jumps at arrivals and follows deterministic curves between arrivals. Intuitively, any arrival is a new piece of information, and the process is updated instantly with this new information by the multiplicative term $\frac{\lambda_1}{\lambda_0}h(Y_i)$.

Next, we define

$$U(q,c) := \inf_{\tau\in\mathcal{F}} \mathbb{E}_0^{(q,c)} \int_0^\tau g(Q_t, C_t)\,dt \qquad \text{where } g(q,c) := q - \frac{1}{\alpha} f_\theta(c), \quad (q,c)\in\mathbb{R}_+^2, \tag{2.4}$$

and $\mathbb{E}_0^{(q,c)}$ denotes the expectation under the measure $\mathbb{P}_0$ with $Q_0 = q, C_0 = c$ with probability one. In terms of the value function, our smallest Bayes risk can be written as

$$1-\pi+\alpha(1-\pi)\,U\big(\frac{\pi}{1-\pi},0\big),$$

and if there is an optimal stopping time for the problem in (2.4), it also gives us an optimal detection rule.

Remark 1. *Because Q depends on $Q_0 = q$ linearly (see 2.1) with a positive multiplying term, it follows that the value function in (2.4) is concave and non-decreasing in q (for every fixed c). Moreover, it is bounded above by 0 (immediate stopping). Hence, the boundary*

$$\beta(c) := \min\{q : U(q,c) = 0\}, \qquad c \geq 0,$$

separates the state space into continuation and stopping regions defined respectively as

$$\mathcal{C} := \{(q,c) : U(q,c) < 0\} \quad \textit{and} \quad \Gamma := \{(q,c) : U(q,c) = 0\}.$$

Above the boundary is the stopping region, and below is the continuation region.

Since we are working with general prior densities, one might ask if there is a finite boundary for the unbounded densities. In other words, when there exists $c_\infty \in \mathbb{R}_+$ such that $f_\theta(c_\infty-) = \infty$ or $f_\theta(c_\infty+) = \infty$. The fact that there isn't is shown after Remark 2.

Remark 2. *For a deterministic time $t \in \mathbb{R}_+$, we have*

$$\mathbb{E}_0^{(q,c)}[Q_t] = q + \int_0^t f_\theta(c+s)ds = q + F_\theta(c+t) - F_\theta(c). \tag{2.5}$$

Moreover, Q is a non-negative submartingale.

Proof. Using the definition of Q, the fact that R is martingale, Fubini's theorem and the tower property, we have

$$\begin{aligned}
&\mathbb{E}_0^{(q,c)}[Q_t] = q\mathbb{E}_0^{(q,c)}[R_t] + \mathbb{E}_0^{(q,c)}\left[\int_0^t f_\theta(c+u)\frac{R_t}{R_u}\,du\right]\\
&= q + \int_0^t f_\theta(c+u)\mathbb{E}_0^{(q,c)}\left[\frac{R_t}{R_u}\right]\,du = q + \int_0^t f_\theta(c+u)\,du\\
&= q + F_\theta(c+t) - F_\theta(c).
\end{aligned}$$

Note that $\mathbb{E}_0^{(q,c)}[Q_t] < \infty$, and for the submartingale property of the process Q, for $t, s \in \mathbb{R}_+$ we write

$$\begin{aligned}
&\mathbb{E}_0^{(q,c)}[Q_{t+s}|\mathcal{F}_t] = \mathbb{E}_0^{(q,c)}\left[qR_{t+s} + \int_0^{t+s} f_\theta(C_u)\frac{R_{t+s}}{R_u}\,du\Big|\mathcal{F}_t\right]\\
&= qR_t + \int_0^t f_\theta(C_u)\mathbb{E}_0^{(q,c)}\left[\frac{R_{t+s}}{R_u}\Big|\mathcal{F}_t\right]\,du + \int_t^{t+s} f_\theta(C_u)\mathbb{E}_0^{(q,c)}\left[\frac{R_{t+s}}{R_u}\Big|\mathcal{F}_t\right]\,du\\
&= \underbrace{qR_t + R_t\int_0^t f_\theta(C_u)\frac{1}{R_u}\,du}_{Q_t} + \int_t^{t+s} f_\theta(C_u)\mathbb{E}_0^{(q,c)}\left[\frac{R_{t+s}}{R_u}\Big|\mathcal{F}_t\right]\,du.
\end{aligned}$$

Since the last term is nonnegative, we are done. □

Back to unbounded densities, observe that the deterministic optimization problem

$$\inf_{t\in\mathbb{R}_+} \mathbb{E}_0^{(q,c)}\int_0^t g(Q_s, C_s)\,ds$$

is a sub-problem of the original problem (2.4) and using similar arguments with Remark 1 it has its own boundary curve. Since deterministic stopping times are subset of the stopping times of the original problem, deterministic problem takes values greater than $U(q,c)$ for $(q,c) \in \mathbb{R}_+^2$. A direct result is that the deterministic boundary curve is bounded above by β.

Using Remark 2 and Fubini's theorem, the deterministic optimization problem can be written as

$$\inf_{t\in\mathbb{R}_+} \mathbb{E}_0^{(q,c)}\int_0^t g(Q_s, C_s)\,ds = \inf_{t\in\mathbb{R}_+}\left\{\mathbb{E}_0^{(q,c)}\left[\int_0^t Q_s\,ds\right] - \frac{1}{\alpha}\int_0^t f_\theta(c+u)\,du\right\}$$

$$= \inf_{t \in \mathbb{R}_+} \left\{ \int_0^t (q + F_\theta(c+s) - F_\theta(c)) \ ds - \frac{1}{\alpha}[F_\theta(c+t) - F_\theta(c)] \right\}$$
$$= \inf_{t \in \mathbb{R}_+} \left\{ [q - F_\theta(c)]t + \int_0^t F_\theta(c+s) \ ds - \frac{1}{\alpha}[F_\theta(c+t) - F_\theta(c)] \right\}$$

and if f_θ is continuous at c, we can differentiate the expression with respect to t and get

$$\frac{\delta}{\delta t}\left[[q - F_\theta(c)]t + \int_0^t F_\theta(c+s) \ ds - \frac{1}{\alpha}[F_\theta(c+t) - F_\theta(c)] \right]$$
$$= [q - F_\theta(c)] + F_\theta(c+t) - \frac{1}{\alpha} f_\theta(c+t).$$

Observe that when $t = 0$, the derivative is $q - \frac{1}{\alpha} f_\theta(c)$. Therefore, when $q - \frac{1}{\alpha} f_\theta(c) < 0$, the function is strictly decreasing and the value function of the deterministic problem for this (q, c) pair is strictly negative, in other words, below the stopping region. Hence, the deterministic boundary curve is bounded below by $\frac{1}{\alpha} f_\theta$ at points of continuity of f_θ.

Corollary 1. *Since β is bounded below by the deterministic boundary curve, and the deterministic boundary curve is bounded below by $\frac{1}{\alpha} f_\theta$, for the original problem, if there exists a continuity point $c_\infty \in \mathbb{R}_+$ of f_θ with $f_\theta(c_\infty -) = \infty$ or $f_\theta(c_\infty +) = \infty$, we have unbounded boundary at c_∞. In other words, $\beta(c_\infty) = \infty$.*

2.2 Truncation of C-Space

Since we are working with general prior densities, for the densities with unbounded support it is important to restrict c values on a compact interval such that the value function takes the value 0 outside of this interval. For a fixed $\bar{C} \in \mathbb{R}_+$, define the deterministic stopping time $\tau_{\bar{C}} \equiv \inf\{t \geq 0 : C_t \geq \bar{C}\} = (\bar{C} - c)^+$, where $c = C_0$. $\tau_{\bar{C}}$ denotes the exit time of the clock process C from the compact interval $[0, \bar{C}]$.

Using $\tau_{\bar{C}}$, truncated problem can be defined as

$$U^{(\bar{C})}(q, c) \equiv \inf_{\tau \in \mathcal{F}} \int_0^{\tau \wedge \tau_{\bar{C}}} g(Q_t, C_t) \ dt. \tag{2.6}$$

Observe that truncated value function $U^{(\bar{C})}$ satisfies the property that it is 0 outside of $[0, \bar{C}]$.

Lemma 1. *For every $(q, c) \in \mathbb{R}^+ \times [0, \bar{C}]$, as $\bar{C} \to \infty$, $U^{(\bar{C})}(q, c)$ converges to $U(q, c)$ uniformly. We have*

$$0 \le U^{(\bar{C})}(q, c) - U(q, c) \le \frac{1}{\alpha}[1 - F_\theta(\bar{C})].$$

Proof. Observe that

$$\begin{aligned}
&\mathbb{E}_0^{(q,c)} \int_0^\tau g(Q_t, C_t)\, dt = \mathbb{E}_0^{(q,c)} \left[\int_0^{\tau \wedge \tau_{\bar{C}}} g(Q_t, C_t)\, dt + 1_{\{\tau > \tau_{\bar{C}}\}} \int_{\tau_{\bar{C}}}^\tau g(Q_t, C_t)\, dt \right] \\
&\ge \mathbb{E}_0^{(q,c)} \left[\int_0^{\tau \wedge \tau_{\bar{C}}} g(Q_t, C_t)\, dt - 1_{\{\tau > \tau_{\bar{C}}\}} \frac{1}{\alpha} \int_{\tau_{\bar{C}}}^\tau f_\theta(C_t)\, dt \right] \\
&\ge \mathbb{E}_0^{(q,c)} \int_0^{\tau \wedge \tau_{\bar{C}}} g(Q_t, C_t)\, dt - \frac{1}{\alpha} \int_{\tau_{\bar{C}}}^\infty f_\theta(C_t)\, dt \\
&= \mathbb{E}_0^{(q,c)} \int_0^{\tau \wedge \tau_{\bar{C}}} g(Q_t, C_t)\, dt - \frac{1}{\alpha}[1 - F_\theta(\tau_{\bar{C}} + c)] \\
&= \mathbb{E}_0^{(q,c)} \int_0^{\tau \wedge \tau_{\bar{C}}} g(Q_t, C_t)\, dt - \frac{1}{\alpha}[1 - F_\theta(\bar{C} \vee c)] \\
&\ge \mathbb{E}_0^{(q,c)} \int_0^{\tau \wedge \tau_{\bar{C}}} g(Q_t, C_t)\, dt - \frac{1}{\alpha}[1 - F_\theta(\bar{C})].
\end{aligned}$$

For every $\epsilon > 0$, if we choose $\bar{C}$ such that $\frac{1}{\alpha}[1 - F_\theta(\bar{C})] \le \epsilon$ and if we take the infimum of the both sides, we get $U(q, c) \ge U^{(\bar{C})}(q, c) - \epsilon$. Moreover, it is clear that $U^{(\bar{C})}(q, c) \ge U(q, c)$, hence we have the inequality $U^{(\bar{C})}(q, c) \ge U(q, c) \ge U^{(\bar{C})}(q, c) - \epsilon$. For $\epsilon > 0$, taking $\bar{C}$ as

$$\bar{C}_\epsilon = \inf\{c \ge 0 : F_\theta(c) \ge 1 - \alpha\epsilon\}$$

gives the required inequality. □

Lemma 2. *The function $U^{(\bar{C})}$ is locally (jointly) continuous. We have*

$$\begin{aligned}
&0 \le U^{(\bar{C})}(q_2, c) - U^{(\bar{C})}(q_1, c) \le (q_2 - q_1)\bar{C}, \quad \textit{for } q_1 < q_2, \\
&\min\{-\bar{C}, -\frac{1}{\alpha}\} \int_0^{\bar{C}} |f_\theta(u) - f_\theta(u - (c_2 - c_1))|\, du \le U^{(\bar{C})}(q, c_2) - U^{(\bar{C})}(q, c_1) \\
&\le \bar{C} \int_0^{\bar{C}} |f_\theta(u) - f_\theta(u - (c_1 - c_2))|\, du + \frac{1}{\alpha}[F_\theta(c_2) - F_\theta(c_1)], \quad \textit{for } c_1 < c_2.
\end{aligned}$$

Proof. The proof is presented at Appendix A.3. $\square$

Corollary 2. *The continuity and the uniform convergence of $U^{\bar{C}}$ to U guarantees that U is also continuous.*

Using Remark 1 and Lemma 1, we can conclude that, solving the truncated optimal stopping problem and stopping when Q enters to the stopping region will provide the optimal alarm time. Next two chapters will explain two different methods for finding the function $U^{\bar{C}}$.

Chapter 3

Successive Approximations

This method constructs the value function as a sequence of converging functions, using the dynamic programming principle. The method was used widely to solve disorder problems. See Dayanık and Sezer [14], Bayraktar and Sezer [18], Ludkovski and Sezer [50], Ürü et al. [24] and Buonaguidi et al. [12].

Let us define

$$U_n(q,c) \equiv \inf_{\tau\in\mathcal{F}} \mathbb{E}_0^{(q,c)}\Big[\int_0^{\tau\wedge W_n} \underbrace{\Big(Q_s - \frac{1}{\alpha} f_\theta(C_s)\Big)}_{g(Q_s,C_s)} ds\Big], \quad (q,c)\in\mathbb{R}_+\times\mathbb{R}_+, n\in\mathbb{N},$$

and the truncated version of it as,

$$\begin{aligned} U_n^{(\bar{C})}(q,c) &\equiv \inf_{\tau\in\mathcal{F}} \mathbb{E}_0^{(q,c)}\Big[\int_0^{\tau\wedge W_n\wedge\tau_{\bar{C}}} \Big(Q_s - \frac{1}{\alpha} f_\theta(C_s)\Big) ds\Big], \quad (q,c)\in\mathbb{R}_+\times[0,\bar{C}], n\in\mathbb{N} \\ &= \inf_{\tau\in\mathcal{F}_{[0,\bar{C}-c]}} \mathbb{E}_0^{(q,c)}\Big[\int_0^{\tau\wedge W_n} \Big(Q_s - \frac{1}{\alpha} f_\theta(C_s)\Big) ds\Big]. \end{aligned}$$

Lemma 3.

$$\frac{F_\theta(c)-F_\theta(\bar{C})}{\alpha} \le U_n^{(\bar{C})}(q,c) \le 0 \quad \textit{for } n\in\mathbb{N}$$

Proof. Observe that, for the truncated problem, 0 is an upper bound for $U_n^{(\bar{C})}$ when $\tau = 0$. On the other hand $Q_s \ge 0$ for all $s\in\mathbb{R}_+$ and $\tau \le \bar{C}-c$ a.s. We have

$$\int_0^{\tau\wedge W_n} Q_s - \frac{1}{\alpha}f_\theta(C_s)ds \geq -\int_0^{\tau\wedge W_n} \frac{1}{\alpha}f_\theta(C_s)ds \geq -\int_0^{\tau} \frac{1}{\alpha}f_\theta(C_s)ds$$
$$\geq -\int_0^{\bar{C}-c} \frac{1}{\alpha}f_\theta(C_s)ds = -\int_0^{\bar{C}-c} \frac{1}{\alpha}f_\theta(c+s)ds = -\int_c^{\bar{C}} \frac{1}{\alpha}f_\theta(u)du$$
$$= -\frac{1}{\alpha}[F_\theta(\bar{C}) - F_\theta(c)] = \frac{F_\theta(c) - F_\theta(\bar{C})}{\alpha},$$

if we take the expectation and infimum of the both sides over $\tau \in \mathcal{F}_{[0,\bar{C}]}$ we see that $\frac{F_\theta(c)-F_\theta(\bar{C})}{\alpha}$ is a lower bound for $U_n^{(\bar{C})}$. □

Moreover, $(U_n^{(\bar{C})}(q,c))_{n\geq 1}$ is decreasing, because $(W_n)_{n\geq 1}$ is an a.s. increasing sequence. Hence, $\lim_{n\to\infty} U_n^{(\bar{C})}(q,c)$ exists everywhere.

Proposition 1. *For every $\bar{C} \in \mathbb{R}_+$, $U_n^{(\bar{C})}(q,c) \to U^{(\bar{C})}(q,c)$ uniformly in $(q,c) \in \mathbb{R} \times [0,\bar{C}]$. Moreover, for every $q \in \mathbb{R}_+$ and $c \in [0,\bar{C}]$ we have*

$$-\frac{1}{\alpha}[1 - \mathbb{E}_0[F_\theta(W_n)]] \leq U^{(\bar{C})}(q,c) - U_n^{(\bar{C})}(q,c) \leq 0$$

where $F_\theta(c) = \mathbb{P}_0\{\theta \leq c | \theta > 0\}$.

Proof.

$$\begin{aligned}
&\mathbb{E}_0^{(q,c)}\Big[\int_0^{\tau} g(Q_s,C_s)ds\Big] \\
&= \mathbb{E}_0^{(q,c)}\Big[\int_0^{\tau\wedge W_n} g(Q_s,C_s)ds + 1_{\{\tau>W_n\}}\int_{W_n}^{\tau} g(Q_s,C_s)ds\Big] \\
&= \mathbb{E}_0^{(q,c)}\Big[\int_0^{\tau\wedge W_n} g(Q_s,C_s)ds\Big] + \mathbb{E}_0^{(q,c)}\Big[1_{\{\tau>W_n\}}\int_{W_n}^{\tau} g(Q_s,C_s)ds\Big] \\
&\geq U_n^{(\bar{C})}(q,c) + \mathbb{E}_0^{(q,c)}\Big[1_{\{\tau>W_n\}}\int_{W_n}^{\tau} \underbrace{g(Q_s,C_s)}_{Q_s-\frac{1}{\alpha}f_\theta(C_s)\geq -\frac{1}{\alpha}f_\theta(C_s)} ds\Big] \\
&\geq U_n^{(\bar{C})}(q,c) + \mathbb{E}_0^{(q,c)}\Big[1_{\{\tau>W_n\}}\int_{W_n}^{\tau} -\frac{1}{\alpha}f_\theta(C_s)ds\Big] \\
&\geq U_n^{(\bar{C})}(q,c) + \mathbb{E}_0^{(q,c)}\Big[\int_{W_n}^{\infty} -\frac{1}{\alpha}f_\theta(C_s)ds\Big] = U_n^{(\bar{C})}(q,c) - \frac{1}{\alpha}\mathbb{E}_0^{(q,c)}\Big[\int_{W_n}^{\infty} f_\theta(c+s)ds\Big] \\
&= U_n^{(\bar{C})}(q,c) - \frac{1}{\alpha}\mathbb{E}_0^{(q,c)}\Big[\int_{W_n+c}^{\infty} f_\theta(s)ds\Big] = U_n^{(\bar{C})}(q,c) - \frac{1}{\alpha}\mathbb{E}_0^{(q,c)}[1 - F_\theta(W_n+c)]]
\end{aligned}$$

$$\geq U_n^{(\bar{C})}(q,c) - \frac{1}{\alpha}\mathbb{E}_0[1 - F_\theta(W_n)]] = U_n^{(\bar{C})}(q,c) - \frac{1}{\alpha}[1 - \mathbb{E}_0[F_\theta(W_n)]].$$

Thus

$$\mathbb{E}_0^{(q,c)}\Big[\int_0^\tau g(Q_s, C_s)ds\Big] \geq U_n^{(\bar{C})}(q,c) - \frac{1}{\alpha}[1 - \mathbb{E}_0[F_\theta(W_n)]]$$

if we take infimum of the both sides over $\tau \in \mathcal{F}_{[0,\bar{C}-c]}$ we get

$$U^{(\bar{C})}(q,c) \geq U_n^{(\bar{C})}(q,c) - \frac{1}{\alpha}[1 - \mathbb{E}_0[F_\theta(W_n)]] \quad \text{or equivalently}$$
$$-\frac{1}{\alpha}[1 - \mathbb{E}_0[F_\theta(W_n)]] \leq U^{(\bar{C})}(q,c) - U_n^{(\bar{C})}(q,c).$$

Thus

$$U_n^{(\bar{C})}(q,c) \geq U^{(\bar{C})}(q,c) \geq U_n^{(\bar{C})}(q,c) - \frac{1}{\alpha}[1 - \mathbb{E}_0[F_\theta(W_n)]] \quad \forall n \in \mathbb{N}.$$

Note that, as $n \to \infty$, $W_n \to \infty$ a.s. and $F_\theta(W_n) \to 1$ a.s. By bounded convergence theorem, $\mathbb{E}_0[F_\theta(W_n)] \to 1$ and the sequence $(U_n^{(\bar{C})}(q,c))_{n\geq 1}$ converges to $U^{(\bar{C})}(q,c)$ uniformly in $(q,c) \in \mathbb{R}_+ \times [0,\bar{C}]$. And we got the inequality

$$-\frac{1}{\alpha}[1 - \mathbb{E}_0[F_\theta(W_n)]] \leq U^{(\bar{C})}(q,c) - U_n^{(\bar{C})}(q,c) \leq 0.$$

□

Thus, if we can construct the sequence of functions $U_n^{(\bar{C})}(q,c)$, we can find the value function $U^{(\bar{C})}(q,c)$ itself. The next section explains the sequential approximation of the value function.

3.1 Defining Successive Approximations

Let us define operators

$$(H_t^{(\bar{C})}v)(q,c) \equiv \mathbb{E}_0^{(q,c)}\Bigg[\int_0^{t\wedge W_1} \overbrace{\Big(\underbrace{x(q,c,s)}_{Q_s} - \frac{1}{\alpha}f_\theta(\underbrace{c+s}_{C_s})\Big)}^{g(Q_s,C_s)}ds \; + 1_{\{t\geq W_1\}}v(Q_{W_1}, C_{W_1})\Bigg]$$

$$= \mathbb{E}_0^{(q,c)}\left[\int_0^{t\wedge W_1} g(Q_s, C_s)ds + 1_{\{t\geq W_1\}} v(Q_{W_1}, C_{W_1})\right],$$

for $t \in [0, \bar{C} - c]$ and

$$(H^{(\bar{C})}v)(q,c) \equiv \inf_{t\in[0,\bar{C}-c]} (H_t^{(\bar{C})}v)(q,c)$$

where v is a Borel function $v : \mathbb{R}_+ \times [0, \bar{C}] \to \mathbb{R}_-$ for $\bar{C} \in \mathbb{R}_+, (q,c) \in \mathbb{R}_+ \times [0, \bar{C}]$. We can write $(H_t^{(\bar{C})}v)(q,c)$ as

$$\begin{aligned}
(H_t^{(\bar{C})}v)(q,c) &= \mathbb{E}_0^{(q,c)}\left[\int_0^t 1_{\{s<W_1\}}\Big(Q_s - \frac{1}{\alpha} f_\theta(C_s)\Big)ds + 1_{\{t\geq W_1\}} v(Q_{W_1}, C_{W_1})\right] \\
&= \mathbb{E}_0^{(q,c)}\Bigg[\int_0^t 1_{\{s<W_1\}}\Big(x(q,c,s) - \frac{1}{\alpha} f_\theta(c+s)\Big)ds + \\
&\qquad 1_{\{t\geq W_1\}} v\Big(x(q,c,W_1)\frac{\lambda_1}{\lambda_0}h(Y_1), c+W_1\Big)\Bigg] \\
&= \int_0^t \mathbb{E}_0^{(q,c)}\left[1_{\{s<W_1\}}\Big(x(q,c,s) - \frac{1}{\alpha} f_\theta(c+s)\Big)\right]ds + \\
&\qquad \mathbb{E}_0^{(q,c)}\left[1_{\{t\geq W_1\}} v\Big(x(q,c,W_1)\frac{\lambda_1}{\lambda_0}h(Y_1), c+W_1\Big)\right] \\
&= \int_0^t \mathbb{P}_0\{s < W_1\}\Big(x(q,c,s) - \frac{1}{\alpha} f_\theta(c+s)\Big)ds + \\
&\qquad \mathbb{E}_0^{(q,c)}\left[1_{\{t\geq W_1\}} v\Big(x(q,c,W_1)\frac{\lambda_1}{\lambda_0}h(Y_1), c+W_1\Big)\right] \\
&= \int_0^t e^{-\lambda_0 s}\Big(x(q,c,s) - \frac{1}{\alpha} f_\theta(c+s)\Big)ds + \\
&\qquad \int_0^t \lambda_0 e^{-\lambda_0 s}\left[\int_{y\in\mathcal{Y}} v\Big(x(q,c,s)\frac{\lambda_1}{\lambda_0}h(y), c+s\Big)\nu_0(dy)\right]ds \\
&= \int_0^t e^{-\lambda_0 s}\Bigg(g\Big(x(q,c,s), c+s\Big) + \lambda_0 (Sv)\Big(x(q,c,s), c+s\Big)\Bigg)ds.
\end{aligned}$$

Where S is the linear operator

$$(Sw)(q,c) \equiv \int_{y\in\mathcal{Y}} w\Big(q\frac{\lambda_1}{\lambda_0}h(y), c\Big)\nu_0(dy)$$

for the function $w : \mathbb{R}_+ \times [0,\bar{C}] \to \mathbb{R}_-$. Moreover, for an ε optimal stopping time $\tau_\varepsilon^{(n-1)}$ we can write

$$
\begin{aligned}
U_n^{(\bar{C})}(q,c) &= \inf_{\tau \in \mathcal{F}_{[0,\bar{C}]}} \mathbb{E}_0^{(q,c)} \left[\int_0^{t \wedge W_n} \underbrace{(Q_s - \frac{1}{\alpha} f_\theta(C_s))}_{g(Q_s,C_s)} ds \right] \\
&\le \mathbb{E}_0^{(q,c)} \left[\int_0^{\tau_\varepsilon^{(n-1)} \wedge W_1} g(Q_s,C_s)ds + 1_{\{\tau_\varepsilon^{(n-1)} > W_1\}} \int_{W_1}^{\tau_\varepsilon^{(n-1)} \wedge W_n} g(Q_s,C_s)ds \right] \\
&= \mathbb{E}_0^{(q,c)} \left[\int_0^{\tau_\varepsilon^{(n-1)} \wedge W_1} g(Q_s,C_s)ds + \right. \\
&\qquad \left. 1_{\{\tau_\varepsilon^{(n-1)} > W_1\}} \int_{W_1}^{(\tau_\varepsilon^{(n-1)} \wedge W_{n-1}) \circ \Theta_{W_1} + W_1} g(Q_s,C_s)ds \right] \\
&= \mathbb{E}_0^{(q,c)} \left[\int_0^{\tau_\varepsilon^{(n-1)} \wedge W_1} g(Q_s,C_s)ds + \right. \\
&\qquad \left. 1_{\{\tau_\varepsilon^{(n-1)} > W_1\}} \int_0^{(\tau_\varepsilon^{(n-1)} \wedge W_{n-1}) \circ \Theta_{W_1}} g(Q_{s+W_1},C_{s+W_1})ds \right] \\
&= \mathbb{E}_0^{(q,c)} \left[\mathbb{E}_0^{(q,c)} \Big(\int_0^{\tau_\varepsilon^{(n-1)} \wedge W_1} g(Q_s,C_s)ds + \right. \\
&\qquad \left. 1_{\{\tau_\varepsilon^{(n-1)} > W_1\}} \Big(\int_0^{\tau_\varepsilon^{(n-1)} \wedge W_{n-1}} g(Q_s,C_s)ds \Big) \circ \Theta_{W_1} \Big| \mathcal{F}_{W_1} \Big) \right] \\
&= \mathbb{E}_0^{(q,c)} \left[\int_0^{\tau_\varepsilon^{(n-1)} \wedge W_1} g(Q_s,C_s)ds + \right. \\
&\qquad \left. 1_{\{\tau_\varepsilon^{(n-1)} > W_1\}} \underbrace{\mathbb{E}_0^{(Q_{W_1},C_{W_1})} \Big(\int_0^{(\tau_\varepsilon^{(n-1)} \wedge W_{n-1})} g(Q_s,C_s)ds \Big)}_{\le U_{n-1}^{(\bar{C})}(Q_{W_1},C_{W_1}) + \varepsilon} \right] \\
&\le \mathbb{E}_0^{(q,c)} \left[\int_0^{\tau_\varepsilon^{(n-1)} \wedge W_1} g(Q_s,C_s)ds + 1_{\{\tau_\varepsilon^{(n-1)} > W_1\}} U_{n-1}^{(\bar{C})} \Big(Q_{W_1}, C_{W_1} \Big) + \varepsilon \right] \\
&\le \mathbb{E}_0^{(q,c)} \left[\int_0^{\tau_\varepsilon^{(n-1)}} 1_{\{s < W_1\}} g(x(q,c,s), c+s)ds + \right. \\
&\qquad \left. 1_{\{\tau_\varepsilon^{(n-1)} > W_1\}} U_{n-1}^{(\bar{C})} \Big(x(q,c,W_1) \frac{\lambda_1}{\lambda_0} h(Y_1), c + W_1 \Big) \right] + \varepsilon
\end{aligned}
$$

$$= \int_0^{\tau_\varepsilon^{(n-1)}} e^{-\lambda_0 s} g(x(q,c,s),c+s)ds +$$
$$\int_0^{\tau_\varepsilon^{(n-1)}} \lambda_0 e^{-\lambda_0 s}\left[\int_{y\in\mathcal{Y}} U_{n-1}^{(\bar{C})}\Big(x(q,c,s)\frac{\lambda_1}{\lambda_0}h(y), c+s\Big)\nu_0(dy)\right]ds + \varepsilon$$
$$= \int_0^{\tau_\varepsilon^{(n-1)}} e^{-\lambda_0 s}\Big[g\Big(x(q,c,s),c+s\Big) + \lambda_0(SU_{n-1}^{(\bar{C})})\Big(x(q,c,s),c+s\Big)\Big]ds + \varepsilon.$$

The fifth equality is due to strong Markov property. Observe that $(H_t^{(\bar{C})}v)(q,c)$ and $U_n^{(\bar{C})}(q,c)$ are of the same form. Using this idea we will now define our successive approximations. Assume that

$$v_0^{(\bar{C})} \equiv 0$$
$$v_n^{(\bar{C})} \equiv (H^{(\bar{C})}v_{n-1}^{(\bar{C})})(q,c).$$

Proposition 2. *For every $n \in \mathbb{N}$, the functions $v_n^{(\bar{C})}$ and $U_n^{(\bar{C})}$ coincide. For every $\varepsilon \geq 0$, let*

$$r_n^{\varepsilon,\bar{C}} \equiv \inf\{s \in [0,\bar{C}] : H_s^{(\bar{C})}v_n^{(\bar{C})}(q,c) \leq H^{(\bar{C})}v_n^{(\bar{C})}(q,c) + \varepsilon\}$$

for $n \in \mathbb{N}_0, \bar{C} \in \mathbb{R}_+, (q,c) \in \mathbb{R}_+ \times [0,\bar{C}]$, and

$$S_1^{\varepsilon,\bar{C}} \equiv r_0^{\varepsilon,\bar{C}}(q,c) \wedge W_1$$

$$S_{n+1}^{\varepsilon,\bar{C}} \equiv \begin{cases} r_n^{\varepsilon/2,\bar{C}}(q,c), & \text{if } W_1 > r_n^{\varepsilon/2,\bar{C}}(q,c) \\ W_1 + S_n^{\varepsilon/2,\bar{C}} \circ \Theta_{W_1}, & \text{if } W_1 \leq r_n^{\varepsilon/2,\bar{C}}(q,c). \end{cases}, n \in \mathbb{N}$$

Then,

$$\mathbb{E}_0^{(q,c)}\Big[\int_0^{S_n^{\varepsilon,\bar{C}}} g(Q_t,C_t)dt\Big] \leq v_n^{(\bar{C})}(q,c) + \varepsilon \quad \forall n \in \mathbb{N}, \bar{C} \in \mathbb{R}_+, \forall \varepsilon \geq 0.$$

Proof. The proof uses Davis [16] and Brémaud's [51] characterization of $\mathcal{F}$-stopping times and is provided in Appendix A.4. □

3.2 Implementation of the Method

Note that we can write $(H_t^{(\bar{C})}v)(q,c)$ as

$$
\begin{aligned}
(H_t^{(\bar{C})}v)(q,c) = & \underbrace{\int_0^t e^{-\lambda_0 s} g(x(q,c,s), c+s) ds}_{\equiv I(t,q,c)} + \\
& \int_0^t \lambda_0 e^{-\lambda_0 s} \left[\int_{y \in \mathcal{Y}} v\Big(x(q,c,s) \frac{\lambda_1}{\lambda_0} h(y), c+s \Big) d\nu_0 \right] ds \\
= & I(t,q,c) + \int_0^t \lambda_0 e^{-\lambda_0 s} \left[\int_{y \in \mathcal{Y}} v\Big(x(q,c,s) \frac{\lambda_1}{\lambda_0} h(y), c+s \Big) d\nu_0 \right] ds \\
= & I(t,q,c) + \int_0^t \lambda_0 e^{-\lambda_0 s} (Sv)\Big(x(q,c,s), c+s \Big) ds \\
= & I(t,q,c) + \int_0^\infty \lambda_0 e^{-\lambda_0 s} 1_{\{s \le t\}} (Sv)\Big(x(q,c,s), c+s \Big) ds,
\end{aligned}
$$

where

$$
\begin{aligned}
I(t,q,c) & \equiv \int_0^t e^{-\lambda_0 s} \Big(x(q,c,s) - \frac{1}{\alpha} f_\theta(c+s) \Big) ds \\
& = \begin{cases}
q\frac{(1-e^{-\lambda_1 t})}{\lambda_1} + \left(\frac{1}{\lambda_1} - \frac{1}{\alpha}\right) \int_0^t e^{-\lambda_0 u} f_\theta(c+u)\ du - \frac{e^{-\lambda_1 t}}{\lambda_1} \int_0^t e^{(\lambda_1 - \lambda_0)u} f_\theta(c+u)\ du, \\
\text{if } \lambda_0 \neq \lambda_1, \\
\\
q\frac{(1-e^{-\lambda_0 t})}{\lambda_0} + \left(\frac{1}{\lambda_0} - \frac{1}{\alpha}\right) \int_0^t e^{-\lambda_0 u} f_\theta(c+u)\ du - \frac{e^{-\lambda_0 t}}{\lambda_0} \int_0^t f_\theta(c+u)\ du, \\
\text{if } \lambda_0 = \lambda_1.
\end{cases}
\end{aligned}
$$

Moreover we can write the integral term as

$$
\begin{aligned}
& \int_0^\infty \lambda_0 e^{-\lambda_0 s} 1_{\{s \le t\}} (Sv)\Big(x(q,c,s), c+s \Big) ds \\
& = \mathbb{E}\left[1_{\{E \le t\}} (Sv)\Big(x(q,c,E), c+E \Big) \right] \\
& = \mathbb{E}\left[(Sv)\Big(x(q,c,E), c+E \Big) \Big| E \le t \right] \mathbb{P}\{E \le t\} \\
& = \mathbb{E}\left[(Sv)\Big(x(q,c,E_t), c+E_t \Big) \right] \mathbb{P}\{E \le t\},
\end{aligned}
$$

where E denotes an exponentially distributed random variable with rate λ_0 and E_t is the truncated exponentially distributed random variable with rate λ_0. To use this method, we first create a rectangular grid on the (Q, C)-space. Since the iterations start from 0, each grid point is assigned to $v_0^{\bar{C}}(q, c) = 0$, initially. Later, the jump operator $(H^{(\bar{C})}v)(q, c) = \inf_{t\in[0,\bar{C}-c]}(H_t^{(\bar{C})}v)(q, c)$ is applied to each grid point iteratively, until the convergence is reached, where

$$(H_t^{(\bar{C})}v)(q, c) = I(t, q, c) + \int_0^\infty \lambda_0 e^{-\lambda_0 s} 1_{\{s\leq t\}}(Sv)\Big(x(q, c, s), c + s\Big)ds.$$

The key point here is that Q-grid must be large enough to contain the whole continuation region. The calculation of the function $I(t, q, c)$ is relatively simple, for the integral term, different approximation methods are available. Ürü et al. [24] used this method with Chebyshev grid points and calculated required expectations using Monte Carlo methods for a uniform prior. However, for complex densities, this method requires large number of grid points and iterations for the convergence, which reduces the performance. For this reason, we come up with a relatively simpler and practical method to solve the problem, this method will be explained in detail in Chapter 4.

Chapter 4

Discretization

For this method, we will discretize the time space to solve a discrete sub-problem of the original/truncated problem. To that end, let us define a grid on the interval $[0, \bar{C}]$ as $\Delta_{\vec{c}}^{(\bar{C})} = \{c_0, c_1, \ldots, c_M\}$, $M \in \mathbb{N}$. c_i is a strictly increasing sequence with $c_0 = 0$ and $c_M = \bar{C}$. If we denote the distance between the consecutive grid points as $\delta_i := c_{i+1} - c_i$ for $0 \leq i < M$, we can define the mesh of this partition as $\delta_{\vec{c}}^{(\bar{C})} = \max_{0 \leq i < M} \delta_i$.

If the clock process starts from c_i, one can define a value function

$$U^{(\bar{C}, \Delta_{\vec{c}}^{(\bar{C})})}(q, c_i) := \inf_{\tau \in \{0, c_{i+1} - c_i, \ldots, \bar{C} - c_i\}} \mathbb{E}_0^{(q, c_i)} \int_0^\tau \big(Q_t - \frac{1}{\alpha} f_\theta(C_t)\big) dt, \tag{4.1}$$

for $q \in \mathbb{R}_+$, $c_i \in \Delta_{\vec{c}}^{(\bar{C})}$ where the optimal stopping time is selected from the discrete set $\{0, c_{i+1} - c_i, \ldots, \bar{C} - c_i\}$. In other words, the process is stopped at one of the points in $\Delta_{\vec{c}}^{(\bar{C})}$. The following lemma shows how to control the discretization error.

Lemma 4. *We have*

$$0 \leq U^{(\bar{C}, \Delta_{\vec{c}}^{(\bar{C})})}(q, c_i) - U^{(\bar{C})}(q, c_i) \leq \delta_{\vec{c}}^{(\bar{C})} \big(q + F_\theta(\bar{C}) - F_\theta(c_i)\big) \tag{4.2}$$

for $q \in \mathbb{R}_+$ and $c_i \in \Delta_{\vec{c}}^{(\bar{C})}$.

Proof. Since the optimal stopping time for the discretized problem is a sub-optimal stopping time for the truncated problem, the first inequality is obvious. For the second problem let us define an ϵ-optimal stopping time for the truncated problem as τ_ϵ, for $\epsilon > 0$ we have

$$\mathbb{E}_0^{(q,c_i)}\left[\int_0^{\tau_\epsilon}\big(Q_t - \frac{1}{\alpha}f_\theta(C_t)\big)dt\right] \leq U^{(\bar{C})}(q,c_i) + \epsilon \quad \text{for} \quad q \in \mathbb{R}, c_i \in \Delta_{\vec{c}}^{(\bar{C})}. \tag{4.3}$$

Let us define another stopping time $\tilde{\tau}_\epsilon$ as

$$\tilde{\tau}_\epsilon \equiv \sum_{j=i}^{M-1} 1_{\{c_j - c_i \leq \tau_\epsilon < c_{j+1} - c_i\}}(c_{j+1} - c_i) + 1_{\{\tau_\epsilon = \bar{C} - c_i\}}(\bar{C} - c_i).$$

Observe that $\tilde{\tau}_\epsilon \in [\tau_\epsilon, \tau_{\bar{C}}]$, and the distance between τ_ϵ and $\tilde{\tau}_\epsilon$ is less than $\delta_{\vec{c}}^{(\bar{C})}$. Using (4.3), we can write

$$\begin{aligned} U^{(\bar{C})}(q,c_i) &\geq \mathbb{E}_0^{(q,c_i)}\left[\int_0^{\tau_\epsilon}\big(Q_t - \frac{1}{\alpha}f_\theta(C_t)\big)dt\right] - \epsilon \qquad (4.4)\\ &= \mathbb{E}_0^{(q,c_i)}\left[\int_0^{\tilde{\tau}_\epsilon}\big(Q_t - \frac{1}{\alpha}f_\theta(C_t)\big)dt\right] - \mathbb{E}_0^{(q,c_i)}\left[\int_{\tau_\epsilon}^{\tilde{\tau}_\epsilon}\big(Q_t - \frac{1}{\alpha}f_\theta(C_t)\big)dt\right] - \epsilon\\ &\geq U^{(\bar{C},\Delta_{\vec{c}}^{(\bar{C})})}(q,c_i) - \mathbb{E}_0^{(q,c_i)}\left[\int_{\tau_\epsilon}^{\tilde{\tau}_\epsilon} Q_t dt\right] - \epsilon. \end{aligned}$$

The last inequality is due to the fact that $\tilde{\tau}_\epsilon \in \{0, c_{i+1} - c_i, \ldots, \bar{C} - c_i\}$.

Using Remark 2, we can write

$$\begin{aligned} &\mathbb{E}_0^{(q,c_i)}\left[\int_{\tau_\epsilon}^{\tilde{\tau}_\epsilon} Q_t\ dt\right] \leq \mathbb{E}_0^{(q,c_i)}\left[Q_{\tau_{\bar{C}}}\int_{\tau_\epsilon}^{\tilde{\tau}_\epsilon}\ dt\right] \leq \mathbb{E}_0^{(q,c_i)}\left[Q_{\tau_{\bar{C}}}\int_{\tau_\epsilon}^{\tau_\epsilon + \delta_{\vec{c}}^{(\bar{C})}}\ dt\right]\\ &= \delta_{\vec{c}}^{(\bar{C})}\mathbb{E}_0^{(q,c_i)}\left[Q_{\tau_{\bar{C}}}\right] = \delta_{\vec{c}}^{(\bar{C})}\left(q + F_\theta(\bar{C}) - F_\theta(c_i)\right). \end{aligned}$$

Then, we can write

$$\begin{aligned} U^{(\bar{C})}(q,c_i) &\geq U^{(\bar{C},\Delta_{\vec{c}}^{(\bar{C})})}(q,c_i) - \mathbb{E}_0^{(q,c_i)}\left[\int_{\tau_\epsilon}^{\tilde{\tau}_\epsilon} Q_t dt\right] - \epsilon\\ &\geq U^{(\bar{C},\Delta_{\vec{c}}^{(\bar{C})})}(q,c_i) - \delta_{\vec{c}}^{(\bar{C})}\left(q + F_\theta(\bar{C}) - F_\theta(c_i)\right) - \epsilon. \end{aligned}$$

By the arbitrariness of ϵ, the second inequality holds. □

The next corollary states how to control the error caused by truncation and discretization.

Corollary 3. *It follows from Lemmas 1 and 4 that, for a given $\epsilon > 0$, if we set $\bar{C}$ and $\Delta_{\vec{c}}^{(\bar{C})}$ so that*

$$\delta_{\vec{c}}^{(\bar{C})}\big(\frac{\pi}{1-\pi} + F_\theta(\bar{C})\big) + \frac{1}{\alpha}[1 - F_\theta(\bar{C})] \leq \epsilon,$$

then $\tau^{(\bar{C},\Delta_{\vec{c}}^{(\bar{C})})}(\frac{\pi}{1-\pi}, 0)$ is an ϵ-optimal stopping time for the detection problem in (2.4).

Using similar arguments with the original and the truncated problem, we observe that for fixed $c_i \in \Delta_{\vec{c}}^{(\bar{C})}$, $q \mapsto U^{(\bar{C},\Delta_{\vec{c}}^{(\bar{C})})}(q, c_i)$ is non-decreasing and concave. Hence, we have the (discrete) boundary curve

$$\beta^{(\bar{C},\Delta_{\vec{c}}^{(\bar{C})})}(c_i) := \min\{q : U^{(\bar{C},\Delta_{\vec{c}}^{(\bar{C})})}(q, c_i) = 0\}, \qquad c_i \in \Delta_{\vec{c}}^{(\bar{C})},$$

at and above which the function $U^{(\bar{C},\Delta_{\vec{c}}^{(\bar{C})})}$ is zero, and below, the function is negative. It follows from standard arguments of the finite horizon discrete-time optimal stopping theory for Markov processes that the function $U^{(\bar{C},\Delta_{\vec{c}}^{(\bar{C})})}$ satisfies the dynamic programming recursions

$$U^{(\bar{C},\Delta_{\vec{c}}^{(\bar{C})})}(q, c_i) = \min\Big\{0\,,\, \mathbb{E}_0^{(q,c_i)}\Big[\int_0^{\delta_i} \big(Q_t - \frac{1}{\alpha} f_\theta(C_t)\big)dt + U^{(\bar{C},\Delta_{\vec{c}}^{(\bar{C})})}(Q_{\delta_i}, c_{i+1})\Big]\Big\} \tag{4.5}$$

for $i \in \{0, 1, \ldots, M-1\}$, with $U^{(\bar{C},\Delta_{\vec{c}}^{(\bar{C})})}(q, \bar{C}) = 0$, and the stopping time at which the process enters to the stopping region $\tau^{(\bar{C},\Delta_{\vec{c}}^{(\bar{C})})}(q, c_i) := \min\{t \in \Delta_{\vec{c}}^{(\bar{C})} : Q_t \geq \beta^{(\bar{C},\Delta_{\vec{c}}^{(\bar{C})})}(C_t)\}$ is an optimal stopping time for $U^{(\bar{C},\Delta_{\vec{c}}^{(\bar{C})})}$.

Recall that, in Remark 1 we have shown that, for the original problem with unbounded densities, we have unbounded boundary curves. We will now investigate if that is the case for the discretized problem. For our calculations, using Fubini's theorem and the expectation in (2.5), we have

$$\begin{aligned} &\mathbb{E}_0^{(q,c)}\Big[\int_0^{\delta} \big(Q_t - \frac{1}{\alpha} f_\theta(C_t)\big)dt\Big] = q\delta + I(c, \delta) \qquad \text{where} \\ &I(c, \delta) := \int_0^{\delta} \big[F_\theta(c+t) - F_\theta(c)\big]\; dt - \frac{1}{\alpha}[F_\theta(c+\delta) - F_\theta(c)]. \end{aligned} \tag{4.6}$$

Observe that for c_{M-1}, we have

$$\begin{aligned}U^{(\bar{C},\Delta_{\vec{c}}^{(\bar{C})})}(q,c_{M-1}) &= \min\Big\{0\,,\ \mathbb{E}_0^{(q,c_{M-1})}\int_0^{\delta_{M-1}}\Big(Q_t-\frac{1}{\alpha}f_\theta(C_t)\Big)dt\Big\}\\ &= \min\Big\{0\,,\ q\delta_{M-1}+I(c_{M-1},\delta_{M-1})\Big\}.\end{aligned}$$

The first equality is due to $U^{(\bar{C},\Delta_{\vec{c}}^{(\bar{C})})}(\cdot,\bar{C})=0$. When $I(c_{M-1},\delta_{M-1})\geq 0$, we need immediate stopping for any q value, thus $\beta^{(\bar{C},\Delta_{\vec{c}}^{(\bar{C})})}(c_{M-1})=0$, in that case. When $I(c_{M-1},\delta_{M-1})<0$, we have $\beta^{(\bar{C},\Delta_{\vec{c}}^{(\bar{C})})}(c_{M-1})=-\frac{1}{\delta_{M-1}}I(c_{M-1},\delta_{M-1})$.

Note that we have $I(c,\delta)\geq-\frac{1}{\alpha}[F_\theta(c+\delta)-F_\theta(c)]$ for $c\geq 0$ and $\delta\geq 0$. Thus, for $I(c_{M-1},\delta_{M-1})<0$, we can write

$$\begin{aligned}\beta^{(\bar{C},\Delta_{\vec{c}}^{(\bar{C})})}(c_{M-1}) = -\frac{1}{\delta_{M-1}}I(c_{M-1},\delta_{M-1}) &\leq \frac{F_\theta(c_{M-1}+\delta_{M-1})-F_\theta(c_{M-1})}{\alpha\delta_{M-1}} \\ &= \frac{F_\theta(\bar{C})-F_\theta(c_{M-1})}{\alpha\delta_{M-1}}.\end{aligned}\tag{4.7}$$

For other points in $\Delta_{\vec{c}}^{(\bar{C})}$, recall that the function $U^{(\bar{C},\Delta_{\vec{c}}^{(\bar{C})})}$ is concave in q. Hence, by Jensen's inequality we can write

$$\begin{aligned}\mathbb{E}_0^{(q,c_i)}\Big[U^{(\bar{C},\Delta_{\vec{c}}^{(\bar{C})})}(Q_{\delta_i},c_{i+1})\Big] &\leq U^{(\bar{C},\Delta_{\vec{c}}^{(\bar{C})})}(\mathbb{E}_0^{(q,c_i)}[Q_{\delta_i}],c_{i+1})\\ &= U^{(\bar{C},\Delta_{\vec{c}}^{(\bar{C})})}(q+F_\theta(c_{i+1})-F_\theta(c_i),c_{i+1}),\end{aligned}$$

and using (4.5)

$$\begin{aligned}U^{(\bar{C},\Delta_{\vec{c}}^{(\bar{C})})}(q,c_i) &= \min\Big\{0\,,\ q\delta_i+I(c_i,\delta_i)+\mathbb{E}_0^{(q,c_i)}\Big[U^{(\bar{C},\Delta_{\vec{c}}^{(\bar{C})})}(Q_{\delta_i},c_{i+1})\Big]\Big\}\\ &\leq \min\Big\{0\,,\ q\delta_i+I(c_i,\delta_i)+U^{(\bar{C},\Delta_{\vec{c}}^{(\bar{C})})}(q+F_\theta(c_{i+1})-F_\theta(c_i),c_{i+1})\Big\}.\end{aligned}$$

When $I(c_i,\delta_i)+U^{(\bar{C},\Delta_{\vec{c}}^{(\bar{C})})}(q+F_\theta(c_{i+1})-F_\theta(c_i),c_{i+1})\geq 0$, we need immediate stopping for any q value, thus $\beta^{(\bar{C},\Delta_{\vec{c}}^{(\bar{C})})}(c_i)=0$, in that case. When $I(c_i,\delta_i)+U^{(\bar{C},\Delta_{\vec{c}}^{(\bar{C})})}(q+F_\theta(c_{i+1})-F_\theta(c_i),c_{i+1})<0$, using the fact that $U^{(\bar{C},\Delta_{\vec{c}}^{(\bar{C})})}(\cdot,c_{i+1})\geq -\frac{1}{\alpha}[F_\theta(\bar{C})-F_\theta(c_{i+1})]$ and $I(c,\delta)\geq-\frac{1}{\alpha}[F_\theta(c+\delta)-F_\theta(c)]$ for $c\geq 0$ and $\delta\geq 0$, we can write

$$\begin{aligned}\beta^{(\bar{C},\Delta_{\vec{c}}^{(\bar{C})})}(c_i) &\leq -\frac{1}{\delta_i}\Big[I(c_i,\delta_i)+U^{(\bar{C},\Delta_{\vec{c}}^{(\bar{C})})}(q+F_\theta(c_{i+1})-F_\theta(c_i),c_{i+1})\Big]\\ &\leq \frac{F_\theta(\bar{C})-F_\theta(c_i)}{\alpha\delta_i}.\end{aligned}\tag{4.8}$$

Corollary 4. *Recall from Remark 1 that for the original problem, the boundary that separates stopping and the continuation region can be unbounded, for certain densities. In the time-discretization method with finite number of points* $0 = c_0 < c_1 < \ldots < c_M = \bar{C}$*, the boundary* $\beta^{(\bar{C},\Delta_{\vec{c}}^{(\bar{C})})}$ *over the grid points is always finite (and therefore uniformly bounded from above); see (4.7,4.8), and the boundary at the terminal point* $c_M = \bar{C}$ *is clearly zero.*

4.1 Implementation of the Method

Recall that we are expecting the dynamic programming recursions

$$U^{(\bar{C},\Delta_{\vec{c}}^{(\bar{C})})}(q,c_i) = \min\Big\{0\,,\,\mathbb{E}_0^{(q,c_i)}\Big[\int_0^{\delta_i}\big(Q_t - \frac{1}{\alpha}f_\theta(C_t)\big)dt + U^{(\bar{C},\Delta_{\vec{c}}^{(\bar{C})})}(Q_{\delta_i},c_{i+1})\Big]\Big\}$$

for $i \in \{0,1,\ldots,M-1\}$, with $U^{(\bar{C},\Delta_{\vec{c}}^{(\bar{C})})}(q,\bar{C}) = 0$, to be satisfied. To implement the method, we start by deciding a δ value for the time-discretization of the C-space. δ denotes the mesh of the grid $\Delta_{\vec{c}}^{\bar{C}} = \{c_0, c_1, ..., c_M\}$ with $c_0 = 0$ and $c_M = \bar{C}$. We then put a grid on the Q-space, as well. The key point here is that, the maximum value for this grid should be large enough that it includes the whole continuation region. In other words, $U^{(\bar{C},\Delta_{\vec{c}}^{(\bar{C})})}(q,c_i)$ must be 0 for some q on this grid. For any q value larger than this, the function will be also 0. For the first iteration, since we have $U^{(\bar{C},\Delta_{\vec{c}}^{(\bar{C})})}(q,\bar{C}) = 0$, the first inequality to be satisfied is

$$\begin{aligned}U^{(\bar{C},\Delta_{\vec{c}}^{(\bar{C})})}(q,c_{M-1}) &= \min\Big\{0\,,\,\mathbb{E}_0^{(q,c_{M-1})}\Big[\int_0^{\delta_i}\big(Q_t - \frac{1}{\alpha}f_\theta(C_t)\big)dt\Big]\Big\}\\ &= \min\Big\{0\,,\,q\delta + I(c_{M-1},\delta)\Big\}.\end{aligned}$$

For the ongoing iterations we need the expectation

$$\mathbb{E}_0^{(q,c_i)}\Big[U^{(\bar{C},\Delta_{\vec{c}}^{(\bar{C})})}(Q_{\delta_i},c_{i+1})\Big].$$

This term can be approximated when δ is small using infinitesimal probabilities of the Poisson process as

$$\mathbb{E}_0^{(q,c_i)}\Big[U^{(\bar{C},\Delta_{\vec{c}}^{(\bar{C})})}(Q_{\delta_i},c_{i+1})\Big] \approx (1-\lambda_0\delta)\cdot U^{(\bar{C},\Delta_{\vec{c}}^{(\bar{C})})}(x(q,c_i,\delta),c_{i+1}) +$$

$$
\begin{aligned}
&\lambda_0\delta \cdot \mathbb{E}_0^{(q,c_i)}\Big[U^{(\bar{C},\Delta_{\vec{c}}^{(\bar{C})})}(\frac{\lambda_1}{\lambda_0}h(Y_1)x(q,c_i,\delta),c_{i+1})\Big] \\
&= (1-\lambda_0\delta)\cdot U^{(\bar{C},\Delta_{\vec{c}}^{(\bar{C})})}(x(q,c_i,\delta),c_{i+1}) + \\
&\lambda_0\delta \cdot \int_{\mathbb{R}^d} U^{(\bar{C},\Delta_{\vec{c}}^{(\bar{C})})}(\frac{\lambda_1}{\lambda_0}h(y)x(q,c_i,\delta),c_{i+1})\ \nu_0(dy).
\end{aligned}
$$

Observe that for the calculation of $U^{(\bar{C},\Delta_{\vec{c}}^{(\bar{C})})}(.,c_i)$ values, we require $U^{(\bar{C},\Delta_{\vec{c}}^{(\bar{C})})}(.,c_{i+1})$ values, which we have already calculated. For this purpose, we use interpolation on the Q-grid of the $U^{(\bar{C},\Delta_{\vec{c}}^{(\bar{C})})}(.,c_{i+1})$ values. On the other hand, for the integral with respect to the ν_0 measure, we use standard approximation methods for the integration. As a final remark, since we don't have a monotone boundary for every prior density, we should check at each step that Q-grid contains the whole continuation region. For each c_i value, the smallest q value that satisfies $U^{(\bar{C},\Delta_{\vec{c}}^{(\bar{C})})}(q,c_i)=0$ is chosen as the boundary for that c_i value. The next section explains how to use this boundary to raise an alarm for a sample path.

4.2 Alarm Rule

For this example, the prior density

$$
f_\theta(x) = \frac{3}{2}\,1_{[0,\frac{1}{3}]\cup[\frac{2}{3},1]}(x)
$$

was used. Other problem parameters are $\lambda_0 = 40$, $\lambda_1 = 50$, $\alpha = 5$, $\nu_0 = \mathcal{N}(0,1)$ and $\nu_1 = \mathcal{N}(0.2,1)$.

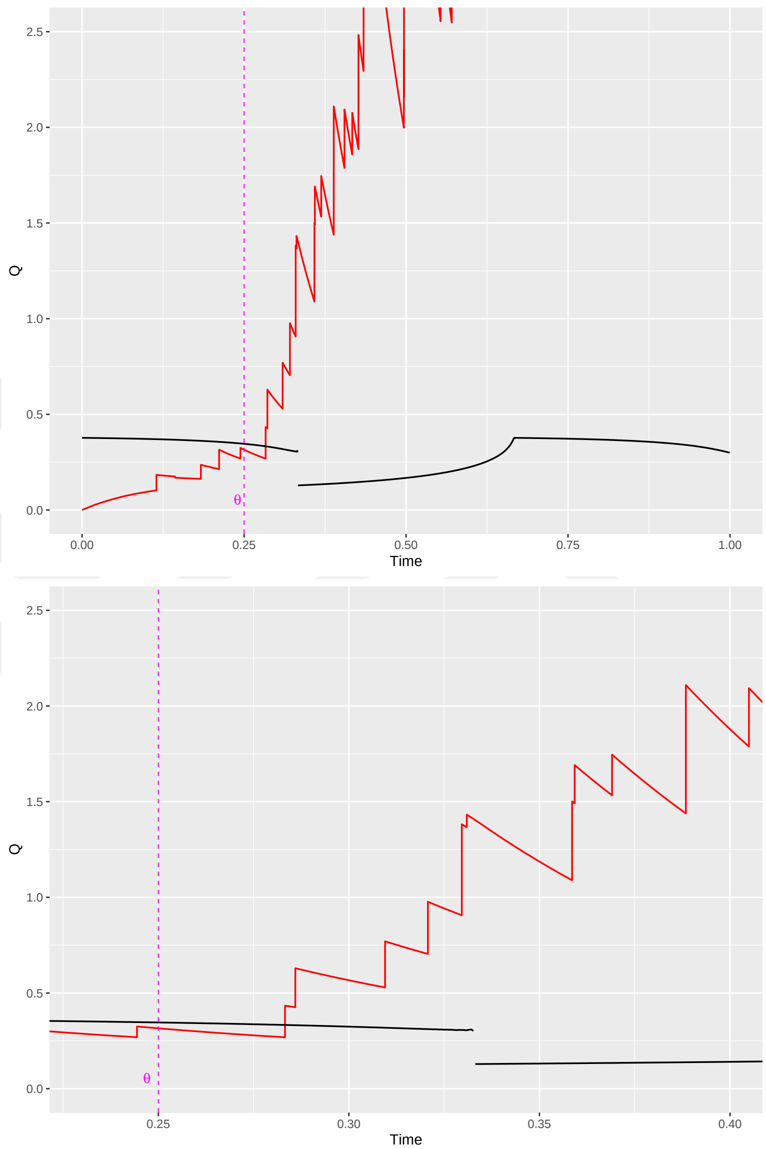


Figure 4.1: Sample path for the disorder problem with the stopping boundary, $\alpha = 5$

Figure 4.1 depicts the sample path of the process Q_t (red) and the boundary curve (black). The disorder time is $\theta = 0.25$, is shown in pink. The boundary curve separates the space into the continuation and the stopping region. Below this curve, the value function is less than zero, and above this curve, it is zero. Thus, when the process Q passes above this curve, we should raise an alarm and stop the system. Observe that, the boundary changes over time. This means that, our threshold, in other words willingness to stop the process changes, as well. In this case, we tend to continue on the interval $[\frac{1}{3}, \frac{2}{3}]$, since there is no new information coming on this interval. This specific disorder density will be discussed in detail later. After the disorder, the arrival frequency and the mean of the marks increase in this problem, resulting Q to grow rapidly and excessively. For this sample path, the user stops the system at the first arrival after the disorder.

For the second example, α value is taken to be 1, instead of 5, for the same sample path. This means that we are penalizing an expected delay with a cost that is five times lower. One intuitive result is that we expect the system to be stopped later than the first case.

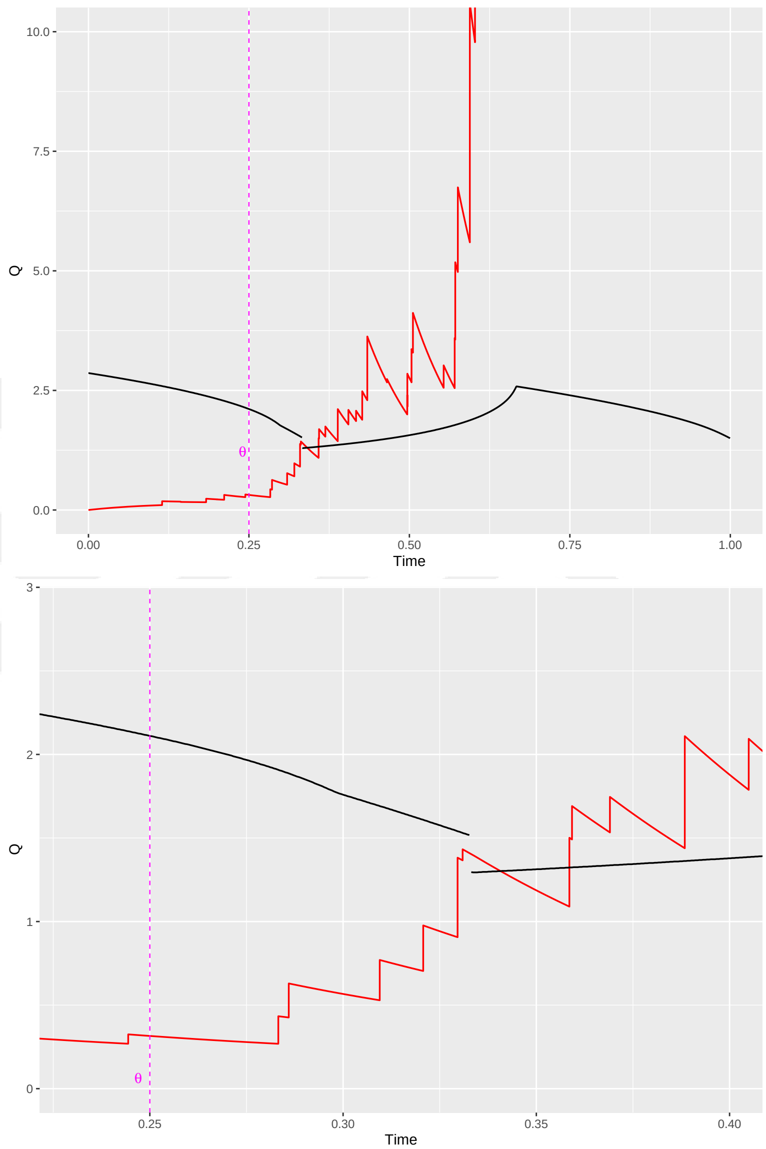


Figure 4.2: Sample path for the disorder problem with the stopping boundary, $\alpha = 1$

This is indeed observed: in the first case, the system is stopped at the first arrival after the disorder, while in the second case, it is stopped at the sixth arrival after the disorder. The reason is, waiting for more information is not as costly as the first case. In the next chapter, we will investigate different prior densities and their various properties.

Chapter 5

Numerical Examples

In this chapter, to investigate boundary behavior of different prior densities, we will provide results for six numeric examples. The first example is a mixed uniform prior where the decision maker knows that no disorder will occur within certain sub-intervals. For the second example, to explore the relationship between the variance of the disorder time and the Bayesian risk, Gamma prior is visited. The third and the fourth problem deals with a phase-type and an asymptotic prior, respectively. For the fifth problem, a mixed uniform approximation for the exponential prior is considered. Lastly, a credit card fraud detection problem is visited, this problem is motivated by the work of Buonaguidi et al. [17]. For the solution of these problems, time discretization method was adopted.

5.1 Mixed Uniform Problem

On a time interval, when the decision maker has a belief that the disorder will certainly not happen on some sub-intervals, but can happen on anywhere else with equal probability, they can use a mixed uniform prior for the disorder. The purely uniform case was studied by Ürü et al. [24]. For this problem, disorder densities

$$f_{\theta,1}(x) = \frac{3}{2}\,1_{[0,\frac{1}{3}]\cup[\frac{2}{3},1]}(x) \qquad \text{and} \qquad f_{\theta,2}(x) = \frac{3}{4}\,1_{[0,\frac{1}{3}]\cup[\frac{2}{3},1]\cup[\frac{4}{3},\frac{5}{3}]\cup[2,\frac{7}{3}]}(x).$$

were used. Other parameters are $\lambda_0 = 5, \lambda_1 = 1, \alpha = 0.2, \nu_0 = \mathcal{N}(0,1)$ and $\nu_1 = \mathcal{N}(0.2,1)$.

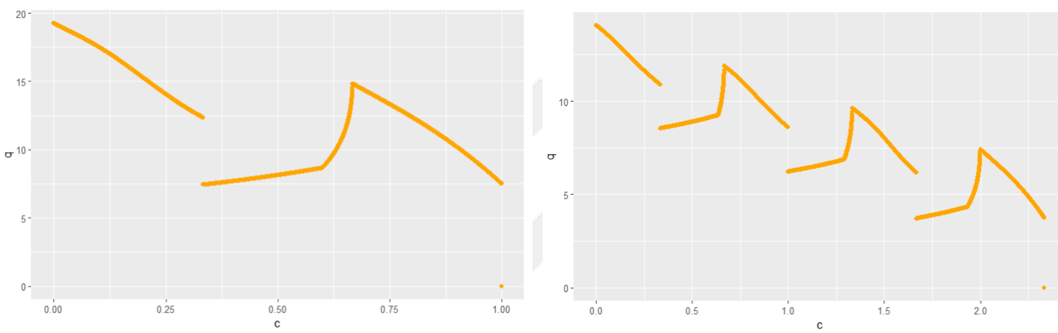


Figure 5.1: Stopping boundaries for two different mixed uniform priors

On the support of the density, the boundary curves show a monotone decreasing behavior, this result is indeed expected from the previous studies. However, where the probability is zero, the opposite behavior is observed. The intuition here is that, during these intervals, decision maker does not receive any additional information since they know that disorder will not happen here. On the other hand, we observed that the Bayesian risk for the second problem is higher. This is again intuitive as the detection becomes more risky when the number of intervals with zero probability increase.

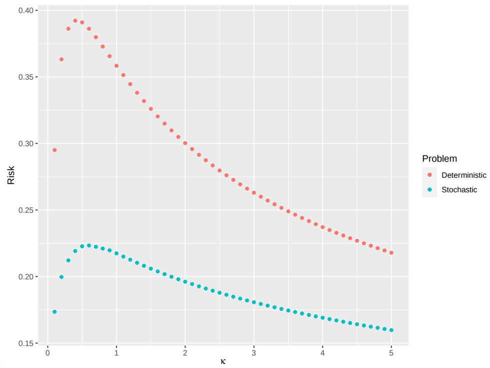
5.2 Gamma Problem

This problem was motivated by the question of whether the detection risk is monotonically related to the variance of the disorder time. We worked on fifty different Gamma priors with constant mean 1 and changing variance. The shape parameter α and the rate parameter β were taken as the same value (denoted by

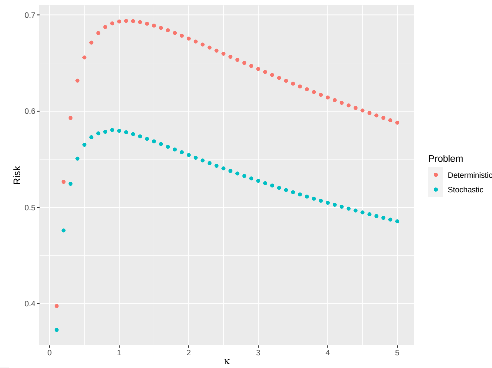
κ) on the interval 0.1, 0.2, . . . , 5. In this case, the density becomes

$$f_\theta(x) = \begin{cases} \frac{\kappa^\kappa}{\Gamma(\kappa)} x^{\kappa-1} e^{-\kappa x}, & \text{if } x \in [0, \infty), \\ 0, & \text{otherwise.} \end{cases}$$

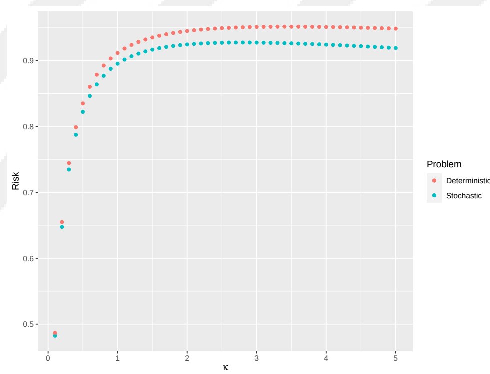
Observe that the mean of the disorder time is $\mathbb{E}[\theta] = 1$ for each of these cases, while the variance is equal to $Var(\theta) = \frac{1}{\kappa}$. The problem was revisited with three different α values 0.2, 1 and 5. Resulting κ vs. Bayes risk plots are shown for each problem below for stochastic and deterministic problems. By deterministic problem, we mean that the stopping times are selected from $\mathcal{F}_0$-measurable deterministic times. Naturally, the risk is higher than the stochastic optimization problem. For all of the Gamma problems, parameters $\lambda_0 = 5$, $\lambda_1 = 1$, $\nu_0 = \mathcal{N}(0, 1)$, $\nu_1 = \mathcal{N}(0.2, 1)$ and $\bar{C}$ was chosen for each problem such that $\mathbb{P}\{\theta > \bar{C}\} < 10^{-4}$.



(a) $\alpha = 0.2$



(b) $\alpha = 1$



(c) $\alpha = 5$

Figure 5.2: Risk for the stochastic and the deterministic problems for three different α values

Intuitively, one would expect the detection to be less risky when the variance decreases, as κ increases, in other words. However, we do not observe a monotone relationship between variance and risk in this case. For κ values less than 1, detection is less risky due to the asymptotic behavior of the gamma density around 0 for these κ values. As κ increases, the shape of the density changes, making detection more risky. However, for higher values of κ, the density concentrates around the mean, leading to a decrease in risk. Moreover, the difference between the risks of deterministic and stochastic optimization problems decreases as α increases. This is also intuitive, as the cost of expected delay increases, waiting for new information becomes more costly for the decision maker.

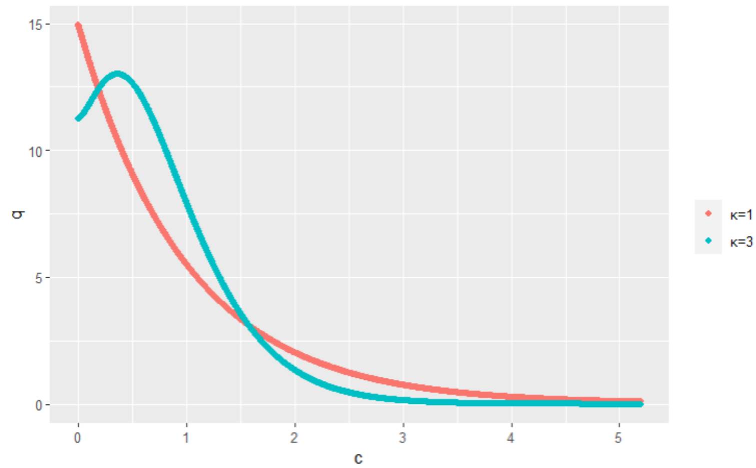


Figure 5.3: Stopping boundaries for two different Gamma priors

Finally, Figure 5.3, depicts the stopping boundaries for the Gamma problem with two different κ values. Shapes of these curves are coherent with their density plots.

5.3 Phase-Type Problem

Phase-type densities for the compound Poisson disorder problem were studied by Bayraktar and Sezer [18]. A phase-type density represents the distribution of time spent in a Markov chain, before absorption. They are useful for approximating other densities but suffer from the high dimensionality problem. We revisited the problem using our time discretization method. Figure 5.4 displays the stopping boundary for the problem, which has a similar shape with the gamma problem above.

The generator matrix $\mathcal{A}$ was used (shown below) for this problem. Moreover, initial state probabilities are $\vec{\pi} = [0.25, 0.55, 0.1, 0.1, 0]$, and problem parameters are $\pi = 0$, $\lambda_0 = 2$, $\lambda_1 = 1$, $\nu_0 = \mathcal{N}(0, 1)$, $\nu_1 = \mathcal{N}(0.2, 1)$, $\alpha = 1$ and $\bar{C} = 20$.

$$\mathcal{A} = \begin{bmatrix} -1 & 0.6 & 0.25 & 0.1 & 0.05 \\ 0.15 & -2 & 1 & 0.5 & 0.35 \\ 0.35 & 1 & -2 & 0.15 & 0.5 \\ 0.1 & 1 & 1 & -3 & 0.9 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

If we denote

$$S = \begin{bmatrix} -1 & 0.6 & 0.25 & 0.1 \\ 0.15 & -2 & 1 & 0.5 \\ 0.35 & 1 & -2 & 0.15 \\ 0.1 & 1 & 1 & -3 \end{bmatrix} \quad \text{and} \quad s = \begin{bmatrix} 0.05 \\ 0.35 \\ 0.5 \\ 0.9 \end{bmatrix},$$

then

$$f_\theta(x) = \vec{\pi} \cdot \exp(xS) \cdot s,$$
$$F_\theta(x) = 1 - \vec{\pi} \cdot \exp(xS) \cdot \vec{1}, \quad x \in [0, \infty).$$

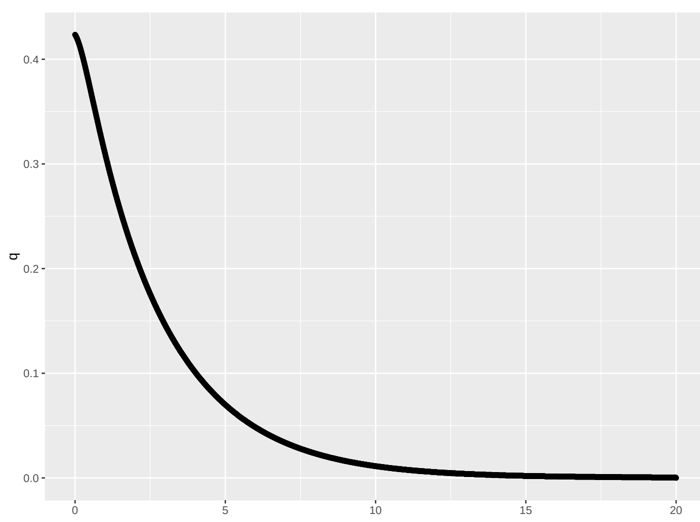


Figure 5.4: Stopping boundary for the phase-type prior

5.4 An Asymptotic Problem

For this problem, the density

$$f_\theta(x) = \begin{cases} \frac{1}{4\sqrt{x}}, & \text{if } x \in [0,1], \\ \frac{1}{4\sqrt{x-1}}, & \text{if } x \in [1,2], \\ 0, & \text{otherwise}, \end{cases}$$

was used. Observe that $f(0+) = f(1+) = \infty$. The stopping boundary for this problem is unbounded, however, as mentioned in Corollary 4. Using the time discretization method, we can find such upper bound for our discrete boundary.

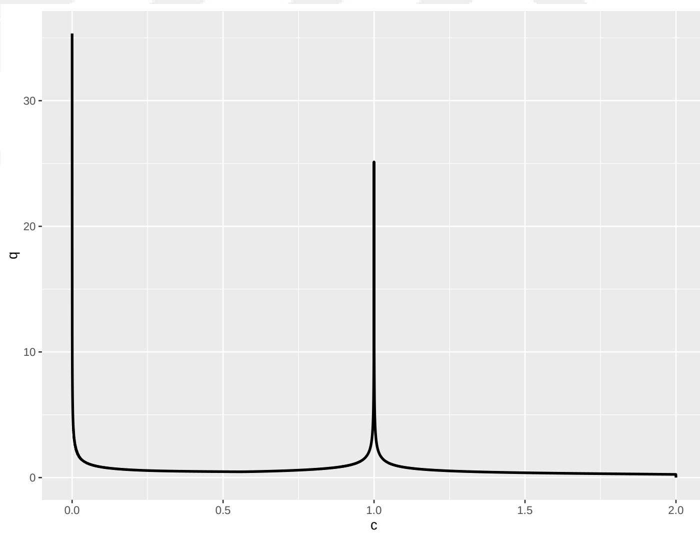


Figure 5.5: Stopping boundary for the asymptotic prior

Due to the asymptotic behavior of the density at points 0 and 1, the boundary curve tends to halt the process at these points.

5.5 Mixed Uniform Distribution for Approximation

For this problem, exponential density

$$f_\theta(x) = \begin{cases} e^{-x}, & \text{if } x \in [0, \infty), \\ 0, & \text{otherwise.} \end{cases}$$

was visited using two different methods. Firstly, exponential density itself, and secondly, a mixed uniform approximation of the density.

The problem parameters are $\pi = 0$, $\lambda_0 = 0.1$, $\lambda_1 = 0.2$, $\nu_0 = \mathcal{N}(0,1)$, $\nu_1 = \mathcal{N}(0.2,1)$, $\alpha = 1$ and $\bar{C} = 20$. For the mixed uniform prior, the interval $[0, 20]$ is divided into 500 sub-intervals $\{[s_0, s_1], ..., [s_{499}, s_{500}]\}$ with equal lengths, 0.04. Moreover, $p_j = \int_{s_{j-1}}^{s_j} e^{-x}\, dx$ for $j = 1, ..., 500$. At each interval $[s_{j-1}, s_j]$, the density takes the value $p_j/(s_j - s_{j-1})$.

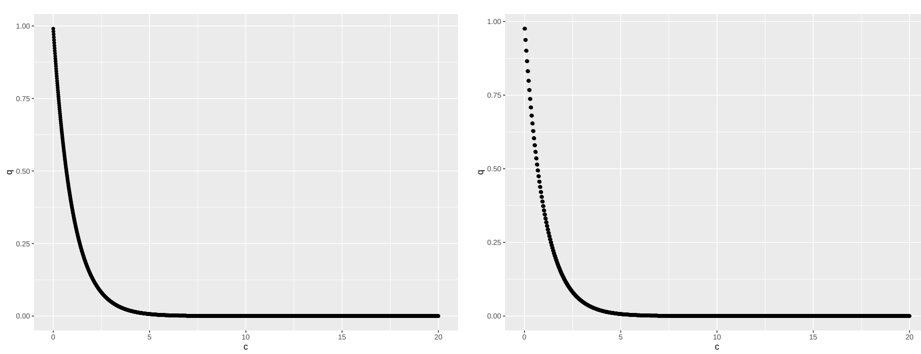


Figure 5.6: Stopping boundaries for the exponential problem

Figure 5.6 displays the stopping boundaries for both cases which look quite similar, as expected. Moreover, the results were satisfying with $U_{approx}(0,0) = -0.3089$ and $U(0,0) = -0.3088$, hence $V_{approx}(0) = 0.6911$ and $V(0) = 0.6912$.

5.6 Fraud Problem

For this case, Buonaguidi et al.'s study [17] was revisited. For the fraud time, instead of an exponential prior, we worked on a special density where we consider two functions

$$h_1(t) = \begin{cases} 1 & \text{for } t \in (0,1), \\ 0 & \text{elsewhere,} \end{cases} \qquad h_2(t) = \begin{cases} 288(t - \frac{11}{12}) & \text{for } t \in (\frac{11}{12}, 1), \\ 0 & \text{elsewhere,} \end{cases}$$

and for a given parameter $\beta \in (0,1)$,

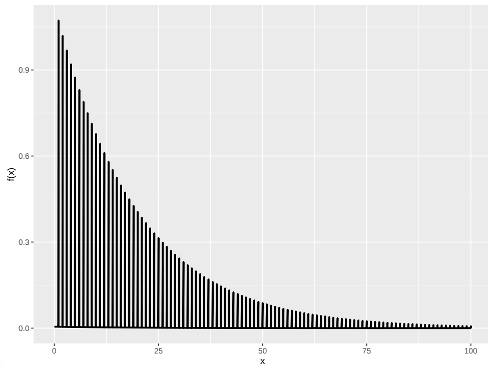
$$h(t) = \beta h_1(t) + (1-\beta) h_2(t).$$

Here, a year is assumed to have a length one. Function h, gives a weighted average of the functions h_1 and h_2, where h_2 denotes the increasing fraud activity at the end of the year, last month, in particular. Increasing shopping pattern due to New Year's Eve or Christmas was motivated this idea.

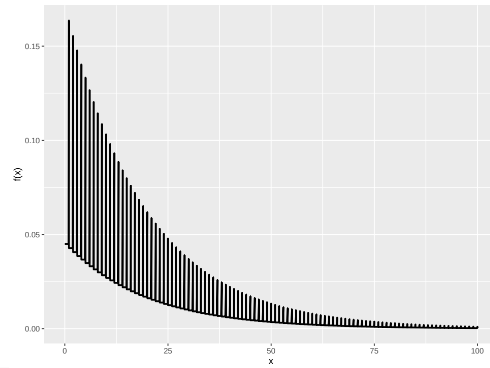
Additionally, let us assume that there is a fraud activity with probability $p \in (0,1)$ on a given year. If there is no fraud activity that year, it will be in the following year independently with probability p again. In other words, the first year where there will be a fraud activity is a geometrically distributed random variable with parameter p. For its corresponding continuous version, where we aim to detect the disorder on a continuous interval, considering the function h, if we denote the number of years elapsed before t, as $I_t := \lfloor t \rfloor$, we obtain the prior density

$$f_\theta(t) = (1-p)^{I_t} \, p \, h(t - \lfloor t \rfloor), \quad t \geq 0,$$

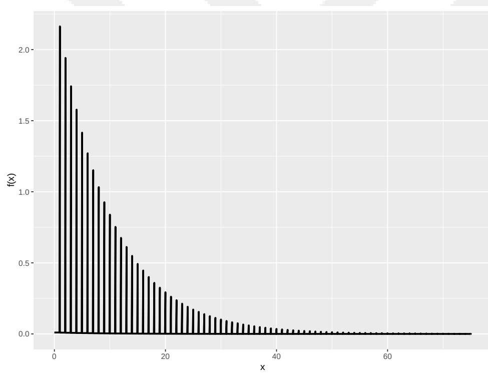
for the time of fraud activity. For illustration, density plots for some (β, p) pairs are given below.



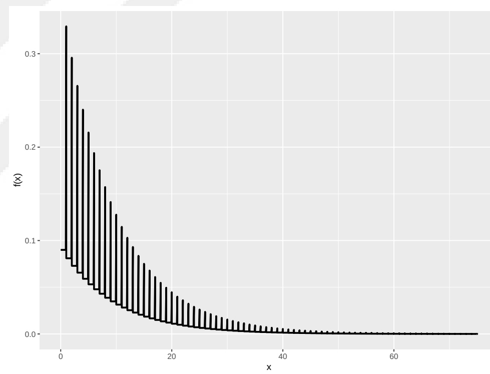
(a) $\beta = 0.1$ and $p = 0.05$



(b) $\beta = 0.9$ and $p = 0.05$



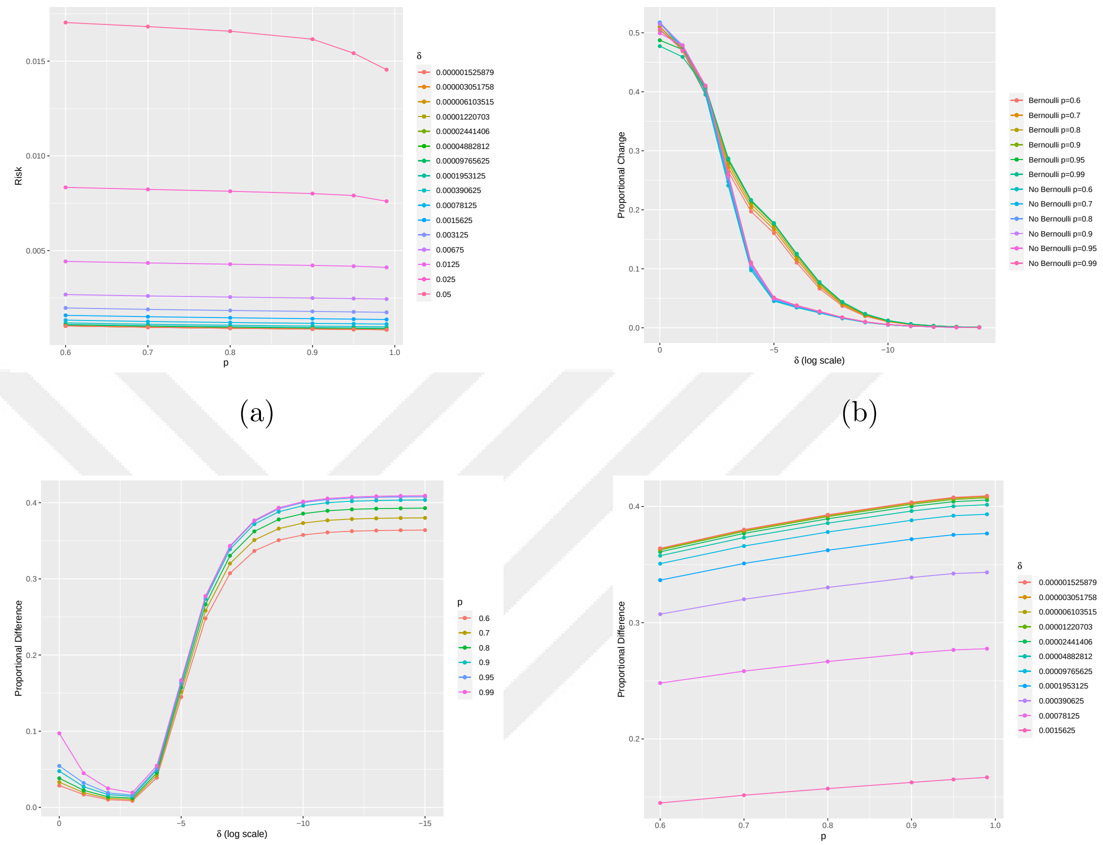
(c) $\beta = 0.1$ and $p = 0.1$



(d) $\beta = 0.9$ and $p = 0.1$

Figure 5.7: Density plots for the fraud problem

For this problem, parameters were obtained from [17], instead of a daily setup, we have worked on an annual setup. $\lambda_0 = 492.75$ $(365 \cdot 1.35)$, $\lambda_1 = 1098.65$ $(365 \cdot 3.01)$, $\nu_0 = \mathcal{N}(2.5, 1.5)$, $\nu_1 = \mathcal{N}(4.1, 3.12)$. Additionally, for an alternative setup, we allowed Poisson arrivals to also have independent Bernoulli(p) distributed marks where $p_0 = 0.1$ and $p_1 = 0.9$. In this context, normally distributed marks can correspond to logarithm of spending amounts, while Bernoulli marks can depict several features, such as whether the spending is online, whether it is within a geographic region, etc. $\bar{C}$ was chosen for each problem such that $\mathbb{P}\{\theta > \bar{C}\} < 5 \cdot 10^{-6}$. α was assumed to be 0.2. For fixed $\beta = 0.5$, p values of $0.6, 0.7, 0.8, 0.9, 0.95, 0.99$ were investigated.



(a) (b) (c) (d)

Figure 5.8: Plots on the fraud problem

Figure 5.8a shows how the Bayes risk changes with δ and p. Observe that, for this problem, as p decreases, the truncated problem's $\bar{C}$ value increases. An increasing time interval increases the Bayesian risk, which the plot also confirms. Figure 5.8b depicts the proportional change of risk with decreasing δ. As it is expected, the risk converges as δ decreases, and the convergence is independent of the p value. One key observation is that, during discretization, we consider that with a probability of $(1 - \lambda_0\delta)$, no jump will occur, and with a probability of $\lambda_0\delta$, a jump will occur over an interval of length δ. However, we do not consider scenarios where two or more jumps occur. Since we have a quite high λ_1 value in our problem, δ should be selected carefully for the numerical methods. Figure 5.8c displays the effect of δ on the proportional difference between problems with and without Bernoulli marks. This plot also supports the argument about the choice of δ, as after the fourth δ value, a monotone increasing and then stabilizing behavior is observed. Figure 5.8d indicates that the difference between two problems increases, as p increases.This result is also intuitive because as the time interval of the problem decreases, the value of the information increases.

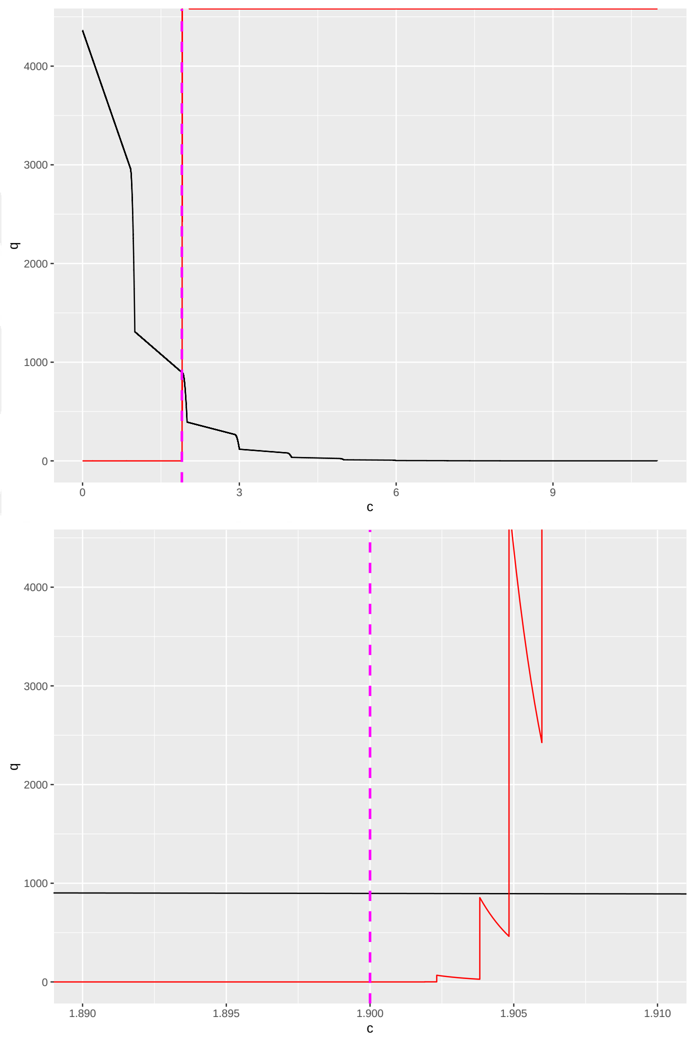


Figure 5.9: A sample path example of the fraud problem

Figure 5.9 depicts a sample path for the Fraud problem. The pink dashed line indicates the disorder time θ. Here, $\beta = 0.5$, $p = 0.7$, $p_0 = 0.1$ and $p_1 = 0.9$. Observe that the system rings an alarm at the third arrival after the disorder. Especially the drastic change in λ and Bernoulli parameter p after the disorder amplifies the post-disorder behavior of the process Q.

Part II

Misspecified Wiener Disorder Problem

In Chapter 6, the Wiener misspecification problem and the required modeling elements will be explained, and the value function for the optimal stopping problem will be constructed. In Chapter 7, the original problem with the true parameters will be discussed, and a solution method for the calculation of the value function will be presented. In Chapter 8, the misspecified problem will be investigated in detail, and in Chapter 9, the solution method for this problem using simulations will be discussed. Finally, in Chapter 10, a numerical example dealing with a real-life scenario will be presented. Long proofs and some other mathematical results will be presented in Appendix B.

Chapter 6

Model and Problem Description

In this chapter, we construct the mathematical model for our asset selling problem, express the value function for the corresponding optimal stopping problem, and later introduce the misspecified problem with incorrect parameters. *Note that, for the second part of the thesis, each notation will be defined separately from the first part. Do not assume any notation in the context of the first study.*

6.1 Model

On a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ hosting a Wiener process W and an independent non-negative random variable θ, we consider a simple market model with a risk-free asset and a risky asset whose market values R and S respectively evolve with the dynamics

$$\frac{dR_t}{R_t} = rdt \quad \text{and} \quad \frac{dS_t}{S_t} = [\mu 1_{\{t<\theta\}} + \nu 1_{\{\theta \le t\}}]dt + \sigma dW_t,$$

with $\sigma > 0$, $r \in \mathbb{R}_+$ and $\nu < r < \mu$. Here, θ denotes the random unknown and unobservable time where the drift of the underlying model changes from μ to ν. θ has the prior density

$$\mathbb{P}\{\theta = 0\} = \pi \quad \mathbb{P}\{\theta > t\} = (1-\pi)e^{-\lambda t} \tag{6.1}$$

with $\lambda > 0$ and $\pi \in [0, 1)$.

If we define the process X as

$$X_t := \mu t + (\nu - \mu)(t - \theta)^+ + \sigma W_t,$$

we can write

$$\frac{dS_t}{S_t} = dX_t.$$

With $R_0 = 1$, we have $R_t = e^{rt}$, and starting from some level S_0, using Itô's lemma, the value of the risky asset can be written as

$$S_t = S_0 \exp\{X_t - \frac{1}{2}\sigma^2 t\}, \quad t \geq 0,$$

and its discounted value is

$$P_t = e^{-rt} S_t = S_0 \exp\{X_t - rt - \frac{1}{2}\sigma^2 t\}.$$

We consider the problem of finding the optimal stopping time to sell a risky asset with the objective of maximizing the expected earning (or linear utility)

$$\sup_{\tau} \mathbb{E}[1_{\{\tau < \infty\}} P_\tau]. \tag{6.2}$$

Let us assume that a probability measure $\mathbb{P}_\infty$ is locally related with the probability measure $\mathbb{P}$ with

$$L_t = \frac{d\mathbb{P}}{d\mathbb{P}_\infty}\bigg|_{\mathcal{F}_t \vee \sigma(\theta)} = 1_{\{t < \theta\}} + \frac{\xi_t}{\xi_\theta} 1_{\{\theta \leq t\}}, \quad t \geq 0,$$

in terms of the process

$$\xi_t = e^{(\frac{\nu - \mu}{\sigma^2}) X_t - \frac{1}{2}(\frac{\nu^2 - \mu^2}{\sigma^2}) t}, \quad t \geq 0.$$

Observe that this new probability measure has the property that change never occurs (or occurs at time ∞), and X and θ are independent stochastic components. Hence, under this measure, X is a Brownian motion with drift μ, and θ has the same zero-modified exponential distribution given in (6.1).

6.2 Problem with Correct Parameters

The problem with correct parameters aims to find an a.s. finite stopping time that maximizes the expected discounted stock price. In this section, we will reformulate the original problem with an optimal stopping problem, using a Markovian process.

Let us now introduce the odds-ratio process

$$\Phi_t := \frac{\mathbb{E}[1_{\{\theta\le t\}}|\mathcal{F}_t]}{\mathbb{E}[1_{\{\theta> t\}}|\mathcal{F}_t]} = e^{\lambda t}\xi_t\left[\frac{\pi}{1-\pi}+\int_0^t \lambda e^{-\lambda u}\frac{1}{\xi_u}\,du\right],$$

with the dynamics

$$d\Phi_t = \lambda(1+\Phi_t)dt + \beta\Phi_t\frac{1}{\sigma}(dX_t-\mu dt) \quad \text{where} \quad \beta = \frac{\nu-\mu}{\sigma}.$$

If we also define a new probability measure $\mathbb{P}_Q$ such that

$$\frac{d\mathbb{P}_Q}{d\mathbb{P}_\infty}\Big|_{\mathcal{F}_t\vee\sigma(\theta)} = \exp\{X_t-\mu t-\frac{1}{2}\sigma^2 t\}, \quad t\ge 0,$$

using Itô's rule, Fubini's theorem and independence of θ and X under $\mathbb{P}_Q$, as shown in Appendix, for $\tau<\infty$, we have

$$\mathbb{E}[P_\tau] = S_0\left[1-(1-\pi)(r-\nu)\mathbb{E}_Q\int_0^\tau e^{-\gamma u}\left[\Phi_u-k\right]\,du\right].$$

So if we define

$$V(\phi) = \inf_\tau \mathbb{E}_Q^\phi\int_0^\tau e^{-\gamma t}\left[\Phi_t-k\right]\,dt, \tag{6.3}$$

where $\phi=\Phi_0$, then

$$\sup_\tau \mathbb{E}[1_{\{\tau<\infty\}}P_\tau] = S_0\left[1-(1-\pi)(r-\nu)V(\frac{\pi}{1-\pi})\right]. \tag{6.4}$$

Observe that in our case, $\Phi_0=\frac{\pi}{1-\pi}$.

Remark 3. *Observe that $\phi\to V(\phi)$, is an increasing concave function of ϕ, which is bounded above by 0. Hence, there exists $\phi^*>0$ such that the problem is separated into a continuation region*

$$\mathcal{C} = \{\phi>0:\phi<\phi^* \quad \textit{and} \quad V(\phi)<0\},$$

and a stopping region

$$\mathcal{S} = \{\phi>0:\phi\ge\phi^* \quad \textit{and} \quad V(\phi)=0\}.$$

6.3 Problem with Incorrect Parameters

Let $\tilde{\pi}$, $\tilde{\lambda}$, $\tilde{\mu}$ and $\tilde{\nu}$ denote the choices by a decision maker for the parameters π, λ, μ and ν respectively. Assume that at least one of their choices is incorrect. In this case, the decision maker will solve an another optimization problem using these parameters, thus processes. Throughout, the tilde symbol ˜ will be also used for the expressions calculated by the decision maker using incorrect parameters. For instance $\tilde{\beta} := \frac{\tilde{\nu}-\tilde{\mu}}{\sigma}$. Moreover, processes ξ and Φ used by the decision maker while solving the problem turn into

$$\tilde{\xi}_t = e^{(\frac{\tilde{\nu}-\tilde{\mu}}{\sigma^2})X_t - \frac{1}{2}(\frac{\tilde{\nu}^2-\tilde{\mu}^2}{\sigma^2})t} \quad \text{and} \quad \tilde{\Phi}_t = e^{\tilde{\lambda}t}\tilde{\xi}_t\left[\frac{\tilde{\pi}}{1-\tilde{\pi}} + \int_0^t \tilde{\lambda}e^{-\tilde{\lambda}u}\frac{1}{\tilde{\xi}_u}\, du\right],$$

with the dynamics

$$\begin{aligned} d\tilde{\xi}_t &= \tilde{\beta}\tilde{\xi}_t\frac{1}{\sigma}(dX_t - \tilde{\mu}dt) \\ &= \tilde{\beta}\tilde{\xi}_t\frac{(\mu-\tilde{\mu})}{\sigma}dt + \tilde{\beta}\tilde{\xi}_t\frac{1}{\sigma}(dX_t - \mu dt) \end{aligned}$$

and

$$\begin{aligned} d\tilde{\Phi}_t &= \tilde{\lambda}(1+\tilde{\Phi}_t)dt + \tilde{\beta}\tilde{\Phi}_t\frac{1}{\sigma}(dX_t - \tilde{\mu}dt) \\ &= \left(\tilde{\lambda} + \left[\tilde{\lambda} + \tilde{\beta}\frac{(\mu-\tilde{\mu})}{\sigma}\right]\tilde{\Phi}_t\right)dt + \tilde{\beta}\tilde{\Phi}_t\frac{1}{\sigma}(dX_t - \mu dt). \end{aligned}$$

As it will be explained in details in Chapter 4, the decision maker finds $\tilde{\phi}^*$ by solving the problem with the incorrect parameters, and uses the corresponding stopping time $\tilde{\tau}^*$ in calculations, while they think they found the optimal one that minimizes (6.3). Hence, what they get is

$$H(\phi, \tilde{\phi}) := \mathbb{E}_Q^{\phi}\int_0^{\tilde{\tau}^*} e^{-\gamma t}\left[\Phi_t - k\right]\, dt. \tag{6.5}$$

Thus in (6.4) they will use $H(\frac{\pi}{1-\pi}, \frac{\tilde{\pi}}{1-\tilde{\pi}})$, instead of $V(\frac{\pi}{1-\pi})$.

Next two chapters will explain the solution method for the problem with both correct and incorrect parameters.

Chapter 7

Problem with Correct Parameters

In this chapter, we will provide the solution of the original problem with correct parameters. To achieve this, we need to understand the behavior of the process Φ under the probability measure $\mathbb{P}_Q$, which turns out to be an Itô process. Subsequently, two solution methods will be presented.

7.1 The Dynamics of Φ under $\mathbb{P}_Q$

Observe that, under $\mathbb{P}_\infty$, the change never occurs. Moreover, we have

$$\frac{d\mathbb{P}_Q}{d\mathbb{P}_\infty}\Big|_{\mathcal{F}_t \vee \sigma(\theta)} = \exp\{X_t - \mu t - \frac{1}{2}\sigma^2 t\}, \quad t \geq 0.$$

Note that, since the change never occurs we can write $X_t - \mu t - \frac{1}{2}\sigma^2 t$ as

$$X_t - \mu t - \frac{1}{2}\sigma^2 t = \int_0^t \sigma \, dW_s^{(\infty)} - \frac{1}{2}\int_0^t \sigma^2 \, ds.$$

where $W^{(\infty)}$ is the Wiener process under the measure $\mathbb{P}_\infty$. By Girsanov's theorem, we have

$$W_t^{(Q)} = W_t^{(\infty)} - \int_0^t \sigma \, ds,$$

$$dW_t^{(Q)} = dW_t^{(\infty)} - \sigma dt.$$

We had stated that

$$d\Phi_t = \lambda(1+\Phi_t)dt + \beta\Phi_t\frac{1}{\sigma}(dX_t - \mu dt)$$

and under $\mathbb{P}_\infty$, $\frac{1}{\sigma}(X_t - \mu t)$ is a standard Brownian motion. Now,

$$d\Phi_t = \lambda(1+\Phi_t)dt + \beta\Phi_t\left[\frac{1}{\sigma}(dX_t - \mu dt) - \sigma dt\right] + \beta\Phi_t\sigma dt$$
$$d\Phi_t = \left[\lambda + (\lambda+\sigma\beta)\Phi_t\right]dt + \beta\Phi_t\left[\frac{1}{\sigma}(dX_t - \mu dt) - \sigma dt\right].$$

Using our previous result we know that under $\mathbb{P}_Q$, $Z_t = \left[\frac{1}{\sigma}(X_t - \mu t) - \sigma t\right]$ is a standard Brownian motion. So under this new measure, Φ is an Itô process with dynamics

$$d\Phi_t = \left[\lambda + (\lambda+\sigma\beta)\Phi_t\right]dt + \beta\Phi_t dZ_t.$$

7.2 Solution of the Original Problem

Remember that our value function was

$$V(\phi) := \inf_\tau \mathbb{E}\int_0^\tau e^{-\gamma t}(\Phi_t - k)\ dt,$$

where the underlying process has the dynamics

$$d\Phi_t = \left[\lambda + (\lambda+\sigma\beta)\Phi_t\right]dt + \beta\Phi_t dZ_t.$$

Now let us consider a general case where

$$d\Phi_t = \ell(\Phi_t)dt + \beta\Phi_t dZ_t$$

for some constant β and some linear function $\ell(\phi) = \ell_0 + \ell_1\phi$ with $\ell_0 > 0$. The other constants k and γ are positive and ϕ denotes the starting point of the process Φ. $\mathbb{E}$ is the expectation with respect to the some probability measure under which Z is a Wiener process.

Lemma 5. *For $0 < a < b$, a continuous function $f_{a,b}$ solving*

$$-\gamma f_{a,b}(\phi) + \ell(\phi) f'_{a,b}(\phi) + \frac{1}{2}\beta^2\sigma^2 f''_{a,b}(\phi) + \phi - k = 0 \quad \textit{for} \quad \phi \in (a,b) \tag{7.1}$$

and $f_{a,b} = 0$ for $\phi \notin (a,b)$, we have

$$f_{a,b}(\phi) = \mathbb{E}\int_0^{\tau_{a,b}} e^{-\gamma t}(\Phi_t - k)\ dt,$$

where $\tau_{a,b}$ is the first exit time of Φ from the interval (a,b). Moreover, if we define $f_b := \lim_{a\searrow 0} f_{a,b}(\phi)$ and extend it at the origin as $f_b(0) = f_b(0+)$, we have

$$f_b(\phi) = \mathbb{E}\int_0^{\tau_b} e^{-\gamma t}(\Phi_t - k)\ dt,$$

where τ_b is the first exit time of Φ from the interval $[0,b)$.

Proof. The proof is given in the Appendix B.4. □

Lemma 6. *With the choice of b for which $f'_b(b-) = 0$, f_b satisfies the variational inequalities*

$$\left.\begin{array}{l} f_b(\phi) < 0 \\ \frac{1}{2}\beta^2\phi^2 f''_b(\phi) + \ell(\phi) f'_b(\phi) - \gamma f_b(\phi) + (\phi - k) = 0 \end{array}\right\}, \quad \phi \in [0,b)$$

$$\left.\begin{array}{l} f_b(\phi) = 0 \\ \frac{1}{2}\beta^2\phi^2 f''_b(\phi) + \ell(\phi) f'_b(\phi) - \gamma f_b(\phi) + (\phi - k) > 0 \end{array}\right\}, \quad \phi \in [b,\infty)$$

Moreover, this b is the unique value which satisfies

$$\int_0^b \frac{2\psi(t)}{\beta^2 t^2 s_\Phi(t)}(t-k)\ dt = 0, \tag{7.2}$$

where ψ is the strictly increasing solution of the homogeneous differential equation in (7.1) and

$$s_\Phi(\phi) := \left(\frac{e^{\frac{\ell_0}{\phi}}}{\phi^{\ell_1}}\right)^{\frac{2}{\beta^2}}.$$

Proof. The proof is given in the Appendix B.5. □

Proposition 3. *If $f : \mathbb{R}_+ \to \mathbb{R}$ satisfies the inequalities*

$$\left.\begin{aligned} &f(\phi) \leq 0 \\ &\frac{1}{2}\beta^2\phi^2 f''(\phi) + \ell(\phi)f'(\phi) - \gamma f(\phi) + (\phi - k) \geq 0 \end{aligned}\right\}, \quad \phi \in [0, \infty)$$

then f minorizes the value function

$$\inf_{\tau} \mathbb{E} \int_0^{\tau} e^{-\gamma t} \left[\Phi_t - k\right] \, dt.$$

Proof. Using Itô's rule and above inequalities, for any stopping time T

$$\begin{aligned} f(\phi) &= -\mathbb{E} \int_0^T e^{-\gamma t} \left[\frac{1}{2}\beta^2\Phi_t^2 f''(\Phi_t) + \ell(\Phi_t)f'(\Phi_t) - \gamma f(\Phi_t)\right] \, dt + \mathbb{E}\left[e^{-\gamma T} f(\Phi_T)\right] \\ &\leq -\mathbb{E} \int_0^T e^{-\gamma t} \left[\frac{1}{2}\beta^2\Phi_t^2 f''(\Phi_t) + \ell(\Phi_t)f'(\Phi_t) - \gamma f(\Phi_t)\right] \, dt \\ &\leq \mathbb{E} \int_0^T e^{-\gamma t} \left[\Phi_t - k\right] \, dt. \end{aligned}$$

□

Corollary 5. *With the choice of b as given above,*

$$\mathbb{E} \int_0^{\tau_b} e^{-\gamma t} \left[\Phi_t - k\right] \, dt = \inf_{\tau} \mathbb{E} \int_0^{\tau} e^{-\gamma t} (\Phi_t - k) \, dt =: V(\phi).$$

Proof. This is a direct result of Lemma 6 and Proposition 3. □

Hence, one needs to solve the Equation (7.2) and find such b, then use it as the exit boundary for the process Φ for the calculation of the expectations. However, finding the function ψ is not an easy task for every differential equation. Dayanik et al. [35] used a sequential approximation method for the calculation of ψ.

7.3 Solution Using Markov Chain Approximation

To solve the original problem, we use the method of Kushner and Dupuis [49] for the Markov chain approximation of the process Φ. To approximate the process Φ, use the fact that we want to find the expectation

$$V_\gamma(\phi) = \inf_\tau \mathbb{E}^\phi \left[\int_0^\tau e^{-\gamma t}(\Phi_t - k)\, dt \right].$$

When $\gamma = 0$,

$$\frac{1}{2}\beta^2\phi^2 V_0''(\phi) + \ell(\phi)V_0'(\phi) + (\phi - k) = 0$$

should be satisfied for $\phi \in [0, \phi^*)$, which we can approximate by

$$\begin{aligned}
&\frac{1}{2}\beta^2\phi^2 \frac{V_0^h(\phi+h) + V_0^h(\phi-h) - 2V_0^h(\phi)}{h^2} + \ell^+(\phi)\frac{V_0^h(\phi+h) - V_0^h(\phi)}{h} \\
&- \ell^-(\phi)\frac{V_0^h(\phi) - V_0^h(\phi-h)}{h} + (\phi - k) = 0
\end{aligned}$$

for a small $h > 0$. If we arrange the terms, we get

$$\begin{aligned}
V_0^h(\phi) = {} & \frac{\frac{1}{2}\beta^2\phi^2 + h\ell^+(\phi)}{\beta^2\phi^2 + h|\ell(\phi)|} V_0^h(\phi+h) + \frac{\frac{1}{2}\beta^2\phi^2 + h\ell^-(\phi)}{\beta^2\phi^2 + h|\ell(\phi)|} V_0^h(\phi-h) + \\
& \frac{h^2}{\beta^2\phi^2 + h|\ell(\phi)|}(\phi - k).
\end{aligned}$$

If we define

$$\begin{aligned}
p^h(\xi^h, \xi^h \pm h) &= \frac{\frac{1}{2}\beta^2 {\xi^h}^2 + h\ell^\pm(\xi^h)}{\beta^2 {\xi^h}^2 + h|\ell(\xi^h)|}, \\
\Delta t^h(\xi^h) &= \frac{h^2}{\beta^2 {\xi^h}^2 + h|\ell(\xi^h)|},
\end{aligned}$$

we can approximate the corresponding process Φ with the Markov chain $\Xi^h_{t_n}$ starting from $\xi^h \in \{0, h, 2h, ...\}$ and uses the above transition probabilities and holding times. Observe that holding times are deterministic for each state.

If we define an auxiliary value function as

$$V^h(\xi^h) = \inf_{N^h} \mathbb{E}^{\xi^h} \left[\sum_{n=0}^{N^h - 1} (\Xi_{t_n} - k)\, e^{-\gamma t_n^h} \frac{1 - e^{-\gamma \Delta t_n^h}}{\gamma} \right],$$

we expect the dynamic programming equation

$$V^h(\xi^h) = \min\Big\{0, (\xi^h - k)\cdot\frac{1-e^{-\gamma\Delta t^h(\xi^h)}}{\gamma} + \\ e^{-\gamma\Delta t^h(\xi^h)}\big[p^h(\xi^h,\xi^h+h)V^h(\xi^h+h) + p^h(\xi^h,\xi^h-h)V^h(\xi^h-h)\big]\Big\} \tag{7.3}$$

to be satisfied. Now, for a Borel function $w : \mathbb{R}_+ \to [-k/\gamma, 0]$ we define the operator

$$(\mathcal{H}^h w)(\xi^h) = \min\Big\{0, (\xi^h - k)\cdot\frac{1-e^{-\gamma\Delta t^h(\xi^h)}}{\gamma} + \\ e^{-\gamma\Delta t^h(\xi^h)}\big[p^h(\xi^h,\xi^h+h)w(\xi^h+h) + p^h(\xi^h,\xi^h-h)w(\xi^h-h)\big]\Big\}. \tag{7.4}$$

Lemma 7. *Let*

$$v_0 \equiv 0$$
$$v_n \equiv \mathcal{H}^h v_{n-1}.$$

as $n \to \infty$, v_n converges to V^h uniformly over any compact interval.

Proof. Observe that in the norm $\sup_{\xi^h \le \bar{\xi}} |\cdot|$, the operator $\mathcal{H}^h$ is a contraction mapping over any finite interval. For bounded Borel functions $w, v : \mathbb{R}_+ \to [-k/\gamma, 0]$ we have

$$\sup_{\xi^h \le \bar{\xi}} |\mathcal{H}^h w - \mathcal{H}^h v| \le \left\{\max_{\xi^h \le \bar{\xi}} e^{-\gamma\Delta t^h(\xi^h)}\right\} \sup_{\xi^h \le \bar{\xi}} |w - v|$$

where $\left\{\max_{\xi^h \le \bar{\xi}} e^{-\gamma\Delta t^h(\xi^h)}\right\} < 1$. Moreover, from (7.3) it is clear that $\mathcal{H}^h V^h = V^h$. Using Banach's fixed-point theorem, we obtain the result. $\square$

Therefore, to find the function V^h, one should select $h > 0$ such that $p^h(h, 2h) >> p^h(h, 0)$. Once h is selected, the Ξ^h space should be divided into an equally spaced grid of length h, where the function value will be initiated to 0 at each grid point. Then, the operator $\mathcal{H}^h$ will be applied to each point iteratively, until a convergence is reached. If $\Xi^h_{\max}$ is selected large enough, the grid should contain the entire continuation region. Otherwise, one should select a larger $\Xi^h_{\max}$. The smallest grid point which takes the value 0 after the convergence will be assigned as ϕ^*, and will be used for calculations later.

Chapter 8

Misspecification

In this chapter, we will provide the dynamics of process $\tilde{\Phi}$ under the measure $\mathbb{P}_Q$, which turns out to be an Itô process, as well. This property of Φ and $\tilde{\Phi}$ pair characterizes them as a 2-dimensional Itô process, run by a 1-dimensional Wiener process.

8.1 The Dynamics of $\tilde{\Phi}$ under $\mathbb{P}_Q$

Note that

$$\begin{aligned} d\tilde{\Phi}_t &= \left(\tilde{\lambda} + \left[\tilde{\lambda} + \tilde{\beta}\frac{(\mu-\tilde{\mu})}{\sigma}\right]\tilde{\Phi}_t\right)dt + \tilde{\beta}\tilde{\Phi}_t\left[\frac{1}{\sigma}(dX_t - \mu dt) - \sigma dt\right] + \tilde{\beta}\tilde{\Phi}_t\sigma dt \\ &= \left(\tilde{\lambda} + \left[\tilde{\lambda} + \tilde{\beta}\left[\sigma + \frac{(\mu-\tilde{\mu})}{\sigma}\right]\right]\tilde{\Phi}_t\right)dt + \tilde{\beta}\tilde{\Phi}_t\left[\frac{1}{\sigma}(dX_t - \mu dt) - \sigma dt\right]. \end{aligned}$$

Moreover, we had stated that, under $\mathbb{P}_Q$, $Z_t = \left[\frac{1}{\sigma}(X_t - \mu t) - \sigma t\right]$ is a standard Brownian motion. Hence, under $\mathbb{P}_Q$, $\tilde{\Phi}_t$ is an Itô process with the dynamics

$$d\tilde{\Phi}_t = \left(\tilde{\lambda} + \left[\tilde{\lambda} + \tilde{\beta}\left[\sigma + \frac{(\mu-\tilde{\mu})}{\sigma}\right]\right]\tilde{\Phi}_t\right)dt + \tilde{\beta}\tilde{\Phi}_t Z_t.$$

So the decision maker solves another optimal stopping problem with the incorrect parameters

$$\tilde{\gamma} = r + \tilde{\lambda} - \tilde{\mu},\ \tilde{k} = \frac{\tilde{\mu} - r}{r - \tilde{\nu}} \text{ and } \tilde{\phi} = \frac{\tilde{\pi}}{1-\tilde{\pi}},$$

using the process $\tilde{\Phi}$ and finds an $\tilde{\phi}^*$. The first time $\tilde{\Phi}_t$ exits from the interval $[0, \tilde{\phi})$ is a stopping time denoted as $\tilde{\tau}^* = \inf\{t \geq 0 : \tilde{\Phi}_t \geq \tilde{\phi}^*\}$. Observe that this is the stopping time the decision maker uses for the calculation of the value function with the correct parameters. This miscalculated function is denoted by $H(\phi, \tilde{\phi})$ and defined as

$$H(\phi, \tilde{\phi}) := \mathbb{E}_Q^{\phi} \int_0^{\tilde{\tau}^*} e^{-\gamma t} \left[\Phi_t - k\right] \, dt.$$

Here, the intuition is that the true process Φ with the true parameters run simultaneously with the incorrect process $\tilde{\Phi}$, the decision maker stops the system when $\tilde{\Phi}$ passes $\tilde{\phi}^*$, however the true integrand (using the process Φ) is accumulated until this time and the incorrectly calculated expected earnings using this value become

$$\begin{aligned}
&S_0 \left[1 - (1-\pi)(r-\nu)\mathbb{E}_Q^{\frac{\pi}{1-\pi}} \int_0^{\tilde{\tau}^*} e^{-\gamma t} \left[\Phi_t - k\right] \, dt\right] \\
&= S_0 \left[1 - (1-\pi)(r-\nu) H\left(\frac{\pi}{1-\pi}, \frac{\tilde{\pi}}{1-\tilde{\pi}}\right)\right] \\
&\leq S_0 \left[1 - (1-\pi)(r-\nu) V\left(\frac{\pi}{1-\pi}\right)\right] \\
&= S_0 \left[1 - (1-\pi)(r-\nu)\mathbb{E}_Q^{\frac{\pi}{1-\pi}} \int_0^{\tau^*} e^{-\gamma t} \left[\Phi_t - k\right] \, dt\right].
\end{aligned}$$

The inequality is an immediate result of the fact that $\tilde{\tau}^*$ is a sub-optimal stopping time for the original problem.

8.2 On the Function $H(\phi, \tilde{\phi})$

Processes Φ and $\tilde{\Phi}$ use the same Wiener process under the measure $\mathbb{P}_Q$. Hence, their dynamics under this measure can be written as

$$d \begin{bmatrix} \Phi_t \\ \tilde{\Phi}_t \end{bmatrix} = \begin{bmatrix} \lambda + (\lambda + \sigma\beta)\Phi_t \\ \tilde{\lambda} + \left[\tilde{\lambda} + \tilde{\beta}\left[\sigma + \frac{(\mu - \tilde{\mu})}{\sigma}\right]\right]\tilde{\Phi}_t \end{bmatrix} dt + \begin{bmatrix} \beta\Phi_t \\ \tilde{\beta}\tilde{\Phi}_t \end{bmatrix} dZ_t$$

where Z is the standard Brownian motion under $\mathbb{P}_Q$ whereas β and $\tilde{\beta}$ are equal to $\frac{\nu-\mu}{\sigma}$ and $\frac{\tilde{\nu}-\tilde{\mu}}{\sigma}$, respectively. Observe that $H(\phi, \tilde{\phi}) = 0$ for $\phi \in \mathbb{R}_+$, $\tilde{\phi} \geq \tilde{\phi}^*$, by definition. For a function $f : \mathbb{R}_+^2 \to \mathbb{R}$ with the boundary condition $f(., \tilde{\phi}^*) = 0$, using Itô's formula we can write

$$\mathbb{E}\left[e^{-\gamma\tilde{\tau}^*} f(\Phi_{\tilde{\tau}^*}, \tilde{\Phi}_{\tilde{\tau}^*}) - f(\phi, \tilde{\phi})\right] =$$
$$\mathbb{E}\left[\int_0^{\tilde{\tau}^*} -\gamma e^{-\gamma t} f(\Phi_t, \tilde{\Phi}_t)\ dt\right] + \mathbb{E}\left[\int_0^{\tilde{\tau}^*} e^{-\gamma t} f_\phi(\Phi_t, \tilde{\Phi}_t) a(\Phi_t)\ dt\right] +$$
$$\mathbb{E}\left[\int_0^{\tilde{\tau}^*} e^{-\gamma t} f_{\tilde{\phi}}(\Phi_t, \tilde{\Phi}_t) \tilde{a}(\tilde{\Phi}_t)\ dt\right] + \frac{1}{2}\mathbb{E}\left[\int_0^{\tilde{\tau}^*} e^{-\gamma t} f_{\phi\phi}(\Phi_t, \tilde{\Phi}_t) b^2(\Phi_t)\ dt\right] +$$
$$\frac{1}{2}\mathbb{E}\left[\int_0^{\tilde{\tau}^*} e^{-\gamma t} f_{\tilde{\phi}\tilde{\phi}}(\Phi_t, \tilde{\Phi}_t) \tilde{b}^2(\tilde{\Phi}_t)\ dt\right] + \mathbb{E}\left[\int_0^{\tilde{\tau}^*} e^{-\gamma t} f_{\phi\tilde{\phi}}(\Phi_t, \tilde{\Phi}_t) b(\Phi_t) \tilde{b}(\tilde{\Phi}_t)\ dt\right].$$

Here

$$a(\phi) = \lambda + (\lambda + \sigma\beta)\phi, \quad \tilde{a}(\tilde{\phi}) = \tilde{\lambda} + \left[\tilde{\lambda} + \tilde{\beta}\left[\sigma + \frac{(\mu - \tilde{\mu})}{\sigma}\right]\right]\tilde{\phi},$$
$$b(\phi) = \beta\phi, \quad \tilde{b}(\tilde{\phi}) = \tilde{\beta}\tilde{\phi},$$

$f_\phi, f_{\tilde{\phi}}$ and $f_{\phi\phi}, f_{\tilde{\phi}\tilde{\phi}}, f_{\phi\tilde{\phi}}$ denote the first and the second order partial derivatives of f, respectively. Using these results we can conclude that to find our function $H(\phi, \tilde{\phi})$, we should solve

$$-\gamma f(\phi, \tilde{\phi}) + a(\phi) f_\phi(\phi, \tilde{\phi}) + \frac{1}{2} b^2(\phi) f_{\phi\phi}(\phi, \tilde{\phi}) +$$
$$\tilde{a}(\tilde{\phi}) f_{\tilde{\phi}}(\phi, \tilde{\phi}) + \frac{1}{2}\tilde{b}^2(\tilde{\phi}) f_{\tilde{\phi}\tilde{\phi}}(\phi, \tilde{\phi}) + b(\phi)\tilde{b}(\tilde{\phi}) f_{\phi\tilde{\phi}}(\phi, \tilde{\phi}) + \phi - k = 0.$$

However, this equation does not yield a general closed-form solution and requires advanced methods to solve. To address this issue, we have used Monte Carlo simulations to calculate expectations, which will be explained in detail in the next chapter.

Chapter 9

Solution of the Misspecified Problem

In this chapter, two methods to find the function $H(\phi, \tilde{\phi})$ will be provided. Both of these methods will use Monte Carlo simulations to find the corresponding expectation. Observe that, the calculation of the function $H(\phi, \tilde{\phi})$ is a two-stage process.

1. Finding $\tilde{\phi}^*$ by solving the original problem with the incorrect parameters
2. Calculating $\mathbb{E}_Q^{\phi} \int_0^{\tilde{\tau}^*} e^{-\gamma t} \left[\Phi_t - k\right] \, dt$ by using the corresponding stoping time $\tilde{\tau}^*$.

For the first part, the solution method mentioned in Chapter 3 was used, the user enters the incorrect parameters to the algorithm and finds $\tilde{\phi}^*$ value. For the second part, Monte Carlo simulations were used. Since $H(\phi, \tilde{\phi})$ is an expectation, the realized values of the random variable

$$\int_0^{\tilde{\tau}^*} e^{-\gamma t} \left[\Phi_t - k\right] \, dt \tag{9.1}$$

were calculated for a large number of sample paths and their mean was used for approximating $H(\phi, \tilde{\phi})$. Two methods were adopted for the simulations: Markov chain approximation, and Euler's method. Moreover, to decrease the variance of the simulations for the Euler's method, antithetic and control variables were used.

Recall that, the dynamics of the processes Φ and $\tilde{\Phi}$ are

$$d\begin{bmatrix}\Phi_t \\ \tilde{\Phi}_t\end{bmatrix} = \begin{bmatrix}\lambda + (\lambda + \sigma\beta)\Phi_t \\ \tilde{\lambda} + \left[\tilde{\lambda} + \tilde{\beta}\left[\sigma + \frac{(\mu - \tilde{\mu})}{\sigma}\right]\right]\tilde{\Phi}_t\end{bmatrix} dt + \begin{bmatrix}\beta\Phi_t \\ \tilde{\beta}\tilde{\Phi}_t\end{bmatrix} dZ_t.$$

Therefore, at each step, the same Wiener sample path should be used for the calculation of (9.1). Once we have the value of $\tilde{\phi}^*$, each simulation is stopped when the process $\tilde{\Phi}$ passes this threshold value, in other words, at the realized value of the random variable $\tilde{\tau}^*$. During the simulation, the integral is calculated using the process Φ as a cumulative sum, and which gives the required result. Next two sections describe two different method of simulations used in this study.

9.1 Markov Chain Approximation Method

Let us define a new process Ψ as $\Psi_t := \tilde{\beta}\log\Phi_t - \beta\log\tilde{\Phi}_t$. Using Itô's formula, we can write its dynamics with $\tilde{\Phi}$ as

$$d\begin{bmatrix}\Psi_t \\ \tilde{\Phi}_t\end{bmatrix} = \begin{bmatrix}\lambda\tilde{\beta}\left[1 + \frac{1}{\Phi_t}\right] - \tilde{\lambda}\beta\left[1 + \frac{1}{\tilde{\Phi}_t}\right] - \beta\tilde{\beta}\left[\frac{\beta - \tilde{\beta}}{2} + \frac{\mu - \tilde{\mu}}{\sigma}\right] \\ \tilde{\lambda} + \left[\tilde{\lambda} + \tilde{\beta}\left[\sigma + \frac{(\mu - \tilde{\mu})}{\sigma}\right]\right]\tilde{\Phi}_t\end{bmatrix} dt + \begin{bmatrix}0 \\ \tilde{\beta}\tilde{\Phi}_t\end{bmatrix} dZ_t,$$

or without using Φ, as

$$d\begin{bmatrix}\Psi_t \\ \tilde{\Phi}_t\end{bmatrix} = \begin{bmatrix}\lambda\tilde{\beta}\left[1 + e^{-\frac{\Psi_t + \beta\log\tilde{\Phi}_t}{\tilde{\beta}}}\right] - \tilde{\lambda}\beta\left[1 + \frac{1}{\tilde{\Phi}_t}\right] - \beta\tilde{\beta}\left[\frac{\beta - \tilde{\beta}}{2} + \frac{\mu - \tilde{\mu}}{\sigma}\right] \\ \tilde{\lambda} + \left[\tilde{\lambda} + \tilde{\beta}\left[\sigma + \frac{(\mu - \tilde{\mu})}{\sigma}\right]\right]\tilde{\Phi}_t\end{bmatrix} dt + \begin{bmatrix}0 \\ \tilde{\beta}\tilde{\Phi}_t\end{bmatrix} dZ_t.$$

The process Ψ is a first order process satisfying

$$d\Psi_t = a(\Phi_t, \tilde{\Phi}_t)\, dt,$$

where

$$a(\Phi_t, \tilde{\Phi}_t) = \lambda\tilde{\beta}\left[1 + \frac{1}{\Phi_t}\right] - \tilde{\lambda}\beta\left[1 + \frac{1}{\tilde{\Phi}_t}\right] - \beta\tilde{\beta}\left[\frac{\beta - \tilde{\beta}}{2} + \frac{\mu - \tilde{\mu}}{\sigma}\right].$$

Hence, we can write the approximation

$$\Psi_{t+\Delta t} \approx \Psi_t + a(\Phi_t, \tilde{\Phi}_t)\Delta t,$$

or equivalently,

$$\tilde{\beta} \log \Phi_{t+\Delta t} - \beta \log \tilde{\Phi}_{t+\Delta t} \approx \tilde{\beta} \log \Phi_t - \beta \log \tilde{\Phi}_t + a(\Phi_t, \tilde{\Phi}_t)\Delta t.$$

Re-arranging the terms we obtain

$$\Phi_{t+\Delta t} \approx \Phi_t \left(\tilde{\Phi}_{t+\Delta t}/\tilde{\Phi}_t\right)^{\beta/\tilde{\beta}} e^{\frac{a(\Phi_t,\tilde{\Phi}_t)}{\tilde{\beta}}}\Delta t.$$

Hence, we can use the Markov chain approximation described in the Section 7.3 for the process $\tilde{\Phi}$, and find the value of the process Φ accordingly, for the calculations. In other words, at time t, the approximation of $\tilde{\Phi}_{t+\Delta t}$, which is $\tilde{\Xi}^h_{t+\Delta t^h(\tilde{\Xi}^h_t)}$ is calculated, and the corresponding approximation of the process Φ, which is Ξ^h, is obtained. Observe that the Δt value in the formula was replaced with $\Delta t^h(\tilde{\Xi}^h_t)$ value of the Markov chain approximation. Each process was assumed to be constant over this period of time, thus

$$\int_t^{t+\Delta t^h(\tilde{\Xi}^h_t)} e^{-\gamma s}\left[\Phi_s - k\right]\, ds \approx (\Xi^h_t - k)\frac{1}{\gamma}e^{-\gamma t}(1 - e^{-\gamma \Delta t^h(\tilde{\Xi}^h_t)}),$$

and

$$\int_0^{\tilde{\tau}^*} e^{-\gamma s}\left[\Phi_s - k\right]\, ds \approx \sum_{n=0}^{N^h-1}(\Xi^h_{t_n} - k)\frac{1}{\gamma}e^{-\gamma t_n}(1 - e^{-\gamma \Delta t^h(\tilde{\Xi}^h_{t_n})}).$$

Where N^h denotes the stopping time defined as $N^h := \inf\{n \in \mathbb{N} : \tilde{\Xi}^h_{t_n} \geq \tilde{\phi}^*\}$.

9.2 Euler's Method

To use this method, we use the fact that for a fixed small $\Delta t > 0$ and $Z \sim \mathcal{N}(0,1)$ we can approximate the change in processes Φ_t and $\tilde{\Phi}_t$ as

$$\Delta \begin{bmatrix} \Phi_t \\ \tilde{\Phi}_t \end{bmatrix} \approx \begin{bmatrix} \lambda + (\lambda + \sigma\beta)\Phi_t \\ \tilde{\lambda} + \left[\tilde{\lambda} + \tilde{\beta}\left[\sigma + \frac{(\mu-\tilde{\mu})}{\sigma}\right]\right]\tilde{\Phi}_t \end{bmatrix} \Delta t + \begin{bmatrix} \beta\Phi_t \\ \tilde{\beta}\tilde{\Phi}_t \end{bmatrix} \sqrt{\Delta t} Z. \tag{9.2}$$

Here $\Delta\Phi_t$ and $\Delta\tilde{\Phi}_t$ denotes $\Phi_{t+\Delta t} - \Phi_t$ and $\tilde{\Phi}_{t+\Delta t} - \tilde{\Phi}_t$ respectively. The idea is that $W_{t+\Delta t} - W_t \sim \mathcal{N}(0, \Delta t)$, for a Wiener process W.

To run the simulation, we initiate the processes Φ and $\tilde{\Phi}$ with ϕ and $\tilde{\phi}$ values. At each step, Δt amount of time is passed and the processes are updated using (9.2) for their new values at time $t + \Delta t$. Observe that the key point here is that two processes are updated using the same Z, as they are running with respect to the same Wiener Process. The calculation of the required function is the same with the Markov chain method. The simulation is stopped when the process $\tilde{\Phi}$ passes $\tilde{\phi}^*$, and the cumulative sum of the integrand gives the result.

We used two methods to reduce the variance of the Euler simulations: antithetic and control variables. Next two subsections explain the application of these methods.

9.2.1 Method of Antithetic Variables

Observe that for the Euler's method, we use standard normal variable Z at each step of the simulation. Because of its symmetric density, we have $\mathbb{P}\{Z \in dz\} = \mathbb{P}\{-Z \in dz\}$. Hence, at each step of the simulation, we can use both Z and $-Z$ to simulate our processes. More precisely,

$$\Delta \begin{bmatrix} \Phi_t^1 \\ \tilde{\Phi}_t^1 \end{bmatrix} \approx \begin{bmatrix} \lambda + (\lambda + \sigma\beta)\Phi_t^1 \\ \tilde{\lambda} + \left[\tilde{\lambda} + \tilde{\beta}\left[\sigma + \frac{(\mu-\tilde{\mu})}{\sigma}\right]\right]\tilde{\Phi}_t^1 \end{bmatrix} \Delta t + \begin{bmatrix} \beta\Phi_t^1 \\ \tilde{\beta}\tilde{\Phi}_t^1 \end{bmatrix} \sqrt{\Delta t} Z$$

$$\Delta \begin{bmatrix} \Phi_t^2 \\ \tilde{\Phi}_t^2 \end{bmatrix} \approx \begin{bmatrix} \lambda + (\lambda + \sigma\beta)\Phi_t^2 \\ \tilde{\lambda} + \left[\tilde{\lambda} + \tilde{\beta}\left[\sigma + \frac{(\mu-\tilde{\mu})}{\sigma}\right]\right]\tilde{\Phi}_t^2 \end{bmatrix} \Delta t - \begin{bmatrix} \beta\Phi_t^2 \\ \tilde{\beta}\tilde{\Phi}_t^2 \end{bmatrix} \sqrt{\Delta t} Z.$$

Here Φ^1, $\tilde{\Phi}^1$ and Φ^2, $\tilde{\Phi}^2$ denote the processes generated simultaneously by Z and $-Z$, respectively. When both $\tilde{\Phi}^1$ and $\tilde{\Phi}^2$ reach to $\tilde{\phi}^*$, we stop calculating the integrals using Φ^1 and Φ^2 respectively. At the end of the simulation, these results' average is calculated and counted as a single simulation result. The reason is that these results are not independent of each other. However, each simulation's averaged result is independent of the other simulations' results.

9.2.2 Method of Control Variables

For this method, we will use an alternative unbiased estimator for our simulations. Recall that we are trying to find the expectation

$$\mathbb{E}_Q^{\phi} \int_0^{\tilde{\tau}^*} e^{-\gamma t} \left[\Phi_t - k\right] \, dt.$$

If we denote the random variable inside of the expectation with X, on our simulations, we use the sample mean $\bar{X}$, which is an unbiased estimator to approximate our random variable X. Now suppose that we have a random variable Y, with a known mean μ_Y. In this case for $\alpha \in \mathbb{R}$ and $n \in \mathbb{N}$

$$\frac{1}{n} \sum_{i=1}^{n} [X_i + \alpha(Y_i - \mu_Y)] \tag{9.3}$$

is also an unbiased estimator for our simulations. Since each simulation is independent, we know that the variance of this estimator is

$$\begin{aligned} \mathrm{Var}\left(\frac{1}{n} \sum_{i=1}^{n} X_i + \alpha(Y_i - \mu_Y)\right) &= \frac{1}{n^2} \sum_{i=1}^{n} \mathrm{Var}\left(X_i + \alpha(Y_i - \mu_Y)\right) = \frac{1}{n} \mathrm{Var}(X_i + \alpha Y_i) \\ &= \frac{1}{n} \left[\sigma_X^2 + \alpha^2 \sigma_Y^2 + 2\alpha \sigma_{X,Y}\right]. \end{aligned}$$

Where σ_X^2 and σ_Y^2 are the variance of X and Y, respectively. $\sigma_{X,Y}$ denotes the covariance of these random variables. We want to minimize our simulation error by choosing α accordingly. Using standard optimization arguments, we know that the α value that minimizes this expression is

$$\alpha^* = -\frac{\sigma_{X,Y}}{\sigma_Y^2}.$$

If we use this value, our resulting variance becomes

$$\mathrm{Var}(X_i + \alpha^* Y_i) = \sigma_X^2 - \frac{\sigma_{X,Y}^2}{\sigma_Y^2} = \sigma_X^2 \left(1 - \frac{\sigma_{X,Y}^2}{\sigma_X^2 \sigma_Y^2}\right) = \sigma_X^2 (1 - \rho_{X,Y}^2). \tag{9.4}$$

Where $\rho_{X,Y}$ denotes the correlation of X and Y. Note that 0 is an alternative for α, so we are guaranteed to decrease our simulation error using this method. Secondly, we should select a Y, that is highly correlated with our random variable X (see (9.4)).

We need to come up with an estimate of α^* to use in our simulations. To that end, before our main simulations, we run a preliminary simulation and estimate α^* with

$$\hat{\alpha}^* = -\frac{\sum_{j=1}^{n_1}(X_i - \bar{X})(Y_i - \bar{Y})}{\sum_{j=1}^{n_1}(Y_i - \bar{Y})^2}.$$

When we use this value on our simulations, we can estimate the error of our estimator with $\frac{\hat{S}}{\sqrt{n}}$, where $\hat{S}$ is the sample standard deviation of our main simulations with $\hat{\alpha}^*$.

Finding a control variable

To implement this method, we need to come up with a random variable Y, which has a known mean and can be calculated during our simulations. Recall that the dynamics of the process $\tilde{\Phi}$ are

$$d\tilde{\Phi}_t = (\tilde{\lambda} + \tilde{\ell}\tilde{\Phi}_t)\, dt + \beta\tilde{\Phi}_t\, dZ_t,$$

where $\tilde{\ell} := \tilde{\lambda} + \tilde{\beta}\left[\sigma + \frac{(\mu-\tilde{\mu})}{\sigma}\right]$ (and we assume $\tilde{\ell} \neq 0$ for the moment). One can verify that the function

$$f(\tilde{\phi}) = 1 + \frac{\tilde{\ell}}{\tilde{\lambda}}\tilde{\phi}$$

solves the equation

$$-\tilde{\ell}f(\tilde{\phi}) + (\tilde{\lambda} + \tilde{\ell}\tilde{\phi})f'(\tilde{\phi}) + \frac{1}{2}\beta^2\tilde{\phi}^2 f''(\tilde{\phi}) = 0.$$

Then for $0 < a < \tilde{\phi} < \tilde{\phi}^*$, if we denote $\tilde{\tau}_a = \inf\{t > 0 : \tilde{\Phi}_t \leq a\}$, using Itô's lemma we have

$$\mathbb{E}\left[e^{-\tilde{\ell}(\tilde{\tau}_a \wedge \tilde{\tau}^*)} f(\tilde{\Phi}_{\tilde{\tau}_a \wedge \tilde{\tau}^*}) - f(\tilde{\phi})\right] = 0.$$

So we have a random variable $Y \equiv e^{-\tilde{\ell}(\tilde{\tau}_a \wedge \tilde{\tau}^*)} f(\tilde{\Phi}_{\tilde{\tau}_a \wedge \tilde{\tau}^*})$ with the known mean $\mu_Y = f(\tilde{\phi})$. Moreover, note that the realizations of Y can be calculated easily during our simulations. This Y is useful when $\tilde{\ell} < 0$. If $\tilde{\ell} > 0$, we can find a simpler Y using dominated convergence theorem and the entry-not-exit property of the origin, we can let $a \searrow 0$ and rewrite

$$\lim_{a \searrow 0} \mathbb{E}\left[e^{-\tilde{\ell}(\tilde{\tau}_a \wedge \tilde{\tau}^*)} f(\tilde{\Phi}_{\tilde{\tau}_a \wedge \tilde{\tau}^*}) - f(\tilde{\phi})\right] = \mathbb{E}\left[\lim_{a \searrow 0} e^{-\tilde{\ell}(\tilde{\tau}_a \wedge \tilde{\tau}^*)} f(\tilde{\Phi}_{\tilde{\tau}_a \wedge \tilde{\tau}^*}) - f(\tilde{\phi})\right]$$

$$= \mathbb{E}\left[e^{-\tilde{\ell}\tilde{\tau}^*} f(\tilde{\Phi}_{\tilde{\tau}^*}) - f(\tilde{\phi})\right] = \mathbb{E}\left[e^{-\tilde{\ell}\tilde{\tau}^*} f(\tilde{\phi}^*) - f(\tilde{\phi})\right] = 0.$$

So when $\tilde{\ell} > 0$, we can use $Y \equiv e^{-\tilde{\ell}\tilde{\tau}^*} f(\tilde{\phi}^*)$ with the known mean $\mu_Y = f(\tilde{\phi})$.

On the other hand, when $\ell = 0$, the dynamics of the process $\tilde{\Phi}$ become

$$d\tilde{\Phi}_t = \tilde{\lambda}\ dt + \beta\tilde{\Phi}_t\ dZ_t.$$

Using Itô's lemma, we can write

$$\mathbb{E}[\tilde{\Phi}_{\tilde{\tau}_a \wedge \tilde{\tau}^*} - \tilde{\phi}] = \tilde{\lambda}\mathbb{E}[\tilde{\tau}_a \wedge \tilde{\tau}^*].$$

Again, using dominated convergence theorem, if we let $a \searrow 0$, we can write

$$\begin{aligned}
&\lim_{a \searrow 0} \mathbb{E}[\tilde{\Phi}_{\tilde{\tau}_a \wedge \tilde{\tau}^*} - \tilde{\phi}] = \mathbb{E}[\lim_{a \searrow 0} \tilde{\Phi}_{\tilde{\tau}_a \wedge \tilde{\tau}^*} - \tilde{\phi}] = \mathbb{E}[\tilde{\Phi}_{\tilde{\tau}^*} - \tilde{\phi}] = \tilde{\phi}^* - \tilde{\phi} \\
&= \tilde{\lambda} \lim_{a \searrow 0} \mathbb{E}[\tilde{\tau}_a \wedge \tilde{\tau}^*] = \tilde{\lambda}\mathbb{E}[\lim_{a \searrow 0} \tilde{\tau}_a \wedge \tilde{\tau}^*] = \tilde{\lambda}\mathbb{E}[\tilde{\tau}^*].
\end{aligned}$$

Hence, when $\tilde{\ell} = 0$, we can use $Y \equiv \tilde{\tau}^*$ with the known mean $\mu_Y = \frac{\tilde{\phi}^* - \tilde{\phi}}{\tilde{\lambda}}$.

For numerical examples, both of these methods were used. Because each simulation is independent, sample paths were created in parallels. Since it is a simpler algorithm, Euler method worked faster than the Markov chain method. Moreover, both antithetic variables and control variables reduced the variance significantly. However, antithetic variables performed slightly better than the control variables. Finally, we also tried implementing both antithetic and control variable methods together, but this did not improve the performance more than antithetics themselves.

9.3 A Sample Path Example

In this section, we will explain how the system is stopped by the decision maker for a sample path. The problem parameters are $\mu = 0.6$, $\nu = -0.5$, $\tilde{\mu} = 0.7$, $\tilde{\nu} = -0.2$, $\lambda = \tilde{\lambda} = 0.9$, $\pi = \tilde{\pi} = 0$, $r = 0.47$ and $\sigma = 0.2$.

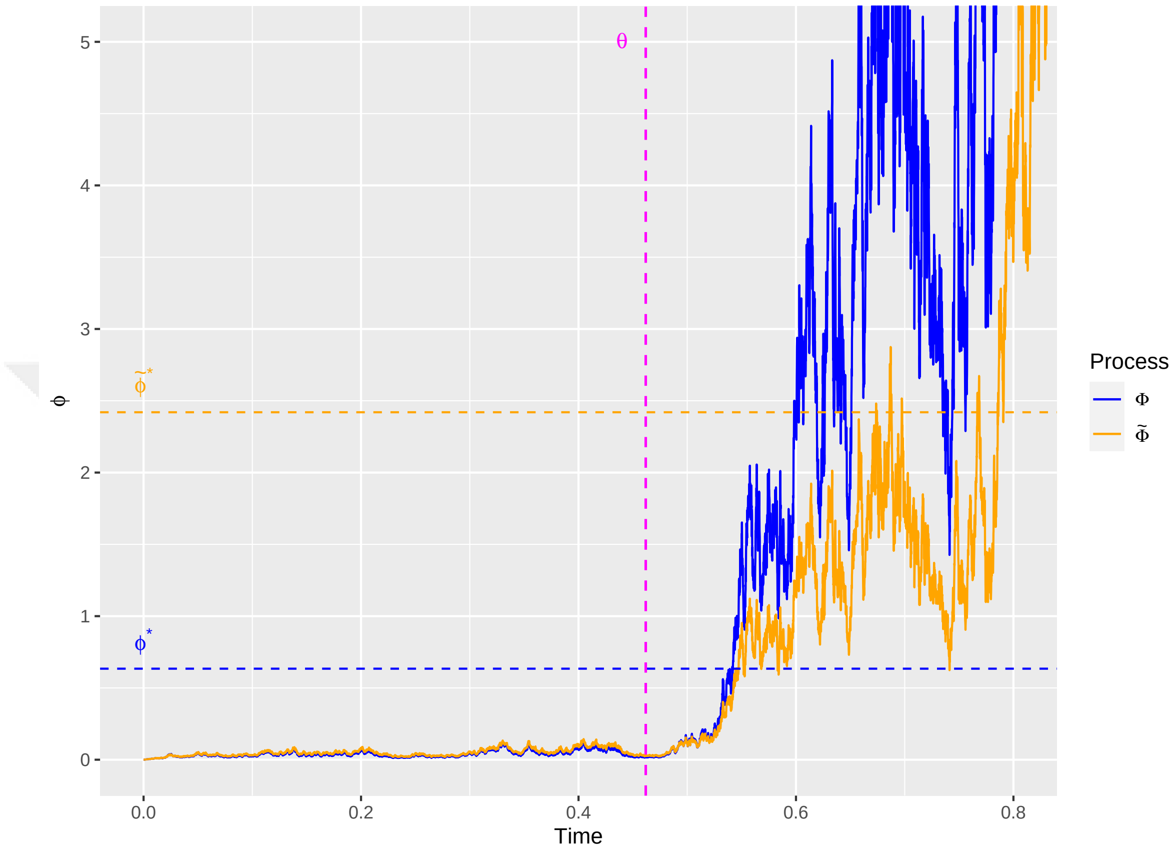


Figure 9.1: Sample path for the Φ and $\tilde{\Phi}$ processes

Figure 9.1 illustrates the realized processes Φ (blue) and $\tilde{\Phi}$ (orange) for a sample path of X. The disorder happens at $\theta = 0.4618$. The decision maker solves the optimal stopping problem with their incorrect parameters and find $\tilde{\phi}^*$ as 2.42 which is depicted with the orange dashed line. While the solution with the correct parameters yields ϕ^* as 0.634, depicted with the blue dashed line. The decision maker stops the process when $\tilde{\Phi}$ passes $\tilde{\phi}^*$, therefore in this case delays to stop the system.

In the next chapter, we will provide a real-life example of the problem from the literature and do a sensitivity analysis on the misspecification of various problem parameters.

Chapter 10

Numerical Examples

In this chapter, we will provide numerical results for the misspecification problem, using Shiryaev et al.'s [42] study on Apple stock price bubble of 2009-2012. This paper deals with the problem of when to sell a stock in the presence of a bubble. Specifically, starting from March 2009, it was assumed that the drift of the underlying Brownian motion that models the stock price will change before December, 2012. For this problem, authors adopted a uniform prior as the disorder density. Moreover, eight different entering dates were considered, each trading day was assumed to have a length of $\Delta t = 0.01$. Each entering date was assumed to be the initial point of the disorder density ($t = 0$) and the distance from the last day of 2012 was denoted by T, thus the disorder was distributed Unif $\sim [0, T]$. For the problem parameters, the drift before the disorder was determined by using historical stock prices. For the post-disorder drift, a scenario analysis was conducted. If the pre-disorder drift is μ, cases where post-disorder rate is $\nu = -0.5\mu, -\mu, -2\mu, -3\mu$ is investigated.

This paper motivated us to ask what happens if a decision maker uses incorrect parameters for solving the disorder problem, turning it into a parameter misspecification problem. We investigated three different scenarios: first, where the post-disorder drift (ν) was misspecified; second, where both the pre- and post-disorder drifts (μ and ν) were misspecified; and lastly, where the density parameter of the underlying prior (λ) was misspecified. The entry date of December 31, 2009, was chosen, making $T = 7.53$. λ was chosen such that $\mathbb{P}\{\theta \leq 7.53\} > 0.999$. For the risk-free rate r, the three-year treasury bond return rate was considered. Finally, for the initial stock price S_0, pre-disorder rate μ, and volatility σ, the parameters specified in the paper were used.

10.1 First Scenario: Misspecification of the Post-Disorder Drift

For the first scenario, the misspecification of the post-disorder rate was of interest. We investigated seven different cases where $\nu = -0.5\mu, -0.75\mu, -\mu, -1.5\mu, -2\mu, -2.5\mu, -3\mu$. For each case, $\tilde{\nu}$ was allowed to take these values, as well. Since the pre-disorder rate can be obtained from historical data, measuring the effect of choosing an incorrect post-disorder rate becomes valuable in this context. Observe that when $\nu = \tilde{\nu}$, the problem turns into the original detection problem. Other parameters are: S_0=\$210.73, $\lambda = 0.92$, $\mu = 0.452$, $\pi = 0$, $\sigma = 0.186$.

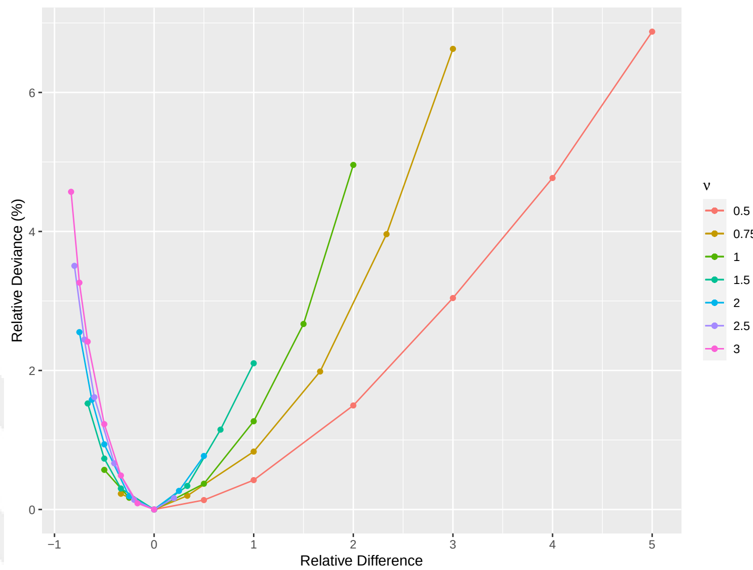


Figure 10.1: The effect of the decision maker's parameter for a fixed correct parameter

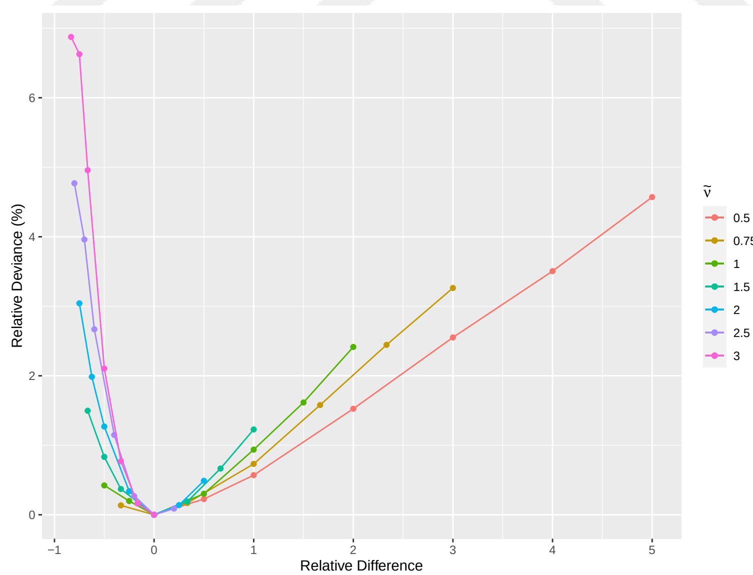


Figure 10.2: The effect of the correct parameter when the decision maker's parameter is fixed

Figure 10.3: The effect of the correct and incorrect parameters for Scenario 1

The two plots in Figure 10.3 depict the effect of the difference between correct and incorrect parameters. More precisely, the x-axes denote

$$\text{Relative Difference}_1 = \frac{\tilde{\nu} - \nu}{\nu} \quad \text{and} \quad \text{Relative Difference}_2 = \frac{\nu - \tilde{\nu}}{\tilde{\nu}},$$

respectively. And y-axes denote how much the discounted prices calculated by incorrect parameters deviate from the discounted prices calculated by correct parameters. More precisely,

$$\begin{aligned}
&\text{Relative Deviance } (\%) \\
&= \left| \frac{S_0 \left[1 - (1-\pi)(r-\nu)H(\phi, \tilde{\phi})\right] - S_0 \left[1 - (1-\pi)(r-\nu)V(\phi)\right]}{S_0 \left[1 - (1-\pi)(r-\nu)V(\phi)\right]} \right| \cdot 100\% \\
&= \left| \frac{-(1-\pi)(r-\nu) \left[H(\phi, \tilde{\phi}) - V(\phi)\right]}{1 - (1-\pi)(r-\nu)V(\phi)} \right| \cdot 100\%.
\end{aligned}$$

As it is expected, when the distance between correct and incorrect parameters increases, the deviance from the true discounted price increases, as well. One key observation here is that, choosing a pessimistic post-disorder rate, then the true post-disorder rate, costs more than the case of choosing an optimistic post-disorder rate. For example, $\nu = -0.5\mu$, $\tilde{\nu} = -3\mu$ case is worse than $\nu = -3\mu$, $\tilde{\nu} = -0.5\mu$ case. The intuition is that the decision maker expects a sharp change when $\tilde{\nu} = -3\mu$, however that doesn't happen, thus the stopping delays. However, when $\tilde{\nu} = -0.5\mu$, the decision maker is ready to stop the process even for a slight change, which is happening because $\nu = -3\mu$.

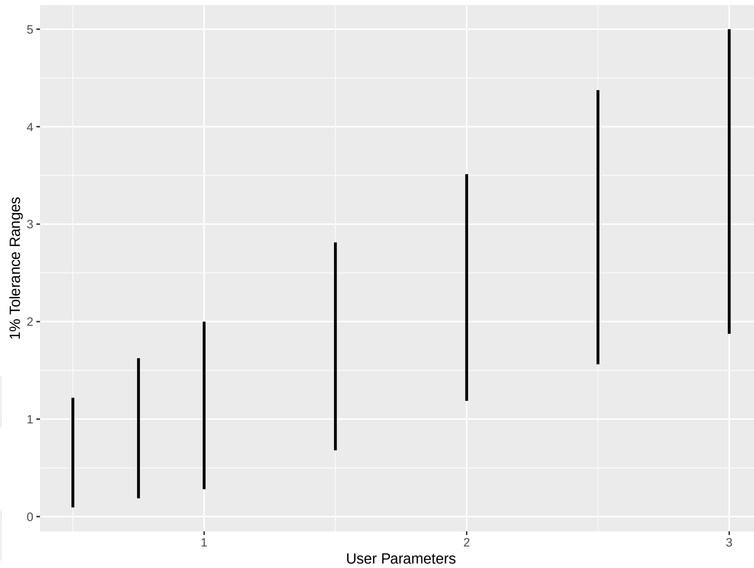


Figure 10.4: 1% Tolerance regions for seven different user parameters

Figure 10.4 depicts the 1% tolerance regions for the seven different cases mentioned above. Tolerance here means that when user specifies its parameter for the model ($\tilde{\nu}$), what is the range of true parameters (ν) such that the deviation from the true discounted price does not exceeds 1%.

10.2 Second Scenario: Misspecification of the Pre- and Post-disorder Drifts

For the second scenario, the decision maker believes that the magnitude of the pre- and post-disorder rates is the same, i.e. $\mu = -\nu$. It was assumed that indeed this is the case, however, the decision maker is not certain about this magnitude. In this case, a natural question arises: whether misspecification of both parameters affects the deviation from the correct objective value more than misspecification of these parameters individually. To investigate this, three different options were considered across two cases. Firstly, keeping μ and ν fixed, variations were made in only $\tilde{\mu}$, only $\tilde{\nu}$, and both $\tilde{\mu}$ and $\tilde{\nu}$. Secondly, with fixed $\tilde{\mu}$ and $\tilde{\nu}$, variations were made in μ, ν, and both μ and ν. $\mu = -\nu = 0.5 \cdot 0.226$ and $\tilde{\mu} = \tilde{\nu} = 0.5 \cdot 0.226$ were used as base cases. Other parameters were the same with the previous example.

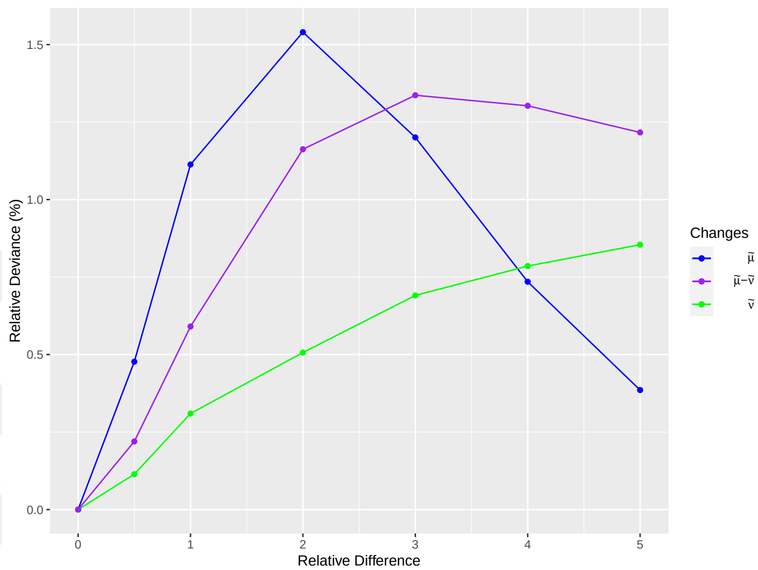


Figure 10.5: The effect of the decision maker's parameters when the correct parameters are $\mu = -\nu = 0.5 \cdot 0.226$

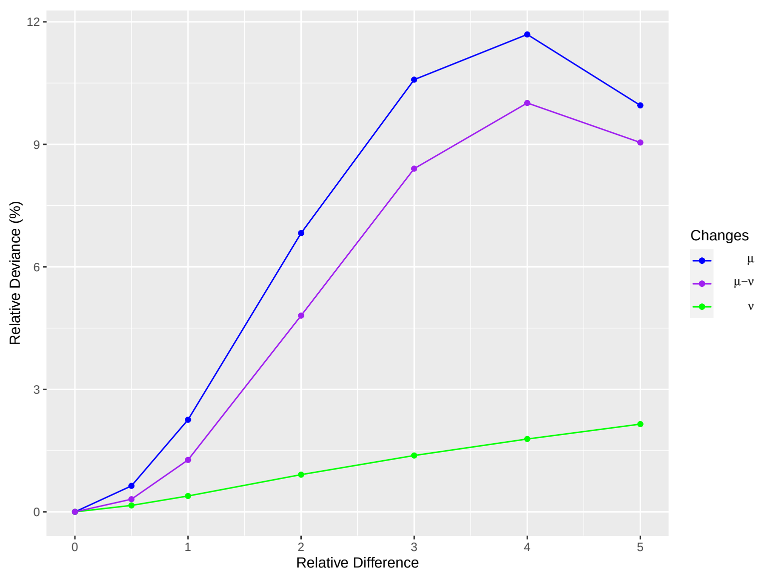


Figure 10.6: The effect of the correct parameters when the decision maker's parameters are $\tilde{\mu} = -\tilde{\nu} = 0.5 \cdot 0.226$

Figure 10.7: The effect of the correct and incorrect parameters for Scenario 2

Figure 10.7 displays the results for these cases. Even ν and $\tilde{\nu}$ behaves monotonically, that is not the case for μ and $\tilde{\mu}$. A possible answer for this interesting result is that, the decision maker solve the optimization problem with parameter $\tilde{\gamma} = r + \tilde{\lambda} - \tilde{\mu}$ to obtain $\tilde{\phi}^*$ and the true problem is solved by $\gamma = r + \lambda - \mu$. So for both cases, a high value of μ and $\tilde{\mu}$ turns the problem into a trivial problem where it is never optimal to stop or it is optimal to stop immediately. Parameter ν or $\tilde{\nu}$ does not have such relation with other problem parameters. Finally, Figure 10.8 displays the 0.5% tolerance interval for the case where $\mu = -\nu$ and $\tilde{\mu} = -\tilde{\nu}$. $\tilde{\mu}$ and $\tilde{\nu}$ were changed to $\pm 5\%$ and $\pm 10\%$ from $0.5 \cdot 0.226$.

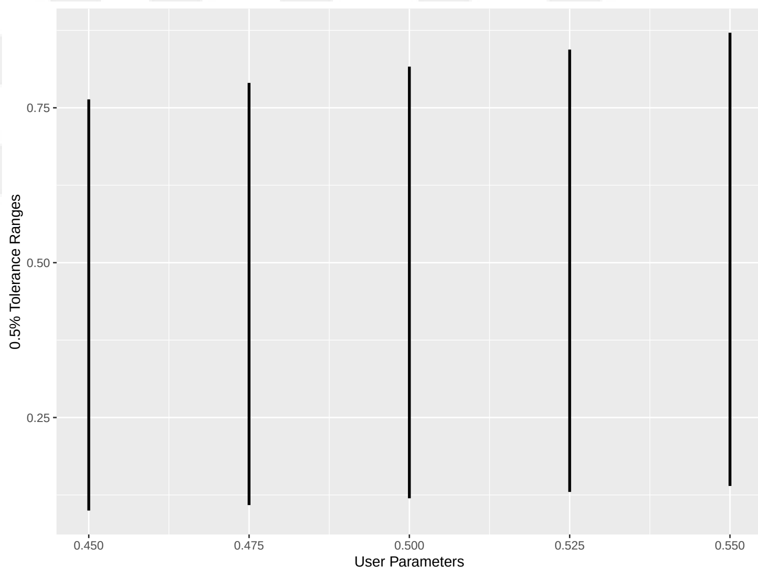


Figure 10.8: 0.5% Tolerance regions for five different user parameters

10.3 Third Scenario: Misspecification of the Mean Time to Disorder

For the last scenario, we considered the misspecification of λ. Here, $\tilde{\lambda}$ denotes the decision maker's belief about where they expect the disorder to most likely occur. In our problems we have used $\lambda = 0.92$ as the base case. λ and $\tilde{\lambda}$ were changed to $\pm 5\%$ and $\pm 10\%$ from 0.92. As it is seen in Figure 10.11, a monotone deviance behavior was observed for fixed λ and $\tilde{\lambda}$.

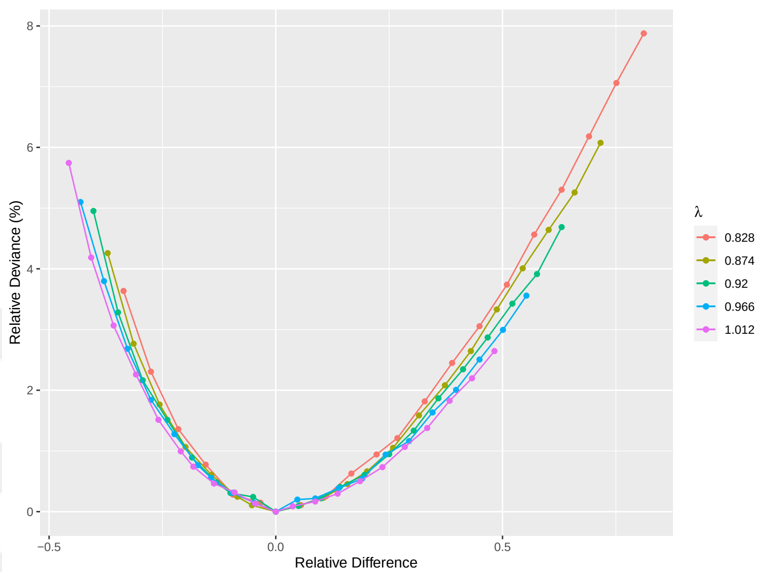


Figure 10.9: The effect of the decision maker's parameter for a fixed correct parameter

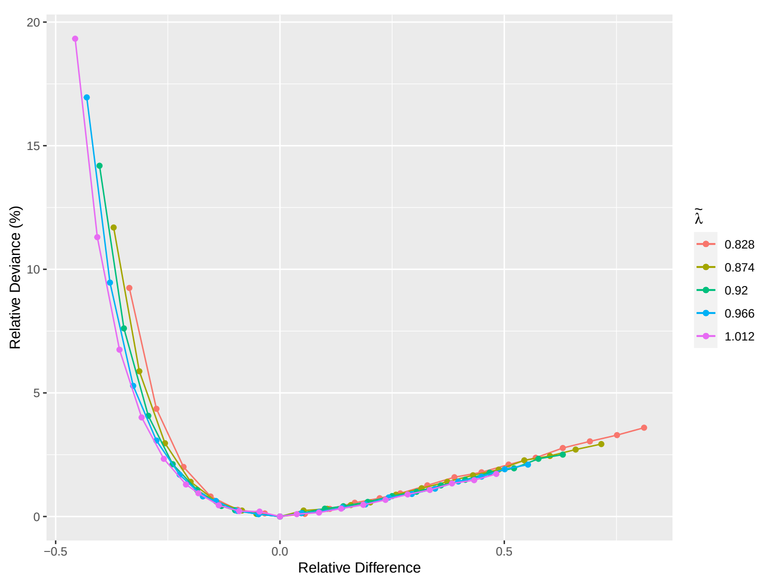


Figure 10.10: The effect of the correct parameter when the decision maker's parameter is fixed

Figure 10.11: The effect of the correct and incorrect parameters for Scenario 3

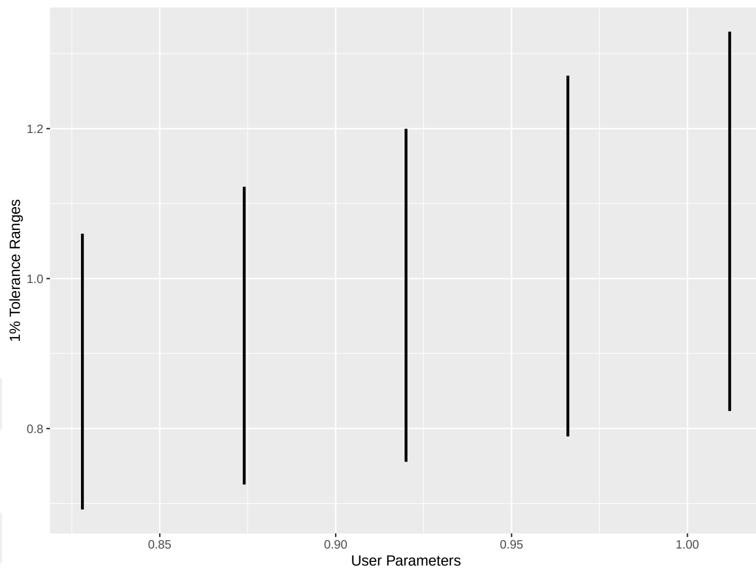


Figure 10.12: 1% Tolerance regions for five different user parameters (a)

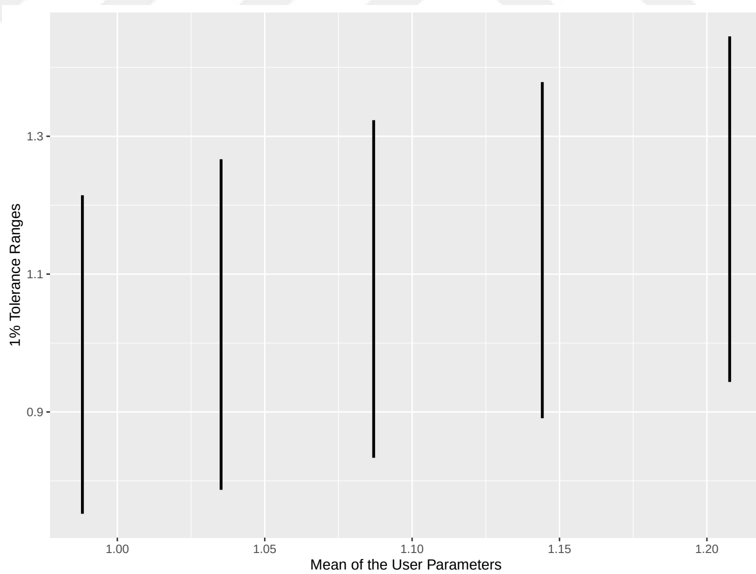


Figure 10.13: 1% Tolerance regions for five different user parameters (b)

Figure 10.14: 1% Tolerance regions for five different user parameters

The first plot in Figure 10.14 displays the 1% tolerance interval for λ parameter. The second plot pictures a similar result in the perspective of the mean of the disorder time, which is $1/\lambda$. Tolerance ranges for these five parameters approximately have the same length, so decision maker can expect a consistent tolerance interval for their choice of λ on the proximity of these values.

Chapter 11

Conclusion

For this thesis, we revisited compound Poisson and misspecified Wiener disorder problems, where the decision maker aims to detect the change in process parameters as quickly as possible.

For the compound Poisson problem, we relaxed the previous assumptions on the prior distribution of the disorder time. Previously, Dayanık and Sezer [14] focused on an exponential prior, whereas Ürü et al. [24] focused on a uniform prior. In our problem, a general prior was considered. To solve the problem, successive approximations and time discretization were used where time discretization offered a faster and a more efficient solution. However, there is a strong assumption that the pre- and post-disorder parameters are known. For future studies, this assumption may also be relaxed by allowing for the randomness of these parameters or considering interval estimations for them. Moreover, like the Wiener disorder problem, a misspecifaction analysis can be hold for the compound Poisson case, as well.

For the Wiener problem, we worked on an asset-selling problem, where the underlying stock was modeled by a geometric Brownian motion. A misspecification analysis was conducted for this problem, where we allowed the decision maker to enter certain problem parameters incorrectly. In this case, it was observed that the process used to solve the original problem also becomes an Itô process under our auxiliary probability measure. For the solution of the original problem, we used Markov chain approximation and a jump operator, and for the incorrect problem, we used Monte Carlo simulations. It was assumed that the drift parameter is constant before the disorder and stays constant after it changes with the disorder. However, stocks usually behave unpredictable and it is generally not possible to model them with fixed values for a long period of time. For future studies, a staged formulation of the problem can be considered where the decision maker updates these parameters.

Bibliography

[1] M. Basseville and I. V. Nikiforov, *Detection of Abrupt Changes - Theory and Application.* Prentice Hall, Inc. - http://people.irisa.fr/Michele.Basseville/kniga/, 1993.

[2] H. V. Poor and O. Hadjiliadis, *Quickest Detection.* Cambridge University Press, 2008.

[3] A. Tartakovsky, I. Nikiforov, and M. Basseville, *Sequential Analysis: Hypothesis Testing and Changepoint Detection.* 08 2014.

[4] A. N. Shiryaev, "On optimum methods in quickest detection problems," *Theory of Probability & Its Applications*, vol. 8, no. 1, pp. 22–46, 1963.

[5] G. Peskir and A. Shiryaev, *Optimal stopping and free-boundary problems.* Springer, 2006.

[6] A. N. Shiryaev, *Optimal stopping rules*, vol. 8. Springer Science & Business Media, 2007.

[7] L. I. Gal'chuk and B. L. Rozovskii, "The "Disorder" Problem for a Poisson Process," *Theory of Probability & Its Applications*, vol. 16, no. 4, pp. 712–716, 1971.

[8] M. Davis, "A note on the Poisson disorder problem," *Banach Center Publications*, vol. 1, no. 1, pp. 65–72, 1976.

[9] G. Peskir and A. N. Shiryaev, *Solving the Poisson Disorder Problem*, pp. 295–312. Berlin, Heidelberg: Springer Berlin Heidelberg, 2002.

[10] E. Bayraktar, S. Dayanik, and I. Karatzas, "The standard Poisson disorder problem revisited," *Stochastic Processes and their Applications*, vol. 115, no. 9, pp. 1437–1450, 2005.

[11] E. Bayraktar, S. Dayanik, and I. Karatzas, "Adaptive Poisson disorder problem," *The Annals of Applied Probability*, vol. 16, no. 3, pp. 1190 – 1261, 2006.

[12] B. Buonaguidi, "The disorder problem for purely jump Lévy processes with completely monotone jumps," *Journal of Statistical Planning and Inference*, vol. 205, pp. 203–218, 2020.

[13] P. V. Gapeev, "The disorder problem for compound Poisson processes with exponential jumps," *The Annals of Applied Probability*, vol. 15, feb 2005.

[14] S. Dayanik and S. Sezer, "Compound Poisson Disorder Problem," *Math. Oper. Res.*, vol. 31, pp. 649–672, 11 2006.

[15] U. Gugerli, "Optimal stopping of a piecewise-deterministic Markov process," *Stochastics*, vol. 19, no. 4, pp. 221–236, 1986.

[16] M. Davis, *Markov models and Optimization.* Chapman and Hall, 1993.

[17] B. Buonaguidi, A. Mira, H. Bucheli, and V. Vitanis, "Bayesian Quickest Detection of Credit Card Fraud," *Bayesian Analysis*, vol. 17, no. 1, pp. 261 – 290, 2022.

[18] E. Bayraktar and S. Sezer, "Online change detection for a Poisson process with a phase-type change-time prior distribution," *Sequential Analysis*, vol. 28, no. 2, pp. 218–250, 2009.

[19] M. F. Neuts, *Structured Stochastic Matrices of M/G/1 Type and Their Applications.* CRC Press, 1st ed., 1989.

[20] T. Altiok, "On the Phase-Type Approximations of General Distributions," *IIE Transactions*, vol. 17, no. 2, pp. 110–116, 1985.

[21] M. Zhitlukhin and A. Shiryaev, "Bayesian disorder problems on filtered probability spaces," *Theory of Probability & Its Applications*, vol. 57, no. 3, pp. 497–511, 2013.

[22] M. Zhitlukhin and A. Shiryaev, "Optimal stopping problems for a Brownian motion with disorder on a segment," *Theory of Probability & Its Applications*, vol. 58, no. 1, pp. 164–171, 2014.

[23] A. Sokko, "Disorder problem for a Brownian motion on a segment in the case of uniformly distributed moment of disorder," *Moscow University Mathematics Bulletin*, vol. 70, no. 3, pp. 103–110, 2015.

[24] C. Uru, S. Dayanik, and S. O. Sezer, "Compound Poisson disorder problem with uniformly distributed disorder time," *Bernoulli*, vol. 29, no. 3, pp. 2272–2294, 2023.

[25] T. Herberts and U. Jensen, "Optimal detection of a change point in a Poisson process for different observation schemes," *Scandinavian Journal of Statistics*, vol. 31, no. 3, pp. 347–366, 2004.

[26] M. Brown, "Bayesian detection of changes of a Poisson process monitored at discrete time points where the arrival rates are unknown," *Sequential Analysis*, vol. 27, no. 1, pp. 68–77, 2008.

[27] M. Brown, "Detecting changes in a Poisson process monitored at random time intervals," *Sequential Analysis*, vol. 35, no. 3, pp. 358–369, 2016.

[28] N. El Karoui, S. Loisel, and Y. Salhi, "Minimax optimality in robust detection of a disorder time in doubly-stochastic Poisson processes," 2017.

[29] J. E. Figueroa-López and S. Ólafsson, "Change-point detection for Lévy processes," *The Annals of Applied Probability*, vol. 29, no. 2, pp. 717–738, 2019.

[30] A. Shiryaev, "Some exact formulas in the "disorder" problem," *Theory of Probability & Its Applications*, vol. 10, pp. 348–354, 1965.

[31] A. Shiryaev, *Optimal Stopping Rules.* New York: Springer-Verlag, 1978.

[32] M. Beibel, "Sequential change-point detection in continuous time when the post-change drift is unknown," *Bernoulli*, vol. 3, no. 4, pp. 457–478, 1997.

[33] M. Beibel, "A note on sequential detection with exponential penalty for the delay," *The Annals of Statistics*, vol. 28, no. 6, pp. 1696–1701, 2000.

[34] P. Gapeev and G. Peskir, "The Wiener disorder problem with finite horizon," *Stochastic Processes and their Applications*, vol. 116, pp. 1770–1791, 2006.

[35] S. Dayanik, H. V. Poor, and S. O. Sezer, "Multisource Bayesian sequential change detection," *The Annals of Applied Probability*, vol. 18, Apr. 2008.

[36] S. O. Sezer, "On the Wiener disorder problem," *The Annals of Applied Probability*, vol. 20, no. 4, pp. 1537–1566, 2010.

[37] S. Dayanik, "Wiener disorder problem with observations at fixed discrete time epochs," *Mathematics of Operations Research*, vol. 35, no. 4, pp. 756–785, 2010.

[38] M. Beibel and H. R. Lerche, "A new look at optimal stopping problems related to mathematical finance," *Statistica Sinica*, vol. 7, no. 1, pp. 93–108, 1997.

[39] A. Shiryaev and A. A. Novikov, "On a stochastic version of the trading rule "buy and hold"," *Statistics & Decisions*, vol. 26, no. 4, pp. 289–302, 2009.

[40] E. Ekström and C. Lindberg, "Optimal closing of a momentum trade," *Journal of Applied Probability*, vol. 50, no. 2, pp. 374–387, 2013.

[41] A. N. Shiryaev and M. Zhitlukhin, "Optimal stopping problems for a Brownian motion with a disorder on a finite interval," *arXiv preprint arXiv:1205.2950*, 2012.

[42] A. N. Shiryaev, M. Zhitlukhin, and W. T. Ziemba, "When to sell Apple and the NASDAQ? Trading bubbles with a stochastic disorder model," *Journal of Portfolio Management*, vol. 40, no. 2, p. 54, 2014.

[43] A. N. Shiryaev, M. V. Zhitlukhin, and W. T. Ziemba, "Land and stock bubbles, crashes and exit strategies in Japan circa 1990 and in 2013," *Quantitative Finance*, vol. 15, no. 9, pp. 1449–1469, 2015.

[44] M. V. Zhitlukhin and W. T. Ziemba, "Exit strategies in bubble-like markets using a changepoint model," *Quantitative Finance Letters*, vol. 4, no. 1, pp. 47–52, 2016.

[45] A. A. Muravlev and A. N. Shiryaev, "Two-sided disorder problem for a Brownian motion in a Bayesian setting," *Proceedings of the Steklov Institute of Mathematics*, vol. 287, no. 1, pp. 202–224, 2015.

[46] E. Ekström and J. Vaicenavicius, "Monotonicity and robustness in Wiener disorder detection," *Sequential Analysis*, vol. 38, no. 1, pp. 57–68, 2019.

[47] C. Blanchet-Scalliet, A. Diop, R. Gibson, D. Talay, and E. Tanré, "Technical analysis techniques versus mathematical models: Boundaries of their validity domains," in *Monte Carlo and Quasi-Monte Carlo Methods 2004* (H. Niederreiter and D. Talay, eds.), (Berlin, Heidelberg), pp. 15–30, Springer Berlin Heidelberg, 2006.

[48] C. Blanchet-Scalliet, A. Diop, R. Gibson, D. Talay, and E. Tanré, "Technical analysis compared to mathematical models based methods under parameters mis-specification," *Journal of Banking & Finance*, vol. 31, no. 5, pp. 1351–1373, 2007.

[49] Kushner, H. J. and Dupuis, P., *Numerical Methods for Stochastic Control Problems in Continuous Time*, vol. 24 of *Applications of Mathematics (New York)*. New York: Springer-Verlag, second ed., 2001. Stochastic Modelling and Applied Probability.

[50] M. Ludkovski and S. O. Sezer, "Finite horizon decision timing with partially observable Poisson processes," *Stochastic Models*, vol. 28, no. 2, pp. 207–247, 2012.

[51] P. Brémaud, "Point processes and queues," *Springer*, 1981.

Appendix A

Detailed Calculations for the Compound Poisson Disorder Problem

A.1 Calculation of the Bayes Risk

The false alarm frequency is

$$
\begin{aligned}
\mathbb{P}\{\theta > \tau\} &= \mathbb{E}[1_{\{\theta>\tau\}}] = \mathbb{E}_0[Z_{\tau\wedge\theta}1_{\{\theta>\tau\}}] \\
&= \mathbb{E}_0[Z_\tau 1_{\{\theta>\tau\}}] = \mathbb{P}_0\{\theta > \tau\} \\
&= \mathbb{E}_0[(1-\pi)\int_\tau^\infty f_\theta(t)dt] = \mathbb{E}_0[(1-\pi)(1-\int_0^\tau f_\theta(t)dt)] \\
&= 1-\pi-(1-\pi)\mathbb{E}_0\int_0^\tau f_\theta(t)dt,
\end{aligned}
$$

and the expected delay is

$$
\begin{aligned}
\mathbb{E}[(\tau-\theta)^+] = \mathbb{E}\Big[1_{\{\tau\geq\theta\}}\int_\theta^\tau dt\Big] &= \mathbb{E}\left[\int_0^\infty 1_{\{\theta\leq t\}}1_{\{\tau>t\}}\ dt\right] \\
&= \int_0^\infty \mathbb{E}[1_{\{\theta\leq t\}}1_{\{\tau>t\}}]\ dt = \int_0^\infty \mathbb{E}_0[Z_t 1_{\{\theta\leq t\}}1_{\{\tau>t\}}]dt \\
&= \int_0^\infty \mathbb{E}_0[\mathbb{E}_0[Z_t 1_{\{\theta\leq t\}}1_{\{\tau>t\}}|\mathcal{F}_t]]\ dt = \int_0^\infty \mathbb{E}_0[1_{\{\tau>t\}}\mathbb{E}_0[Z_t 1_{\{\theta\leq t\}}|\mathcal{F}_t]]dt
\end{aligned}
$$

$$= \mathbb{E}_0\left[\int_0^\infty 1_{\{\tau>t\}}\mathbb{E}_0[Z_t 1_{\{\theta\le t\}}|\mathcal{F}_t]\ dt\right] = \mathbb{E}_0\left[\int_0^\tau \mathbb{E}_0[Z_t 1_{\{\theta\le t\}}|\mathcal{F}_t]\ dt\right]$$
$$= (1-\pi)\mathbb{E}_0\int_0^\tau \frac{1}{1-\pi}\mathbb{E}_0[Z_t 1_{\{\theta\le t\}}|\mathcal{F}_t]dt.$$

In this case, the Bayesian risk becomes

$$\begin{aligned}
\mathcal{R}_\tau(\pi) &= \mathbb{P}\{\theta>\tau\} + \alpha\mathbb{E}[(\tau-\theta)^+] \\
&= 1-\pi-(1-\pi)\mathbb{E}_0\int_0^\tau f_\theta(t)dt + \\
&\qquad \alpha(1-\pi)\mathbb{E}_0\int_0^\tau \frac{1}{1-\pi}\mathbb{E}_0[Z_t 1_{\{\theta\le t\}}|\mathcal{F}_t]dt \\
&= (1-\pi)+\alpha(1-\pi)\mathbb{E}_0\int_0^\tau \Big(\frac{1}{1-\pi}\mathbb{E}_0[Z_t 1_{\{\theta\le t\}}|\mathcal{F}_t] - \frac{1}{\alpha}f_\theta(t)\Big)dt.
\end{aligned}$$

A.2 Dynamics of Q and C

Recall that,

$$Q_t = R_t\Big[q+\int_0^t f_\theta(C_u)\frac{1}{R_u}\ du\Big] \quad \text{and} \quad C_t = c+t,$$

where $q = Q_0, c = C_0$.
For $t = W_i$

$$\begin{aligned}
dQ_t^{\text{jump}} &= Q_t - Q_{t-} \\
&= R_t\Big[q+\int_0^t f_\theta(C_u)\frac{1}{R_u}\ du\Big] - R_{t-}\Big[q+\int_0^t f_\theta(C_u)\frac{1}{R_u}\ du\Big] \\
&= R_{t-}\Big[\frac{\lambda_1}{\lambda_0}h(Y_i)\Big]\Big[q+\int_0^t f_\theta(C_u)\frac{1}{R_u}\ du\Big] - R_{t-}\Big[q+\int_0^t f_\theta(C_u)\frac{1}{R_u}\ du\Big] \\
&= R_{t-}\Big[q+\int_0^t f_\theta(C_u)\frac{1}{R_u}\ du\Big]\Big[\frac{\lambda_1}{\lambda_0}h(Y_i)-1\Big] \\
&= Q_{t-}\Big[\frac{\lambda_1}{\lambda_0}h(Y_i)-1\Big]
\end{aligned}$$

For $W_i \le t < W_{i+1}$

$$dQ_t = (dR_t)\Big[q+\int_0^t f_\theta(C_u)\frac{1}{R_u}\ du\Big] + R_t d\Big[q+\int_0^t f_\theta(C_u)\frac{1}{R_u}\ du\Big]$$

$$
\begin{aligned}
&= -(\lambda_1 - \lambda_0) R_t dt \Big[q + \int_0^t f_\theta(C_u) \frac{1}{R_u}\ du \Big] + R_t \Big[f_\theta(C_t) \frac{1}{R_t} \Big] dt \\
&= -(\lambda_1 - \lambda_0) Q_t dt + f_\theta(C_t) dt \\
&= [f_\theta(C_t) - (\lambda_1 - \lambda_0) Q_t] dt.
\end{aligned}
$$

Hence

$$
dQ_t = \begin{cases} [f_\theta(C_t) - (\lambda_1 - \lambda_0) Q_t] dt, & \text{if } t \in (W_i, W_{i+1}), \\ Q_{t-}\Big[\frac{\lambda_1}{\lambda_0} h(Y_{i+1}) - 1\Big], & \text{if } t = W_{i+1}, \end{cases}
$$

and

$$
dC_t = dt.
$$

When $W_i < t < W_{i+1}$

$$
\begin{aligned}
&\frac{dQ_t}{dt} = f_\theta(C_t) - (\lambda_1 - \lambda_0) Q_t \\
&Q_t' + (\lambda_1 - \lambda_0) Q_t = f_\theta(C_t) \\
&d(Q_t e^{(\lambda_1 - \lambda_0) t}) = e^{(\lambda_1 - \lambda_0) t} f_\theta(C_t)
\end{aligned}
$$

if we integrate on $[W_i, t]$,

$$
Q_t e^{(\lambda_1 - \lambda_0) t} - Q_{W_i} e^{(\lambda_1 - \lambda_0) W_i} = \begin{cases} \int_{W_i}^t e^{(\lambda_1 - \lambda_0) u} f_\theta(C_u) du & \text{if } \lambda_0 \neq \lambda_1 \\ \int_{W_i}^t f_\theta(C_u) du & \text{if } \lambda_0 = \lambda_1 \end{cases}
$$

and

$$
Q_t = \begin{cases} Q_{W_i} e^{-(\lambda_1 - \lambda_0)(t - W_i)} + e^{-(\lambda_1 - \lambda_0) t} \int_{W_i}^t e^{(\lambda_1 - \lambda_0) u} f_\theta(C_u) du, & \text{if } \lambda_0 \neq \lambda_1, \\ Q_{W_i} + \int_{W_i}^t f_\theta(C_u) du, & \text{if } \lambda_0 = \lambda_1, \end{cases}
$$

or equivalently

$$
Q_t = \begin{cases} e^{-(\lambda_1 - \lambda_0)(t - W_i)} \Big[Q_{W_i} + \int_0^{t - W_i} e^{(\lambda_1 - \lambda_0) u} f_\theta(C_{W_i} + u) du \Big], & \text{if } \lambda_0 \neq \lambda_1, \\ Q_{W_i} + \int_0^{t - W_i} f_\theta(C_{W_i} + u) du, & \text{if } \lambda_0 = \lambda_1. \end{cases}
$$

When $t = W_{i+1}$

$$
dQ_t = Q_{t-} \Big[\frac{\lambda_1}{\lambda_0} h(Y_{i+1}) - 1 \Big]
$$

$$Q_{W_{i+1}} = Q_{W_{i+1}-}\Big[\frac{\lambda_1}{\lambda_0}h(Y_{i+1})\Big]$$

$$Q_{W_{i+1}} = \begin{cases} \frac{\lambda_1}{\lambda_0}h(Y_{i+1})\Big[e^{-(\lambda_1-\lambda_0)(W_{i+1}-W_i)}\Big(Q_{W_i} + \int_0^{W_{i+1}-W_i} e^{(\lambda_1-\lambda_0)u} f_\theta(C_{W_i}+u)du\Big)\Big], \\ \text{if } \lambda_0 \neq \lambda_1, \\ \\ h(Y_{i+1})\Big[Q_{W_i} + \int_0^{W_{i+1}-W_i} f_\theta(C_{W_i}+u)du\Big], \\ \text{if } \lambda_0 = \lambda_1. \end{cases}$$

A.3 Proof of Lemma 2

Proof. For $q_1 < q_2$ and fixed c, let $\tilde{\tau}_\epsilon$ denote an ϵ-optimal rule starting from (q_1, c), which is also feasible rule starting from (q_2, c). We have

$$U^{(\bar{C})}(q_2, c) \leq \mathbb{E}_0^{(q_2,c)} \int_0^{\tau_\epsilon} Q_t \ dt - \frac{1}{\alpha}\mathbb{E}_{0,c}\int_0^{\tau_\epsilon} f_\theta(C_t) \ dt,$$
$$\mathbb{E}_0^{(q_1,c)} \int_0^{\tau_\epsilon} Q_t \ dt - \frac{1}{\alpha}\mathbb{E}_{0,c}\int_0^{\tau_\epsilon} f_\theta(C_t) \ dt \leq U^{(\bar{C})}(q_1, c) + \epsilon.$$

Thus we have

$$\begin{aligned} &U^{(\bar{C})}(q_2, c) - U^{(\bar{C})}(q_1, c) \leq \mathbb{E}_0^{(q_2,c)} \int_0^{\tau_\epsilon} Q_t \ dt - \mathbb{E}_0^{(q_1,c)} \int_0^{\tau_\epsilon} Q_t \ dt \\ &= \mathbb{E}_0 \int_0^{\tau_\epsilon} R_t(q_2 + \int_0^t \frac{1}{R_u} f_\theta(c+u) \ du) \ dt - \mathbb{E}_0 \int_0^{\tau_\epsilon} R_t(q_1 + \int_0^t \frac{1}{R_u} f_\theta(c+u) \ du) \ dt \\ &= \mathbb{E}_0 \int_0^{\tau_\epsilon} R_t(q_2 - q_1) \ dt = (q_2 - q_1)\mathbb{E}_0 \int_0^{\tau_\epsilon} R_t \ dt \leq (q_2 - q_1)\mathbb{E}_0 \int_0^{\bar{C}} R_t \ dt \\ &= (q_2 - q_1)\int_0^{\bar{C}} \mathbb{E}_0[R_t] \ dt = (q_2 - q_1)\bar{C}. \end{aligned}$$

Where we have used Fubini's theorem and the martingale property of R_t in the last two equations. Moreover for a fixed c, $U^{(\bar{C})}(q, c)$ is non-decreasing with respect to increasing q. Thus we can write,

$$0 \leq U^{(\bar{C})}(q_2, c) - U^{(\bar{C})}(q_1, c) \leq (q_2 - q_1)\bar{C},$$

by the arbitrariness of ϵ.

For $c_1 < c_2$ and fixed q, let $\tau_\epsilon \leq \bar{C} - c_1$ denote an ϵ-optimal rule starting from (q, c_1), then $\tau_\epsilon \wedge (\bar{C} - c_2)$ is also a feasible rule starting from (q, c_2). We have

$$U^{(\bar{C})}(q, c_2) \leq \mathbb{E}_0^{(q,c_2)} \int_0^{\tau_\epsilon \wedge (\bar{C} - c_2)} (Q_t - \frac{1}{\alpha} f_\theta(C_t))dt$$
$$\mathbb{E}_0^{(q,c_1)} \int_0^{\tau_\epsilon} (Q_t - \frac{1}{\alpha} f_\theta(C_t))dt \leq U^{(\bar{C})}(q, c_1) + \epsilon.$$

Let us try to bound the difference $\mathbb{E}_0^{(q,c_2)} \int_0^{\tau_\epsilon \wedge (\bar{C}-c_2)} (Q_t - \frac{1}{\alpha} f_\theta(C_t))dt - \mathbb{E}_0^{(q,c_1)} \int_0^{\tau_\epsilon} (Q_t - \frac{1}{\alpha} f_\theta(C_t))dt$. Observe that for a deterministic time t,

$$\mathbb{E}_0^{(q,c)}[Q_t] = \mathbb{E}_{0,c}[R_t(q + \int_0^t \frac{1}{R_s} f_\theta(C_s)ds)]$$
$$= q + \int_0^t f_\theta(C_s)ds = q + \int_0^t f_\theta(c + s)ds.$$

Using the linearity of integral,

$$\mathbb{E}_0^{(q,c_2)} \int_0^{\tau_\epsilon \wedge (\bar{C}-c_2)} (Q_t - \frac{1}{\alpha} f_\theta(C_t))dt - \mathbb{E}_0^{(q,c_1)} \int_0^{\tau_\epsilon} (Q_t - \frac{1}{\alpha} f_\theta(C_t))dt$$
$$= \mathbb{E}_0^{(q,c_2)} \int_0^{\tau_\epsilon \wedge (\bar{C}-c_2)} Q_t \; dt - \mathbb{E}_0^{(q,c_1)} \int_0^{\tau_\epsilon} Q_t \; dt -$$
$$\mathbb{E}_{0,c_2} \int_0^{\tau_\epsilon \wedge (\bar{C}-c_2)} \frac{1}{\alpha} f_\theta(C_t) \; dt + \mathbb{E}_{0,c_1} \int_0^{\tau_\epsilon} \frac{1}{\alpha} f_\theta(C_t) \; dt.$$

Using the fact that $\tau_\epsilon \wedge (\bar{C} - c_2) \leq \tau_\epsilon$, non-negativity and the martingale property of the process R,

$$\mathbb{E}_0^{(q,c_2)} \int_0^{\tau_\epsilon \wedge (\bar{C}-c_2)} Q_t \; dt - \mathbb{E}_0^{(q,c_1)} \int_0^{\tau_\epsilon} Q_t \; dt$$
$$= \mathbb{E}_0\Big[\int_0^{\tau_\epsilon \wedge (\bar{C}-c_2)} R_t\Big(q + \int_0^t \frac{1}{R_u} f_\theta(c_2 + u) \; du\Big) \; dt\Big] -$$
$$\mathbb{E}_0\Big[\int_0^{\tau_\epsilon} R_t\Big(q + \int_0^t \frac{1}{R_u} f_\theta(c_1 + u) \; du\Big) \; dt\Big]$$
$$= q\Big(\mathbb{E}_0 \int_0^{\tau_\epsilon \wedge (\bar{C}-c_2)} R_t \; dt - \mathbb{E}_0 \int_0^{\tau_\epsilon} R_t \; dt\Big) +$$
$$\mathbb{E}_0\Big[\int_0^{\tau_\epsilon \wedge (\bar{C}-c_2)} R_t \int_0^t \frac{1}{R_u} f_\theta(c_2 + u) \; du \; dt\Big]$$
$$- \mathbb{E}_0\Big[\int_0^{\tau_\epsilon} R_t \int_0^t \frac{1}{R_u} f_\theta(c_1 + u) \; du \; dt\Big]$$

$$
\begin{aligned}
&\leq \mathbb{E}_0\Big[\int_0^{\tau_\epsilon} R_t \int_0^t \frac{1}{R_u} f_\theta(c_2+u)\ du\ dt\Big] - \mathbb{E}_0\Big[\int_0^{\tau_\epsilon} R_t \int_0^t \frac{1}{R_u} f_\theta(c_1+u)\ du\ dt\Big] \\
&= \mathbb{E}_0\Big[\int_0^{\tau_\epsilon} R_t \int_0^t \frac{1}{R_u}[f_\theta(c_2+u) - f_\theta(c_2+u)]\ du\ dt\Big] \\
&\leq \mathbb{E}_0\Big[\int_0^{\tau_\epsilon} R_t \int_0^t \frac{1}{R_u}|f_\theta(c_1+u) - f_\theta(c_2+u)|\ du\ dt\Big] \\
&\leq \int_0^{\bar{C}-c_1} \int_0^t |f_\theta(c_1+u) - f_\theta(c_2+u)|\ du\ dt \\
&\leq \bar{C} \int_0^{\bar{C}-c_1} |f_\theta(c_1+u) - f_\theta(c_2+u)|\ du \\
&= \bar{C} \int_{c_1}^{\bar{C}} |f_\theta(u) - f_\theta(u-(c_1-c_2))|\ du \\
&\leq \bar{C} \int_0^{\bar{C}} |f_\theta(u) - f_\theta(u-(c_1-c_2))|\ du.
\end{aligned}
$$

Moreover using the fact that $c_1 < c_2$,

$$
\begin{aligned}
&- \mathbb{E}_{0,c_2} \int_0^{\tau_\epsilon \wedge (\bar{C}-c_2)} \frac{1}{\alpha} f_\theta(C_t)\ dt + \mathbb{E}_{0,c_1} \int_0^{\tau_\epsilon} \frac{1}{\alpha} f_\theta(C_t)\ dt \\
&= -\frac{1}{\alpha}\mathbb{E}_{0,c_2} \int_0^{\tau_\epsilon \wedge (\bar{C}-c_2)} f_\theta(C_t)\ dt + \frac{1}{\alpha}\mathbb{E}_{0,c_1} \int_0^{\tau_\epsilon} f_\theta(C_t)\ dt \\
&= -\frac{1}{\alpha}\mathbb{E}_0 \int_0^{\tau_\epsilon \wedge (\bar{C}-c_2)} f_\theta(c_2+t)\ dt + \frac{1}{\alpha}\mathbb{E}_0 \int_0^{\tau_\epsilon} f_\theta(c_1+t)\ dt \\
&= -\frac{1}{\alpha}\mathbb{E}_0 \left[\int_0^{\tau_\epsilon \wedge (\bar{C}-c_2)} f_\theta(c_2+t)\ dt - \int_0^{\tau_\epsilon} f_\theta(c_1+t)\ dt \right] \\
&= -\frac{1}{\alpha}\mathbb{E}_0 \left[\int_{c_2}^{(\tau_\epsilon+c_2) \wedge \bar{C}} f_\theta(t)\ dt - \int_{c_1}^{\tau_\epsilon+c_1} f_\theta(t)\ dt \right] \\
&= -\frac{1}{\alpha}\mathbb{E}_0 \left[F_\theta((\tau_\epsilon+c_2)\wedge\bar{C}) - F_\theta(c_2) - F_\theta(\tau_\epsilon+c_1) + F_\theta(c_1) \right] \\
&= \frac{1}{\alpha}\mathbb{E}_0 \left[F_\theta(c_2) - F_\theta(c_1) + F_\theta(\tau_\epsilon+c_1) - F_\theta((\tau_\epsilon+c_2)\wedge\bar{C}) \right] \\
&\leq \frac{1}{\alpha}[F_\theta(c_2) - F_\theta(c_1)].
\end{aligned}
$$

For $c_1 < c_2$ and fixed q, let $\tilde{\tau}_\epsilon \leq \bar{C} - c_2$ denote an ϵ-optimal rule starting from (q, c_2), then $\tilde{\tau}_\epsilon$ is also a feasible rule starting from (q, c_1). We have

$$
\mathbb{E}_0^{(q,c_2)} \int_0^{\tilde{\tau}_\epsilon} g(Q_t, C_t)\ dt \leq U^{(\bar{C})}(q, c_2) + \epsilon,
$$

$$U^{(\bar{C})}(q,c_1) \le \mathbb{E}_0^{(q,c_1)} \int_0^{\tilde{\tau}_\epsilon} g(Q_t,C_t)\,dt,$$

hence

$$\mathbb{E}_0^{(q,c_2)} \int_0^{\tilde{\tau}_\epsilon} g(Q_t,C_t)\,dt - \mathbb{E}_0^{(q,c_1)} \int_0^{\tilde{\tau}_\epsilon} g(Q_t,C_t)\,dt \le U^{(\bar{C})}(q,c_2) - U^{(\bar{C})}(q,c_1) + \epsilon.$$

Let us try to find a lower bound for the term on the left.

$$\begin{aligned}
&\mathbb{E}_0^{(q,c_2)} \int_0^{\tilde{\tau}_\epsilon} g(Q_t,C_t)\,dt - \mathbb{E}_0^{(q,c_1)} \int_0^{\tilde{\tau}_\epsilon} g(Q_t,C_t)\,dt \\
&= \mathbb{E}_0^{(q,c_2)} \int_0^{\tilde{\tau}_\epsilon} Q_t\,dt - \frac{1}{\alpha}\mathbb{E}_{0,c_2} \int_0^{\tilde{\tau}_\epsilon} f_\theta(C_t)\,dt \\
&\qquad - \mathbb{E}_0^{(q,c_1)} \int_0^{\tilde{\tau}_\epsilon} Q_t\,dt + \frac{1}{\alpha}\mathbb{E}_{0,c_1} \int_0^{\tilde{\tau}_\epsilon} f_\theta(C_t)\,dt.
\end{aligned}$$

Firstly,

$$\begin{aligned}
&\mathbb{E}_0^{(q,c_2)} \int_0^{\tilde{\tau}_\epsilon} Q_t\,dt - \mathbb{E}_0^{(q,c_1)} \int_0^{\tilde{\tau}_\epsilon} Q_t\,dt \\
&= \mathbb{E}_0 \int_0^{\tilde{\tau}_\epsilon} R_t\Big[q + \int_0^t f_\theta(c_2+u)\frac{1}{R_u}\,du\Big]dt - \mathbb{E}_0 \int_0^{\tilde{\tau}_\epsilon} R_t\Big[q + \int_0^t f_\theta(c_1+u)\frac{1}{R_u}\,du\Big]dt \\
&= \mathbb{E}_0 \int_0^{\tilde{\tau}_\epsilon} R_t \int_0^t \frac{1}{R_u}[f_\theta(c_2+u) - f_\theta(c_1+u)]\,du\,dt \\
&= \mathbb{E}_0 \int_0^{\tilde{\tau}_\epsilon} \int_0^t \frac{R_t}{R_u}[f_\theta(c_2+u) - f_\theta(c_1+u)]^+\,du\,dt - \\
&\qquad \mathbb{E}_0 \int_0^{\tilde{\tau}_\epsilon} \int_0^t \frac{R_t}{R_u}[f_\theta(c_2+u) - f_\theta(c_1+u)]^-\,du\,dt \\
&\ge -\mathbb{E}_0 \int_0^{\tilde{\tau}_\epsilon} \int_0^t \frac{R_t}{R_u}[f_\theta(c_2+u) - f_\theta(c_1+u)]^-\,du\,dt \\
&\ge -\mathbb{E}_0 \int_0^{\bar{C}-c_2} \int_0^t \frac{R_t}{R_u}[f_\theta(c_2+u) - f_\theta(c_1+u)]^-\,du\,dt \\
&= -\int_0^{\bar{C}-c_2} \int_0^t [f_\theta(c_2+u) - f_\theta(c_1+u)]^-\,du\,dt \\
&\ge -\bar{C} \int_0^{\bar{C}-c_2} [f_\theta(c_2+u) - f_\theta(c_1+u)]^-\,du.
\end{aligned}$$

Where the last equality is because of the Fubini's theorem and martingale property of R. Secondly,

$$\frac{1}{\alpha}\mathbb{E}_{0,c_1} \int_0^{\tilde{\tau}_\epsilon} f_\theta(C_t)\,dt - \frac{1}{\alpha}\mathbb{E}_{0,c_2} \int_0^{\tilde{\tau}_\epsilon} f_\theta(C_t)\,dt$$

$$
\begin{aligned}
&= \frac{1}{\alpha}\mathbb{E}_0\Big[\int_0^{\tilde{\tau}_\epsilon} f_\theta(c_1+t) - f_\theta(c_2+t)\; dt\Big] \\
&= \frac{1}{\alpha}\mathbb{E}_0\Big[\int_0^{\tilde{\tau}_\epsilon} [f_\theta(c_1+t) - f_\theta(c_2+t)]^+\; dt\Big] - \frac{1}{\alpha}\mathbb{E}_0\Big[\int_0^{\tilde{\tau}_\epsilon} [f_\theta(c_1+t) - f_\theta(c_2+t)]^-\; dt\Big] \\
&\geq -\frac{1}{\alpha}\mathbb{E}_0\Big[\int_0^{\tilde{\tau}_\epsilon} [f_\theta(c_1+t) - f_\theta(c_2+t)]^-\; dt\Big] \\
&\geq -\frac{1}{\alpha}\mathbb{E}_0\Big[\int_0^{\bar{C}-c_2} [f_\theta(c_1+t) - f_\theta(c_2+t)]^-\; dt\Big] \\
&= -\frac{1}{\alpha}\int_0^{\bar{C}-c_2} [f_\theta(c_1+t) - f_\theta(c_2+t)]^-\; dt.
\end{aligned}
$$

Note that

$$
\begin{aligned}
&-\bar{C}\int_0^{\bar{C}-c_2} [f_\theta(c_2+u) - f_\theta(c_1+u)]^-\; du - \frac{1}{\alpha}\int_0^{\bar{C}-c_2} [f_\theta(c_1+u) - f_\theta(c_2+u)]^-\; du \\
&\geq \min\{-\bar{C}, -\frac{1}{\alpha}\}\Big[\int_0^{\bar{C}-c_2} [f_\theta(c_2+u) - f_\theta(c_1+u)]^-\; du + \\
&\qquad \int_0^{\bar{C}-c_2} [f_\theta(c_1+u) - f_\theta(c_2+u)]^-\; du\Big] \\
&= \min\{-\bar{C}, -\frac{1}{\alpha}\}\int_0^{\bar{C}-c_2} |f_\theta(c_2+u) - f_\theta(c_1+u)|\; du \\
&= \min\{-\bar{C}, -\frac{1}{\alpha}\}\int_{c_2}^{\bar{C}} |f_\theta(u) - f_\theta(u-(c_2-c_1))|\; du \\
&\geq \min\{-\bar{C}, -\frac{1}{\alpha}\}\int_0^{\bar{C}} |f_\theta(u) - f_\theta(u-(c_2-c_1))|\; du.
\end{aligned}
$$

Combining these two inequalities, by the arbitrariness of ϵ we get

$$
\begin{aligned}
&\min\{-\bar{C}, -\frac{1}{\alpha}\}\int_0^{\bar{C}} |f_\theta(u) - f_\theta(u-(c_2-c_1))|\; du \leq U^{(\bar{C})}(q, c_2) - U^{(\bar{C})}(q, c_1) \\
&\leq \bar{C}\int_0^{\bar{C}} |f_\theta(u) - f_\theta(u-(c_1-c_2))|\; du + \frac{1}{\alpha}[F_\theta(c_2) - F_\theta(c_1)], \quad \text{for } c_1 < c_2.
\end{aligned}
$$

Observe that,

$$
\lim_{\delta\to 0}\int_0^{\bar{C}} |f_\theta(t) - f_\theta(t-\delta)|\; dt = 0,
$$

hence expressions $\int_0^{\bar{C}} |f_\theta(u) - f_\theta(u-(c_2-c_1))|\; du$ and $\int_0^{\bar{C}} |f_\theta(u) - f_\theta(u-(c_1-c_2))|\; du$ goes to 0, as $c_1 \to c_2$.

□

A.4 Proof of Proposition 2

Proof. We will use a characterization of $\mathcal{F}$-stopping times by Davis [16] and Brémaud [51].

Lemma 8. *For every $\mathcal{F}$-stopping time τ and every $n \in \mathbb{N}_0$, there is an $\mathcal{F}_{W_n}$ measurable random variable $R_n : \Omega \to [0, \infty]$ such that $\tau \wedge W_{n+1} = (W_n + R_n) \wedge W_{n+1}$ $\mathbb{P}_0$-a.s. on the event $\{\tau \geq W_n\}$.*

Firstly we will show that

$$\mathbb{E}_0^{(q,c)}\Big[\int_0^{\tau\wedge\tau_{\bar{C}}\wedge W_n} g(Q_t, C_t)dt\Big] \geq v_n^{(\bar{C})}(q, c) \quad \text{for every } n \in \mathbb{N}_0. \tag{A.1}$$

If we can show

$$\mathbb{E}_0^{(q,c)}\Big[\int_0^{\tau\wedge\tau_{\bar{C}}\wedge W_n} g(Q_t, C_t)dt\Big] \geq \mathbb{E}_0^{(q,c)}\Big[\int_0^{\tau\wedge\tau_{\bar{C}}\wedge W_{n-k+1}} g(Q_t, C_t)dt + 1_{\{\tau\wedge\tau_{\bar{C}}\geq W_{n-k+1}\}} v_{k-1}^{(\bar{C})}(Q_{W_{n-k+1}}, C_{W_{n-k+1}})\Big] \tag{A.2}$$

for $k \in \mathbb{N}$, we are done. Observe that when $k = n + 1$, (A.1) holds.

We will use induction on k to prove (A.2). When $k = 1$, equality holds as $v_0^{(\bar{C})} \equiv 0$. If we denote the inequality in (A.2) as p_k, we will show that p_{k+1} also holds.

We rewrite (A.2) as

$$\begin{aligned}\mathbb{E}_0^{(q,c)}\Big[\int_0^{\tau\wedge\tau_{\bar{C}}\wedge\tau_{\bar{C}}\wedge W_n} g(Q_t, C_t)dt\Big] &\geq \mathbb{E}_0^{(q,c)}\Big[\int_0^{\tau\wedge\tau_{\bar{C}}\wedge W_{n-k}} g(Q_t, C_t)dt\Big] \\ &+ \mathbb{E}_0^{(q,c)}\Big[1_{\{\tau\wedge\tau_{\bar{C}}\geq W_{n-k}\}} \int_{W_{n-k}}^{\tau\wedge W_{n-k+1}} g(Q_t, C_t)dt\Big] \\ &+ \mathbb{E}_0^{(q,c)}\Big[1_{\{\tau\wedge\tau_{\bar{C}}\geq W_{n-k}\}} 1_{\{\tau\wedge\tau_{\bar{C}}\geq W_{n-k+1}\}} v_{k-1}^{(\bar{C})}(Q_{W_{n-k+1}}, C_{W_{n-k+1}})\Big].\end{aligned} \tag{A.3}$$

We will work on each term on the right hand side of the inequality separately. Using Lemma 8, we know that on the event $\tau \geq W_{n-k}$, there exists $\mathcal{F}_{W_{n-k}}$ measurable random variable R_{n-k} such that $\tau \wedge W_{n-k+1} = (W_{n-k} + R_{n-k}) \wedge W_{n-k+1}$

$\mathbb{P}_0$-a.s. Observe that we can write $W_{n-k+1} = (W_1 \circ \theta_{W_{n-k}}) + W_{n-k}$ where θ is the shift operator. Moreover, using strong Markov property of the process (Q, C)

$$\begin{aligned}
&\mathbb{E}_0^{(q,c)}\Big[1_{\{\tau\wedge\tau_{\bar{C}}\geq W_{n-k}\}}\int_{W_{n-k}}^{\tau\wedge\tau_{\bar{C}}\wedge W_{n-k+1}} g(Q_t, C_t)dt\Big] \\
&= \mathbb{E}_0^{(q,c)}\Big[1_{\{\tau\wedge\tau_{\bar{C}}\geq W_{n-k}\}}\int_{W_{n-k}}^{(W_{n-k}+R_{n-k})\wedge\tau_{\bar{C}}\wedge W_{n-k+1}} g(Q_t, C_t)dt\Big] \\
&= \mathbb{E}_0^{(q,c)}\Big[1_{\{\tau\wedge\tau_{\bar{C}}\geq W_{n-k}\}}\int_{0}^{R_{n-k}\wedge(\tau_{\bar{C}}-W_{n-k})\wedge W_1\circ\theta_{W_{n-k}}} g(Q_{t+W_{n-k}}, C_{t+W_{n-k}})dt\Big] \\
&= \mathbb{E}_0^{(q,c)}\Bigg[\mathbb{E}_0^{(q,c)}\Big[1_{\{\tau\wedge\tau_{\bar{C}}\geq W_{n-k}\}} \\
&\qquad \int_{0}^{R_{n-k}\wedge(\tau_{\bar{C}}-W_{n-k})\wedge W_1\circ\theta_{W_{n-k}}} g(Q_t, C_t)\circ\theta_{W_{n-k}}dt\Big|\mathcal{F}_{W_{n-k}}\Big]\Bigg] \\
&= \mathbb{E}_0^{(q,c)}\Bigg[1_{\{\tau\wedge\tau_{\bar{C}}\geq W_{n-k}\}} \\
&\qquad \mathbb{E}_0^{(q,c)}\Big[\Big\{\int_{0}^{r\wedge(\tau_{\bar{C}}-W_{n-k})\wedge W_1} g(Q_t, C_t)dt\Big\}\circ\theta_{W_{n-k}}\Big|\mathcal{F}_{W_{n-k}}\Big]\Big|_{r=R_{n-k}}\Bigg] \\
&= \mathbb{E}_0^{(q,c)}\Bigg[1_{\{\tau\wedge\tau_{\bar{C}}\geq W_{n-k}\}}\mathbb{E}_0^{(\hat{q},\hat{c})}\Big(\int_{0}^{r\wedge\hat{\tau}_{\bar{C}}\wedge W_1} g(Q_t, C_t)dt\Big)\Big|_{r=R_{n-k},\hat{q}=Q_{W_{n-k}},\hat{c}=C_{W_{n-k}}}\Bigg].
\end{aligned}$$

Here, the updated stopping time $\hat{\tau}_{\bar{C}}$ is equal to $(\bar{C}-\hat{c})^+$.

For the next expression, observe that on the event $\tau \geq W_{n-k}$, using Lemma 8 and shift operator, we have the equivalence $\tau \geq W_{n-k+1} \Leftrightarrow R_{n-k} \geq W_1 \circ \theta_{W_{n-k}}$. Using strong Markov property, we write

$$\begin{aligned}
&\mathbb{E}_0^{(q,c)}\Big[1_{\{\tau\wedge\tau_{\bar{C}}\geq W_{n-k}\}}1_{\{\tau\wedge\tau_{\bar{C}}\geq W_{n-k+1}\}}v_{k-1}^{(\bar{C})}(Q_{W_{n-k+1}}, C_{W_{n-k+1}})\Big] \\
&= \mathbb{E}_0^{(q,c)}\Bigg[\mathbb{E}_0^{(q,c)}\Big[1_{\{\tau\wedge\tau_{\bar{C}}\geq W_{n-k}\}} \\
&\qquad 1_{\{R_{n-k}\wedge(\bar{C}-W_{n-k})\geq W_1\circ\theta_{W_{n-k}}\}}\Big\{v_{k-1}^{(\bar{C})}(Q_{W_1}, C_{W_1})\Big\}\circ\theta_{W_{n-k}}\Big|\mathcal{F}_{W_{n-k}}\Big]\Bigg] \\
&= \mathbb{E}_0^{(q,c)}\Bigg[1_{\{\tau\wedge\tau_{\bar{C}}\geq W_{n-k}\}} \\
&\qquad \mathbb{E}_0^{(q,c)}\Big[\Big\{1_{\{r\wedge(\bar{C}-W_{n-k})\geq W_1\}}v_{k-1}^{(\bar{C})}(Q_{W_1}, C_{W_1})\Big\}\circ\theta_{W_{n-k}}\Big|\mathcal{F}_{W_{n-k}}\Big]\Big|_{r=R_{n-k}}\Bigg]
\end{aligned}$$

$$= \mathbb{E}_0^{(q,c)}\Bigg[1_{\{\tau\wedge\tau_{\bar{C}}\geq W_{n-k}\}}\mathbb{E}_0^{(\hat{q},\hat{c})}\Big[1_{\{r\wedge\hat{\tau}_{\bar{C}}\geq W_1\}}v_{k-1}^{(\bar{C})}(Q_{W_1},C_{W_1})\Big]\Bigg|_{r=R_{n-k},\hat{q}=Q_{W_{n-k}},\hat{c}=C_{W_{n-k}}}\Bigg].$$

If we combine three expressions and rewrite (A.2), we have

$$\mathbb{E}_0^{(q,c)}\Big[\int_0^{\tau\wedge\tau_{\bar{C}}\wedge\tau_{\bar{C}}\wedge W_n} g(Q_t,C_t)dt\Big]\geq \mathbb{E}_0^{(q,c)}\Bigg[\int_0^{\tau\wedge\tau_{\bar{C}}\wedge W_{n-k}} g(Q_t,C_t)dt$$
$$+ 1_{\{\tau\wedge\tau_{\bar{C}}\geq W_{n-k}\}}\mathbb{E}_0^{(\hat{q},\hat{c})}\Big[\int_0^{r\wedge\hat{\tau}_{\bar{C}}\wedge W_1} g(Q_t,C_t)dt+$$
$$1_{\{r\wedge\hat{\tau}_{\bar{C}}\geq W_1\}}v_{k-1}^{(\bar{C})}(Q_{W_1},C_{W_1})\Big]\Bigg|_{r=R_{n-k},\hat{q}=Q_{W_{n-k}},\hat{c}=C_{W_{n-k}}}\Bigg].$$

Observe that, by the definition of the operators $(H_t^{(\bar{C})}w^{(\bar{C})})$ and $(H^{(\bar{C})}w^{(\bar{C})})$, and successive approximations $v_n^{(\bar{C})}$, we have

$$\mathbb{E}_0^{(\hat{q},\hat{c})}\Big[\int_0^{r\wedge\hat{\tau}_{\bar{C}}\wedge W_1} g(Q_t,C_t)dt + 1_{\{r\wedge\hat{\tau}_{\bar{C}}\geq W_1\}}v_{k-1}^{(\bar{C})}(Q_{W_1},C_{W_1})\Big]\Bigg|_{r=R_{n-k},\hat{q}=Q_{W_{n-k}},\hat{c}=C_{W_{n-k}}}$$
$$= (H_r^{(\bar{C})}v_{k-1}^{(\bar{C})})(Q_{W_{n-k}},C_{W_{n-k}}) \geq (H^{(\bar{C})}v_{k-1}^{(\bar{C})})(Q_{W_{n-k}},C_{W_{n-k}}) = v_k^{(\bar{C})}(Q_{W_{n-k}},C_{W_{n-k}}).$$

So we can write

$$\mathbb{E}_0^{(q,c)}\Big[\int_0^{\tau\wedge\tau_{\bar{C}}\wedge\tau_{\bar{C}}\wedge W_n} g(Q_t,C_t)dt\Big]\geq$$
$$\mathbb{E}_0^{(q,c)}\Big[\int_0^{\tau\wedge\tau_{\bar{C}}\wedge W_{n-k}} g(Q_t,C_t)dt + 1_{\{\tau\wedge\tau_{\bar{C}}\geq W_{n-k}\}}v_k^{(\bar{C})}(Q_{W_{n-k}},C_{W_{n-k}})\Big].$$

This inequality is indeed p_{k+1}, induction is completed. (A.2) holds, thus (A.1) holds. If we take the infimum of both sides, we get $U_n^{(\bar{C})} \geq v_n^{(\bar{C})}$, $n \in \mathbb{N}_0$.

To show the reverse inequality $U_n^{(\bar{C})} \leq v_n^{(\bar{C})}$, we will use the characterization of the stopping times $S_n^{\varepsilon,\bar{C}}$. Observe that $S_n^{\varepsilon,\bar{C}} \leq W_n$ $\mathbb{P}_0$ a.s. for $n \in \mathbb{N}_0$. Thus, we can write

$$U_n^{(\bar{C})}(q,c) \leq \mathbb{E}_0^{(q,c)}\Big[\int_0^{S_n^{\varepsilon,\bar{C}}\wedge\tau_{\bar{C}}\wedge W_n} g(Q_t,C_t)dt\Big]= \mathbb{E}_0^{(q,c)}\Big[\int_0^{S_n^{\varepsilon,\bar{C}}\wedge\tau_{\bar{C}}} g(Q_t,C_t)dt\Big].$$

Hence, if we can show the inequality

$$\mathbb{E}_0^{(q,c)}\Big[\int_0^{S_n^{\varepsilon,\bar{C}}\wedge\tau_{\bar{C}}} g(Q_t,C_t)dt\Big]\leq v_n^{(\bar{C})}(q,c) + \varepsilon \quad \forall n \in \mathbb{N}, \varepsilon > 0, \qquad \text{(A.4)}$$

we are done.

We will use induction on n. When $n=1$, we have

$$\mathbb{E}_0^{(q,c)}\Big[\int_0^{S_1^{\varepsilon,\bar{C}}\wedge\tau_{\bar{C}}} g(Q_t,C_t)dt\Big] = \mathbb{E}_0^{(q,c)}\Big[\int_0^{r_0^{\varepsilon,\bar{C}}(q,c)\wedge W_1\wedge\tau_{\bar{C}}} g(Q_t,C_t)dt\Big]$$
$$= (H^{(\bar{C})}_{r_0^{\varepsilon,\bar{C}}(q,c)} v_0^{(\bar{C})})(q,c).$$

By the construction of stopping times $r_n^{\varepsilon,\bar{C}}$, definition of the operators $(H_t^{(\bar{C})}w^{(\bar{C})})$ and $(H^{(\bar{C})}w^{(\bar{C})})$ and successive approximations $v_n^{(\bar{C})}$, we have

$$(H^{(\bar{C})}_{r_0^{\varepsilon,\bar{C}}(q,c)} v_0^{(\bar{C})})(q,c) \le (H^{(\bar{C})}v_0^{(\bar{C})})(q,c)+\varepsilon = v_1^{(\bar{C})}(q,c)+\varepsilon.$$

Hence,

$$\mathbb{E}_0^{(q,c)}\Big[\int_0^{S_1^{\varepsilon,\bar{C}}} g(Q_t,C_t)dt\Big] \le v_1^{(\bar{C})}(q,c)+\varepsilon.$$

Now assume that (A.4) holds for $\forall\varepsilon>0$ and $n\in\mathbb{N}$. We will prove that it also holds for $n+1$. Observe that $S_{n+1}^{\varepsilon,\bar{C}} = r_n^{\varepsilon/2,\bar{C}}(q,c)1_{\{W_1>r_n^{\varepsilon/2,\bar{C}}(q,c)\}} + (W_1+S_n^{\varepsilon/2,\bar{C}}\circ\theta_{W_1})1_{\{W_1\le r_n^{\varepsilon/2,\bar{C}}(q,c)\}}$. Thus, we have $S_{n+1}^{\varepsilon,\bar{C}}\wedge\tau_{\bar{C}}\wedge W_1 = r_n^{\varepsilon/2,\bar{C}}(q,c)\wedge\tau_{\bar{C}}\wedge W_1$, and using strong Markov property we can write

$$\mathbb{E}_0^{(q,c)}\Big[\int_0^{S_{n+1}^{\varepsilon,\bar{C}}\wedge\tau_{\bar{C}}} g(Q_t,C_t)dt\Big]$$
$$= \mathbb{E}_0^{(q,c)}\Big[\int_0^{S_{n+1}^{\varepsilon,\bar{C}}\wedge\tau_{\bar{C}}\wedge W_1} g(Q_t,C_t)dt + 1_{\{S_{n+1}^{\varepsilon,\bar{C}}\wedge\tau_{\bar{C}}\ge W_1\}}\int_{W_1}^{S_{n+1}^{\varepsilon,\bar{C}}\wedge\tau_{\bar{C}}} g(Q_t,C_t)dt\Big]$$
$$= \mathbb{E}_0^{(q,c)}\Big[\int_0^{r_n^{\varepsilon/2,\bar{C}}(q,c)\wedge\tau_{\bar{C}}\wedge W_1} g(Q_t,C_t)dt\Big] +$$
$$\mathbb{E}_0^{(q,c)}\left[1_{\{r_n^{\varepsilon/2,\bar{C}}(q,c)\wedge\tau_{\bar{C}}\ge W_1\}}\mathbb{E}_0^{(q,c)}\Big[\int_0^{S_n^{\varepsilon/2,\bar{C}}\circ\theta_{W_1}\wedge(\tau_{\bar{C}}-W_1)} g(Q_{t+W_1},C_{t+W_1})dt\Big|\mathcal{F}_{W_1}\Big]\right]$$
$$= \mathbb{E}_0^{(q,c)}\Big[\int_0^{r_n^{\varepsilon/2,\bar{C}}(q,c)\wedge\tau_{\bar{C}}\wedge W_1} g(Q_t,C_t)dt\Big] +$$
$$\mathbb{E}_0^{(q,c)}\left[1_{\{r_n^{\varepsilon/2,\bar{C}}(q,c)\wedge\tau_{\bar{C}}\ge W_1\}}\mathbb{E}_0^{(q,c)}\Big[\Big\{\int_0^{S_n^{\varepsilon/2,\bar{C}}\circ\theta_{W_1}\wedge(\tau_{\bar{C}}-W_1)} g(Q_t,C_t)dt\Big\}\circ\theta_{W_1}\Big|\mathcal{F}_{W_1}\Big]\right]$$
$$= \mathbb{E}_0^{(q,c)}\Big[\int_0^{r_n^{\varepsilon/2,\bar{C}}(q,c)\wedge\tau_{\bar{C}}\wedge W_1} g(Q_t,C_t)dt\Big] +$$

$$\mathbb{E}_0^{(q,c)}\left[1_{\{r_n^{\varepsilon/2,\bar{C}}(q,c)\wedge\tau_{\bar{C}}\geq W_1\}}\mathbb{E}_0^{(\hat{q},\hat{c})}\Big[\int_0^{S_n^{\varepsilon/2,\bar{C}}\wedge\hat{\tau}_{\bar{C}}} g(Q_t,C_t)dt\Big]\bigg|_{\hat{q}=Q_{W_1},\hat{c}=C_{W_1}}\right].$$

Observe that using our inductive argument on (A.4), we can state that

$$\mathbb{E}_0^{(\hat{q},\hat{c})}\Big[\int_0^{S_n^{\varepsilon/2,\bar{C}}\wedge\hat{\tau}_{\bar{C}}} g(Q_t,C_t)dt\Big]\bigg|_{\hat{q}=Q_{W_1},\hat{c}=C_{W_1}} \leq v_n^{(\bar{C})}(Q_{W_1},C_{W_1})+\varepsilon/2.$$

Hence by the construction of stopping times $r_n^{\varepsilon,\bar{C}}$, definition of the operators $(H_t^{(\bar{C})}w^{(\bar{C})})$ and $(H^{(\bar{C})}w^{(\bar{C})})$ and successive approximations $v_n^{(\bar{C})}$, we have

$$\begin{aligned}
&\mathbb{E}_0^{(q,c)}\Big[\int_0^{S_{n+1}^{\varepsilon,\bar{C}}\wedge\tau_{\bar{C}}} g(Q_t,C_t)dt\Big] \\
&\leq \mathbb{E}_0^{(q,c)}\Big[\int_0^{r_n^{\varepsilon/2,\bar{C}}(q,c)\wedge\tau_{\bar{C}}\wedge W_1} g(Q_t,C_t)dt+ \\
&\qquad 1_{\{r_n^{\varepsilon/2,\bar{C}}(q,c)\wedge\tau_{\bar{C}}\geq W_1\}}\Big(v_n^{(\bar{C})}(Q_{W_1},C_{W_1})+\varepsilon/2\Big)\Big] \\
&\leq \mathbb{E}_0^{(q,c)}\Big[\int_0^{r_n^{\varepsilon/2,\bar{C}}(q,c)\wedge\tau_{\bar{C}}\wedge W_1} g(Q_t,C_t)dt+1_{\{r_n^{\varepsilon/2,\bar{C}}(q,c)\wedge\tau_{\bar{C}}\geq W_1\}}v_n^{(\bar{C})}(Q_{W_1},C_{W_1})\Big]+\varepsilon/2 \\
&= (H^{(\bar{C})}_{r_n^{\varepsilon/2,\bar{C}}(q,c)}v_n^{(\bar{C})})(q,c)+\varepsilon/2 \leq (H^{(\bar{C})}v_n^{(\bar{C})})(q,c)+\varepsilon = v_{n+1}^{(\bar{C})}(q,c)+\varepsilon.
\end{aligned}$$

Thus

$$\mathbb{E}_0^{(q,c)}\Big[\int_0^{S_{n+1}^{\varepsilon,\bar{C}}\wedge\tau_{\bar{C}}} g(Q_t,C_t)dt\Big]\leq v_{n+1}^{(\bar{C})}(q,c)+\varepsilon, \quad \forall\varepsilon>0, n\in\mathbb{N},$$

and the induction is proved, hence $U_n^{(\bar{C})}\leq v_n^{(\bar{C})}$. We proved that $U_n^{(\bar{C})}=v_n^{(\bar{C})}$. $\square$

A.5 Variational Inequalities

Remember that

$$V(\pi)=(1-\pi)+\alpha(1-\pi)\underbrace{\inf_{\tau\in\mathcal{F}}\mathbb{E}_0^{(q,c)}\Big[\int_0^\tau Q_t-\frac{1}{\alpha}f_\theta(C_t)dt\Big]}_{U(q,c)}$$

where

$$U(q,c)\equiv\inf_{\tau\in\mathcal{F}}\mathbb{E}_0^{(q,c)}\Big[\int_0^\tau Q_t-\frac{1}{\alpha}f_\theta(C_t)dt\Big], \quad q=Q_0\in\mathbb{R}_+\ c=C_0\in\mathbb{R}_+,$$

$$U(q,c) \equiv \inf_{\tau \in \mathcal{F}} \mathbb{E}_0^{(q,c)} \Big[\int_0^{\tau} g(Q_t, C_t) dt \Big].$$

If we have not stopped before t, we can write

$$\begin{aligned} U(Q_t, C_t) &= \text{minimum cost stop action} \\ &= \min \Big\{ 0, \mathbb{E}_0^{(q,c)} \Big[\int_t^{t+dt} g(Q_s, C_s) ds + U(Q_{t+dt}, C_{t+dt}) \Big| \mathcal{F}_t \Big] \Big\}. \end{aligned}$$

Here, assuming U is continuously differentiable

$$\begin{aligned} U(Q_{t+dt}, C_{t+dt}) &= U(Q_t, C_t) + \int_t^{t+dt} \Big[\frac{\partial}{\partial q} U(Q_s, C_s) dQ_s + \frac{\partial}{\partial c} U(Q_s, C_s) dC_s \Big] \\ &\approx U(Q_t, C_t) + \frac{\partial}{\partial q} U(Q_t, C_t) dQ_t + \frac{\partial}{\partial c} U(Q_t, C_t) dC_t. \end{aligned}$$

If t is a jump point

$$U(Q_{t+dt}, C_{t+dt}) \approx U(Q_t, C_t) + \frac{\partial}{\partial q} U(Q_{t-}, C_{t-}) \Big[Q_{t-} \Big(\frac{\lambda_1}{\lambda_0} h(Y_{N_t}) - 1 \Big) \Big] N(dt),$$

if t is a continuity point

$$\begin{aligned} U(Q_{t+dt}, C_{t+dt}) &\approx U(Q_t, C_t) + \frac{\partial}{\partial q} U(Q_t, C_t) \Big[f_\theta(C_t) - (\lambda_1 - \lambda_0) Q_t \Big] dt \\ &\quad + \frac{\partial}{\partial c} U(Q_t, C_t) dt \\ &= U(Q_t, C_t) + \Bigg[\frac{\partial}{\partial q} U(Q_t, C_t) \Big[f_\theta(C_t) - (\lambda_1 - \lambda_0) Q_t \Big] \\ &\quad + \frac{\partial}{\partial c} U(Q_t, C_t) \Bigg] dt. \end{aligned}$$

Thus variational inequalities become

$$U(Q_t, C_t) \leq 0$$

$$U(Q_t, C_t) \leq \mathbb{E}_0^{(q,c)} \Bigg[g(Q_t, C_t) dt + U(Q_t, C_t) + \Bigg\{ \frac{\partial}{\partial q} U(Q_t, C_t) \Big[f_\theta(C_t) - (\lambda_1 - \lambda_0) Q_t \Big] + \frac{\partial}{\partial c} U(Q_t, C_t) \Bigg\} dt + \frac{\partial}{\partial q} U(Q_{t-}, C_{t-}) \Big[Q_{t-} \Big(\frac{\lambda_1}{\lambda_0} h(Y_{N_t}) - 1 \Big) \Big] N(dt) \Bigg| \mathcal{F}_t \Bigg].$$

$$0 \leq \mathbb{E}_0^{(q,c)} \Bigg[\Bigg\{ g(Q_t, C_t) + \frac{\partial}{\partial q} U(Q_t, C_t) \Big[f_\theta(C_t) - (\lambda_1 - \lambda_0) Q_t \Big]$$

$$+ \frac{\partial}{\partial c}U(Q_t, C_t)\Big\}dt + \frac{\partial}{\partial q}U(Q_{t-}, C_{t-})\Big[Q_{t-}\Big(\frac{\lambda_1}{\lambda_0}h(Y_{N_t}) - 1\Big)\Big]N(dt)\Big|\mathcal{F}_t\Big]$$

$$= \Big\{g(Q_t, C_t) + \frac{\partial}{\partial q}U(Q_t, C_t)\Big[f_\theta(C_t) - (\lambda_1 - \lambda_0)Q_t\Big]$$

$$+ \frac{\partial}{\partial c}U(Q_t, C_t)\Big\}dt$$

$$+ \frac{\partial}{\partial q}U(Q_{t-}, C_{t-})\Big[\int_{y\in\mathcal{Y}} Q_{t-}\Big(\frac{\lambda_1}{\lambda_0}h(y) - 1\Big)\lambda_0\nu_0(dy)\Big]dt.$$

If we divide both sides by dt and let $dt \downarrow 0$

$$0 \le g(Q_t, C_t) + \frac{\partial}{\partial q}U(Q_t, C_t)\Big[f_\theta(C_t) - (\lambda_1 - \lambda_0)Q_t\Big]$$
$$+ \frac{\partial}{\partial c}U(Q_t, C_t) + \frac{\partial}{\partial q}U(Q_{t-}, C_{t-})Q_{t-}\int_{y\in\mathcal{Y}}\Big(\frac{\lambda_1}{\lambda_0}h(y) - 1\Big)\lambda_0\nu_0(dy).$$

Finally, the variational inequality needs to be satisfied by value function $U(Q_t, C_t)$ is

$$0 = \min\Big\{ -U(Q_t, C_t), g(Q_t, C_t) + \frac{\partial}{\partial q}U(Q_t, C_t)\Big[f_\theta(C_t) - (\lambda_1 - \lambda_0)Q_t\Big]$$
$$+ \frac{\partial}{\partial c}U(Q_t, C_t) + \frac{\partial}{\partial q}U(Q_{t-}, C_{t-})Q_{t-}\int_{y\in\mathcal{Y}}\Big(\frac{\lambda_1}{\lambda_0}h(y) - 1\Big)\lambda_0\nu_0(dy)\Big\}.$$

Appendix B

Detailed Calculations for the Wiener Disorder Problem

B.1 The Dynamics of ξ and Φ

We have

$$\xi_t = e^{(\frac{\nu-\mu}{\sigma^2})X_t - \frac{1}{2}(\frac{\nu^2-\mu^2}{\sigma^2})t}, \quad t \geq 0.$$

Using Itô's lemma, we have

$$\begin{aligned} d\xi_t &= \left[-\frac{1}{2}\left(\frac{\nu^2-\mu^2}{\sigma^2}\right) + \frac{1}{2}\sigma^2 \left(\frac{\nu-\mu}{\sigma^2}\right)^2 \right] \xi_t dt + \left(\frac{\nu-\mu}{\sigma^2}\right) \xi_t dX_t \\ &= \left(\frac{\nu-\mu}{\sigma}\right) \xi_t \frac{1}{\sigma}(dX_t - \mu dt) = \beta \xi_t \frac{1}{\sigma}(dX_t - \mu dt). \end{aligned}$$

And for the process Φ defined as

$$\begin{aligned} \Phi_t &= \frac{\mathbb{E}[1_{\{\theta \leq t\}} | \mathcal{F}_t]}{\mathbb{E}[1_{\{\theta > t\}} | \mathcal{F}_t]} = \frac{\mathbb{E}_\infty[L_t 1_{\{\theta \leq t\}} | \mathcal{F}_t]}{\mathbb{E}_\infty[L_t 1_{\{\theta > t\}} | \mathcal{F}_t]} \\ &= \frac{\pi \xi_t + (1-\pi) \int_0^t \frac{\xi_t}{\xi_u} \lambda e^{-\lambda u} \, du}{(1-\pi) e^{-\lambda t}} = e^{\lambda t} \xi_t \left[\frac{\pi}{1-\pi} + \int_0^t \lambda e^{-\lambda u} \frac{1}{\xi_u} \, du \right], \end{aligned}$$

again using Itô's lemma, we have

$$d\Phi_t = (d e^{\lambda t} \xi_t) \left[\frac{\pi}{1-\pi} + \int_0^t \lambda e^{-\lambda u} \frac{1}{\xi_u} \, du \right] + e^{\lambda t} \xi_t \lambda e^{-\lambda t} \frac{1}{\xi_t} dt$$

$$
\begin{aligned}
&= (\lambda e^{\lambda t}\xi_t dt + e^{\lambda t} d\xi_t)\left[\frac{\pi}{1-\pi} + \int_0^t \lambda e^{-\lambda u}\frac{1}{\xi_u}\ du\right] + \lambda dt \\
&= \left(\lambda e^{\lambda t}\xi_t dt + e^{\lambda t}\beta\xi_t\frac{1}{\sigma}(dX_t - \mu dt)\right)\left[\frac{\pi}{1-\pi} + \int_0^t \lambda e^{-\lambda u}\frac{1}{\xi_u}\ du\right] + \lambda dt \\
&= \lambda\Phi_t dt + \beta\frac{1}{\sigma}\Phi_t(dX_t - \mu dt) + \lambda dt \\
&= \lambda(1+\Phi_t)dt + \beta\Phi_t\frac{1}{\sigma}(dX_t - \mu dt).
\end{aligned}
$$

B.2 Independence of θ and X under $\mathbb{P}_Q$

Note that θ and X are independent under a measure $\tilde{\mathbb{P}}$ if and only if

$$
\mathbb{E}_{\tilde{\mathbb{P}}}[f_1(\theta)f_2(X_t)] = \mathbb{E}_{\tilde{\mathbb{P}}}[f_1(\theta)]\mathbb{E}_{\tilde{\mathbb{P}}}[f_2(X_t)]
$$

for $t \geq 0$ and for every choice of functions $f_1 : \mathbb{R}_+ \to \bar{\mathbb{R}}_+$ and $f_2 : \mathbb{R} \to \bar{\mathbb{R}}_+$. We know that θ and X are independent under the measure $\mathbb{P}_\infty$. Using Itô's rule, we have

$$
\begin{aligned}
\mathbb{E}_\infty[f_1(\theta)e^{X_t-\mu t-\frac{1}{2}\sigma^2 t}] &= \mathbb{E}_\infty[f_1(\theta)]\mathbb{E}_\infty[e^{X_t-\mu t-\frac{1}{2}\sigma^2 t}] \\
\mathbb{E}_Q[f_1(\theta)] &= \mathbb{E}_\infty[f_1(\theta)]\mathbb{E}_\infty[e^{X_t-\mu t-\frac{1}{2}\sigma^2 t}] \\
\mathbb{E}_Q[f_1(\theta)] &= \mathbb{E}_\infty[f_1(\theta)]
\end{aligned}
$$

for $t \geq 0$ and for every choice of function $f_1 : \mathbb{R}_+ \to \bar{\mathbb{R}}_+$. Finally,

$$
\begin{aligned}
\mathbb{E}_Q[f_1(\theta)f_2(X_t)] &= \mathbb{E}_\infty[f_1(\theta)f_2(X_t)e^{X_t-\mu t-\frac{1}{2}\sigma^2 t}] \\
&= \mathbb{E}_\infty[f_1(\theta)]\mathbb{E}_\infty[f_2(X_t)e^{X_t-\mu t-\frac{1}{2}\sigma^2 t}] \\
&= \mathbb{E}_\infty[f_1(\theta)]\mathbb{E}_Q[f_2(X_t)] \\
&= \mathbb{E}_Q[f_1(\theta)]\mathbb{E}_Q[f_2(X_t)]
\end{aligned}
$$

for $t \geq 0$ and for every choice of functions $f_1 : \mathbb{R}_+ \to \bar{\mathbb{R}}_+$ and $f_2 : \mathbb{R} \to \bar{\mathbb{R}}_+$, where the fourth equality is because of the previous equality. Thus, θ and X are independent under the measure $\mathbb{P}_Q$.

B.3 Derivation of the Objective Function

For $t \geq 0$ and a stopping time τ, we have

$$
\begin{aligned}
\mathbb{E}[P_{\tau\wedge t}] &= S_0\left[1+\int_0^t \mathbb{E}\left[1_{\{\tau>u\}}\left[\mu 1_{\{u<\theta\}}+\nu 1_{\{\theta\leq u\}}-r\right]e^{X_u-ru-\frac{1}{2}\sigma^2 u}\right]\, du\right]\\
&= S_0\left[1+\int_0^t \mathbb{E}_\infty\left[1_{\{\tau>u\}}\left[(\mu-r)1_{\{u<\theta\}}+(\nu-r)1_{\{\theta\leq u\}}\frac{\xi_u}{\xi_\theta}\right]e^{X_u-ru-\frac{1}{2}\sigma^2 u}\right]\, du\right]\\
&= S_0\left[1+\int_0^t \mathbb{E}_Q\left[1_{\{\tau>u\}}\left[(\mu-r)1_{\{u<\theta\}}+(\nu-r)1_{\{\theta\leq u\}}\frac{\xi_u}{\xi_\theta}\right]\right]e^{(\mu-r)u}\, du\right]\\
&= S_0\Bigg[1+\int_0^t \mathbb{E}_Q\Bigg[1_{\{\tau>u\}}\Bigg[(\mu-r)(1-\pi)e^{-\lambda u}+\\
&\qquad (\nu-r)\Big(\pi\xi_u+(1-\pi)\int_0^u \lambda e^{-\lambda s}\frac{\xi_u}{\xi_s}\, ds\Big)\Bigg]\Bigg]e^{(\mu-r)u}\, du\Bigg]\\
&= S_0\Bigg[1+\int_0^t \mathbb{E}_Q\Bigg[1_{\{\tau>u\}}\Bigg[(\mu-r)(1-\pi)e^{-\lambda u}+\\
&\qquad (\nu-r)(1-\pi)\xi_u\Big(\frac{\pi}{1-\pi}+\int_0^u \lambda e^{-\lambda s}\frac{1}{\xi_s}\, ds\Big)\Bigg]\Bigg]e^{(\mu-r)u}\, du\Bigg]\\
&= S_0\left[1+\int_0^t \mathbb{E}_Q\left[1_{\{\tau>u\}}\left[(\mu-r)(1-\pi)e^{-\lambda u}+(\nu-r)(1-\pi)e^{-\lambda u}\Phi_u\right]\right]e^{(\mu-r)u}\, du\right]\\
&= S_0\left[1+\mathbb{E}_Q\int_0^{\tau\wedge t}\left[(\mu-r)(1-\pi)e^{-\lambda u}+(\nu-r)(1-\pi)e^{-\lambda u}\Phi_u\right]e^{(\mu-r)u}\, du\right]\\
&= S_0\left[1+(1-\pi)\mathbb{E}_Q\int_0^{\tau\wedge t} e^{-(r+\lambda-\mu)u}\left[(\mu-r)+(\nu-r)\Phi_u\right]\, du\right]\\
&= S_0\left[1+(1-\pi)(\nu-r)\mathbb{E}_Q\int_0^{\tau\wedge t} e^{-(r+\lambda-\mu)u}\left[\frac{\mu-r}{\nu-r}+\Phi_u\right]\, du\right]\\
&= S_0\left[1-(1-\pi)(r-\nu)\mathbb{E}_Q\int_0^{\tau\wedge t} e^{-(r+\lambda-\mu)u}\left[\Phi_u-\frac{\mu-r}{r-\nu}\right]\, du\right]\\
&= S_0\left[1-(1-\pi)(r-\nu)\mathbb{E}_Q\int_0^{\tau\wedge t} e^{-\gamma u}\left[\Phi_u-k\right]\, du\right] \quad .
\end{aligned}
$$

where $\gamma = r+\lambda-\mu$ and $k=\frac{\mu-r}{r-\nu}$.

Under the condition $\{\tau<\infty\}$, using monotone convergence theorem, as $t\to\infty$, we have

$$\mathbb{E}[P_\tau] = S_0\left[1-(1-\pi)(r-\nu)\mathbb{E}_Q\int_0^{\tau} e^{-\gamma u}\left[\Phi_u-k\right]\, du\right].$$

B.4 Proof of Lemma 5

Proof. Let us first find the $f_{a,b}$ itself. We are dealing with the differential equation

$$\frac{1}{2}\beta^2\phi^2 f''_{a,b}(\phi) + \ell(\phi) f'_{a,b}(\phi) - \gamma f_{a,b}(\phi) + (\phi - k) = 0 \quad \text{for } \phi \in (a,b). \tag{B.1}$$

Equivalently,

$$f''_{a,b}(\phi) + \frac{2\ell(\phi)}{\beta^2\phi^2} f'_{a,b}(\phi) - \frac{2\gamma}{\beta^2\phi^2} f_{a,b}(\phi) = \frac{2(k-\phi)}{\beta^2\phi^2} \quad \text{for } \phi \in (a,b), \tag{B.2}$$

and $f_{a,b} = 0$ for $\phi \notin (a,b)$. Let us denote the derivative of the scale function as

$$s_\Phi(\phi) := \left(\frac{e^{\frac{\ell_0}{\phi}}}{\phi^{\ell_1}}\right)^{\frac{2}{\beta^2}}.$$

There are two solutions $\psi(\phi)$ and $\varphi(\phi)$ for the homogeneous second order differential equation in (B.1) (where ψ is increasing and ϕ is decreasing) we know that

$$\begin{aligned}\varphi(\phi) &= \psi(\phi)\int_\phi^\infty \frac{e^{-\int \frac{2\ell(x)}{\beta^2 x^2}\,dx}}{\psi^2(x)}\,dx \\ &= \psi(\phi)\int_\phi^\infty \frac{s_\Phi(x)}{\psi^2(x)}\,dx.\end{aligned}$$

Observe that the Wronskian of $\psi(\phi)$ and $\varphi(\phi)$ is $s_\Phi(\phi)$. Particular solution is $p(\phi) = \nu_1(\phi)\psi(\phi) + \nu_2(\phi)\varphi(\phi)$ where

$$\nu_1(\phi) = \int \frac{-g(t)\varphi(t)}{s_\Phi(t)}\,dt \quad \text{and } \nu_2(\phi) = \int \frac{g(t)\psi(t)}{s_\Phi(t)}\,dt \quad \text{where } g(t) = \frac{2(k-t)}{\beta^2 t^2}.$$

Hence

$$p(\phi) = \psi(\phi)\int \frac{2\varphi(t)}{\beta^2 t^2 s_\Phi(t)}(t-k)\,dt - \varphi(\phi)\int \frac{2\psi(t)}{\beta^2 t^2 s_\Phi(t)}(t-k)\,dt.$$

And general solution is

$$f_{a,b}(\phi) = \psi(\phi)\int_\phi^b \frac{2\varphi(t)}{\beta^2 t^2 s_\Phi(t)}(t-k)\,dt - \varphi(\phi)\int_\phi^b \frac{2\psi(t)}{\beta^2 t^2 s_\Phi(t)}(t-k)\,dt +$$
$$C_1\psi(\phi) + C_2\varphi(\phi)$$

where $C_1, C_2 \in \mathbb{R}$.

We want this function to be 0, outside of (a,b). Therefore, if we plug a and b into the function. We get

$$C_1\psi(a) + C_2\varphi(a) = -\psi(a)\int_a^b \frac{2\varphi(t)}{\beta^2 t^2 s_\Phi(t)}(t-k)\ dt + \varphi(a)\int_a^b \frac{2\psi(t)}{\beta^2 t^2 s_\Phi(t)}(t-k)\ dt,$$

$$C_1\psi(b) + C_2\varphi(b) = 0.$$

If we solve these equations, we find C_1 and C_2 as

$$\begin{aligned} C_1 &= \frac{\varphi(b)}{\varphi(a)\psi(b) - \varphi(b)\psi(a)}\Bigg[\psi(a)\int_a^b \frac{2\varphi(t)}{\beta^2 t^2 s_\Phi(t)}(t-k)\ dt - \\ &\qquad \varphi(a)\int_a^b \frac{2\psi(t)}{\beta^2 t^2 s_\Phi(t)}(t-k)\ dt\Bigg], \\ C_2 &= -\frac{\psi(b)}{\varphi(a)\psi(b) - \varphi(b)\psi(a)}\Bigg[\psi(a)\int_a^b \frac{2\varphi(t)}{\beta^2 t^2 s_\Phi(t)}(t-k)\ dt - \\ &\qquad \varphi(a)\int_a^b \frac{2\psi(t)}{\beta^2 t^2 s_\Phi(t)}(t-k)\ dt\Bigg]. \end{aligned}$$

And our function becomes

$$\begin{aligned} f_{a,b}(\phi) &= \psi(\phi)\int_\phi^b \frac{2\varphi(t)}{\beta^2 t^2 s_\Phi(t)}(t-k)\ dt - \varphi(\phi)\int_\phi^b \frac{2\psi(t)}{\beta^2 t^2 s_\Phi(t)}(t-k)\ dt + \\ &\frac{\psi(a)\int_a^b \frac{2\varphi(t)}{\beta^2 t^2 s_\Phi(t)}(t-k)\ dt - \varphi(a)\int_a^b \frac{2\psi(t)}{\beta^2 t^2 s_\Phi(t)}(t-k)\ dt}{\varphi(a)\psi(b) - \varphi(b)\psi(a)}[\varphi(b)\psi(\phi) - \psi(b)\varphi(\phi)]. \end{aligned}$$

We are asked to prove that

$$f_{a,b}(\phi) = \mathbb{E}\int_0^{\tau_{a,b}} e^{-\gamma t}\left[\Phi_t - k\right]\ dt,$$

where $\tau_{a,b}$ is the first exit time of Φ from the interval (a,b). One quick observation is that when $t \in [0, \tau_{a,b})$, $\Phi_t \in (a,b)$, therefore the equation in (B.2) holds and

$$\Phi_t - k = -\left(\frac{1}{2}\beta^2\Phi_t^2 f''_{a,b}(\Phi_t) + \ell(\phi)f'_{a,b}(\Phi_t) - \gamma f_{a,b}(\Phi_t)\right) \quad \text{for } t \in [0, \tau_{a,b}).$$

Using Itô's rule we obtain that, for $\tau_{a,b}$

$$\mathbb{E}[e^{-\gamma\tau_{a,b}} f_{a,b}(\Phi_{\tau_{a,b}}) - f_{a,b}(\phi)]$$

$$
\begin{aligned}
&= \mathbb{E}\int_0^{\tau_{a,b}} e^{-\gamma t}\left[\frac{1}{2}\beta^2\Phi_t^2 f''_{a,b}(\Phi_t) + \ell(\Phi_t) f'_{a,b}(\Phi_t) - \gamma f_{a,b}(\Phi_t)\right]\ dt \\
&= -\mathbb{E}\int_0^{\tau_{a,b}} e^{-\gamma t}\left[\Phi_t - k\right]\ dt.
\end{aligned}
$$

Moreover, $\Phi_{\tau_{a,b}} = a$ or b, hence $f_{a,b}(\Phi_{\tau_{a,b}}) = 0$. Thus, we get

$$
f_{a,b}(\phi) = \mathbb{E}\int_0^{\tau_{a,b}} e^{-\gamma t}\left[\Phi_t - k\right]\ dt.
$$

Finally, to prove

$$
f_b(\phi) = \mathbb{E}\int_0^{\tau_b} e^{-\gamma t}\left[\Phi_t - k\right]\ dt,
$$

where τ_b is the first exit time of Φ from the interval $[0, b)$, we use the fact that $\{0\}$ is an entry-not-exit boundary, we have $\tau_{a,b} \nearrow \tau_b$ as $a \searrow 0$. Observe that

$$
\begin{aligned}
\left|\int_0^{\tau_{a,b}} e^{-\gamma t}\left[\Phi_t - k\right]\ dt\right| &\leq \int_0^{\tau_{a,b}} \left|e^{-\gamma t}\left[\Phi_t - k\right]\right|\ dt \\
&= \int_0^{\tau_{a,b}} e^{-\gamma t}\left|\Phi_t - k\right|\ dt \\
&\leq \int_0^{\tau_{a,b}} \left|\Phi_t - k\right|\ dt \\
&\leq \int_0^{\tau_b} \left|\Phi_t - k\right|\ dt \\
&\leq \max\{k, |b-k|\}\tau_b
\end{aligned}
$$

Furthermore, we know that $\mathbb{E}[\tau_b] < \infty$. Hence, by dominated convergence theorem we have

$$
\lim_{a\searrow 0}\mathbb{E}\int_0^{\tau_{a,b}} e^{-\gamma t}\left[\Phi_t - k\right]\ dt = \mathbb{E}\int_0^{\tau_b} e^{-\gamma t}\left[\Phi_t - k\right]\ dt.
$$

□

B.5 Proof of Lemma 6

Proof. We will take the limit of $f_{a,b}(\phi)$ as $a \searrow 0$.

$$\lim_{a\searrow 0} f_{a,b}(\phi) = \psi(\phi)\int_{\phi}^{b} \frac{2\varphi(t)}{\beta^2 t^2 s_\Phi(t)}(t-k)\ dt - \varphi(\phi)\int_{\phi}^{b} \frac{2\psi(t)}{\beta^2 t^2 s_\Phi(t)}(t-k)\ dt +$$
$$[\varphi(b)\psi(\phi)-\psi(b)\varphi(\phi)] \lim_{a\searrow 0} \frac{\psi(a)\int_a^b \frac{2\varphi(t)}{\beta^2 t^2 s_\Phi(t)}(t-k)\ dt - \varphi(a)\int_a^b \frac{2\psi(t)}{\beta^2 t^2 s_\Phi(t)}(t-k)\ dt}{\varphi(a)\psi(b)-\varphi(b)\psi(a)}.$$

For convenience, it is assumed that $\varphi(b) = \psi(b) = 1$. Moreover, we have $0 < \psi(0+) < \infty$, $\varphi(0+) = \infty$ and $lim_{\phi\to 0+} \frac{\psi'(\phi)}{s_\Phi(\phi)} = 0$. Using L'Hôspital's rule we get

$$\lim_{a\searrow 0} f_{a,b}(\phi) = \psi(\phi)\int_{\phi}^{b} \frac{2\varphi(t)}{\beta^2 t^2 s_\Phi(t)}(t-k)\ dt - \varphi(\phi)\int_{\phi}^{b} \frac{2\psi(t)}{\beta^2 t^2 s_\Phi(t)}(t-k)\ dt +$$
$$[\psi(\phi)-\varphi(\phi)] \lim_{a\searrow 0} \frac{\psi'(a)\int_a^b \frac{2\varphi(t)}{\beta^2 t^2 s_\Phi(t)}(t-k)\ dt - \varphi'(a)\int_a^b \frac{2\psi(t)}{\beta^2 t^2 s_\Phi(t)}(t-k)\ dt}{\varphi'(a)-\psi'(a)}.$$

To get $f_b'(\phi)$, let us differentiate the function with respect to ϕ

$$\begin{aligned}
f_b'(\phi) &= \psi'(\phi)\int_{\phi}^{b} \frac{2\varphi(t)}{\beta^2 t^2 s_\Phi(t)}(t-k)\ dt - \varphi'(\phi)\int_{\phi}^{b} \frac{2\psi(t)}{\beta^2 t^2 s_\Phi(t)}(t-k)\ dt + \\
&[\psi'(\phi)-\varphi'(\phi)] \lim_{a\searrow 0} \frac{\psi'(a)\int_a^b \frac{2\varphi(t)}{\beta^2 t^2 s_\Phi(t)}(t-k)\ dt - \varphi'(a)\int_a^b \frac{2\psi(t)}{\beta^2 t^2 s_\Phi(t)}(t-k)\ dt}{\varphi'(a)-\psi'(a)} \\
&= \psi'(\phi)\int_{\phi}^{b} \frac{2\varphi(t)}{\beta^2 t^2 s_\Phi(t)}(t-k)\ dt - \varphi'(\phi)\int_{\phi}^{b} \frac{2\psi(t)}{\beta^2 t^2 s_\Phi(t)}(t-k)\ dt + \\
&[\psi'(\phi)-\varphi'(\phi)] \lim_{a\searrow 0} \frac{\frac{\psi'(a)}{s_\Phi(a)}\int_a^b \frac{2\varphi(t)}{\beta^2 t^2 s_\Phi(t)}(t-k)\ dt - \frac{\varphi'(a)}{s_\Phi(a)}\int_a^b \frac{2\psi(t)}{\beta^2 t^2 s_\Phi(t)}(t-k)\ dt}{\frac{\varphi'(a)}{s_\Phi(a)}-\frac{\psi'(a)}{s_\Phi(a)}} \\
&= \psi'(\phi)\int_{\phi}^{b} \frac{2\varphi(t)}{\beta^2 t^2 s_\Phi(t)}(t-k)\ dt - \varphi'(\phi)\int_{\phi}^{b} \frac{2\psi(t)}{\beta^2 t^2 s_\Phi(t)}(t-k)\ dt + \\
&[\psi'(\phi)-\varphi'(\phi)] \lim_{a\searrow 0} \frac{-\frac{\varphi'(a)}{s_\Phi(a)}\int_a^b \frac{2\psi(t)}{\beta^2 t^2 s_\Phi(t)}(t-k)\ dt}{\frac{\varphi'(a)}{s_\Phi(a)}} \\
&= \psi'(\phi)\int_{\phi}^{b} \frac{2\varphi(t)}{\beta^2 t^2 s_\Phi(t)}(t-k)\ dt - \varphi'(\phi)\int_{\phi}^{b} \frac{2\psi(t)}{\beta^2 t^2 s_\Phi(t)}(t-k)\ dt - \\
&[\psi'(\phi)-\varphi'(\phi)] \int_{0}^{b} \frac{2\psi(t)}{\beta^2 t^2 s_\Phi(t)}(t-k)\ dt.
\end{aligned}$$

We were asked to find b value such that $f_b'(b-) = 0$. If we plug $b-$ to the previous function, we require

$$[\psi'(b-)-\varphi'(b-)] \int_{0}^{b} \frac{2\psi(t)}{\beta^2 t^2 s_\Phi(t)}(t-k)\ dt = 0.$$

However, as ψ is strictly increasing and φ is strictly decreasing, the first term is strictly positive. Thus, the b value that we want satisfies

$$\int_0^b \frac{2\psi(t)}{\beta^2 t^2 s_\Phi(t)}(t-k)\ dt = 0,$$

where $b > k$ and with this choice of b, $f_b(\phi)$ becomes

$$f_b(\phi) = \psi(\phi) \int_{\phi\wedge b}^b \frac{2\varphi(t)}{\beta^2 t^2 s_\Phi(t)}(t-k)\ dt - \varphi(\phi) \int_{\phi\wedge b}^b \frac{2\psi(t)}{\beta^2 t^2 s_\Phi(t)}(t-k)\ dt.$$

To show that such b exists, we can show that the limit of the derivative of this integrand goes to ∞ as $t \to \infty$. So we are looking for

$$\begin{aligned}\lim_{t\to\infty} \frac{\psi(t)}{t^2 s_\Phi(t)}(t-k) &= \lim_{t\to\infty} \frac{\psi(t)(t-k)}{t^{2-\frac{2\ell_1}{\beta^2}}} e^{-\frac{2\ell_0}{\beta^2 t}} \\ &= \lim_{t\to\infty} \psi(t) \frac{t-k}{t^{2-\frac{2\ell_1}{\beta^2}}} e^{-\frac{2\ell_0}{\beta^2 t}}.\end{aligned}$$

If $2 - \frac{2\ell_1}{\beta^2} \leq 1$ then the limit is ∞. Otherwise we can use L'Hôspital's rule and get

$$\begin{aligned}\lim_{t\to\infty} \frac{\psi(t)}{t^2 s_\Phi(t)}(t-k) &= \lim_{t\to\infty} \frac{\psi'(t)(t-k)+\psi(t)}{s_\Phi(t)(t-\ell_0)} \\ &= \lim_{t\to\infty} \frac{\psi'(t)}{s_\Phi(t)} \frac{(t-k)}{(t-\ell_0)} + \lim_{t\to\infty} \frac{\psi_t}{s_\Phi(t)(t-\ell_0)}.\end{aligned}$$

On the RHS, we know that the first limit goes to infinity, as $lim_{\phi\to\infty} \frac{\psi'(\phi)}{s_\Phi(\phi)} = \infty$, and the second limit is non negative.

We also want to show that

$$\left.\begin{aligned} & f_b(\phi) < 0 \\ & \frac{1}{2}\beta^2\phi^2 f_b''(\phi) + \ell(\phi) f_b'(\phi) - \gamma f_b(\phi) + (\phi - k) = 0 \end{aligned}\right\}, \quad \phi \in [0, b)$$

$$\left.\begin{aligned} & f_b(\phi) = 0 \\ & \frac{1}{2}\beta^2\phi^2 f_b''(\phi) + \ell(\phi) f_b'(\phi) - \gamma f_b(\phi) + (\phi - k) > 0 \end{aligned}\right\}, \quad \phi \in [b, \infty)$$

By the way we constructed f_b function, the equalities are obvious. For the first inequality, let us observe that

$$f_b'(\phi) = \psi'(\phi) \int_{\phi\wedge b}^b \frac{2\varphi(t)}{\beta^2 t^2 s_\Phi(t)}(t-k)\ dt - \varphi'(\phi) \int_{\phi\wedge b}^b \frac{2\psi(t)}{\beta^2 t^2 s_\Phi(t)}(t-k)\ dt.$$

For $\phi \in [k,b)$ $f_b'(\phi) > 0$ and we know that $f_b(b) = 0$. Hence $f_b(\phi) < 0$ for $\phi \in [k,b)$.

For $\phi \in [0,k)$, using the alternative representation of $f_b(\phi)$ which we have shown in 5 and the Markov property of the process Φ_t we can write

$$\begin{aligned}
f_b(\phi) &= \mathbb{E}^{\phi}\left[\int_0^{\tau_k} e^{-\gamma t}\left[\Phi_t - k\right]\ dt\right] + \mathbb{E}^{\phi}\left[\int_{\tau_k}^{\tau_b} e^{-\gamma t}\left[\Phi_t - k\right]\ dt\right] \\
&= \mathbb{E}^{\phi}\left[\int_0^{\tau_k} e^{-\gamma t}\left[\Phi_t - k\right]\ dt\right] + \mathbb{E}^{\phi}\left[e^{-\gamma\tau_k}\int_0^{\tau_b-\tau_k} e^{-\gamma t}\left[\Phi_{t+\tau_k} - k\right]\ dt\right] \\
&= \mathbb{E}^{\phi}\left[\int_0^{\tau_k} e^{-\gamma t}\left[\Phi_t - k\right]\ dt\right] + \mathbb{E}^{\phi}\left[e^{-\gamma\tau_k}\right]\mathbb{E}^{k}\left[\int_0^{\tau_b} e^{-\gamma t}\left[\Phi_t - k\right]\ dt\right] \\
&= \mathbb{E}^{\phi}\left[\int_0^{\tau_k} e^{-\gamma t}\left[\Phi_t - k\right]\ dt\right] + \mathbb{E}^{\phi}\left[e^{-\gamma\tau_k}\right] f_b(k).
\end{aligned}$$

We have already shown that $f_b(\phi) < 0$ for $\phi \in [k,b)$. Moreover $\Phi_t - k < 0$ for $t \in [0,\tau_k)$. This proves that $f_b(\phi) < 0$ for $\phi \in [0,b)$.

The second inequality is again obvious as $f_b(\phi) = 0$ for $\phi \in [b,\infty)$ and also we know that $b > k$. Hence,

$$\frac{1}{2}\beta^2\phi^2 f_b''(\phi) + \ell(\phi) f_b'(\phi) - \gamma f_b(\phi) + (\phi - k) = \phi - k > 0, \quad \phi \in [b,\infty).$$

$\square$