Three Problems in Stochastic Control and Applications

by

Jia Guo

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
(Applied and Interdisciplinary Mathematics)
in The University of Michigan
2020

Doctoral Committee:

Professor Erhan, Bayraktar, Chair
Assistant Professor Asaf Cohen
Professor Stilian Stoev
Professor Virginia Young

Jia Guo
guojia@umich.edu
ORCID iD:0000-0002-5549-0647

To my mom and dad

ACKNOWLEDGEMENTS

This thesis would not have been possible without the enormous support from my Ph.D. advisor Professor Erhan Bayraktar. Not only did he teach me mathematics, but also influenced me in his attitude when facing difficulties. His great passion toward research encourages me a lot.

I would like thank my collaborators Professors Yan Dolinsky and Erik Ekström. They gave me a great motivation and lots of significant suggestions. I could not complete the work well without their help.

I am very grateful for my committee members for their great help. I would like thank Professor Stilian Stoev, Virginia Young, and Asaf Cohen. Thanks for serving on my dissertation committee and reading my thesis carefully.

I am really proud of being part of the Department of Mathematics at the University of Michigan. I thank you for providing me an academic home and the financial support during last five years. I also thank all the participants at the AIM program and Financial Math program for providing me a great atmosphere to study and work.

I would like to express my appreciation for the staff of the Department of Mathematics at the University of Michigan, particularly Tara McQueen, and Teresa Stokes. Their assistance considerably simplified my routine work. Their advice regarding campus events and resources helped me a lot.

I also would like to thank my friends, Anthony Della Pella, Jonathan Guzman, Yiwang Chen, Ruian Chen, Yifeng Huang, Yuchen Liao, Ningyuan Wang, Zhan Jiang, Jingjie Zhang, Xin Zhang, and many others who have have been by my side encouraging me during difficult times, and who have also been sharing many joyfulness in life in the past five years.

I would like to express my special appreciation to my friends Ketian Yu and Changjiang Liu. Without their help, I could not have gone through this tough year. Many thanks and touched in my heart.

I cannot finish without saying how grateful I am to my family: my mother Jiayun Gong, my father Jinhui Guo. Their unconditional love is the most important support for me.

TABLE OF CONTENTS

LIST OF FIGURES

Figure

LIST OF TABLES

Table

ABSTRACT

This thesis mainly summarizes three different projects that I am devoted to: Recombining Tree Approximations for Optimal Stopping for Diffusions (Chapter II), Continuity of Utility Maximization under Weak Convergence (Chapter III) and Disorder Detection with Costly Observations (Chapter IV). The first two projects are related work. The third one is based on [16].

In Chapter II, we develop two numerical methods for optimal stopping in the framework of one dimensional diffusion. Both of the methods use the Skorohod embedding in order to construct *recombining tree* approximations for diffusions with general coefficients. This technique allows us to determine convergence rates and construct nearly optimal stopping times which are optimal at the same rate. Finally, we demonstrate the efficiency of our schemes on several models.

In Chapter III, we find sufficient conditions for the continuity of the utility maximization problem from terminal wealth under convergence in distribution of the underlying processes. We provide several examples which illustrate that without these conditions, we cannot generally expect continuity to hold. Finally, we apply our continuity results to numerical computations of the shortfall risk in the Heston model.

In Chapter IV, we study the Wiener disorder detection problem where each observation is associated with a positive cost. In this setting, a strategy is a pair consisting of a sequence of observation times and a stopping time corresponding to the declaration of disorder. We characterize the minimal cost of the disorder problem with costly observations as the unique fixed point of a certain jump operator, and we determine

the optimal strategy.

CHAPTER I

Introduction

The Motivation of Chapter II comes from the calculations of American options prices in local volatility models we develop numerical schemes for optimal stopping in the framework of one dimensional diffusion. In particular, we develop tree based approximations. In general for non-constant volatility models the nodes of the tree approximation trees do not recombine and this fact results in an exponential and thus a computationally explosive tree that cannot be used especially in pricing American options. We will present two novel ways of constructing recombining trees.

The first numerical scheme we propose, a recombining binomial tree (on a uniform time and space lattice), is based on correlated random walks. A correlated random walk is a generalized random walk in the sense that the increments are not identically and independently distributed and the one we introduce has one step memory. The idea to use correlated random walks for approximating diffusion processes goes back to [43], where the authors studied the weak convergence of correlated random walks to Markov diffusions. The disadvantage of the weak convergence approach is that it can not provide error estimates. Moreover, the weak convergence result can not be applied for numerical computations of the optimal stopping time. In order to obtain error estimates for the approximations and calculate numerically the optimal control we should consider all the processes on the same probability space, and so methods based on strong approximation theorems come into picture. In this chapter we apply the Skorokhod embedding technique for "small" perturbations of the correlated random walks. Our approach can be seen as an extension of recombining tree approximations from the case when a stock evolves according to the geometrical Brownian motion to the case of a more general diffusion evolution. This particular case required the Skorokhod embedding into the Brownian motion (with a constant variance) and it was treated in the more general game options case in [51, 32].

Under boundedness and Lipschitz conditions on the drift and volatility of the diffusion, we obtain error estimates of order $O(n^{-1/4})$. Moreover we show how to construct a stopping time which is optimal up to the same order. In fact, we consider a more general setup where the diffusion process might have absorbing barriers. Clearly, most of the local volatility modes which are used in practice (for instance, the CEV model [23] and the CIR model [25]) do not satisfy the above conditions. Still, by choosing an appropriate absorbing barriers for these models, we can efficiently approximate the original models, and for the absorbed diffusions, which satisfy the above conditions, we apply our results.

Our second numerical scheme is a recombining trinomial tree, which is obtained by directly sampling the original process at suitable chosen random intervals. In this method we relax the continuity assumption and work with diffusions with measurable coefficients, e.g. the coefficients can be discontinuous, while we keep the boundedness conditions. (See Section 2.3.5 for an example). The main idea is to construct a sequence of random times such that the increment of the diffusion between two sequel times belongs to some fixed set of the form $\left\{-\bar{\sigma}\sqrt{\frac{T}{n}}, 0, \bar{\sigma}\sqrt{\frac{T}{n}}\right\}$ and the expectation of the difference between two sequel times equals to $\frac{T}{n}$. Here n is the discretization parameter and T is the maturity date. This idea is inspired by the recent work [2, 1] where the authors applied Skorohod embedding in order to obtain an Euler approximation of irregular one dimensional diffusions. Following their methodology we construct an exact scheme along stopping times. The constructions are different and in particular, in contrast to the above papers we introduce a recombining tree approximation. The above papers do provide error estimate (in fact they are of the same order as our error estimates) for the expected value of a function of the terminal value (from mathematical finance point of view, this can be viewed as European Vanilla options). Since we deal with American options, our proof requires additional machinery which allows to treat stopping times.

The second method is more direct and does not require the construction of a diffusion perturbation. In particular it can be used for the computations of Barrier options prices; see Remark 2.2.1. As we do for the first method, we obtain error estimates of order $O(n^{-1/4})$ and construct a stopping time which is optimal up to the same order.

Comparison with other numerical schemes. The most well-known approach to evaluate American options in one dimension is the finite difference scheme called the projected SOR method, see e.g. [74]. As opposed to a tree based scheme, the finite difference scheme needs to artificially restrict the domain and impose boundary conditions to fill in the entire lattice. The usual approach is the so-called *far-field* boundary condition (as is done in [74]). See the discussion in [75]. The main problem is that it is not known how far is sufficient and the error is hard to quantify. This problem was addressed recently by [75] for put options written on CEV models by what they call an artificial boundary method, which is model dependent in that an exact boundary condition is derived that is then embedded into the numerical scheme. (In fact, the PDE is transformed into a new PDE which satisfies a Neumann boundary condition.) This technique requires having an explicit expression of a Sturm-Lioville equation associated with the Laplace transform of the pricing equation. As an improvement over [75] the papers [76, 53] propose a numerical method based on solving the integral equations that the optimal exercise boundary satisfies. None of these papers discuss the convergence rate or prove that their proposed optimal exercise times are nearly optimal. We have the advantage of having an optimal stopping problem at the discrete level and without much additional effort are able to show that these stopping times are actually approximately optimal when applied to the continuous time problem. We compare our numerical results to these papers in Table 2.3.2.

Another popular approach is the Euler scheme or other Monte-Carlo based schemes (see e.g. [39, 3]). The grid structure allows for an efficient numerical computations of stochastic control problems via dynamical programming since computation of conditional expectation simplifies considerably. Compare this to the least squares method in [59] (also see [39]). Besides the computational complexity, the bias in the least squares is hard to characterize since it also relies on the choice of bases functions. Recently a hybrid approach was developed in [14]. (The scope of this paper is more general and is in the spirit of [59]). Our discussion above in comparing our proposed method to Monte-Carlo applies here. Moreover, although [14] provides a convergence rate analysis using PDE methods (requiring regularity assumptions on the coefficients, which our scheme does not need), our approach in addition can prove that the proposed stopping times from our numerical approximations are nearly optimal.

To summarize, our contributions are

- We build two recombining tree approximations for local volatility models, which considerably simplifies the computation of conditional expectations.
- By using the Skorohod embedding technique, we give a proof of the rate of convergence of the schemes.
- This technique in particular allows us to show that the stopping times constructed from the numerical approximations are nearly optimal and the error is of the same magnitude as the convergence rate.

These two novel discretizations gives us idea to construct approximating models for the Heston model. This leads to further study. Chapter III deals with the following question. Given a utility function and a sequence of financial markets with underlying assets $S^{(n)}$, $n \in \mathbb{N}$, which converges weakly to an underlying asset S. Under which condition the values of the utility maximization problem from terminal wealth in the approximating sequence converge to the corresponding value for the model given by S.

Although the utility maximization problem was largely studied (see, for instance, [60, 61, 48, 54, 45, 69, 17]), to the best of our knowledge, the continuity under weak convergence was studied only in [70] in an essentially complete market setup.

In this work we consider a general incomplete framework with a continuous state dependent utility. We divide the problem into two problems. The first one studies lower semi–continuity under weak convergence, namely the conditions that the value of the utility maximization problem in the limit model is less or equal than the lower limit of the converging models. The second problem deals with upper semi–continuity, i.e. conditions that the value of the utility maximization problem in the limit model will dominate from above the upper limit of the approximating sequence.

We show that for the lower semi–continuity to hold, it is sufficient that the approximating sequence $S^{(n)}$, $n \in \mathbb{N}$, has a bounded jump activity. The formal condition is given in Assumption 3.1.4. In particular, we do not require concavity of the utility function. The main idea is to prove that an admissible integral of the form $\int \gamma dS$ can be approximated in the weak sense by admissible integrals of the form $\int \gamma^{(n)} dS^{(n)}$, $n \in \mathbb{N}$. The assumption on the jump activity is essential for the admissability of the approximating sequence.

The upper semi–continuity is a more delicate issue. Roughly speaking, we prove that if the utility function is concave and the state price densities in the limit model can be approximated by state price densities in the approximating sequence (see Assumption 3.1.5) then upper semi–continuity holds. The proof relies on the optional decomposition theorem. We provide two examples which illustrate that these assumptions are essential.

We apply our continuity results in order to construct an approximating sequence for the Heston model. For technical reasons we truncate the model in such a way that the volatility is bounded. The novelty of our construction is that the approximating sequence lies on a grid and satisfies the assumptions required for the continuity of utility maximization from terminal wealth. The grid structure enables efficient numerical computations of stochastic control problems via dynamical programming since the computation of conditional expectations simplifies considerably.

Our last contribution is the implementation of the constructed approximating models for the numerical computations of the shortfall risk measure in the Heston model. We focus on European call options. It is well known (see [26, 38, 33, 65]) that in the Heston model the super–replication price is prohibitively high and lead to buy–and–hold strategies. Namely, the cheapest way to super–hedge a European call option is to buy one stock at the initial time and keep that position till maturity. For a given initial capital which is less than the initial stock price we want to compute the corresponding shortfall risk. This cannot be done analytically and so numerical schemes come into picture.

It is important to mention the series of papers [58, 15, 57, 63, 62] where the authors studied the stability and the corresponding expansion of the utility maximization problem in terms of a perturbation of the model. The main difference is that in these papers the stochastic base is fixed while in our setup each financial model is defined on its own probability space. As a result, while their approach deals with the stability of the models with respect small perturbations, we are able to obtain numerical approximations using discrete models.

The idea of Chapter IV is similar to paper [36]. Both deal with sequential testing with costly observations. We aim to find a strategy that consists of a decision whether to stop or not, and with a rule specifying how long to wait for the next observation. The main result states that the value function of the problem can be characterized as the unique fixed point of this operator and that the value function can be determined by an iterative procedure involving the operator. However, due to the different operators constructed by different problems, the main properties of our problem seems more challenging. For example, we use Girsanov theorem to prove the operator keeps concavity of functions. On the other hand, unlike paper [36], it is not straightforward that the continuation region is an open interval based on concavity and monotonity of the operator. Although it is seems true in the numerical examples, we leave it as an open problem for further study.

Comparing with paper [16], we add extra cost for observations. The idea and computations between them are approximiate. Our work gives a more concise way to consider the observation cost in practice and makes it more useful in further study.

We actually compare the numerical results for correctness checking(by assuming $d = 0$).

The fixed-point strategy is a quite common tool in this kind of problems, however, different constructions of the operators give different difficulities in studying. We refer to paper [9]. Here the observing processs N is a Poisson process with arriving rate sudden changes in an unkown time. By assuming the "changing" arriving rate is Bernoulli distributed, they solve the problem by using functional iterations. Due to the complexity of the operator, their work is subtle and abstract. In our current work, we show more detailed properties of the operator with less assumptions. Anyway, although the fixing point strategy is not a novel strategy in this kind of problems, different types of operators will generate different constructions.

CHAPTER II

Recombining Tree Approximations for Optimal Stopping for Diffusions

In this chapter we develop two numerical methods for optimal stopping in the framework of one dimensional diffusion. See paper [10]. We introduce the setup and introduce the two methods we propose and present the main theoretical results of the paper, namely Theorems 2.1.1 and 2.1.2 (see the last statements in subsections 2.1.2 and 2.1.3). Sections 2.1.4 and 2.2 are devoted to the proof of these respective results. In Section 2.3 we provide a detailed numerical analysis for both of our methods by applying them to various local volatility models.

2.1 Preliminaries and Main Results

2.1.1 The Setup

Let $\{W_t\}_{t=0}^{\infty}$ be a standard one dimensional Brownian motion, and let Y_t, $t \geq 0$ be a one dimensional diffusion

$$dY_t = \sigma(Y_t)dW_t + \mu(Y_t)dt, \quad Y_0 = x. \tag{2.1}$$

Let $\{\mathcal{F}_t\}_{t=0}^{T}$ be a filtration which satisfies the usual conditions such that Y_t, $t \geq 0$ is an adapted process and W_t, $t \geq 0$ is a Brownian motion with respect to this filtration. We assume that the SDE (2.1) has a weak solution that is unique in law.

Introduce absorbing barriers $B < C$ where $B \in [-\infty, \infty)$ and $C \in (-\infty, \infty]$. We assume that $x \in (B, C)$ and look at the absorbed stochastic process

$$X_t = \mathbb{I}_{t<\inf\{s:Y_s \notin (B,C)\}} Y_t + \mathbb{I}_{t \geq \inf\{s:Y_s \notin (B,C)\}} Y_{\inf\{s:Y_s \notin (B,C)\}}, \quad t \geq 0.$$

Motivated by the valuation of American options prices we study an optimal stopping problem with maturity date T and a reward $f(t, X_t)$, $t \in [0, T]$ where $f : [0, T] \times \mathbb{R} \to \mathbb{R}_+$ satisfies

(2.2)

$$|f(t_1, x_1) - f(t_2, x_2)| \leq L\left((1 + |x_1|)|t_2 - t_1| + |x_2 - x_1|\right), \quad t_1, t_2 \in [0, T], \quad x_1, x_2 \in \mathbb{R}$$

for some constant L. Clearly, payoffs of the form $f(t, x) = e^{-rt}(x-K)^+$ (call options) and $f(t, x) = e^{-rt}(K - x)^+$ (put options) fit in our setup.

The optimal stopping value is given by

$$V = \sup_{\tau \in \mathcal{T}_{[0,T]}} \mathbb{E}[f(\tau, X_\tau)], \tag{2.3}$$

where $\mathcal{T}_{[0,T]}$ is the set of stopping times in the interval $[0, T]$, with respect to the filtration $\mathcal{F}_t$, $t \geq 0$.

2.1.2 Binomial Approximation Scheme Via Random Walks

In this section we adopt the following.

Assumption 2.1.1. *The functions $\mu, \sigma : (B, C) \to \mathbb{R}$ are bounded and Lipschitz continuous. Moreover, $\sigma : (B, C) \to \mathbb{R}$ is strictly positive and uniformly bounded away from zero.*

Binomial Approximation of the State Process. Fix $n \in \mathbb{N}$ and let $h = h(n) := \frac{T}{n}$ be the time discretization. Since the meaning is clear we will use h instead of $h(n)$. Let $\xi_1^{(n)}, ..., \xi_n^{(n)} \in \{-1, 1\}$ be a sequence of random variables. In the sequel, we always use the initial data $\xi_0^{(n)} \equiv 1$. Consider the random walk

$$X_k^{(n)} = x + \sqrt{h}\sum_{i=1}^{k} \xi_i^{(n)} \quad k = 0, 1, ..., n$$

with finite absorbing barriers $B_n < C_n$ where

$$B_n = x + \sqrt{h}\min\{k \in \mathbb{Z} \cap [-n-1, \infty) : x + \sqrt{h}k > B + h^{1/3}\}$$

$$C_n = x + \sqrt{h}\max\{k \in \mathbb{Z} \cap (-\infty, n+1] : x + \sqrt{h}k < C - h^{1/3}\}.$$

Clearly, the process $\{X_k^{(n)}\}_{k=0}^n$ lies on the grid $x + \sqrt{h}\{-n, 1-n, ..., 0, 1, ..., n\}$ (in fact because of the barriers it lies on a smaller grid).

We aim to construct a probabilistic structure so the pair $\{X_k^{(n)}, \xi_k^{(n)}\}_{k=0}^n$ forms a Markov chain weakly approximating (with the right time change) the absorbed diffusion X. We look for a predictable (with respect the filtration generated by $\xi_1^{(n)}, ..., \xi_n^{(n)}$) process $\alpha^{(n)} = \{\alpha_k^{(n)}\}_{k=0}^n$ which is uniformly bounded (in space and n) and a probability measure $\mathbb{P}_n$ on $\sigma\{\xi_1^{(n)}, ..., \xi_n^{(n)}\}$ such that the perturbation defined by

$$\hat{X}_k^{(n)} := X_k^{(n)} + \sqrt{h}\alpha_k^{(n)}\xi_k^{(n)}, \quad k = 0, 1, ..., n \tag{2.4}$$

is matching the first two moments of the diffusion X. Namely, we require that for any $k = 1, ..., n$ (on the event $B_n < X_{k-1}^{(n)} < C_n$)

$$\mathbb{E}_n\left(\hat{X}_k^{(n)} - \hat{X}_{k-1}^{(n)} | \xi_1^{(n)}, ..., \xi_{k-1}^{(n)}\right) = h\mu(X_{k-1}^{(n)}) + o(h), \tag{2.5}$$

$$\mathbb{E}_n\left((\hat{X}_k^{(n)} - \hat{X}_{k-1}^{(n)})^2 | \xi_1^{(n)}, ..., \xi_{k-1}^{(n)}\right) = h\sigma^2(X_{k-1}^{(n)}) + o(h), \tag{2.6}$$

where we use the standard notation $o(h)$ to denote a random variable that converge to zero (as $h \downarrow 0$) a.s. after dividing by h. We also use the convention $O(h)$ to denote a random variable that is uniformly bounded after dividing by h.

From (2.4) it follows that $X_k^{(n)} - \hat{X}_k^{(n)} = o(1)$. Hence the convergence of $X^{(n)}$ is equivalent to the convergence of $\hat{X}^{(n)}$. It remains to solve the equations (2.5)–(2.6). By applying (2.4)–(2.5) together with the fact that $\alpha^{(n)}$ is predictable and $(\xi^{(n)})^2 \equiv 1$ we get

$$\mathbb{E}_n \left((\hat{X}_k^{(n)} - \hat{X}_{k-1}^{(n)})^2 | \xi_1^{(n)}, ..., \xi_{k-1}^{(n)} \right) = h(1 + 2\alpha_k^{(n)}) + h \left((\alpha_k^{(n)})^2 - (\alpha_{k-1}^{(n)})^2 \right) + o(h).$$

This together with (2.6) gives that

$$\alpha_k^{(n)} = \frac{\sigma^2(X_{k-1}^{(n)}) - 1}{2}, \quad k = 0, ..., n$$

is a solution, where we set $X_{-1}^{(n)} \equiv x - \sqrt{h}$. Next, (2.5) yields that

$$\mathbb{E}_n \left(\xi_k^{(n)} | \xi_1^{(n)}, ..., \xi_{k-1}^{(n)} \right) = \frac{\alpha_{k-1}^{(n)} \xi_{k-1}^{(n)} + \sqrt{h} \mu(X_{k-1}^{(n)})}{1 + \alpha_k^{(n)}}.$$

Recall that $\xi_k^{(n)} \in \{-1, 1\}$. We conclude that the probability measure $\mathbb{P}_n$ is given by (prior to absorbing time)

(2.7)

$$\mathbb{P}_n \left(\xi_k^{(n)} = \pm 1 | \xi_1^{(n)}, ..., \xi_{k-1}^{(n)} \right) = \frac{1}{2} \left(1 \pm \frac{\alpha_{k-1}^{(n)} \xi_{k-1}^{(n)} + \sqrt{h} \mu(X_{k-1}^{(n)})}{1 + \alpha_k^{(n)}} \right), \quad k = 1, ..., n.$$

In view of Assumption 2.1.1 we assume that n is sufficiently large so $\mathbb{P}_n$ is indeed a probability measure. Moreover, we notice that $\alpha_{k-1}^{(n)} = \frac{\sigma^2(X_{k-1}^{(n)} - \sqrt{h}\xi_{k-1}^{(n)}) - 1}{2}$. Thus the right hand side of (2.7) is determined by $X_{k-1}^{(n)}, \xi_{k-1}^{(n)}$, and so $\{X_k^{(n)}, \xi_k^{(n)}\}_{k=0}^n$ is indeed a Markov chain.

Optimal Stopping Problem on a Binomial Tree. For any n denote by $\mathcal{T}_n$ the (finite) set of all stopping times with respect to filtration $\sigma\{\xi_1^{(n)}, ..., \xi_k^{(n)}\}$, $k = 0, ..., n$. Introduce the optimal stopping value

(2.8)
$$V_n := \max_{\eta \in \mathcal{T}_n} \mathbb{E}_n [f(\eta h, X_\eta^{(n)})].$$

By using standard dynamical programming for optimal stopping (see [67] Chapter I) we can calculate V_n and the rational stopping times by the following backward recursion. For any $k = 0, 1, ..., n$ denote by $G_k^{(n)}$ all the points on the grid $x + \sqrt{h}\{-k, 1-k, ..., 0, 1, ..., k\}$ which lie in the interval $[B_n, C_n]$.

Define the functions

$$\mathfrak{J}_k^{(n)} : G_k^{(n)} \times \{-1, 1\} \to \mathbb{R}, \quad k = 0, 1, ..., n$$

$$\mathfrak{J}_n^{(n)}(z, y) = f(T, z).$$

For $k = 0, 1, ..., n-1$, if $z \in (B_n, C_n)$,

$$\mathfrak{J}_k^{(n)}(z, y) = \max\left\{ f(kh, z), \sum_{i=1}^{2} \frac{1}{2}\left(1 + (-1)^i \frac{\alpha' y + \sqrt{h}\mu(z)}{1+\alpha}\right) \mathfrak{J}_{k+1}^{(n)}(z + (-1)^i\sqrt{h}, i)\right\},$$

where $\alpha' = \frac{\sigma^2(z - y\sqrt{h}) - 1}{2}$, $\alpha = \frac{\sigma^2(z)-1}{2}$, and

$$\mathfrak{J}_k^{(n)}(z, y) = \max_{k \le m \le n} f(mh, z) \quad \text{if } z \in \{B_n, C_n\}.$$

We get that

$$V_n = \mathfrak{J}_0^{(n)}(x, 1)$$

and the stopping time

$$\eta_n^* = n \wedge \min\left\{k : \mathfrak{J}_k^{(n)}(X_k^{(n)}, \xi_k^{(n)}) = f(kh, X_k^{(n)})\right\} \tag{2.9}$$

is a rational stopping time. Namely,

$$V_n = \mathbb{E}_n[f(\eta_n^* h, X_{\eta_n^*}^{(n)})].$$

Skorohod Embedding. Before we formulate our first result let us introduce the Skorokhod embedding which allows to consider the random variables $\{\xi_k^{(n)}, X_k^{(n)}\}_{k=0}^n$ (without changing their joint distribution) and the diffusion X on the same probability space.

Set,

$$M_t := x + \int_0^t \sigma(Y_s)dW_s = Y_t - \int_0^t \mu(Y_s)ds, \quad t \geq 0.$$

The main idea is to embed the process

$$\tilde{X}_k^{(n)} := X_k^{(n)} + \sqrt{h}(\alpha_k^{(n)}\xi_k^{(n)} - \alpha_0^{(n)}\xi_0^{(n)}) - h\sum_{i=0}^{k-1}\mu(X_i^{(n)}), \quad k = 0, 1, ..., n$$

into the martingale $\{M_t\}_{t=0}^{\infty}$. Observe that prior the absorbing time, the process $\{\tilde{X}_k^{(n)}\}_{k=0}^{n}$ is a martingale with respect to the measure $\mathbb{P}_n$, and $\tilde{X}_0^{(n)} = x$.

Set,

$$\alpha_0^{(n)} = \frac{\sigma^2(x - \sqrt{h}) - 1}{2}, \quad \theta_0^{(n)} = 0, \quad \xi_0^{(n)} = 1, \quad X_0^{(n)} = x.$$

For $k = 0, 1, ..., n-1$ define by recursion the following random variables

$$\alpha_{k+1}^{(n)} = \frac{\sigma^2(X_k^{(n)}) - 1}{2},$$

If $X_k^{(n)} \in (B_n, C_n)$ then

(2.10)

$$\theta_{k+1}^{(n)} = \inf\left\{t > \theta_k^{(n)} : |M_t - M_{\theta_k^{(n)}} + \sqrt{h}\alpha_k^{(n)}\xi_k^{(n)} + h\mu(X_k^{(n)})| = \sqrt{h}(1 + \alpha_{k+1}^{(n)})\right\},$$

(2.11) $$\xi_{k+1}^{(n)} = \mathbb{I}_{\theta_{k+1}^{(n)} < \infty} sgn\left(M_{\theta_{k+1}^{(n)}} - M_{\theta_k^{(n)}} + \sqrt{h}\alpha_k^{(n)}\xi_k^{(n)} + h\mu(X_k^{(n)})\right),$$

where we put $sgn(z) = 1$ for $z > 0$ and $= -1$ otherwise, and

(2.12) $$X_{k+1}^{(n)} = X_k^{(n)} + \sqrt{h}\xi_{k+1}^{(n)}.$$

If $X_k^{(n)} \notin (B_n, C_n)$ then $\theta_{k+1}^{(n)} = \theta_k^{(n)} + h$ and $X_{k+1}^{(n)} = X_k^{(n)}$. Set, $\Theta_n := n \wedge \min\{k : X_k^{(n)} \notin (B_n, C_n)\}$ and observe that on the event $k < \Theta_n$, $\theta_{k+1}^{(n)}$ is the stopping time which corresponds to the Skorokhod embedding of the binary random variable with values in the (random) set $\{\pm\sqrt{h}(1 + \alpha_{k+1}^{(n)}) - \sqrt{h}\alpha_k^{(n)}\xi_k^{(n)} - h\mu(X_k^{(n)})\}$, into the martingale $\{M_t - M_{\theta_k^{(n)}}\}_{t \geq \theta_k^{(n)}}$. Moreover, the grid structure of $X^{(n)}$ implies that if $X_k^{(n)} \notin (B_n, C_n)$ then $X_k^{(n)} \in \{B_n, C_n\}$.

Lemma 2.1.1. *The stopping times* $\{\theta_k^{(n)}\}_{k=0}^n$ *have a finite mean and the random variables* $\{\xi_k^{(n)}, X_k^{(n)}\}_{k=0}^n$ *satisfy* (2.7).

Proof. For sufficiently large n we have that for any $k < n$

$$-\sqrt{h}(1+\alpha_{k+1}^{(n)})-\sqrt{h}\alpha_k^{(n)}\xi_k^{(n)}-h\mu(X_k^{(n)}) < 0 < \sqrt{h}(1+\alpha_{k+1}^{(n)})-\sqrt{h}\alpha_k^{(n)}\xi_k^{(n)}-h\mu(X_k^{(n)}).$$

Thus by using the fact that volatility of the martingale M is bounded away from zero, we conclude $\mathbb{E}(\theta_{k+1}^{(n)} - \theta_k^{(n)}) < \infty$ for all $k < n$. Hence the stopping times $\{\theta_k^{(n)}\}_{k=0}^n$ have a finite mean.

Next, we establish (2.7) for the redefined $\{\xi_k^{(n)}, X_k^{(n)}\}_{k=0}^n$. Fix k and consider the event $k < \Theta_n$. From (2.10) we get

$$M_{\theta_{k+1}^{(n)}} - M_{\theta_k^{(n)}} = \sqrt{h}(1 + \alpha_{k+1}^{(n)})\xi_{k+1}^{(n)} - \sqrt{h}\alpha_k^{(n)}\xi_k^{(n)} - h\mu(X_k^{(n)}). \tag{2.13}$$

The stochastic process

$$\{M_t - M_{\theta_k^{(n)}}\}_{t=\theta_k^{(n)}}^{\theta_{k+1}^{(n)}}$$

is a bounded martingale and so $\mathbb{E}(M_{\theta_{k+1}^{(n)}} - M_{\theta_k^{(n)}}|\mathcal{F}_{\theta_k^{(n)}}) = 0$. Hence, from (2.13)

$$\mathbb{E}(\xi_{k+1}^{(n)}|\mathcal{F}_{\theta_k^{(n)}}) = \frac{\alpha_k^{(n)}\xi_k^{(n)} + \sqrt{h}\mu(X_k^{(n)})}{1 + \alpha_{k+1}^{(n)}}.$$

Since $\xi_{k+1}^{(n)} \in \{-1, 1\}$ we arrive at

$$\mathbb{P}(\xi_{k+1}^{(n)} = \pm 1|\mathcal{F}_{\theta_k^{(n)}}) = \frac{1}{2}\left(1 \pm \frac{\alpha_k^{(n)}\xi_k^{(n)} + \sqrt{h}\mu(X_k^{(n)})}{1 + \alpha_{k+1}^{(n)}}\right)$$

and conclude that (the above right hand side is $\sigma\{\xi_1^{(n)}, ..., \xi_k^{(n)}\}$ measurable) (2.7) holds true. □

The first Main Result.

Theorem 2.1.1. *The values* V *and* V_n *defined by (2.3) and (2.8), respectively satisfy*

$$|V_n - V| = O(n^{-1/4}).$$

Moreover, if we consider the random variables $\xi_0^{(n)}, ..., \xi_n^{(n)}$ *defined by (2.10)–(2.12) and denote* $\tau_n^* \in \mathcal{T}_{[0,T]}$ *by* $\tau_n^* = T \wedge \theta_{\eta_n^*}^{(n)}$, *in which* η_n^* *is from (2.9) and* $\theta_k^{(n)}$ *is from (2.10), then*

$$V - \mathbb{E}[f(\tau_n^*, X_{\tau_n^*})] = O(n^{-1/4}).$$

2.1.3 Trinomial Tree Approximation

In this section we relax the Lipschitz continuity requirement and assume the following.

Assumption 2.1.2. *The functions* $\mu, \sigma, \frac{1}{\sigma} : (B, C) \to \mathbb{R}$ *are bounded and measurable.*

As our assumption indicates the results in this section apply for diffusions with discontinuous coefficients. See Section 2.3.5 for a pricing problem for a regime switching volatility example. See also [4] for other applications of such models.

Remark 2.1.1. *From the general theory of one dimensional, time–homogeneous SDE (see Section 5.5 in [49]) it follows that if* $\sigma, \mu : \mathbb{R}$ *are measurable functions such that* $\sigma(z) \neq 0$ *for all* $z \in \mathbb{R}$ *and the function* $|\mu(z)| + |\sigma(z)| + |\sigma^{-1}(z)|$ *is uniformly bonded, then the SDE (2.1) has a unique weak solution. Since the distribution of* X *is determined only by the values of* μ, σ *in the interval* (B, C) *we obtain (by letting* $\sigma, \mu \equiv 1$ *outside of the interval* (B, C)*) that Assumption 2.1.2 above is sufficient for an existence and uniqueness in law, of the absorbed diffusion* X*. Clearly, Assumption 2.1.1 is stronger than Assumption 2.1.2. A relevant reference here is [4] which not only considers the existence of weak solutions to SDEs but also of their Malliavin differentiability.*

In this section we assume that $x, B, C \in \mathbb{Q} \cup \{-\infty, \infty\}$ (recall that the barriers B and C can take the values $-\infty$ and ∞, respectively), and so for sufficiently large n, we can choose a constant $\bar{\sigma} = \bar{\sigma}(n) > \sup_{y\in(B,C)}|\sigma(y)| + \sqrt{h}\sup_{y\in(B,C)}|\mu(y)|$ which satisfies $\frac{C-x}{\bar{\sigma}\sqrt{h}}, \frac{x-B}{\bar{\sigma}\sqrt{h}} \in \mathbb{N}\cup\{\infty\}$.

Trinomial Approximation of the State Process. The main idea in this section is to find stopping times $0 = \vartheta_0^{(n)} < \vartheta_1^{(n)} < ... < \vartheta_n^{(n)}$, such that for any $k = 0, 1, ..., n-1$

$$X_{\vartheta_{k+1}^{(n)}} - X_{\vartheta_k^{(n)}} \in \{-\bar{\sigma}\sqrt{h}, 0, \bar{\sigma}\sqrt{h}\} \quad \text{and} \quad \mathbb{E}(\vartheta_{k+1}^{(n)} - \vartheta_k^{(n)}|\mathcal{F}_{\vartheta_k^{(n)}}) = h + O(h^{3/2}). \tag{2.1.4}$$

In this case the random variables $\{X_{\vartheta_k^{(n)}}\}_{k=0}^n$ lie on the grid $x + \bar{\sigma}\sqrt{h}\{-b_n, 1-b_n, ..., 0, 1,$
$..., c_n\}$ where $b_n = n \wedge \frac{x-B}{\bar{\sigma}\sqrt{h}}$ and $c_n = n \wedge \frac{C-x}{\bar{\sigma}\sqrt{h}}$. Moreover, we will see that

$$\max_{0\le k\le n}|\vartheta_k^{(n)} - kh| = O(\sqrt{h}).$$

Skorohod Embedding. Next, we describe the construction. For any initial position $B + \bar{\sigma}\sqrt{h} \le X_0 \le C - \bar{\sigma}\sqrt{h}$ and $A \in [0, \bar{\sigma}\sqrt{h}]$ consider the stopping times

$$\rho_A^{X_0} = \inf\{t : |X_t - X_0| = A\} \quad \text{and}$$

$$\kappa_A^{X_0} = \sum_{i=1}^2 \mathbb{I}_{X_{\rho_A^{X_0}} = X_0 + (-1)^i A} \inf\{t \ge \rho_A^{X_0} : X_t = X_0 \text{ or } X_t = X_0 + (-1)^i\bar{\sigma}\sqrt{h}\}.$$

Observe that $X_{\kappa_A^{X_0}} - X_0 \in \{-\bar{\sigma}\sqrt{h}, 0, \bar{\sigma}\sqrt{h}\}$. Let us prove the following lemma.

Lemma 2.1.2. *There exists a unique $\hat{A} = \hat{A}(X_0, n) \in (0, \bar{\sigma}\sqrt{h}]$ such that $\mathbb{E}(\kappa_{\hat{A}}^{X_0}) = h$.*

Proof. Clearly, for any $A_1 < A_2$ $\kappa_{A_1}^{X_0} \le \kappa_{A_2}^{X_0}$ a.s. Since $\kappa_{A_1}^{X_0} \ne \kappa_{A_2}^{X_0}$ for $A_1 \ne A_2$ we conclude that the function $g(A) := \mathbb{E}(\kappa_A^{X_0})$ is strictly increasing which satisfies $g(0) = 0$. Thus in order to complete the proof it remains to show that $g(\bar{\sigma}\sqrt{h}) \ge h$. Observe that $g(\bar{\sigma}\sqrt{h}) = \mathbb{E}(\rho_{\bar{\sigma}\sqrt{h}}^{X_0})$. Assume (by contradiction) that $\mathbb{E}(\rho_{\bar{\sigma}\sqrt{h}}^{X_0}) < h$. From the Itô isometry and the Jensen Inequality we obtain

$$\bar{\sigma}\sqrt{h} = \mathbb{E}|X_{\rho^{X_0}_{\bar{\sigma}\sqrt{h}}} - X_0| \leq \mathbb{E}\left|\int_0^{\rho^{X_0}_{\bar{\sigma}\sqrt{h}}} \sigma(X_t)dW_t\right| + \mathbb{E}\left|\int_0^{\rho^{X_0}_{\bar{\sigma}\sqrt{h}}} \mu(X_t)dt\right|$$
$$\leq \sup_{y\in\mathbb{R}}|\sigma(y)|\sqrt{\mathbb{E}(\rho^{X_0}_{\bar{\sigma}\sqrt{h}})} + \sup_{y\in\mathbb{R}}|\mu(y)|\mathbb{E}(\rho^{X_0}_{\bar{\sigma}\sqrt{h}}) < \bar{\sigma}\sqrt{h}.$$

This clearly a contradiction, and the result follows. □

Next, we recall the theory of exit times of Markov diffusions (see Section 5.5 in [49]). Set,

(2.15)
$$p(y) = \int_{X_0}^{y} \exp\left(-2\int_{X_0}^{z} \frac{\mu(w)}{\sigma^2(w)}dw\right)dz,$$
$$G_{a,b}(y,z) = \frac{(p(y\wedge z)-p(a))(p(b)-p(y\vee z))}{p(b)-p(a)}, \quad a \leq y, z \leq b,$$
$$M_{a,b}(y) = \int_a^b \frac{2G_{a,b}(y,z)}{p'(z)\sigma^2(z)}dz, \quad y \in [a,b].$$

Then for any $A \in [0, \bar{\sigma}\sqrt{h}]$

(2.16)
$$\mathbb{E}(\kappa^{X_0}_A) = \mathbb{E}(\rho^{X_0}_A) + \mathbb{P}(X_{\rho^{X_0}_A} = X_0 + A)M_{X_0,X_0+\bar{\sigma}\sqrt{h}}(X_0 + A)$$
$$+ \mathbb{P}(X_{\rho^{X_0}_A} = X_0 - A)M_{X_0-\bar{\sigma}\sqrt{h},X_0}(X_0 - A)$$
$$= M_{X_0-A,X_0+A}(X_0) + \frac{p(X_0) - p(X_0 - A)}{p(X_0 + A) - p(X_0 - A)}M_{X_0,X_0+\bar{\sigma}\sqrt{h}}(X_0 + A)$$
$$+ \frac{p(X_0 + A) - p(X_0)}{p(X_0 + A) - p(X_0 - A)}M_{X_0-\bar{\sigma}\sqrt{h},X_0}(X_0 - A)$$

and

$$q^{(1)}(X_0, A) := \mathbb{P}(X_{\kappa^{X_0}_A} = X_0 + \bar{\sigma}\sqrt{h}) = \frac{(p(X_0)-p(X_0-A))(p(X_0+A)-p(X_0))}{(p(X_0+A)-p(X_0-A))\left(p(X_0+\bar{\sigma}\sqrt{h})-p(X_0)\right)},$$
$$q^{(-1)}(X_0, A) := \mathbb{P}(X_{\kappa^{X_0}_A} = X_0 - \bar{\sigma}\sqrt{h}) = \frac{(p(X_0+A)-p(X_0))(p(X_0)-p(X_0-A))}{(p(X_0+A)-p(X_0-A))\left(p(X_0)-p(X_0-\bar{\sigma}\sqrt{h})\right)},$$
$$q^{(0)}(X_0, A) := \mathbb{P}(X_{\kappa^{X_0}_A} = X_0) = 1 - q^{(1)}(X_0, A) - q^{(-1)}(X_0, A).$$

We aim to find numerically $\hat{A} = \hat{A}(X_0, n) \in [0, \bar{\sigma}\sqrt{h}]$ which satisfies $\mathbb{E}(\kappa^{X_0}_A) = h + O(h^{3/2})$. Observe that $p'(X_0) = 1$. From the Mean value theorem and the fact that

$\frac{\mu}{\sigma^2}$ is uniformly bonded we obtain that for any $X_0 - \bar{\sigma}\sqrt{h} \le a \le y, z \le b \le X_0 + \bar{\sigma}\sqrt{h}$

$$\begin{aligned}\frac{G_{a,b}(y,z)}{p'(z)} =& \frac{\Big((1+O(\sqrt{h}))(y\wedge z - a)\Big)\Big((1+O(\sqrt{h}))(b - y\vee z)\Big)}{(1+O(\sqrt{h}))^2(b-a)} \\ =&(1+O(\sqrt{h}))\frac{(y\wedge z - a)(b - y\vee z)}{(b-a)}.\end{aligned}$$

Hence, for any $X_0 - \bar{\sigma}\sqrt{h} \le a \le y \le b \le X_0 + \bar{\sigma}\sqrt{h}$

$$M_{a,b}(y) = 2\int_a^b \frac{(y\wedge z - a)(b - y\vee z)}{(b-a)\sigma^2(z)}dz + O(h^{3/2}).$$

This together with (2.16) yields

(2.17)

$$\begin{aligned}\mathbb{E}(\kappa_A^{X_0}) =& \int_{X_0-A}^{X_0+A} \frac{(X_0\wedge z + A - X_0)(X_0 + A - X_0\vee z)}{A\sigma^2(z)}dz \\ &+ \int_{X_0}^{X_0+\bar{\sigma}\sqrt{h}} \frac{((X_0+A)\wedge z - X_0)(X_0 + \bar{\sigma}\sqrt{h} - (X_0+A)\vee z)}{\bar{\sigma}\sqrt{h}\sigma^2(z)}dz \\ &+ \int_{X_0-\bar{\sigma}\sqrt{h}}^{X_0} \frac{((X_0-A)\wedge z + \bar{\sigma}\sqrt{h} - X_0)(X_0 - (X_0-A)\vee z)}{\bar{\sigma}\sqrt{h}\sigma^2(z)}dz + O(h^{3/2}).\end{aligned}$$

Thus, $\hat{A} = \hat{A}(X_0, n)$ can be calculated numerically by applying the bisection method and (2.17).

Remark 2.1.2. *If in addition to Assumption 2.1.2 we assume that σ is Lipschitz then (2.17) implies*

$$\begin{aligned}\sigma^2(X_0)\mathbb{E}(\kappa_A^{X_0}) =& A^2 + 2\left(\frac{A^2(\bar{\sigma}\sqrt{h} - A) + A(\bar{\sigma}\sqrt{h} - A)^2}{2\bar{\sigma}\sqrt{h}}\right) + O(h^{3/2}) \\ =& \bar{\sigma}A\sqrt{h} + O(h^{3/2}).\end{aligned}$$

Thus for the case where σ is Lipschitz we set

$$\hat{A}(X_0, n) = \frac{\sigma^2(X_0)}{\bar{\sigma}}\sqrt{h}.$$

Now, we define the Skorokhod embedding by the following recursion. Set $\vartheta_0^{(n)} = 0$ and for $k = 0, 1, ..., n-1$

$$\vartheta_{k+1}^{(n)} = \mathbb{I}_{X_{\vartheta_k^{(n)}} \in (B,C)} \, \kappa^{X_{\vartheta_k^{(n)}}}_{\hat{A}(X_{\vartheta_k^{(n)}}, n)} + \mathbb{I}_{X_{\vartheta_k^{(n)}} \notin (B,C)} (\vartheta_k^{(n)} + h). \tag{2.18}$$

From the definition of κ and $\hat{A}(\cdot, n)$ it follows that (2.14) holds true.

Optimal Stopping of the Trinomial Model. Denote by $\mathcal{S}_n$ the set of all stopping times with respect to the filtration $\{\sigma(X_{\vartheta_1^{(n)}}, ..., X_{\vartheta_k^{(n)}})\}_{k=0}^n$, with values in the set $\{0, 1, ..., n\}$. Introduce the corresponding optimal stopping value

$$\tilde{V}_n := \max_{\eta \in \mathcal{S}_n} \mathbb{E}[f(\eta h, X_{\vartheta_\eta^{(n)}})]. \tag{2.19}$$

As before, $\tilde{V}_n$ and the rational stopping times can be found by applying dynamical programming. Thus, define the functions

$$\mathcal{J}_k^{(n)} : \{x + \bar{\sigma}\sqrt{h}\{-(k \wedge b_n), 1 - (k \wedge b_n), ..., 0, 1, ..., k \wedge c_n\}\} \to \mathbb{R}, \quad k = 0, 1, ..., n$$

$$\mathcal{J}_n^{(n)}(z) = f(T, z).$$

For $k = 0, 1, ..., n-1$

$$\mathcal{J}_k^{(n)}(z) = \max\left(f(kh, z), \sum_{i=-1,0,1} q^{(i)}(z, \hat{A}(z,n)) \mathcal{J}_{k+1}^{(n)}(z + i\bar{\sigma}\sqrt{h})\right) \text{ if } z \in (B, C)$$

and

$$\mathcal{J}_k^{(n)}(z) = \max_{k \le m \le n} f(mh, z) \text{ if } z \in \{B, C\}.$$

We get that

$$\tilde{V}_n = \mathcal{J}_0^{(n)}(x)$$

and the stopping times given by

$$\tilde{\eta}_n^* = n \wedge \min\left\{k : \mathcal{J}_k^{(n)}(X_{\vartheta_k^{(n)}}) = f(kh, X_{\vartheta_k^{(n)}})\right\} \tag{2.20}$$

satisfies

$$\tilde{V}_n = \mathbb{E}[f(\tilde{\eta}^*_n h, X_{\vartheta^{(n)}_{\tilde{\eta}^*_n}})].$$

The second main result.

Theorem 2.1.2. *The values V and $\tilde{V}_n$ defined by (2.3) and (2.19) satisfy*

$$|V - \tilde{V}_n| = O(n^{-1/4}).$$

Moreover, if we denote $\tilde{\tau}^*_n = T \wedge \vartheta^{(n)}_{\tilde{\eta}^*_n}$, where $\vartheta^{(n)}$ is defined by (2.18) and $\tilde{\eta}^*_n$ by (2.20), then

$$V - \mathbb{E}[f(\tau^*_n, X_{\tau^*_n})] = O(n^{-1/4}).$$

Remark 2.1.3. *Theorems 2.1.1 and 2.1.2 can be extended with the same error estimates to the setup of Dynkin games which are corresponding to game options (see [50, 52]). The dynamical programming in discrete time can be done in a similar way by applying the results from [66]. Moreover, as in the American options case, the Skorokhod embedding technique allows to lift the rational times from the discrete setup to the continuous one. Since the proof for Dynkin games is very similar to our setup, then for simplicity, in this work we focus on optimal stopping and the pricing of American options.*

2.1.4 Proof of Theorem 2.1.1

Fix $n \in \mathbb{N}$. Recall the definition of $\theta^{(n)}_k, \mathcal{T}_n, \eta^*_n$ from Section 2.1.2. Denote by $\mathbb{T}_n$ the set of all stopping times with respect to the filtration $\{\mathcal{F}_{\theta^{(n)}_k}\}_{k=0}^n$, with values in $\{0, 1, ..., n\}$. Clearly, $\mathcal{T}_n \subset \mathbb{T}_n$. From the strong Markov property of the diffusion X it follows that

$$V_n = \sup_{\eta \in \mathbb{T}_n} \mathbb{E}[f(\eta_n h, X^{(n)}_{\eta_n})] = \mathbb{E}[f(\eta^*_n h, X^{(n)}_{\eta^*_n})]. \tag{2.21}$$

Define the function $\phi_n : \mathcal{T}_{[0,T]} \to \mathbb{T}_n$ by $\phi_n(\tau) = n \wedge \min\{k : \theta_k^{(n)} \geq \tau\}$. The equality (2.21) implies that $V_n \geq \sup_{\tau \in \mathcal{T}_{[0,T]}^X} \mathbb{E}[f(\phi_n(\tau)h, X_{\phi_n(\tau)}^{(n)})]$. Hence, from (2.2) we obtain

$$\begin{aligned}(2.22) \qquad V \leq V_n + O(1) \sup_{\tau \in \mathcal{T}_{[0,T]}} \mathbb{E}|X_\tau - X_{\theta_{\phi_n(\tau)}^{(n)}}| + O(1)\mathbb{E}\left(\max_{0 \leq i \leq n} |X_i^{(n)} - X_{\theta_i^{(n)}}|\right) \\ + O(1)\mathbb{E}\left(\sup_{0 \leq t \leq T} X_t \left(\max_{1 \leq i \leq n} |\theta_i^{(n)} - ih| + \max_{1 \leq i \leq n} (\theta_i^{(n)} - \theta_{i-1}^{(n)})\right)\right).\end{aligned}$$

Next, recall the definition of τ_n^* and observe that

$$\begin{aligned}(2.23) \quad V \geq & \mathbb{E}[f(\tau_n^*, X_{\tau_n^*})] \geq V_n - O(1)\mathbb{E}\left(\max_{0 \leq i \leq n} |X_i^{(n)} - X_{\theta_i^{(n)}}|\right) \\ & - O(1) \sup_{\theta_n^{(n)} \wedge T \leq t \leq \theta_n^{(n)} \vee T} |X_t - X_{\theta_n^{(n)} \wedge T}| - O(1)\mathbb{E}\left(\sup_{0 \leq t \leq T} X_t \max_{1 \leq i \leq n} |\theta_i^{(n)} - ih|\right).\end{aligned}$$

Let $\Theta := \inf\{t : X_t \notin (B, C)\}$ be the absorbing time. From the Burkholder–Davis–Gundy inequality and the inequality $(a+b)^m \leq 2^m(a^m + b^m)$, $a, b \geq 0$, $m \geq 1$ it follows that for any $m > 1$ and stopping times $\varsigma_1 \leq \varsigma_2$

(2.24)

$$\begin{aligned}\mathbb{E}\left(\sup_{\varsigma_1 \leq t \leq \varsigma_2} |X_t - X_s|^m\right) &\leq 2^m \mathbb{E}\left(\sup_{\varsigma_1 \wedge \Theta \leq t \leq \varsigma_2 \wedge \Theta} |M_t - M_{\varsigma_1 \wedge \Theta}|^m + ||\mu||_\infty^m (\varsigma_2 \wedge \Theta - \varsigma_1 \wedge \Theta)^m\right) \\ &= O(1)\mathbb{E}\left(\left|\int_{\varsigma_1 \wedge \Theta}^{\varsigma_2 \wedge \Theta} \sigma^2(X_t) dt\right|^{m/2} + (\varsigma_2 - \varsigma_1)^m\right) \\ &= O(1)\mathbb{E}\left((\varsigma_2 - \varsigma_1)^{m/2} + (\varsigma_2 - \varsigma_1)^m\right).\end{aligned}$$

Observe that for any stopping time $\tau \in \mathcal{T}_{[0,T]}$, $|\tau - \theta_{\phi_n(\tau)}^{(n)}| \leq \max_{1 \leq i \leq n}(\theta_i^{(n)} - \theta_{i-1}^{(n)}) + |T - \theta_n^{(n)}|$. Moreover, $2\max_{1 \leq i \leq n} |\theta_i^{(n)} - ih| + h \geq \max_{1 \leq i \leq n}(\theta_i^{(n)} - \theta_{i-1}^{(n)})$. Thus Theorem 2.1.1 follows from (2.22)–(2.24), the Cauchy–Schwarz inequality, the Jensen inequality and Lemmas 2.1.4–2.1.5 below.

2.1.5 Technical estimates for the proof of Theorem 2.1.1

The next lemma is a technical step in proving Lemmas 2.1.4–2.1.5 which are the main results of this subsection, which are then used for the proof of Theorem 2.1.1.

Lemma 2.1.3. *Recall the definition of Θ_n given after (2.12). For any $m > 0$*

$$\mathbb{E}\left(\max_{0\le k\le\Theta_n} |Y_{\theta^{(n)}_k} - X^{(n)}_k|^m\right) = O(h^{m/2}).$$

Proof. From the Jensen inequality it follows that it is sufficient to prove the claim for $m > 2$. Fix $m > 2$ and $n \in \mathbb{N}$. We will apply the discrete version of the Gronwall inequality. Introduce the random variables $U_k := \mathbb{I}_{k\le\Theta_n}|X^{(n)}_k - Y_{\theta^{(n)}_k}|$. Since the process $\alpha^{(n)}$ is uniformly bounded and μ is Lipschitz continuous, then from (2.4) and (2.13) we obtain that

$$\begin{aligned}
U_k =& O(\sqrt{h}) + \mathbb{I}_{k\le\Theta_n}\left|h\sum_{i=0}^{k-1}\mu(X^{(n)}_i) - \sum_{i=0}^{k-1}\int_{\theta^{(n)}_i}^{\theta^{(n)}_{i+1}}\mu(Y_t)dt\right| \\
\le& O(\sqrt{h}) + \mathbb{I}_{k\le\Theta_n}\left|h\sum_{i=0}^{k-1}\mu(X^{(n)}_i) - \sum_{i=0}^{k-1}\mu(Y_{\theta^{(n)}_i})(\theta^{(n)}_{i+1} - \theta^{(n)}_i)\right| \\
&+O(1)\sum_{i=0}^{k-1}\mathbb{I}_{i<\Theta_n}\sup_{\theta^{(n)}_i\le t\le\theta^{(n)}_{i+1}}|Y_t - Y_{\theta^{(n)}_i}|(\theta^{(n)}_{i+1} - \theta^{(n)}_i) \\
\le& O(\sqrt{h}) + O(1)\sum_{i=0}^{k-1}L_i + O(h)\sum_{i=0}^{k-1}U_i + |\sum_{i=0}^{k-1}(I_i + J_i)|
\end{aligned} \tag{2.25}$$

where

$$\begin{aligned}
I_i :=& \mathbb{I}_{i<\Theta_n}\mu(Y_{\theta^{(n)}_i})\left(\theta^{(n)}_{i+1} - \theta^{(n)}_i - \mathbb{E}(\theta^{(n)}_{i+1} - \theta^{(n)}_i|\mathcal{F}_{\theta^{(n)}_i})\right), \\
J_i :=& \mathbb{I}_{i<\Theta_n}\mu(Y_{\theta^{(n)}_i})\left(\mathbb{E}(\theta^{(n)}_{i+1} - \theta^{(n)}_i|\mathcal{F}_{\theta^{(n)}_i}) - h\right), \\
L_i :=& \mathbb{I}_{i<\Theta_n}\sup_{\theta^{(n)}_i\le t\le\theta^{(n)}_{i+1}}|Y_t - Y_{\theta^{(n)}_i}|(\theta^{(n)}_{i+1} - \theta^{(n)}_i).
\end{aligned}$$

Next, from (2.7), (2.13) and the Itô Isometry it follows that on the event $i < \Theta_n$

$$\begin{aligned}
&\mathbb{E}\left(\int_{\theta^{(n)}_i}^{\theta^{(n)}_{i+1}}\sigma^2(Y_t)dt|\mathcal{F}_{\theta^{(n)}_i}\right) = \mathbb{E}\left((M_{\theta^{(n)}_{i+1}} - M_{\theta^{(n)}_i})^2|\mathcal{F}_{\theta^{(n)}_i}\right) \\
=& h(1 + 2\alpha^{(n)}_{i+1}) + h\left((\alpha^{(n)}_{i+1})^2 - (\alpha^{(n)}_i)^2\right) + O(h^{3/2}) = h\sigma^2(X^{(n)}_i) + O(h^{3/2}).
\end{aligned} \tag{2.26}$$

On the other hand, the function σ is bounded and Lipschitz, and so

(2.27)

$$\begin{aligned}\mathbb{E}\left(\int_{\theta_i^{(n)}}^{\theta_{i+1}^{(n)}} \sigma^2(Y_t)dt|\mathcal{F}_{\theta_i^{(n)}}\right) =&\sigma^2(X_i^{(n)})\mathbb{E}(\theta_{i+1}^{(n)}-\theta_i^{(n)}|\mathcal{F}_{\theta_i^{(n)}}) \\ &+O(1)\mathbb{E}\left(\sup_{\theta_i^{(n)}\le t\le\theta_{i+1}^{(n)}} |Y_t-Y_{\theta_i^{(n)}}|(\theta_{i+1}^{(n)}-\theta_i^{(n)})|\mathcal{F}_{\theta_i^{(n)}}\right) \\ &+O(1)\left|X_i^{(n)}-Y_{\theta_i^{(n)}}\right|\mathbb{E}\left(\theta_{i+1}^{(n)}-\theta_i^{(n)}|\mathcal{F}_{\theta_i^{(n)}}\right).\end{aligned}$$

From (2.26)–(2.27) and the fact that σ bounded away from zero we get that on the event $i<\Theta_n$

$$\begin{aligned}(2.28)\qquad \mathbb{E}&\ (\theta_{i+1}^{(n)}-\theta_i^{(n)}|\mathcal{F}_{\theta_i^{(n)}}) \\ &= h+O(h^{3/2})+O(1)\mathbb{E}\left(\sup_{\theta_i^{(n)}\le t\le\theta_{i+1}^{(n)}} |Y_t-Y_{\theta_i^{(n)}}|(\theta_{i+1}^{(n)}-\theta_i^{(n)})|\mathcal{F}_{\theta_i^{(n)}}\right) \\ &+ O(1)\left|X_i^{(n)}-Y_{\theta_i^{(n)}}\right|\mathbb{E}\left(\theta_{i+1}^{(n)}-\theta_i^{(n)}|\mathcal{F}_{\theta_i^{(n)}}\right).\end{aligned}$$

Clearly, (2.26) implies that (σ is bounded away from zero) on the event $i<\Theta_n$, $\mathbb{E}(\theta_{i+1}^{(n)}-\theta_i^{(n)}|\mathcal{F}_{\theta_i^{(n)}})=O(h)$. This together with (2.28) gives (μ is bounded)

$$(2.29)\qquad |J_i|=O(h^{3/2})+O(h)U_i+O(1)\mathbb{E}(L_i|\mathcal{F}_{\theta_i^{(n)}}).$$

From (2.25) and (2.29) we obtain that for any $k=1,...,n$

$$\begin{aligned}\max_{0\le j\le k} U_j &= O(\sqrt{h})+O(h)\sum_{i=0}^{k-1}\max_{0\le j\le i} U_j \\ &+ \max_{0\le k\le n-1}|\sum_{i=0}^{k} I_i|+O(1)\sum_{i=0}^{n-1} L_i+O(1)\sum_{i=0}^{n-1}\mathbb{E}(L_i|\mathcal{F}_{\theta_i^{(n)}}).\end{aligned}$$

Next, recall the following inequality, which is a direct consequence of the Jensen's inequality,

$$(2.30)\qquad (\sum_{i=1}^{n} a_i)^{\tilde{m}}\le n^{\tilde{m}-1}\sum_{i=1}^{n} a_i^{\tilde{m}},\quad a_1,...,a_n\ge 0,\ \tilde{m}\ge 1.$$

Using the above inequality along with Jensen's inequality we arrive at

$$\begin{aligned}\mathbb{E}\left(\max_{0\le j\le k} U_j^m\right) &= O(h^{m/2}) + O(h)\sum_{i=0}^{k-1}\mathbb{E}\left(\max_{0\le j\le i} U_j^m\right)\\ &+ O(1)\mathbb{E}\left(\max_{0\le k\le n-1}|\sum_{i=0}^{k} I_i|^m\right) + O(n^{m-1})\sum_{i=0}^{n-1}\mathbb{E}(L_i^m).\end{aligned}$$

From the discrete version of Gronwall's inequality (see [21])

$$\begin{aligned}&(2.31)\quad \mathbb{E}\left(\max_{0\le i\le n} U_i^m\right)\\ &= (1+O(h))^n \times \left(O(h^{m/2}) + O(1)\mathbb{E}\left(\max_{0\le k\le n-1}|\sum_{i=0}^{k} I_i|^m\right) + O(n^{m-1})\sum_{i=0}^{n-1}\mathbb{E}(L_i^m)\right)\\ &= O(h^{m/2}) + O(1)\mathbb{E}\left(\max_{0\le k\le n-1}|\sum_{i=0}^{k} I_i|^m\right) + O(n^{m-1})\sum_{i=0}^{n-1}\mathbb{E}(L_i^m).\end{aligned}$$

Next, we estimate $\mathbb{E}\left(\max_{0\le k\le n-1}|\sum_{i=0}^{k} I_i|^m\right)$ and $\mathbb{E}(L_i^m|\mathcal{F}_{\theta_i^{(n)}})$, $i = 0,1,...,n-1$. By applying the Burkholder—Davis—Gundy inequality for the martingale $\{M_t - M_{\theta_i^{(n)}}\}_{t=\theta_i^{(n)}}^{\theta_{i+1}^{(n)}}$ it follows that for any $\tilde{m} > 1/2$

$$\begin{aligned}&\mathbb{I}_{i<\Theta_n}\mathbb{E}\left(\left(\int_{\theta_i^{(n)}}^{\theta_{i+1}^{(n)}} \sigma^2(Y_t)dt\right)^{\tilde{m}} \middle| \mathcal{F}_{\theta_i^{(n)}}\right)\\ &=O(1)\mathbb{I}_{i<\Theta_n}\mathbb{E}\left(\max_{\theta_i^{(n)}\le t\le\theta_{i+1}^{(n)}}\left(M_t - M_{\theta_i^{(n)}}\right)^{2\tilde{m}} \middle| \mathcal{F}_{\theta_i^{(n)}}\right) = O(h^{\tilde{m}})\mathbb{I}_{i<\Theta_n}\end{aligned}$$

where the last equality follows from the fact that

$$\mathbb{I}_{i<\Theta_n}\max_{\theta_i^{(n)}\le t\le\theta_{i+1}^{(n)}}|M_t - M_{\theta_i^{(n)}}| = O(\sqrt{h})\mathbb{I}_{i<\Theta_n}.$$

Since σ is bounded away from zero we get

$$(2.32)\qquad \mathbb{I}_{i<\Theta_n}\mathbb{E}\left((\theta_{i+1}^{(n)} - \theta_i^{(n)})^{\tilde{m}}|\mathcal{F}_{\theta_i^{(n)}}\right) = O(h^{\tilde{m}})\mathbb{I}_{i<\Theta_n}, \quad \tilde{m} > 1/2.$$

Next, observe that $\sum_{i=0}^{k} I_i$, $k = 0,...,n-1$ is a martingale. From the Burkholder-

Davis-Gundy inequality, (2.30), (2.32) and the fact that μ is bounded we conclude

$$
\begin{aligned}
(2.33)\ \mathbb{E}\left(\max_{0\le k\le n-1}|\sum_{i=0}^{k} I_i|^m\right) &\le O(1)\mathbb{E}\left(\left(\sum_{i=0}^{n-1} I_i^2\right)^{m/2}\right) \\
&\le O(1)n^{m/2-1}\sum_{i=0}^{n-1}\mathbb{E}[|I_i|^m] = O(1)n^{m/2-1}nO(h^m) \\
&= O(h^{m/2}).
\end{aligned}
$$

Finally, we estimate $\mathbb{E}(L_i^m|\mathcal{F}_{\theta_i^{(n)}})$ for $i = 0, 1, ..., n-1$. Clearly, on the event $i < \Theta_n$,

$$
\begin{aligned}
\sup_{\theta_i^{(n)}\le t\le\theta_{i+1}^{(n)}} |Y_t - Y_{\theta_i^{(n)}}| \le &\sup_{\theta_i^{(n)}\le t\le\theta_{i+1}^{(n)}} |M_t - M_{\theta_i^{(n)}}| + ||\mu||_\infty(\theta_{i+1}^{(n)} - \theta_i^{(n)}) \\
=&O(\sqrt{h}) + ||\mu||_\infty(\theta_{i+1}^{(n)} - \theta_i^{(n)}).
\end{aligned}
$$

Hence, from (2.32) we get

$$
\begin{aligned}
(2.34)\mathbb{E}(L_i^m|\mathcal{F}_{\theta_i^{(n)}}) &\le \mathbb{I}_{i<\Theta_n}\mathbb{E}\left(O(h^{m/2})(\theta_{i+1}^{(n)} - \theta_i^{(n)})^m + O(1)(\theta_{i+1}^{(n)} - \theta_i^{(n)})^{2m}|\mathcal{F}_{\theta_i^{(n)}}\right) \\
&= O(h^{3m/2}).
\end{aligned}
$$

This together with (2.31) and (2.33) yields $\mathbb{E}\left(\max_{0\le i\le n} U_i^m\right) = O(h^{m/2})$, and the result follows. □

Lemma 2.1.4. *For any $m > 0$*

$$
\mathbb{E}\left(\max_{1\le k\le n}|\theta_k^{(n)} - kh|^m\right) = O(h^{m/2}).
$$

Proof. Fix $m > 2$ and $n \in \mathbb{N}$. Observe (recall the definition after (2.12)) that for $k \ge \Theta_n$ we have $\theta_k^{(n)} - kh = \theta_{\Theta_n}^{(n)} - \Theta_n h$. Hence

$$
(2.35) \qquad \max_{1\le k\le n}|\theta_k^{(n)} - kh| = \max_{1\le k\le\Theta_n}|\theta_k^{(n)} - kh|.
$$

Redefine the terms I_i, J_i from Lemma 2.1.3 as following

$$
I_i := \mathbb{I}_{i<\Theta_n}\left(\theta_{i+1}^{(n)} - \theta_i^{(n)} - \mathbb{E}(\theta_{i+1}^{(n)} - \theta_i^{(n)}|\mathcal{F}_{\theta_i^{(n)}})\right),
$$

$$J_i := \mathbb{I}_{i<\Theta_n}\left(\mathbb{E}(\theta^{(n)}_{i+1}-\theta^{(n)}_i|\mathcal{F}_{\theta^{(n)}_i})-h\right).$$

Then similarly to (2.32)–(2.33) it follows that

$$\mathbb{E}\left(\max_{0\leq k\leq n-1}|\sum_{i=0}^{k} I_i|^m\right)=O(h^{m/2}).$$

Moreover, by applying Lemma 2.1.3, (2.29) and (2.34) we obtain that for any i, $\mathbb{E}[|J_i|^m]=O(h^{3m/2})$. We conclude that

$$\begin{aligned}\mathbb{E}\left(\max_{1\leq k\leq\Theta_n}|\theta^{(n)}_k-kh|^m\right) &= O(1)\mathbb{E}\left(\max_{0\leq k\leq n-1}|\sum_{i=0}^{k} I_i|^m\right)+O(1)\mathbb{E}\left(\max_{0\leq k\leq n-1}|\sum_{i=0}^{k} J_i|^m\right)\\ &= O(h^{m/2}).\end{aligned}$$

This together with (2.35) completes the proof. □

We end this section with establishing the following estimate.

Lemma 2.1.5.

$$\mathbb{E}\left(\max_{0\leq k\leq n}|X_{\theta^{(n)}_k}-X^{(n)}_k|\right)=O(h^{1/4}).$$

Proof. Set $\Gamma_n=\sup_{0\leq t\leq\theta^{(n)}_n}|Y_t|+\max_{0\leq k\leq n}|X^{(n)}_k|$, $n\in\mathbb{N}$. From Lemma 2.1.4 it follows that for any $m>1$, $\sup_{n\in\mathbb{N}}\mathbb{E}[(\theta^{(n)}_n)^m]<\infty$. Thus, from the Burkholder–Davis–Gundy inequality and the fact that μ,σ are bounded we obtain

$$\sup_{n\in\mathbb{N}}\mathbb{E}[\sup_{0\leq t\leq\theta^{(n)}_n}|Y_t|^m]<\infty.$$

This together with Lemma 2.1.3 gives that (recall that $X^{(n)}$ remains constant after Θ_n)

$$\sup_{n\in\mathbb{N}}\mathbb{E}[\Gamma^m_n]<\infty\quad\forall m>1. \tag{2.36}$$

We start with estimating $\mathbb{E}\left(\max_{0\leq k\leq\Theta_n}|X_{\theta^{(n)}_k}-X^{(n)}_k|\right)$. Fix n and introduce the events

$$O=\{\max_{0\le k<\Theta_n}\sup_{\theta^{(n)}_k\le t\le \theta^{(n)}_{k+1}}|Y_t-X^{(n)}_k|\ge n^{-1/3}\},$$

$$O_1=\{\max_{0\le k<\Theta_n}\sup_{\theta^{(n)}_k\le t\le \theta^{(n)}_{k+1}}|Y_t-Y_{\theta^{(n)}_k}|\ge n^{-1/3}/2\},$$

$$O_2=\{\max_{1\le k\le\Theta_n}|X^{(n)}_k-Y_{\theta^{(n)}_k}|\ge n^{-1/3}/2\}.$$

From Lemma 2.1.3 (for $m=6$) and the Markov inequality we get

$$\mathbb{P}(O_2)=O(h).$$

Observe that on the event $k<\Theta_n$, $\sup_{\theta^{(n)}_k\le t\le \theta^{(n)}_{k+1}}|M_t-M_{\theta^{(n)}_k}|=O(\sqrt{h})$. Hence, from the fact that μ is uniformly bounded we obtain that for sufficiently large n

$$\mathbb{P}(O_1)\le\sum_{i=0}^{n-1}\mathbb{P}\left(\mathbb{I}_{i<\Theta_n}(\theta^{(n)}_{i+1}-\theta^{(n)}_i)>h^{1/2}\right)\le O(n)h^4/h^2=O(h)$$

where the last inequality follows form the Markov ineqaulity and (2.32) for $\tilde{m}=4$. Clearly,

$$\mathbb{P}(O)\le\mathbb{P}(O_1)+\mathbb{P}(O_2)=O(h). \tag{2.37}$$

From the simple inequalities $B_n-B, C-C_n\ge n^{-1/3}$ it follows that $\{\exists k\le\Theta_n: X_{\theta^{(n)}_k}\neq Y_{\theta^{(n)}_k}\}\subset O$. Thus from Lemma 2.1.3, (2.36)–(2.37) and the Cauchy–Schwarz inequality it follows

$$\begin{aligned}\mathbb{E}\left(\max_{0\le k\le\Theta_n}|X_{\theta^{(n)}_k}-X^{(n)}_k|\right) &\le \mathbb{E}\left(\max_{0\le k\le\Theta_n}|Y_{\theta^{(n)}_k}-X^{(n)}_k|\right)+\mathbb{E}[2\Gamma_n\mathbb{I}_O]\\ &\le O(\sqrt{h})+O(1)\sqrt{\mathbb{P}(O)}=O(\sqrt{h}).\end{aligned}$$

It remains to estimate $\mathbb{E}\left(\max_{\Theta_n<k\le n}|X_{\theta^{(n)}_k}-X^{(n)}_k|\right)$. Let $\mathbb{Q}\sim\mathbb{P}$ be the probability measure given by

$$Z_t:=\frac{d\mathbb{Q}}{d\mathbb{P}}|\mathcal{F}_t=\exp\left(-\int_0^t\frac{\mu(Y_u)}{\sigma(Y_u)}dW_u-\int_0^t\frac{\mu^2(Y_u)}{2\sigma^2(Y_u)}du\right),\quad t\ge 0.$$

From Girsanov's theorem it follows that under the measure $\mathbb{Q}$, Y is a martingale. Assume that the martingale Y starts at some $Y_0 \in (B, C)$ and we want to estimate $\mathbb{Q}(\max_{0\leq t\leq T} X_t - Y_0 > \epsilon)$ for $\epsilon > 0$. Define a stopping time $\varsigma = T \wedge \inf\{t : Y_t = Y_0 + \epsilon\} \wedge \inf\{t : Y_t = B\}$. Then from the Optional stopping theorem

$$Y_0 = \mathbb{E}_{\mathbb{Q}}[Y_\varsigma] \geq B(1 - \mathbb{Q}(Y_\varsigma = Y_0 + \epsilon)) + (Y_0 + \epsilon)\mathbb{Q}(Y_\varsigma = Y_0 + \epsilon).$$

Hence, (B is an absorbing barrier for X)

$$\mathbb{Q}(\max_{0\leq t\leq T} X_t - Y_0 > \epsilon) \leq \mathbb{Q}(Y_\varsigma = Y_0 + \epsilon) \leq \frac{Y_0 - B}{Y_0 + \epsilon - B}. \tag{2.38}$$

Similarly,

$$\mathbb{Q}(\max_{0\leq t\leq T} Y_0 - X_t > \epsilon) \leq \frac{C - Y_0}{C + \epsilon - Y_0}. \tag{2.39}$$

Next, recall the event O from the beginning of the proof. Consider the event $\tilde{O} = \{\Theta_n < n\} \setminus O$ and choose $\epsilon > 3n^{-1/3}$. Observe that $\theta^{(n)}_n - \theta^{(n)}_{\Theta_n} = (n - \Theta_n)h \leq T$. Moreover on the event $\tilde{O}$, $\min(C - X_{\theta^{(n)}_{\Theta_n}}, X_{\theta^{(n)}_{\Theta_n}} - B) \leq 3n^{-1/3}$ (for sufficiently large n). Thus, from (2.38)–(2.39)

$$\mathbb{Q}\left(\tilde{O} \bigcap \left(\sup_{\theta^{(n)}_{\Theta_n} < t \leq \theta^{(n)}_n} |X_{\theta^{(n)}_{\Theta_n}} - X_t| \geq \epsilon\right) \Bigg| \mathcal{F}_{\theta^{(n)}_{\Theta_n}}\right) \leq \frac{3n^{-1/3}}{3n^{-1/3} + \epsilon} \quad \text{a.s.} \tag{2.40}$$

We conclude that

$$\begin{aligned} &\mathbb{E}_{\mathbb{Q}}\left(\mathbb{I}_{\tilde{O}}\left(1 \wedge \sup_{\theta^{(n)}_{\Theta_n} < t \leq \theta^{(n)}_n} |X_{\theta^{(n)}_{\Theta_n}} - X_t|\right) \Big| |\mathcal{F}_{\theta^{(n)}_{\Theta_n}}\right) \leq \\ &3n^{-1/3} + \int_{3n^{-1/3}}^{1} \tfrac{3n^{-1/3}}{3n^{-1/3} + \epsilon} d\epsilon = O(n^{-1/3} \ln n) \quad \text{a.s.}, \end{aligned} \tag{2.41}$$

where $\mathbb{E}_{\mathbb{Q}}$ denotes the expectation with respect to $\mathbb{Q}$.

Next, denote $\mathcal{Z}_n = \frac{Z_{\theta^{(n)}_n}}{Z_{\theta^{(n)}_{\Theta_n}}}$. Observe that for a random variable $\mathcal{X}$,

$$\mathbb{E}(\mathcal{X} | \mathcal{F}_{\theta^{(n)}_{\Theta_n}}) = \mathbb{E}_{\mathbb{Q}}(\mathcal{Z}_m \mathcal{X} | \mathcal{F}_{\theta^{(n)}_{\Theta_n}}).$$

From the relations $\theta^{(n)}_n - \theta^{(n)}_{\Theta_n} = (n-\Theta_n)h \le T$ it follows that for any $m \in \mathbb{R}$, $\mathbb{E}_Q[\mathcal{Z}^m_n]$ is uniformly bounded (in n). This together with the (conditional) Holder inequality, (2.36)–(2.37), and (2.40)–(2.41) gives

$$\begin{aligned}
\mathbb{E}\left(\sup_{\Theta_n<k\le n}|X_{\theta^{(n)}_k} - X^{(n)}_k|\right) &\le \mathbb{E}[2\Gamma_n \mathbb{I}_O] + \mathbb{E}\left(\mathbb{E}_{\mathbb{Q}}\left(2\mathcal{Z}_n\Gamma_n\mathbb{I}_{\tilde{O}}\mathbb{I}_{\sup_{\theta^{(n)}_{\Theta_n}<t\le\theta^{(n)}_n}|X_{\theta^{(n)}_{\Theta_n}}-X_t|>1}\middle|\mathcal{F}_{\theta^{(n)}_{\Theta_n}}\right)\right)\\
&+ \mathbb{E}\left(\mathbb{E}_{\mathbb{Q}}\left(\mathcal{Z}_n\mathbb{I}_{\tilde{O}}\left(1\wedge\sup_{\theta^{(n)}_{\Theta_n}<t\le\theta^{(n)}_n}|X_{\theta^{(n)}_{\Theta_n}} - X_t|\right)\middle|\mathcal{F}_{\theta^{(n)}_{\Theta_n}}\right)\right)\\
&+ \mathbb{E}\left(\mathbb{I}_{\tilde{O}}|X^{(n)}_{\Theta_n} - X_{\theta^{(n)}_{\Theta_n}}|\right) \le O(\sqrt{h}) + \mathbb{E}\left(O(1)\left(\frac{3n^{-1/3}}{3n^{-1/3}+1}\right)^{4/5}\right)\\
&+ \mathbb{E}\left(O(1)\left(O(n^{-1/3}\ln n)\right)^{4/5}\right) + O(n^{-1/3}) = O(h^{1/4})
\end{aligned}$$

and the result follows. □

2.2 Proof of Theorem 2.1.2

The proof of Theorem 2.1.2 is less technical than the proof of Theorem 2.1.1. We start with the following Lemma.

Lemma 2.2.1. *Recall the stopping time $\rho^{X_0}_A$ from Section 2.1.3. For $X_0 \in (B, C)$ and $0 < \epsilon < \min(X_0 - B, C - X_0)$ we have*

$$\mathbb{E}\left((\rho^{X_0}_\epsilon)^m\right) = O(\epsilon^{2m}) \quad \forall m > 0.$$

Proof. Choose $m > 1$, $X_0 \in (B, C)$ and $0 < \epsilon < \min(X_0 - B, C - X_0)$. Define the stochastic process $\mathcal{M}_t = p(Y_t)$, $t \ge 0$ where recall p is given in (2.15). It is easy to see that p is strictly increasing function and $\{\mathcal{M}_t\}_{t\ge0}$ is a local–martingale which satisfies $\mathcal{M}_t = \int_0^t p'(Y_u)\sigma(Y_u)dW_u$. Hence $\rho^{X_0}_\epsilon = \inf\{t : \mathcal{M}_t = p(X_0 \pm \epsilon)\}$. From the Burkholder–Davis–Gundy inequality it follows that there exists a constant $C > 0$

such that

$$\begin{aligned}
\max\left\{|p(X_0+\epsilon)|^{2m},|p(X_0-\epsilon)|^{2m}\right\} &\geq \mathbb{E}\left(\max_{0\leq t\leq \rho_\epsilon^{X_0}}|\mathcal{M}_t|^{2m}\right)\\
&\geq C\mathbb{E}\left(\left(\int_0^{\rho_\epsilon^{X_0}}|p'(Y_u)\sigma(Y_u)|^2du\right)^m\right)\\
&\geq C\left(\inf_{y\in[X_0-\epsilon,X_0+\epsilon]}|p'(y)\sigma(y)|\right)^{2m}\mathbb{E}\left((\rho_\epsilon^{X_0})^m\right). \qquad (2.42)
\end{aligned}$$

Since $\frac{\mu}{\sigma^2}$ is uniformly bounded, then for $y\in[X_0-\epsilon,X_0+\epsilon]$ we have $p(y)=O(\epsilon)$ and $p'(y)=1-O(\epsilon)$. This together with (2.42) completes the proof. □

Now we are ready to prove Theorem 2.1.2.

Proof of Theorem 2.2. The proof follows the steps of the proof of Theorem 2.1.1, however the current case is much less technical. The reason is that we do not take a perturbation of the process X_t, $t\geq 0$ and so we have an exact scheme at the random times $\{\vartheta_k^{(n)}\}_{k=0}^n$. Thus, we can write similar inequalities to (2.22)–(2.23), only now the analogous term to $\max_{0\leq i\leq n}|X_i^{(n)}-X_{\theta_i^{(n)}}|$ is vanishing. Hence, we can skip Lemmas 2.1.3 and 2.1.5. Namely, in order to complete the proof of Theorem 2.1.2 it remains to show the following analog of Lemma 2.1.4:

$$\mathbb{E}\left(\max_{1\leq k\leq n}|\vartheta_k^{(n)}-kh|^m\right)=O(h^{m/2}),\quad \forall m>0. \qquad (2.43)$$

Without loff of generality, we assume that $m>1$. Set,

$$H_i=\vartheta_i^{(n)}-\vartheta_{i-1}^{(n)}-\mathbb{E}\left(\vartheta_i^{(n)}-\vartheta_{i-1}^{(n)}|\mathcal{F}_{\vartheta_{i-1}^{(n)}}\right),\quad i=1,...,n.$$

Clearly the stochastic process $\{\sum_{i=1}^k H_i\}_{k=1}^n$ is a martingale, and from (2.14) we get that for all k, $\vartheta_k^{(n)}-kh=\sum_{i=1}^k H_i+O(\sqrt{h})$. Moreover, on the event $X_{\vartheta_{i-1}^{(n)}}\in(B,C)$ $\vartheta_i^{(n)}-\vartheta_{i-1}^{(n)}\leq \rho_{\bar\sigma\sqrt{h}}^{X_{\vartheta_{i-1}^{(n)}}}$ and so from Lemma 2.2.1, $\mathbb{E}[H_i^m]=O(h^m)$. This together with the Burkholder–Davis–Gundy inequality and (2.30) yields

$$
\begin{aligned}
\mathbb{E}\left(\max_{1\le k\le n}|\vartheta_k^{(n)}-kh|^m\right) &\le O(h^{m/2})+O(1)\mathbb{E}\left(\max_{1\le k\le n}|\sum_{i=1}^{k}H_i|^m\right)\\
&\le O(h^{m/2})+O(1)\mathbb{E}\left(\left(\sum_{i=1}^{n}H_i^2\right)^{m/2}\right)\\
&\le O(h^{m/2})+O(1)n^{m/2-1}\sum_{i=1}^{n}\mathbb{E}[|H_i|^m]=O(h^{m/2})
\end{aligned}
$$

and the result follows. □

2.2.1 Some remarks on the proofs

Remark 2.2.1. *Theorem 2.1.2 can be easily extended to American barrier options which we study numerically in Sections 5.4–5.5.*

Namely, let $-\infty\le B<x<C\le\infty$ *and let* Y *be the unique (in law) solution of (2.1). Let* $\tau_{B,C}=T\wedge\inf\{t:Y_t\notin(B,C)\}$. *Consider the optimal stopping value*

$$V=\sup_{\tau\in\mathcal{T}_{[0,T]}}\mathbb{E}[\mathbb{I}_{\tau\le\tau_{B,C}}f(\tau,X_\tau)]$$

which corresponds to the price of American barrier options. Let us notice that if we change in the above formula, the indicator $\mathbb{I}_{\tau\le\tau_{B,C}}$ *to* $\mathbb{I}_{\tau<\tau_{B,C}}$ *the value remains the same. As before, we choose* n *sufficiently large and a constant* $\bar\sigma=\bar\sigma(n)>\sup_{y\in(B,C)}|\sigma(y)|+\sqrt{h}\sup_{y\in(B,C)}|\mu(y)|$ *which satisfies* $\frac{C-x}{\bar\sigma\sqrt{h}},\frac{x-B}{\bar\sigma\sqrt{h}}\in\mathbb{N}\cup\{\infty\}$. *Define*

$$\tilde V_n=\max_{\eta\in\mathcal{S}_n}\mathbb{E}[\mathbb{I}_{\eta\le\eta^{(n)}_{B,C}}f(\eta h,X_{\vartheta^{(n)}_\eta})]$$

where $\eta^{(n)}_{B,C}=n\wedge\min\{k:Y_{\vartheta^{(n)}_k}\in\{B,C\}\}$. *In this case, we observe that on the interval* $[0,\vartheta^{(n)}_n]$ *the process* Y *can reach the barriers* B,C *only at the moments* $\vartheta^{(n)}_i$ $i=0,1,...,n$. *Hence, by similar arguments, Theorem 2.1.2, can be extended to the current case as well.*

Remark 2.2.2. *Let us notice that in both of the proofs we get that the random maturity dates* $\theta_n^{(n)}, \vartheta_n^{(n)}$ *are close to the real maturity date* T*, and the error estimates are of order* $O(n^{-1/2})$*. Hence, even for the simplest case where the diffusion process* Y *is the standard Brownian motion we get that the random variables* $|Y_T - Y_{\theta_n^{(n)}}|, |Y_T - Y_{\vartheta_n^{(n)}}|$ *are of order* $O(n^{-1/4})$*. This is the reason that we can not expect better error estimates if the proof is done via Skorohod embedding. In [56] the authors got error estimates of order* $O(n^{-1/2})$ *for optimal stopping approximations via Skorokhod embedding, however they assumed very strong assumptions on the payoff, which excludes even call and put type of payoffs.*

Our numerical results in the next section suggest that the order of convergence we predicted is better than we prove here. In fact, the constant multiplying the power of n *is small, which we think is due to the fact that we have a very efficient way of calculating conditional expectations. We will leave the investigation of this phenomena for future research.*

2.3 Numerical examples

In the first subsection we consider two toy models, which are variations of geometric Brownian motion, the first one with capped coefficients and the other one with absorbing boundaries. In Subsections 2.3.2, we consider the CEV model and in 2.3.3, we consider the CIR model. In Subsection 2.3.4 we consider the European capped barrier of [29] and its American counterpart. We close the paper by considering another modification of geometric Brownian motion with discontinuous coefficients, see Subsection 2.3.5. In these sections we report the values of the option prices, optimal exercise boundaries, the time/accuracy performance of the schemes and numerical convergence rates. All computations are implemented in Matlab R2014a on Intel(R)

Core(TM) i5-5200U CPU @2.20 GHz, 8.00 GB installed RAM PC.

2.3.1 Approximation of one dimensional diffusion with bounded μ and σ by using both two schemes

We consider the following model:

$$dS_t = [(A_1 \wedge S_t) \vee B_1]dt + [(A_2 \wedge S_t) \vee B_2]dW_t, \quad S_0 = x, \tag{2.44}$$

where A_1, B_1, A_2, B_2 are constants, and the American put option with strike price $\$K$ and expiry date T:

$$v(x) = \sup_{\tau \in \mathcal{T}_{[0,T]}} \mathbb{E}\left[e^{-r\tau}(K - S_\tau)^+\right]. \tag{2.45}$$

Figure 2.1 shows the value function v and the optimal exercise boundary curve which we obtained using both schemes. Table 2.1 reports the schemes take and Figure 2.2 shows the rate of convergence of both models.

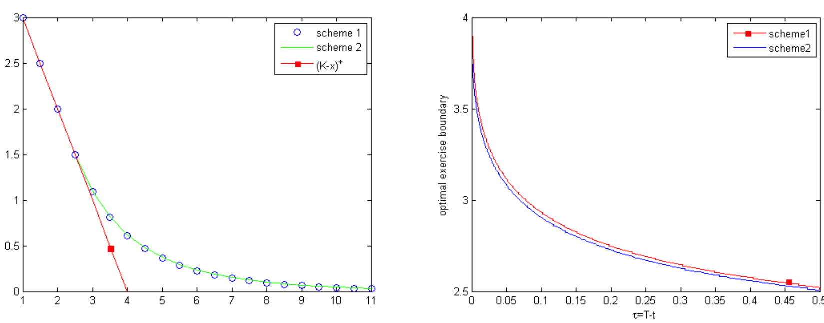

Figure 2.1: The left figure shows that value function of American put (2.45) under model (2.44) with parameters: $n = 8000, r = 0.1, K = 4, A_1 = 10, A_2 = 10, B_1 = 2, B_2 = 2, T = 0.5$. The right figure is the optimal exercise boundary curve. Here τ is the time to maturity.

Table 2.1: Put option prices(2.45) under model (2.44):

	scheme 1	scheme 2		scheme 1	scheme 2
n=1000	0.5954	0.6213	n=2000	0.6042	0.6213
error(%)	3.64	0.23	error(%)	2.21	0.08
CPU	0.540s	0.374s	CPU	2.097s	1.467s
	scheme 1	scheme 2		scheme 1	scheme 2
n=3000	0.6079	0.6214	n=4000	0.6100	0.6216
error(%)	1.61	0.06	error(%)	1.27	0.03
CPU	4.749s	3.249s	CPU	8.726s	5.747s
	scheme 1	scheme 2		scheme 1	scheme 2
n=5000	0.6114	0.6216	n=6000	0.6124	0.6216
error(%)	1.04	0.03	error(%)	0.88	0.02
CPU	14.301s	9.145s	CPU	24.805s	14.080s

Parameters used in computation are: $r = 0.1, K = 4, A_1 = 10, A_2 = 10, B_1 = 2, B_2 = 2, T = 0.5, x = 4$. Error is computed by taking absolute difference between $v_n(x)$ and $v_{30000}(x)$ and then dividing by $v_{30000}(x)$.

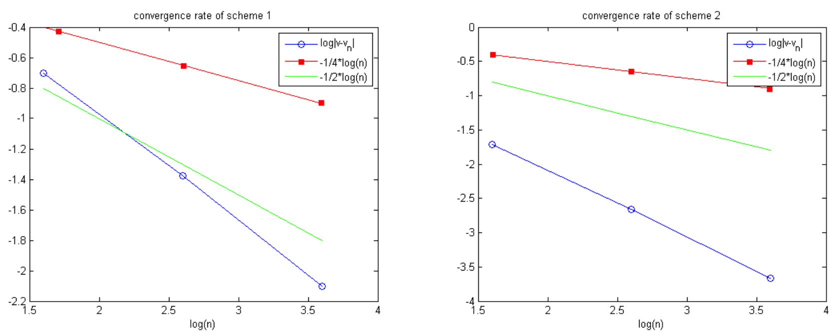

Figure 2.2: Convergence rate figure of both two schemes in the implement of American put (2.45) under (2.44). We draw $\log|v - v_n|$ vs $\log n$ for both schemes to show the convergence speed. Here we pick $v(x) = v_{30000}(x)$ for $x = 4$ and $n \in \{40, 400, 4000\}$. The slope of the left line given by linear regression is -0.69981, and the right is -0.97422. Parameters used in computation are: $r = 0.1, K = 4, A_1 = 10, A_2 = 10, B_1 = 2, B_2 = 2, T = 0.5, x = 4$.

For comparison we will also consider a geometric Brownian motion with absorbing barriers. Let

$$dY_t = Y_t dt + Y_t dW_t, \quad Y_0 = x, \tag{2.46}$$

and $B, C \in [-\infty, \infty)$ with $B < x < C$ and denote

$$X_t = \mathbb{I}_{t<\inf\{s: Y_s \notin (B,C)\}} Y_t + \mathbb{I}_{t \geq \inf\{s: Y_s \notin (B,C)\}} Y_{\inf\{s: Y_s \notin (B,C)\}}, \quad X_0 = x. \tag{2.47}$$

As in the previous example, we show below the value function, times the schemes take and the numerical convergence rate.

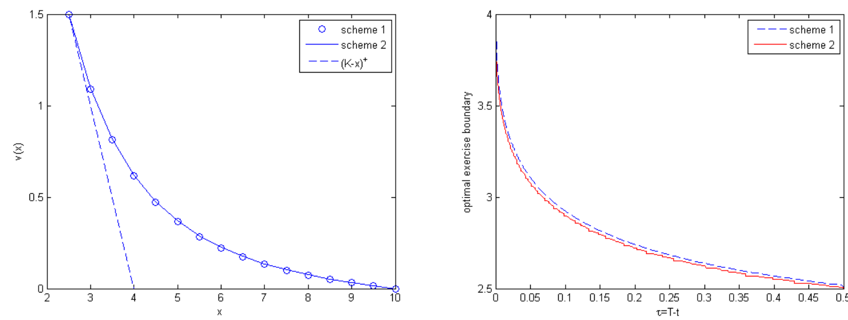

Figure 2.3: The left figure shows the value function of American put (2.45) under model (2.47) with parameters: $n = 30000, r = 0.1, K = 4, C = 10, B = 2, T = 0.5$. The right figure is the optimal exercise boundary curve.

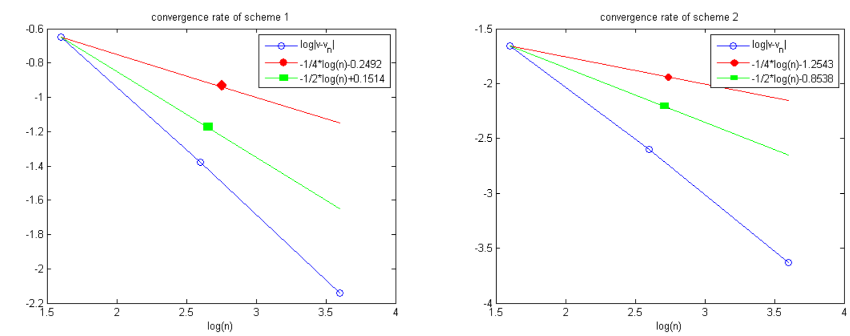

Figure 2.4: Convergence rate figure of both two schemes in the implement of American put (2.45) under (2.47). We draw $\log|v - v_n|$ vs $\log n$ for both schemes to show the convergence speed. Here we pick, $v(x) = v_{30000}(x)$ for $x = 4$ and $n \in \{40, 400, 4000\}$. The slope of the left line given by linear regression is -0.74513, the right one is -0.98927.

Table 2.2: put option prices(2.45) under (2.47)

	scheme 1	scheme 2		scheme 1	scheme 2
n=1000	0.5943	0.6175	n=2000	0.6029	0.6185
error(%)	3.44	0.24	error	2.05	0.09
CPU	0.240s	0.029s	CPU	0.667s	0.091s
	scheme 1	scheme 2		scheme 1	scheme 2
n=3000	0.6063	0.6184	n=4000	0.6082	0.6188
error(%)	1.49	0.09	error	1.17	0.03
CPU	1.220s	0.157s	CPU	1.927s	0.254s
	scheme 1	scheme 2		scheme 1	scheme 2
n=5000	0.6095	0.6188	n=6000	0.6104	0.6189
error(%)	0.96	0.04	error	0.81	0.02
CPU	2.720s	0.359s	CPU	3.723s	0.457s

Note: Parameters used in computation: $r = 0.1, K = 4, C = 10, B = 2, T = 0.5, x = 4$. Error is computed by taking absolute difference between $v_n(x)$ and $v_{30000}(x)$ and then dividing by $v_{30000}(x)$.

2.3.2 Pricing American put option under CEV

Set $\beta \leq 0$ and consider the CEV model:

$$dS_t = rS_t + \delta S_t^{\beta+1} dW_t, \quad S_0 = x, \tag{2.48}$$

and let $\delta = \sigma_0 x^{-\beta}$ (to be consistent with the notation of [29]). Consider the American put option pricing problem:

$$v_A(x) = \sup_{\tau \in \mathcal{T}_{[0,T]}} \mathbb{E}\left[e^{-r\tau}(K - S_\tau)^+\right]. \tag{2.49}$$

Clearly, the CEV model given by (2.48) does not satisfies our assumptions. However, this limitation can be solved by truncating the model. The next result shows that if we choose absorbing barriers $B, C > 0$ such that B is small enough and C is large enough, then the optimal stopping problem does not change much. Hence, we can apply our numerical scheme for the absorbed diffusion and still get error estimates of order $O(n^{-1/4})$.

Lemma 2.3.1. *Choose $0 < B < C$. Consider the CEV model given by (2.48) and let X be the absorbed process*

$$X_t = \mathbb{I}_{t < \inf\{s: S_s \notin (B,C)\}} S_t + \mathbb{I}_{t \geq \inf\{s: S_s \notin (B,C)\}} S_{\inf\{s: S_s \notin (B,C)\}}, \quad t \geq 0.$$

Then

(2.50)

$$\sup_{\tau\in\mathcal{T}_{[0,T]}} \mathbb{E}[e^{-r\tau}(K-S_\tau)^+] - \sup_{\tau\in\mathcal{T}_{[0,T]}} \mathbb{E}[e^{-r\tau}(K-X_\tau)^+] = \mathbb{I}_{\tau^*\leq\tau_B}B + O(C^{-k}), \quad \forall k\geq 1,$$

where τ^* is the optimal time for first the expression on the left and τ_B is the first time S_t is below the level B.[1]

Proof. Let us briefly argue that $\mathbb{E}\left(\max_{0\leq t\leq T} S_t^k\right) < \infty$ for all k. Clearly, if $\beta = 0$ then S is a geometric Brownian motion and so the statement is clear. For $\beta < 0$ we observe that the process $Y = S^{-\beta}$ belongs to CIR family of diffusions (see equation (1) in [40]). Such process can be represented in terms of BESQ (squared Bessel processes); see equation (4) in [40]. Finally, the moments of the running maximum of the (absolute value) Bessel process can be estimated by Theorem 4.1 in [42]. Hence from the Markov inequality it follows that $\mathbb{P}(\max_{0\leq t\leq T} S_t \geq M) = O(M^{-k})$. Next, consider the stopping time $\tau_B = \inf\{t : S_t = B\} \wedge T$. If $\tau_B < T$ then the payoff of the put option is $K - B$, and since S is non–negative its B optimal to stop at this time. It is clear also that one would make no error if B is always in the stopping region. □

[1]It might seem to be a moot point to have $\mathbb{I}_{\tau^*\leq\tau_B}$ since this involves determining the optimal stopping boundary. But one can actually easily determine B's satisfying this condition through through setting $B \leq b$, the perpetual version of the problem whose stopping boundary, which can be determined by finding the unique root of

(2.51)
$$F(x) = \phi'(x)\frac{K-x}{\phi(x)} + 1,$$

in which the function ϕ is given by [29, Equation (38)] (with $\lambda = r$), see e.g. [8]. The perpetual Black-Scholes boundary,

(2.52)
$$b = 2rK/(2r+\delta^2)$$

is in fact is a lower bound to b for $\beta < 0$, which can be proved using the comparison arguments in [7].

Table 2.3: Pricing American puts (2.49) under the CEV (2.48) model with different β's by using the first scheme

K	$\beta = -1$	$\beta = -\frac{1}{3}$
90	1.4373(1.5122)	1.3040(1.3844)
100	4.5359(4.6390)	4.5244(4.6491)
110	10.6502(10.7515)	10.7716(10.8942)

Parameters used in computation are: $\sigma_0 = 0.2, x = 100, T = 0.5, r = 0.05, n = 15000, B = 0.01, C = 200$. The values in the parentheses are the results from last row of Tables 4.1, 4.2, 4.3 from paper [75], which carries out the artificial boundary method discussed in the introduction. This is a finite difference method which relies on artificially introducing an exact boundary condition.

Table 2.4: Prices of the American puts (2.49) under (2.48) with different β's by using the second scheme

K	$\beta = -1$	$\beta = -\frac{1}{3}$
90	1.5123(1.5122)	1.3845(1.3844)
100	4.6392(4.6390)	4.6492(4.6489)
110	10.7517(10.7515)	10.8943(10.8942)

Parameters used in computation are: $\sigma_0 = 0.2, x = 100, T = 0.5, r = 0.05, n = 15000, B = 0.01, C = 200$. The values in the parentheses are the results from last row of Tables 4.1, 4.2, 4.3 from paper [75], which carries out the artificial boundary method discussed in the introduction. This is a finite difference method which relies on artificially introducing an exact boundary condition.

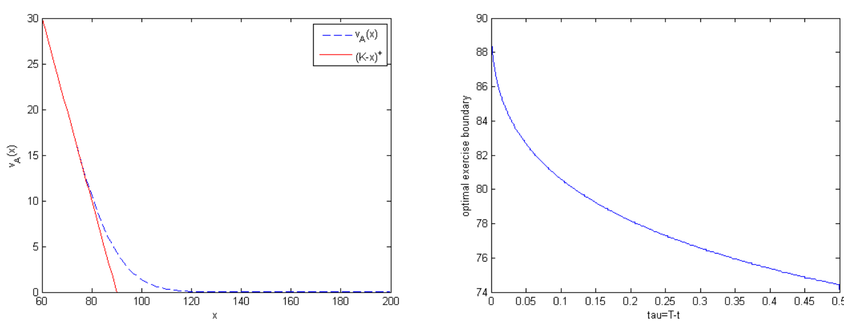

Figure 2.5: The left figure shows that value function of American put (2.49) under CEV (2.48) with parameters: $n = 2000, \beta = -\frac{1}{3}, \delta = \sigma_0(100)^{-\beta}, \sigma_0 = 0.2, T = 0.5, r = 0.05, B = 0.01, C = 200, K = 90$. The right figure is the optimal exercise boundary curve which $\tau = T - t$ denotes time to maturity. With these parameters $F(60) = -0.3243, F(70) = 0.1790$ (where F is defined in (2.51)), from which we conclude that the unique root is in (60,70). So there in fact is no error in introducing a lower barrier. In this case $b = 9.3577$ (where b is defined in (2.52)).

Since we observed that second scheme exhibits faster convergence, we report the CPU time and errors for second scheme.

Table 2.5: Prices of the American puts under (2.48) with different *beta*'s (second scheme)

	$\beta = -1$	$\beta = -\frac{1}{3}$
K=90	1.5147	1.3862
CPU	0.335s	0.259s
error(%)	0.16	0.12
K=100	4.6414	4.6383
CPU	0.314s	0.256s
error(%)	0.05	0.23
K=110	10.7523	10.8925
CPU	0.338s	0.253s
error(%)	0.005	0.02

Parameters used in computation are: $\sigma_0 = 0.2, x = 100, T = 0.5, r = 0.05, B = 0.01, C = 200$. The Error is computed by taking absolute difference between $v_n(x)$ and $v_{15000}(x)$ and then dividing by $v_{15000}(x)$.

Table 2.6: Pricing American puts (2.49) under (2.48) with different σ_0 (second scheme)

	$\sigma_0 = 0.2$	$\sigma_0 = 0.3$	$\sigma_0 = 0.4$
K=35	1.8606(1.8595)	4.0424(4.0404)	6.4017(6.3973)
CPU	0.247s	0.171s	0.136s
error(%)	0.04	0.03	0.06
K=40	3.3965(3.3965)	5.7920(5.7915)	8.2584(8.2574)
CPU	0.237s	0.172s	0.154s
error(%)	0.01	0.002	0.004
K=45	5.9205(5.9204)	8.1145(8.1129)	10.5206(10.5167)
CPU	0.203s	0.170s	0.146s
error(%)	0.01	0.02	0.03

Parameters used in computation are: $T = 3, r = 0.05, x = 40, \beta = -1, n = 100, B = 0.01, C = 100$. Error is computed by taking absolute difference between $v_n(x)$ and $v_{30000}(x)$ and then dividing by $v_{30000}(x)$. The values in the parentheses are the results computed by FDM:1024×1024 reported in Table 1 of [76].

In Table 2.3.2 we compare our numerical results with the ones reported in Table 1 of [76], the numerical values here are closer to results computed by FDM:1024×1024, which is the traditional Crank-Nicolson finite difference scheme (the 1024×1024 refer to the size of the lattice). The CPU time of FDM:1024×1024 is 6.8684s. Table 1 of [76] also reports the performance of the Laplace-Carson transform and the artificial boundary method of [75] (both of which we briefly described in the introduction). The CPU for Laplace-Carson method is 0.609s, ABC:512×512 is 0.9225s. If we consider FDM:1024 $\times$ 1024 as a Benchmark, our method is more than **30 times** faster with accuracy up to 4 decimal points. Comparing with ABC:512$\times$512 method, our method is about **5 times** faster with the same accuracy. Our method is about **3 times** faster than what the Laplace transform method produced when it had accuracy up to 3 decimal points. (Here accuracy represents the absolute difference between the computed numerical values and the FDM:1024 $\times$ 1024 divided by the FDM:1024 $\times$ 1024).

2.3.3 Pricing American put option under CIR

In this subsection, we use numerical scheme to evaluate American put options $v_A(x)$ (recall formula (2.49)) under CIR model. (This model is used for pricing VIX options, see e.g. [53]). Let the volatility follow

$$dY_t = (\beta - \alpha Y_t)dt + \sigma\sqrt{Y_t}dW_t, \quad Y_0 = y, \tag{2.53}$$

Consider the absorbed process

$$X_t = \mathbb{I}_{t<\inf\{s:Y_s\notin(B,C)\}}Y_t + \mathbb{I}_{t\geq\inf\{s:Y_s\notin(B,C)\}}Y_{\inf\{s:Y_s\notin(B,C)\}}, \quad t \geq 0,$$

for $0 < B < C$. Same arguments as in Section 2.3.2 give that for CIR process $\mathbb{P}(\max_{0\leq t\leq T} Y_t \geq C) = O(C^{-k})$, $k \geq 1$. Moreover, from Theorem 2 in [40] it follows that $\mathbb{P}(\min_{0\leq t\leq T} Y_t \leq B) = O(B^{2\nu})$ for $\nu = \frac{2\beta}{\sigma^2} - 1 > 0$. Thus if we consider the absorbed diffusion X then the change of the value of the option price is bounded by $O(C^{-k}) + O(B^{2\nu})$ for all k. (In fact, potentially we do not make any error by having an absorbing boundary if B is small enough, as we argued in the previous section.)

Table 2.7: Prices of American puts (2.49) under (2.53) with different K

	scheme 1	scheme 2
$K = 35$	4.4654	4.5223
$K = 40$	8.1285	8.1932
$K = 45$	12.4498	12.5167

Parameters used in computation are: $\sigma = 2, y = 40, T = 0.5, r = 0.1, B = 0.01, C = 200, \beta = 2, \alpha = 0.5, n = 30000$.

As the previous example, since the second scheme looks more efficient, we only report its time with the accuracy performance here.

Table 2.8: American put prices under (2.53) for a variety of K's using the second scheme

	$K = 35$	$K = 40$	$K = 45$
$n = 100$	4.5140	8.1834	12.5127
CPU	0.245s	0.202s	0.198s
error(%)	0.18	0.12	0.03
$n = 500$	4.5215	8.1918	12.5163
CPU	0.600s	0.585s	0.579s
error(%)	0.02	0.02	0.003
$n = 1000$	4.5238	8.1925	12.5170
CPU	0.807s	0.813s	0.839s
error(%)	0.03	0.01	0.002

Parameters used in computation are: $\sigma = 2, y = 40, T = 0.5, r = 0.1, B = 0.01, C = 200, \beta = 2, \alpha = 0.5$. Error is computed by taking absolute difference between $v_n(x)$ and $v_{30000}(x)$ and then dividing by $v_{30000}(x)$.

2.3.4 Pricing double capped barrier options under CEV by using second scheme

In this subsection, we use second scheme to evaluate double capped barrier call options under the CEV model given by (2.48). Let us denote the European option value by

$$v_E(x) = e^{-rT}\mathbb{E}\left[\mathbb{I}_{T<\tau^*}(S_T - K)^+\right], \tag{2.54}$$

and the American option value by

$$v_A(x) = \sup_{\tau \in \mathcal{T}_{[0,T]}} \mathbb{E}\left[\mathbb{I}_{\tau<\tau^*}e^{-r\tau}(S_\tau - K)^+\right], \tag{2.55}$$

where $\tau^* = \inf\{t \geq 0; S_t \neq (L, U)\}$. We use the parameters values given in [29] and report our results in Table 2.3.4 and Figure 2.3.4.

Table 2.9: Double capped European Barrier Options (2.54) under CEV(2.48) (second scheme)

U	L	K	$\beta = -0.5$	$\beta = 0$	$\beta = -2$	$\beta = -3$
120	90	95	1.9012	1.7197	2.5970	3.2717
	CPU		0.970s	1.026s	6.203s	1.209s
	error		0.96%	1.2%	0.4%	2.7%
120	90	100	1.1090	0.9802	1.6101	2.0958
	CPU		0.957s	1.102s	6.066s	1.205s
	error		0.98%	1.2%	0.44%	2.3%
120	90	105	0.5201	0.4473	0.8142	1.1072
	CPU		0.961s	1.023s	6.205s	1.187s
	error		1.00%	1.3%	0.53%	2.2%
	n		2000	2000	5000	2000

Table 2.10: Double capped American Barrier Options (2.55) under CEV(2.48) (second scheme)

U	L	K	$\beta = -0.5$	$\beta = 0$	$\beta = -2$	$\beta = -3$
120	90	95	9.8470	9.8271	9.9826	10.0586
	CPU		0.965s	1.007s	1.248s	1.170s
	error		0.55%	0.63%	0.56%	0.81%
120	90	100	7.4546	7.4522	7.5118	7.5203
	CPU		0.942s	1.017s	1.239s	1.211s
	error		0.70%	0.8%	0.47%	0.57%
120	90	105	5.2612	5.2788	5.2345	5.1669
	CPU		0.968s	1.044s	1.145s	1.188s
	error		1.00%	1.2%	0.35%	0.17%
	n		2000	2000	2000	2000

Parameters used in calculations are: $x = 100, \sigma(100) = 0.25$ (denoting $\sigma(S) = \delta S^{1+\beta}$), $r = 0.1, T = 0.5$. Error is computed by taking the absolute difference between v_n and v_{40000} and then dividing by v_{40000}.

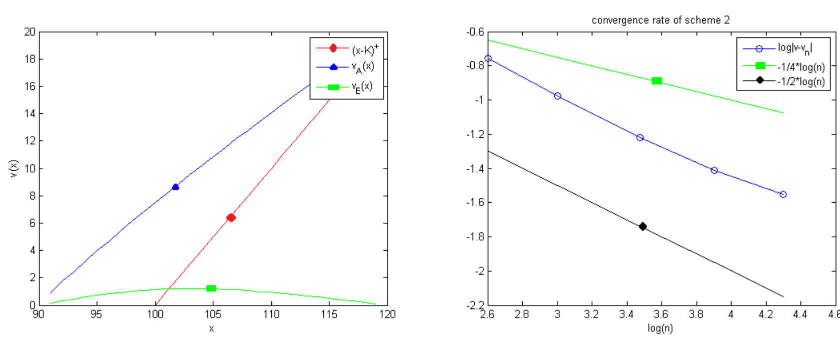

Figure 2.6: The left figure shows the value function $v_A(x)$(2.55) and $v_E(x)$(2.54) with Parameters: $\delta = 2.5, r = 0.1, T = 0.5, \beta = -0.5, n = 5000, L = 90, U = 120, K = 100,$. In the right figure, we show the $\log|v - v_n|$ vs $\log n$ picking $v(x) = v_{40000}(x)$, $x = 100$, and $n \in \{400, 1000, 3000, 8000, 20000\}$. The slope of the blue line given by linear regression is -0.47178.

2.3.5 Pricing double capped American barrier options with jump volatility by using second scheme

In this subsection, we show the numerical results for pricing double capped American barrier option when the volatility has a jump. Consider the following modification of a geometric Brownian motion:

$$dS_t = rS_t dt + \sigma(S_t)S_t dW_t, S_0 = x \tag{2.56}$$

where $\sigma(S_t) = \sigma_1$, for $S_t \in [0, S_1]$, $\sigma(S_t) = \sigma_2$, for $S_t \in [S_1, \infty)$. We will compute the American option price

$$v(x) = \sup_{\tau \in \mathcal{T}_{[0,T]}} \mathbb{E}\left[e^{-r\tau}\mathbb{I}_{\tau<\tau^*}(S_\tau - K)^+\right], \tag{2.57}$$

where $\tau^* = \inf\{t \geq 0; S_t \neq (L, U)\}$. We use scheme 2 here, since it can handle discontinuous coefficients; see Figures 2.3.5 and 2.3.5. We compare it to the geometric Brownian motion (gBm) problem with different volatility parameters. In particular, we observe that although the jump model's price and the gBm model with average volatility are close in price, the optimal exercise boundaries differ significantly.

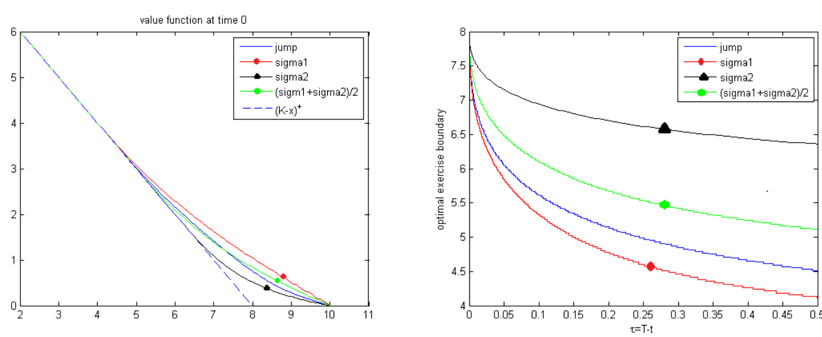

Figure 2.7: There are four curves in the left figure. They are value functions (2.57) under model (2.56) with different values of $\sigma(S_t)$ as specified in the legend. The parameters used in computations are: $n = 4000, r = 0.1, S_1 = K = 8, L = 2, U = 10, \sigma_1 = 0.7, \sigma_2 = 0.3, T = 0.5$. Here $\sigma_2 < \frac{\sigma_1+\sigma_2}{2} < \sigma_1$, and as expected the gBm option price decreases as σ decreases but the gBm price with jump volatility and one corresponding to gBm with $\sigma = \sigma_2$ intersect in the interval $[7, 8]$. The right figure is of the optimal exercise boundary curves.

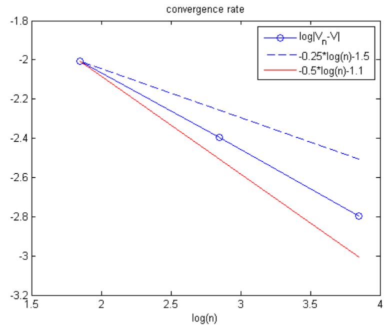

Figure 2.8: Convergence rate figure of scheme 2 in the implement of double capped American put (2.57) under (2.56). We pick $v(x) = v_{40000}(x)$ to be the actual price, and the other three values are $v_{70}(x), v_{700}(x), v_{7000}(x), x = 8$. The slope of the blue line given by linear regression approach is -0.39499. Parameters used in computation: $n = 4000, r = 0.1, K = 8, L = 2, U = 10, \sigma_1 = 0.7, \sigma_2 = 0.3, S1 = 8, T = 0.5$.

CHAPTER III

Continuity of Utility Maximization under Weak Convergence

The main work in this Chapter is based on paper [11]. I also refer to paper [12]. It studied extended weak convergence and utility maximization with proportional transaction costs. The outline of the Chapter is the following. We introduce the setup and formulate the main results. In Section 3.2 we discuss Assumptions 3.1.4,3.1.5,3.1.6 and demonstrate their necessity. In Section 3.3 we prove the lower semi–continuity. In Section 3.4 we prove the upper semi–continuity. In Section 3.4.1 we establish Theorem 3.1.2. Section 3.5 is devoted to the construction of an approximating sequence for the Heston model. In Section 3.6 we provide a detailed numerical analysis for shortfall risk minimization.

3.1 Preliminaries and Main Results

We consider a model of a security market which consists of d risky assets which we denote by $S = (S_t^{(1)}, ..., S_t^{(d)})_{0\le t\le T}$, where $T < \infty$ is the time horizon. We assume that the investor has a bank account that, for simplicity, bears no interest. The process S is assumed to be a continuous semi–martingale on a filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t^S)_{0\le t\le T}, \mathbb{P})$ where the filtration $(\mathcal{F}_t^S)_{0\le t\le T}$ is the usual filtration generated by S. Namely, the filtration $\{\mathcal{F}_t^S\}_{t=0}^T$ is the minimal filtration which is complete, right continuous and satisfies $\mathcal{F}_t \supset \sigma\{S_u : u \leq t\}$. Without a loss of generality we take $\mathcal{F} := \mathcal{F}_T^S$.

A (self–financing) portfolio π is defined as a pair $\pi = (x, \gamma)$ where the constant x is the initial value of the portfolio and $\gamma = (\gamma^{(i)})_{1 \leq i \leq d}$ is a predictable S–integrable process specifying the amount of each asset held in the portfolio. The corresponding portfolio value process is given by

$$V_t^\pi := x + \int_0^t \gamma_u dS_u, \quad t \in [0, T].$$

Observe that the continuity of S implies that the wealth process $\{V_t^\pi\}_{t=0}^T$ is continuous as well. We say that a trading strategy π is admissible if $V_t^\pi \geq 0$, $\forall t \geq 0$. For any $x > 0$ we denote by $\mathcal{A}(x)$ the set of all admissible trading strategies.

Denote by $\mathcal{M}(S)$ the set of all equivalent (to $\mathbb{P}$) local martingale measures. We assume that $\mathcal{M}(S) \neq \emptyset$. This condition is intimately related to the absence of arbitrage opportunities on the security market. See [31] for a precise statement and references.

Next, we introduce our utility maximization problem. Consider a continuous function $U : (0, \infty) \times \mathbb{D}([0, T]; \mathbb{R}^d) \to \mathbb{R}$. As usual, $\mathbb{D}([0, T]; \mathbb{R}^d)$ denotes the space of all RCLL (right continuous with left limits) functions $f : [0, T] \to \mathbb{R}^d$ equipped with the Skorokhod topology (for details see [18]).

Assumption 3.1.1.

(i) For any $s \in \mathbb{D}([0, T]; \mathbb{R}^d)$ the function $U(\cdot, s)$ is non–decreasing.

(ii) For any $x > 0$ we have $\mathbb{E}_{\mathbb{P}}[U(x, S)] > -\infty$.

We extend U to $\mathbb{R}_+ \times \mathbb{D}([0, T]; \mathbb{R}^d)$ by $U(0, s) := \lim_{v \downarrow 0} U(v, s)$. In view of Assumption 3.1.1(i) the limit exists (might be $-\infty$).

For a given initial capital $x > 0$ consider the optimization problem

$$u(x) := \sup_{\pi \in \mathcal{A}(x)} \mathbb{E}_{\mathbb{P}}[U(V_T^{\pi}, S)],$$

where we set $-\infty + \infty = -\infty$. Namely, for a random variable X which satisfies $\mathbb{E}_{\mathbb{P}}[\max(-X, 0)] = \infty$ we set $\mathbb{E}_{\mathbb{P}}[X] := -\infty$.

Let us notice that Assumption 3.1.1(ii) implies $u(x) > -\infty$.

Assumption 3.1.2. *The function $u : (0, \infty) \to \mathbb{R} \cup \{\infty\}$ is continuous. Namely, for any $x > 0$ we have $u(x) = \lim_{y \to x} u(y)$ where a priori the joint value can be equal to ∞.*

Next, for any n, let $S^{(n)} = (S_t^{n,1}, ..., S_t^{n,d})_{0 \le t \le T}$ be a RCLL semi–martingale defined on some filtered probability space $(\Omega_n, \mathcal{F}^{(n)}, (\mathcal{F}_t^{(n)})_{0 \le t \le T}, \mathbb{P}_n)$ where the filtration $(\mathcal{F}_t^{(n)})_{0 \le t \le T}$ satisfies the usual assumptions (right continuity and completeness). For the n–th model we define $\mathcal{A}_n(x)$ as the set of all pairs $\pi_n = (x, \gamma^{(n)})$ such that $\gamma^{(n)}$ is a predictable $S^{(n)}$–integrable process and the resulting portfolio value process

$$V_t^{\pi_n} := x + \int_0^t \gamma_u^{(n)} dS_u^{(n)} \ge 0, \quad t \in [0, T],$$

is non-negative. Set,

$$u_n(x) := \sup_{\pi_n \in \mathcal{A}_n(x)} \mathbb{E}_{\mathbb{P}_n}[U(V_T^{\pi_n}, S^{(n)})].$$

We assume the weak convergence $S^{(n)} \Rightarrow S$ on the space $\mathbb{D}([0, T]; \mathbb{R}^d)$ equipped with the Skorokhod topology. Moreover, we assume the following uniform integrability assumptions.

Assumption 3.1.3.

(i) For any $x > 0$ the family of random variables $\{U^-(x, S^{(n)})\}_{n \in \mathbb{N}}$ is uniformly integrable where $U^- := \max(-U, 0)$.

(ii) For any $x > 0$ *the family of random variables* $\{U^+(V_T^{\pi_n}, S^{(n)})\}_{n\in\mathbb{N},\pi_n\in\mathcal{A}_n(x)}$ *is uniformly integrable, where* $U^+ := \max(U, 0)$.

Remark 3.1.1. *The verification of Assumption 3.1.2 and Assumption 3.1.3(ii) can be a difficult task. In Section 3.1.1 we provide quite general and easily verifiable conditions which are sufficient for the above assumptions to hold true.*

Due to the admissibility requirements we will need the following assumption which bounds the uncertainty of the jump activity. This assumption will be discussed in details in Section 3.2.1.

Assumption 3.1.4. *For any* $n \in \mathbb{N}$ *consider the non-decreasing RCLL process given by* $D_t^{(n)} := \sup_{0\le u\le t} |S_u^{(n)} - S_{u-}^{(n)}|$, $t \in [0, T]$ *where* $|\cdot|$ *denotes the Euclidean norm in* $\mathbb{R}^d$. *For any* n, *there exists an adapted (to* $(\mathcal{F}_t^{(n)})_{0\le t\le T}$*) left continuous process* $\{J_t^{(n)}\}_{t=0}^T$, $n \in \mathbb{N}$ *such that* $\inf_{0\le t\le T}\left(J_t^{(n)} - D_t^{(n)}\right) \ge 0$ *a.s. and* $J_T^{(n)} \to 0$ *in probability.*

Now, we ready to formulate our first result (lower semi–continuity) which will be proved in Section 3.3.

Proposition III.1. *Under Assumptions 3.1.1–3.1.2, Assumption 3.1.3(i) and Assumption 3.1.4 we have*

$$u(x) \le \liminf_{n\to\infty} u_n(x), \quad \forall x > 0.$$

Next, we treat upper semi–continuity.

Assumption 3.1.5. *Recall the set* $\mathcal{M}(S)$ *of all equivalent local martingale measures. Denote by* $\mathcal{M}(S^{(n)})$, $n \in \mathbb{N}$ *the set of all equivalent local martingale measures for the* n*–th model. For any* $\mathbb{Q} \in \mathcal{M}(S)$ *there exists a sequence of probability measures* $\mathbb{Q}_n \in \mathcal{M}(S^{(n)})$, $n \in \mathbb{N}$ *such that under* $\mathbb{P}_n$ *the joint distribution of* $\left(\{S_t^{(n)}\}_{t=0}^T, \frac{d\mathbb{Q}_n}{d\mathbb{P}_n}\right)$

on the space $\mathbb{D}([0,T];\mathbb{R}^d)\times\mathbb{R}$ *converges to the joint distribution of* $\left(\{S_t\}_{t=0}^T, \frac{d\mathbb{Q}}{d\mathbb{P}}\right)$ *under* $\mathbb{P}$*. We denote this relation by*

$$\left(\left(S^{(n)}, \frac{d\mathbb{Q}_n}{d\mathbb{P}_n}\right); \mathbb{P}_n\right) \Rightarrow \left(\left(S, \frac{d\mathbb{Q}}{d\mathbb{P}}\right); \mathbb{P}\right). \tag{3.1}$$

Remark 3.1.2. *The verification of Assumption 3.1.5 requires a convenient representation of the corresponding local martingale measures. This is the case for tree based approximations of diffusion processes. In Section 3.5.2 we illustrate the verification of Assumption 3.1.5 for tree based approximations of the Heston model.*

We do notice that in order to verify Assumption 3.1.5 it is sufficient to establish (3.1) for a dense subset of $\left\{\frac{d\mathbb{Q}}{d\mathbb{P}} : \mathbb{Q}\in\mathcal{M}(S)\right\}$*. This simplification will be used in Section 3.5.2.*

Assumption 3.1.6. *For any* $s\in\mathbb{D}([0,T];\mathbb{R}^d)$*, the function* $U(\cdot, s)$ *is concave.*

Assumption 3.1.6 says that the investor can not gain from additional randomization.

The following upper semi–continuity result will be proved in Section 3.4.

Proposition III.2. *Under Assumption 3.1.1(i), Assumption 3.1.3(ii) and Assumptions 3.1.5,3.1.6 we have*

$$u(x) \geq \limsup_{n\to\infty} u_n(x), \quad \forall x>0.$$

We now combine the statements of the above propositions and state them as the main theorem of our paper:

Theorem 3.1.1. *Under Assumptions 3.1.1–3.1.3,3.1.4,3.1.5,3.1.6 we have*

$$u(x) = \lim_{n\to\infty} u_n(x), \quad \forall x>0. \tag{3.2}$$

Proof. Follows from Proposition III.1 and Proposition III.2. □

Remark 3.1.3. *Observe that in view of Assumption 3.1.3 we have*

$$-\infty < \liminf_{n\to\infty} u_n(x) \le \limsup_{n\to\infty} u_n(x) < \infty, \quad \forall x > 0.$$

We conclude that the joint value in (3.2) is finite.

Next, we establish the weak convergence for the optimal terminal wealths.

Theorem 3.1.2. *Assume that Assumptions 3.1.1–3.1.3,3.1.4,3.1.5,3.1.6 hold true. Moreover, assume that for any $s \in \mathbb{D}([0,T];\mathbb{R}^d)$ the function $U(\cdot, s)$ is strictly concave. Let $x > 0$ and $\hat{\pi}_n \in \mathcal{A}_n(x)$, $n \in \mathbb{N}$ be a sequence of asymptotically optimal portfolios, namely*

$$\lim_{n\to\infty} \left(u_n(x) - \mathbb{E}_{\mathbb{P}_n}[U(V_T^{\hat{\pi}_n}, S^{(n)})]\right) = 0. \tag{3.3}$$

Then

$$\left(S^{(n)}, V_T^{\hat{\pi}_n}\right) \Rightarrow \left(S, V_T^{\hat{\pi}}\right),$$

where $\hat{\pi} \in \mathcal{A}(x)$ is the unique portfolio that satisfies $u(x) = \mathbb{E}_{\mathbb{P}}[U(V_T^{\hat{\pi}}, S)]$.

The proof of Theorem 3.1.2 will be given in Section 3.4.1.

Remark 3.1.4. *It is well known (see Theorem 2.2 in [54]) that for a utility function which is strictly concave there exists a unique optimizer. Although in [54] the authors do not consider a random utility, their argument can be without much effort extended to our setup.*

3.1.1 On the verification of Assumption 3.1.2 and Assumption 3.1.3(ii)

The following result provides a simple and quite general condition which implies Assumption 3.1.2.

Lemma 3.1.1. *Assume that Assumption 3.1.1 holds true and there exist continuous functions* $m_1, m_2 : [0,1) \to \mathbb{R}_+$ *with* $m_1(0) = m_2(0) = 0$ *(modulus of continuity) and a non-negative random variable* $\zeta \in L^1(\Omega, \mathcal{F}, \mathbb{P})$ *such that for any* $\lambda \in (0,1)$ *and* $v > 0$

$$U((1-\lambda)v, S) \geq (1 - m_1(\lambda))U(v, S) - m_2(\lambda)\zeta.$$

Then Assumption 3.1.2 holds true.

Proof. In view of the fact that u is a non-decreasing function (follows from Assumption 3.1.1(i)) it sufficient to prove that for any $x > 0$

$$\lim_{\alpha \downarrow 0} u((1-\alpha)x) \geq \lim_{\alpha \downarrow 0} u((1+\alpha)x).$$

For any $\beta, y > 0$ the map $(y, \{\gamma_t\}_{t=0}^T) \to (\beta y, \{\beta\gamma_t\}_{t=0}^T)$ is a bijection between $\mathcal{A}(y)$ and $\mathcal{A}(\beta y)$. Thus,

$$\begin{aligned}
&\lim_{\alpha \downarrow 0} u((1-\alpha)x) \\
&\geq \lim_{\alpha \downarrow 0} \left(\left(1 - m_1\left(1 - \frac{1-\alpha}{1+\alpha}\right)\right) u((1+\alpha)x) - m_2\left(1 - \frac{1-\alpha}{1+\alpha}\right) \mathbb{E}_{\mathbb{P}}[\zeta] \right) \\
&= \lim_{\alpha \downarrow 0} u((1+\alpha)x).
\end{aligned}$$

□

Remark 3.1.5. *We notice that the power and the log utility satisfy the assumptions of Lemma 3.1.1. On the other hand for these utility functions Assumption 3.1.2 is straightforward.*

A "real" application of Lemma 3.1.1 is the case which corresponds to the utility function given by (3.1). In this case, if $v \geq \frac{S_T}{1-\lambda}$ *then* $U((1-\lambda)v, S) = U(v, S) = 0$*. If* $v < \frac{S_T}{1-\lambda}$ *then* $|U((1-\lambda)v, S) - U(v, S)| \leq \lambda v \leq \frac{\lambda}{1-\lambda} S_T$*. Thus, for* $m_1(\lambda) := 0$, $m_2(\lambda) := \frac{\lambda}{1-\lambda}$ *and* $\zeta := S_T$ *the assumptions of Lemma 3.1.1 hold true (provided that* $\mathbb{E}_{\mathbb{P}}[S_T] < \infty$*).*

Next, we treat Assumption 3.1.3(ii).

Lemma 3.1.2. *Suppose there exist constants $C > 0$, $0 < \gamma < 1$ and $q > \frac{1}{1-\gamma}$ which satisfy the following.*

(I) For all $(v, s) \in (0, \infty) \times \mathbb{D}([0, T]; \mathbb{R}^d)$,

$$U(v, s) \leq C(1 + v^\gamma). \tag{3.4}$$

(II) For any $n \in \mathbb{N}$ there exists a local martingale measure $\mathbb{Q}_n \in \mathcal{M}(S^{(n)})$ such that

$$\sup_{n \in \mathbb{N}} \mathbb{E}_{\mathbb{Q}_n}\left[\left(\frac{d\mathbb{P}_n}{d\mathbb{Q}_n}\right)^q\right] < \infty. \tag{3.5}$$

Then Assumption 3.1.3(ii) holds true.

Proof. Let $p = \frac{q}{q-1}$. Clearly $\frac{1}{p} > \gamma$. Thus in view of (3.4), in order to prove that Assumption 3.1.3(ii) holds true, it suffices to show that for any $x > 0$

$$\sup_{n \in \mathbb{N}} \sup_{\pi_n \in \mathcal{A}_n(x)} \mathbb{E}_{\mathbb{P}_n}[(V_T^{\pi_n})^{1/p}] < \infty.$$

For any $n \in \mathbb{N}$ and $\pi_n \in \mathcal{A}_n(x)$, $\{V_t^{\pi_n}\}_{t=0}^T$ is a $\mathbb{Q}_n$ super–martingale. Hence, from the Holder inequality (observe that $\frac{1}{p} + \frac{1}{q} = 1$) we get

$$\begin{aligned}
&\sup_{n \in \mathbb{N}} \sup_{\pi_n \in \mathcal{A}_n(x)} \mathbb{E}_{\mathbb{P}_n}[(V_T^{\pi_n})^{1/p}] \\
&\quad = \sup_{n \in \mathbb{N}} \sup_{\pi_n \in \mathcal{A}_n(x)} \mathbb{E}_{\mathbb{Q}_n}\left[(V_T^{\pi_n})^{1/p} \frac{d\mathbb{P}_n}{d\mathbb{Q}_n}\right] \\
&\quad \leq \sup_{n \in \mathbb{N}} \sup_{\pi_n \in \mathcal{A}_n(x)} (\mathbb{E}_{\mathbb{Q}_n}[V_T^{\pi_n}])^{1/p} \sup_{n \in \mathbb{N}} \left(\mathbb{E}_{\mathbb{Q}_n}\left[\left(\frac{d\mathbb{P}_n}{d\mathbb{Q}_n}\right)^q\right]\right)^{1/q} \\
&\quad \leq x^{1/p} \sup_{n \in \mathbb{N}} \left(\mathbb{E}_{\mathbb{Q}_n}\left[\left(\frac{d\mathbb{P}_n}{d\mathbb{Q}_n}\right)^q\right]\right)^{1/q} < \infty,
\end{aligned}$$

and the result follows. □

3.2 The necessity of Assumptions 3.1.4,3.1.5,3.1.6

3.2.1 On the necessity of Assumption 3.1.4

Let us explain by example why Assumption 3.1.4 is essential for the lower semi–continuity to hold.

Example 3.2.1. ***Naive discretization does not work.***

Let $d = 1$. Consider a random utility which corresponds to shortfall risk minimization for a call option with strike price $K > 0$. Namely, we set

$$U(v, s) := -((s_T - K)^+ - v)^+. \tag{3.1}$$

We have,

$$u(x) = - \inf_{\pi \in \mathcal{A}(x)} \mathbb{E}_{\mathbb{P}} \left[\left((S_T - K)^+ - V_T^\pi \right)^+ \right].$$

Consider the Black–Scholes model

$$S_t = S_0 e^{\sigma W_t - \sigma^2 t/2}, \quad t \in [0, T]$$

where $\sigma > 0$ is a constant volatility and $W = \{W_t\}_{t=0}^T$ is a Brownian motion (under $\mathbb{P}$).

We take the naive discretization and define the processes $S^{(n)}$, $n \in \mathbb{N}$, by

$$S_t^{(n)} := S_{\frac{kT}{n}}, \quad kT/n \le t < (k+1)T/n.$$

Let $\mathcal{F}^{(n)}$ the usual filtration which is generated by $S^{(n)}$. Namely,

$$\mathcal{F}_t^{(n)} := \sigma \left\{ S_{\frac{T}{n}}, ..., S_{\frac{kT}{n}}, \mathcal{N} \right\}, \quad kT/n \le t < (k+1)T/n$$

where $\mathcal{N}$ is the collection of all null sets. We also set $\mathbb{P}_n := \mathbb{P}$.

It is easy to see that $S^{(n)} \Rightarrow S$ and Assumptions 3.1.1–3.1.3 hold true (for Assumption 3.1.2 see Remark 3.1.5).

Next, we check Assumption 3.1.4. Fix n. Recall the processes $D^{(n)}, J^{(n)}$ from Assumption 3.1.4. First, observe that if $J^{(n)}$ is an adapted left continuous process, then for all $k < n$ $J^{(n)}_{\frac{(k+1)T}{n}}$ is $\mathcal{F}^{(n)}_{\frac{kT}{n}}$ measurable. Notice that for or all $k < n$,

$$ess \sup \left(S^{(n)}_{\frac{(k+1)T}{n}} - S^{(n)}_{\frac{kT}{n}} | \mathcal{F}^{(n)}_{\frac{kT}{n}} \right) = \infty \ a.s.$$

As usual $\text{ess}\sup(Y|\mathcal{G})$ *is the minimal random variable (which may take the value* ∞*) that is* $\mathcal{G}$ *measurable and* $\geq Y$ *a.s. These two simple observations yield that there is no (finite) adapted left continuous process* $\{J^{(n)}_t\}_{t=0}^T$ *which satisfy* $J^{(n)}_{\frac{(k+1)T}{n}} \geq D^{(n)}_{\frac{(k+1)T}{n}}$. *Thus, Assumption 3.1.4 is not satisfied.*

In [64] (see Section 6.1.2) it was proved that for the processes $S^{(n)}$, $n \in \mathbb{N}$ *defined above and the initial capital* $x := \mathbb{E}_{\mathbb{P}}[(S_T - K)^+]$ *(i.e. the Black–Scholes price) we have*

$$\liminf_{n\to\infty} \inf_{\pi_n \in \mathcal{A}_n(x)} \mathbb{E}_{\mathbb{P}}\left[\left((S_T - K)^+ - V_T^{\pi_n}\right)^+\right] > 0.$$

Clearly, the fact that x *is the Black–Scholes price implies that*

$$\inf_{\pi \in \mathcal{A}(x)} \mathbb{E}_{\mathbb{P}}\left[\left((S_T - K)^+ - V_T^{\pi}\right)^+\right] = 0.$$

We get

$$u(x) = 0 > \limsup_{n\to\infty} u_n(x),$$

and as a result Proposition III.1 does not hold true.

Example 3.2.2. ***Discrete approximations with vanishing growth rates do work.***

Consider a setup where for any n, $S^{(n)}$ *is a pure jump process of the form*

$$S^{(n)}_t = \sum_{i=1}^{m_n} S^{(n)}_{\tau_i^{(n)}} \mathbb{I}_{\tau_i^{(n)} \leq t < \tau_{i+1}^{(n)}} + S^{(n)}_T \mathbb{I}_{t=T}$$

where $m_n \in \mathbb{N}$ *and* $0 = \tau_1^{(n)} < \tau_2^{(n)} < ... < \tau_{m_n+1}^{(n)} = T$ *are stopping times with respect to* $\{\mathcal{F}^{(n)}_t\}_{t=0}^T$.

Assume that there exists a deterministic sequence $a_n > 0$, $n \in \mathbb{N}$ *such that* $\lim_{n\to\infty} a_n = 0$ *and*

$$|S^{(n)}_{\tau_{i+1}^{(n)}} - S^{(n)}_{\tau_i^{(n)}}| \leq a_n |S^{(n)}_{\tau_i^{(n)}}| \quad a.s, \quad \forall i, n.$$

Then Assumption 3.1.4 holds true with the processes

$$J_t^{(n)} := a_n \left(\sum_{i=1}^{m_n} \max_{1 \le j \le i} |S^{(n)}_{\tau_j^{(n)}}| \ \mathbb{I}_{\tau_i^{(n)} < t \le \tau_{i+1}^{(n)}} \right), \quad n \in \mathbb{N}.$$

In other words, if the growth rates go to zero uniformly then Assumption 3.1.4 holds true. This is exactly the case for binomial approximations of diffusion models with bounded volatility.

3.2.2 On the necessity of Assumption 3.1.5

A natural question to ask is whether Assumption 3.1.5 can be replaced by a simpler one.

In [46] the authors analyzed when weak convergence implies the convergence of option prices. Roughly speaking, the main result was that under contiguity properties of the sequences of physical measures with respect to the martingale measures there is a convergence of prices of derivative securities. The contiguity assumption (for the exact definition see [46]) is simpler than Assumption 3.1.5 and deals only with the approximating sequence. The main advantage of such assumption that it does not require establishing weak convergence (unlike Assumption 3.1.5). However, this classical result assumes that the limit model is complete. In general, in incomplete markets “strange phenomena” can happen as we will demonstrate in Example 3.2.3.

In Example 3.2.3 we construct a sequence of binomial (discrete) martingales $S^{(n)}$ considered with their natural filtrations that converge weakly to a continuous martingale S (the contiguity assumption trivially holds true). Surprisingly, the limiting model, which is given by the martingale S, is a fully incomplete market (see Definition 2.1 in [33]) and the set of all equivalent martingale measures is dense in the set of all martingale measures (for a precise formulation see Lemma 8.1 in [33]). We use this construction to illustrate that Assumption 3.1.5 is the “right” assumption to make.

The cornerstone of our construction is the following result which was established in [22] (see Theorem 8 there). For the reader's convenience we provide a short self-contained proof.

Lemma 3.2.1. *Let $\xi_i = \pm 1$, $i \in \mathbb{N}$ be i.i.d. and symmetric. Define the processes $W^{(n)}_t, \hat{W}^{(n)}_t$, $t \in [0,T]$ by*

$$W^{(n)}_t := \sqrt{\tfrac{T}{n}}\textstyle\sum_{i=1}^{k}\xi_i, \quad \tfrac{kT}{n} \le t < \tfrac{(k+1)T}{n},$$
$$\hat{W}^{(n)}_t := \sqrt{\tfrac{T}{n}}\textstyle\sum_{i=1}^{k}\prod_{j=1}^{i}\xi_j, \quad \tfrac{kT}{n} \le t < \tfrac{(k+1)T}{n}$$

where $\sum_{i=1}^{0} \equiv 0$. Then, we have the weak convergence

$$(W^{(n)}, \hat{W}^{(n)}) \Rightarrow (W, \hat{W}),$$

where W and $\hat{W}$ are independent Brownian motions.

Proof. We apply the martingale invariance principle given by Theorem 2.1 in [73]. For any n define the filtration $\{\mathcal{G}^{(n)}_t\}_{t=0}^{T}$ by $\mathcal{G}^{(n)}_t = \sigma\{\xi_1, ..., \xi_k\}$ for $kT/n \le t < (k+1)T/n$. Observe that $W^{(n)}, \hat{W}^{(n)}$ are martingales with respect to the filtration $\mathcal{G}^{(n)}$. Thus it remains to establish (2)–(3) in [73]. Clearly,

$$\sup_{0\le t\le T}|W^{(n)}_t - W^{(n)}_{t-}| = \sup_{0\le t\le T}|\hat{W}^{(n)}_t - \hat{W}^{(n)}_{t-}| = \sqrt{\frac{T}{n}},$$

and so the maximal jump size goes to zero as $n \to \infty$. Moreover, $[W^{(n)}]_t = [\hat{W}^{(n)}]_t = kT/n$ for $kT/n \le t < (k+1)T/n$. Thus, $[W^{(n)}]_t \to t$ and $[\hat{W}^{(n)}]_t \to t$ as $n \to \infty$.

It remains to show that for all $t \in [0,T]$

$$[W^{(n)}, \hat{W}^{(n)}]_t \to 0 \text{ in probability.} \tag{3.2}$$

Indeed, let $n \in \mathbb{N}$ and $kT/n \le t < (k+1)T/n$. Clearly,

$$[W^{(n)}, \hat{W}^{(n)}]_t = \frac{T}{n}\sum_{i=1}^{k}\prod_{j=1}^{i-1}\xi_j, \quad \frac{kT}{n} \le t < \frac{(k+1)T}{n},$$

where $\prod_{i=1}^{0} \equiv 1$. Observe that the random variables $\prod_{j=1}^{m} \xi_j = \pm 1$, $m \in \mathbb{N}$ are i.i.d. and symmetric. Thus,

$$\mathbb{E}\left(\left([W^{(n)}, \hat{W}^{(n)}]_t\right)^2\right) = \frac{T^2}{n^2}k \le \frac{Tt}{n}$$

and (3.2) follows. This completes the proof. $\square$

Example 3.2.3. ***Binomial models can converge weakly to fully incomplete markets.***

Let $d = 1$. For any $n \in \mathbb{N}$ define the stochastic processes $\{\nu_t^{(n)}\}_{t=0}^{T}$ and $\{S_t^{(n)}\}_{t=0}^{T}$ by

$$\nu_t^{(n)} := \prod_{i=1}^{k}\left(1 + \sqrt{\tfrac{T}{n}}\xi_i\right), \quad \tfrac{kT}{n} \le t < \tfrac{(k+1)T}{n},$$
$$S_t^{(n)} := \prod_{i=1}^{k}\left(1 + \min(\nu^{(n)}_{\frac{(i-1)T}{n}}, \ln n)\sqrt{\tfrac{T}{n}}\prod_{j=1}^{i}\xi_j\right), \quad \tfrac{kT}{n} \le t < \tfrac{(k+1)T}{n},$$

where $\xi_i = \pm 1$, $i \in \mathbb{N}$ are i.i.d. and symmetric. Let $\mathbb{P}_n$ be the corresponding probability measure.

We assume that n is sufficiently large so that $S^{(n)}$ and $\min(\nu^{(n)}, \ln n)$ are strictly positive. Let $\mathcal{F}^{(n)}$ be the filtration which is generated by $S^{(n)}$,

$$\mathcal{F}_t^{(n)} := \sigma\left\{S_{\frac{T}{n}}, ..., S_{\frac{kT}{n}}\right\}, \quad kT/n \le t < (k+1)T/n.$$

Observe that $\mathcal{F}_t^{(n)} = \sigma\{\xi_1, ..., \xi_k\}$ for $kT/n \le t < (k+1)T/n$. Moreover, the conditional support of supp $\left(S^{(n)}_{\frac{(k+1)T}{n}} | S^{(n)}_{\frac{T}{n}}, ..., S^{(n)}_{\frac{kT}{n}}\right)$ consists of exactly two points, and so the physical measure $\mathbb{P}_n$ is the unique martingale measure for $S^{(n)}$.

From Theorems 4.3–4.4 in [35] and Lemma 3.2.1 we obtain the weak convergence $(S^{(n)}, \nu^{(n)}) \Rightarrow (S, \nu)$ where (S, ν) is the (unique strong) solution of the SDE

$$dS_t = \nu_t S_t d\hat{W}_t, \quad S_0 = 1$$

(3.3)

$$d\nu_t = \nu_t dW_t, \quad \nu_0 = 1$$

where W and $\hat{W}$ are independent Brownian motions (under $\mathbb{P}$).

Namely, for the complete binomial models $S^{(n)}$, $n \in \mathbb{N}$ we have the weak convergence $S^{(n)} \Rightarrow S$ where S is the distribution of the stochastic volatility model given by (3.3). This is a specific case of the Hull–White model which was introduced in [47]. From Theorem 3.3 in [72] it follows that $\{S_t\}_{t=0}^T$ is a true martingale. Hence,

$$\mathbb{E}_{\mathbb{P}}[S_T] = S_0 = 1 = S_0^{(n)} = \mathbb{E}_{\mathbb{P}_n}[S_T^{(n)}].$$

This together with Theorem 3.6 in [18] gives that the random variables $\{S_T^{(n)}\}_{n\in\mathbb{N}}$ are uniformly integrable.

Let us observe that Assumption 3.1.5 does not hold true. Indeed, for any n we have the equality $\mathcal{M}(S^{(n)}) = \{\mathbb{P}_n\}$. Hence, $((S,1);\mathbb{P})$ is the only cluster point for the distributions $\left(\left(S^{(n)}, \frac{d\mathbb{Q}_n}{d\mathbb{P}_n}\right);\mathbb{P}_n\right)$, $\mathbb{Q}_n \in \mathcal{M}(S^{(n)})$. Since the set $\mathcal{M}(S)$ is not a singleton then clearly Assumption 3.1.5 is not satisfied.

Next, let $K > 0$. Consider a call option with strike price K and the utility function given by (3.1). Obviously, Assumption 3.1.1(i) and Assumption 3.1.3(ii) ($U^+ \equiv 0$) are satisfied. We want to demonstrate that Proposition III.2 does not hold true.

For any $n \in \mathbb{N}$ let $\mathbb{V}_n$ be the unique arbitrage free price of the above call option in the (complete) model given by $S^{(n)}$. From the weak convergence $S^{(n)} \Rightarrow S$ and the uniform integrability of $\{S_T^{(n)}\}_{n\in\mathbb{N}}$ we get

$$\lim_{n\to\infty} \mathbb{V}_n = \mathbb{E}_{\mathbb{P}}\left[(S_T - K)^+\right] < S_0 = 1.$$

In particular there exists $\epsilon > 0$ such that for sufficiently large n we have $\mathbb{V}_n < 1 - \epsilon$. Thus,

$$\lim_{n\to\infty} u_n(1-\epsilon) = 0.$$

On the other hand, the model given by S is a fully incomplete market (see Definition 2.1 and Example 2.5 in [33]). In [33, 65] it was proved that in fully incomplete *markets the super–replication price is prohibitively high and lead to buy–and–hold strategies. Namely, the super–hedging price of a call option is equal to the initial stock price $S_0 = 1$. Thus $u(1-\epsilon) < 0$ and so Proposition III.2 does not hold.*

3.2.3 On the necessity of Assumption 3.1.6

Example 3.2.4. ***Non-concave utility.***

Let $d = 1$. Assume that the investor utility function is given by

$$U(v, s) := \min(2, \max(v, 1)),$$

and depends only on the wealth. We notice that the function U does not satisfy Assumption 3.1.6.

For any $n \in \mathbb{N}$ consider the binomial model given by

$$S_t^{(n)} := \prod_{i=1}^{k}\left(1 + \frac{\xi_i}{n^2}\right), \quad \frac{kT}{n} \leq t < \frac{(k+1)T}{n},$$

where $\xi_i = \pm 1$, $i \in \mathbb{N}$ are i.i.d. and symmetric. Namely, $\mathbb{P}_n$ is the unique martingale measure for the n–th model. Clearly, for the constant process $S \equiv 1$ we have the weak convergence $S^{(n)} \Rightarrow S$. Thus, Assumption 3.1.1(i), Assumption 3.1.3(ii) and Assumption 3.1.5 are satisfied.

Next, consider the initial capital $x := 1$. Observe that for any n, there is a set $A_n \in \sigma\{\xi_1, ..., \xi_n\}$ with $\mathbb{P}_n(A_n) = 1/2$. Thus, from the completeness of the binomial models we get that there exists $\pi_n \in \mathcal{A}_n(1)$ such that $V_T^{\pi_n} = 2\mathbb{I}_{A_n}$. In particular,

$$u_n(1) \geq \mathbb{E}_{\mathbb{P}_n}[\min(2, \max(2\mathbb{I}_{A_n}, 1))] = 3/2, \quad n \in \mathbb{N}.$$

On the other hand, trivially $u(1) = 1$, which means that Proposition III.2 does not hold true.

The paper [70] studies the continuity of the value of the utility maximization problem from terminal wealth (under convergence in distribution) in a complete market. The author does not assume that the utility function is concave. The main result says that if the limit probability space is atomless and the atoms in approximating sequence of models are vanishing (see Assumption 2.1 in [70]) then continuity holds. Clearly, this is not satisfied in the Example 3.2.4 above where the filtration generated by the limit process is trivial.

An open question is to understand whether the continuity result from [70] can be extended to the incomplete case.

3.3 The Lower Semi–Continuity under Weak Convergence

In this section we prove Proposition III.1. We start by establishing a general result.

For any $M > 0$ and $n \in \mathbb{N}$ introduce the set $\Gamma_M^{(n)}$ of all simple predictable integrands of the from

$$\gamma_t^{(n)} = \sum_{i=1}^{k} \beta_i \mathbb{I}_{t_i < t \le t_{i+1}}$$

where $k \in \mathbb{N}$, $0 = t_1 < t_2 < .... < t_{k+1} = T$ is a deterministic partition and

$$\beta_i = \psi_i(S_{a_{i,1}}^{(n)}, ..., S_{a_{i,m_i}}^{(n)}), \quad i = 1, ..., k,$$

for a deterministic partition $0 = a_{i,1} < ... < a_{i,m_i} = t_i$ and a continuous function $\psi_i : (\mathbb{R}^d)^{m_i} \to \mathbb{R}^d$ that satisfies $|\psi_i| \le M$.

Lemma 3.3.1. *Let γ be a predictable process (with respect to $(\mathcal{F}_t^S)_{0 \le t \le T}$) with $|\gamma| \le M$ for some constant M. Then there exists a sequence $\gamma^{(n)} \in \Gamma_M^{(n)}$, $n \in \mathbb{N}$ such that we have the weak convergence*

$$\left(\{S_t^{(n)}\}_{t=0}^T, \left\{ \int_0^t \gamma_u^{(n)} dS_u^{(n)} \right\}_{t=0}^T \right) \Rightarrow \left(\{S_t\}_{t=0}^T, \left\{ \int_0^t \gamma_u dS_u \right\}_{t=0}^T \right) \tag{3.1}$$

on the space $\mathbb{D}([0,T];\mathbb{R}^d) \times \mathbb{D}([0,T];\mathbb{R})$.

Proof. On the space $(\Omega, \mathcal{F}, (\mathcal{F}_t^S)_{0\le t\le T}, \mathbb{P})$, let Γ_M be the set of all integrands of the form

$$\gamma_t = \sum_{i=1}^{k} \beta_i \mathbb{I}_{t_i < t \le t_{i+1}} \tag{3.2}$$

where $k \in \mathbb{N}$, $0 = t_1 < t_2 < .... < t_{k+1} = T$ is a deterministic partition and

$$\beta_i = \psi_i(S_{a_{i,1}}, ..., S_{a_{i,m_i}}), \quad i = 1, ..., k \tag{3.3}$$

for a deterministic partition $0 = a_{i,1} < ... < a_{i,m_i} = t_i$ and a continuous function $\psi_i : (\mathbb{R}^d)^{m_i} \to \mathbb{R}^d$ which satisfy $|\psi_i| \le M$. From standard density arguments it follows that for any $\epsilon > 0$ we can find $\gamma' \in \Gamma_M$ which satisfy

$$\mathbb{P}\left(\sup_{0\le t\le T}\left|\int_0^t \gamma_u dS_u - \int_0^t \gamma'_u dS_u\right| > \epsilon\right) < \epsilon.$$

Hence, without loss of generality we can assume that $\gamma \in \Gamma_M$. Thus, let γ be given by (3.2)–(3.3).

For any $n \in \mathbb{N}$ define $\gamma^{(n)} \in \Gamma_M^{(n)}$ by

$$\gamma_t^{(n)} := \sum_{i=1}^{k} \psi_i\left(S^{(n)}_{a_{i,1}}, ..., S^{(n)}_{a_{i,m_i}}\right) \mathbb{I}_{t_i < t \le t_{i+1}}, \quad t \in [0, T]. \tag{3.4}$$

It is well known that there exists a metric d on the Skorokhod space $\mathbb{D}([0,T];\mathbb{R}^d)$ that induces the Skorokhod topology and such that $\mathbb{D}([0,T];\mathbb{R}^d)$ is separable under d (for details see Chapter 3 in [18]). From the weak convergence $S^{(n)} \Rightarrow S$ and the Skorokhod representation theorem (see Theorem 3 in [34]) it follows that we can redefine the stochastic processes $S^{(n)}$, $n \in \mathbb{N}$ and S on the same probability space such that $\lim_{n\to\infty} d(S^{(n)}, S) = 0$ a.s. Recall that if $\lim_{n\to\infty} d(z^{(n)}, z) = 0$ and $z : [0,T] \to \mathbb{R}^d$ is a continuous function then $\lim_{n\to\infty} \sup_{0\le t\le T} |z_t^{(n)} - z_t| = 0$ (see e.g. Chapter 3 in [18]). We conclude that

$$\sup_{0\le t\le T} |S_t^{(n)} - S_t| \to 0 \text{ a.s.} \tag{3.5}$$

Next, recall the partition $0 = t_1 < t_2 < .... < t_{k+1} = T$ and redefine (on the common probability space) the integrands γ, $\gamma^{(n)}$ by the relations (3.2)–(3.4). Since these integrands are simple then the corresponding stochastic integrals $\int_0^t \gamma_u dS_u, \int_0^t \gamma_u^{(n)} dS_u^{(n)}$, $t \in [0,T]$ can be redefined as finite sums.

From (3.5) and the continuity of ψ_i, $i = 1, ..., k$ we get that

$$\sup_{0\le t\le T} |\gamma_t^{(n)} - \gamma_t| \to 0 \ \text{ a.s.}$$

Thus,

$$\begin{aligned}
&\sup_{0\le t\le T}\left|\int_0^t \gamma_u^{(n)} dS_u^{(n)} - \int_0^t \gamma_u dS_u\right| \\
&\quad = \sup_{0\le t\le T}\left|\sum_{i=1}^k \left(\gamma_{t_{i+1}}^{(n)}(S_{t_{i+1}\wedge t}^{(n)} - S_{t_i\wedge t}^{(n)}) - \gamma_{t_{i+1}}(S_{t_{i+1}\wedge t} - S_{t_i\wedge t})\right)\right| \\
&\quad \le \sup_{0\le t\le T}\left|\sum_{i=1}^k \gamma_{t_{i+1}}^{(n)}\left((S_{t_{i+1}\wedge t}^{(n)} - S_{t_i\wedge t}^{(n)}) - (S_{t_{i+1}\wedge t} - S_{t_i\wedge t})\right)\right| \\
&\quad + \sup_{0\le t\le T}\left|\sum_{i=1}^k (\gamma_{t_{i+1}}^{(n)} - \gamma_{t_{i+1}})(S_{t_{i+1}\wedge t} - S_{t_i\wedge t})\right| \\
&\quad \le 2Mkd \sup_{0\le t\le T}|S_t^{(n)} - S_t| \\
&\quad + 2kd \sup_{0\le t\le T}|\gamma_t^{(n)} - \gamma_t| \sup_{0\le t\le T}|S_t| \to 0 \ \text{ a.s.}
\end{aligned}$$

and the proof is completed. □

Now, we are ready to prove Proposition III.1.

Proof of Proposition III.1.

The proof will be done in two steps.

Step I: For any $x > 0$ let $\mathcal{A}_0(x) \subset \mathcal{A}(x)$ be the set of all admissible portfolios $\pi = (x, \gamma)$ such that γ is predictable, uniformly bounded and of bounded variation. In this step we show that for any $x_1 > x_2 > 0$

$$u(x_2) \le \sup_{\pi\in\mathcal{A}_0(x_1)} \mathbb{E}_{\mathbb{P}}[U(V_T^\pi, S)]. \tag{3.6}$$

A priori the left hand side and the right hand side of (3.6) can be both equal to ∞.

Let $\bar{\pi} = (x_2, \bar{\gamma}) \in \mathcal{A}(x_2)$ be an arbitrary portfolio. By applying the density argument given by Theorem 3.4 in [6] we obtain that there exists an adapted continuous process of bounded variation $\tilde{\gamma} = \{\tilde{\gamma}_t\}_{t=0}^T$ such that

$$\sup_{0\leq t\leq T}\left|\int_0^t \tilde{\gamma}_u dS_u - \int_0^t \bar{\gamma}_u dS_u\right| \leq \frac{x_1 - x_2}{2} \quad \text{a.s.}$$

We conclude that the portfolio which is given by $\tilde{\pi} := (x_1, \tilde{\gamma})$ satisfies

$$V_t^{\tilde{\pi}} \geq V_t^{\bar{\pi}} + \frac{x_1 - x_2}{2} \geq \frac{x_1 - x_2}{2}, \quad t \in [0, T]. \tag{3.7}$$

Next, for the continuous process $\tilde{\gamma}$ define the stopping times

$$\theta_n := T \wedge \inf\{t : |\tilde{\gamma}_t| = n\}, \quad n \in \mathbb{N}$$

and the trading strategies

$$\tilde{\gamma}_t^{(n)} := \mathbb{I}_{t\leq\theta_n}\tilde{\gamma}_t, \quad t \in [0, T].$$

Set $\tilde{\pi}_n = (x_1, \tilde{\gamma}^{(n)})$. Clearly, $|\tilde{\gamma}^{(n)}| \leq n$ and from (3.7) we have

$$V_t^{\tilde{\pi}_n} = V_{t\wedge\theta_n}^{\tilde{\pi}} \geq \frac{x_1 - x_2}{2}, \quad t \in [0, T].$$

Hence, $\tilde{\pi}_n \in \mathcal{A}_0(x_1)$. Observe that $\theta_n \uparrow T$ a.s., and so

$$\lim_{n\to\infty} V_T^{\tilde{\pi}_n} = \lim_{n\to\infty} V_{\theta_n}^{\tilde{\pi}} = V_T^{\tilde{\pi}}.$$

This together with Fatou's Lemma, Assumption 3.1.1 (notice that $V_T^{\tilde{\pi}_n} \geq \frac{x_1-x_2}{2} > 0$), the fact that U is continuous and (3.7) gives

$$\mathbb{E}_{\mathbb{P}}[U(V_T^{\bar{\pi}}, S)] \leq \mathbb{E}_{\mathbb{P}}[U(V_T^{\tilde{\pi}}, S)] \leq \liminf_{n\to\infty} \mathbb{E}_{\mathbb{P}}[U(V_T^{\tilde{\pi}_n}, S)].$$

Since $\bar{\pi} \in \mathcal{A}(x_2)$ was arbitrary we complete the proof of (3.6).

Step II: In view of (3.6) and Assumption 3.1.2, in order to prove Proposition III.1

it sufficient to show that for any initial capital $x > 0$, $0 < \epsilon < \frac{x}{2}$ and $\pi \in \mathcal{A}_0(x - 2\epsilon)$ there exists a sequence $\pi_n \in \mathcal{A}_n(x)$, $n \in \mathbb{N}$ which satisfies

$$\liminf_{n\to\infty} \mathbb{E}_{\mathbb{P}_n}[U(V_T^{\pi_n}, S^{(n)})] \geq \mathbb{E}_{\mathbb{P}}[U(V_T^{\pi}, S)]. \tag{3.8}$$

Let $0 < \epsilon < \frac{x}{2}$ and $\pi = (x - 2\epsilon, \gamma) \in \mathcal{A}_0(x - 2\epsilon)$. Let $M > 0$ such that $|\gamma| \leq M$. Lemma 3.3.1 provides the existence of simple integrands $\gamma^{(n)} \in \Gamma_M^{(n)}$, $n \in \mathbb{N}$ which satisfy (3.1).

For a given n, the portfolio $(x, \gamma^{(n)})$ might fail to be admissible and so a modification is needed. Recall Assumption 3.1.4 and the stochastic processes $J^{(n)}$, $n \in \mathbb{N}$. For any $n \in \mathbb{N}$ introduce the stopping time

$$\Theta_n := T \wedge \inf\left\{t : x + \int_0^{t-} \gamma_u^{(n)} dS_u^{(n)} < \epsilon + MdJ_t^{(n)}\right\}, \tag{3.9}$$

and define the portfolio $\pi_n = (x, \bar{\gamma}^{(n)})$ by $\bar{\gamma}_t^{(n)} := \mathbb{I}_{t\leq\Theta_n}\gamma_t^{(n)}$. Let us show that $V_t^{\pi_n} \geq \epsilon$ for all $t \in [0, T]$. Indeed,

$$\begin{aligned} V_t^{\pi_n} &= x + \int_0^{t\wedge\Theta_n} \gamma_u^{(n)} dS_u^{(n)} \\ &\geq x + \int_0^{t\wedge\Theta_n -} \gamma_u^{(n)} dS_u^{(n)} - Md|S_{t\wedge\Theta_n -}^{(n)} - S_{t\wedge\Theta_n}^{(n)}| \\ &\geq \epsilon + MdJ_{t\wedge\Theta_n}^{(n)} - Md|S_{\Theta}^{(n)} - S_{\Theta -}^{(n)}| \geq \epsilon \end{aligned}$$

as required. The first inequality follows from $|\gamma^{(n)}| \leq M$. The second inequality follows from the fact that on the time interval $[0, \Theta_n)$ se have $x + \int_0^{\cdot-} \gamma_u^{(n)} dS_u^{(n)} \geq \epsilon + MdJ_{\cdot}^{(n)}$. The last inequality is due to $J_{\Theta}^{(n)} \geq |S_{\Theta}^{(n)} - S_{\Theta-}^{(n)}|$. We conclude that $\pi_n \in \mathcal{A}_n(x)$ and

$$V_T^{\pi_n} = x + \int_0^{\Theta_n} \gamma_u^{(n)} dS_u^{(n)} \geq \epsilon. \tag{3.10}$$

Next, we apply the Skorokhod representation theorem. Recall that the processes $\{J_t^{(n)}\}_{t=0}^T$, $n \in \mathbb{N}$ are non–negative, non decreasing and $J_T^{(n)} \to 0$ in probability. This together with (3.1) implies that we have the weak convergence

$$\left(\{J_t^{(n)}\}_{t=0}^T, \{S_t^{(n)}\}_{t=0}^T, \left\{\int_0^t \gamma_u^{(n)} dS_u^{(n)}\right\}_{t=0}^T\right) \Rightarrow \left(0, \{S_t\}_{t=0}^T, \left\{\int_0^t \gamma_u dS_u\right\}_{t=0}^T\right) \tag{3.11}$$

on the space $\mathbb{D}([0,T];\mathbb{R}) \times \mathbb{D}([0,T];\mathbb{R}^d) \times \mathbb{D}([0,T];\mathbb{R})$.

For any $n \in \mathbb{N}$ the integrand $\gamma^{(n)}$ is of the form (3.4). Hence the integrand $\gamma^{(n)}$ and the corresponding stochastic integral $\int_0^\cdot \gamma_u^{(n)} dS_u^{(n)}$ are determined pathwise by $S^{(n)}$. Since γ is of bounded variation then we have

$$\int_0^t \gamma_u dS_u = \gamma_t S_t - \gamma_0 S_0 - \int_0^t S_u d\gamma_u,$$

where the last term is the pathwise Riemann–Stieltjes integral. We conclude that γ and the corresponding stochastic integral $\int_0^\cdot \gamma_u dS_u$ are determined pathwise by S.

Thus, from the Skorokhod representation theorem and (3.11) it follows that we can redefine the stochastic processes $\gamma^{(n)}, S^{(n)}, J^{(n)}$, $n \in \mathbb{N}$ and γ, S on the same probability space such that (3.5) holds true,

$$\sup_{0\le t\le T} J_t^{(n)} \to 0 \ \text{ a.s.} \tag{3.12}$$

and

$$\sup_{0\le t\le T} \left| \int_0^t \gamma_u^{(n)} dS_u^{(n)} - \int_0^t \gamma_u dS_u \right| \to 0 \text{ a.s.} \tag{3.13}$$

As in (3.5) the uniform convergence is due to the fact that the limit processes are continuous. By applying (3.9) we redefine Θ_n, $n \in \mathbb{N}$ on the common probability space. With abuse of notations we denote by $\mathbb{P}$ and $\mathbb{E}$ the probability and the expectation on the common probability space, respectively.

First, we argue that

$$\lim_{n\to\infty} \mathbb{P}(\Theta_n = T) = 1. \tag{3.14}$$

Recall, the admissible portfolio $\pi = (x - 2\epsilon, \gamma)$. From (3.13) it follows that

$$\lim_{n\to\infty}\inf \inf_{0\le t\le T}\left(x + \int_0^t \gamma_u^{(n)} dS_u^{(n)}\right) = x + \inf_{0\le t\le T}\int_0^t \gamma_u dS_u \ge 2\epsilon.$$

In particular

$$\lim_{n\to\infty}\mathbb{P}\left(\inf_{0\le t\le T}\left(x + \int_0^t \gamma_u^{(n)} dS_u^{(n)}\right) > \frac{3\epsilon}{2}\right) = 1.$$

This together with (3.12) gives (3.14).

Finally, from Fatou's Lemma, the continuity of U, Assumption 3.1.1(i), Assumption 3.1.3(i) (recall that $V_T^{\pi_n} \ge \epsilon$), (3.5), (3.10) and (3.13)–(3.14) we obtain

$$\begin{aligned}\lim_{n\to\infty}\inf \mathbb{E}_{\mathbb{P}_n}[U(V_T^{\pi_n}, S^{(n)})] &= \lim_{n\to\infty}\inf \mathbb{E}\left[U\left(x + \int_0^{\Theta_n} \gamma_u^{(n)} dS_u^{(n)}, S^{(n)}\right)\right] \\ &\ge \mathbb{E}\left[U\left(x + \int_0^T \gamma_u dS_u, S\right)\right] \ge \mathbb{E}_{\mathbb{P}}[U(V_T^{\pi}, S)],\end{aligned}$$

and (3.8) follows. □

3.4 The Upper Semi–Continuity under Weak Convergence

In this section we prove Proposition III.2.

Proof. Let $x > 0$. From Assumption 3.1.3(ii) it follows that for any $n \in \mathbb{N}$ $u_n(x) < \infty$. Hence, we can choose a sequence $\hat{\pi}_n \in \mathcal{A}_n(x)$, $n \in \mathbb{N}$ which satisfy (3.3). Without loss of generality (by passing to a subsequence) we assume that the limit $\lim_{n\to\infty} \mathbb{E}_{\mathbb{P}_n}[U(V_T^{\hat{\pi}_n}, S^{(n)})]$ exists. We will prove that there exists $\hat{\pi} \in \mathcal{A}(x)$ such that

(3.1) $$\mathbb{E}_{\mathbb{P}}[U(V_T^{\hat{\pi}}, S)] \ge \lim_{n\to\infty}\mathbb{E}_{\mathbb{P}_n}[U(V_T^{\hat{\pi}_n}, S^{(n)})]$$

and this will give Proposition III.2. The proof will be done in two steps.

Step I: Choose $\mathbb{Q} \in \mathcal{M}(S)$ (recall that we assume $\mathcal{M}(S) \neq \emptyset$) and set $Z := \frac{d\mathbb{Q}}{d\mathbb{P}}$. From Assumption 3.1.5 it follows that there exists a sequence $\mathbb{Q}_n \in \mathcal{M}(S^{(n)})$, $n \in \mathbb{N}$ for which (3.1) holds true. For any n, $\{V_t^{\hat{\pi}_n}\}_{t=0}^T$ is a $\mathbb{Q}_n$ super–martingale. Hence,

$$\mathbb{E}_{\mathbb{P}_n}\left(V_T^{\hat{\pi}_n}\frac{d\mathbb{Q}_n}{d\mathbb{P}_n}\right) = \mathbb{E}_{\mathbb{Q}_n}[V_T^{\hat{\pi}_n}] \le V_0^{\hat{\pi}_n} = x.$$

We conclude that the sequence $\left(V_T^{\hat{\pi}_n}\frac{d\mathbb{Q}_n}{d\mathbb{P}_n};\mathbb{P}_n\right)$, $n\in\mathbb{N}$ is tight. This together with (3.1) yields that the sequence $\left(\left(S^{(n)},\frac{d\mathbb{Q}_n}{d\mathbb{P}_n},V_T^{\hat{\pi}_n}\frac{d\mathbb{Q}_n}{d\mathbb{P}_n}\right);\mathbb{P}_n\right)$, $n\in\mathbb{N}$ is tight on the space $\mathbb{D}([0,T];\mathbb{R}^d)\times\mathbb{R}^2$. From Prohorov's theorem it follows that there exists a subsequence $\left(\left(S^{(n)},\frac{d\mathbb{Q}_n}{d\mathbb{P}_n},V_T^{\hat{\pi}_n}\frac{d\mathbb{Q}_n}{d\mathbb{P}_n}\right);\mathbb{P}_n\right)$ (for simplicity the subsequence is still denoted by n) which converge weakly. From (3.1) we obtain that

$$\left(\left(S^{(n)},\frac{d\mathbb{Q}_n}{d\mathbb{P}_n},V_T^{\hat{\pi}_n}\frac{d\mathbb{Q}_n}{d\mathbb{P}_n}\right);\mathbb{P}_n\right)\Rightarrow(S,Z,Y), \tag{3.2}$$

where Y is some random variable. In particular we have the weak convergence

$$\left(\left(S^{(n)},V_T^{\hat{\pi}_n}\right);\mathbb{P}_n\right)\Rightarrow\left(S,\frac{Y}{Z}\right). \tag{3.3}$$

The random vector (S,Z,Y) is defined on a new probability space $(\tilde{\Omega},\tilde{\mathcal{F}},\tilde{\mathbb{P}})$, which might be different from the original probability space $(\Omega,\mathcal{F},\mathbb{P})$. We redefine the filtration $\mathcal{F}^S$ (the usual filtration which is generated by S) and the sets $\mathcal{M}(S),\mathcal{A}(\cdot)$ (as before, these sets defined with respect to $\mathcal{F}^S$) on the new probability space $(\tilde{\Omega},\tilde{\mathcal{F}},\tilde{\mathbb{P}})$.

Set (notice that $\frac{Y}{Z}\geq 0$)

$$V:=\mathbb{E}_{\tilde{\mathbb{P}}}\left(\frac{Y}{Z}\mid\mathcal{F}_T^S\right) \tag{3.4}$$

where a priori V can be equal to ∞ with finite probability. In order to prove (3.1) it is sufficient to show that there exists $\hat{\pi}\in\mathcal{A}(x)$ such that

$$V_T^{\hat{\pi}}\geq V \text{ a.s.} \tag{3.5}$$

Indeed, if (3.5) holds true (in particular $V<\infty$ a.s.), then from the Jensen inequality, the continuity of U, Assumption 3.1.1(i), Assumption 3.1.3(ii), Assumption 3.1.6 and (3.3) we obtain

$$\mathbb{E}_{\mathbb{P}}[U(V_T^{\hat{\pi}},S)]\geq\mathbb{E}_{\tilde{\mathbb{P}}}[U(V,S)]\geq\mathbb{E}_{\tilde{\mathbb{P}}}\left[U\left(\frac{Y}{Z},S\right)\right]\geq\lim_{n\to\infty}\mathbb{E}_{\mathbb{P}_n}[U(V_T^{\hat{\pi}_n},S^{(n)})] \tag{3.6}$$

as required.

This brings us to the second step.

Step II: In this step we establish (3.5). In view of the optional decomposition theorem (Theorem 3.2 in [55]) it is sufficient to show that the super-hedging price which is given by $\sup_{\hat{\mathbb{Q}}\in\mathcal{M}(S)} \mathbb{E}_{\hat{\mathbb{Q}}}[V]$ is less or equal than x. From (3.4) we obtain

$$\sup_{\hat{\mathbb{Q}}\in\mathcal{M}(S)} \mathbb{E}_{\hat{\mathbb{Q}}}[V] = \sup_{\hat{\mathbb{Q}}\in\mathcal{M}(S)} \mathbb{E}_{\tilde{\mathbb{P}}}\left[\frac{Y}{Z}\frac{d\hat{\mathbb{Q}}}{d\tilde{\mathbb{P}}}\right].$$

Hence, it remains to prove that for any $\hat{\mathbb{Q}} \in \mathcal{M}(S)$

$$x \geq \mathbb{E}_{\tilde{\mathbb{P}}}\left[\frac{Y}{Z}\frac{d\hat{\mathbb{Q}}}{d\tilde{\mathbb{P}}}\right]. \tag{3.7}$$

From Assumption 3.1.5 we get a sequence $\hat{\mathbb{Q}}_n \in \mathcal{M}(S^{(n)})$, $n \in \mathbb{N}$ for which

$$\left(\left(S^{(n)}, \frac{d\hat{\mathbb{Q}}_n}{d\mathbb{P}_n}\right); \mathbb{P}_n\right) \Rightarrow \left(\left(S, \frac{d\hat{\mathbb{Q}}}{d\mathbb{P}}\right); \mathbb{P}\right). \tag{3.8}$$

This together with (3.2) yields that the sequence

$$\left(\left(S^{(n)}, \frac{d\mathbb{Q}_n}{d\mathbb{P}_n}, V_T^{\pi_n}\frac{d\mathbb{Q}_n}{d\mathbb{P}_n}, \frac{d\hat{\mathbb{Q}}_n}{d\mathbb{P}_n}\right); \mathbb{P}_n\right), \quad n \in \mathbb{N},$$

is tight on the space $\mathbb{D}([0,T];\mathbb{R}^d) \times \mathbb{R}^3$. From Prohorov's Theorem and (3.2) there is a subsequence which converge weakly

$$\left(\left(S^{(n)}, \frac{d\mathbb{Q}_n}{d\mathbb{P}_n}, V_T^{\hat{\pi}_n}\frac{d\mathbb{Q}_n}{d\mathbb{P}_n}, \frac{d\hat{\mathbb{Q}}_n}{d\mathbb{P}_n}\right); \mathbb{P}_n\right) \Rightarrow (S, Z, Y, X) \tag{3.9}$$

for some random variable X.

Once again, the random vector (S, Z, Y, X) is defined on a new probability space $(\tilde{\tilde{\Omega}}, \tilde{\tilde{\mathcal{F}}}, \tilde{\tilde{\mathbb{P}}})$, on which we redefine the filtration $\mathcal{F}^S$ and the sets $\mathcal{M}(S), \mathcal{A}(\cdot)$.

Observe that $\frac{d\hat{\mathbb{Q}}}{d\mathbb{P}}$ is determined by S. Hence, there exists a measurable function $g : \mathbb{D}([0,T];\mathbb{R}^d) \to \mathbb{R}$ such that $\frac{d\hat{\mathbb{Q}}}{d\mathbb{P}} = g(S)$ $\mathbb{P}$ a.s, i.e. $\mathbb{E}_{\mathbb{P}}|\frac{d\hat{\mathbb{Q}}}{d\mathbb{P}} - g(S)| = 0$. From (3.8)–(3.9) we get that the distribution of (S, X) equals to $\left(\left(S, \frac{d\hat{\mathbb{Q}}}{d\mathbb{P}}\right); \mathbb{P}\right)$. Thus, $\mathbb{E}_{\tilde{\tilde{\mathbb{P}}}}|X - g(S)| = 0$. We conclude that $X = g(S)$ $\tilde{\tilde{\mathbb{P}}}$ a.s.

Finally, from Fatou's Lemma, (3.9) and the fact that $\{V_t^{\hat{\pi}_n}\}_{t=0}^T$ is a $\hat{\mathbb{Q}}_n$ super–martingale it follows that

$$\mathbb{E}_{\tilde{\mathbb{P}}}\left(\frac{Y}{Z}\frac{d\hat{\mathbb{Q}}}{d\tilde{\mathbb{P}}}\right) = \mathbb{E}_{\tilde{\mathbb{P}}}\left(\frac{Y}{Z}g(S)\right) = \mathbb{E}_{\tilde{\tilde{\mathbb{P}}}}\left(\frac{Yg(S)}{Z}\right) = \mathbb{E}_{\tilde{\tilde{\mathbb{P}}}}\left(\frac{YX}{Z}\right)$$

$$\leq \lim_{n\to\infty}\inf \mathbb{E}_{\mathbb{P}_n}\left(V_T^{\hat{\pi}_n}\frac{d\mathbb{Q}_n}{d\mathbb{P}_n}\frac{\frac{d\hat{\mathbb{Q}}_n}{d\mathbb{P}_n}}{\frac{d\mathbb{Q}_n}{d\mathbb{P}_n}}\right) = \lim_{n\to\infty}\inf \mathbb{E}_{\hat{\mathbb{Q}}_n}[V_T^{\hat{\pi}_n}] \leq x,$$

from which we get (3.7). □

Next, we prove Theorem 3.1.2.

3.4.1 Proof of Theorem 3.1.2

In order to prove the statement it is sufficient to show the for any sub–sequence of laws $\left(S^{(n)}, V_T^{\hat{\pi}_n}\right)$ there is a further subsequence which converge weakly to $\left(S, V_T^{\hat{\pi}}\right)$. We stay with notation of the proof of Proposition III.2.

Following the same arguments as in the proof of Proposition III.2 we obtain for any sub–sequence of laws $\left(S^{(n)}, V_T^{\hat{\pi}_n}\right)$ that there is a further sequence which satisfies (3.3). Moreover, there exists $\hat{\pi} \in \mathcal{A}(x)$ such that (3.5)–(3.6) hold true.

From (3.6), Theorem 3.1.1 (holds true because the required Assumptions are satisfied) and the fact that $\hat{\pi}_n \in \mathcal{A}_n(x)$ are asymptotically optimal we get

$$u(x) \geq \mathbb{E}_{\mathbb{P}}[U(V_T^{\hat{\pi}}, S)] \geq \mathbb{E}_{\tilde{\mathbb{P}}}\left[U\left(\mathbb{E}_{\tilde{\mathbb{P}}}\left(\frac{Y}{Z} \mid \mathcal{F}_T^S\right), S\right)\right] \geq \mathbb{E}_{\tilde{\mathbb{P}}}\left[U\left(\frac{Y}{Z}, S\right)\right] \geq u(x).$$

We conclude that all the above inequalities are in fact equalities. This together with (3.5) and the assumption that U is strictly concave and strictly increasing in the first variable (follows from Assumption 3.1.1(i) and the strict concavity) implies that

$$V_T^{\hat{\pi}} = \mathbb{E}_{\tilde{\mathbb{P}}}\left(\frac{Y}{Z} \mid \mathcal{F}_T^S\right) = \frac{Y}{Z}$$

and $\hat{\pi} \in \mathcal{A}(x)$ is the unique optimal portfolio. This completes the proof. □

We end this section with a remark on how our results can be generalized.

Remark 3.4.1. *Consider the case where the filtration* $\mathcal{F} := \mathcal{F}^{S,Y}$ *is the usual filtration generated by* S *and another RCLL process* $R = (R^{(1)}_t, ..., R^{(m)}_t)_{0 \le t \le T}$. *The process* R *can be viewed as a collection of non traded assets.*

For the approximate model we take $(S^{(n)}, R^{(n)})$ *and a filtration which satisfies the usual assumptions and makes both* $S^{(n)}$ *and* $R^{(n)}$ *adapted. Once again* $R^{(n)} = (R^{n,1}_t, ..., R^{n,m}_t)_{0 \le t \le T}$ *is the collection of non traded assets. Consider a continuous utility function* $U : (0, \infty) \times \mathbb{D}([0, T]; \mathbb{R}^d) \times \mathbb{D}([0, T]; \mathbb{R}^m) \to \mathbb{R}$ *and assume the weak convergence* $(S^{(n)}, R^{(n)}) \Rightarrow (S, R)$ *and an analogous assumptions to those in Section 3.1. Of course, as before the martingale measures are with respect to the traded assets. Then, by using similar arguments as in Sections 3.3–3.4 we can extend the main results Theorems 3.1.1-3.1.2 to this setup as well.*

3.5 Lattice Based Approximations of the Heston Model

Consider the Heston model [44] given by

$$d\hat{S}_t = \hat{S}_t(\mu dt + \sqrt{\hat{\nu}_t} dW_t),$$

$$d\hat{\nu}_t = \kappa(\theta - \hat{\nu}_t)dt + \sigma\sqrt{\hat{\nu}_t} d\tilde{W}_t,$$

where $\mu \in \mathbb{R}$, $\kappa, \theta, \sigma > 0$ are constants and $W, \tilde{W}$ are two standard Brownian motions with a constant correlation $\rho \in (-1, 1)$. The initial values $\hat{S}_0, \hat{\nu}_0 > 0$ are given. We assume the condition $2\kappa\theta > \sigma^2$ which guarantees that $\hat{\nu}$ does not touch zero (see [24]).

For technical reasons our approximations require that the volatility will lie in an interval of the form $[\underline{\sigma}, \overline{\sigma}]$ for some $0 < \underline{\sigma} < \overline{\sigma}$. Thus, we modify the Heston model as following. Fix two barriers $0 < \underline{\sigma} < \overline{\sigma}$ and define the function

$$h(z) := \max(\underline{\sigma}^2, \min(z, \overline{\sigma}^2)), \quad z \in \mathbb{R}.$$

Consider the SDE

$$dS_t = S_t(\mu dt + \sqrt{h(\nu_t)}dW_t)$$

(3.10)

$$d\nu_t = \kappa(\theta - h(\nu_t))dt + \sigma\sqrt{h(\nu_t)}d\tilde{W}_t$$

where the initial values are $S_0 := \hat{S}_0$, $\nu_0 := \hat{\nu}_0$. Observe that $\sqrt{h}, h$ are Lipschitz continuous, and so (3.10) has a unique solution.

We expect that if $\underline{\sigma}$ is small and $\overline{\sigma}$ is large then the value of the utility maximization problem in the Heston model will be close to the one in the model given by (3.10). For the shortfall risk measure we provide an error estimate in Lemma 3.6.1.

3.5.1 Discretization

In this section we construct discrete time lattice based approximations for the model given by (3.10). The novelty of our constructions is that the approximating sequence satisfies Assumptions 3.1.4,3.1.5.

It is more convenient to work with a transformed system of equations driven by independent Brownian motions. Therefore, we set

$$\Phi_t := \ln S_t, \quad \Psi_t := \frac{\nu_t}{\sigma} - \rho\Phi_t.$$

From Itô's formula we obtain that

$$d\Phi_t = \mu_\Phi(\Phi_t, \Psi_t)dt + \sigma_\Phi(\Phi_t, \Psi_t)dW_t$$

$$d\Psi_t = \mu_\Psi(\Phi_t, \Psi_t)dt + \sigma_\Psi(\Phi_t, \Psi_t)d\hat{W}_t$$

where

$$\mu_\Phi(y, z) := \mu - h\left(\sigma(\rho y + z)\right)/2, \quad \sigma_\Phi(y, z) := \sqrt{h\left(\sigma(\rho y + z)\right)},$$

$$\mu_\Psi(y, z) := \tfrac{\kappa}{\sigma}\left(\theta - h\left(\sigma(\rho y + z)\right)\right) - \rho\mu_\Phi(y, z), \quad \sigma_\Psi := \sqrt{(1 - \rho^2)}\sigma_\Phi,$$

and $\hat{W} := \frac{\tilde{W}-\rho W}{\sqrt{1-\rho^2}}$ is a Brownian motion independent of W.

Next, we define lattice based approximations for the process (Φ, Ψ). Choose $\tilde{\sigma} \geq \overline{\sigma}$. For any $n \in \mathbb{N}$ define the stochastic processes $\Phi^{(n)}_t, \Psi^{(n)}_t$, $t \in [0,T]$ by

$$\Phi^{(n)}_t := \Phi_0 + \tilde{\sigma}\sqrt{\tfrac{T}{n}} \textstyle\sum_{i=1}^k \xi_i, \quad \tfrac{kT}{n} \leq t < \tfrac{(k+1)T}{n}$$
$$\Psi^{(n)}_t := \Psi_0 + \tilde{\sigma}\sqrt{\tfrac{T}{n}} \textstyle\sum_{i=1}^k \hat{\xi}_i, \quad \tfrac{kT}{n} \leq t < \tfrac{(k+1)T}{n}$$

where $\xi_i, \hat{\xi}_i \in \{-1,0,1\}$. Observe that the processes $\Phi^{(n)} - \Phi_0$, $\Psi^{(n)} - \Psi_0$ lie on the grid $\tilde{\sigma}\sqrt{\frac{T}{n}}\{-n, 1-n, ..., n\}$.

Let $\mathcal{F}^{(n)}_t$, $t \leq T$ be the piece wise constant filtration generated by the processes $\Phi^{(n)}, \Psi^{(n)}$. Namely,

$$\mathcal{F}^{(n)}_t := \sigma\left\{\xi_1, ..., \xi_k, \hat{\xi}_1, ..., \hat{\xi}_k\right\}, \quad kT/n \leq t < (k+1)T/n.$$

It remains to define the probability measure $\mathbb{P}_n$. First since W and $\hat{W}$ are independent Brownian motions we require that for all $a, b \in \{-1,0,1\}$ and $k \geq 1$

$$\mathbb{P}_n\left(\xi_k = a, \hat{\xi}_k = b | \mathcal{F}^{(n)}_{\frac{(k-1)T}{n}}\right) := \mathbb{P}_n\left(\xi_k = a | \mathcal{F}^{(n)}_{\frac{(k-1)T}{n}}\right)\mathbb{P}_n\left(\hat{\xi}_k = b | \mathcal{F}^{(n)}_{\frac{(k-1)T}{n}}\right).$$

In order to match the drift and the volatility, we set,

$$\mathbb{P}_n\left(\xi_k = \pm 1 | \mathcal{F}^{(n)}_{\frac{(k-1)T}{n}}\right) := \frac{\sigma^2_\Phi\left(\Phi^{(n)}_{\frac{(k-1)T}{n}}, \Psi^{(n)}_{\frac{(k-1)T}{n}}\right)}{2\tilde{\sigma}^2} \pm \sqrt{\frac{T}{n}}\frac{\mu_\Phi\left(\Phi^{(n)}_{\frac{(k-1)T}{n}}, \Psi^{(n)}_{\frac{(k-1)T}{n}}\right)}{2\tilde{\sigma}},$$
$$\mathbb{P}_n\left(\xi_k = 0 | \mathcal{F}^{(n)}_{\frac{(k-1)T}{n}}\right) := 1 - \frac{\sigma^2_\Phi\left(\Phi^{(n)}_{\frac{(k-1)T}{n}}, \Psi^{(n)}_{\frac{(k-1)T}{n}}\right)}{\tilde{\sigma}^2},$$

and

$$\mathbb{P}_n\left(\hat{\xi}_k = \pm 1 | \mathcal{F}^{(n)}_{\frac{(k-1)T}{n}}\right) := \frac{\sigma^2_\Psi\left(\Phi^{(n)}_{\frac{(k-1)T}{n}}, \Psi^{(n)}_{\frac{(k-1)T}{n}}\right)}{2\tilde{\sigma}^2} \pm \sqrt{\frac{T}{n}}\frac{\mu_\Psi\left(\Phi^{(n)}_{\frac{(k-1)T}{n}}, \Psi^{(n)}_{\frac{(k-1)T}{n}}\right)}{2\tilde{\sigma}},$$
$$\mathbb{P}_n\left(\hat{\xi}_k = 0 | \mathcal{F}^{(n)}_{\frac{(k-1)T}{n}}\right) := 1 - \frac{\sigma^2_\Psi\left(\Phi^{(n)}_{\frac{(k-1)T}{n}}, \Psi^{(n)}_{\frac{(k-1)T}{n}}\right)}{\tilde{\sigma}^2}.$$

Observe that for sufficiently large n, the right hand side of the above equations all lie in the interval $[0,1]$.

Proposition III.3. *For any $n \in \mathbb{N}$ (sufficiently large) consider the financial market given by $S^{(n)} := e^{\Phi^{(n)}}$ and the filtration $\mathcal{F}^{(n)}$ defined above. Then, the following holds true.*

(I) We have the weak convergence $S^{(n)} \Rightarrow S$ to the modified Heston model.

(II) Assumption 3.1.4 holds true.

Proof.

Proof of (I). Let us prove that

$$(\Phi^{(n)}, \Psi^{(n)}) \Rightarrow (\Phi, \Psi). \tag{3.11}$$

Clearly, (3.11) implies that $S^{(n)} \Rightarrow S$.

From the definition of $\mathbb{P}_n$ we have

$$\mathbb{E}_{\mathbb{P}_n}\left(\Phi^{(n)}_{\frac{kT}{n}} - \Phi^{(n)}_{\frac{(k-1)T}{n}} \Big| \mathcal{F}^{(n)}_{\frac{(k-1)T}{n}}\right) = \frac{T}{n}\mu_\Phi\left(\Phi^{(n)}_{\frac{(k-1)T}{n}}, \Psi^{(n)}_{\frac{(k-1)T}{n}}\right), \tag{3.12}$$

$$\mathbb{E}_{\mathbb{P}_n}\left(\Psi^{(n)}_{\frac{kT}{n}} - \Psi^{(n)}_{\frac{(k-1)T}{n}} \Big| \mathcal{F}^{(n)}_{\frac{(k-1)T}{n}}\right) = \frac{T}{n}\mu_\Psi\left(\Phi^{(n)}_{\frac{(k-1)T}{n}}, \Psi^{(n)}_{\frac{(k-1)T}{n}}\right), \tag{3.13}$$

$$\mathbb{E}_{\mathbb{P}_n}\left((\Phi^{(n)}_{\frac{kT}{n}} - \Phi^{(n)}_{\frac{(k-1)T}{n}})^2 \Big| \mathcal{F}^{(n)}_{\frac{(k-1)T}{n}}\right) = \frac{T}{n}\sigma^2_\Phi\left(\Phi^{(n)}_{\frac{(k-1)T}{n}}, \Psi^{(n)}_{\frac{(k-1)T}{n}}\right),$$
$$\mathbb{E}_{\mathbb{P}_n}\left((\Psi^{(n)}_{\frac{kT}{n}} - \Psi^{(n)}_{\frac{(k-1)T}{n}})^2 \Big| \mathcal{F}^{(n)}_{\frac{(k-1)T}{n}}\right) = \frac{T}{n}\sigma^2_\Psi\left(\Phi^{(n)}_{\frac{(k-1)T}{n}}, \Psi^{(n)}_{\frac{(k-1)T}{n}}\right)$$

and

$$\mathbb{E}_{\mathbb{P}_n}\left((\Phi^{(n)}_{\frac{kT}{n}} - \Phi^{(n)}_{\frac{(k-1)T}{n}})(\Psi^{(n)}_{\frac{kT}{n}} - \Psi^{(n)}_{\frac{(k-1)T}{n}}) \Big| \mathcal{F}^{(n)}_{\frac{(k-1)T}{n}}\right) = O(n^{-2}).$$

Thus, (3.11) follows from the the martingale convergence result Theorem 7.4.1 in [37].

Proof of II. The statement follows from applying Example 3.2.2 for $m_n = n$, $\tau_i^{(n)} = (i-1)T/n$ and $a_n = e^{\tilde{\sigma}\sqrt{\frac{T}{n}}} - 1$. □

3.5.2 Verification of Assumption 3.1.5

We start with some preparations. Denote by $\mathcal{D}$ the set of all stochastic processes $\Upsilon = \{\Upsilon_t\}_{t=0}^T$ of the form $\Upsilon = F(\Phi)$ where $F : \mathbb{D}([0,T];\mathbb{R}) \to \mathbb{D}([0,T];\mathbb{R})$ is a bounded, continuous function (we take the Skorokhod topology on the space $\mathbb{D}([0,T];\mathbb{R})$ and F is a progressively measurable map. Namely, for any $t \in [0,T]$ and $f^{(1)}, f^{(2)} \in \mathbb{D}([0,T];\mathbb{R})$, $f^{(1)}_{[0,t]} = f^{(2)}_{[0,t]}$ implies that $F_t(f^{(1)}) = F_t(f^{(2)})$.

Define the set

$$\mathcal{M}^d(S) :=$$
$$\left\{\mathbb{Q} : \ \exists \Upsilon \in \mathcal{D}, \ \frac{d\mathbb{Q}}{d\mathbb{P}}|\mathcal{F}^S_T = e^{\int_0^T \frac{-\mu}{\sqrt{h(\nu_t)}} dW_t + \int_0^T \Upsilon_t d\hat{W}_t - \int_0^T \frac{\mu^2}{2h(\nu_t)} dt - \int_0^T \frac{1}{2}\Upsilon_t^2 dt}\right\}.$$

From the Girsanov theorem it follows that $\mathcal{M}^d(S) \subset \mathcal{M}(S)$. Moreover, since $\Phi = \ln S$ then the usual filtration which is generated by S equals to the usual filtration which is generated by Φ. Hence standard arguments yield that $\mathcal{M}^d(S) \subset \mathcal{M}(S)$ is dense.

Choose an arbitrary $\Upsilon = F(\Phi) \in \mathcal{D}$ and denote

$$Z_t := e^{\int_0^t \frac{-\mu}{\sqrt{h(\nu_u)}} dW_u + \int_0^t \Upsilon_u d\hat{W}_u - \int_0^t \frac{\mu^2}{2h(\nu_u)} du - \int_0^t \frac{1}{2}\Upsilon_u^2 du}, \quad t \in [0,T]. \tag{3.14}$$

It is sufficient to prove that (recall Remark 3.1.2) there exists a sequence of probability measures $\mathbb{Q}_n \in \mathcal{M}(S^{(n)})$, $n \in \mathbb{N}$, such that for the processes $Z^{(n)}_t := \frac{d\mathbb{Q}_n}{d\mathbb{P}_n}|\mathcal{F}^{(n)}_t$, $t \in [0,T]$, we have the weak convergence

$$(S^{(n)}, Z^{(n)}) \Rightarrow (S, Z). \tag{3.15}$$

For any $n \in \mathbb{N}$ (sufficiently large) define the probability measure $\mathbb{Q}_n$ by the following relations

$$\mathbb{Q}_n\left(\xi_k = a, \hat{\xi}_k = b|\mathcal{F}^{(n)}_{\frac{(k-1)T}{n}}\right) := \mathbb{Q}_n\left(\xi_k = a|\mathcal{F}^{(n)}_{\frac{(k-1)T}{n}}\right)\mathbb{Q}_n\left(\hat{\xi}_k = b|\mathcal{F}^{(n)}_{\frac{(k-1)T}{n}}\right),$$

$$
(3.16)\qquad
\begin{aligned}
\mathbb{Q}_n\left(\xi_k = \pm 1 | \mathcal{F}^{(n)}_{\frac{(k-1)T}{n}}\right) &:= \frac{\sigma^2_\Phi\left(\Phi^{(n)}_{\frac{(k-1)T}{n}}, \Psi^{(n)}_{\frac{(k-1)T}{n}}\right)}{\tilde{\sigma}^2\left(1+e^{\pm\tilde{\sigma}\sqrt{\frac{T}{n}}}\right)},\\
\mathbb{Q}_n\left(\xi_k = 0 | \mathcal{F}^{(n)}_{\frac{(k-1)T}{n}}\right) &:= 1 - \frac{\sigma^2_\Phi\left(\Phi^{(n)}_{\frac{(k-1)T}{n}}, \Psi^{(n)}_{\frac{(k-1)T}{n}}\right)}{\tilde{\sigma}^2},
\end{aligned}
$$

and

$$
\begin{aligned}
\mathbb{Q}_n\left(\hat{\xi}_k = \pm 1 | \mathcal{F}^{(n)}_{\frac{(k-1)T}{n}}\right) &:= \frac{\sigma^2_\Psi\left(\Phi^{(n)}_{\frac{(k-1)T}{n}}, \Psi^{(n)}_{\frac{(k-1)T}{n}}\right)}{2\tilde{\sigma}^2}\\
&\pm \sqrt{\frac{T}{n}} \frac{F_{\frac{(k-1)T}{n}}(\Phi^{(n)})\sigma_\Psi\left(\Phi^{(n)}_{\frac{(k-1)T}{n}}, \Psi^{(n)}_{\frac{(k-1)T}{n}}\right) + \mu_\Psi\left(\Phi^{(n)}_{\frac{(k-1)T}{n}}, \Psi^{(n)}_{\frac{(k-1)T}{n}}\right)}{2\tilde{\sigma}},\\
\mathbb{Q}_n\left(\hat{\xi}_k = 0 | \mathcal{F}^{(n)}_{\frac{(k-1)T}{n}}\right) &:= 1 - \frac{\sigma^2_\Psi\left(\Phi^{(n)}_{\frac{(k-1)T}{n}}, \Psi^{(n)}_{\frac{(k-1)T}{n}}\right)}{\tilde{\sigma}^2}.
\end{aligned}
$$

Observe that (3.16) implies $\mathbb{Q}_n \in \mathcal{M}(S^{(n)})$.

Lemma 3.5.1. *We have the weak convergence*

$$
(\Phi^{(n)}, \Psi^{(n)}, Z^{(n)}) \Rightarrow (\Phi, \Psi, Z).
$$

Proof. In order to prove the lemma it suffices to show that for any subsequence there exists a further subsequence (still denoted by n) such that

$$
(3.17)\qquad (\Phi^{(n)}, \Psi^{(n)}, Z^{(n)}) \Rightarrow (\Phi, \Psi, Z).
$$

Fix $n \in \mathbb{N}$. By applying Taylor's expansion we obtain that there exist uniformly bounded (in n) processes $E^{n,1}_k, E^{n,2}_k$, $k = 0, 1, ..., n$ such that

$$
\frac{\mathbb{Q}_n\left(\xi_k | \mathcal{F}^{(n)}_{\frac{(k-1)T}{n}}\right)}{\mathbb{P}_n\left(\xi_k | \mathcal{F}^{(n)}_{\frac{(k-1)T}{n}}\right)} = 1 - \tilde{\sigma}\xi_k\sqrt{\frac{T}{n}}\left(\frac{1}{2} + \frac{\mu_\Phi\left(\Phi^{(n)}_{\frac{(k-1)T}{n}}, \Psi^{(n)}_{\frac{(k-1)T}{n}}\right)}{\sigma^2_\Phi\left(\Phi^{(n)}_{\frac{(k-1)T}{n}}, \Psi^{(n)}_{\frac{(k-1)T}{n}}\right)}\right) + \frac{E^{n,1}_k}{n} + o(1/n)
$$

and

$$
\frac{\mathbb{Q}_n\left(\hat{\xi}_k | \mathcal{F}^{(n)}_{\frac{(k-1)T}{n}}\right)}{\mathbb{P}_n\left(\hat{\xi}_k | \mathcal{F}^{(n)}_{\frac{(k-1)T}{n}}\right)} = 1 + \tilde{\sigma}\hat{\xi}_k\sqrt{\frac{T}{n}}\frac{F_{\frac{(k-1)T}{n}}(\Phi^{(n)})}{\sigma_\Psi\left(\Phi^{(n)}_{\frac{(k-1)T}{n}}, \Psi^{(n)}_{\frac{(k-1)T}{n}}\right)} + \frac{E^{n,2}_k}{n} + o(1/n).
$$

We conclude that there exists a uniformly bounded (in n) process $E^{(n)}_k$, $k = 0, 1..., n$ such that

$$\frac{Z^{(n)}_{\frac{kT}{n}} - Z^{(n)}_{\frac{(k-1)T}{n}}}{Z^{(n)}_{\frac{(k-1)T}{n}}} = \frac{\mathbb{Q}_n\left(\xi_k|\mathcal{F}^{(n)}_{\frac{(k-1)T}{n}}\right)}{\mathbb{P}_n\left(\xi_k|\mathcal{F}^{(n)}_{\frac{(k-1)T}{n}}\right)} \frac{\mathbb{Q}_n\left(\hat{\xi}_k|\mathcal{F}^{(n)}_{\frac{(k-1)T}{n}}\right)}{\mathbb{P}_n\left(\hat{\xi}_k|\mathcal{F}^{(n)}_{\frac{(k-1)T}{n}}\right)} - 1$$

$$(3.18) \qquad = -(\Phi^{(n)}_{\frac{kT}{n}} - \Phi^{(n)}_{\frac{(k-1)T}{n}})\left(\frac{1}{2} + \frac{\mu_\Phi\left(\Phi^{(n)}_{\frac{(k-1)T}{n}}, \Psi^{(n)}_{\frac{(k-1)T}{n}}\right)}{\sigma^2_\Phi\left(\Phi^{(n)}_{\frac{(k-1)T}{n}}, \Psi^{(n)}_{\frac{(k-1)T}{n}}\right)}\right)$$

$$+ (\Psi^{(n)}_{\frac{kT}{n}} - \Psi^{(n)}_{\frac{(k-1)T}{n}})\frac{F_{\frac{(k-1)T}{n}}(\Phi^{(n)})}{\sigma_\Psi\left(\Phi^{(n)}_{\frac{(k-1)T}{n}}, \Psi^{(n)}_{\frac{(k-1)T}{n}}\right)} + \frac{E^{(n)}_k}{n} + o(1/n).$$

In particular, $\left(\frac{Z^{(n)}_{\frac{kT}{n}} - Z^{(n)}_{\frac{(k-1)T}{n}}}{Z^{(n)}_{\frac{(k-1)T}{n}}}\right)^2$ is of order $O(1/n)$. Since $Z^{(n)}$ is a martingale, then by taking conditional expectation we arrive to

$$\mathbb{E}_{\mathbb{P}_n}\left([Z^{(n)}_{\frac{kT}{n}}]^2|\mathcal{F}^{(n)}_{\frac{(k-1)T}{n}}\right) = [Z^{(n)}_{\frac{(k-1)T}{n}}]^2(1 + O(1/n)).$$

By taking expectation we obtain

$$\mathbb{E}_{\mathbb{P}_n}\left([Z^{(n)}_{\frac{kT}{n}}]^2\right) = \mathbb{E}_{\mathbb{P}_n}\left([Z^{(n)}_{\frac{(k-1)T}{n}}]^2\right)(1 + O(1/n)).$$

This together with the Doob–Kolmogorov inequality gives

$$(3.19) \qquad \sup_{n\in\mathbb{N}} \mathbb{E}_{\mathbb{P}_n}\left(\sup_{0\leq t\leq T}[Z^{(n)}_t]^2\right) \leq 4 \sup_{n\in\mathbb{N}} \mathbb{E}_{\mathbb{P}_n}\left([Z^{(n)}_T]^2\right) < \infty.$$

Next, define $\hat{E}^{(n)}_k := \mathbb{E}_{\mathbb{P}_n}\left(E^{(n)}_k|\mathcal{F}^{(n)}_{\frac{(k-1)T}{n}}\right)$, $k = 1, ..., n$ and consider the martingale

$$\hat{M}^{(n)}_k := \frac{1}{n}\sum_{i=1}^k (E^{(n)}_i - \hat{E}^{(n)}_i), \quad k = 0, 1, ..., n.$$

Since $E^{(n)}$, $n \in \mathbb{N}$, are uniformly bounded then

$$\mathbb{E}_{\mathbb{P}_n}\left(\max_{0\leq k\leq n}|\hat{M}^{(n)}_k|^2\right) \leq 4\mathbb{E}_{\mathbb{P}_n}\left(|\hat{M}^{(n)}_n|^2\right)$$
$$= \tfrac{4}{n^2}\sum_{i=1}^n \mathbb{E}_{\mathbb{P}_n}\left[\left(E^{(n)}_i - \hat{E}^{(n)}_i\right)^2\right] = O(1/n).$$

Thus,

$$\max_{0\le k\le n} |\hat{M}_k^{(n)}| \to 0 \ \text{ in probability.} \tag{3.20}$$

Introduce the adapted (to $\mathcal{F}^{(n)}$) processes

$$\Xi_t^{(n)} := \int_0^t \hat{E}^{(n)}_{\lfloor * \rfloor nu/T} du, \quad t\in[0,T]$$

$$M_t^{(n)} := \hat{M}^{(n)}_{\lfloor * \rfloor nt/T}, \quad t\in[0,T]$$

where $\lfloor * \rfloor \cdot$ is the integer part of $\cdot$ and $\hat{E}_0^{(n)} := E_0^{(n)}$.

Again, $E^{(n)}$, $n\in\mathbb{N}$, are uniformly bounded, and so $\Xi^{(n)}$, $n\in\mathbb{N}$, is tight. From (3.11) and (3.20) we conclude that the sequence $(\Phi^{(n)},\Psi^{(n)},\Xi^{(n)},M^{(n)})$, $n\in\mathbb{N}$, is tight as well. Thus, from Prohorov's Theorem, (3.11) and (3.20) it follows that for any subsequence there exists a further subsequence such that

$$(\Phi^{(n)},\Psi^{(n)},\Xi^{(n)},M^{(n)}) \Rightarrow (\Phi,\Psi,\Xi,0) \tag{3.21}$$

for some absolutely continuous process $\Xi=\{\Xi_t\}_{t=0}^T$. From Theorems 4.3–4.4 in [35], (3.18), (3.21) and the equality $\frac{E_k^{(n)}}{n} = \frac{\hat{E}_k^{(n)}}{n} + \hat{M}_k^{(n)} - \hat{M}_{k-1}^{(n)}$ we obtain that

$$(\Phi^{(n)},\Psi^{(n)},\Xi^{(n)},M^{(n)},Z^{(n)}) \Rightarrow (\Phi,\Psi,\Xi,0,\hat{Z})$$

where $\hat{Z}$ is the solution of the SDE

$$\frac{d\hat{Z}_t}{\hat{Z}_t} = -\left(\frac{1}{2} + \frac{\mu_\Phi(\Phi_t,\Psi_t)}{\sigma_\Phi^2(\Phi_t,\Psi_t)}\right) d\Phi_t + \frac{\Upsilon_t}{\sigma_\Psi(\Phi_t,\Psi_t)} d\Psi_t + \frac{d\Xi_t}{T} \tag{3.22}$$

with the initial condition $\hat{Z}_0 = 1$.

Finally, (3.19) implies that for any $t\in[0,T]$ the random variables $\{Z_t^{(n)}\}_{n\in\mathbb{N}}$ are uniformly integrable. This together with the fact that for any n, $Z^{(n)}$ is a martingale with respect to the filtration generated by $\Phi^{(n)},\Psi^{(n)},\Xi^{(n)},M^{(n)},Z^{(n)}$ gives that $\hat{Z}$ is a martingale with respect to the filtration generated by $\Phi,\Psi,\Xi,\hat{Z}$. Moreover, from

(3.12)–(3.13) we get that $\{\Phi_t - \int_0^t \mu_\Phi(\Phi_u, \Psi_u)du\}_{t=0}^T$ and $\{\Psi_t - \int_0^t \mu_\Psi(\Phi_u, \Psi_u)du\}_{t=0}^T$ are martingales with respect to the filtration generated by $\Phi, \Psi, \Xi, \hat{Z}$. In particular, from Lévy's Theorem it follows that the stochastic processes W and $\hat{W}$ which we redefine by

$$W_t := \frac{\Phi_t - \int_0^t \mu_\Phi(\Phi_u, \Psi_u)du}{\sigma_\Phi(\Phi_t, \Psi_t)}, \quad \hat{W}_t := \frac{\Psi_t - \int_0^t \mu_\Psi(\Phi_u, \Psi_u)du}{\sigma_\Psi(\Phi_t, \Psi_t)}$$

are (independent) Brownian motions with respect to the filtration generated by $\Phi, \Psi, \Xi, \hat{Z}$. We conclude that the drift of the right hand side of (3.22) is equal to zero. Namely,

$$\frac{d\hat{Z}_t}{\hat{Z}_t} = -\left(\frac{1}{2} + \frac{\mu_\Phi(\Phi_t, \Psi_t)}{\sigma^2_\Phi(\Phi_t, \Psi_t)}\right)\sigma_\Phi(\Phi_t, \Psi_t)dW_t + \Upsilon_t d\hat{W}_t = \frac{dZ_t}{Z_t},$$

where the last equality follows from (3.14). Hence, $\hat{Z} = Z$ and (3.17) follows. $\square$

Clearly, Lemma 3.5.1 implies (3.15). This gives us the following result.

Proposition III.4. *Consider the set-up of Proposition III.3. Assumption 3.1.5 holds true.*

We end this section by addressing condition (II) in Lemma 3.1.2.

Remark 3.5.1. *Consider the martingale measures $\mathbb{Q}_n \in \mathcal{M}(S^{(n)})$, $n \in \mathbb{N}$ which were defined before Lemma 3.5.1 for $\Upsilon \equiv 0$. Since $\mu^\Phi, \sigma^\Phi, \frac{1}{\sigma^\Phi}$ are uniformly bounded, then standard arguments yield that for any $q > 0$ (3.5) holds true.*

3.6 Approximations of the Shortfall Risk in the Heston Model

In this section we focus on shortfall risk minimization for European call options (which corresponds to U given by (3.1)) in the Heston model. We start with the following estimate.

Lemma 3.6.1. *For an initial capital x let $\hat{R}(x)$ be the shortfall risk in the Heston model and let $R(x)$ be the shortfall risk in the model given by (3.10). Then for any $m \in \mathbb{N}$*

$$|\hat{R}(x) - R(x)| \leq O(\underline{\sigma}^{2\kappa\theta/\sigma^2 - 1}) + O(1/\overline{\sigma}^m),$$

where the O terms do not depend on x.

Proof. Define the stopping time

$$\Theta_{\underline{\sigma},\overline{\sigma}} := T \wedge \inf\{t : \sqrt{\hat{\nu}_t} \notin (\underline{\sigma}, \overline{\sigma})\}.$$

Observe that on the event $\Theta_{\underline{\sigma},\overline{\sigma}} = T$ the processes $\hat{S}$ and S coincide. Hence,

$$(3.23)\qquad |\hat{R}(x) - R(x)| \leq \mathbb{E}_{\mathbb{P}}[(\hat{S}_T + S_T)\mathbb{I}_{\Theta_{\underline{\sigma},\overline{\sigma}} < T}] \leq 2e^{\mu T}\mathbb{E}_{\mathbb{P}}[e^{-\mu\Theta_{\underline{\sigma},\overline{\sigma}}}\hat{S}_{\theta_{\underline{\sigma},\overline{\sigma}}}\mathbb{I}_{\Theta_{\underline{\sigma},\overline{\sigma}} < T}]$$

where the last inequality is due to the fact that the processes $e^{-\mu t}\hat{S}_t, e^{-\mu t}S_t$, $t \in [0, T]$ are martingales.

Introduce the probability measure $\mathbf{P}$ by $\frac{d\mathbf{P}}{d\mathbb{P}}|\mathcal{F}^S_T := \frac{e^{-\mu\Theta_{\underline{\sigma},\overline{\sigma}}}\hat{S}_{\Theta_{\underline{\sigma},\overline{\sigma}}}}{S_0}$. Then by Girsanov theorem the process $\mathbf{W}_t := \tilde{W}_t - \rho\int_0^{t \wedge \Theta_{\underline{\sigma},\overline{\sigma}}}\sqrt{\hat{\nu}_u}$, $t \in [0, T]$, is a Brownian motion with respect to $\mathbf{P}$. Let $\{\alpha_t\}_{t=0}^T$ be the unique strong solution of the SDE

$$d\alpha_t = (\kappa(\theta - \alpha_t) + \sigma\rho\alpha_t)\,dt + \sigma\sqrt{\alpha_t}d\mathbf{W}_t, \quad \alpha_0 = \hat{\nu}_0.$$

Observe that

$$(3.24)\qquad \alpha_{[0,\Theta_{\underline{\sigma},\overline{\sigma}}]} = \hat{\nu}_{[0,\Theta_{\underline{\sigma},\overline{\sigma}}]}.$$

Clearly, for any $m \in \mathbb{N}$ we have

$$\mathbb{E}_{\mathbf{P}}\left(\sup_{0 \leq t \leq T}[\sqrt{\alpha_t}]^m\right) < \infty.$$

Thus, from the Markov inequality we get

$$(3.25)\qquad \mathbf{P}\left(\sup_{0 \leq t \leq T}\sqrt{\alpha_t} \geq \overline{\sigma}\right) = O(1/\overline{\sigma}^m), \quad \forall m \in \mathbb{N}.$$

Moreover, from Theorem 2 in [41] it follows that

$$\mathbf{P}\left(\inf_{0\le t\le T}\sqrt{\alpha_t}\le\underline{\sigma}\right)=O(\underline{\sigma}^{2\kappa\theta/\sigma^2-1}). \tag{3.26}$$

By combining (3.23)–(3.26) we conclude that

$$\begin{aligned}|\hat{R}(x)-R(x)| &\le 2S_0e^{\mu T}\left(\mathbf{P}\left(\inf_{0\le t\le T}\sqrt{\alpha_t}\le\underline{\sigma}\right)+\mathbf{P}\left(\sup_{0\le t\le T}\sqrt{\alpha_t}\ge\overline{\sigma}\right)\right)\\ &\le O(\underline{\sigma}^{2\kappa\theta/\sigma^2-1})+O(1/\overline{\sigma}^m)\end{aligned}$$

as required. □

Next, we focus on approximating the shortfall risk in the model given by (3.10). In order to apply Theorem 3.1.1 we need to verify the required Assumptions. Observe that Assumption 3.1.1, Assumption 3.1.3(ii) ($U^+\equiv 0$) and Assumption 3.1.6 trivially hold true. Moreover, from Remark 3.1.5 we obtain Assumption 3.1.2. Since the drift and the volatility are uniformly bounded we get that the random variables $\{S_T^{(n)}\}_{n\in\mathbb{N}}$ are uniformly integrable, which gives Assumption 3.1.3(i). In view of Propositions III.3,III.4 we conclude that our Assumptions are satisfied and so Theorem 3.1.1 holds true.

Thus, fix $n\in\mathbb{N}$ and recall the discrete models introduced in Section 3.5.1. The stock price process $S^{(n)}$ is piece wise constant and so the investor trades only at the jump times $\frac{kT}{n}$, $k=0,1...,n$. Notice that $\left\{\sum_{m=1}^k\xi_m,\sum_{m=1}^k\hat{\xi}_m\right\}_{k=0}^n$ is a lattice valued Markov chain (with respect to $\mathbb{P}_n$). Hence, we introduce the functions $J_k^{(n)}(i,j,\lambda)$, $k=0,1...,n$ such that $J_k^{(n)}(i,j,\lambda)$ measures the shortfall risk at time kT/n given that $\sum_{m=1}^k\xi_m=i$, $\sum_{m=1}^k\hat{\xi}_m=j$, and λ is the ratio of the portfolio value and the stock price. The stock price is recovered by

$$S^{(n)}_{\frac{kT}{n}}=S_0e^{\tilde{\sigma}\sqrt{\frac{T}{n}}\sum_{m=1}^k\xi_m}=S_0e^{i\tilde{\sigma}\sqrt{\frac{T}{n}}}.$$

Clearly, if $\lambda \geq 1$, then the shortfall risk is zero because we can buy the stock and hold it until maturity. Namely, $J_k^{(n)}(i,j,\lambda) = 0$ for $\lambda \geq 1$. Hence, we assume that $\lambda \in [0,1]$.

Next, we describe the dynamic programming principle to solve the discrete control-problem. At time kT/n the investor decides about his investment policy. Assume that the investor portfolio value is $\lambda S^{(n)}_{\frac{kT}{n}}$. We have a trinomial model with growth rates $\left\{e^{-\tilde{\sigma}\sqrt{\frac{T}{n}}}, 1, e^{\tilde{\sigma}\sqrt{\frac{T}{n}}}\right\}$. From the binomial representation theorem we easily deduce that the set of replicable portfolios at time $(k+1)T/n$ are of the form $\Lambda(\xi_{k+1})S^{(n)}_{\frac{(k+1)T}{n}}$ where $\Lambda : \{-1,0,1\} \to \mathbb{R}$ satisfies $\Lambda(0) = \lambda$ and

$$\frac{\Lambda(-1) + \Lambda(1)e^{\tilde{\sigma}\sqrt{\frac{T}{n}}}}{1 + e^{\tilde{\sigma}\sqrt{\frac{T}{n}}}} = \lambda.$$

Thus, if $\Lambda(-1)$ is known then we set

$$\Lambda(1) := 1 \wedge \left(\lambda(1 + e^{-\tilde{\sigma}\sqrt{\frac{T}{n}}}) - \Lambda(-1)e^{-\tilde{\sigma}\sqrt{\frac{T}{n}}}\right). \tag{3.27}$$

We take a truncation in order to have $\Lambda(1) \in [0,1]$. In view of our admissibility condition, we denote by $\mathcal{A}(\lambda)$ the set of all $\Lambda(-1) \in [0,1]$ for which the right hand side of (3.27) is non-negative.

We arrive to the following recursive relations. Define

$$J_k^{(n)}(i,j,\lambda) : \{-k, 1-k, ..., k\} \times \{-k, 1-k, ..., k\} \times [0,1] \to \mathbb{R}_+, \quad k = 0,1,...,n$$

by

$$J_n^{(n)}(i,j,\lambda) := U\left(\lambda S_0 \exp\left(i\tilde{\sigma}\sqrt{\frac{T}{n}}\right), S_0 \exp\left(i\tilde{\sigma}\sqrt{\frac{T}{n}}\right)\right),$$

and for $k < n$,

(3.28)

$$J_k^{(n)}(i,j,\lambda) :=$$

$$\sup_{\Lambda(-1)\in\mathcal{A}(\lambda)} \mathbb{E}_{\mathbb{P}_n}\left(J_{k+1}^{(n)}\left(i+\xi_{m+1}, j+\hat{\xi}_{m+1}, \Lambda(\xi_{m+1})\right) \middle| \sum_{m=1}^{k}\xi_m = i, \ \sum_{m=1}^{k}\hat{\xi}_m = j\right)$$

where $\Lambda(0)=\lambda$ and $\Lambda(1)$ is given by (3.27). For $k=0$ we have $J^{(n)}_0(x/S_0)=u_n(x)$.

Observe that the functions $J^{(n)}_k(i,j,\lambda)$ are piece wise linear and continuous in λ, and so they can be represented by an array which consists of the slope values and the slope jump points. This together with the condition $J^{(n)}_k(i,j,1)=0$ is sufficient to recover the function. Of course the array will depend on time kT/n and the states i,j. Thus, theoretically, the dynamic programming given by (3.28) can be implemented using a computer. However, from practical point of view the number of the slope points of the function $J^{(n)}_k$ grows exponentially (in $n-k$), and so for large n it cannot be implemented. Hence, we need to introduce a grid structure for the portfolio value as well.

Thus, choose $M\in\mathbb{N}$ and consider the grid

$$GR:=\left\{0,\frac{1}{M},\frac{2}{M},...,1\right\}. \tag{3.29}$$

For a given $\Lambda(-1)\in GR$ we define two grid values for $\Lambda(1)$. The first value is

$$\Lambda^{-}(1):=1\wedge\frac{\lfloor*\rfloor\left(\lambda(1+e^{-\tilde{\sigma}\sqrt{\frac{T}{n}}})-\Lambda(-1)e^{-\tilde{\sigma}\sqrt{\frac{T}{n}}}\right)M}{M} \tag{3.30}$$

where, recall that $\lfloor*\rfloor\cdot$ is the integer part of $\cdot$. The second value is

$$\Lambda^{+}(1):=1\wedge\frac{\lceil*\rceil\left(\lambda(1+e^{-\tilde{\sigma}\sqrt{\frac{T}{n}}})-\Lambda(-1)e^{-\tilde{\sigma}\sqrt{\frac{T}{n}}}\right)M+1}{M} \tag{3.31}$$

where $\lceil*\rceil\cdot=\min\{n\in\mathbb{Z}:n\geq\cdot\}$. Define two value functions

(3.32)

$$J^{(n)}_k(\pm,i,j,\lambda):\{-k,1-k,...,k\}\times\{-k,1-k,...,k\}\times GR\to\mathbb{R}_+,\quad k=0,1,...,n$$

as following. The terminal condition is

$$J^{(n)}_n(\pm,i,j,\lambda):=U\left(\lambda S_0\exp\left(i\tilde{\sigma}\sqrt{\frac{T}{n}}\right),S_0\exp\left(i\tilde{\sigma}\sqrt{\frac{T}{n}}\right)\right).$$

For $k < n$,

$$J_k^{(n)}(\pm, i, j, \lambda)$$
$$:= \max_{\Lambda(-1)\in\mathcal{A}(\lambda)\bigcap GR} \mathbb{E}_{\mathbb{P}_n}\Bigg(J_{k+1}^{(n)}\Bigg(\pm, i+\xi_{m+1}, j+\hat{\xi}_{m+1},$$
$$\Lambda^{\pm}(\xi_{m+1})\Bigg)\Bigg| \sum_{m=1}^{k}\xi_m = i,\ \sum_{m=1}^{k}\hat{\xi}_m = j\Bigg)$$

where $\Lambda^{\pm}(-1) = \Lambda(-1)$, $\Lambda^{\pm}(0) = \lambda$ and $\Lambda^{\pm}(1)$ are given by (3.30)–(3.31).

For $k = 0$ we obtain two values $J_0^{(n)}(+, x/S_0)$ and $J_0^{(n)}(-, x/S_0)$. Observe that the complexity of the above dynamic programming is polynomial in M, n. For the exact value $u_n(x) = J_0^{(n)}(x/S_0)$ we have the following simple lemma.

Lemma 3.6.2. *Assume that* $\frac{x}{S_0} \in GR$. *Then*

$$J_n(x/S_0) \in [J_0^{(n)}(-, x/S_0), J_0^{(n)}(+, x/S_0)].$$

Proof. The inequality $J_0^{(n)}(-, x/S_0) \leq J_0^{(n)}(x/S_0)$ is obvious. Let us prove that $J_0^{(n)}(x/S_0) \leq J_0^{(n)}(+, x/S_0)$. Choose $\lambda \in GR$ and $\tilde{\Lambda}(-1), \tilde{\Lambda}(1) \in [0,1]$ which satisfy (3.27). Define $\Lambda(-1) := 1 \wedge \frac{[*]\tilde{\Lambda}(-1)M}{M}$ and let $\Lambda^{+}(1)$ be given by (3.31). Then it is straightforward to check that $\Lambda(-1) \geq \tilde{\Lambda}(-1)$ and $\Lambda^{+}(1) \geq \tilde{\Lambda}(1)$. Hence, by applying backward induction (on k) and the fact that $J_k^{(n)}(i, j, \lambda)$ is non-decreasing in λ we get that for any k, $J_k^{(n)}(\cdot) \leq J_k^{(n)}(+, \cdot)$ where we take the restriction of $J_k^{(n)}(\cdot)$ to $\{-k, 1-k, ..., k\} \times \{-k, 1-k, ..., k\} \times GR$. For $k = 0$, we obtain $J_0^{(n)}(x/S_0) \leq J_0^{(n)}(+, x/S_0)$ as required. $\square$

Remark 3.6.1. *By using the fact that* U *is Lipschitz continuous in the first variable, it can be shown that the difference* $J_0^{(n)}(+, x/S_0) - J_0^{(n)}(-, x/S_0)$ *is of order* $O(n/M)$. *In practice this difference goes to zero much faster (in* M*). As we will see in the following numerical results, already for* M *"close" to* n *the difference* $J_0^{(n)}(+, x/S_0) - J_0^{(n)}(-, x/S_0)$ *becomes very small.*

3.6.1 Numerical Results

In this section we implement numerically the above described procedure. In Table 3.1 and in the corresponding Figure 3.1 we compute the functions defined in (3.32). To serve as a reference we also evaluate the function $\overline{u}(x) = -\mathbb{E}_P\left[((S_T - K)^+ - x)^+\right]$, a lower bound, which corresponds to the value of spending no extra effort in reducing the shortfall.

	$J_0^{(n)}(-, x/S_0)$	$J_0^{(n)}(+, x/S_0)$	$\overline{u}_0^{(n)}(x)$
x=0	-24.5421	-24.0371	-24.6095
x=5	-18.4702	-17.7050	-21.4086
x=10	-12.3159	-11.6165	-18.2077
x=15	-7.0529	-6.3398	-16.3018
x=20	-2.7913	-2.2453	-14.3959
x=25	-0.6802	-0.4201	-12.4901
x=30	-0.0825	-0.0274	-10.5842
x=35	-0.0043	-0.0004	-8.6783
x=40	0	0	-7.1540
x=45	0	0	-6.4423
x=50	0	0	-5.7306
x=55	0	0	-5.0190
x=60	0	0	-4.3073
x=70	0	0	-2.8840
x=80	0	0	-2.0045
x=90	0	0	-1.6418
x=100	0	0	-1.2792

Table 3.1: Shortfall risk minimization for call options. Parameters used in computation: $K = 90, \underline{\sigma} = 1, \tilde{\sigma} = 5, \overline{\sigma} = 0.0001, \sigma = 0.39, \rho = -0.64, \kappa = 1.15, \theta = 0.348, \mu = 0.05, S_0 = 100, T = 1, \nu_0 = 0.09, n = 400, M = 400$

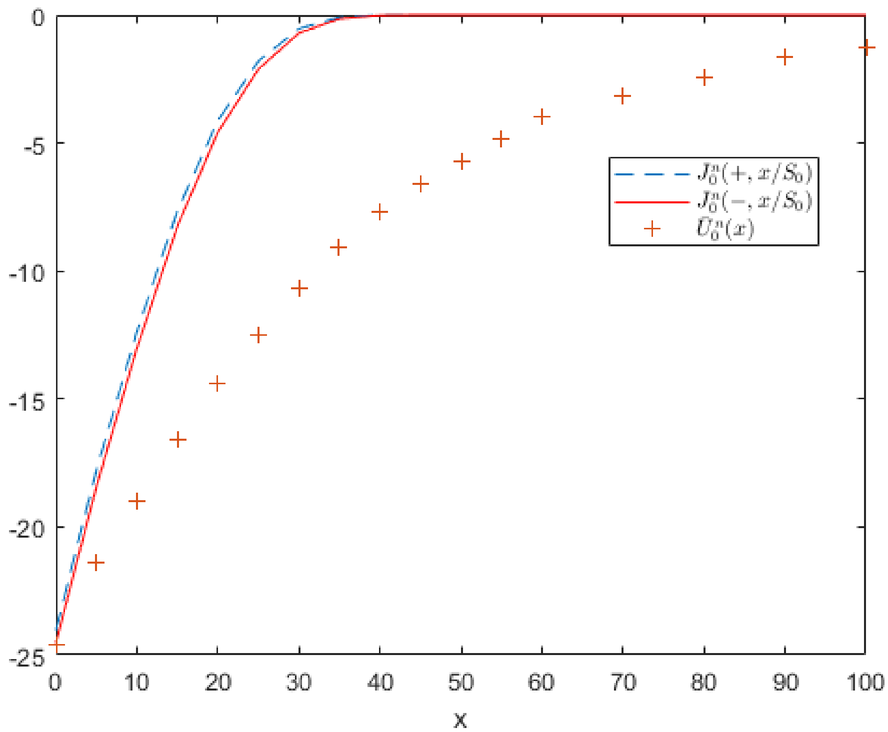

Figure 3.1: Plot of the values reported in Table 3.1.

In the next table we analyze the sensitivity of the problem to $\overline{\sigma}$. The smaller this parameter, the faster the algorithm takes. Although, Lemma 3.6.1 indicates an error bound for large $\overline{\sigma}$ (which was obtained by an application of Markov's inequality), we observe that we can in practice take $\overline{\sigma} = 1$ for our parameters.

	$\overline{\sigma} = 0.4(0.1757)$	$\overline{\sigma} = 0.6(0.8085)$	$\overline{\sigma} = 0.8(0.9939)$	$\overline{\sigma} = 1$	$\overline{\sigma} = 2$
x=0	-15.3139	-23.1077	-22.0861	-24.5421	-24.5421
x=10	-4.1129	-9.6884	-10.9334	-12.3159	-12.3159
x=20	-0.1435	-4.5287	-1.9145	-2.7913	-2.7913

Table 3.2: Variation with respect to $\overline{\sigma}$. Parameters are the same as in Table 3.1. The values in the parentheses represent $\mathbb{P}(\Theta_{\underline{\sigma},\overline{\sigma}} < T)$ rounded to 4 decimals points. We did not indicate these values when this probability is extremely close to 1.

In Table 3.3 we analyze the sensitivity of solution to the grid size of the control variable defined in (3.29). We observe, as stated in Remark 3.6.1, that the we can actually take $M = kn$, where $k < 1$. In this table, we determine the range of k we can choose. We observe that choosing n larger leads to more error reduction than choosing k larger. We have also checked this for values of $k > 1$.

	M=n/4	M=n/2	M=n
n=50	-9.2138	-6.6971	-6.6586
n=100	-5.4667	-5.4282	-5.4238
n=200	-3.7184	-3.6541	-3.6448
n=400	-2.9834	-2.8392	-2.7913
n=800	-2.6675	-2.5299	-2.4833

Table 3.3: Variation with respect to M. $x = 20$. Other parameters are the same as in Table 3.1.

Table 3.4 and the corresponding Figure 3.4 demonstrate the convergence with respect to n. We observe that the convergence rate is a power of n. We leave the rigorous demonstration of this result for future work.

	M=n/4	$\frac{J_0^n(-,x/S_0)-J_0^{n/2}(-,x/S_0)}{\lvert J_0^{n/2}(-,x/S_0)\rvert}$
n=50	-9.2138	–
n=100	-5.4667	0.4067
n=200	-3.7184	0.3198
n=400	-2.9834	0.1977
n=800	-2.6675	0.1059
n=1600	-2.6171	0.0189

Table 3.4: $x = 20$. Other parameters are the same as in Table 3.1.

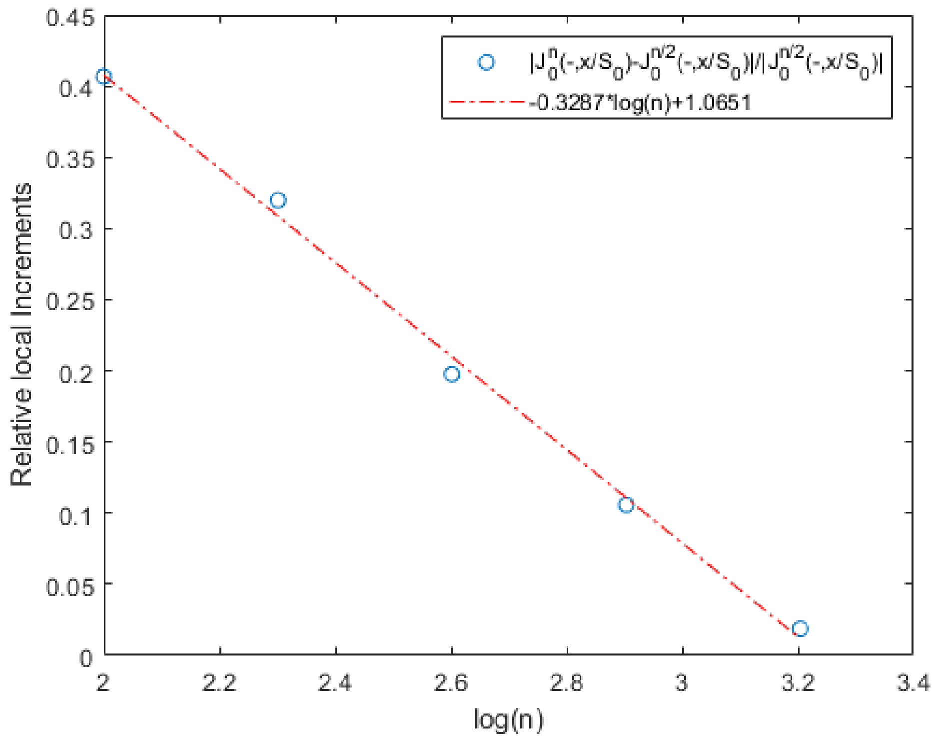

Figure 3.2: Plot of the values in Table 3.4.

CHAPTER IV

Disorder Detection with Costly Observations

In this chapter we study the Wiener disorder detection problem where each observation is associated with a positive cost. This is a foumous problem which has been studied for much of twentieth century. In our setting, a strategy is a pair consisting of a sequence of observation times and a stopping time corresponding to the declaration of disorder. We characterize the minimal cost of the disorder problem with costly observations as the unique fix-point of a certain jump operator, and we determine the optimal strategy. More details can be found in paper [13].

4.1 Problem Formulation

Let $(\Omega, \mathcal{F}, \mathbb{P}_\pi)$ be a probability space hosting a Brownian motion W and an independent random variable Θ having distribution

$$\mathbb{P}_\pi\{\Theta = 0\} = \pi, \quad \mathbb{P}_\pi\{\Theta > t\} = (1-\pi)e^{-\lambda t}, \quad t \geq 0,$$

where $\pi \in [0,1]$. We assume that the observation process $(X_t)_{t\geq 0}$ is given by

$$X_t = \alpha(t-\Theta)^+ + W_t, \tag{4.1}$$

i.e. a Brownian motion which after the random (disorder) time Θ drifts at rate α. Our objective is to detect the unknown disorder time Θ based on the observations of X_t as quickly after its occurrence as possible, but at the same time with a small proportion of false alarms. A classical Bayes' risk associated with a stopping strategy τ (where τ is a stopping time with respect to some appropriate filtration) is given by

$$\mathbb{P}_\pi(\tau < \Theta) + c\mathbb{E}_\pi[(\tau - \Theta)^+], \tag{4.2}$$

where $c > 0$ is a cost associated to the detection delay.

In the classical version of the detection problem, see [71], observations of the underlying process are costless, and a solution can be obtained by making use of the associated formulation in terms of a free-boundary problem. Subsequent literature has, among different things, focused on the case of costly observations. In [5] and [27], a version of the problem was considered in which observations of increments of the underlying process are costly, and where the cost is proportional to the length of the observation time. An alternative set-up was considered in [16], where the number of observations of the underlying process is limited.

In the current chapter, we consider a model in which observations of X are unrestricted, but where each observation is associated with an observation cost $d > 0$. We stress the fact that we assume that the controller observes values of the process X, as opposed to increments of X as in [5] and [27]. In a related work [36], the sequential hypothesis testing problem for the drift of a Wiener process was considered under the same assumption of costly observations.

Due to the discrete cost of each observation, our observation strategies will consist of finitely many samples; this motivates the following definition.

Definition 4.1.1. *A strictly increasing sequence $\hat{\tau} = \{\tau_1, \tau_2, \cdots\}$ of random variables is said to belong to $\mathcal{T}$ if τ_1 is positive and deterministic and if τ_j is measurable with respect to $\sigma(X_{\tau_1}, \cdots, X_{\tau_{j-1}}, \tau_1, \cdots, \tau_{j-1})$, $j = 2, 3, \cdots$. For a given sequence $\hat{\tau} \in \mathcal{T}$,* *let*

$$\mathcal{F}_t^{\hat{\tau}} = \sigma(X_{\tau_1}, \cdots, X_{\tau_j}, \tau_1, \cdots, \tau_j;\ \textit{where } j = \sup\{k : \tau_k \le t\}),$$

let $\mathbb{F}^{\hat{\tau}} = (\mathcal{F}_t^{\hat{\tau}})_{t\ge 0}$, *and denote by* $\mathcal{S}^{\hat{\tau}}$ *the stopping times with respect to* $\mathbb{F}^{\hat{\tau}}$.

A useful result regarding the structure of the stopping times is the following result which is presented as Proposition 2.1 in [16].

Lemma 4.1.1. *Let* $\hat{\tau} \in \mathcal{T}$, *and let* S *be an* $\mathbb{F}^{\hat{\tau}}$*-stopping time. Then for each* $j \ge 1$, *both* $S1_{\{\tau_j \le S < \tau_{j+1}\}}$ *and* $1_{\{\tau_j \le S < \tau_{j+1}\}}$ *are* $\mathbb{F}_{\tau_j}^{\hat{\tau}}$*-measurable.*

We generalize the Bayes' risk defined in (4.2) by formulating the quickest detection problem with observation costs as

$$V(\pi) = \inf_{\hat{\tau}\in\mathcal{T}} \inf_{\tau\in\mathcal{S}^{\hat{\tau}}} \left\{ \mathbb{P}_\pi(\tau < \Theta) + \mathbb{E}_\pi\left[c\,(\tau-\Theta)^+ + d\sum_{k=1}^{\infty} \mathbb{1}_{\{\tau_k \le \tau\}} \right] \right\}. \tag{4.3}$$

Here the positive constant c represents the cost of detection delay, and the positive constant d represents the cost for each observation. Note that the observer has two controls: she controls the observation sequence $\hat{\tau}$, and also needs to decide when the change happened, which is the role of τ.

Problem (4.3) can be formulated as a control problem in terms of the a posteriori probability process

$$\Pi_t^{\hat{\tau}} := \mathbb{P}_\pi(\Theta \le t \big| \mathcal{F}_t^{\hat{\tau}}) \tag{4.4}$$

as

$$V(\pi) = \inf_{\hat{\tau}\in\mathcal{T}} \inf_{\tau\in\mathcal{S}^{\hat{\tau}}} \rho^\pi(\hat{\tau}, \tau), \tag{4.5}$$

where

$$\rho^\pi(\hat{\tau}, \tau) := \mathbb{E}_\pi\left[1 - \Pi_\tau^{\hat{\tau}} + c\int_0^\tau \Pi_s^{\hat{\tau}} ds + d\sum_{k=1}^{\infty} \mathbb{1}_{\{\tau_k \le \tau\}} \right].$$

The computations are analogous to, e.g., [68, Proposition 5.8]. Observe that we can restrict ourselves to stopping times with $\mathbb{E}[\tau] < \infty$.

Remark 4.1.1. *Clearly, $V(\pi) \geq 0$. Moreover, choosing $\tau = 0$ yields $V(\pi) \leq 1 - \pi$.*

For $\pi = 1$, the a posteriori probability process $\Pi_t^{\hat{\tau}}$ is constantly equal to 1. If $\pi \in [0, 1)$, then $\Pi_t^{\hat{\tau}}$ can (see [16] and [30]) be expressed recursively as

$$\Pi_t^{\hat{\tau}} = \begin{cases} \pi & t = 0, \\ 1 - (1 - \Pi_{\tau_{k-1}}^{\hat{\tau}})e^{-\lambda(t-\tau_{k-1})} & \tau_{k-1} < t < \tau_k, \\ \frac{j(\tau_k - \tau_{k-1}, \Pi_{\tau_{k-1}}^{\hat{\tau}}, X_{\tau_k} - X_{\tau_{k-1}})}{1 + j(\tau_k - \tau_{k-1}, \Pi_{\tau_{k-1}}^{\hat{\tau}}, X_{\tau_k} - X_{\tau_{k-1}})} & t = \tau_k, \end{cases} \tag{4.6}$$

where $k \geq 1$, $\tau_0 := 0$, and

$$j(t, \pi, x) = \exp\left\{\alpha x + (\lambda - \frac{\alpha^2}{2})t\right\} \frac{\pi}{1 - \pi} + \lambda \int_0^t \exp\left\{(\lambda + \frac{\alpha x}{t})u - \frac{\alpha^2 u^2}{2t}\right\} du.$$

Thus at an observation time τ_k, the process $\Pi^{\hat{\tau}}$ jumps from

$$1 - (1 - \Pi_{\tau_{k-1}}^{\hat{\tau}})e^{-\lambda(\tau_k - \tau_{k-1})}$$

to

$$\frac{j(\tau_k - \tau_{k-1}, \Pi_{\tau_{k-1}}^{\hat{\tau}}, X_{\tau_k} - X_{\tau_{k-1}})}{1 + j(\tau_k - \tau_{k-1}, \Pi_{\tau_{k-1}}^{\hat{\tau}}, X_{\tau_k} - X_{\tau_{k-1}})}.$$

Moreover, $(t, \Pi_t^{\hat{\tau}})$ with respect to $\mathbb{F}^{\hat{\tau}}$ is a piece-wise deterministic Markov process in the sense of [28, Section 2] and therefore has the strong Markov property.

At time $t = 0$, the observer could decide that he will not be making any observations (by setting $\tau_1 = \infty$). Then $\Pi^{\hat{\tau}}$ evolves deterministically (see (4.6)), and the corresponding cost of following that strategy is thus given by

$$\begin{aligned} F(\pi) &= \inf_{t \geq 0}\left\{1 - \Pi_t^{\hat{\tau}} + c\int_0^t \Pi_s^{\hat{\tau}} ds\right\} \\ &= \inf_{t \geq 0}\left\{(1 - \pi)e^{-\lambda t} + ct - c(1 - \pi)\frac{1 - e^{-\lambda t}}{\lambda}\right\} \\ &= \begin{cases} \frac{c}{\lambda}\left(\pi + \log\frac{(\lambda + c)(1 - \pi)}{c}\right) & \pi < \frac{\lambda}{c + \lambda}; \\ 1 - \pi & \pi \geq \frac{\lambda}{c + \lambda}. \end{cases} \end{aligned}$$

Moreover, the optimizer t^* is given by

$$t^*(\pi) = \begin{cases} \frac{1}{\lambda}\log\frac{(\lambda+c)(1-\pi)}{c} & \pi < \frac{\lambda}{c+\lambda}; \\ 0 & \pi \geq \frac{\lambda}{c+\lambda}. \end{cases} \tag{4.7}$$

For a given sequence $\hat{\tau} \in \mathcal{T}$ of observations, let $\mathcal{S}_0^{\hat{\tau}} \subseteq \mathcal{S}^{\hat{\tau}}$ denote the set of $\mathbb{F}^{\hat{\tau}}$-stopping times τ such that $\mathbb{P}_\pi$-a.s. $\tau = \tau_k$ for some $k = k(\omega)$.

Proposition IV.1. *The quickest detection problem with costly observations $V(\pi)$ in (4.3) can be represented as*

$$\begin{aligned} V(\pi) = \inf_{\hat{\tau}\in\mathcal{T}} \inf_{\tau\in\mathcal{S}_0^{\hat{\tau}}} \mathbb{E}_\pi\Big[F(\Pi_\tau^{\hat{\tau}}) + c\tau - \frac{c}{\lambda}\sum_{k=0}^{\infty}(1-\Pi_{\tau_k}^{\hat{\tau}})(1-e^{-\lambda(\tau_{k+1}-\tau_k)})\mathbb{1}_{\{\tau_{k+1}\leq\tau\}} \\ + d\sum_{k=1}^{\infty}\mathbb{1}_{\{\tau_k\leq\tau\}}\Big], \end{aligned} \tag{4.8}$$

i.e. the value function is a combined optimal stopping and impulse control problem.

Proof. It follows from Lemma 4.1.1 that any stopping time $\bar{\tau} \in \mathcal{S}^{\hat{\tau}}$ can be written as $\bar{\tau} = \tau + \bar{t}$, for $\tau \in \mathcal{S}_0^{\hat{\tau}}$ and for some $\mathbb{F}_\tau^{\hat{\tau}}$-measurable random variable $\bar{t}$. Then by conditioning at τ first, optimizing over the stopping times in $\mathcal{S}^{\hat{\tau}}$ and then taking expectations we obtain

$$V(\pi) = \inf_{\hat{\tau}\in\mathcal{T}} \inf_{\tau\in\mathcal{S}_0^{\hat{\tau}}} \mathbb{E}_\pi\left[F(\Pi_\tau) + c\int_0^\tau \Pi_s^{\hat{\tau}} ds + d\sum_{k=1}^{\infty}\mathbb{1}_{\{\tau_k\leq\tau\}}\right]. \tag{4.9}$$

The rest of the proof can be done using (4.6) and partitioning the integral into integrals over $[\tau_i, \tau_{i+1})$. □

4.2 A functional characterization of the value function

In this section we study the value function V and its relation to a certain operator $\mathcal{J}$. To define the operator $\mathcal{J}$, let

$$\mathbb{F} := \{f : [0,1] \to [0,1] \text{ measurable and such that } f(\pi) \leq 1-\pi\}$$

and set

$$\mathcal{J}f(\pi) = \min\{F(\pi), \inf_{t>0} \mathcal{J}_0 f(\pi, t)\}$$

for $f \in \mathbb{F}$, where

$$\mathcal{J}_0 f(\pi, t) := d + \mathbb{E}_\pi \left[f\left(\frac{j(t, \pi, X_t)}{1 + j(t, \pi, X_t)} \right) + c(t - \Theta)^+ \right].$$

Note that

$$\mathbb{E}_\pi \left[(t - \Theta)^+ \right] = t - (1 - \pi) \frac{1 - e^{-\lambda t}}{\lambda},$$

so

$$\mathcal{J}_0 f(\pi, t) = d + \mathbb{E}_\pi \left[f\left(\frac{j(t, \pi, X_t)}{1 + j(t, \pi, X_t)} \right) + ct - c(1 - \pi) \frac{1 - e^{-\lambda t}}{\lambda} \right]. \quad (4.10)$$

Proposition IV.2. *The operator* $\mathcal{J}$

(i) is monotone: $f_1 \leq f_2 \implies \mathcal{J}f_1 \leq \mathcal{J}f_2$*;*

(ii) is concave: $\mathcal{J}(af_1 + (1-a)f_2) \geq a\mathcal{J}f_1 + (1-a)\mathcal{J}f_2$ *for* $a \in [0, 1]$*;*

(iii) satisfies $\mathcal{J}_0(\pi) = \min\{F(\pi), d\}$*;*

(iv) has at most one fixed point $f \in \mathbb{F}$ *such that* $f = \mathcal{J}f$*;*

(v) is concavity preserving: if $f \in \mathbb{F}$ *is concave, then also* $\mathcal{J}f$ *is concave.*

Proof. (i) and (iii) are immediate. For (ii), let $f_1, f_2 \in \mathbb{F}$ and let $a \in [0, 1]$. Then

$$\begin{aligned} \mathcal{J}(af_1 + (1-a)f_2) &= \min\left\{ F, \inf_{t>0} \{a\mathcal{J}_0 f_1 + (1-a)\mathcal{J}_0 f_2\} \right\} \\ &\geq \inf_{t>0} \{a \min\{F, \mathcal{J}_0 f_1\} + (1-a) \min\{F, \mathcal{J}_0 f_2\}\} \\ &\geq a\mathcal{J}f_1 + (1-a)\mathcal{J}f_2. \end{aligned}$$

For (iv) we argue as in [28, Lemma 54.21]; assume that there exist two distinct fixed points of $\mathcal{J}$, i.e. $f_1 = \mathcal{J}f_1$ and $f_2 = \mathcal{J}f_2$ for $f_1, f_2 \in \mathbb{F}$ such that $f_1(\pi) < f_2(\pi)$ (without loss of generality) for some $\pi \in [0, 1)$. Let $a_0 := \sup\{a \in [0, 1] : af_2 \leq f_1\}$, and note that $a_0 \in [0,1)$. From (iii) it follows that there exists $\kappa > 0$ such that $\kappa \mathcal{J}0 \geq 1-\pi$, $\pi \in [0,1]$, so using (i) and (ii) we get

$$f_1 = \mathcal{J}f_1 \geq \mathcal{J}(a_0 f_2) \geq a_0 \mathcal{J}f_2 + (1-a_0)\mathcal{J}0 \geq (a_0 + (1-a_0)\kappa^{-1})f_2,$$

which contradicts the definition of a_0.

For (v), first note that F is concave. Since the infimum of concave functions is again concave, it therefore follows from (4.10) that it suffices to check that

$$\mathbb{E}_\pi\left[f\left(\frac{j(t,\pi,X_t)}{1+j(t,\pi,X_t)}\right)\right]$$

is concave in π for any $t > 0$ given and fixed. To do that, define measures $\mathbb{Q}_\pi$, $\pi \in [0,1)$, on $\sigma(X_t)$ by

$$d\mathbb{Q}_\pi := \frac{e^{\lambda t}}{(1-\pi)(1+j(t,\pi,X_t))} d\mathbb{P}_\pi.$$

Then

$$\mathbb{E}_\pi\left[\frac{d\mathbb{Q}_\pi}{d\mathbb{P}_\pi}\right] = \frac{e^{\lambda t}}{1-\pi}\int_{\mathbb{R}} \frac{1}{1+j(t,\pi,y)} \mathbb{P}_\pi(X_t \in dy).$$

Denoting by φ the density of the standard normal distribution, we have

$$\begin{aligned}
\frac{\mathbb{P}_\pi(X_t \in dy)}{1-\pi} &= \frac{\pi}{1-\pi}\mathbb{P}_\pi(X_t \in dy|\Theta = 0) + \lambda\int_0^t \mathbb{P}_\pi(X_t \in dy|\Theta = s)e^{-\lambda s}ds \\
&\quad + \mathbb{P}_\pi(X_t \in dy|\Theta > t)e^{-\lambda t} \\
&= \frac{\pi}{(1-\pi)\sqrt{t}}\varphi\left(\frac{y-\alpha t}{\sqrt{t}}\right) + \frac{\lambda}{\sqrt{t}}\int_0^t e^{-\lambda s}\varphi\left(\frac{y-\alpha(t-s)}{\sqrt{t}}\right)ds \\
&\quad + \frac{e^{-\lambda t}}{\sqrt{t}}\varphi\left(\frac{y}{\sqrt{t}}\right) \\
&= e^{-\lambda t}(1+j(t,\pi,y))\varphi\left(\frac{y}{\sqrt{t}}\right).
\end{aligned}$$

Thus

$$\mathbb{E}_\pi\left[\frac{d\mathbb{Q}_\pi}{d\mathbb{P}_\pi}\right] = \frac{1}{\sqrt{t}}\int_{\mathbb{R}}\varphi\left(\frac{y}{\sqrt{t}}\right)dy = 1$$

so $\mathbb{Q}_\pi$ is a probability measure. Furthermore, the random variable X_t is $N(0,t)$-distributed under $\mathbb{Q}_\pi$; in particular, the $\mathbb{Q}_\pi$-distribution of X_t does not depend on π.

Since $j(t,\pi,x)$ is affine in $\pi/(1-\pi)$, the function

$$\pi \mapsto (1-\pi) f\left(\frac{j(t,\pi,x)}{1+j(t,\pi,x)}\right)(1+j(t,\pi,x))$$

is concave if f is concave. It thus follows from

$$\begin{aligned} & \mathbb{E}_\pi\left[f\left(\frac{j(t,\pi,X_t)}{1+j(t,\pi,X_t)}\right)\right] \\ = \quad & (1-\pi)\exp\{-\lambda t\}\mathbb{E}^{\mathbb{Q}_\pi}\left[f\left(\frac{j(t,\pi,X_t)}{1+j(t,\pi,X_t)}\right)(1+j(t,\pi,X_t))\right] \end{aligned}$$

and (4.10) that $\pi \mapsto \mathcal{J}_0 f(\pi,t)$ is concave, which completes the proof. □

Next we define a sequence $\{f_n\}_{n=0}^\infty$ of functions on $[0,1]$ by setting

$$f_0(\pi) = F(\pi), \quad f_{n+1}(\pi) = \mathcal{J} f_n(\pi), \quad n \geq 0.$$

Proposition IV.3. *For $\{f_n\}_{n=1}^\infty$ we have that*

(i) the sequence is decreasing;

(ii) each f_n is concave.

Proof. Clearly, $f_1 \leq F = f_0$, so Proposition IV.2 (i) and a straightforward induction argument give that f_n is decreasing in n. Hence the pointwise limit $f_\infty := \lim_{n\to\infty} f_n$ exists. Furthermore, since F is concave, each f_n is concave by Proposition IV.2 (v). □

Thus the pointwise limit $f_\infty := \lim_{n\to\infty} f_n$ exists. Since the pointwise limit of concave functions is concave, it follows that also f_∞ is concave.

Lemma 4.2.1. *Let $f \in \mathbb{F}$ be continuous. For fixed $\pi \in [0,1]$, the function $t \mapsto \mathcal{J}_0 f(\pi,t)$ attains its minimum for some point $t \in [0,\infty)$. Denote the first of these minimums by $t(\pi,f)$, i.e.*

$$t(\pi,f) := \inf\{t \geq 0 : \inf_s \mathcal{J}_0 f(\pi,s) = \mathcal{J}_0 f(\pi,t)\}. \tag{4.11}$$

Then $\pi \mapsto t(\pi,f)$ is measurable.

Proof. Observe that $(t,\pi) \mapsto \mathcal{J}_0 f(\pi,t)$ is a finite continuous function which approaches ∞ as $t \to \infty$. It follows that $t(\pi,f)$ is finite.

We will prove the measurability of $\pi \mapsto t(\pi,f)$ by showing that it is lower semi-continuous. Let $\pi_i \to \pi_\infty$ and let $t_i = t(\pi_i,f)$. Because $t \to ct$ is the dominating term in $t \mapsto \mathcal{J}_0 f(\pi,t)$, it is clear that the sequence $\{t_i\}_{i\in\mathbb{N}}$ is bounded. It follows that $t_\infty := \liminf t_i < \infty$; let $\{t_{i_j}\}_{j=1}^\infty$ be a subsequence such that $t_{i_j} \to t_\infty$. Then, by the Fatou lemma,

$$\mathcal{J}_0 f(\pi_\infty,t_\infty) \leq \liminf_{j\to\infty} \mathcal{J}_0 f(\pi_{i_j},t_{i_j}) = \lim_{j\to\infty} \mathcal{J}_0 f(\pi_{i_j},t_{i_j}) = \mathcal{J}_0 f(\pi_\infty,t_\infty).$$

Thus

$$t(\pi_\infty,f) \leq t_\infty = \liminf_{i\to\infty} t(\pi_i,f),$$

which establishes the desired lower semi-continuity. □

Proposition IV.4. *The function f_∞ is the unique fixed point of the operator $\mathcal{J}$.*

Proof. Since the operator $\mathcal{J}$ is monotone and $f_n \geq f_\infty$, it is clear that $f_\infty \geq \mathcal{J} f_\infty$. On the other hand,

$$f_{n+1}(\pi) = \mathcal{J} f_n(\pi) \leq \min\{F(\pi), \mathcal{J}_0 f_n(\pi, t(\pi,f_\infty))\},$$

where $t(\pi,f_\infty)$ is defined as in (4.11). Letting $n \to \infty$ and using the monotone convergence theorem we obtain that f_∞ is a fixed point. Since uniqueness is established in Proposition IV.2, this completes the proof. □

Next we introduce the problem of an agent who is allowed to make at most n observations:

$$(4.12) \qquad V_n(\pi) := \inf_{\hat{\tau}\in\mathcal{T}} \inf_{\tau\in\mathcal{S}_0^{\hat{\tau}}, \tau\le\tau_n} \rho^\pi(\hat{\tau},\tau).$$

These functions can be sequentially generated using the integral operator $\mathcal{J}$.

Proposition IV.5. *We have $V_n = f_n$, $n \geq 0$.*

Proof. First note that $V_0 = f_0 = F$. Now assume that $V_{n-1} = f_{n-1}$ for some $n \geq 1$.

Step 1: $V_n(\pi) \geq f_n(\pi)$.

For any $\hat{\tau} \in \mathcal{T}$ and $\tau \in \mathcal{S}_0^{\hat{\tau}}$ we have

$$
\begin{aligned}
(4.13) \quad & \mathbb{E}_\pi \left[F(\Pi_{\tau\wedge\tau_n}) + c\int_0^{\tau\wedge\tau_n} \Pi_s^{\hat{\tau}} ds + d\sum_{k=1}^{\infty} \mathbb{1}_{\{\tau_k\le\tau\wedge\tau_n\}} \right] \\
&= \mathbb{E}_\pi \left[\mathbb{1}_{\{\tau_1=0\}} F(\pi) \right] \\
&\quad + \mathbb{E}_\pi \left[\mathbb{1}_{\{\tau_1>0\}} \left(F(\Pi_{\tau\wedge\tau_n}) + c\int_0^{\tau\wedge\tau_n} \Pi_s^{\hat{\tau}} ds + d\sum_{k=1}^{\infty} \mathbb{1}_{\{\tau_k\le\tau\wedge\tau_n\}} \right) \right] \\
&\geq \mathbb{1}_{\{\tau_1=0\}} F(\pi) + \mathbb{1}_{\{\tau_1>0\}} \mathbb{E}_\pi \left[\left(d + c\int_0^{\tau_1} \Pi_s^{\hat{\tau}} ds + V_{n-1}(\Pi_{\tau_1}) \right) \right] \\
&= \mathbb{1}_{\{\tau_1=0\}} F(\pi) + \mathbb{1}_{\{\tau_1>0\}} \mathbb{E}_\pi \left[\left(d + c\int_0^{\tau_1} \Pi_s^{\hat{\tau}} ds + f_{n-1}(\Pi_{\tau_1}) \right) \right] \\
&\geq \mathcal{J} f_{n-1}(\pi) = f_n(\pi),
\end{aligned}
$$

where we used the fact that τ_1 is deterministic and the Markov property of $\Pi^{\hat{\tau}}$. We obtain the desired result from (4.13) by taking the infimum over strategy pairs $(\hat{\tau}, \tau)$.

Step 2: $V_n(\pi) \leq f_n(\pi)$.

We only need to prove this for the case $\mathcal{J} f_{n-1}(\pi) < F(\pi)$ (since otherwise $f_n(\pi) = \mathcal{J} f_{n-1}(\pi) = F(\pi) \geq V_n(\pi)$ already).

Note that $V_0 = F = f_0$. We will assume that the assertion holds for $n-1$ and then prove it for n. We will follow ideas used in the proof of Theorem 4.1 in [19].

Denoting $t_n := t(\pi, f_{n-1})$, let us introduce a sequence $\hat{\tau}$ of stopping times

$$\tau_1 = t_n, \quad \tau_{i+1} = \sum_k \tau_i^k \circ \theta_{t_n} \mathbb{1}_{\{\Pi^{\hat{\tau}}_{t_n} \in B_k\}}, \quad i = 1, \cdots, n-1, \tag{4.14}$$

where $(B_k)_k$ is a finite partition of $[0,1)$ by intervals and τ^k are ϵ-optimal observation times for when the process Π starts from the centre of these intervals. [1]

Since V_{n-1} is continuous, and the expected value (before optimizing) is a continuous function of the initial starting point for any strategy choice, which is due to the continuity of Π with respect to its starting point, the above sequence is a $O(\varepsilon)$ if the intervals are chosen to be fine enough.

Now we can write

$$\begin{aligned}
f_n(\pi) &= ct_n + d - \frac{c}{\lambda}(1-\pi)(1-e^{-\lambda t_n}) + \mathbb{E}_\pi[V_{n-1}(\Pi^{\hat{\tau}}_{t_n})] \\
&\geq ct_n + d - \frac{c}{\lambda}(1-\pi)(1-e^{-\lambda t_n}) - O(\epsilon) \\
&+\mathbb{E}_\pi\left[\mathbb{E}_\pi\left[\left(F(\Pi^{\hat{\tau}}_{\tau\wedge\tau_{n-1}}) + \int_0^{\tau_{n-1}\wedge\tau} \Pi^{\hat{\tau}}_s ds + d\sum_{k=1}^{\infty} \mathbb{1}_{\{\tau_k \leq \tau\wedge\tau_{n-1}\}}\right)\circ\theta_{t_n}\middle|\mathcal{F}^{\hat{\tau}}_{t_n}\right]\right] \\
&= \mathbb{E}_\pi\Big[F(\Pi^{\hat{\tau}}_{\tau\wedge\tau_n}) + \int_0^{\tau\wedge\tau_n} \Pi^{\hat{\tau}}_s ds + d\sum_{k=1}^{\infty}\mathbb{1}_{\{\tau_k\leq\tau\wedge\tau_n\}}\Big] - O(\epsilon) \\
&\geq V_n(\pi) - O(\epsilon),
\end{aligned}$$

where we used the fact that

$$c\int_0^{t_n} \Pi^{\hat{\tau}}_s ds = ct_n - \frac{c}{\lambda}(1-\pi)(1-e^{-\lambda t_n}).$$

Since $\epsilon > 0$ can be made arbitrary small, this shows that $V_n(\pi) \leq f_n(\pi)$. □

Theorem 4.2.1. *We have that $V = f_\infty$, i.e., V is the unique fixed point of $\mathcal{J}$.*

Proof. Since $V_n = f_n \to f_\infty$, it suffices to show $\lim_{n\to\infty} V_n = V$. It follows by definition that $V(\pi) \leq V_n(\pi)$ for any $n \geq 1$ and $\pi \in [0,1]$. We thus only need to

[1] θ is the shift operator in the Markov process theory, see e.g. [20]

prove that $\lim_n V_n(\pi) \le V(\pi)$. Assume that a pair $(\hat{\tau}, \tau)$ where $\hat{\tau} \in \mathcal{T}$ and $\tau \in \mathcal{S}_0^{\hat{\tau}}$ is an ϵ-optimizer for (4.9). Then

$$
\begin{aligned}
(4.15) \qquad V_n(\pi) &\le \mathbb{E}\left[F(\Pi^{\hat{\tau}}_{\tau\wedge\tau_n}) + \int_0^{\tau\wedge\tau_n} \Pi^{\hat{\tau}}_s ds + d\sum_{k=1}^{\infty} \mathbb{1}_{\{\tau_k \le \tau\wedge\tau_n\}}\right] \\
&\le \mathbb{E}\left[F(\Pi^{\hat{\tau}}_{\tau\wedge\tau_n}) + \int_0^{\tau} \Pi^{\hat{\tau}}_s ds + d\sum_{k=1}^{\infty} \mathbb{1}_{\{\tau_k \le \tau\}}\right].
\end{aligned}
$$

Note that since $\tau(\omega) = \tau_k(\omega)$ for some $k = k(\omega)$, we have $\Pi^{\hat{\tau}}_{\tau\wedge\tau_n}(w) = \Pi^{\hat{\tau}}_{\tau}(\omega)$ if $n \ge k(\omega)$. As a result, and since F is bounded and continuous, the bounded convergence theorem applied to (4.15) gives

$$\lim_{n\to\infty} V_n(\pi) \le V(\pi) + \epsilon.$$

Since $\epsilon > 0$ is arbitrary, this completes the proof. □

4.3 The optimal strategy

In this section we study the optimal strategy for the detection problem with costly observations. More precisely, we seek to determine an optimal distribution of observation times $\hat{\tau}$ and an optimal stopping time τ. The optimal strategy is determined in terms of the continuation region

$$\mathcal{C} := \{\pi \in [0,1] : V(\pi) < F(\pi)\}.$$

Note that for $\pi \in \mathcal{C}$ we have

$$V(\pi) = \inf_{t\ge 0} \mathcal{J}_0 V(\pi, t)$$

by the definition of $\mathcal{J}$. Denote by $t(\pi) := t(\pi, f_\infty) = t(\pi, V)$, and note that since $\mathcal{J}_0 V(\pi, 0) = d + V(\pi)$, we have $t(\pi) > 0$ on $\mathcal{C}$.

Moreover, define t^* by

$$t^*(\pi) = \begin{cases} t(\pi) & \text{for } \pi \in \mathcal{C} \\ \infty & \text{for } \pi \notin \mathcal{C} \end{cases}$$

Using the function t^*, we construct recursively an observation sequence $\hat{\tau}^*$ and a stopping time τ^* as follows.

Denote by $\tau_0^* = 0$ and $\Pi_0 = \pi$. For $k = 1, 2...$, define recursively

$$\tau_k^* := \tau_{k-1}^* + t^*(\Pi_{\tau_{k-1}^*})$$

and

$$\Pi_{\tau_k^*} := \frac{j(\tau_k^* - \tau_{k-1}^*, \Pi_{\tau_{k-1}^*}, X_{\tau_k^*} - X_{\tau_{k-1}^*})}{1 + j(\tau_k^* - \tau_{k-1}^*, \Pi_{\tau_{k-1}^*}, X_{\tau_k^*} - X_{\tau_{k-1}^*})}.$$

Then $\hat{\tau}^* := \{\tau_k^*\}_{k=1}^\infty \in \mathcal{T}$. Moreover, let

$$n^* := \min\{k \geq 0 : \Pi_{\tau_k^*} \notin \mathcal{C}\} = \min\{k \geq 0 : \tau_k^* = \infty\},$$

and define $\tau^* := \tau_{n^*}^*$. Then $\tau^* \in \mathcal{S}^{\hat{\tau}^*}$, and n^* is the total number of finite observation times in $\hat{\tau}^*$.

Theorem 4.3.1. *The strategy pair* $(\hat{\tau}^*, \tau^*)$ *is an optimal strategy.*

Proof. Denote by

$$V^*(\pi) = \mathbb{E}_\pi\Big[F(\Pi_{\tau^*}) + c\tau^* - \frac{c}{\lambda}\sum_{k=0}^{n^*-1}(1 - \Pi_{\tau_k^*})(1 - e^{-\lambda(\tau_{k+1}^* - \tau_k^*)}) + dn^*\Big],$$

Clearly, by the definition of V, we have $V^*(\pi) \geq V(\pi)$. It thus remains to show $V \geq V^*(\pi)$.

For $n \geq 0$, let $\tau_n' := \tau_n^* \wedge \tau^* = \tau_{n\wedge n^*}^*$.

Claim: We have

$$\begin{aligned} V(\pi) \quad &= \quad \mathbb{E}_\pi\left[V(\Pi_{\tau_n'}) + c\tau_n' - \frac{c}{\lambda}\sum_{k=0}^{n\wedge n^*-1}(1 - \Pi_{\tau_k^*})(1 - e^{-\lambda(\tau_{k+1}^* - \tau_k^*)})\right] \\ &\quad + \mathbb{E}_\pi\left[d(n \wedge n^*)\right] \\ &=: \quad RHS(n) \end{aligned} \tag{4.16}$$

for all $n \geq 0$.

To prove the claim, first note that $\tau_0' = 0$, so $V(\pi) = RHS(0)$. Furthermore, by the Markov property we have

$$\begin{aligned}
&RHS(n+1) - RHS(n) \\
&= \mathbb{E}_\pi \Big[\Big(V(\Pi_{\tau^*_{n+1}}) - V(\Pi_{\tau^*_n}) + c(\tau^*_{n+1} - \tau^*_n) \\
&\qquad - \frac{c}{\lambda}(1 - \Pi_{\tau^*_n})(1 - e^{-\lambda(\tau^*_{n+1} - \tau^*_n)}) + d \Big) \mathbb{1}_{\{n^* \geq n+1\}} \Big] \\
&= \mathbb{E}_\pi \Big[\Big(\mathbb{E}_{\Pi_{\tau^*_n}} \left[V(\Pi_{\tau^*_1}) + c\tau^*_1 \right] - V(\Pi_{\tau^*_n}) \\
&\qquad - \frac{c}{\lambda}(1 - \Pi_{\tau^*_n}) \mathbb{E}_{\Pi_{\tau^*_n}} \left[1 - e^{-\lambda \tau^*_1} \right] + d \Big) \mathbb{1}_{\{n^* > n\}} \Big] \\
&= 0,
\end{aligned}$$

which shows that (4.16) holds for all $n \geq 0$.

Note that it follows from (4.16) that $n^* < \infty$ a.s. (since otherwise the term $\mathbb{E}_\pi[d(n \wedge n^*)]$ would explode as $n \to \infty$). Therefore, letting $n \to \infty$ in (4.16), using bounded convergence and monotone convergence, we find that

$$\begin{aligned}
V(\pi) &= \mathbb{E}_\pi \left[V(\Pi_{\tau^*}) + c\tau^* - \frac{c}{\lambda} \sum_{k=0}^{n^*-1} (1 - \Pi_{\tau^*_k})(1 - e^{-\lambda(\tau^*_{k+1} - \tau^*_k)}) + dn^* \right] \\
&= \mathbb{E}_\pi \left[F(\Pi_{\tau^*}) + c\tau^* - \frac{c}{\lambda} \sum_{k=0}^{n^*-1} (1 - \Pi_{\tau^*_k})(1 - e^{-\lambda(\tau^*_{k+1} - \tau^*_k)}) + dn^* \right] \\
&= V^*(\pi),
\end{aligned}$$

which completes the proof. □

4.4 Numerical Examples

In Figure 4.1, we illustrate Proposition IV.3. We use the same parameters that were used for Figure 2 in [16], where $d = 0$.

Clearly, the value functions V_n increase in the cost parameters. Figure 4.2 displays the value functions $V_1, ..., V_{10}$ for the same parameters as in Figure 4.1 but for a larger cost c. Similarly, the sensitivity with respect to the observation cost parameter d

is pictured in Figure 4.3. x In Figure 4.4 we compute the function t defined in (4.11), when f in the definition is replaced by V_n, for various values of n. While it appears that $t(\pi, V_n)$ is decreasing in n (the more observation rights one has, the more inclined one is to make early observations) and decreasing in π, we have not been able to prove these monotonicities.

Finally, in Figure 4.5 we determine $\pi^*(n) = \inf\{\pi : t^*(\pi, V_n) = \infty\}$. Our observations consistently indicate that the continuation region for taking observations is an interval of the form $[0, \pi^*(n))$; also here, an analytical proof of this remains to be found.

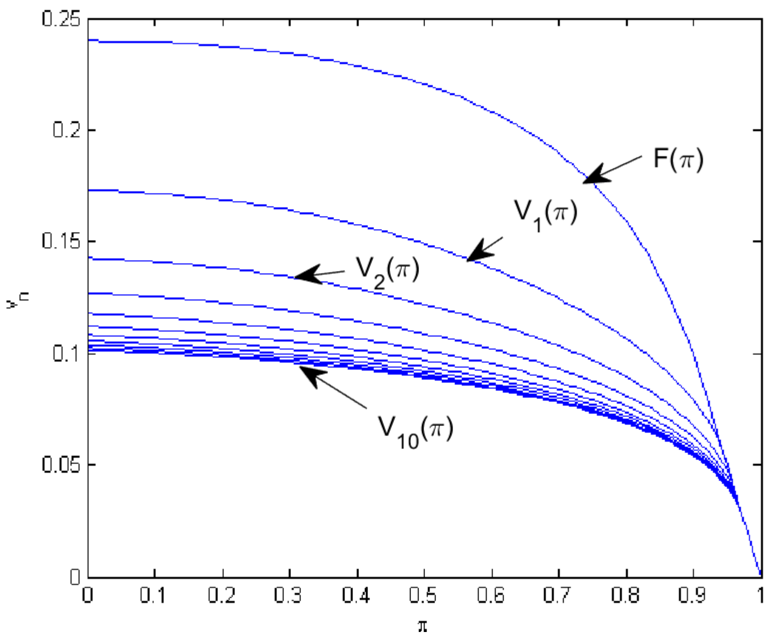

Figure 4.1: $c = 0.01, \lambda = 0.1, \alpha = 1, d = 0.001, n = 0, 1, \cdots, 10.$

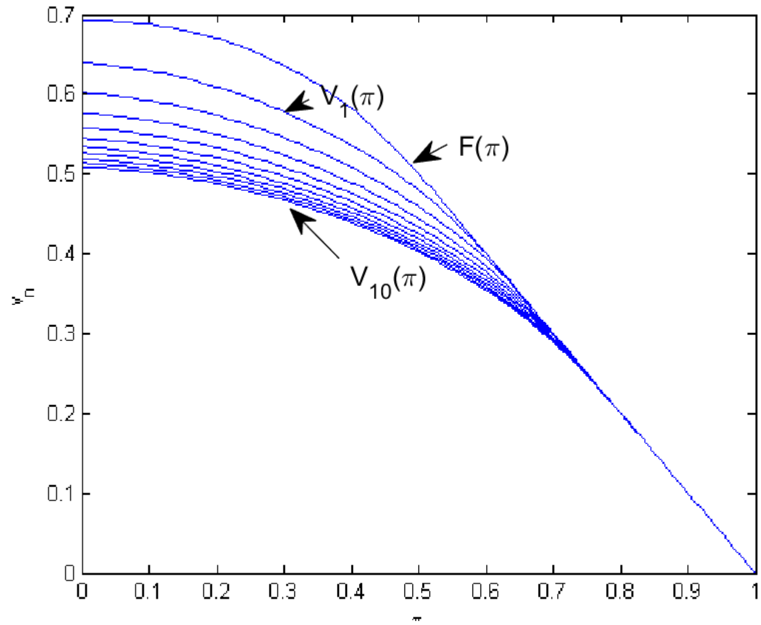

Figure 4.2: $c = 0.1, \lambda = 0.1, \alpha = 1, d = 0.001, n = 0, 1, \cdots, 10.$

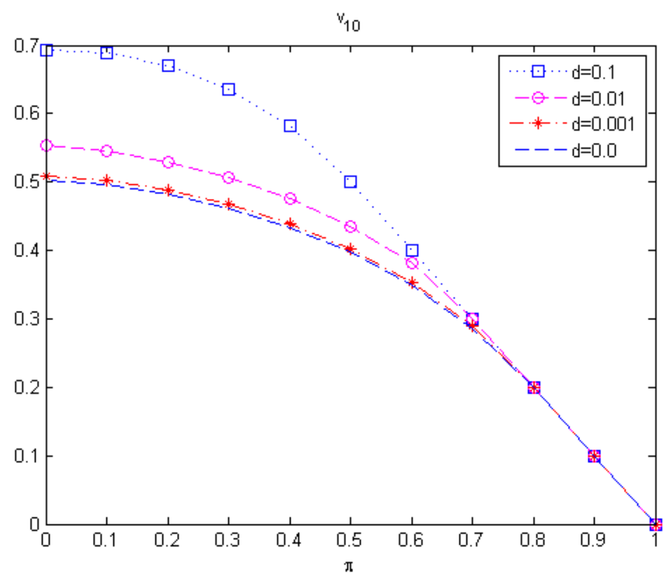

Figure 4.3: $c = 0.1, \lambda = 0.1, \alpha = 1.$

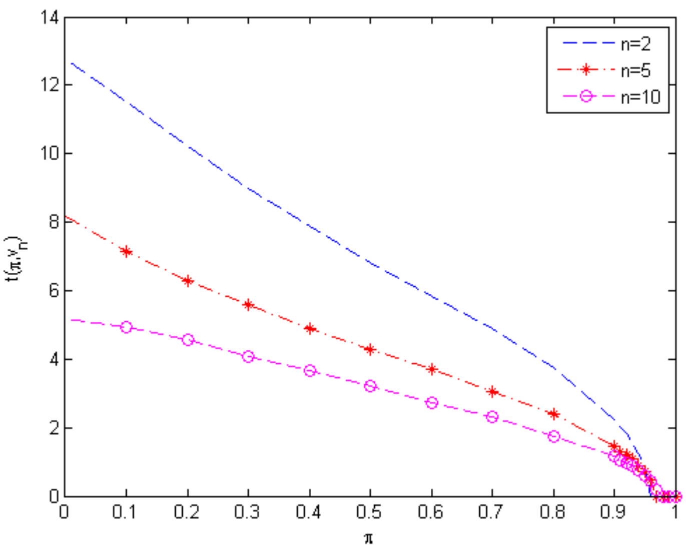

Figure 4.4: $c = 0.01, \lambda = 0.1, \alpha = 1, d = 0.001$.

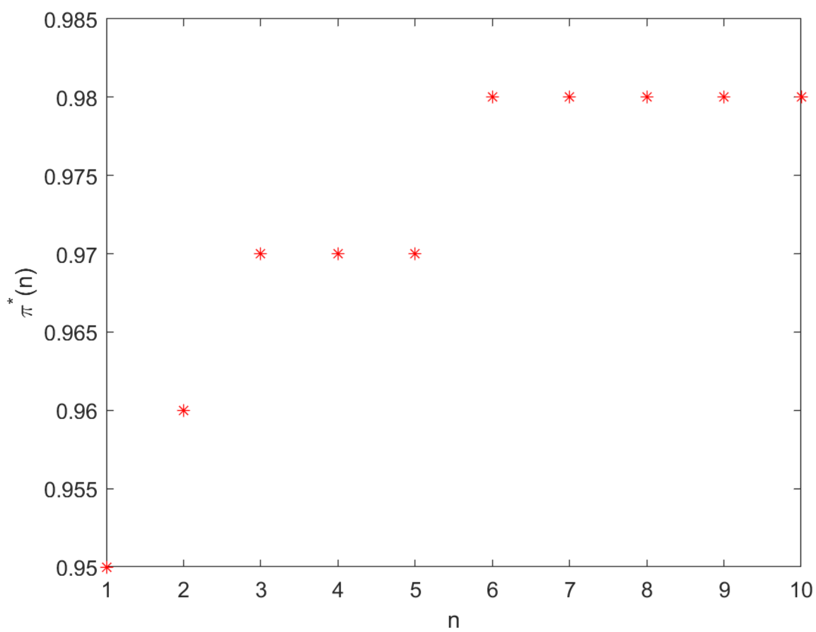

Figure 4.5: $c = 0.01, \lambda = 0.1, \alpha = 1, d = 0.001$.

BIBLIOGRAPHY

BIBLIOGRAPHY

[1] S. Ankirchner, T. Kruse, and M. Urusov. A functional limit theorem for irregular SDEs. *To appear in Annales de l'Institut Henri Poincaré - Probabilités et Statistiques.*

[2] Stefan Ankirchner, Thomas Kruse, and Mikhail Urusov. Numerical approximation of irregular SDEs via Skorokhod embeddings. *J. Math. Anal. Appl.*, 440(2):692–715, 2016.

[3] Sø ren Asmussen and Peter W. Glynn. *Stochastic simulation: algorithms and analysis*, volume 57 of *Stochastic Modelling and Applied Probability.* Springer, New York, 2007.

[4] D. Baños, T. Meyer-Brandis, F. Proske, and S. Duedahl. Computing deltas without derivatives. *Finance Stoch.*, 21(2):509–549, 2017.

[5] D. W. Balmer. On a quickest detection problem with costly information. *J. Appl. Probability*, 12:87–97, 1975.

[6] Peter Bank and Dietmar Baum. Hedging and portfolio optimization in financial markets with a large trader. *Math. Finance*, 14(1):1–18, 2004.

[7] Nicole Bäuerle and Erhan Bayraktar. A note on applications of stochastic ordering to control problems in insurance and finance. *Stochastics*, 86(2):330–340, 2014.

[8] Erhan Bayraktar. On the perpetual American put options for level dependent volatility models with jumps. *Quantitative Finance*, 11(3):335–341, 2011.

[9] Erhan Bayraktar, Savas Dayanik, and Ioannis Karatzas. Adaptive poisson disorder problem. *The Annals of Applied Probability*, 16(3):1190–1261, Aug 2006.

[10] Erhan Bayraktar, Yan Dolinsky, and Jia Guo. Recombining tree approximations for optimal stopping for diffusions. *SIAM Journal on Financial Mathematics*, 9(2):602–633, 2018.

[11] Erhan Bayraktar, Yan Dolinsky, and Jia Guo. Continuity of utility maximization under weak convergence. *Mathematics and Financial Economics*, 14(4):725–757, 2020.

[12] Erhan Bayraktar, Leonid Dolinskyi, and Yan Dolinsky. Extended weak convergence and utility maximisation with proportional transaction costs. *Finance and Stochastics*, 24(4):1013–1034, 2020.

[13] Erhan Bayraktar, Erik Ekstrom, and Jia Guo. Disorder detection with costly observations. *arXiv preprint arXiv:2002.00930*, 2020.

[14] Erhan Bayraktar and Arash Fahim. A stochastic approximation for fully nonlinear free boundary parabolic problems. *Numer. Methods Partial Differential Equations*, 30(3):902–929, 2014.

[15] Erhan Bayraktar and Ross Kravitz. Stability of exponential utility maximization with respect to market perturbations. *Stochastic Process. Appl.*, 123(5):1671–1690, 2013.

[16] Erhan Bayraktar and Ross Kravitz. Quickest detection with discretely controlled observations. *Sequential Anal.*, 34(1):77–133, 2015.

[17] Sara Biagini and Marco Frittelli. A unified framework for utility maximization problems: an Orlicz space approach. *Ann. Appl. Probab.*, 18(3):929–966, 2008.

[18] Patrick Billingsley. *Convergence of probability measures*. Wiley Series in Probability and Statistics: Probability and Statistics. John Wiley & Sons, Inc., New York, second edition, 1999. A Wiley-Interscience Publication.

[19] Bruno Bouchard and Nizar Touzi. Weak dynamic programming principle for viscosity solutions. *SIAM J. Control Optim.*, 49(3):948–962, 2011.

[20] Erhan Çınlar. *Probability and stochastics*, volume 261 of *Graduate Texts in Mathematics*. Springer, New York, 2011.

[21] Dean S. Clark. Short proof of a discrete Gronwall inequality. *Discrete Appl. Math.*, 16(3):279–281, 1987.

[22] Andrea Collevecchio, Kais Hamza, and Meng Shi. Bootstrap random walks. *Stochastic Process. Appl.*, 126(6):1744–1760, 2016.

[23] J. C. Cox. The Constant Elasticity of Variance Option Pricing model. *Journal of Portfolio Management*, 22:15–17, 1996.

[24] J.C. Cox, J.E. Ingersoll, and S.A. Ross. A theory of the trerm structure of interest rates. *Econometrica*, 53:385–407, 1985.

[25] John C. Cox, Jonathan E. Ingersoll, Jr., and Stephen A. Ross. A theory of the term structure of interest rates. *Econometrica*, 53(2):385–407, 1985.

[26] Jakša Cvitanić, Huyên Pham, and Nizar Touzi. Super-replication in stochastic volatility models with portfolio constraints. *Journal of Applied Probability*, 36:523–545, 1999.

[27] Robert C. Dalang and Albert N. Shiryaev. A quickest detection problem with an observation cost. *Ann. Appl. Probab.*, 25(3):1475–1512, 2015.

[28] M. H. A. Davis. *Markov models and optimization*, volume 49 of *Monographs on Statistics and Applied Probability*. Chapman & Hall, London, 1993.

[29] Dmitry Davydov and Vadim Linetsky. Pricing and Hedging Path-Dependent Options Under the CEV Process. *Management Science*, 47(7):949–965, 2001.

[30] Savas Dayanik. Wiener disorder problem with observations at fixed discrete time epochs. *Math. Oper. Res.*, 35(4):756–785, 2010.

[31] Freddy Delbaen and Walter Schachermayer. A general version of the fundamental theorem of asset pricing. *Math. Ann.*, 300(3):463–520, 1994.

[32] Y. Dolinsky and K. Kifer. Binomial approximations for barrier options of Israeli style. *Advances in Dynamic Games*, 11:447–469, 2011.

[33] Yan Dolinsky and Ariel Neufeld. Super-replication in fully incomplete markets. *Math. Finance*, 28(2):483–515, 2018.

[34] R. M. Dudley. Distances of probability measures and random variables. *Ann. Math. Statist*, 39:1563–1572, 1968.

[35] Darrell Duffie and Philip Protter. From discrete– to continuous–time finance: Weak convergence of the financial gain process. *Math. Finance*, 2:1–15, 1992.

[36] Hannah Dyrssen and Erik Ekström. Sequential testing of a Wiener process with costly observations. *Sequential Anal.*, 37(1):47–58, 2018.

[37] Stewart N. Ethier and Thomas G. Kurtz. *Markov processes*. Wiley Series in Probability and Mathematical Statistics: Probability and Mathematical Statistics. John Wiley & Sons, Inc., New York, 1986. Characterization and convergence.

[38] Rüdiger Frey and Carlos A. Sin. Bounds on European option prices under stochastic volatility. *Math. Finance*, 9(2):97–116, 1999.

[39] Paul Glasserman. *Monte Carlo methods in financial engineering*, volume 53 of *Applications of Mathematics (New York)*. Springer-Verlag, New York, 2004. Stochastic Modelling and Applied Probability.

[40] Anja Göing-Jaeschke and Marc Yor. A survey and some generalizations of Bessel processes. *Bernoulli*, 9(2):313–349, 2003.

[41] Anja Göing-Jaeschke and Marc Yor. A survey and some generalizations of Bessel processes. *Bernoulli*, 9(2):313–349, 2003.

[42] S. E. Graversen and G. Peskir. Maximal inequalities for Bessel processes. *J. Inequal. Appl.*, 2(2):99–119, 1998.

[43] Urs Gruber and Martin Schweizer. A diffusion limit for generalized correlated random walks. *J. Appl. Probab.*, 43(1):60–73, 2006.

[44] Steven L. Heston. A closed-form solution for options with stochastic volatility with applications to bond and currency options. *Review of Financial Studies*, 6:327–343, 1993.

[45] Ying Hu, Peter Imkeller, and Matthias Müller. Utility maximization in incomplete markets. *Ann. Appl. Probab.*, 15(3):1691–1712, 2005.

[46] Friedrich Hubalek and Walter Schachermayer. When does convergence of asset price processes imply convergence of option prices? *Math. Finance*, 8(4):385–403, 1998.

[47] J. Hull and A. White. Pricing of options on assets with stochastic volatility. *Journal of Finance*, 42:281–300, 1987.

[48] Ioannis Karatzas, John P. Lehoczky, Steven E. Shreve, and Gan-Lin Xu. Martingale and duality methods for utility maximization in an incomplete market. *SIAM J. Control Optim.*, 29(3):702–730, 1991.

[49] Ioannis Karatzas and Steven E. Shreve. *Brownian motion and stochastic calculus*, volume 113 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, 1988.

[50] K. Kifer. Game Options. *Finance and Stochastics*, 4:443–463, 2000.

[51] K. Kifer. Error Estimates for Binomial Approximations of Game Options. *Annals of Applied Probability*, 18:1737–1770, 2006.

[52] Y. Kifer. Dynkin games and Israeli options. *ISRN Probability and Statistics*, 2013.

[53] Y. Kitapbayev and J. Detemple. On American VIX options under the generalized 3/2 and 1/2 models. *ArXiv e-prints*, June 2016.

[54] D. Kramkov and W. Schachermayer. The asymptotic elasticity of utility functions and optimal investment in incomplete markets. *Ann. Appl. Probab.*, 9(3):904–950, 1999.

[55] D. O. Kramkov. Optional decomposition of supermartingales and hedging contingent claims in incomplete security markets. *Probab. Theory Related Fields*, 105(4):459–479, 1996.

[56] D. Lamberton and L. C. G. Rogers. Optimal stopping and embedding. *J. Appl. Probab.*, 37(4):1143–1148, 2000.

[57] Kasper Larsen, Oleksii Mostovyi, and Gordan Žitković. An expansion in the model space in the context of utility maximization. *Finance Stoch.*, 22(2):297–326, 2018.

[58] Kasper Larsen and Gordan Žitković. Stability of utility-maximization in incomplete markets. *Stochastic Process. Appl.*, 117(11):1642–1662, 2007.

[59] Francis A. Longstaff and Eduardo S. Schwartz. Valuing American Options by Simulation: A Simple Least-Squares approach. *The Review of Financial Studies*, 14(1):113, 2001.

[60] R.C. Merton. Lifetime portoflio selection under uncertainty: The continuous time case. *Rev. Econ. Stat*, 51:247–257, 1969.

[61] R.C. Merton. Optimum consumption and portfolio rules in a contiunous-time model. *J. Econ. Th*, 3:373–413, 1971.

[62] O. Mostovyi. Asymptotic analysis of the expected utility maximization problem with respect to perturbations of the numéraire. *preprint, arXiv: 1805.1142*, 2018.

[63] O. Mostovyi and M. Sîrbu. Sensitivity analysis of the utility maximization problem with respect to model perturbations. *preprint, arXiv: 1705.08291*, 2018.

[64] Sabrina Mulinacci. The efficient hedging problem for American options. *Finance Stoch.*, 15(2):365–397, 2011.

[65] Ariel Neufeld. Buy-and-hold property for fully incomplete markets when super-replicating markovian claims. *International Journal of Theoretical and Applied Finance*, 21(08):1850051, 2018.

[66] Yoshio Ohtsubo. Optimal stopping in sequential games with or without a constraint of always terminating. *Math. Oper. Res.*, 11(4):591–607, 1986.

[67] Goran Peskir and Albert Shiryaev. *Optimal stopping and free-boundary problems*. Lectures in Mathematics ETH Zürich. Birkhäuser Verlag, Basel, 2006.

[68] H. Vincent Poor and Olympia Hadjiliadis. *Quickest detection*. Cambridge University Press, Cambridge, 2009.

[69] Miklós Rásonyi and Lukasz Stettner. On utility maximization in discrete-time financial market models. *Ann. Appl. Probab.*, 15(2):1367–1395, 2005.

[70] Christian Reichlin. Behavioral portfolio selection: asymptotics and stability along a sequence of models. *Math. Finance*, 26(1):51–85, 2016.

[71] A. N. Shiryaev. Two problems of sequential analysis. *Cybernetics*, 3(2):63–69 (1969), 1967.

[72] Carlos A. Sin. Complications with stochastic volatility models. *Adv. in Appl. Probab.*, 30(1):256–268, 1998.

[73] Ward Whitt. Proofs of the martingale FCLT. *Probab. Surv.*, 4:268–302, 2007.

[74] Paul Wilmott, Sam Howison, and Jeff Dewynne. *The mathematics of financial derivatives*. Cambridge University Press, Cambridge, 1995. A student introduction.

[75] Hoi Ying Wong and Jing Zhao. An artificial boundary method for American option pricing under the CEV model. *SIAM J. Numer. Anal.*, 46(4):2183–2209, 2008.

[76] Hoi Ying Wong and Jing Zhao. Valuing American options under the CEV model by Laplace-Carson transforms. *Oper. Res. Lett.*, 38(5):474–481, 2010.