Problems in Optimal Stopping and Control

by

Ross Daniel Kravitz

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
(Mathematics)
in The University of Michigan
2013

Doctoral Committee:

Associate Professor Erhan Bayraktar, Chair
Professor Joseph G. Conlon
Professor Robert W. Keener
Assistant Professor Sergey Nadtochiy
Professor Virginia R. Young

To Mom, Dad, and Anni.

ACKNOWLEDGEMENTS

I would like to thank the members of my committee for the effort they have put in while reviewing my thesis. Special thanks go to my second reader, Sergey Nadtochiy, for his careful reading of my thesis. Above all, I thank my adviser, Erhan Bayraktar, for his assistance and guidance throughout the course of my studies. I wish to thank the Department of Defense for supporting me through the NDSEG fellowship, and the National Science Foundation for supporting me through grant DMS-1118673.

CONTENTS

LIST OF FIGURES

LIST OF TABLES

CHAPTER I

Introduction and Background

In this thesis, we will study three separate problems, all dealing with different aspects of the optimal stopping and control of stochastic processes. We give general introductions and background information here, and refer the reader to the individual sections for more detailed introductions to the subject matter.

The first problem belongs to theoretical mathematical finance, and takes up the question of how stable exponential utility maximization is with respect to market perturbations. The problem of utility maximization in mathematical finance has a long history, dating from the work of Merton in [37]. In the abstract, the problem can be stated as follows: an agent has initial wealth x_0, and available to him is a set $\mathcal{X}(x_0)$ of random variables, which model the possible terminal wealths which he may attain through investing in some risky assets. Typically, each $X_T \in \mathcal{X}(x_0)$ has the form

$$X_T = x_0 + \int_0^T H_u dS_u, \tag{1.1}$$

which models the gains from trading on the asset S with the trading strategy H. S is a stochastic process, and so the integral in (1.1) is an Itô Integral. The agent's goal is to choose $\widehat{X}_T \in \mathcal{X}(x_0)$ so that

$$E\left[U\left(\widehat{X}_T\right)\right] = \sup_{X_T \in \mathcal{X}(x_0)} E\left[U\left(X_T\right)\right]. \tag{1.2}$$

Here $U : \mathbb{R} \to \mathbb{R}$ is a function which quantifies the agent's attitude towards risk. It is always increasing and concave, so that the agent always prefers more wealth to less, and prefers a deterministic payoff to a random one if their expected payoffs are the same.

As in any optimization problem, the first questions to ask concern existence and uniqueness of an optimizer in (1.2). In the papers [37], [29], [34], [50], these questions are answered in the affirmative for all reasonable market models and utility functions. The next question one might ask concerns the stability of the problem: if the inputs, or parameters of the optimization problem change slightly, then will the output of the problem, the maximal expected utility, also change by a small amount? There are many questions to ask in this direction, and stability properties with respect to time horizon, information structure, and utility function have been studied in [57], [31], and [56]. Here we study a particularly natural question in this vein, namely the stability with respect to perturbations in the market. More precisely, suppose that there exists a sequence of markets, parametrized by their minimal martingales Z^n. When $Z^n_T \to Z^\infty_T$, we seek sufficient conditions ensuring that for all $x \in Dom(U)$,

$$\underbrace{u^n(x)}_{\sup_{X_T \in \mathcal{X}^n(x)} E[U(X_T)]} \quad \to \quad \underbrace{u^\infty(x).}_{\sup_{X_T \in \mathcal{X}^\infty(x)} E[U(X_T)]} \tag{1.3}$$

This problem was first studied in [36] for utility functions on $\mathbb{R}_+$, like power utilities. There, a condition known as V-compactness was shown to be a sufficient condition for stability. Let V be the convex dual of U:

$$V(y) \triangleq \sup_{x \in Dom(U)} U(x) - xy.$$

The V-compactness hypothesis states that

$$\{V(Z^n_T) : n \geq 0\} \text{ is uniformly integrable.}$$

Here, we attempt to extend (1.3) to utility functions on $\mathbb{R}$, working with the typical example of exponential utility. We show that with some mild strengthening of V-compactness, we can ensure stability in this setting with respect to market perturbations. Our arguments make heavy use of the *bmo*-theory of martingales, which has a natural role in implying the existence of near optimal wealth processes with nice regularity properties. To prove our main Theorem II.5, our basic strategy is to bootstrap from the continuity results of [36] by establishing conditions under which the utility of the optimal wealth process can be approximated, uniformly over markets, by the utility of wealth processes which are bounded from below. This allows us to build stability results for utility functions on $\mathbb{R}$ from their analogs for utilities on $\mathbb{R}_+$, as in [36].

The second problem in this thesis comes from mathematical statistics. It involves an extension of the classical sequential analysis problem, found in [55], from a single channel of statistical information to multiple independent channels, subject to the restriction that only one channel can be observed at a time. In the most classical formulation of the sequential analysis problem, an agent in discrete time monitors a channel of information, represented by a sequence of independent random variables

$$X_1, X_2, \ldots.$$

The agent must decide between two possible statistical hypotheses about the X_i: either

$$H_0 : X_i \sim Q_0 \text{ holds for each } i \qquad \text{or} \qquad H_1 : X_i \sim Q_1 \text{ holds for each } i,$$

where Q_0 and Q_1 represent two distinct, but equivalent distributions. In a typical example, Q_0 may correspond to a $N(0,1)$ random variable, and Q_1 to a $N(1,1)$ random variable. The sequential analysis problem is usually formulated in terms of the posterior probability,

$$\pi_t \triangleq P\left(H_1 | \mathcal{F}_t\right).$$

Here π_t represents the posterior likelihood that H_1 is satisfied, given the observed values of $\{X_i : i \leq t\}$. The technical device which makes this comparison possible is Bayes' Theorem, which gives a formula for updating the posterior after each observation:

$$\pi_{n+1} = \frac{\pi_n q_1(X_{n+1})}{\pi_n q_1(X_{n+1}) + (1-\pi_n) q_0(X_{n+1})},$$

where q_0 and q_1 are the respective density functions of the X_i's under Q_0 and Q_1. Thanks to this reformulation, the sequential analysis problem can be restated as an optimal stopping problem on $\{\pi_n\}_{n\geq 0}$. When π_n is sufficiently large, we should accept H_1, and when it is sufficiently small, we should accept H_0. The exact optimal thresholds for accepting either hypothesis can be computed using optimal stopping theory.

In continuous time, suppose that our agent now observes a standard one-dimensional Brownian Motion, and must decide whether or not it has a known non-zero drift. In this case, the posterior probability π_t of non-zero drift follows the stochastic differential equation (SDE)

$$d\pi_t = \pi_t(1-\pi_t)dW_t, \pi_0 = \widehat{\pi}, \tag{1.4}$$

and the sequential analysis problem becomes one of optimal stopping on (1.4).

The formulations above both correspond to a single channel. Suppose instead that there is an infinite number of channels, and that our objective is only to find any channel, as soon as possible, which satisfies H_1. From a practical viewpoint, this formulation captures many physical problems well. Typical examples include screening for an open frequency on which to broadcast, searching for extraterrestrial radio signals in different sectors of space, or determining the source of a chemical or biological attack after it is known that such an attack has occurred.

This multi-channel search problem was studied in discrete time in [35]. Here, we tackle the continuous time version, finding analytical expressions for the value function and optimal threshold, which quantifies the confidence level one should attain before accepting H_1 for the channel under inspection. These expressions can be evaluated efficiently, and may lead to efficient implementations as approximate solutions in discrete time.

The end goal of our analysis is an SDE with a reflecting boundary at $\widehat{\pi} \in (0,1)$:

$$d\pi_t^r = \pi_t^r(1-\pi_t^r)dW_t + A_t, \pi_0^r = \widehat{\pi}, \tag{1.5}$$

where A_t is an increasing process, starting from zero, which satisfies $\int_0^\infty (\pi_s^r - \widehat{\pi})\, dA_s = 0$. Here, the reflecting boundary at $\widehat{\pi}$ captures the ability of the agent to switch to a new channel at any time when it seems H_1 is unlikely, which corresponds to replacing a current posterior value $\pi_t < \widehat{\pi}$ with $\widehat{\pi}$, the prior probability of H_1 in each channel. The agent's objective is to find a stopping time τ *and* observation strategy Φ (determining which channel will be viewed at a given moment in time) which minimizes some linear combination of

$$E\left[1-\pi_\tau^\Phi\right] \text{ and } E\left[\tau\right],$$

the former modeling the probability of false alarm, and the latter capturing the amount of time spent in observation. The choice of Φ behaves like an impulse control on the process π^Φ, but the filtration of observed information additionally depends on the choice of Φ. After we establish the equivalence of our search problem to an optimal stopping problem on (1.5), we use a verification theorem to deduce the explicit form of our value function, (3.23).

The last problem in this thesis also comes from mathematical statistics. Instead of determining which of two statistical hypotheses a channel satisfies, we are confronted with a channel whose signal's statistical properties change abruptly at an unknown disorder time. From the observed signal, we must sound the alarm as soon as possible after the disorder starts, while avoiding penalties for sounding the alarm prematurely. Furthermore, and this is the aspect of the problem which we investigate, information about the channel is scarce, in the sense that we have a finite number of observation rights, which we must adaptively choose how to allocate.

Problems in this field, known as quickest detection or online change-point analysis, have been studied for much of the twentieth century. The mathematical theory of quickest detection begins with [52], who studied a Bayesian formulation of the problem in discrete time, in which the disorder time possesses some prior probability distribution. To describe this original problem more precisely, we suppose that on some probability space we have a random variable Θ taking nonnegative integer values, modeling the disorder time, and a sequence of independent random variables $Z_1, Z_2, \ldots$ modeling the observations. Conditionally on the set $\{\Theta = i\}$, the random variables $Z_1, Z_2, \ldots, Z_{i-1}$ have the distribution P_0, and the random variables $Z_i, Z_{i+1}, \ldots$ have the distribution P_1, where P_0 and P_1 are distinct but equivalent probabilities. The random variable Θ is supposed to have a geometric distribution. With probability π, $\Theta = 0$, and $P(\Theta = n) = (1-\pi)(1-p)^{n-1}p$, for some constant p between zero and one.

Let $\mathbb{F}$ be the filtration generated by $Z_1, Z_2, \ldots$. For a $\mathbb{F}$-stopping time τ, the risk is

$$\rho^{\pi}(\tau) \triangleq P^{\pi}(\tau < \Theta) + cE\left[(\tau - \Theta)^{+}\right].$$

Here, $P^{\pi}(\tau < \Theta)$ represents the probability of false alarm, $E\left[(\tau - \Theta)^{+}\right]$ is the expected delay time, and c is a positive constant which weighs the relative importance of these two terms. The goal in the quickest detection problem is to find a stopping time τ which minimizes the collective risk $\rho^{\pi}(\tau)$.

The first formulation of a quickest detection problem in continuous time is also found in [52]. Here, one observes a process X with $dX_t = dB_t + \alpha 1_{\{t \geq \Theta\}} dt$, where B_t is a standard Brownian motion, and α is constant. Θ is a random variable which is zero with probability π, and with probability $1 - \pi$ it is exponentially distributed with parameter λ. Both B_t and Θ are not directly observable. As in the discrete-time formulation, the goal is to estimate the value of Θ as a result of observing X. Our objective is to minimize a weighted average of the probability of false alarm and the detection delay time. The set over which we optimize is the set of stopping times for the filtration generated by X.

As in the sequential analysis problem, quickest detection is often formulated as an optimal stopping problem. In place of the posterior process, we use the closely related odds process

$$\Phi_t \triangleq \frac{P\left(\Theta \leq t | \mathcal{F}_t\right)}{P\left(\Theta > t | \mathcal{F}_t\right)}.$$

In the classical, full information setting of [52], Φ_t satisfies the SDE

$$d\Phi_t^c = \lambda\left(1 + \Phi_t^c\right) dt + \alpha \Phi_t^c dW_t.$$

In contrast, we consider here a model in which the channel X is not observed continuously, and instead we must adaptively choose n discrete points in time when X should be observed. For modeling purposes, this formulation is a natural one. For example, if a battery-powered sensor array is placed in a remote location, we must conserve energy by taking observations infrequently, so that the sensor can operate for an extended amount of time. In general, it is often the case that observation rights are in finite supply because of constraints on energy, funding, or time.

In a discrete observation model, new information comes in only when an observation is made, and the posterior process should reflect this. We refer the reader to (4.1) for the exact definition, but the basic structure of the posterior process is one of deterministic evolution between observations, along with random jumps induced by observations. This piece-wise deterministic structure allows for a recursive, jump-operator based approach to the problem. Essentially, we may compute the value function for n observations, v_n, by a deterministic optimization involving the value function for $n-1$ observations. v_{n-1}. In one model which we consider, this takes the form

$$v_n(\phi) = \inf_{t\geq 0} \int_0^t e^{-\lambda s} \left(\Phi_s - \frac{\lambda}{c} \right) ds + e^{-\lambda t} K v_{n-1}(t, \phi), \tag{1.6}$$

where K is an operator encoding the random result of an observation.

In contrast to [15], we consider a model in which observation times are chosen adaptively, as opposed to deterministically, which may realize large efficiency gains. We consider two models of observation rights, one in which all n observations are immediately available, and one in which these n rights arrive sporadically from an independent Poisson process. Each problem requires a different jump-operator framework to solve. In particular, (1.6) corresponds to the "Lump Sum" observation problem, while the "Stochastic Arrival Rate" problem involves three separate jump operators; this is necessary because the agent faces three separate scenarios:

- All observation rights have been received (in this case, the jump operator is very similar to (1.6))

- Observation rights are still arriving, and the agent has observation rights stockpiled, and so can observe at any time
- Observation rights are still arriving, but the agent has no observation rights stockpiled, and so cannot observe until one arrives

Additionally, we study asymptotics as the number of observation rights tends to infinity, and use the theory of Extended Weak Convergence from [1] to prove the convergence of the discrete observation value functions to the value function of the classical continuous observation problem of [52].

CHAPTER II

Stability of Exponential Utility Maximization

II.1 Introduction

In this chapter we provide stability results for the problem of maximizing expected exponential utility. We give conditions under which convergence of markets implies the convergence of optimal terminal wealths as well as their expected utility. Specifically, for markets of the form $S = M + \int \lambda d\langle M \rangle$, our regularity condition consists of two complementary hypotheses: the first, the familiar V-compactness assumption of [36], is used to establish lower semi-continuity, while the second, a new condition related to a local *bmo* hypothesis, is used to establish upper semi-continuity. Both the V-compactness and local *bmo* conditions originally arose as consequences of our original regularity condition, a uniform bound on the bmo_2 norm of $\lambda \cdot M$. This type of hypothesis is a natural one in mathematical finance and has, for example, appeared in [18] and [24], where it was used in connection with establishing closedness properties of the space of attainable terminal wealths. In the current setting, the *bmo* hypothesis allows us to find wealth processes which are simultaneously near optimal and bounded from below. This is useful because the optimal wealth process is in general unbounded, meaning that it may go arbitrarily far into the red. With this approximation result in hand, we may use the stability results of [36] for utility functions on $\mathbb{R}_+$ to obtain convergence under *bmo* regularity. From there, we can prove our most general continuity result, which builds upon the *bmo* arguments to establish upper semi-continuity.

In comparison with [36], dealing with the stability problem for utility functions on $\mathbb{R}_+$, our regularity assumption is of course stronger than the notion of V-compactness alone, since we impose the additional local *bmo* condition. We will show, however, that this condition is not especially stringent, and that on some level it is already implicit in the setting of [36]: indeed, the basic purpose of the assumption is to guarantee that over all markets, the optimal expected utility $E\left[U\left(\widehat{X}_T^n\right)\right]$ may be uniformly approximated by payoffs of the form $E\left[U\left(\widehat{X}_T^{(n,k)}\right)\right]$, with the processes $\widehat{X}^{(n,k)}$ satisfying $\sup_{0\leq t\leq T} U^-\left(\widehat{X}_t^{(n,k)}\right) \in L^\infty$. For utility functions defined on $(0,\infty)$, this property is guaranteed as soon as there is V-compactness.

In comparison with two other extant stability results in the literature, we see that our regularity hypothesis is weaker than in either of those papers, although they provide additional convergence results that are beyond the scope of this article. In [22], the stability of quadratic BSDE's is studied with respect to, among other things, perturbation of the driver. From the natural connection between this class of BSDE's and exponential utility maximization (see [38]), the results from [22] allow one to recover stability results about exponential utility maximization, but only under the more restrictive assumption of a uniform bound on $\lambda \cdot M$ in the Hardy Space H^∞, i.e. $||\lambda \cdot M||_{H^\infty} \triangleq ||\lambda^2 \cdot \langle M\rangle_T||_{L^\infty}$. Additionally, it is assumed that the filtration is continuous.

In [60], very strong convergence results are obtained in a narrow class of utility maximization problems, with equilibrium problems in mind. In order to use PDE methods, the setting is exclusively Markov, and the assumptions on market convergence are quite stringent: given a sequence $(\lambda^n)_{n=1,\dots,\infty}$ of drift parameters, essentially $\lambda^n(t) \to \lambda^\infty(t)$ in $L^\infty([0,T])$. These strong hypotheses are necessary to deduce quantitative estimates about the stability of exponential utility maximization.

Here, we take a different approach. We consider simply the continuity of exponential utility maximization in a general filtration, and are interested in finding minimal regularity conditions under which continuity will hold true. The outline of the chapter is as follows. In Section II.2, we provide the necessary background definitions to state our main result. In Section II.3, we present some preliminaries on the theory of bmo martingales. In Section II.4, we apply this theory to give a proof of an intermediate result. In Section II.5, we prove the main results of the chapter, using the results of Section II.4 to establish upper semi-continuity. Finally, in Section II.6, we discuss our second assumption in the context of [36] and discuss its economic significance, in connection with the opportunity process of [41] and [40]; the necessity of the first assumption is also addressed. We close with two appendices, Appendix A and Appendix B, which contain auxiliary results.

II.2 Setup and Main Results

Let $(\Omega, \mathcal{F}, P, (\mathcal{F}_t)_{t\in[0,T]})$ be a filtered probability space satisfying the usual conditions. We assume that $\mathcal{F}_T = \mathcal{F}$. Let M be a continuous local martingale, and let

$$\Lambda \triangleq \left\{ \lambda : \lambda \text{ is a predictable process satisfying } \int_0^T \lambda_u^2 d\langle M\rangle_u < \infty \right\}.$$

For $\lambda \in \Lambda$, define

$$S_t^\lambda \triangleq M_t + \int_0^t \lambda_u d\langle M\rangle_u, \tag{2.1}$$

where $\langle M\rangle = (\langle M\rangle_t)_{t\in[0,T]}$ denotes the quadratic variation of the local martingale M. Along with a numéraire bond, identically equal to 1, each S^λ defines a stock market, in which S^λ is interpreted as the discounted price of a tradeable asset.

We let $S^n \triangleq M + \int \lambda^n \, d\langle M \rangle$, $n = 1, \ldots, \infty$, describe a sequence of markets, and

$$\begin{aligned} Z^n &\triangleq \mathcal{E}(-\lambda^n \cdot M) \\ &= \exp\left(- \int_0^\cdot \lambda^n dM - \frac{1}{2} \int_0^\cdot (\lambda^n)^2 d\langle M \rangle \right) \end{aligned}$$

is the *nth* minimal martingale measure.

In the exponential utility maximization problem, an agent with utility function $U(x) \triangleq -\exp(-x)$ seeks to maximize $E\left[U(x + X_T)\right]$ over a set of admissible wealth processes X that start from initial capital zero. We set $V(y) \triangleq y \log y - y$ for $y > 0$, so that V is the convex dual of U. To define our regularity assumptions, we need the notion of *bmo* martingales.

Definition II.1. Let $1 \leq p < \infty$. A not necessarily continuous martingale R is in bmo_p, with $||R||_{bmo_p} = r$, if there is a minimal constant r such that

$$E\left[|R_T - R_{\tau-}|^p \mid \mathcal{F}_\tau\right]^{\frac{1}{p}} \leq r,$$

for all stopping times τ taking values in $[0, T]$. We will occasionally abbreviate bmo_1 to bmo.

For $p = 2$, if $||R||_{bmo_2} < \infty$, then $||R||_{bmo_2}$ also has the representation

$$\operatorname*{ess\,sup}_\tau \; E\left[\langle R \rangle_T - \langle R \rangle_{\tau-} \mid \mathcal{F}_\tau\right]^{\frac{1}{2}}.$$

The equivalence of this representation is derived from considering the martingale $R^2 - \langle R \rangle$.

Now we can state our two-pronged regularity assumption on a sequence of markets:

Assumption II.2. *[Regularity Assumption 1: V-compactness] The set $\{V(Z_T^n) : n \in \mathbb{N}\}$ is uniformly integrable.*

Assumption II.3. *[Regularity Assumption 2] There exists a sequence of stopping times* $(\tau_j) \uparrow T$ *such that* $\sup_n ||(\lambda^n \cdot M)^{\tau_j}||_{bmo_2} < \infty$ *for each* j.

We continue on with our description of the utility maximization problem. In comparison to utilities on $\mathbb{R}_+$, defining the right notion of admissibility is more complicated when U is finite-valued over the whole real line. We state here the most common definition of admissibility at this level of generality, for which we refer to [51]. Let $\mathcal{M}^n$ denote the set of equivalent local martingale measures for S^n.

Definition II.4. For any n, let H be predictable and S^n-integrable. We say that $H \cdot S^n \in \mathcal{A}^n$ if $H \cdot S^n$ is a $\mathbb{Q}$-martingale for every $\mathbb{Q} \in \mathcal{M}^n$ with finite entropy, that is, $E\left[V\left(\frac{d\mathbb{Q}}{dP}\right)\right] < \infty$.

The primal value function u^n, $n = 1, \ldots, \infty$, is defined as

$$u^n(x) \triangleq \sup_{X \in \mathcal{A}^n} E\left[U(x + X_T)\right], x \in \mathbb{R}.$$

In the stability problem for utility maximization, we seek assumptions on the processes Z^n that ensure the convergence of $u^n(\cdot)$ towards $u^\infty(\cdot)$. We can now state the main results of the chapter:

Theorem II.5. *Suppose that* $Z^n_T \to Z^\infty_T$ *in probability,* Z^∞ *is a martingale, and that Assumptions II.2 and II.3 are satisfied. Then* $u^n(\cdot) \to u^\infty(\cdot)$ *pointwise, hence locally uniformly.*

Theorem II.6. *Suppose that* $Z^n_T \to Z^\infty_T$ *in probability,* Z^∞ *is a martingale, and that Assumptions II.2 and II.3 are satisfied. Then for all* x *the optimal terminal wealths* $\widehat{X}^n_T(x)$ *converge to* $\widehat{X}^\infty_T(x)$ *in probability as* $n \to \infty$.

A crucial intermediate step in establishing these theorems lies in first establishing them under a stronger *bmo*-type hypothesis. This is the main intermediate theorem: we remark that under these assumptions, Z^∞ is automatically a martingale.

Theorem II.7. *Suppose that $Z_T^n \to Z_T^\infty$ in probability and that $\sup_n ||\lambda^n \cdot M||_{bmo_2} < \infty$. Then $u^n(\cdot) \to u^\infty(\cdot)$ pointwise, hence locally uniformly.*

II.3 BMO Preliminaries

Definition II.8. A positive martingale Y satisfies the Reverse Hölder Inequality $\mathcal{R}_p(\mathbb{P})$ for $p > 1$ with constant K_p and with respect to the measure $\mathbb{P}$, if there exists minimal K_p such that

$$E^{\mathbb{P}}\left[\frac{Y_T^p}{Y_\tau^p}\,\middle|\,\mathcal{F}_\tau\right] \le K_p$$

for all stopping times τ in $[0,T]$.

The following lemma is found in the appendix of [25], and originally in Propositions 5 and 6 of [21].

Lemma II.9. *Suppose that the collection $(\lambda^n \cdot M)_{n\ge1}$ is bounded in the bmo_2 norm. Then for some $p > 1$ which depends only on this uniform bound, the collection $(Z^n = \mathcal{E}(\lambda^n \cdot M))_{n\ge1}$ satisfies $\mathcal{R}_p(\mathbb{P})$ with, respectively, uniformly bounded constants C_p^n.*

Definition II.10. A positive martingale Y satisfies $\mathcal{R}_{LLogL}$ with constant K_{LLogL} if there exists minimal K_{LLogL} such that

$$E\left[V\left(\frac{Y_T}{Y_\tau}\right)\,\middle|\,\mathcal{F}_\tau\right] \le K_{LLogL}$$

for all stopping times τ in $[0,T]$.

Definition II.11. A positive càdlàg process Y satisfies condition (S) if there exist constants $0 < c \le 1 \le C$ such that $cY_- \le Y \le CY_-$.

The following proposition is mostly in the literature:

Proposition II.12. *Let R be a martingale such that $Y = \mathcal{E}(R)$ is a strictly positive martingale. Then $R \in bmo_2$ and there exists $h > 0$ such that $\Delta R \geq h - 1$ if and only if Y satisfies $\mathcal{R}_{LLogL}$ and condition (S). The constants K_{LLogL} and C of Y can be bounded as a function of $||R||_{bmo_2}$.*

Proof. In the ($\Leftarrow$) direction, Lemma 2.2 of [25] establishes that $R \in bmo_2$. Now $dY = Y_- dR$ and $\Delta Y = Y_- \Delta R$. By the first inequality of condition (S), $(c-1)Y_- \leq Y - Y_- = Y_- \Delta R$, implying that $\Delta R \geq c - 1 > -1$.

Now the ($\Rightarrow$) direction. Since R is in bmo_2 it is locally bounded; indeed, for $n \in \mathbb{N}$, let $\tau_n = \inf\{t : \Delta R_t \geq n\} \wedge T$, and let $r \triangleq ||R||_{bmo_2}$. Then $||R^{\tau_n}||_{bmo_2} \leq r$, so that

$$
\begin{aligned}
(\Delta R_{\tau_n})^2 &= \Delta \langle R \rangle_{\tau_n} \\
&= E\left[\langle R \rangle_{\tau_n} - \langle R \rangle_{\tau_n -} \mid \mathcal{F}_{\tau_n}\right] \\
&\leq r,
\end{aligned}
$$

so that the jumps of R are bounded in magnitude by $\sqrt{r}$. This implies that R is locally bounded. Then $\Delta Y = Y_- \Delta R \leq \sqrt{r} Y_-$. Hence $Y \leq Y_- + \sqrt{r} Y_-$. Additionally, $\Delta R \geq h - 1$ implies that

$$
\begin{aligned}
Y - Y_- &= \Delta Y \\
&Y_- \Delta R \\
&\geq Y_-(h-1),
\end{aligned}
$$

so $Y \geq h Y_-$. This establishes condition (S), with $C = 1 + \sqrt{r}$, which is bounded as a function of $||R||_{bmo_2}$.

By Lemma II.9, Y satisfies the reverse Hölder inequality for some $p > 1$. Since $x \log x \leq K' x^p$ for some constant K', it follows that Y satisfies $\mathcal{R}_{LLogL}$. Additionally, it is evident that Lemma II.9 also implies that Y satisfies $\mathcal{R}_{LLogL}$ with constant K_{LLogL} only depending on $||R||_{bmo_2}$. $\square$

Definition II.13. For each market n, let $\widehat{Z}^n$ be the minimal entropy martingale measure. Its existence and uniqueness is established in Theorem 2.2 of [50].

The next lemma is precisely Lemma 3.1 of [17]. We give a proof for the reader's convenience.

Lemma II.14. *For any n, if Z^n satisfies $\mathcal{R}_{LLogL}$, with constant K, then $\widehat{Z}^n$ satisfies $\mathcal{R}_{LLogL}$ with a constant less than or equal K.*

Proof. By hypothesis, $E\left[\frac{Z_T^n}{Z_\tau^n}\log\frac{Z_T^n}{Z_\tau^n}\Big|\mathcal{F}_\tau\right] \leq K$ for all stopping times τ less than than or equal to T. Suppose that $\widehat{Z}^n$ does not satisfy $\mathcal{R}_{LLogL}$ with a constant less than or equal K. Then there exists $\epsilon > 0$, a stopping time σ less than or equal to T, and a set $A \in \mathcal{F}_\sigma$ with $P(A) > 0$ such that

$$E\left[\frac{\widehat{Z}_T^n}{\widehat{Z}_\sigma^n}\log\frac{\widehat{Z}_T^n}{\widehat{Z}_\sigma^n}\Big|\mathcal{F}_\sigma\right] \geq K + \epsilon$$

on the set A. Let $\tilde{Z}_t^n \triangleq 1_{\{t<\sigma\}}\widehat{Z}_t^n + 1_{\{t\geq\sigma\}}\left(1_A\frac{Z_t^n}{Z_\sigma^n}\widehat{Z}_\sigma^n + 1_{A^c}\widehat{Z}_t^n\right)$ for $t \in [0,T]$. Then $\tilde{Z}^n$ is the density process of an element of $\mathcal{M}^n$ and satisfies $\tilde{Z}_T^n = 1_A\widehat{Z}_\sigma^n\frac{Z_T^n}{Z_\sigma^n} + 1_{A^c}\widehat{Z}_T^n$. Thus,

$$\tilde{Z}_T^n\log\tilde{Z}_T^n = 1_{A^c}\widehat{Z}_T^n\log\widehat{Z}_T^n + 1_A\left(\widehat{Z}_\sigma^n\frac{Z_T^n}{Z_\sigma^n}\log\frac{Z_T^n}{Z_\sigma^n} + \widehat{Z}_\sigma^n\frac{Z_T^n}{Z_\sigma^n}\log\widehat{Z}_\sigma^n\right).$$

Therefore,

$$\begin{aligned}
&E\left[\widetilde{Z}_T^n\log\widetilde{Z}_T^n|\mathcal{F}_\sigma\right] - E\left[\widehat{Z}_T^n\log\widehat{Z}_T^n|\mathcal{F}_\sigma\right]\\
&\quad= 1_A\left(\widehat{Z}_\sigma^n E\left[\frac{Z_T^n}{Z_\sigma^n}\log\frac{Z_T^n}{Z_\sigma^n}|\mathcal{F}_\sigma\right] + \widehat{Z}_\sigma^n\log\widehat{Z}_\sigma^n - E\left[\widehat{Z}_T^n\log\widehat{Z}_T^n|\mathcal{F}_\sigma\right]\right)\\
&\quad= 1_A\widehat{Z}_\sigma^n\left(E\left[\frac{Z_T^n}{Z_\sigma^n}\log\frac{Z_T^n}{Z_\sigma^n}|\mathcal{F}_\sigma\right] - E\left[\frac{\widehat{Z}_T^n}{\widehat{Z}_\sigma^n}\log\frac{\widehat{Z}_T^n}{\widehat{Z}_\sigma^n}|\mathcal{F}_\sigma\right]\right)\\
&\quad\leq -\epsilon 1_A\widehat{Z}_\sigma^n.
\end{aligned}$$

Taking expectations, this contradicts the fact that $\widehat{Z}_T^n$ has minimal entropy. □

We now show that the bmo_2 hypothesis of (II.7) implies the V-compactness condition of Assumption II.2, which plays a prominent role in [36]. The next proposition is proven for continuous martingales in [32].

Proposition II.15. *Suppose* $\sup_n ||\lambda^n \cdot M||_{bmo_2} < \infty$*. Then there exists* $p > 1$ *such that* $\sup_n E\left[(Z_T^n)^p\right] < \infty$*.*

Proof. By the conditional form of Jensen's inequality, the norm $||\cdot||_{bmo_1} \leq ||\cdot||_{bmo_2}$. Let R be an arbitrary element of bmo_2, and let $n(R) = 2||R||_{bmo_1} + ||R||_{bmo_2}^2$. Without loss of generality, we assume that $||R||_{bmo_2} > 0$, and show that the L^p norm of $\mathcal{E}(R)_T$ has an upper bound that only depends on $n(R)$ for some $p > 1$.

Let $\delta = \exp(-pn(R)) < 1$ (so $\log 1/\delta = pn(R)$), and let $\tau = \inf\{t : \mathcal{E}(R)_t > \lambda\}$ for $\lambda > 1$. Considering time $\tau-$ instead of τ and arguing as in [32], we obtain the inequality $P(\mathcal{E}(R)_T/\mathcal{E}(R)_{\tau-} \geq \delta \mid \mathcal{F}_\tau) \geq 1 - \frac{1}{2p}$; indeed,

$$
\begin{aligned}
&P\left(\mathcal{E}(R)_T/\mathcal{E}(R)_{\tau-} < \delta \mid \mathcal{F}_\tau\right) \\
&\quad = P\left(1/\delta < \mathcal{E}(R)_{\tau-}/\mathcal{E}(R)_T \mid \mathcal{F}_\tau\right) \\
&\quad = P\left(pn(R) < R_{\tau-} - R_T + \frac{1}{2}(\langle R\rangle_T - \langle R\rangle_{\tau-}) \mid \mathcal{F}_\tau\right) \\
&\quad \leq \frac{1}{2pn(R)} E\left[2|R_T - R_{\tau-}| + (\langle R\rangle_T - \langle R\rangle_{\tau-}) \mid \mathcal{F}_\tau\right] \\
&\quad \leq \frac{n(R)}{2pn(R)} \\
&\quad = \frac{1}{2p},
\end{aligned}
$$

with the first inequality following from Markov's inequality. This implies that $P\left(\mathcal{E}(R)_T/\mathcal{E}(R)_{\tau-} \geq \delta \mid \mathcal{F}_\tau\right) \geq 1 - \frac{1}{2p}$.

By Proposition II.12, $\mathcal{E}(R)$ satisfies the upper bound of condition (S) with a constant C whose size is controlled by $n(R)$, and we have $\mathcal{E}(R)_{\tau-} \geq \frac{1}{C}\mathcal{E}(R)_\tau \geq \frac{1}{C}\lambda$

on $\{\tau < \infty\}$. This yields $P\left(\mathcal{E}(R)_T \geq \frac{\delta\lambda}{C} \mid \mathcal{F}_\tau\right) \geq \frac{2p-1}{2p} 1_{\{\tau<\infty\}}$. Thus,

$$
\begin{aligned}
&E\left[\mathcal{E}(R)_T 1_{\{\mathcal{E}(R)_T > \lambda\}}\right] \\
&\quad \leq E\left[\mathcal{E}(R)_T 1_{\{\tau<\infty\}}\right] \\
&\quad = E\left[\mathcal{E}(R)_\tau 1_{\{\tau<\infty\}}\right] \\
&\quad \leq E\left[C\mathcal{E}(R)_{\tau-} 1_{\{\tau<\infty\}}\right] \\
&\quad \leq C\lambda P(\tau < \infty) \\
&\quad \leq \frac{2C\lambda p}{2p-1} P\left(\mathcal{E}(R)_T \geq \frac{\delta\lambda}{C}\right),
\end{aligned}
$$

where the equality above follows from the optional sampling theorem.

Take the inequality $E\left[\mathcal{E}(R)_T 1_{\{\mathcal{E}(R)_T > \lambda\}}\right] \leq \frac{2C\lambda p}{2p-1} P\left(\mathcal{E}(R)_T \geq \frac{\delta\lambda}{C}\right)$, multiply both sides by $(p-1)\lambda^{p-2}$ and integrate with respect to λ from 1 to ∞:

$$
\int_1^\infty (p-1)\lambda^{p-2} E\left[\mathcal{E}(R)_T 1_{\{\mathcal{E}(R)_T > \lambda\}}\right] d\lambda \tag{2.2}
$$

$$
\leq \int_1^\infty (p-1)\lambda^{p-2} \frac{2C\lambda p}{2p-1} P\left(\mathcal{E}(R)_T \geq \frac{\delta\lambda}{C}\right) d\lambda. \tag{2.3}
$$

Applying Fubini's Theorem to the left hand side (2.2), we get

$$
\begin{aligned}
&\int_1^\infty (p-1)\lambda^{p-2} E\left[\mathcal{E}(R)_T 1_{\{\mathcal{E}(R)_T > \lambda\}}\right] d\lambda \\
&\quad = E\left[\int_1^\infty (p-1)\lambda^{p-2} \mathcal{E}(R)_T 1_{\{\mathcal{E}(R)_T > \lambda\}} d\lambda\right] \\
&\quad = E\left[\mathcal{E}(R)_T \int_1^\infty (p-1)\lambda^{p-2} 1_{\{\mathcal{E}(R)_T > \lambda\}} d\lambda\right] \\
&\quad = E\left[\mathcal{E}(R)_T 1_{\{\mathcal{E}(R)_T > 1\}} \int_1^{\mathcal{E}(R)_T} (p-1)\lambda^{p-2} d\lambda\right] \\
&\quad = E\left[\mathcal{E}(R)_T \left(\mathcal{E}(R)_T^{p-1} - 1\right) 1_{\{\mathcal{E}(R)_T > 1\}}\right].
\end{aligned}
$$

After a similar computation for the right hand side (2.3), this yields

$$
\begin{aligned}
&E\left[(\mathcal{E}(R)_T^p - \mathcal{E}(R)_T) 1_{\{\mathcal{E}(R)_T > 1\}}\right] \\
&\quad \leq \frac{2C(p-1)}{2p-1} E\left[\left(\left(\frac{C}{\delta}\mathcal{E}(R)_T\right)^p - 1\right) 1_{\{\mathcal{E}(R)_T > \frac{\delta}{C}\}}\right].
\end{aligned}
$$

Grouping the terms with $\mathcal{E}(R)_T^p$ together on the left hand side, we obtain

$$\begin{aligned}
&\left(1-\frac{2C(p-1)}{2p-1}\frac{C^p}{\delta^p}\right)E\left[\mathcal{E}(R)_T^p 1_{\{\mathcal{E}(R)_T>1\}}\right]\\
&\quad\le E\left[\mathcal{E}(R)_T\right]-\frac{2C(p-1)}{2p-1}E\left[1_{\{\mathcal{E}(R)_T>\frac{\delta}{C}\}}\right]\\
&\quad\le 1,
\end{aligned}$$

for any $p>1$. Hence, by choosing p close enough to 1 so that $\frac{2C(p-1)}{2p-1}\frac{C^p}{\delta^p}<1$, we establish an upper bound for $E[\mathcal{E}(R)_T^p]$ which depends only on $n(R)$. Note that the choice of C depends on $n(R)$. □

Corollary II.16. *Suppose that* $\sup_n ||\lambda^n\cdot M||_{bmo_2}<\infty$. *Then* $\{V(Z_T^n):n\in\mathbb{N}\}$ *is uniformly integrable.*

Proof. By Proposition II.15, $\sup_n E[(Z_T^n)^p]<\infty$ for some $p>1$. As $x^{\widetilde{p}}/V(x)\to\infty$ as $x\to\infty$, for any $\widetilde{p}>1$, the claim follows from the de la Vallée-Poussin criterion. □

We make one last digression to the theory of bmo martingales. Specifically, we need the bmo theory of weighted norm inequalities. The following theorem is stated as Theorem 2.16 of [19] without mentioning that the constant C_p in (2.4) can be chosen as the same constant associated with the reverse Hölder inequality. For this fact, we refer to Proposition 2 of [21]. For a càdlàg process Y, let $Y^*\triangleq\sup_{t\in[0,T]}|Y_t|\in\mathcal{F}_T$.

Proposition II.17. *Let* $Y=\mathcal{E}(R)$ *be a continuous martingale and* $\frac{d\mathbb{Q}}{dP}=Y_T$. *Then if* Y *satisfies* $\mathcal{R}_p(P)$ *with constant* C_p, *then for each* $\mathbb{Q}$*-martingale* X *and* $q=\frac{p}{p-1}$,

$$\lambda^q P\left(X^*>\lambda\right)\le C_p E\left[|X_T|^q\right]. \tag{2.4}$$

II.4 Approximation of Optimal Wealth

In [50], U is approximated by auxiliary utility functions defined on a half axis. For $k\in\mathbb{N}$, we define utility functions $U^{(k)}$ as follows: $U^{(k)}=U$ on $[-k,\infty)$, $U(x)\ge$

$U^{(k)}(x) > -\infty$ for $x > -k-1$, and $\lim\limits_{x \downarrow -k-1} U^{(k)}(x) = -\infty$. Each $U^{(k)}$ is assumed C^1, concave, satisfying the Inada conditions, and having reasonable asymptotic elasticity. For details on these assumptions, see [50]. $V^{(k)}$ is the convex conjugate of $U^{(k)}$. Since $U^{(k)} \leq U$, $V^{(k)} \leq V$.

For $n = 1, \ldots, \infty$, v^n is the dual value function associated to V in market number n:

$$v^n(y) \triangleq \inf_{\mathbb{Q} \in \mathcal{M}^n} E\left[V\left(y\frac{d\mathbb{Q}}{dP}\right)\right],\ y > 0. \tag{2.5}$$

For $n = 1, \ldots, \infty$ and $k \in \mathbb{N}$, $v^{(n,k)}$ is the dual value function associated to $V^{(k)}$ in market number n:

$$v^{(n,k)}(y) \triangleq \inf_{Y \in \mathcal{Y}^n} E\left[V^{(k)}(yY_T)\right],\ y > 0,$$

where $\mathcal{Y}^n$ is the set of supermartingale deflators for S^n:

Definition II.18. $\mathcal{Y}^n$ is the set of càdlàg processes Y such that $Y_0 = 1$ and $Y(H \cdot S^n)$ is a supermartingale whenever H is predictable, S^n-integrable, such that $H \cdot S^n$ is bounded from below by a constant.

Let $\mathcal{A}_b^n$ be the set of wealth processes $H \cdot S^n$ where H is predictable and S^n-integrable, and $H \cdot S^n$ is bounded from below by a constant. The value functions $u^{(n,k)}$ are defined as follows:

$$u^{(n,k)}(x) \triangleq \sup_{X \in \mathcal{A}_b^n} E\left[U^{(k)}(x + X_T)\right],\ x > -k-1.$$

By a shift on the real line (see [50]), one can identify the value functions $v^{(n,k)}$, $u^{(n,k)}$ with an equivalent optimization problem which uses a utility function $\widetilde{U}^{(k)}$ defined on $\mathbb{R}_+$. We copy verbatim this procedure here.

Let $\widetilde{U}^{(k)}(x) \triangleq U^{(k)}(x - (k+1))$, which is finitely valued for $x > 0$. Then $\widetilde{U}^{(k)}$ is a utility function of the type encountered in [34], and so there is a unique optimal solution $\overline{X}^{(n,k)}(x) = x + H^{(n,k)} \cdot S^n$ to the optimization problem

$$\widetilde{u}^{(n,k)}(x) \triangleq \sup_{X \in \mathcal{A}_b^n} E\left[\widetilde{U}^{(k)}(X_T)\right],\ x > -k-1.$$

Then, for $x > -k-1$, $\widehat{X}^{(n,k)}(x) \triangleq \overline{X}^{(n,k)}(x+k+1) - (k+1)$ is the optimal solution to the optimization problem

$$u^{(n,k)}(x) = \sup_{X \in \mathcal{A}_b^n} E\left[U^{(k)}(x + X_T)\right],\ x > 0.$$

It follows that $u^{(n,k)}(x) = \widetilde{u}^{(n,k)}(x+k+1)$ for $x > -k-1$. Let $\widetilde{V}^{(k)}$ be the convex conjugate of $\widetilde{U}^{(k)}$. Then the convex conjugate $\widetilde{v}^{(n,k)}$ of $\widetilde{u}^{(n,k)}$ has the form

$$\begin{aligned}\widetilde{v}^{(n,k)}(y) &= \inf_{Y \in \mathcal{Y}^n} E\left[\widetilde{V}^{(k)}(yY_T)\right] \\ &= E\left[\widetilde{V}^{(k)}\left(y\widetilde{Y}_T^{(n,k)}\right)\right],\ y > 0;\end{aligned}$$

Here, $\widetilde{Y}^{(n,k)} = \widetilde{Y}^{(n,k)}(y)$ is the dual minimizer, which in general depends on y; the existence of such minimizers is established in [34]. We also have

$$V^{(k)}(y) = \widetilde{V}^{(k)}(y) + (k+1)y \tag{2.6}$$

and $v^{(n,k)}(y) = \widetilde{v}^{(n,k)}(y) + (k+1)y$. The main result of [36] implies that for each k, $\lim_{n\to\infty} \widetilde{u}^{(n,k)} = \widetilde{u}^{(\infty,k)}$ under the $\widetilde{V}^k$-compactness condition: $\{\widetilde{V}^k(Z_T^n) : n \in \mathbb{N}\}$ is uniformly integrable.

Lemma II.19. *Suppose that $Z_T^n \to Z_T^\infty$ in probability and $\{Z_T^n : n \in \mathbb{N}\}$ is V-compact. Then for each k, $\lim_{n\to\infty} u^{(n,k)}(x) = u^{(\infty,k)}(x)$.*

Proof. For each k, $V^{(k)} \leq V$ and $V^{(k)}$ is bounded from below, so $\{V^{(k)}(Z_T^n) : n \in \mathbb{N}\}$ is uniformly integrable. Since $V(x)/x \to \infty$ as $x \to \infty$, it is also true that $\{Z_T^n : n \in \mathbb{N}\}$ is uniformly integrable. Given the form of $\widetilde{V}^{(k)}$ in (2.6), it now follows that $\{\widetilde{V}^{(k)}(Z_T^n) : n \in \mathbb{N}\}$ is uniformly integrable. Hence the main theorem of [36] implies that $\widetilde{u}^{(n,k)}(x) \to \widetilde{u}^{(\infty,k)}(x)$ as $n \to \infty$. It immediately follows that $u^{(n,k)}(x) \to u^{(\infty,k)}(x)$ as $n \to \infty$. $\square$

Lemma II.20. *Suppose that* $v^*(y) \triangleq \sup_n v^n(y) < \infty$ *for all* $y > 0$*. Then for all* $x \in \mathbb{R}$, $u^*(x) \triangleq \sup_n u^n(x) < 0$.

Proof. By passing to a subsequence, we can assume that $u^n(0) \to u^*(0)$. For each n, $u^n(x) = \exp(-x)u^n(0)$, and similarly for $u^*(x)$. Hence, $u^n \to u^*$ locally uniformly and u^* is concave. Let $\overline{v}$ be the convex dual of u^*. Since v^n and u^n are convex duals, then $\lim_n v^n$ exists and is the convex dual of u^*, and hence is equal to $\overline{v}$. By definition, $\overline{v} \le v^*$. Suppose that for some x, $u^*(x) = 0$. Then $u^* \equiv 0$. But, if $u^* \equiv 0$, then it would be that $\overline{v}(y) = \sup_{x\in\mathbb{R}}[u^*(x) - xy] \equiv \infty$, which contradicts the finiteness of $v^*(y)$. Thus, $u^*(x)$ is bounded away from zero. □

Let

$$x + \widehat{X}^n \triangleq x + \widehat{X}^n(0) = \widehat{X}^n(x)$$

be the optimal wealth process in market n from initial capital x. This special form for the optimal wealth processes is due to the wealth homogeneity of the exponential utility. Let $\mathcal{T}$ be the set of $[0, T]$-valued stopping times.

Proposition II.21. *Suppose that* $\sup_n ||\lambda^n \cdot M||_{bmo_2} < \infty$*. Then* $\{\exp(-\widehat{X}^n_\tau) : n \in \mathbb{N}, \tau \in \mathcal{T}\}$ *is uniformly integrable.*

Proof. Recall $\widehat{Z}^n$ is the density of the minimal entropy martingale measure for S^n, which we denote by $\widehat{\mathbb{Q}}^n$. From Theorem 2.2 of [50], $\widehat{X}^n$ is a true $\widehat{\mathbb{Q}}^n$-martingale for each n. From Theorem 2.2 of [50] again, we have $c_n e^{-\widehat{X}^n_T} = \widehat{Z}^n_T$ for some constant c_n.

Taking conditional expectations under $\widehat{\mathbb{Q}}^n$ via Bayes' rule, and using the fact that

$\widehat{X}^n$ is a $\widehat{\mathbb{Q}}^n$-martingale, we obtain

$$
\begin{aligned}
&\log c_n - \widehat{X}^n_\tau \\
&\quad = E^{\widehat{\mathbb{Q}}^n}\left[\log c_n - \widehat{X}^n_T | \mathcal{F}_\tau\right] \\
&\quad = E^{\widehat{\mathbb{Q}^n}}\left[\log \widehat{Z}^n_T | \mathcal{F}_\tau\right] \\
&\quad = E\left[\frac{\widehat{Z}^n_T}{\widehat{Z}^n_\tau}\log \widehat{Z}^n_T \,\middle|\, \mathcal{F}_\tau\right] \\
&\quad = E\left[\frac{\widehat{Z}^n_T}{\widehat{Z}^n_\tau}\left(\log \frac{\widehat{Z}^n_T}{\widehat{Z}^n_\tau} + \log \widehat{Z}^n_\tau\right) \,\middle|\, \mathcal{F}_\tau\right] \\
&\quad = E\left[\frac{\widehat{Z}^n_T}{\widehat{Z}^n_\tau}\log \frac{\widehat{Z}^n_T}{\widehat{Z}^n_\tau} \,\middle|\, \mathcal{F}_\tau\right] + \log \widehat{Z}^n_\tau.
\end{aligned}
$$

Exponentiating the previous inequality, we obtain

$$
\begin{aligned}
&\exp(-\widehat{X}^n_\tau) \\
&\quad = \frac{1}{c_n}\widehat{Z}^n_\tau \exp\left(E\left[\frac{\widehat{Z}^n_T}{\widehat{Z}^n_\tau}\log \frac{\widehat{Z}^n_T}{\widehat{Z}^n_\tau} \,|\mathcal{F}_\tau\right]\right) \\
&\quad \leq \frac{1}{c_n} e^{\widehat{K}^n_{LLogL}+1}\widehat{Z}^n_\tau,
\end{aligned}
$$

where $\widehat{K}^n_{LLogL}$ is the $\mathcal{R}_{LLogL}$ constant of $\widehat{Z}^n$. According to Proposition II.12 and Lemma II.14, $\sup\limits_n \widehat{K}^n_{LLogL} < \infty$. By Corollary II.16, $v^*(y) < \infty$, and so Lemma II.20 implies that $u^* < 0$. Note that $c_n = -u^n(0)$. Thus, $\inf\limits_n c_n > 0$, so that $\sup\limits_n \frac{1}{c_n} < \infty$. We may then write

$$
\exp(-\widehat{X}^n_\tau) \leq \mathbf{C}\widehat{Z}^n_\tau \tag{2.7}
$$

for some constant $\mathbf{C}$, so that the inequality is valid for all n and all τ. In what follows we will show that the right-hand-side of (2.7) is uniformly integrable, which completes the proof. Since $\sup\limits_n E\left[V(\widehat{Z}^n_T)\right] < \infty$ (thanks to V-compactness and Lemma II.14) and $V(x)/x \to \infty$ as $x \to \infty$, the de la Vallée-Poussin criterion implies that $\{\widehat{Z}^n_T : n \in \mathbb{N}\}$ is uniformly integrable. Since each $\widehat{Z}^n$ is a martingale, this extends to the uniform integrability of $\{\widehat{Z}^n_\tau : n \in \mathbb{N}, \tau \in \mathcal{T}\}$. □

Remark II.22. In the literature (see [26]), admissible wealth processes are sometimes defined directly to be those satisfying the conclusion of Proposition II.21, i.e. having uniformly integrable utility over all stopping times.

For $i \in \mathbb{Z}$, let $\widehat{\tau}^{(n,i)} \triangleq \inf\{t : \widehat{X}^n_t = i\}$, and let $\widehat{X}^{(n,i)} \triangleq (\widehat{X}^n)^{\widehat{\tau}^{(n,i)}} = \left(\widehat{X}^n_{\widehat{\tau}^{(n,i)} \wedge t}\right)_{t\in[0,T]}$.

Lemma II.23. *Suppose that* $\sup_n ||\lambda^n \cdot M||_{bmo_2} < \infty$. *Then for each* $i \in \mathbb{N}$*, the collection* $\{(\widehat{X}^{(n,i)})^* : n \in \mathbb{N}\}$ *is bounded in probability.*

Remark II.24. The conclusions of Lemma II.23 and Proposition II.21 will be shown to be sufficient for obtaining continuity of the utility maximization problems. Given the strength of the *bmo* hypothesis, it is natural to ask whether these conditions are also necessary. In Appendix A, it is shown that the conclusion of Lemma II.23 is indeed necessary. The conclusion of Proposition II.21, however, is not, and it in fact may fail within a *single* market. We give an example of this in Appendix B. Note that this market, and indeed all continuous markets, still satisfy the local *bmo* hypothesis of Assumption II.3.

Proof. Let $\mathbb{Q}^n$ be the probability measure associated to the minimal martingale Z^n, which is continuous. By Corollary II.16, each $\mathbb{Q}^n$ has finite entropy. Theorem 1 of [51] implies that $\widehat{X}^n$ is a $\mathbb{Q}^n$-martingale for each n. Then it is also true that $\widehat{X}^{(n,i)}$ is a $\mathbb{Q}^n$-martingale for each n. Since $\sup_n ||\lambda^n \cdot M||_{bmo_2} < \infty$, Lemma II.9 implies that there exists a $p > 1$ such that each Z^n satisfies the Reverse Hölder inequality $\mathcal{R}_p(P)$ with uniformly bounded constant C_p.

By Proposition II.17, for $q = \frac{p}{p-1}$,

$$\begin{aligned}
&\lambda^q P\left((\widehat{X}^{(n,i)})^* > \lambda\right) \\
&\quad \leq C_p E\left[\left|\widehat{X}_T^{(n,i)}\right|^q\right] \\
&\quad \leq C_p \left(i^q + \mathbf{C}_i E\left[\exp\left(-\widehat{X}_T^{(n,i)}\right)\right]\right) \\
&\quad \leq C_p(i^q + \widetilde{\mathbf{C}}_i),
\end{aligned}$$

for constants $\mathbf{C}_i, \widetilde{\mathbf{C}}_i$ independent of n, with the third inequality a consequence of Proposition II.21. □

Proposition II.25. *Suppose* $\sup_n ||\lambda^n \cdot M||_{bmo_2} < \infty$*. Then* $u^{(n,k)} \to u^n$ *as* $k \to \infty$*, uniformly over the markets* n*.*

Remark II.26*.* As indicated by Proposition II.34 in Appendix A, the uniform approximation condition given above is both necessary and sufficient for convergence of the utility maximization problem.

Proof. Let $\epsilon > 0$. Fix $i \in \mathbb{N}$ large enough so that $0 > -\exp(-i) > -\epsilon$. Then

$$\begin{aligned}
u^n(0) &\geq E\left[U(\widehat{X}_T^{(n,i)})\right] \\
&> u^n(0) - \epsilon \text{ for all } n \in \mathbb{N}.
\end{aligned} \tag{2.8}$$

For $k \in \mathbb{N}$, let $\widehat{X}^{(n,i,-k)} \triangleq (\widehat{X}^n)^{\widehat{\tau}^{(n,-k)} \wedge \widehat{\tau}^{(n,i)}} = (\widehat{X}^{(n,i)})^{\widehat{\tau}^{(n,-k)}}$. We claim that

$$\lim_{k\to\infty} \sup_{n \in \mathbb{N}} P\left(\widehat{\tau}^{(n,-k)} < \widehat{\tau}^{(n,i)}\right) = 0. \tag{2.9}$$

Indeed, Lemma II.23 implies that the collection $\{(\widehat{X}^{(n,i)})^* : n \in \mathbb{N}\}$ is bounded in probability. Therefore, $\lim_{k\to\infty} \sup_n P((\widehat{X}^{(n,i)})^* \geq k) = 0$. But $P(\widehat{\tau}^{(n,-k)} < \widehat{\tau}^{(n,i)}) \leq P((\widehat{X}^{(n,i)})^* \geq k)$, which establishes (2.9). We next claim that

$$\lim_{k\to\infty} \sup_n \left| E\left[U(\widehat{X}_T^{(n,i)})\right] - E\left[U(\widehat{X}_T^{(n,i,-k)})\right] \right| = 0. \tag{2.10}$$

Let $\epsilon_2 > 0$. Write

$$
\begin{aligned}
&E\left[U(\widehat{X}_T^{(n,i,-k)})\right] \\
&\quad = E\left[U(\widehat{X}_T^{(n,i)})1_{\{\widehat{\tau}^{(n,-k)} \geq \widehat{\tau}^{(n,i)}\}} + U(\widehat{X}_T^{(n,i,-k)})1_{\{\widehat{\tau}^{(n,-k)} < \widehat{\tau}^{(n,i)}\}}\right] \\
&\quad = E\left[U(\widehat{X}_T^{(n,i)}) - U(\widehat{X}_T^{(n,i)})1_{\{\widehat{\tau}^{(n,-k)} < \widehat{\tau}^{(n,i)}\}} + U(\widehat{X}_T^{(n,i,-k)})1_{\{\widehat{\tau}^{(n,-k)} < \widehat{\tau}^{(n,i)}\}}\right].
\end{aligned}
$$

According to Proposition II.21, the set $\{\exp(-\widehat{X}_\tau^n) : n \in \mathbb{N}, \tau \in \mathcal{T}\}$ is uniformly integrable, which immediately implies that the set $\{\exp(-\widehat{X}_T^{(n,i)}), \exp(-\widehat{X}_T^{(n,i,-k)}) : n, k \in \mathbb{N}\}$ is uniformly integrable. Therefore, there exists $\delta = \delta(\epsilon_2) > 0$ such that for any set A, $P(A) < \delta$ implies that $\max\left\{E[U(\widehat{X}_T^{(n,i)})1_A], E[U(\widehat{X}_T^{(n,i,-k)})1_A]\right\} < \epsilon_2$. According to (2.9), there exists $k_0 \in \mathbb{N}$ such that for $k \geq k_0$ and all $n \in \mathbb{N}$, the sets $\{\widehat{\tau}^{(n,-k)} < \widehat{\tau}^{(n,i)}\}$ have probability less than δ. Therefore, for $k \geq k_0$ and all $n \in \mathbb{N}$, $\max\left\{E\left[U(\widehat{X}_T^{(n,i)})1_{\{\widehat{\tau}^{(n,-k)} < \widehat{\tau}^{(n,i)}\}}\right], E\left[U(\widehat{X}_T^{(n,i,-k)})1_{\{\widehat{\tau}^{(n,-k)} < \widehat{\tau}^{(n,i)}\}}\right]\right\} < \epsilon_2$. Thus, for $k \geq k_0$ and all $n \in \mathbb{N}$, we have

$$
\left| E\left[U(\widehat{X}_T^{(n,i)})\right] - E\left[U(\widehat{X}_T^{(n,i,-k)})\right] \right| < 2\epsilon_2,
$$

and (2.10) is established. Then (2.8) and (2.10) imply that

$$
\lim_{k\to\infty} \sup_{n\in\mathbb{N}} \left| u^n(0) - E\left[U(\widehat{X}_T^{(n,i,-k)})\right] \right| \leq \epsilon. \tag{2.11}
$$

Since $\widehat{X}^{(n,i,-k)} > -k-1$, by definition, $u^{(n,k)}(0) \geq \left[U(\widehat{X}_T^{(n,i,-k)})\right]$. Then (2.11) and the fact that $u^{(n,k)} \leq u^n$ imply that for any $\epsilon > 0$,

$$
\lim_{k\to\infty} \sup_{n\in\mathbb{N}} |u^n(0) - u^{(n,k)}(0)| \leq \epsilon, \tag{2.12}
$$

implying that $\lim\limits_{k\to\infty} \sup\limits_{n\in\mathbb{N}} |u^{(n,k)}(0) - u^n(0)| = 0$, i.e. that $u^{(n,k)} \to u^n$ as $k \to \infty$, uniformly over n. $\square$

II.4.1 Proof of the intermediate theorem

Proof of Theorem II.7. It follows from Corollary II.16 and Lemma II.19 that for each k, $\lim_{n\to\infty} u^{(n,k)} = u^{(\infty,k)}$. Proposition II.25, on the other hand, states that $\lim_{k\to\infty} u^{(n,k)} = u^n$, uniformly over n. These facts together imply that $\lim_{n\to\infty} u^n = u^\infty$. $\square$

II.5 Proofs of the Main Theorems

We establish the main Theorem II.5 in pieces, establishing lower semi-continuity and upper semi-continuity separately. The proof of lower semi-continuity is the easier of the two, and indeed is not dependent on the special structure of the exponential utility function.

Lemma II.27. *Suppose that $Z_T^n \to Z_T^\infty$ in probability and that $\{Z_T^n : n \in \mathbb{N}\}$ is V-compact, i.e. Assumption II.2 holds. Then $u^\infty(x) \leq \liminf_{n\to\infty} u^n(x)$.*

Proof. As in the proof of Lemma II.19, the V-compactness of $\{Z_T^n : n \in \mathbb{N}\}$ implies that this set is also $V^{(k)}$-compact, where $V^{(k)}$ is the dual of the "truncated" utility function $U^{(k)} \leq U$, defined at the beginning of Section II.4. By Lemma II.19 and the main theorem of [36], $\lim_{n\to\infty} u^{(n,k)}(x) = u^{(\infty,k)}(x)$ for each $k \in \mathbb{N}$. By Step 1 of Theorem 2.2 of [50], $u^n(x) = \sup_{k\in\mathbb{N}} u^{(n,k)}(x)$ for each n. Therefore

$$\begin{aligned}
\liminf_{n\to\infty} u^n(x) &= \liminf_{n\to\infty} \sup_{k\in\mathbb{N}} u^{(n,k)}(x) \\
&\geq \sup_{k\in\mathbb{N}} \liminf_{n\to\infty} u^{(n,k)}(x) \\
&= \sup_{k\in\mathbb{N}} u^{(\infty,k)}(x) \\
&= u^\infty(x).
\end{aligned}$$

$\square$

We now establish upper semi-continuity.

Proposition II.28. *Suppose that there exists a sequence of stopping times* $(\tau_j) \uparrow T$ *such that for each* j, $\sup_n ||(\lambda^n \cdot M)^{\tau_j}||_{bmo} < \infty$, *i.e. Assumption II.3 holds. Additionally, suppose that* $V(Z_T^\infty) \in L^1$ *and* Z^∞ *is a martingale. Then* $u^\infty(x) \geq \liminf_{n\to\infty} u^n(x)$.

Proof. For $j, n = 1, \ldots, \infty$, let $u^{n,j}$ denote the indirect utility arising from trading in market n up until time τ_j, where $u^n = u^{n,\infty}$. Since all trading opportunities arising on $[0, \tau_j)$ are also available over the whole time period $[0, T]$, we know that $u^{n,j} \leq u^{n,j+1} \leq u^{n,\infty}$. We claim that in addition,

$$u^{\infty,j} \uparrow u^\infty \tag{2.13}$$

as $j \to \infty$. As Z^∞ is a martingale and V is convex, $V(Z^\infty)$ is a submartingale (whose terminal value is integrable). As V is bounded from below, this implies that $V(Z^\infty)$ is of Class D, as defined, in [20], p.11, for example. In particular, the set $\left\{V\left(Z_{\tau_j}^\infty\right) : j \in \mathbb{N}\right\}$ is uniformly integrable. In the context of Lemma II.27, set $Z^j \triangleq (Z^\infty)^{\tau_j}$, so that $Z_{\tau_j}^\infty = Z_T^j$. So, applying Lemma II.27 to the sequence $\{Z^j\}$, it follows that $u^\infty \leq \liminf_{j\to\infty} u^{\infty,j}$. As $u^{\infty,j} \leq u^{\infty,j+1} \leq u^\infty$, (2.13) now follows.

We now claim that for each $j < \infty$, $u^{n,j} \to u^{\infty,j}$. First, $Z_T^n \to Z_\infty^n$ in L^1 by Scheffe's Lemma, and hence $Z^n \to Z^\infty$ in ucp, which follows by applying Doob's weak L^1 inequality. In particular, $Z_{\tau_j}^n \to Z_{\tau_j}^\infty$ in probability. We are now in the setting of Theorem II.7: considering τ_j as our terminal time, we have $Z_{\tau_j}^n \to Z_{\tau_j}^\infty$ in probability. Note that Theorem II.7 can be applied to the terminal time $\tau_j \leq T$ by considering, for example, $(Z^n)^{\tau_j}$ defined on the time interval $[0, T]$. So, applying Theorem II.7, we deduce that for each $j < \infty$, $u^{n,j} \to u^{\infty,j}$.

Next, we claim that

$$\liminf_{n\to\infty} u^n \geq u^\infty. \tag{2.14}$$

First choose $\epsilon > 0$ and $J = J(\epsilon)$ sufficiently large so that $u^{\infty,J} \geq u^{\infty} - \epsilon$. Next choose $N = N(J)$ such that, for $n \geq N$, $|u^{n,J} - u^{\infty,J}| < \epsilon$. The triangle inequality implies that $|u^{n,j} - u^{\infty}| < 2\epsilon$. Since $u^n \geq u^{n,j}$, it follows that for $n \geq N$, $u^n \geq u^{\infty} - 2\epsilon$. In other words, $\liminf_{n\to\infty} u^n \geq u^{\infty}$, and (2.14) is established. $\square$

We now obtain:

Proof of Theorem II.5. By Lemma II.27, the mapping is lower semi-continuous, and by Proposition II.28, the mapping is upper semi-continuous. Together, these imply the theorem. $\square$

Proof of Theorem II.6. As before, for each $k \in \mathbb{N}$, let $\widehat{X}_T^{n,k}$ be the optimal terminal wealth in the *nth* market that satisfies the constraint $\widehat{X}_T^{n,k} > -k$. By Step 7 in the proof of Theorem 2.2 of [50], we know that as $k \to \infty$, $U(\widehat{X}_T^{n,k}) \to U(\widehat{X}_T^n)$ in L^1 for each $n \in \mathbb{N}$. As a consequence of Proposition II.25, $E[U(\widehat{X}_T^{n,k})] \uparrow E[U(\widehat{X}_T^n)]$ as $k \to \infty$, and the convergence is uniform over n. As $U(\cdot)$ is nonpositive, Scheffe's Lemma then implies that $U(\widehat{X}_T^{n,k}) \to U(\widehat{X}_T^n)$ in L^1 as $k \to \infty$, *uniformly over n*. L^1 convergence being stronger than L^0 convergence, we also have that $U(\widehat{X}_T^{n,k}) \to U(\widehat{X}_T^n)$ in probability as $k \to \infty$, uniformly over n. Since $\widehat{X}_T^{n,k} \to \widehat{X}_T^{\infty,k}$ in probability as $n \to \infty$ for all k, then by Lemma 3.10 of [36], it follows that $U(\widehat{X}_T^n) \to U(\widehat{X}_T^{\infty})$ in probability. Since U is bounded from above, we need a little more work to show that $\widehat{X}_T^n \to \widehat{X}_T^{\infty}$ in probability.

We claim now that $\{\widehat{X}_T^n\}_{n\in\mathbb{N}}$ is bounded in probability. Note that $Z^n\widehat{X}^n$ is a martingale, so $E[Z_T^n\widehat{X}_T^n] = 0$. By Proposition II.21, $\{U(\widehat{X}_T^n)\}_{n\in\mathbb{N}}$ is uniformly integrable, and V-compactness implies that $\{V(Z_T^n)\}_{n\in\mathbb{N}}$ is uniformly integrable. The duality relationship $Z_T^n\widehat{X}_T^n \geq U(\widehat{X}_T^n) - V(Z_T^n)$ now implies that the negative parts $\{(Z_T^n\widehat{X}_T^n)^-\}_{n\in\mathbb{N}}$ are uniformly integrable. Hence $\{Z_T^n\widehat{X}_T^n\}_{n\in\mathbb{N}}$ is bounded in L^1, and also in L^0. But $Z_T^n \to Z_T^\infty$ in probability, and Z_T^∞ is strictly positive. Hence $\{Z_T^n\}_{n\in\mathbb{N}}$ is bounded away from zero in probability, and it follows that $\{\widehat{X}_T^n\}_{n\in\mathbb{N}}$ is bounded in probability.

Suppose now that $\widehat{X}_T^n$ does not converge to $\widehat{X}_T^\infty$ in probability. Then there exists an $\epsilon > 0$ such that for infinitely many n, $P(|\widehat{X}_T^n - \widehat{X}_T^\infty| > \epsilon) > \epsilon$. Now, choose a compact set K such that $P(\widehat{X}_T^n \notin K) < \frac{\epsilon}{4}$ for all n. Then $P\left(|\widehat{X}_T^n - \widehat{X}_T^\infty| > \epsilon, \text{ and } \widehat{X}_T^n, \widehat{X}_T^\infty \in K\right) > \frac{\epsilon}{2}$. For $x, y \in K$, there exists a constant $c > 0$ such that $|U(x) - U(y)| > c|x - y|$, due to the fact that $U'(x)$ is positive and bounded away from zero on the compact set K. Thus, it follows that for infinitely many n, $P(|U(\widehat{X}_T^n) - U(\widehat{X}_T^\infty)| > c\epsilon) > \frac{\epsilon}{2}$, contradicting the fact that $U(\widehat{X}_T^n) \to U(\widehat{X}_T^\infty)$ in probability. $\square$

II.6 On Assumptions II.2 and II.3

II.6.1 Comparison of II.3 with the Half-Line setting

Recall that in addition to the V-compactness assumption II.2, we required Assumption II.3, which has no direct analog in [36]. Reviewing Proposition II.21, one sees that the purpose of this assumption was to ensure that (locally) the set $\left\{\exp\left(-\widehat{X}_\tau^n\right) : n \in \mathbb{N}, \tau \in \mathcal{T}\right\}$ is uniformly integrable. More precisely, when we say locally, we mean that there exists a sequence of stopping times $\tau_j \uparrow T$ such that $\left\{\exp\left(-\widehat{X}_{\tau\wedge\tau_j}^n\right) : n \in \mathbb{N}, \tau \in \mathcal{T}\right\}$ is uniformly integrable for each j.

Indeed, Assumption II.3 could be weakened so that it is exactly this condition: let $C_j \triangleq \operatorname*{ess\,sup}_{\tau_j \geq \tau \in \mathcal{T}, n\in\mathbb{N}} E\left[\frac{Z_{\tau_j}^n}{Z_\tau^n} \log \frac{Z_{\tau_j}^n}{Z_\tau^n} | \mathcal{F}_\tau\right]$, which is uniformly bounded thanks to Assumption II.3 and Proposition II.12. We have, as in the proof of Proposition II.21, for $\tau \leq \tau_j$, the duality relationship

$$(2.15) \qquad \exp(-\widehat{X}^n_\tau) = \underbrace{\frac{1}{c_n}\widehat{Z}^n_\tau \exp\left(E\left[\frac{\widehat{Z}^n_{\tau_j}}{\widehat{Z}^n_\tau}\log\frac{\widehat{Z}^n_{\tau_j}}{\widehat{Z}^n_\tau}\ |\mathcal{F}_\tau\right]\right)}_{\leq \frac{1}{c_n}\widehat{Z}^n_\tau \exp(C_j) \text{ by } \textit{Assumption II.3}}.$$

Now, the actual structural condition we need to prove our main results is that $\left\{\exp\left(-\widehat{X}^n_{\tau\wedge\tau_j}\right) : n \in \mathbb{N}, \tau \in \mathcal{T}\right\}$ is uniformly integrable. By V-compactness, the set $\{\widehat{Z}^n_\tau : \tau \in \mathcal{T}, n \in \mathbb{N}\}$ is uniformly integrable, as in Proposition II.21. Therefore, from considering (2.15), we see that Assumption II.3 implies the uniform integrability of $\left\{\exp\left(-\widehat{X}^n_{\tau\wedge\tau_j}\right) : n \in \mathbb{N}, \tau \in \mathcal{T}\right\}$; additionally, we see that Assumption II.3 is a slightly stronger hypothesis than the required property of uniform integrability of $\left\{\exp\left(-\widehat{X}^n_{\tau\wedge\tau_j}\right) : n \in \mathbb{N}, \tau \in \mathcal{T}\right\}$ for each j.

This uniform integrability condition is useful because it allows for processes $\widehat{X}^{n,k}$ such that $\lim_{k\to\infty} E\left[-\exp\left(-\widehat{X}^{n,k}_T\right)\right] = E\left[-\exp\left(-\widehat{X}^n_T\right)\right]$, uniformly over n, and $\sup_{0\leq t\leq T, n\in\mathbb{N}} \exp\left(-\widehat{X}^{n,k}_t\right) \in L^\infty$. Here, we show that, in the setting of utility maximization with a utility function $\widetilde{U}$ defined on $\mathbb{R}_+$, this uniform approximation property is already implied by the $\widetilde{V}$-compactness assumption, with $\widetilde{V}$ the conjugate of $\widetilde{U}$. The proof of this fact is interesting because it mirrors, in our opinion, the essential technical step of [36], Corollary 3.4.

Proposition II.29. *In the setting of [36], let $\{Z^n : n = 1, 2, \ldots, \infty\}$ define a $\widetilde{V}$-compact sequence of markets with $Z^n_T \to Z^\infty_T$ in probability. Fix an initial wealth x_0 , and let $\widetilde{X}^n$ be the optimal wealth process starting from x_0 in the n^{th} market. Then there exist wealth processes $\widetilde{X}^{n,k}$, each defined in the n^{th} market, such that $E\left[\widetilde{U}\left(\widetilde{X}^{n,k}_T\right)\right] \to E\left[\widetilde{U}\left(\widetilde{X}^n_T\right)\right]$ as $k \to \infty$, uniformly over all n, and $\sup_{0\leq t\leq T, n\in\mathbb{N}} \widetilde{U}^-\left(\widetilde{X}^{n,k}_t\right) \in L^\infty$.*

Proof. The proposition uses a simple construction, inspired by [36]. Given $\widetilde{X}^n$, define

$$\widetilde{X}^{n,k} \triangleq \frac{1}{k}x_0 + \frac{k-1}{k}\widetilde{X}^n.$$

In other words, the wealth process $\widetilde{X}^{n,k}$ follows the optimal trajectory, except that a small portion is set aside and left in the riskless asset. The concavity of $\widetilde{U}$ implies that

$$\begin{aligned} &\frac{1}{k}\widetilde{U}(x_0) + \frac{k-1}{k}E\left[\widetilde{U}\left(\widetilde{X}_T^n\right)\right] \\ &\leq E\left[\widetilde{U}\left(\widetilde{X}_T^{n,k}\right)\right] \\ &\leq E\left[\widetilde{U}\left(\widetilde{X}_T^n\right)\right]. \end{aligned} \tag{2.16}$$

Since $\widetilde{V}$-compactness implies that the collection $\left\{\widetilde{U}\left(\widetilde{X}_T^n\right) : n \in \mathbb{N}\right\}$ is bounded in L^1, the uniform approximation property is established in (2.16). Next, each wealth process $\widetilde{X}^n$ is strictly positive, and therefore $\widetilde{X}^{n,k} > \frac{1}{k}$. Consequently

$$\sup_{0\leq t\leq T, n\in\mathbb{N}} \widetilde{U}^-\left(\widetilde{X}_t^{n,k}\right) < \widetilde{U}^-\left(\frac{1}{k}\right).$$

□

II.6.2 Economic interpretation of Assumption II.3

Consider a generic market with dynamics $S = M + \int \lambda d\langle M\rangle$ and associated minimal martingale measure $Z = \mathcal{E}(-\lambda\cdot M)$. In this market, consider the opportunity process L_t^{exp}, introduced in [40], and used in [41]. It is the utility value process normalized by the optimal wealth process, and it exists as a consequence of the homogeneity of power and exponential utilities, and their associated optimal wealth processes. In the notation of [40], the opportunity process L_t^{exp} satisfies

$$V_t(\theta) = \exp(-G_t(\theta))L_t^{exp},$$

where $V_t(\theta)$ represents the indirect utility arising from following the trading strategy θ up to time t, and $G_t(\theta)$ is the wealth resulting from trading according to θ up to time t. As the name suggests, L_t^{exp} describes how much utility can be attained per unit of wealth. Then equation (6.6) of [40] establishes a relationship between L^{exp} and $\widehat{Z}$:

$$-\log(L_t^{exp}) = E\left[V\left(\frac{\widehat{Z}_T}{\widehat{Z}_t}\right)|\mathcal{F}_t\right]. \tag{2.17}$$

Frequently in this chapter, we have concerned ourselves with the size of the right hand side of (2.17): specifically, the *bmo* hypothesis has been used to establish a uniform upper bound on this term over t. This implies that the value processes L_t^{exp} are uniformly bounded away from zero. In economic terms, this puts a constraint on how attractive the investment opportunities can be in our sequence of markets. If the opportunity process is close to zero, this means that an optimal investing agent is relatively unconcerned with having very negative wealth, in that L_t^{exp} counteracts the size of $\exp(-G_t(\theta))$ (note that we wish to maximize $V_t(\theta)$, which is negative).

Note, however, that in (2.17), what matters is the $\mathcal{R}_{LLogL}$ constant $K_{LLogL}(\widehat{Z})$ for the optimal dual variable $\widehat{Z}$, while our regularity Assumption II.3 involves the $\mathcal{R}_{LLogL}$ constant $K_{LLogL}(Z)$ for Z, the minimal martingale measure. By Lemma II.14, we know that $K_{LLogL}(\widehat{Z}) \leq K_{LLogL}(Z)$. More interesting is the claim that the sizes of $K_{LLogL}(\widehat{Z})$ in fact control the sizes of $K_{LLogL}(Z)$, a result which we will establish in Proposition II.30 below. Thus, the $\mathcal{R}_{LLogL}$ constants bind the sizes of the minimal entropy martingale and minimal martingale in a substantive way. In general, the dual object we are interested in is the minimal entropy martingale, while the dual object which we can describe most explicitly is the minimal martingale. The claim above, however, implies that the ostensibly more restrictive act of placing a regularity assumption on the minimal martingales is essentially equivalent to placing one on the minimal entropy martingales, both implying control over the size of the opportunity process.

Proposition II.30. *Let S^n, $n \geq 1$, describe a sequence of markets, with minimal martingales Z^n and minimal entropy martingales $\widehat{Z}^n$. Then*

$$\sup_n K_{LLogL}(Z^n) < \infty \textit{ if and only if } \sup_n K_{LLogL}(\widehat{Z}^n) < \infty.$$

Proof. The "$\Rightarrow$" direction is trivial, given Lemma II.14. We therefore address the "$\Leftarrow$" condition. Write $\widehat{Z}^n = \mathcal{E}(\widehat{R}^n) = \mathcal{E}(-\lambda^n \cdot M + \widehat{L}^n)$, where $\widehat{L}^n$ is a local martingale orthogonal to M. Thus, $\langle \widehat{R}^n \rangle = \langle -\lambda \cdot M \rangle + \langle \widehat{L}^n \rangle$. As a consequence,

$$|| - \lambda \cdot M||_{bmo_2} \leq ||\widehat{R}^n||_{bmo_2}.$$

According to the proof of Lemma II.9, found in Propositions 5 and 6 of [21], there exists an increasing function $f : \mathbb{R}_+ \to \mathbb{R}_+$ such that for a continuous martingale M, $||M||_{bmo_2} \leq x$ implies that $K_{LLogL}(\mathcal{E}(M)) \leq f(x)$. Therefore, for each n, $K_{LLogL}(Z^n) \leq f(||\lambda^n \cdot M||_{bmo_2}) \leq f(||\widehat{R}^n||_{bmo_2})$. Taking suprema over n, we have

$$\mathop{sup}_n K_{LLogL}(Z^n) \leq \mathop{sup}_n f(||\widehat{R}^n||_{bmo_2}) \triangleq R^* < \infty,$$

with the finite constant R^* existing by hypothesis.

□

II.6.3 On Assumption II.2

Here, we illustrate the necessity of the V-compactness hypothesis with a few examples.

Lemma II.31. *Suppose that $Z^n_T \to Z^\infty_T$ in probability and that $\{Z^n_T : n \in \mathbb{N} \cup \{\infty\}\}$ is V-compact. Additionally, suppose that $Z^\infty_T = \widehat{Z}^\infty_T$, i.e. the terminal values of the minimal martingale measure and minimal entropy martingale measure coincide. Then $\lim_{n\to\infty} v^n(y) = v^\infty(y)$. Hence, $\lim_{n\to\infty} u^n(x) = u^\infty(x)$.*

Proof. Thanks to Lemma II.27, it suffices to show that $\limsup_{n\to\infty} v^n(y) \le v^\infty(y)$. By hypothesis, $E[V(yZ_T^n)] \to E[V(yZ_T^\infty)] = v^\infty(y)$ as $n \to \infty$. But $v^n(y) \le E[V(yZ_T^n)]$. Therefore, $\limsup_{n\to\infty} v^n(y) \le \lim_{n\to\infty} E[V(yZ_T^n)] = v^\infty(y)$. The last claim in the lemma, that $u^n \to u^\infty$, follows from the duality between v^n and u^n, see Proposition 3.9 of [36]. □

Corollary II.32. *Suppose that $Z_T^n \to Z_T^\infty$ in probability and that the limiting market is complete. Then $\lim_{n\to\infty} u^n(x) = u^\infty(x)$ if and only if $\{Z_T^n : n \in \mathbb{N}\}$ is V-compact.*

Proof. For the "if" direction, note that in a complete market there is only one equivalent martingale measure, and hence trivially the minimal martingale measure and minimal entropy martingale must agree. Therefore, by Lemma II.31, $\lim_{n\to\infty} u^n(x) = u^\infty(x)$. The "only if" direction is identical to the proof of Proposition 2.13 of [36]. □

Remark II.33. We also note that there are examples of incomplete markets where the minimal martingale and minimal entropy martingale agree; in these cases it is also clear that V-compactness is necessary and sufficient. This is the case in a market when one tries to hedge an option written on a non-tradeable asset using a geometric Brownian motion correlated with that asset; see e.g. Section 4 of [28].

Appendix A Continuity and Uniform Approximation

In this appendix, we address the first claim made in Remark II.24. Its proof requires a bit of preparatory work.

Proposition II.34. *$u^n \to u^\infty$ if and only if $u^{(n,k)} \to u^n$ as $k \to \infty$, uniformly over n.*

Proof. The " $\Leftarrow$ " implication was the content of Theorem II.7. For the other direction, let $\mathbb{N}^*$ be the space $\{1, 2, \ldots, \infty\}$, whose topology is the one point compactification of $\mathbb{N}$ with the discrete topology; the open sets of $\mathbb{N}^*$ are the finite subsets of $\mathbb{N}$ and cofinite subsets containing ∞. This space is compact.

For each $k \in \mathbb{N}$, the map $n \mapsto u^{(n,k)}(0)$ is continuous by Lemma II.19. By construction, $u^{(n,k)} \leq u^{(n,k+1)}$ for all n, k, and $u^{(n,k)} \to u^n$ as $k \to \infty$ for all n. Therefore, supposing that $n \mapsto u^n(0)$ is continuous, we apply Dini's Theorem to get the desired result. $\square$

Lemma II.35. *Suppose that $u^n \to u^\infty$. Then $\widehat{X}_T^n \to \widehat{X}_T^\infty$ in probability. Furthermore, $U(\widehat{X}_T^n) \to U(\widehat{X}_T^\infty)$ in L^1.*

Proof. The proof of the first claim is identical to Lemma 3.10 of [36], which establishes the result in the positive wealth case. The second claim follows from Scheffe's Lemma.

$\square$

Let $d(\cdot, \cdot)$ be a metric whose topology is associated with the one corresponding to convergence in probability, i.e.

$$d(X_n, X) \to 0 \text{ if and only if } P(|X_n - X| > \epsilon) \to 0 \text{ for all } \epsilon > 0.$$

Recall that $\widehat{X}^{(n,k)}$ is the optimal wealth process in market n satisfying the constraint $\widehat{X}^{(n,k)} > -k$.

Lemma II.36. *Suppose that $u^n \to u^\infty$. Then*

$$\lim_{k \to \infty} \sup_n d\left(\widehat{X}_T^{(n,k)}, \widehat{X}_T^n\right) = 0.$$

Proof. On p. 708, Step 2 of [50], it is established that, for fixed n, $d\left(\widetilde{Y}_T^{(n,k)}, \widehat{Y}_T^n\right) \to 0$ as $k \to \infty$; recall that $\widehat{Y}^n$ is the minimal dual variable arising from utility maximization with $U : \mathbb{R} \to \mathbb{R}$ in the n^{th} market, and $\widetilde{Y}^{(n,k)}$ is the minimal dual variable arising from utility maximization in the n^{th} market with $\widetilde{U}^{(k)}:\mathbb{R}_+\to\mathbb{R}$, as defined in Section 4. For details, we refer the reader to [50]. A careful reading of this proof yields the fact that the rate of this convergence, for each n, is governed by the rate at which $v^{(n,k)}$ converges to v^n. The hypothesis that $u^n\to u$ is equivalent, by Proposition II.34, to the uniform convergence of $u^{(n,k)}$ to u^n as $k\to\infty$, over all n. By a standard duality argument, this is equivalent to $v^{(n,k)}$ converging to v^n as $k\to\infty$, uniformly over all n. Applying a standard argument based on optimality and strict convexity (see Lemma 3.6 of [34]), it therefore follows that

$$\lim_{k\to\infty}\sup_n d\left(\widetilde{Y}_T^{(n,k)},\widehat{Y}_T^n\right)=0.$$

By duality, we have $U^{(k)\prime}(\widehat{X}_T^{(n,k)})=u^{(n,k)\prime}(0)\widetilde{Y}_T^{(n,k)}$, and the lemma follows. □

Corollary II.37. *Suppose that $u^n\to u^\infty$. Then*

$$\lim_{k\to\infty}\sup_n ||U(\widehat{X}_T^{(n,k)})-U(\widehat{X}_T^n)||_{L^1}=0.$$

Proof. The result follows by applying Lemma II.36 and Proposition II.34, along with Scheffe's Lemma. □

Corollary II.38. *Suppose $u^n\to u^\infty$. Then the set $\{U(\widehat{X}_T^{(n,k)}):n,k\}$ is uniformly integrable.*

Proof. The result follows by applying Lemma II.35 and Corollary II.37. □

Lemma II.39. *Suppose that $u^n\to u^\infty$. Then*

$$\lim_{k\to\infty}\sup_n d\left((\widehat{X}^{(n,k)}-\widehat{X}^n)^*,0\right)=0. \tag{2.18}$$

Proof. Suppose that (2.18) does not hold. Then, there exists a sequence $(n_m,k_m)_{m\geq 1}$ and $\alpha>0$ such that

$$P\left((\widehat{X}^{(n_m,k_m)}-\widehat{X}^{n_m})^*>\alpha\right)>\alpha$$

for each m. Let $\tau_m = \inf\{t \geq 0 : \widehat{X}_t^{(n_m,k_m)} \geq \widehat{X}_t^{n_m} + \alpha\} \wedge T$, and let $\widetilde{\tau}_m = \inf\{t \geq 0 : \widehat{X}_t^{(n_m,k_m)} \leq \widehat{X}_t^{n_m} - \alpha\} \wedge T$. It must be the case that either $P(\tau_m < T) > \frac{\alpha}{2}$ or $P(\widetilde{\tau}_m < T) > \frac{\alpha}{2}$. The treatment of each contingency is similar, and so without loss of generality, we assume that $P(\tau_m < T) > \frac{\alpha}{2}$. Consider the concatenated wealth process $\widetilde{X}_t^{n_m} \triangleq \widehat{X}_{t\wedge\tau_m}^{(n_m,k_m)} + (\widehat{X}_{t\vee\tau_m}^{n_m} - \widehat{X}_{\tau_m}^{n_m})$. For any $\mathbb{Q} \in \mathcal{M}^n$ with finite entropy, $\widehat{X}^{(n_m,k_m)}$ and $\widehat{X}^{n_m}$ are $\mathbb{Q}$-martingales, since they are admissible wealth processes in the sense of Definition II.4. Since the concatenation of martingales yields a martingale, $\widetilde{X}^{n_m}$ is a $\mathbb{Q}$-martingale for any $\mathbb{Q} \in \mathcal{M}^n$ with finite entropy, so this concatenated strategy is still admissible. On the set $\{\tau_m < T\}$, $\widetilde{X}_T^{n_m} \geq \alpha + \widehat{X}_T^{n_m}$, and on the set $\{\tau_m < T\}^c$, $\widetilde{X}_T^{n_m} = \widehat{X}_T^{(n_m,k_m)}$. As in the proof of Theorem II.6, for any $\epsilon > 0$, there is a compact subset $K = K(\epsilon)$ of $\mathbb{R}$ such that

$$\max\left\{P(\widehat{X}_T^{n_m} - \alpha \notin K), P(\widehat{X}_T^{(n_m,k_m)} - \alpha \notin K)\right\} < \epsilon \tag{2.19}$$

for all m, and $U'(x) \geq c = c(\epsilon)$ for $x \in K$. We will fix some $\epsilon < \frac{\alpha}{2}$.

Thus,

$$\begin{aligned}
&E\left[U\left(\widetilde{X}_T^{n_m}\right)\right] \\
&\quad \geq E\left[1_{\{\tau_m<T\}}U\left(\widehat{X}_T^{n_m}+\alpha\right)\right] + E\left[1_{\{\tau_m=T\}}U\left(\widehat{X}_T^{(n_m,k_m)}\right)\right] \\
&\quad \geq E\left[1_{\{\tau_m<T\}}U'\left(\widehat{X}_T^{n_m}+\alpha\right)\cdot\alpha\right] + E\left[1_{\{\tau_m<T\}}U\left(\widehat{X}_T^{n_m}\right)\right] + E\left[1_{\{\tau_m=T\}}U\left(\widehat{X}_T^{(n_m,k_m)}\right)\right].
\end{aligned}$$

By Corollary II.37, we have

$$E\left[1_{\{\tau_m<T\}}U\left(\widehat{X}_T^{n_m}\right)\right] + E\left[1_{\{\tau_m=T\}}U\left(\widehat{X}_T^{(n_m,k_m)}\right)\right] \to E\left[U\left(\widehat{X}_T^{n_m}\right)\right],$$

as $m \to \infty$.

From (2.19), we know that $U'\left(\widehat{X}_T^{n_m}+\alpha\right) \geq c$ up to a set of measure ϵ. We then have

$$\liminf_{m\to\infty}\ \left(E\left[U\left(\widetilde{X}_T^{n_m}\right)\right] - E\left[U\left(\widehat{X}_T^{n_m}\right)\right]\right) \geq \left(\frac{\alpha}{2}-\epsilon\right)c\alpha > 0.$$

This, however, contradicts the optimality of $\widehat{X}_T^{n_m}$ when m is sufficiently large. $\square$

Now we can prove the main result of this section.

Proposition II.40. *Suppose that $u^n \to u^\infty$. Then for each $i > 0$, the set $\{(\widehat{X}^{(n,i)})^* : n \in \mathbb{N}\}$ is bounded in probability.*

Proof. It is true by construction that for each k, $\left\{ \inf_{0 \le t \le T} \widehat{X}_t^{(n,k)} : n \in \mathbb{N} \right\}$ is bounded in probability, since $\widehat{X}^{(n,k)} > -k$. To conclude, it only remains to apply Lemma II.39.

□

Appendix B Uniformly Integrable Wealth Processes: A Brief Counterexample

The next proposition is based directly from an example of [49], which can be easily modified to fit the setting of this chapter. It addresses the second claim made in Remark II.24.

Proposition II.41. *There exists a single market for which the optimal wealth process $(\widehat{X}_t)_{0 \le t \le T}$ does not have $\{\exp(-\widehat{X}_\tau) : \tau \in \mathcal{T}\}$ uniformly integrable.*

Proof. Consider the example introduced on p. 13 of [49]. In that market, it is shown on p. 19 that the optimal wealth process $\widehat{X}$ satisfies $\lim_{t \uparrow T} E[-\exp(-\widehat{X}_t)] = -\infty$. This clearly is not possible if $\{\exp(-\widehat{X}_\tau) : \tau \in \mathcal{T}\}$ is uniformly integrable. □

CHAPTER III

Quickest Search over Brownian Channels

III.1 Introduction

In [35], the following problem is studied: consider countably many sequences $\{Y_k^i : k = 1, 2, \ldots\}$, $i = 1, \ldots$. For each i, $\{Y_k^i : k = 1, 2, \ldots\}$ are random variables which obey one of two hypotheses: under H_0, the Y_k^i are i.i.d. and $Y_k^i \sim Q_0, k = 1, 2, \ldots$, and under H_1, the Y_k^i are i.i.d. and $Y_k^i \sim Q_1, k = 1, 2, \ldots$, where Q_0 and Q_1 are two distinct, but equivalent, distributions. At each discrete time k, one can take one of three actions: stop sampling and choose a channel which is believed to satisfy H_1, continue observation of the *current* channel, or continue observation in a *new* channel. Over all possible observation strategies and their associated stopping times τ, our goal is to minimize $P(H^{s_\tau} = H_0) + cE[\tau]$, where here H^{s_τ} is the true condition of the channel observed at time τ, and c represents the cost of making one observation.

The authors of [35] solve the problem above, in the sense that they find an optimal observation strategy and stopping time, both of which can be computed as hitting times of an underlying posterior process. In this same paper, the authors ask for a solution to the corresponding problem in continuous time, and it is this task which we take up. In the quickest search problem in continuous time, the basic object of study is a sequence of observable processes $\xi_t^i = \theta^i t + B_t^i$, $i \in \mathbb{N}$, where each B^i is an independent Brownian Motion and each θ^i is an independent Bernoulli random variable. For some prior $\hat{\pi} \in (0,1)$ and each i, $P_{\hat{\pi}}(\theta^i = 1) = \hat{\pi}$, and $P_{\hat{\pi}}(\theta^i = 0) = 1-\hat{\pi}$. We say that ξ^i satisfies hypothesis H_0 if $\theta^i = 0$ and ξ^i satisfies hypothesis H_1 if $\theta^i = 1$. The general objective in the quickest search problem is to find, as quickly as possible, a process ξ^i which satisfies H_1. Observing any process for t units of time incurs a cost of ct, where $c > 0$ is a constant. One may observe only one ξ^i at a time, but one can instantaneously change between observed processes at any time. In discrete time, the problem may be approximated by a finite horizon version solvable by backwards induction, but this method is not available in continuous time. Instead, the problem is solved by formulating it as a free boundary ordinary differential equation.

In the literature, there has been a large amount of research into quickest search problems, although the majority of it has been in discrete time. Since the focus of this chapter is a continuous time problem, we refer the reader to the references in [35] for an excellent description of the discrete-time literature, as well as [58] for a multi-channel quickest detection problem in discrete time. The continuous time literature is sparser, but there are several papers addressing a problem similar to ours. The three papers [59], [45], and [33] all consider the quickest search variant in which there are a finite number of channels, and exactly one channel satisfies H_1. For comparison, in the problem we study, there are infinitely many channels and no knowledge of how many channels satisfy H_1. Also in continuous time, in [16], the authors study a multi-channel problem where all channels are observed simultaneously: here the goal is to find the common intensity rate of a collection of Poisson processes. In [10], two Poisson channels are observed simultaneously to determine the one in which disruption occurs first. Other problems involving multiple stopping times include

[14] and [13], although our problem is closer in spirit to a multiple switching problem than one of multiple stopping.

The technical details of continuous time formulations are somewhat subtle, and indeed, in the three papers closest to ours, [59], [45], and [33] , each purports to fix an error in the previous one. For example, in [59], an optimal switching strategy is described in the following way: consider N diffusions γ_t^n, $1 \leq n \leq N$, which are supposed to represent the posterior for each channel, and take the strategy that when γ_t^i is largest among the N processes, observe channel i. The inherent problem in such a strategy is that when, for example, $\gamma_t^i = \gamma_t^j$, the sign of $\gamma_s^i - \gamma_s^j$ will oscillate infinitely many times when s is in a neighborhood of t. So, such a strategy would necessarily switch between channels i and j infinitely many times in that neighborhood, and this is physically unfeasible.

Indeed, even in the discrete time case, certain technical details have not been completely developed. To rigorously describe the set of all observation strategies in our problem, one must talk of different filtrations, since by choosing which channel to observe at different times, we are modulating the exact information which we receive. Therefore, one of the aims of this chapter is describe a continuous time quickest search problem in a mathematically rigorous way. At the same time, however, we can obtain a closed form analytic solution to our quickest search problem, both for the value function and the optimal stopping time, and so from a practical perspective, our results may be useful in the discrete time case if calculating the solution to that problem is too expensive. The theory of extended weak convergence, in [1], provides a possible way to translate insights from continuous time optimal stopping back to discrete time.

The outline of this chapter is as follows: we will first describe how problems in continuous time sequential analysis can be formulated as optimal stopping problems, using the building block of the one channel problem. Much of this material is derived from parts of [43], Chapter 6, and [53], Chapter 4. We will then state an optimal stopping/switching problem which models our quickest detection problem. Its basic feature is that it involves many different filtrations, corresponding to the different ways in which one may observe the processes. We show how this problem may be reduced to an impulse control/stopping problem with a single filtration and a single Brownian Motion, so that the tools of optimal stopping may be applied.

Let us describe this impulse control/stopping problem more closely, since here the exact structure of the problem is evident. Consider a sequence of channels which all have prior probability $\hat{\pi}$ of satisfying H_1. By observing one channel at the time, the posterior process of precisely one of these channels will evolve in time according to a stochastic differential equation, and all the rest will simply stay at $\hat{\pi}$. The goal of the agent is essentially to find some channel and some time when the respective posterior process is very close to 1, ensuring that H_1 is almost certainly satisfied. Given this goal, the agent, when faced with a posterior which has dropped below $\hat{\pi}$, will want to immediately move on to the next channel. The agent must take a finite time to react, so after some amount of time ϵ, or more realistically, when the posterior process hits the level $\hat{\pi} - \epsilon$, he switches channels, and the effect of this is that the posterior process is reset to $\hat{\pi}$. These are the impulse controls available to the agent: at any time, he can move "his" posterior process back to $\hat{\pi}$. Now, in the limit, the agent ideally wants to react as quickly as possible, which means $\epsilon \downarrow 0$. In this limit, we show that these impulsed processes converge in an appropriate sense to a diffusion with reflecting boundary at $\hat{\pi}$, so that the original problem may be stated as an optimal stopping problem on this reflected process.

Finally, we solve this optimal stopping problem using a standard verification theorem. Using the results derived thus far, we outline ϵ-optimal algorithms for the quickest search problem, and provide some computations of optimal threshold levels for different parameters.

III.2 From Sequential Analysis to Optimal Stopping

Let $(\Omega, \mathcal{F}, P_{\hat{\pi}})$ be a probability space supporting a Brownian Motion B and an independent Bernoulli random variable θ, with $P_{\hat{\pi}}(\theta = 1) = \hat{\pi}$. We consider two statistical hypotheses

$$H_1 : \theta = 1, \text{ and } H_0 : \theta = 0,$$

which have respective prior probabilities $\hat{\pi}$ and $1 - \hat{\pi}$. We let $\xi_t = \theta t + B_t$ model the observed process, with induced filtration $\mathbb{F}^\xi = \left(\mathcal{F}_t^\xi\right)_{t\geq 0}$. We will find it useful to write $P_{\hat{\pi}} = \hat{\pi} P_1 + (1 - \hat{\pi}) P_0$, where under P_1, ξ_t is a Brownian Motion with unit drift, and under P_0 it is a standard Brownian Motion with zero drift.

Based on the continuous observation of ξ, our goal is to choose a sequential decision rule (τ, d), where τ is a $\mathbb{F}^\xi$ stopping time, and d is $\mathcal{F}_\tau^\xi$-measurable and equal to one or zero. Taking $d = 1$ models accepting H_1 at time τ, and taking $d = 0$ models accepting H_0. The goal is to minimize the risk function

$$V(\pi) \triangleq \inf_{(\tau,d)} E_\pi \left[\tau + a 1_{\{d=0,\theta=1\}} + b 1_{\{d=1,\theta=0\}}\right], \tag{3.1}$$

where a and b are used to weight the importance of each misidentification. Let $\pi_t \triangleq P_{\hat{\pi}}(\theta = 1 | \mathcal{F}_t^\xi)$ be the posterior process, which models our belief that H_1 is satisfied in the current channel, based on observations up to time t, $\mathcal{F}_t^\xi$. Given a stopping time τ, the choice of whether to set $d = 0$ or $d = 1$ is completely described by the value of π_τ. For $c = \frac{b}{a+b}$, if $\pi_\tau \leq c$, $d = 0$, and $d = 1$ otherwise. Therefore, (3.1) may be restated as an optimal stopping problem on π_t:

$$V(\pi) = \inf_{\tau} E_\pi \left[\tau + a\pi_\tau \wedge b(1-\pi_\tau)\right].$$

We have reduced the sequential analysis problem to one of optimal stopping, but it remains to understand the time dynamics of the posterior process π_t. To that end, we introduce the odds process

$$\Phi_t \triangleq \frac{dP_1}{dP_0}\Big|_{\mathcal{F}_t^\xi}(\omega).$$

By an application of Bayes' rule, Φ_t can be calculated, and is equal to $\exp\left\{\xi_t - \frac{t}{2}\right\}$. From [53], p. 181, the posterior process π_t satisfies

$$\pi_t = \hat{\pi}\frac{dP_1}{d[\hat{\pi}P_1 + (1-\hat{\pi})P_0]}\Big|_{\mathcal{F}_t^\xi}(\omega);$$

using this equation, we can relate π_t to Φ_t.

We have

$$\pi_t = \frac{\frac{\hat{\pi}}{1-\hat{\pi}}\Phi_t}{1 + \frac{\hat{\pi}}{1-\hat{\pi}}\Phi_t}.$$

Now Φ_t clearly satisfies $d\Phi_t = \Phi_t d\xi_t$ by Itô's Lemma. We have a formula for π_t in terms of Φ_t, so using Itô's Lemma again, we derive

$$d\pi_t = -(\pi_t)^2(1-\pi_t)dt + \pi_t(1-\pi_t)d\xi_t \tag{3.2}$$

with initial condition $\pi_0 = \hat{\pi}$. Consider the process $W_t \triangleq \xi_t - \int_0^t E_{\hat{\pi}}[\theta|\mathcal{F}_s^\xi]ds = \xi_t - \int_0^t \pi_s ds$. This process is immediately seen to be a martingale, and it has continuous paths by construction. It is clear that $\langle W\rangle_t = t$, and so Lévy's Theorem implies that W is a Brownian Motion.

We may therefore rewrite (3.2) as

$$d\pi_t = \pi_t(1-\pi_t)dW_t, \pi_0 = \hat{\pi}. \tag{3.3}$$

The optimal stopping problem on (3.3) will play a central role in our analysis.

III.3 Reduction to a Problem with a Single Filtration

Let Ω^i, $i \in \mathbb{N}$, denote $C[0,\infty)$. We consider a sequence of independent Brownian Motions $W^1, W^2, \ldots$, defined canonically as the coordinate process on $\Omega^{\mathbb{N}} = \Omega^1 \times \Omega^2 \times \cdots$: for $(\omega^1, \omega^2, \ldots) \in \Omega^{\mathbb{N}}$, $W^i_t(\omega^1, \omega^2, \ldots) = \omega^i_t$. Let $\mathbb{F}$ be the filtration generated by the canonical coordinate process. Let $P^{\mathbb{N}}$ denote the product measure of Wiener measures on $\Omega^{\mathbb{N}}$.

We define, for each i, and for any random time ϕ independent of W^i, the process $\pi^{i,\phi}$ satisfying

$$d\pi^{i,\phi}_t = \pi^{i,\phi}_t(1-\pi^{i,\phi}_t)dW^i_t, \text{ for } t \geq \phi, \pi^{i,\phi}_\phi = \hat{\pi}. \tag{3.4}$$

The value $\pi^{i,\phi}_t$ represents the posterior probability, based on observing the history of Channel i from time ϕ up until time t, that Channel i satisfies H_1. Although ϕ here is arbitrary, it will always be, for our purposes, a switching time between two channels. We will take these stochastic differential equations as our main object of study. Now, given that the single channel sequential analysis problem involves the optimal stopping of a single copy of the SDE (3.3), it stands to reason that a multi channel problem should involve a sequence of the processes in (3.4) which are concatenated together according to the way in which we observe each channel. We describe this procedure now.

Let $\mathbb{F}^{(1)}$ be the filtration on $\Omega^{\mathbb{N}}$ generated by W^1, which coincides with the filtration generated by $\pi^1 \triangleq \pi^{1,0}$. We let $\mathcal{T}^{(1)}$ be the set of $\mathbb{F}^{(1)}$-stopping times.

Let $\mathfrak{S}$ be the set of admissible switching controls, which will determine the set of times when the currently observed channel is changed. Elements of $\mathfrak{S}$ will consist of sequences of increasing random times $\{\phi_1, \phi_2, \ldots\}$ with $\phi_1 = 0$. The time ϕ_i, $i \geq 2$, may be interpreted as the time when observation of Channel $i-1$ stops and observation of Channel i begins. The main property that each ϕ_i should have is that it should be measurable with respect to the information gathered before it: our decision to switch should be based on what we have seen thus far. In order to make this precise, we will define elements of $\mathfrak{S}$ inductively: given the first n switching times $\phi_1, \phi_2, \ldots, \phi_n$, we will define an allowed $(n+1)^{st}$ switching time. First, the base case. We let $\phi_1 = 0$.

The first possible switching time ϕ_2 is any strictly positive $\mathbb{F}^{(1)}$-stopping time. So, given ϕ_2, we define the process $\pi^{(2),\phi_2}$ as follows:

$$\pi_t^{(2),\phi_2} = \pi_t^1 1_{\{t<\phi_2\}} + \pi_t^{2,\phi_2} 1_{\{t\geq\phi_2\}}.$$

The process $\pi^{(2),\phi_2}$ generates a filtration $\mathbb{F}^{(2),\phi_2}$. We let $\mathcal{T}^{(2),\phi_2}$ denote the set of $\mathbb{F}^{(2),\phi_2}$-stopping times.

Now, we define what the switching time ϕ_3 may look like, given that ϕ_2 has already been chosen. Such a switching time is any $\phi_3 \in \mathcal{T}^{(2),\phi_2}$ such that $\phi_3 > \phi_2$. Given ϕ_2 and ϕ_3, we define the process $\pi^{(3),\phi_2,\phi_3}$ as follows:

$$\pi_t^{(3),\phi_2,\phi_3} = \pi_t^{(2),\phi_2} 1_{\{t<\phi_3\}} + \pi_t^{3,\phi_3} 1_{\{t\geq\phi_3\}}.$$

The process $\pi^{(3),\phi_2,\phi_3}$ generates a filtration $\mathbb{F}^{(3),\phi_2,\phi_3}$, and $\mathcal{T}^{(3),\phi_2,\phi_3}$ denotes the set of $\mathbb{F}^{(3),\phi_2,\phi_3}$-stopping times. Proceeding in this way, we define, for each $n \in \mathbb{N}$, $\pi^{(n),\phi_2,\ldots,\phi_n}, \mathbb{F}^{(n),\phi_2,\ldots,\phi_n}$, and $\mathcal{T}^{(n),\phi_2,\ldots,\phi_n}$. These are, respectively, the posterior process, filtration, and stopping times which result from switching channels at times $\phi_2, \phi_3, \ldots, \phi_n$.

Definition III.1. Let $\Phi = \{\phi_1, \phi_2, \ldots\}$ be a sequence of random times such that $\phi_i > \phi_{i-1}$ on the set $\{\phi_{i-1} < \infty\}$, and such that $\lim_i \phi_i = \infty$. We say that Φ is an admissible switching control, writing $\Phi \in \mathfrak{S}$, if $\phi_1 = 0$, $\phi_2 \in \mathcal{T}^{(1)}$, and for $n \geq 2$, $\phi_n \in \mathcal{T}^{(n-1),\phi_2,\ldots,\phi_{n-1}}$.

Each $\Phi = \{\phi_1, \phi_2, \ldots\} \in \mathfrak{S}$ induces an observed posterior process π^Φ, defined as follows:

$$\pi_t^\Phi = \pi_t^{(n),\phi_2,\ldots,\phi_n} \text{ on the set } \{\phi_n \le t < \phi_{n+1}\}. \tag{3.5}$$

Intuitively, in comparison with (3.4), π_t^Φ represents the posterior probability that at time t, the channel currently being observed under the observation strategy Φ satisfies hypothesis H_1. If the same channel was always observed, π^Φ would behave exactly like $\pi^{1,\phi=0}$. As it is, when the channel is switched, the effect on the posterior is a sudden jump back to the original level $\hat{\pi}$.

The process π^Φ induces the filtration $\mathbb{F}^\Phi$ along with $\mathcal{T}^\Phi$, the set of $\mathbb{F}^\Phi$-stopping times. We define the value function as follows:

$$\begin{aligned} V_{\hat{\pi}} &\triangleq \inf_{\Phi\in\mathfrak{S}} V_{\hat{\pi}}^\Phi \\ &\triangleq \inf_{\Phi\in\mathfrak{S}} \inf_{\tau\in\mathcal{T}^\Phi} E[c\tau + (1-\pi_\tau^\Phi)]. \end{aligned} \tag{3.6}$$

Note that in the value function (3.6), there is only a $(1-\pi_t)$ term, instead of $\pi_t \wedge (1-\pi_t)$. This reflects the fact that instead of deciding whether a single channel satisfies H_1 or H_0, one is looking only for a channel which satisfies H_1. Now, the Φ also induces a process W^Φ, which is defined as follows on $\Omega^{\mathbb{N}}$:

$$W_t^\Phi(\omega^1, \omega^2, \ldots) \triangleq \omega_t^1 \text{ on the set } \{t < \phi_2\}, \tag{3.7}$$

and for $n \ge 2$

$$W_t^\Phi(\omega^1, \omega^2, \ldots) \triangleq W_{\phi_n-}^\Phi(\omega^1, \omega^2, \ldots) + (\omega_t^n - \omega_{\phi_n}^n) \text{ on the set } \{\phi_n \le t < \phi_{n+1}\}. \tag{3.8}$$

We prove the following standard fact:

Lemma III.2. *For each $\Phi \in \mathfrak{S}$, W^Φ is a Brownian Motion.*

Proof. We first prove that W^{Φ} is a martingale. Note that W_t^{Φ} can be written as

$$W_t^{\Phi} = \sum_{i=1}^{\infty} \left(W^i_{t \wedge \phi_{i+1}} - W^i_{t\wedge \phi_i} \right).$$

Additionally,

$$\begin{aligned} E[(W_t^{\Phi})^2] &= \sum_{i=1}^{\infty} E\left[\left(W^i_{t \wedge \phi_{i+1}} - W^i_{t\wedge \phi_i} \right)^2 \right] \\ &= \sum_{i=1}^{\infty} (E[t \wedge \phi_{i+1}] - E[t \wedge \phi_i]) \\ &= t, \end{aligned}$$

since $\lim_i \phi_i = \infty$. Therefore, by the Dominated Convergence Theorem,

$$\begin{aligned} E[W_t^{\Phi} | \mathcal{F}_s] &= E\left[\sum_{i=1}^{\infty} \left(W^i_{t \wedge \phi_{i+1}} - W^i_{t\wedge \phi_i} \right) | \mathcal{F}_s \right] \\ &= \sum_{i=1}^{\infty} \left(W^i_{s \wedge \phi_{i+1}} - W^i_{s\wedge \phi_i} \right) \\ &= W_s^{\Phi}, \end{aligned}$$

the second inequality following from Optional Sampling and the fact that each W^i is a martingale. Since $\langle W^i \rangle_t = t$ for all t, a.s., it follows that $\langle W^{\Phi} \rangle_t = t$ for all t, a.s. It is also clear by construction that W^{Φ} has continuous paths. Thus by Lévy's characterization of Brownian Motion, W^{Φ} is a Brownian Motion for each $\Phi \in \mathfrak{S}$. $\square$

The process π^{Φ} has continuous paths, with the exception of jump times at $\phi_2, \phi_3, \ldots$.

Lemma III.3. $\pi_t^{\Phi} = \hat{\pi} + \int_0^t \pi_s^{\Phi}(1-\pi_s^{\Phi}) dW_s^{\Phi} + \sum_{i=1}^{\infty} (\hat{\pi} - \pi^{\Phi}_{\phi_i -}) 1_{\{t \geq \phi_i\}}$

Proof. By (3.5) and (3.4), on $\{\phi_n \leq t < \phi_{n+1}\}$, $\pi_t^{\Phi} = \pi^{\Phi}_{\phi_n} + \int_{\phi_n}^t \pi_s^{\Phi} \left(1 - \pi_s^{\Phi}\right) dW_s^n$. Furthermore, on $\{\phi_n \leq t < \phi_{n+1}\}$, $W_t^{\Phi} - W^{\Phi}_{\phi_n} = W_t^n - W^n_{\phi_n}$ by construction of W^{Φ}. Using the locality of stochastic integration (see [46], the Corollary on p. 62), this implies that $\pi_t^{\Phi} = \pi^{\Phi}_{\phi_n} + \int_{\phi_n}^t \pi_s^{\Phi}(1-\pi_s^{\Phi}) dW_s^{\Phi}$. In particular, $\pi^{\Phi}_{\phi_{n+1}-} = \pi^{\Phi}_{\phi_n} +$

$\int_{\phi_n}^{\phi_{n+1}} \pi_s^\Phi(1-\pi_s^\Phi)dW_s^\Phi$. Finally, $\Delta\pi_{\phi_n}^\Phi \triangleq \pi_{\phi_n}^\Phi - \pi_{\phi_n-}^\Phi = \hat{\pi} - \pi_{\phi_n-}^\Phi$. Then, on $\{\phi_n \le t < \phi_{n+1}\}$,

$$\begin{aligned}
\pi_t^\Phi &= \pi_{\phi_n}^\Phi + \int_{\phi_n}^t \pi_s^\Phi(1-\pi_s^\Phi)dW_s^\Phi \\
&= \pi_{\phi_n-}^\Phi + \Delta\pi_{\phi_n}^\Phi + \int_{\phi_n}^t \pi_s^\Phi(1-\pi_s^\Phi)dW_s^\Phi \\
&= \pi_{\phi_{n-1}}^\Phi + \int_{\phi_{n-1}}^{\phi_n} \pi_s^\Phi(1-\pi_s^\Phi)dW_s^\Phi + \pi_{\phi_n}^\Phi - \pi_{\phi_n-}^\Phi + \int_{\phi_n}^t \pi_s^\Phi(1-\pi_s^\Phi)dW_s^\Phi \\
&= \pi_{\phi_{n-1}}^\Phi + \int_{\phi_{n-1}}^t \pi_s^\Phi(1-\pi_s^\Phi)dW_s^\Phi + \hat{\pi} - \pi_{\phi_n-}^\Phi
\end{aligned}$$

Now, if we repeated apply this procedure, reducing the index n by one each time, we obtain, on $\{\phi_n \le t < \phi_{n+1}\}$,

$$\pi_0^\Phi + \int_0^t \pi_s^\Phi(1-\pi_s^\Phi)dW_s^\Phi + \sum_{i=1}^n(\hat{\pi} - \pi_{\phi_i-}^\Phi) = \hat{\pi} + \int_0^t \pi_s^\Phi(1-\pi_s^\Phi)dW_s^\Phi + \sum_{i=1}^n(\hat{\pi} - \pi_{\phi_i-}^\Phi).$$

□

Let $\overline{\Omega}$ be another copy of the canonical space $C[0,\infty)$ with coordinate process $\overline{W}_t$ and filtration $\overline{\mathbb{F}}$ generated by $\overline{W}$. Let $\overline{P}$ denote Wiener measure on this space. Also, let $\overline{\mathcal{T}}$ denote the set of $\overline{\mathbb{F}}$-stopping times. We would like to reduce the original problem $V_{\hat{\pi}}$ to one where everything uses the *same* Brownian Motion ($\overline{W}$) and *same* filtration ($\overline{\mathbb{F}}$).

Lemma III.4. *Let $\Phi \in \mathfrak{S}$. For any $\mathbb{F}^\Phi$-stopping time τ, there exists a $\overline{\tau} \in \overline{\mathcal{T}}$ such that $W_{\cdot\wedge\tau}^\Phi$ and $\overline{W}_{\cdot\wedge\overline{\tau}}$ are identically distributed as processes. Conversely, for any $\overline{\tau} \in \overline{\mathcal{T}}$, there exists a $\mathbb{F}^\Phi$-stopping time τ such that $W_{\cdot\wedge\tau}^\Phi$ and $\overline{W}_{\cdot\wedge\overline{\tau}}$ are identically distributed as processes.*

Proof. Let $\Phi \in \mathfrak{S}$, and let $\tau \in \mathcal{T}^\Phi$. We have a mapping $W^\Phi : \Omega^{\mathbb{N}} \to \overline{\Omega}$ which is defined according to (3.7) and (3.8). Since $\tau \in \mathcal{T}^\Phi$, it is in particular measurable with respect to the filtration $\mathbb{F}^\Phi$ generated by W^Φ on $\Omega^{\mathbb{N}}$. This implies that for any $\overline{\omega} \in \overline{\Omega}$, τ is constant on $(W^\Phi)^{-1}(\overline{\omega})$.

Thus, we define $\overline{\tau} : \overline{\Omega} \to \mathbb{R}$ as follows. For $\overline{\omega} \in \overline{\Omega}$, choose any $\omega \in (W^{\Phi})^{-1}(\overline{\omega})$, and set $\overline{\tau}(\overline{\omega}) \triangleq \tau(\omega)$. From the discussion in the above paragraph, this definition is well-defined.

We claim that $\overline{\tau}$ is an $\overline{\mathbb{F}}$-stopping time. For $t \in \mathbb{R}$, we have

$$\begin{aligned}\{\overline{\omega} : \overline{\tau}(\overline{\omega}) \leq t\} &= \{\overline{\omega} : \tau(\omega) \leq t \text{ for } \omega \in (W^{\Phi})^{-1}(\overline{\omega})\} \\ &= W^{\Phi}\left(\{\omega : \tau(\omega) \leq t\}\right).\end{aligned}$$

Since $\tau \in \mathcal{T}^{\Phi}$, the set $\{\omega : \tau(\omega) \leq t\} \in \mathcal{F}_t^{\Phi}$. By construction, the mapping $W^{\Phi} : \Omega^{\mathbb{N}} \to \overline{\Omega}$ takes $\mathcal{F}_t^{\Phi}$-measurable sets into $\overline{\mathcal{F}}_t$-measurable sets. Therefore, $\overline{\tau} \in \overline{\mathcal{T}}$.

Next, we claim that $W^{\Phi}_{\cdot\wedge\tau}$ and $\overline{W}_{\cdot\wedge\overline{\tau}}$ are distributed identically as processes. As before, this is essentially a tautology. Let $A \in \overline{\mathcal{F}}_{\infty}$. We have

$$\begin{aligned}\{\overline{\omega} : \overline{W}_{\cdot\wedge\overline{\tau}(\overline{\omega})}(\overline{\omega}) \in A\} &= \{\overline{\omega} : W^{\Phi}_{\cdot\wedge\tau(\omega)}(\omega) \in A \text{ for } \omega \in (W^{\Phi})^{-1}(\overline{\omega})\} \\ &= W^{\Phi}\left(\{\omega : W^{\Phi}_{\cdot\wedge\tau(\omega)}(\omega) \in A\}\right).\end{aligned}$$

Thus, $\overline{P}(\{\overline{\omega} : \overline{W}_{\cdot\wedge\overline{\tau}(\overline{\omega})} \in A\}) = \overline{P}\left(W^{\Phi}\left(\{\omega : W^{\Phi}_{\cdot\wedge\tau(\omega)}(\omega) \in A\}\right)\right)$. Since W^{Φ} is a Brownian Motion, the measure P^{Φ} which W^{Φ} induces on $\overline{\Omega}$ agrees with $\overline{P}$. Thus,

$$\begin{aligned}\overline{P}(\{\overline{\omega} : \overline{W}_{\cdot\wedge\overline{\tau}(\overline{\omega})} \in A\}) &= P^{\Phi}\left(W^{\Phi}\left(\{\omega : W^{\Phi}_{\cdot\wedge\tau(\omega)}(\omega) \in A\}\right)\right) \\ &= P^{\mathbb{N}}\left(\{\omega : W^{\Phi}_{\cdot\wedge\tau(\omega)}(\omega) \in A\}\right).\end{aligned}$$

Thus, $\overline{W}_{\cdot\wedge\overline{\tau}}$ and $W^{\Phi}_{\cdot\wedge\tau}$ are identically distributed.

Conversely, suppose that $\overline{\tau}$ is an $\overline{\mathbb{F}}$-stopping time. Define $\tau : \Omega^{\mathbb{N}} \to \mathbb{R}$ by $\tau = \overline{\tau} \circ W^{\Phi}$. We claim that τ is a stopping time. Let $t \in \mathbb{R}$. Then

$$\begin{aligned}\{\omega : \tau(\omega) \leq t\} &= \{\omega : \overline{\tau}(W^{\Phi}(\omega)) \leq t\} \\ &= (W^{\Phi})^{-1}\left(\{\overline{\omega} : \overline{\tau}(\overline{\omega}) \leq t\}\right).\end{aligned}$$

Since $\overline{\tau}$ is an $\overline{\mathbb{F}}$-stopping time, the set $\{\overline{\omega} : \overline{\tau}(\overline{\omega}) \leq t\} \in \overline{\mathcal{F}}_t$, and $(W^{\Phi})^{-1}\left(\{\overline{\omega} : \overline{\tau}(\overline{\omega}) \leq t\}\right) \in \mathcal{F}_t^{\Phi}$. So, $\tau \in \mathcal{T}^{\Phi}$.

Now we claim that $W^{\Phi}_{\cdot\wedge\tau}$ and $\overline{W}_{\cdot\wedge\overline{\tau}}$ are identically distributed as processes. Let $A \in \overline{\mathcal{F}}^{\infty}$.

$$\begin{aligned}\{\omega : W^{\Phi}_{\cdot\wedge\tau(\omega)} \in A\} &= \{\omega : \overline{W}_{\cdot\wedge\overline{\tau}(\overline{\omega})}(\overline{\omega}) \in A \text{ for } \overline{\omega} \text{ such that } W^{\Phi}(\omega) = \overline{\omega}\} \\ &= (W^{\Phi})^{-1}\left(\{\overline{\omega} : \overline{W}_{\cdot\wedge\overline{\tau}(\overline{\omega})}(\overline{\omega}) \in A\}\right).\end{aligned}$$

As before,

$$\begin{aligned}P^{\mathbb{N}}\left(\{\omega : W^{\Phi}_{\cdot\wedge\tau(\omega)} \in A\}\right) &= P^{\mathbb{N}}\left((W^{\Phi})^{-1}\left(\{\overline{\omega} : \overline{W}_{\cdot\wedge\overline{\tau}(\overline{\omega})}(\overline{\omega}) \in A\}\right)\right) \\ &= P^{\Phi}\left(\{\overline{\omega} : \overline{W}_{\cdot\wedge\overline{\tau}(\overline{\omega})}(\overline{\omega}) \in A\}\right) \\ &= \overline{P}\left(\{\overline{\omega} : \overline{W}_{\cdot\wedge\overline{\tau}(\overline{\omega})}(\overline{\omega}) \in A\}\right),\end{aligned}$$

and so the processes are identically distributed. $\square$

Lemma III.5. *For each $\Phi \in \mathfrak{S}$, there exists a sequence of $\overline{\mathbb{F}}$-stopping times $\overline{\Phi} = \{\overline{\phi_1} = 0, \overline{\phi_2}, \ldots\}$ such that for*

$$\pi^{\overline{\Phi}}_t \triangleq \hat{\pi} + \int_0^t \pi^{\overline{\Phi}}_s(1 - \pi^{\overline{\Phi}}_s)d\overline{W}_s + \sum_{i=1}^{\infty}(\hat{\pi} - \pi^{\overline{\Phi}}_{\overline{\phi}_i-})1_{\{t \geq \overline{\phi_i}\}}, \tag{3.9}$$

$\pi^{\overline{\Phi}}$ is identically distributed with π^{Φ}.

Proof. Let N^{Φ} be the simple point process on $\Omega^{\mathbb{N}}$ which jumps at the $\mathbb{F}^{\Phi}$-stopping times $\phi_1, \phi_2, \ldots$. Let $\overline{\Phi} = \{\overline{\phi}_1, \overline{\phi}_2, \ldots\}$ be a sequence of $\overline{\mathbb{F}}$-stopping times whose existence is guaranteed by Lemma III.4, and let $\overline{N}$ be the simple point process on $\overline{\Omega}$ which jumps at the $\overline{\mathbb{F}}$-stopping times $\overline{\phi}_1, \overline{\phi}_2, \ldots$. According to Lemma III.4, (W^{Φ}, N^{Φ}) and $(\overline{W}, \overline{N})$ are identically distributed as processes. Let $f(x) = x(1-x)$ and let $g(x) = \hat{\pi} - x$. Then π^{Φ} and $\pi^{\overline{\Phi}}$ satisfy the SDE's

$$d\pi^{\Phi}_t = f(\pi^{\Phi}_t)dW^{\Phi}_t + g(\pi^{\Phi}_{t-})dN^{\Phi}_t,$$

and

$$d\pi^{\overline{\Phi}}_t = f(\pi^{\overline{\Phi}}_t)d\overline{W}_t + g(\pi^{\overline{\Phi}}_{t-})d\overline{N}_t.$$

Note that if π^{Φ} starts inside the interval $(0,1)$, then it stays there for all time, and similarly for $\overline{\pi}^{\overline{\Phi}}$. On the interval $(0,1)$, $f(x)$ is bounded and Lipschitz, and the same goes for $g(x)$. By Theorem 9.1 of [27] (see p. $245-6$), the above SDE's have uniqueness in law. Consequently, π^{Φ} and $\overline{\pi}^{\overline{\Phi}}$ are identically distributed. $\square$

The converse is proven similarly using Lemma III.4.

Lemma III.6. *Let $\overline{\Phi} = \{\overline{\phi}_1, \overline{\phi}_2, \ldots\}$ be a collection of $\overline{\mathbb{F}}$-stopping times which increase to infinity, and let $\overline{\Phi}$ induce $\overline{\pi}^{\overline{\Phi}}$ as in (3.9). Then there exists $\Phi \in \mathfrak{S}$ such that π^{Φ} is identically distributed to $\overline{\pi}^{\overline{\Phi}}$.*

Lemma III.7. *We have*

$$V_{\hat{\pi}} = \inf_{\overline{\Phi} \in \overline{\mathfrak{S}}} \inf_{\tau \in \overline{\mathcal{T}}} E[c\tau + (1 - \overline{\pi}^{\overline{\Phi}}_{\tau})], \tag{3.10}$$

where $\overline{\mathfrak{S}}$ is the set of all sequences $\overline{\Phi} = \{\overline{\phi}_1, \overline{\phi}_2, \ldots\}$ of $\overline{\mathbb{F}}$-stopping times which increase to infinity.

Proof. Denote by $\overline{V}_{\hat{\pi}}$ the right side of (3.10). Let $\Phi \in \mathfrak{S}$, and consider the optimal stopping problem $\inf_{\tau \in \mathcal{T}^{\Phi}} E[c\tau + (1 - \pi^{\Phi}_{\tau})]$. By Lemma III.5, the process π^{Φ}_t is distributed identically to $\overline{\pi}^{\overline{\Phi}}$. According to Lemma 2.3 of [39], the value function associated to the optimal stopping of a process in its natural filtration depends only on that process's distribution. Therefore $\inf_{\tau \in \mathcal{T}^{\Phi}} E[c\tau + (1 - \pi^{\Phi}_{\tau})] = \inf_{\tau \in \overline{\mathcal{T}}} E[c\tau + (1 - \overline{\pi}^{\overline{\Phi}}_{\tau})] \geq \overline{V}_{\hat{\pi}}$. Taking the infimum over all $\Phi \in \mathfrak{S}$, we obtain

$$V_{\hat{\pi}} \geq \overline{V}_{\hat{\pi}}.$$

Now, let $\overline{\Phi} \in \overline{\mathfrak{S}}$. By Lemma III.6, there exists $\Phi \in \mathfrak{S}$ such that π^{Φ} is identically distributed with $\overline{\pi}^{\overline{\Phi}}$. So, using the same reasoning as above and taking the infimum over all $\overline{\Phi}$, we obtain

$$\overline{V}_{\hat{\pi}} \geq V_{\hat{\pi}}.$$

□

III.4 Working with the New Problem, and Reduction to an Optimal Stopping Problem

From now on, we will drop the overline notation, and simply write $\mathfrak{S}, \Phi, \pi_t, W, \mathbb{F}, \mathcal{T}$, for, respectively, the set of allowed switching strategies, an arbitrary switching strategy, the posterior process, the single Brownian Motion W, the filtration induced by W, and the stopping times for that filtration.

Let π^0 denote the posterior process when there is no switching. In other words, π^0 satisfies the SDE $d\pi_t^0 = \pi_t^0(1-\pi_t^0)dW_t$ along with $\pi^0 = \hat{\pi}$. We next define the reflected process π^r with boundary at $\hat{\pi}$:

$$d\pi_t^r = \pi_t^r(1-\pi_t^r)dW_t + dA_t, \tag{3.11}$$

where A_t is continuous, non-decreasing, flat off of $\pi^r = \hat{\pi}$, $A_0 = 0$.

We also have an optimal stopping problem associated with π^r:

$$V_{\hat{\pi}}^r \triangleq \inf_{\tau \in \mathcal{T}} E[c\tau + (1-\pi_\tau^r)] \tag{3.12}$$

Lemma III.8. $V_{\hat{\pi}}^r \leq V_{\hat{\pi}}$.

Proof. Let $\Phi = \{\phi_1, \phi_2, \ldots\} \in \mathfrak{S}$. Fix i; we will show that $\pi_t^r \geq \pi_t^\Phi$ on $[\phi_i, \phi_{i+1})$. By construction, $\pi_{\phi_i}^\Phi = \hat{\pi}$, and on the interval $[\phi_i, \phi_{i+1})$, the dynamics of π^Φ are described by the diffusion $d\pi_t^\Phi = \pi_t^\Phi(1-\pi_t^\Phi)dW_t$. Let π^{0,ϕ_i} be the un-switched diffusion starting from $\pi_{\phi_i}^{0,\phi_i} = \hat{\pi}$, so that $\pi^{0,\phi_i} = \pi^\Phi$ on $[\phi_i, \phi_{i+1})$. Note that, by construction, $\pi_{\phi_i}^r \geq \hat{\pi}$. Then (3.11) and the comparison theorem for SDE's (i.e. Theorem 54 p. 324 of [46]) imply that $\pi^r \geq \pi^{0,\phi_i}$ on $[\phi_i, \phi_{i+1})$, and so $\pi^r \geq \pi^\Phi$ on $[\phi_i, \phi_{i+1})$. It now follows that $\pi_t^r \geq \pi_t^\Phi$ for all t, a.s. Consequently, for any $\tau \in \mathcal{T}$, $E[(1-\pi_\tau^r)] \leq E[(1-\pi_\tau^\Phi)]$, implying that $V_{\hat{\pi}}^r \leq V_{\hat{\pi}}$. □

Following [48], p. 146, we give the Skorokhod representation of π^r. Given a process Y and $\hat{\pi} \in \mathbb{R}$, the Skorokhod representation consists in finding a process X and an increasing process A such that $X = Y + A$, $X \geq \hat{\pi}$, and $\int_0^\infty (X_s - \hat{\pi}) dA_s = 0$, i.e. A only increases when $X = \hat{\pi}$.

Let $\sigma(x) = (1-x)x$, and let Y solve the SDE

$$Y_t \triangleq \hat{\pi} + \int_0^t \sigma(Y_s + A_s(Y)) dW_s, \tag{3.13}$$

$$A_t(Y) \triangleq \sup_{0 \leq s \leq t} \left\{ (Y_s - \hat{\pi})^- \right\}. \tag{3.14}$$

As in [48], the SDE (3.13)-(3.14) does in fact have a unique strong solution. Then, if we set $X_t \triangleq Y_t + A_t(Y)$, it is clear that $\pi^r = X$.

Let $\epsilon > 0$. We outline a parametrized family of switching strategies (impulse controls). Let Φ^ϵ denote the strategy that switches channels whenever the observed posterior process hits the level $\hat{\pi} - \epsilon$. Φ^ϵ induces the process π^ϵ, starting from $\pi_0^\epsilon = \hat{\pi}$, which diffuses according to $d\pi_t^\epsilon = \pi_t^\epsilon(1 - \pi_t^\epsilon) dW_t$ on $(\hat{\pi} - \epsilon, 1)$. When it reaches the level $\hat{\pi} - \epsilon$, it is instantaneously brought back to $\hat{\pi}$ (i.e. switched). We wish to give a Skorokhod type representation of π^ϵ. Consider the SDE

$$Y_t^\epsilon \triangleq \hat{\pi} + \int_0^t \sigma(Y_s^\epsilon + A_s^\epsilon(Y^\epsilon)) dW_s, \tag{3.15}$$

where

$$A_s^\epsilon(Y^\epsilon) \triangleq \epsilon \left\lfloor \frac{1}{\epsilon} \sup_{0 \leq s \leq t} \left\{ (Y_s^\epsilon - \hat{\pi})^- \right\} \right\rfloor. \tag{3.16}$$

Note that $A_s^\epsilon(\cdot)$ is not even continuous with respect to the uniform norm on continuous paths. Therefore, the standard theory does not imply that the SDE (3.15)-(3.16) has a strong solution. We can, however, show that a solution exists by a piecewise construction.

Lemma III.9. *For each $\epsilon > 0$, the SDE* (3.15)-(3.16) *has a strong solution. Moreover, for $X_t^\epsilon = Y_t^\epsilon + A_t^\epsilon(Y^\epsilon)$, $X^\epsilon = \pi^\epsilon$.*

Proof. Consider the SDE

$$Y_t^1 = Y_t^{\epsilon,1} \triangleq \hat{\pi} + \int_0^t \sigma(Y_s^{\epsilon,1})dW_s. \tag{3.17}$$

As $\sigma(\cdot)$ is Lipschitz and bounded on the interval $(0,1)$, it is known (see Theorem 11.5 of [48]) that (3.17) has a strong solution. Let $\tau^{\epsilon,0} \triangleq 0$, and $\tau^{\epsilon,1} \triangleq \inf\{t \geq 0 : Y_t^{\epsilon,1} = \hat{\pi} - \epsilon\}$. Note that on the random time interval $[0, \tau^{\epsilon,1})$, $Y^{\epsilon,1}$ solves the SDE (3.15). For $t \geq \tau^{\epsilon,1}$, consider next the SDE

$$Y_t^{\epsilon,2} \triangleq Y_{\tau^{\epsilon,1}}^{\epsilon,1} + \int_{\tau^{\epsilon,1}}^t \sigma(Y_s^{\epsilon,2} + \epsilon)dW_s. \tag{3.18}$$

As before, (3.18) has a strong solution. Let $\tau^{\epsilon,2} \triangleq \inf\{t \geq \tau^{\epsilon,1} : Y_t^{\epsilon,2} = \hat{\pi} - 2\epsilon\}$. Then on $[\tau^{\epsilon,1}, \tau^{\epsilon,2})$, $Y^{\epsilon,2}$ solves (3.15). Arguing inductively, we define $Y_t^{\epsilon,n}$, for $t \geq \tau^{\epsilon,n-1}$, by

$$Y_t^{\epsilon,n} \triangleq Y_{\tau^{\epsilon,n-1}}^{\epsilon,n-1} + \int_{\tau^{\epsilon,n-1}}^t \sigma(Y_s^{\epsilon,n} + (n-1)\epsilon)dW_s, \tag{3.19}$$

which has a strong solution as before, and the stopping time $\tau^{\epsilon,n} \triangleq \inf\{t \geq \tau^{\epsilon,n-1} : Y^{\epsilon,n} = \hat{\pi} - n\epsilon\}$. Defining the process Y^ϵ by $Y_t^\epsilon \triangleq Y_t^{\epsilon,n}$ for $t \in [\tau^{\epsilon,n-1}, \tau^{\epsilon,n})$, $n \geq 1$, it is apparent that Y^ϵ solves the SDE (3.15).

For the last claim, that $X^\epsilon = \pi^\epsilon$, note that when $A^\epsilon(Y^\epsilon)$ is constant, $dX_t^\epsilon = \sigma(Y_t^\epsilon + A_t^\epsilon(Y^\epsilon))dW_t = \sigma(X_t^\epsilon)dW_t$. The times when $A^\epsilon(Y^\epsilon)$ jumps (by ϵ) correspond to the impulses from $\hat{\pi} - \epsilon$ to $\hat{\pi}$. □

Lemma III.10. *For any $t, \epsilon > 0$, we have $E\left[(Y - Y^\epsilon)_t^{*2}\right] \leq 8te^{32t}\epsilon^2$. In particular, for any $t \geq 0$, $(Y - Y^\epsilon)_t^* \to 0$ in L^2 as $\epsilon \to 0$.*

Proof. Let K be the Lipschitz constant of $\sigma(\cdot)$ on $(0,1)$. Write

$$
\begin{aligned}
&E(Y^\epsilon - Y)_t^{*2} \\
&\quad = E\left[\left(\hat{\pi} + \int_0^t \sigma(Y_s^\epsilon + A_s^\epsilon(Y^\epsilon))dW_s - \hat{\pi} - \int_0^t \sigma(Y_s + A_s(Y))dW_s\right)^{*2}\right] \\
&\quad = E\left[\left(\sigma(Y_\cdot^\epsilon + A_\cdot^\epsilon(Y^\epsilon)) \cdot W - \sigma(Y_\cdot^\epsilon + A_\cdot(Y^\epsilon)) \cdot W + \sigma(Y_\cdot^\epsilon + A_\cdot(Y^\epsilon)) \cdot W - \sigma(Y_\cdot + A_\cdot(Y)) \cdot W\right)_t^{*2}\right] \\
&\quad \le E\Bigg[\Big(\big[\sigma(Y_\cdot^\epsilon + A_\cdot^\epsilon(Y^\epsilon)) \cdot W - \sigma(Y_\cdot^\epsilon + A_\cdot(Y^\epsilon)) \cdot W\big]_t^* \\
&\qquad + \big[\sigma(Y_\cdot^\epsilon + A_\cdot(Y^\epsilon)) \cdot W - \sigma(Y_\cdot + A_\cdot(Y)) \cdot W\big]_t^*\Big)^2\Bigg] \\
&\quad \le 2E\left[\left[(\sigma(Y_\cdot^\epsilon + A_\cdot^\epsilon(Y^\epsilon)) - \sigma(Y_\cdot^\epsilon + A_\cdot(Y^\epsilon))) \cdot W\right]_t^{*2}\right] \\
&\qquad + 2E\left[\left[(\sigma(Y_\cdot^\epsilon + A_\cdot(Y^\epsilon)) - \sigma(Y_\cdot + A_\cdot(Y))) \cdot W\right]_t^{*2}\right] \\
&\quad \triangleq (1) + (2),
\end{aligned}
$$

with the second inequality above following from $(a+b)^2 \le 2a^2 + 2b^2$. Next, using the Burkholder-Davis-Gundy Inequality for the first inequality below and the K-Lipschitzness of $\sigma(\cdot)$ for the second,

$$
\begin{aligned}
(1) &\le 2C_2 E \int_0^t \left(\sigma(Y_s^\epsilon + A_s^\epsilon(Y^\epsilon)) - \sigma(Y_s^\epsilon + A_s(Y^\epsilon))\right)^2 ds \\
&\le 2C_2 K^2 E \int_0^t (Y_s^\epsilon + A_s^\epsilon(Y^\epsilon) - Y_s^\epsilon - A_s(Y^\epsilon))^2 ds \\
&\le 2C_2 K^2 \epsilon^2 t;
\end{aligned}
$$

the last inequality follows from the fact that for a given path ω, $\sup\limits_{0\le s\le t} |A_s(\omega) - A_s^\epsilon(\omega)| \le \epsilon$. The constant C_2 is a universal constant arising from the Burkholder-Davis-Gundy Inequality. The L^2 version used above actually can be proven using Doob's L^2-inequality for martingales, and from this the explicit formula $C_2 = 4$ can be derived. For details, see p. 14 of [30].

Next, we note that $A_\cdot(\cdot)$ is Lipschitz continuous with respect to the uniform norm on continuous paths, with Lipschitz constant 1. Applying this fact for the third inequality below, Burkholder-Davis-Gundy Inequality for the first inequality, the K-Lipschitz continuity of $\sigma(\cdot)$ for the second inequality, and Fubini's Theorem in the last inequality, we obtain

$$\begin{aligned}(2) &\le 2C_2E\int_0^t (\sigma(Y_s^\epsilon + A_s(Y^\epsilon)) - \sigma(Y_s + A_s(Y)))^2\, ds\\ &\le 2C_2K^2E\int_0^t (Y_s^\epsilon + A_s(Y^\epsilon) - Y_s + A_s(Y))^2 ds\\ &\le 8C_2K^2E\int_0^t (Y_s^\epsilon - Y_s)^2 ds\\ &\le 8C_2K^2\int_0^t E(Y^\epsilon - Y)_s^{*2} ds.\end{aligned}$$

For each $\epsilon > 0$, define $f^\epsilon : \mathbb{R}_+ \to \mathbb{R}_+$ by $f^\epsilon(s) = E(Y^\epsilon - Y)_s^{*2}$. According to the above reasoning, $f^\epsilon(t) \le 2C_2K^2t\epsilon^2 + 8C_2K^2\int_0^t f^\epsilon(s)ds$. By Gronwall's Lemma, it follows then that $f^\epsilon(t) \le 2C_2K^2t\epsilon^2 e^{8C_2K^2t}$. Since all processes in question live in the interval $(0,1)$, we may assume that $K = 1$. Therefore, $f^\epsilon(t) \le 8te^{32t}\epsilon^2$. In particular, for fixed t, $(Y^\epsilon - Y)_t^* \to 0$ in L^2 as $\epsilon \to 0$. □

Corollary III.11. *For any $t, \epsilon \ge 0$, $E\left[(\pi^\epsilon - \pi^r)_t^{*2}\right] \le 16te^{32t}\epsilon^2 + \epsilon$. In particular, as $\epsilon \to 0$, $(\pi^\epsilon - \pi^r)_t^* \to 0$ in L^2.*

Proof. Write $\pi_t^r = Y_t + A_t(Y)$ and $\pi_t^\epsilon = Y_t^\epsilon + A_t^\epsilon(Y^\epsilon)$. We have shown in Lemma III.10 that $(Y^\epsilon - Y)_t^* \to 0$ in L^2 as $\epsilon \to 0$. Therefore, it suffices to show that $(A_\cdot(Y) - A_\cdot^\epsilon(Y^\epsilon))_t^* \to 0$ in L^2 as $\epsilon \to 0$. So, for any $s \ge 0$,

$$\begin{aligned}|A_s(Y) - A_s^\epsilon(Y^\epsilon)| &\le |A_s(Y) - A_s(Y^\epsilon)| + |A_s(Y^\epsilon) - A_s^\epsilon(Y^\epsilon)|\\ &\le (Y - Y^\epsilon)_s^* + \epsilon,\end{aligned}$$

where we have used the Lipschitz continuity of $A_s(\cdot)$ with respect to the uniform norm. Therefore, $(A_\cdot(Y) - A_\cdot^\epsilon(Y^\epsilon))_t^* \le (Y - Y^\epsilon)_t^* + \epsilon$, which converges to 0 in L^2 as $\epsilon \to 0$. The quantitative estimate is also clear. □

Lemma III.12. $V_{\hat\pi}^r = V_{\hat\pi}$

Proof. In light of Lemma III.8, it suffices to show that $V_{\hat\pi}^r \ge V_{\hat\pi}$. Without loss of generality, we assume that $V_{\hat\pi}^r < \infty$; otherwise, there is nothing to show. Let

$\{\tau_n : n \in \mathbb{N}\}$ be a sequence of stopping times such that $E[c\tau_n + (1 - \pi^r_{\tau_n})] \downarrow V^r_{\hat{\pi}}$. Fix $\delta > 0$, and choose n sufficiently large so that $E\left[c\tau_n + (1 - \pi^r_{\tau_n})\right] < V^r_{\hat{\pi}} + \delta$.

Next, we note that the processes π^r and π^ϵ are all bounded, so that in particular, they are uniformly of Class D. Therefore, for a suitably large t, it is the case that $\max\left\{E\left[\pi^r_{\tau_n} 1_{\{\tau_n > t\}}\right], E\left[\pi^\epsilon_{\tau_n} 1_{\{\tau_n > t\}}\right]\right\} < \delta$ for each $\epsilon > 0$. By Corollary III.11, for ϵ sufficiently small, $\left|E\left[\pi^r_{\tau_n} 1_{\{\tau_n \le t\}}\right] - E\left[\pi^\epsilon_{\tau_n} 1_{\{\tau_n \le t\}}\right]\right| < \delta$. Thus,

$$\begin{aligned}
&\left|V^r_{\hat{\pi}} - E\left[c\tau_n + (1 - \pi^\epsilon_{\tau_n})\right]\right| \\
&\quad \le \left|V^r_{\hat{\pi}} - E\left[c\tau_n + (1 - \pi^r_{\tau_n})\right]\right| + \left|E\left[c\tau_n + (1 - \pi^r_{\tau_n})\right] - E\left[c\tau_n + (1 - \pi^\epsilon_{\tau_n})\right]\right| \\
&\quad = \left|V^r_{\hat{\pi}} - E\left[c\tau_n + (1 - \pi^r_{\tau_n})\right]\right| + \left|E\left[\pi^r_{\tau_n}\right] - E\left[\pi^\epsilon_{\tau_n}\right]\right| \\
&\quad < \delta + \left|E\left[\pi^r_{\tau_n} 1_{\{\tau_n > t\}}\right] - E\left[\pi^\epsilon_{\tau_n} 1_{\{\tau_n > t\}}\right]\right| + \left|E\left[\pi^r_{\tau_n} 1_{\{\tau_n \le t\}}\right] - E\left[\pi^\epsilon_{\tau_n} 1_{\{\tau_n \le t\}}\right]\right| \\
&\quad < \delta + 2\delta + \delta \\
&\quad = 4\delta.
\end{aligned}$$

Since $V_{\hat{\pi}} \le V^{\Phi^\epsilon}_{\hat{\pi}} \le E\left[c\tau_n + (1 - \pi^\epsilon_{\tau_n})\right]$, it now follows that $V_{\hat{\pi}} \le V^r_{\hat{\pi}}$. □

III.5 Optimal Stopping of the Reflected Diffusion

We wish to relate the optimal stopping problem $V^r_{\hat{\pi}} = \inf_{\tau \in \mathcal{T}} E[c\tau + (1 - \pi^r_\tau)]$ to an ODE with a free boundary. First, we look for $f : [\hat{\pi}, 1] \to \mathbb{R}$ and $\pi^* \in (\hat{\pi}, 1)$ that satisfy:

$$\frac{1}{2}[x(1-x)]^2 \frac{d^2 f}{dx^2} = -c, \hat{\pi} < x < \pi^*, \tag{3.20}$$

$$f(x) = 1 - x, \pi^* \le x \le 1, \tag{3.21}$$

$$f'(\hat{\pi}) = 0, f'(\pi^*) = -1. \tag{3.22}$$

Notice in particular that we require f to be C^1 at $\hat{\pi}$ and π^*; π^* must be chosen to ensure that this happens.

Lemma III.13. *The problem* (3.20),(3.21),(3.22) *has a unique solution, for precisely one* $\pi^* \in [\hat{\pi}, 1)$.

Proof. One may verify directly that the function $\Psi_{A,B}(x) \triangleq 2c(1-2x)\log\frac{x}{1-x} + Ax + B \triangleq \Psi(x) + Ax + B$, for constants A and B, is the general solution of (3.20). We will show that the boundary conditions are satisfied for precisely one π^*, $\overline{A}$, and $\overline{B}$.

The condition $f'(\hat{\pi}) = 0$ forces $\overline{A} = -\Psi'(\hat{\pi}) = -2c\left[\frac{2\hat{\pi}-2(\hat{\pi}-1)\hat{\pi}\log\left(\frac{\hat{\pi}}{1-\hat{\pi}}\right)-1}{(\hat{\pi}-1)\hat{\pi}}\right]$. Since $\Psi(x) + \overline{A}x$ is strictly concave and $\Psi'(x) + \overline{A}$ is continuous on $[\hat{\pi}, 1)$, $\Psi'(\hat{\pi}) + \overline{A} = 0$, and $\lim_{x\uparrow 1} \Psi(x) + \overline{A}x = -\infty$ (so $\lim_{x\uparrow 1} \Psi'(x) + \overline{A} = -\infty$), it follows that there is a unique $\pi^* \in (\hat{\pi}, 1)$ such that $\Psi'(\pi^*) + \overline{A} = -1$. Define $\overline{B}$ so that $\overline{B}$ satisfies the equality $1 - \pi^* = \Psi_{\overline{A},\overline{B}}(\pi^*) = \Psi(\pi^*) + \overline{A}\pi^* + \overline{B}$. Taking $f(x) = \Psi_{\overline{A},\overline{B}}(x)$ for $x \in [\hat{\pi}, \pi^*)$ and $f(x) = 1 - x$ for $x \in [\pi^*, 1]$ yields the unique solution to (3.20),(3.21),(3.22). $\square$

Our candidate for the value function is therefore

$$f(x) \triangleq \begin{cases} 2c(1-2x)\log\frac{x}{1-x} + \overline{A}x + \overline{B} & \text{if } \widehat{\pi} \le x \le \pi^* \\ 1 - x & \text{if } \pi^* \le x \le 1 \end{cases} \tag{3.23}$$

with

$$\overline{A} = -2c\left[\frac{2\hat{\pi} - 2(\hat{\pi}-1)\hat{\pi}\log\left(\frac{\hat{\pi}}{1-\hat{\pi}}\right) - 1}{(\hat{\pi}-1)\hat{\pi}}\right],$$

$$\pi^* \text{ satisfying } 2c\left[\frac{2\pi^* - 2(\pi^*-1)\pi^*\log\left(\frac{\pi^*}{1-\pi^*}\right) - 1}{(\pi^*-1)\pi^*}\right] + \overline{A} = -1, \tag{3.24}$$

$$\overline{B} \text{ satisfying } 1 - \pi^* = 2c(1-2\pi^*) + \overline{A}\pi^* + \overline{B}.$$

Remark III.14. It is interesting to consider the dependence of π^* on $\widehat{\pi}$. Note that $\overline{A}(\widehat{\pi}) \triangleq -\Psi'(\widehat{\pi})$ is increasing as a function of $\widehat{\pi}$, since $\Psi(x)$ is concave. Rewriting (3.24) as

$$\text{Find } \pi^*(\widehat{\pi}) \text{ such that } \Psi'(\pi^*(\widehat{\pi})) = -1 - \overline{A}(\widehat{\pi}), \tag{3.25}$$

we see that the left hand side of (3.25) must be decreasing in $\widehat{\pi}$, and this corresponds to $\pi^*(\widehat{\pi})$ increasing in $\widehat{\pi}$. This reflects the fact that, as the prior probability of success in each channel increases, we should become more selective in accepting H_1 for a given channel.

For each $x \in [\hat{\pi}, 1]$, we set $V^r_{\hat{\pi}}(x) = \inf_{\tau \in \mathcal{T}} E_x\left[c\tau + (1 - \pi^r_\tau)\right]$, where the expectation $E_x[\cdot]$ denotes expectation under the probability P_x, i.e. $P_x(\pi^r_0 = x) = 1$. We now claim that $f(x)$ is equal to the value function $V^r_{\hat{\pi}}(x)$. Consider the set $\mathcal{D} \triangleq \{f \in C^2_b([\hat{\pi}, 1)) : f'(\hat{\pi}) = 0\}$. The infinitesimal generator $\mathcal{L}^r$ of π^r satisfies, for $f \in \mathcal{D}$, $\mathcal{L}^r f(x) = \frac{1}{2}x^2(1-x)f''(x)$.

Lemma III.15. *For $f(x)$ as above, $V^r_{\hat{\pi}}(x) = f(x)$.*

Proof. We wish to apply a verification theorem for optimal stopping problems, Theorem 10.4.1 of [42], p. 225. We must check that $f(x)$ defined by (3.20),(3.21),(3.22) satisfies the nine hypotheses of that theorem. Note that several inequalities are reversed because our problem involves a minimization over all stopping times. Let $G = [\hat{\pi}, 1]$, and let $D = \{x \in G : f(x) < 1 - x\}$.

(i) $f \in C^1(G)$: This is true by construction.

(ii) $f(x) \leq 1 - x$ on G: At π^*, $f(\pi^*) = 1 - \pi^*$. Since $f(x)$ is concave down, $f(x) \leq 1 - x$ for $x \in [\hat{\pi}, \pi^*]$, and by construction $f(x) = 1 - x$ on $[\pi^*, 1]$.

(iii) $E_x\left[\int_0^\infty 1_{\{\pi^*\}}(\pi^r_s)ds\right] = 0$: This follows from the fact that the speed measure of π^r_s is $m^r(dx) \triangleq \frac{dx}{x^2(1-x)^2}$. Now, i.e. Proposition 3.10 of [47], p. 307 may be applied.

(iv) ∂D is Lipschitz: This is trivial in the one-dimensional problem here.

(v) $f \in C^2(G \backslash \{\pi^*\})$ and the second order derivatives of f are bounded near π^*: For

$x \in (\hat{\pi}, \pi^*)$, $f''(x) = \frac{-2c}{x^2(1-x)^2}$, which is bounded on $(\hat{\pi}, \pi^*)$, and for $x \in (\pi^*, 1)$, $f''(x) = 0$.

(vi) $\mathcal{L}^r f + c \geq 0$ on $G \setminus D$: For $x \in G \setminus D$, $\mathcal{L}^r f + c = 0 + c \geq 0$.

(vii) $\mathcal{L}^r f + c = 0$ on D: For $x \in D$, $\mathcal{L}^r f + c = \frac{1}{2}x^2(1-x)^2 \left(\frac{-2c}{x^2(1-x)^2}\right) + c = 0$.

(viii) $\tau_D \triangleq \inf\{t > 0 : \pi_t^r \notin D\} < \infty$, P_x-a.s. for each $x \in G$: Using the same argument as in (iii), Proposition 3.10 of [47] implies that $E_x[\tau_D] < \infty$ for each $x \in G$.

(ix) The family $\{\pi_\tau^r : \tau \leq \tau_D, \tau \in \mathcal{T}\}$ is P_x-uniformly integrable for any $x \in G$: This is immediate, using the fact that π^r is bounded.

Having checked all the hypotheses of the verification theorem, we deduce that $f(x) = V_{\hat{\pi}}^r(x)$. □

III.6 A Rough Algorithm for Quickest Search

Using the methods of the previous sections, we can describe near optimal algorithms for quickly finding a channel which satisfies hypothesis H_1. We outline a procedure below for finding an $\boldsymbol{\epsilon}$-optimal strategy.

(1) Fix $\boldsymbol{\epsilon} > 0$.

(2) For given values of $c, \hat{\pi}$, calculate the threshold $\pi^* = \pi^*(c, \hat{\pi})$ via Lemma III.10. Let $\tau \triangleq \inf\{t \geq 0 : \pi_t^r = \pi^*\}$ be defined for any version of π^r.

(3) Choose $t > 0$ sufficiently large so that $P(\tau > t) < \frac{\boldsymbol{\epsilon}}{4}$. This can be done, for example, by calculating $E[\tau]$ via the speed measure of π^r.

(4) Choose ϵ_2 sufficiently small so that $16te^{32t}\epsilon_2^2 + \epsilon_2 < \frac{\boldsymbol{\epsilon}}{2}$.

(5) Adopt the switching strategy Φ^{ϵ_2}, in which the observed channel is switched whenever the posterior level hits $\hat{\pi} - \epsilon_2$.

(6) The switching strategy Φ^{ϵ_2} induces the observed Brownian Motion $W = W^{\Phi^{\epsilon_2}}$. Using W, construct the solution to the SDE $Y_t = \hat{\pi} + \int_0^t \sigma(Y_s + A_s(Y))dW_s$, and set $X_t = Y_t + A_t(Y)$. Let $\tau^* = \inf\{t \geq 0 : X_t = \pi^*\}$.

(7) At time τ^*, accept hypothesis H_1 for the channel which is currently being observed.

Applying the reasoning of Lemmas III.10 and III.12, we may deduce that this observation/stopping strategy will be $\boldsymbol{\epsilon}$-optimal.

III.7 Numerical Results

In this section, we illustrate our previous results by computing the optimal threshold level for various levels of the observation cost c and prior $\hat{\pi}$. The data below, which can be found in Appendix A, is directly calculated from the value function established in Section III.5. We first plot the threshold levels against the observation cost c, when the prior $\hat{\pi}$ is fixed. As indicated by Tables 3.1 and 3.2 below, for fixed $\hat{\pi}$, $\pi^*(c)$ decreases with c. This is not surprising, because the higher the running cost for observations, the lower one's standards will be for accepting the hypothesis that a channel satisfies H_1.

Next, we plot the threshold levels against the prior $\hat{\pi}$, when the observation cost c is fixed. As indicated by Tables 3.3 and 3.4 below, for fixed c, $\pi^*(\hat{\pi})$ increases with $\hat{\pi}$. This fact was proven analytically in Remark III.14.

Appendix A Tables of Data

Table 3.1: Optimal thresholds $\pi^*(c)$, for $\hat{\pi} = .5$.

c	$\pi^*(c)$	c	$\pi^*(c)$
.0025	.995	.05	.865
.005	.989	.06	.840
.01	.977	.07	.815
.02	.950	.08	.793
.03	.922	.09	.773
.04	.893	.1	.755

Table 3.2: Optimal thresholds $\pi^*(c)$, for $\hat{\pi} = .75$.

c	$\pi^*(c)$	c	$\pi^*(c)$
.0025	.995	.05	.911
.005	.990	.06	.899
.01	.979	.07	.889
.02	.959	.08	.879
.03	.941	.09	.871
.04	.925	.1	.864

Table 3.3: Optimal thresholds $\pi^*(\hat{\pi})$, for $c = .01$

$\hat{\pi}$	$\pi^*(\hat{\pi})$	$\hat{\pi}$	$\pi^*(\hat{\pi})$
.025	.704	.5	.977
.05	.951	.6	.978
.1	.968	.7	.979
.2	.973	.8	.980
.3	.975	.9	.982
.4	.976	.95	.985

Table 3.4: Optimal thresholds $\pi^*(\hat{\pi})$, for $c = .03$

$\hat{\pi}$	$\pi^*(\hat{\pi})$	$\hat{\pi}$	$\pi^*(\hat{\pi})$
.025	.041	.5	.922
.05	.164	.6	.930
.1	.690	.7	.937
.2	.867	.8	.946
.3	.898	.9	.960
.4	.913	.95	.972

CHAPTER IV

Quickest Detection with Discretely Controlled Observations

IV.1 Introduction

In this chapter, we study a quickest detection problem with imperfect information. Since their original formulation, there has been an extensive literature on quickest detection problems which modify or generalize the classical assumptions of [52]. We mention a few of these papers, although it is impossible to describe the entire literature on the subject. In [11], the deterministic drift term α received after disorder is replaced with a random variable. In [12], the "linear" penalty for delay, $E\left[(\tau-\Theta)^{+}\right]$, is replaced with an exponential one, $E\left[e^{(\tau-\Theta)^{+}}-1\right]$. In [23], the continuous time problem is solved on a finite time horizon, in comparison with the classical infinite horizon case. In continuous time, it is also natural to study a Poisson process whose intensity rate suddenly changes, and this problem, known as Poisson disorder, has been extensively studied in, for example, [44], [9], [8], and [7].

In all of the above problems, there is perfect observation. Another interesting generalization of the quickest detection problem occurs when we place restrictions on our ability to make observations. The literature on this problem is relatively sparser, but there is a strand of research from the 1960's and 1970's in which observations are available at all times, but must be purchased. In [2], the costly information quickest detection problem for Brownian Motions is first studied; solutions are not obtained, but some qualitative properties of the value function are established. In [6], the author studies a problem with both costly information and imperfect observations; again, no explicit solutions are found. Finally, in [3], a more thorough analysis of a costly information problem is attempted, but there is a problem with the author's analysis: when the process X is not continuously observed, the posterior process (e.g. $\pi_t \triangleq P\left(\Theta \leq t | \mathcal{F}_t^X\right)$) ceases to become a sufficient statistic, and the time elapsed since an observation was last made must also be tracked. More recently, in [4], a minimax problem is studied in discrete time, where the expected detection delay is minimized subject to constraints on false alarm probability *and* expected number of observations taken before the disorder time, and in [5] the same problem is studied in a Bayesian framework.

Alternatively, one may formulate a quickest detection problem in which observations of the process X can be made only at discrete time periods, and such that the agent possesses, no matter what, a fixed amount of observation rights. For example, we could imagine a remote battery powered sensor which has enough power to make a fixed number of observations and must be active for an extended amount of time. Such a problem was first studied in [15], in which it is assumed that observations fall on a grid which is determined exogenously. If observations can be made only at discrete time periods, it makes sense to consider the case when there is control over when observations can be used: if observations are a limited resource, then the judicious use of them should increase efficiency significantly. In such a scenario, the observation times will no longer be exogenously given, but will be determined adaptively within the problem as part of the optimal strategy. In [15], an infinite sequence of of observation times is given. If we allow the controller to choose when observations are made, it does not make sense to allow him infinitely many observation rights, because as we will see, such a problem is degenerate, and equivalent to the classical continuously observed case. Therefore, if observation times can be chosen, there must be some limit on how they can be spent. In this chapter, we therefore study such a problem: an agent seeks to determine when a disorder occurs, but his ability to observe is constrained. In the first variant of the problem, we assume that the agent receives a lump sum of n observation rights which he may use as he sees fit. In the second variant, we assume that an independent Poisson process regulates the times at which new observation rights become available.

We now outline the structure of the chapter. In Section IV.2, we formulate the lump sum n-observation problem, and establish the theoretical existence of optimal strategies. In Section IV.3, we demonstrate that as $n \to \infty$, this n-observation problem converges to the classical continuous observation problem. In Section IV.4, we formulate the stochastic arrival rate n-observation problem, and establish the theoretical existence of optimal strategies. In Sections IV.5 and IV.6, we give a numerical algorithm for computing the value functions and optimal strategies in the lump sum n-observation problem, and illustrate some results from the implementation of this algorithm. In Section IV.7, we describe a heuristic algorithm for computing the value functions and optimal strategies in the stochastic arrival rate problem, and illustrate a result from a partial implementation of this algorithm. Sections IV.8, IV.9, and IV.10 contain the technical proofs of the results in Sections IV.2, IV.4, and IV.3. Finally, Appendix A establishes the dynamics of the posterior process under discrete observations, and Appendix B contains source code for our algorithms in Octave and for graphing the data in R.

IV.2 The Lump Sum n-Observation Problem: Setup, Existence of Optimal Strategies

Our basic setup is a probability space $(\Omega, \mathcal{F}, P')$, which supports a Wiener process $X = \{X_t\}_{t\geq 0}$ and an independent random variable Θ, which has the same distribution as before: with probability p it is zero, and with probability $1-p$ it is exponentially distributed with parameter λ.

In [9], observation times are determined exogenously, and therefore the information flow is a fixed aspect of the problem. In contrast, when the agent must decide when to make observations, the information flow is itself variable. In other words, the filtration is dependent on the observation strategy used. We therefore have to be somewhat technical in our definition of observation strategies. We will now inductively define elements in the set of allowed observation strategies, denoted by $\mathfrak{O}^n$.

Definition IV.1. We say that a sequence of random variables $\Psi = \{\psi_1, \psi_2, \ldots, \psi_n\} \in \mathfrak{O}^n$ if $\psi_1 \leq \psi_2 \leq \cdots \leq \psi_n$, ψ_1 deterministic, and for $1 \leq j \leq n$, $\psi_j \in$ $m\,\sigma(X_{\psi_1}, \ldots, X_{\psi_{j-1}}, \psi_1, \ldots, \psi_{j-1})$, i.e. ψ_j is measurable with respect to the sigma algebra generated by $X_{\psi_1}, \ldots, X_{\psi_{j-1}}, \psi_1, \ldots, \psi_{j-1}$. We set $\psi_0 = 0$, and for convenience take $\psi_{n+1} = \infty$.

For each $\Psi \in \mathfrak{O}^n$, let $\mathcal{F}^{\Psi}_{\psi_j} = \sigma(X_{\psi_1}, \ldots, X_{\psi_j}, \psi_1, \ldots, \psi_j)$. Ψ generates a continuous time filtration $\mathbb{F}^{\Psi} = (\mathcal{F}^{\Psi}_t)_{t\geq 0}$ in the following way. We say that $A \in \mathcal{F}^{\Psi}_t$ if and only if for each $1 \leq j \leq n$, $A \cap \{\psi_j \leq t\} \in \mathcal{F}^{\Psi}_{\psi_j}$. Intuitively, this just means that the set A is known at time t if, for any j, it is known at the time of the j^{th} observation when this observation comes before t. Let $\mathcal{T}^{\Psi}$ be the set of $\mathbb{F}^{\Psi}$-stopping times which are a.s. finite.

Let Φ^{Ψ} be the conditional odds-ratio ratio process that the disorder has occurred, supposing that the observation strategy Φ has been used. In other words

$$\Phi_t^\Psi \triangleq \frac{P(\Theta \le t | \mathcal{F}_t^\Psi)}{P(\Theta > t | \mathcal{F}_t^\Psi)}.$$

The posterior process Φ^Ψ can be calculated recursively by the following formula, starting from $\Phi_0^\Psi = \frac{p}{1-p}$: for more details, please see Appendix A, which follows the derivation on p.32-33 of [9].

$$\Phi_t^\Psi = \begin{cases} \varphi(t - \psi_{n-1}, \Phi_{\psi_{n-1}}^\Psi) & \text{if } \psi_{n-1} \le t < \psi_n \\ j\left(\Delta\psi_n, \Phi_{\psi_{n-1}}^\Psi, \frac{\Delta X_{\psi_n}}{\sqrt{\Delta\psi_n}}\right) & \text{if } t = \psi_n, \end{cases} \tag{4.1}$$

where $\Delta\psi_n = \psi_n - \psi_{n-1}$, $\Delta X_{\psi_n} = X_{\psi_n} - X_{\psi_{n-1}}$, $\varphi(t, \phi) = e^{\lambda t}(\phi + 1) - 1$, and

$$\begin{aligned} & j(\Delta t, \phi, z) \\ & = \exp\left\{\alpha z \sqrt{\Delta t} + \left(\lambda - \frac{\alpha^2}{2}\right)\Delta t\right\}\phi + \int_0^{\Delta t} \lambda \exp\left\{\left(\lambda + \frac{\alpha z}{\sqrt{\Delta t}}\right)u - \frac{\alpha^2 u^2}{2\Delta t}\right\} du. \end{aligned}$$

According to Lemma 3.1 of [9], the minimum Bayes risk equals $R_n(p) = 1 - p + (1-p)cV_n(p/(1-p))$, where

$$V_n(\phi) \triangleq \inf_{\Psi \in \mathfrak{O}^n} \inf_{\tau \in \mathcal{T}^\Psi} E^\phi\left[\int_0^\tau e^{-\lambda t}\left(\Phi_t^\Psi - \frac{\lambda}{c}\right) dt\right], \tag{4.2}$$

and the expectation $E^\phi[\cdot]$ is with respect to a probability measure P under which X is a standard Weiner process and $\Phi_0^\Psi = \phi$. Consequently, (4.2) is the problem we will focus on.

Proposition IV.2. *Let $\Psi \in \mathfrak{O}^n$, and let τ be an $\mathbb{F}^\Psi$-stopping time. Then for each $0 \le j \le n$, $\tau 1_{\{\psi_j \le \tau < \psi_{j+1}\}}$ and $\{\psi_j \le \tau < \psi_{j+1}\}$ are both $\mathcal{F}_{\psi_j}^\Psi$-measurable.*

Proof. The proof is done by a basic modification of Proposition 3.1 and Theorem 3.2 of [15]. The essential property here is that between observations, there is no flow of new information. $\square$

Define

$$\mathcal{T}_o^{\Psi} \triangleq \left\{\tau \in \mathcal{T}^{\Psi} : \text{for } \omega \in \Omega \text{ with } \tau(\omega) \le \psi_n(\omega), \tau(\omega) = \psi_j(\omega) \text{ for some } 0 \le j \le n\right\},$$

i.e., those $\mathbb{F}^{\Psi}$-stopping times that do not stop between observations. The following proposition says that, in contrast with [9], it is never optimal to stop between observations: if one has a total of n observations at their disposal, he may as well use all of them.

Proposition IV.3. $V_n(\phi) = \inf\limits_{\Psi \in \mathfrak{O}^n} \inf\limits_{\tau \in \mathcal{T}_o^{\Psi}} E^{\phi}\left[\int_0^{\tau} e^{-\lambda t}\left(\Phi_t^{\Psi} - \frac{\lambda}{c}\right) dt\right].$

Proof. See Section IV.8. □

Note that for each Ψ, Φ^{Ψ} evolves deterministically between observations. This means that between observations, there is no additional information being accrued. Therefore, upon making an observation, one may as well determine in that instant when to make the next observation, as opposed to waiting to see what happens ϵ seconds in the future; no additional information is gained by waiting. Therefore, the problem is amenable to study by the recursive use of jump operators. We lay out this strategy now.

For bounded $w : \mathbb{R}_+ \to \mathbb{R}$, define the operators

$$Kw(t,\phi) \triangleq \int_{-\infty}^{\infty} w(j(t,\phi,z)) \frac{\exp(-z^2/2)}{\sqrt{2\pi}} dz, \tag{4.3}$$

$$Jw(t,\phi) \triangleq \int_0^t e^{-\lambda u}\left(\varphi(u,\phi) - \frac{\lambda}{c}\right) du + 1_{\{t>0\}} e^{-\lambda t} Kw(t,\phi), \tag{4.4}$$

and

$$J_0 w(\phi) \triangleq \inf_{t \ge 0} Jw(t,\phi). \tag{4.5}$$

Set $v_0(\phi) \triangleq J_0 0$. Inductively define the value functions, for $j \ge 1$,

$$v_j(\phi) \triangleq J_0 v_{j-1}(\phi).$$

Let $0 \le k \le n$, and let $\Psi^k = \{\psi_1^k, \ldots, \psi_k^k\} \in \mathfrak{O}^k$. We set

$$\mathfrak{O}^n(\Psi^k) \triangleq \left\{\Psi = \{\psi_1, \ldots, \psi_n\} \in \mathfrak{O}^n : \psi_i = \psi_i^k,\ 1 \le i \le k\right\}. \tag{4.6}$$

These are the observation strategies whose first k observation times agree with those in Ψ^k. We define the following conditional value functions:

$$\gamma_k^n(\Psi^k) \triangleq \operatorname*{ess\,inf}_{\Psi \in \mathfrak{O}^n(\Psi^k)} \operatorname*{ess\,inf}_{\tau \in \mathcal{T}_o^\Psi, \tau \ge \psi_k^k} E\left[\int_{\psi_k^k}^{\tau} e^{-\lambda(t-\psi_k^k)} \left(\Phi_t^\Psi - \frac{\lambda}{c}\right) dt \middle| \mathcal{F}_{\psi_k^k}^{\Psi^k}\right].$$

Propositions IV.4 and IV.5 allow us to describe the optimization problem in terms of functions defined by the jump operator J_0. It also establishes that the optimization problem is Markov.

Proposition IV.4. *For any n, $0 \le k \le n$, and $\Psi^k \in \mathfrak{O}^k$,*

$$\gamma_k^n(\Psi^k) \ge v_{n-k}\left(\Phi_{\psi_k^k}^{\Psi^k}\right).$$

Proof. See Section IV.8. □

Proposition IV.5. *For any n, $0 \le k \le n$, and $\Psi^k \in \mathfrak{O}^k$,*

$$\gamma_k^n(\Psi^k) \le v_{n-k}\left(\Phi_{\psi_k}^{\Psi^k}\right).$$

Hence, $\gamma_k^n(\Psi^k) = v_{n-k}\left(\Phi_{\psi_k}^{\Psi^k}\right)$. In particular, $V_n(\phi) = v_n(\phi)$.

Proof. See Section IV.8. □

For $0 \le k < n$ and $\epsilon \ge 0$, define

$$h_{n-k}^\epsilon(\phi) \triangleq \min\{s \ge 0 : Jv_{n-k}(s, \phi) \le J_0 v_{n-k}(\phi) + \epsilon\}.$$

These functions are used to construct the near optimal strategies needed for the proof of Proposition IV.5, but we must show that they are measurable. Note that in the definition of h_{n-k}^ϵ, we require the first time s such that $Jv_{n-k}(s, \phi) \le J_0 v_{n-k}(\phi) + \epsilon$.

If we simply required any such ϵ-optimal time, we could use a Measurable Selection Theorem, as in [54], to imply the measurability of h^{ϵ}_{n-k}. Such a theorem, however, would provide only abstract existence. For computational reasons it is preferable to take the first optimal time.

Lemma IV.6. *For $0 \le k < n$ and $\epsilon \ge 0$, $\widehat{\psi}^{k+1} \triangleq \psi^k_k + h^{\epsilon}_{n-k}\left(\Phi^{\Psi^k}_{\psi^k_k}\right)$ is a stopping time, i.e. it is measurable with respect to $\mathcal{F}^{\Psi^k}_{\psi^k_k}$.*

Proof. See Section IV.8. $\square$

Corollary IV.7. *Fix $n \ge 1$ and $\epsilon \ge 0$. Consider the observation strategy $\widehat{\Psi}^{\epsilon} \triangleq \widehat{\Psi}^{\epsilon}(\phi) \triangleq \{\widehat{\psi}^{\epsilon}_1, \ldots, \widehat{\psi}^{\epsilon}_n\}$, defined inductively by $\widehat{\psi}^{\epsilon}_1 \triangleq h^{\epsilon}_n(\phi)$, and for $2 \le j \le n$, $\widehat{\psi}^{\epsilon}_j \triangleq \widehat{\psi}^{\epsilon}_{j-1} + h^{\epsilon}_{n-j}\left(\Phi^{\widehat{\Psi}^{\epsilon}}_{\widehat{\psi}^{\epsilon}_{j-1}}\right)$. Let $\widehat{\tau}^{\epsilon}(\omega) \triangleq \inf\{\widehat{\psi}^{\epsilon}_j(\omega) : \widehat{\psi}^{\epsilon}_j(\omega) = \widehat{\psi}^{\epsilon}_{j+1}(\omega), 0 \le j \le n-1\} \wedge \left(\widehat{\psi}^{\epsilon}_n(\omega) + t^*_0\left(\Phi^{\widehat{\Psi}^{\epsilon}}_{\widehat{\psi}^{\epsilon}_n}(\omega)\right)\right) \in \mathcal{T}^{\widehat{\Psi}^{\epsilon}}_o$, where $t^*_0(\phi)$ is defined to satisfy $\varphi(t^*_0(\phi), \phi) = \frac{\lambda}{c}$. Then*

$$V_n(\phi) \ge E^{\phi}\left[\int_0^{\widehat{\tau}^{\epsilon}} e^{-\lambda t}\left(\Phi^{\widehat{\Psi}^{\epsilon}} - \frac{\lambda}{c}\right) dt\right] - n\epsilon.$$

IV.3 Convergence to the Continuous Observation Problem

In this section, we will show the extended weak convergence of a discretized quickest detection problem to the (classical) continuous observation quickest detection problem, as formulated in [43], Chapter 4. In all of these problems, the cost functional has the same form, while the dynamics of the underlying odds processes capture the effect of different observation procedures. The theory of extended weak convergence, as developed by Aldous in [1], provides a metric under which convergence of optimal stopping problems and their value functions are guaranteed.

To be more precise, we will show that in a sequence of discrete-time problems, the odds processes $\widetilde{\Phi}^n_t$ extended weak converge to the continuous observation odds process Φ_t^c. We consider two discrete time problems which are essentially equivalent, one of which fits the model of [15]. Studying these problems will give upper bounds for the value function of our (adaptive) n-observation problem because they are more restrictive with respect to admissible observation and stopping strategies: in our n-observation problem, there is complete freedom over both observation and stopping times, whereas in [15] there is freedom over the stopping time but observations are confined to a preset grid. As we will see, the value function $v_c(\phi)$ in the continuous observation problem always gives a lower bound for our value functions $v_n(\phi)$. Therefore, we can construct a sequence of functions $\{\widetilde{v}_n^{\mathfrak{D}}\}_{n\geq 1}$ such that $\widetilde{v}_n^{\mathfrak{D}}(\phi) \geq v_n(\phi) \geq v_c(\phi)$ and $\widetilde{v}_n^{\mathfrak{D}}(\phi) \to v_c(\phi)$, which suffices to show that $v_n(\phi) \downarrow v_c(\phi)$.

IV.3.1 Review of the Continuous Observation Problem and Comparison to the Lump Sum n-Observation Problem

As before, Let X be a standard Brownian motion which gains drift α at the unobservable time Θ, satisfying $P(\Theta = 0) = p$, $P(\Theta \in dt | \Theta > 0) = e^{-\lambda t}$. Let $\mathbb{F}^c$ be the filtration generated by X, and $\mathcal{S}^c$ the set of associated stopping times. In the quickest detection problem with continuous observation, the minimization problem is

$$R^c(p) \triangleq \inf_{\tau \in \mathcal{S}^c} P(\tau < \Theta) + cE[(\tau - \Theta)^+].$$

For details on this problem, see [43], Chapter 4. Here, P and E refer to a probability measure under which $P(\Theta = 0) = p$. As in Proposition 2.1 of [9], we may write

$$R^c(p) = 1 - p + (1-p)cv_c\left(\frac{p}{1-p}\right),$$

where $v_c\left(\frac{p}{1-p}\right) = v_c(\phi) = \inf_{\tau \in \mathcal{S}^c} E\left[\int_0^\infty \left(\Phi_t^c - \frac{\lambda}{c}\right) dt\right]$, and Φ_t^c is the odds process under continuous observation. The dynamics of Φ_t^c are given by the following stochastic differential equation, whose derivation is in [43]:

$$d\Phi_t^c \triangleq \lambda(1+\Phi_t^c)dt + \alpha\Phi_t^c dW_t,$$

with W a standard Brownian Motion. The quickest detection problem is therefore reformulated as an optimal stopping problem on the diffusion Φ^c. The following proposition is intuitively clear, since continuous observation is certainly preferable to being limited to a finite set of observation times. Recall the value function $v_n(\phi)$ of Section IV.2.

Proposition IV.8. *For each n, $v_n(\phi) \geq v_c(\phi)$.*

Proof. Let $n \geq 0$ be fixed. Let $\Psi = \{\psi_1, \ldots, \psi_n\}$ be an admissible observation strategy, as described in Section IV.2. Ψ induces the filtration $\mathbb{F}^\Psi$, along with its set of stopping times $\mathcal{S}^\Psi$. As in [15], the optimal stopping problem associated with the observation strategy Ψ is

$$R^\Psi(p) \triangleq \inf_{\tau\in\mathcal{S}^\Psi} R_\tau^\Psi(p),$$

where $R_\tau^\Psi(p) = P(\tau < \Theta) + cE[(\tau-\Theta)^+]$. By definition, $\left(\mathcal{F}_t^\Psi\right)_{t\geq 0} = \mathbb{F}^\Psi \subset \mathbb{F}^c = (\mathcal{F}_t^c)_{t\geq 0}$, in the sense that $\mathcal{F}_t^\Psi \subset \mathcal{F}_t^c$ for each time t. It follows then that $\mathcal{S}^\Psi \subset \mathcal{S}^c$. Therefore, $R^\Psi(p) \geq R^c(p)$. Writing $R^\Psi(p) = 1 - p + (1-p)cv_\Psi\left(\frac{p}{1-p}\right)$, it follows that $v_\Psi\left(\frac{p}{1-p}\right) \geq v_c\left(\frac{p}{1-p}\right)$. Let $\mathfrak{O}^n$ denote the set of all admissible n-observation strategies. Then

$$v_n(\phi) = \inf_{\Psi\in\mathfrak{O}^n} v_\Psi(\phi) \geq v_c(\phi).$$

$\square$

IV.3.2 Defining the Discretized Problem, and the Convergence Result

We will define two closely related processes, $\widetilde{\Phi}^n$, and $\widetilde{\Phi}^{\mathfrak{O},n}$. Let $\Delta t = \frac{1}{n}$. The process $\widetilde{\Phi}^n$ will be defined on the grid points $\{0, \Delta t, 2\Delta t, \ldots\}$, and then it will be extended to $\mathbb{R}_+$ as a piecewise constant function. Let $\{Z_1, Z_2, \ldots\}$ be a sequence of i.i.d $N(0,1)$ random variables. We define $\widetilde{\Phi}^n$ and $\widetilde{\Phi}^{\mathfrak{D},n}$ recursively, so that they only differ in between grid points.

Definition IV.9. Define the process $\widetilde{\Phi}^n$:

$$\begin{cases} \widetilde{\Phi}^n_0 = \phi, \\ \widetilde{\Phi}^n_{k\Delta t} = j\left(\Delta t, Z_k, \widetilde{\Phi}^n_{(k-1)\Delta t}\right) & \text{for } k \in \mathbb{N}, \\ \widetilde{\Phi}^n_t = \widetilde{\Phi}^n_{(k-1)\Delta t} & \text{for } (k-1)\Delta t \leq t < k\Delta t. \end{cases}$$

Definition IV.10. Define the process $\widetilde{\Phi}^{\mathfrak{D},n}$:

$$\begin{cases} \widetilde{\Phi}^{\mathfrak{D},n}_0 = \phi, \\ \widetilde{\Phi}^{\mathfrak{D},n}_{k\Delta t} = j\left(\Delta t, Z_k, \widetilde{\Phi}^{\mathfrak{D},n}_{(k-1)\Delta t}\right) & \text{for } k \in \mathbb{N}, \\ \widetilde{\Phi}^{\mathfrak{D},n}_t = \varphi\left(\lambda(t-(k-1)\Delta t), \widetilde{\Phi}^{\mathfrak{D},n}_{(k-1)\Delta t}\right) & \text{for } (k-1)\Delta t \leq t < k\Delta t, \end{cases}$$

where $\varphi\left(\lambda(t-(k-1)\Delta t), \widetilde{\Phi}^{\mathfrak{D},n}_{(k-1)\Delta t}\right) = e^{\lambda(t-(k-1)\Delta t)}\left(\widetilde{\Phi}^{\mathfrak{D},n}_{(k-1)\Delta t} + 1\right) - 1$. We remark that the dynamics of $\widetilde{\Phi}^{\mathfrak{D},n}$ are precisely those of our n-observation problem when observations are taken every $\frac{1}{n}$ units of time. Since the gaps between observations are deterministic, they are also the typical example of the model in [15]. The dynamics of $\widetilde{\Phi}^n$ are modified to make computations more tractable. Notice also that $\widetilde{\Phi}^n$ and $\widetilde{\Phi}^{\mathfrak{D},n}$ induce the same filtration. We take $\widetilde{\mathbb{F}}^n$ to be the (continuous time) natural filtration generated by $\widetilde{\Phi}^n_t$, and $\widetilde{\mathcal{T}}^n$ the set of $\widetilde{\mathbb{F}}^n$-stopping times. We set

$$\widetilde{v}^{\mathfrak{D}}_n(\phi) \triangleq \inf_{\tau \in \widetilde{\mathcal{T}}^n} E\left[\int_0^\tau e^{-\lambda s}\left(\widetilde{\Phi}^{\mathfrak{D},n}_s - \frac{\lambda}{c}\right) ds\right].$$

To properly state our result, we need the concept of extended weak convergence, from [1]. We state the definition for the sake of completeness, but we will essentially only need the fact that extended weak convergence implies convergence of optimal stopping problems.

Definition IV.11. Let $(X, \mathbb{F})$ be a random process, considered as a random element in $D(\mathbb{R}_+)$, the set of càdlàg paths on $\mathbb{R}_+$. For each t, there exists a conditional distribution Z_t for X, given $\mathcal{F}_t$, and Z_t may be viewed as a random element of $\mathcal{P}(D(\mathbb{R}_+))$, the set of probabilities on $D(\mathbb{R}_+)$. It is a fact (see Theorem 13.1 of [1]) that these Z_t can be combined to form a càdlàg process taking values in $\mathcal{P}(D(\mathbb{R}_+))$. This process Z is referred to as the prediction process. For processes $(X^n, \mathbb{F}^n)$ and $(X, \mathbb{F})$, we say that X^n extended converges to X, writing $X^n \Rrightarrow X$, if the associated prediction processes Z^n converge weakly to Z, i.e. weak convergence of their induced measures on $\mathcal{P}(D(\mathbb{R}_+))$.

Our principal interest in extended weak convergence is derived from the following (in a slightly weakened form) theorem in [1]. Let $\gamma : [0, \infty) \times \mathbb{R} \to \mathbb{R}$ be bounded and continuous. Given a process $(X, \mathbb{F})$, let $\mathcal{T}_L$ denote its stopping times bounded in size by L, and define

$$\Gamma(L) \triangleq \sup_{T \in \mathcal{T}_L} E\left[\gamma(T, X_T)\right].$$

Proposition IV.12. *[Theorem 17.2, (Aldous)] Suppose $(X^n, \mathbb{F}^n) \Rrightarrow (X^\infty, \mathbb{F}^\infty)$. Suppose $(X^\infty, \mathbb{F}^\infty)$ is quasi left continuous (or continuous), and suppose that $\mathbb{F}^\infty$ is the usual filtration for X^∞. Then $\Gamma_n(L) \to \Gamma_\infty(L)$.*

Our goal, therefore, is to show that $\widetilde{\Phi}^n \Rrightarrow \Phi^c$. The following two results from [1] yield a feasible strategy for establishing extended weak convergence to a diffusion. The effectiveness of this method lies in the fact that the (complicated) limiting process never needs to be directly studied; this is the basic property of establishing weak convergence. The message of these two Propositions is the following: the standard way that one shows weak convergence is Proposition IV.13, but in fact weak convergence is strictly weaker than the conditions in this Proposition. On the other hand, it turns out that these conditions are exactly equivalent to extended weak convergence. Thus, by following the "standard" method for establishing weak convergence, one obtains the more powerful extended weak convergence for free.

Proposition IV.13 (Theorem 8.22, (Aldous))**.** *Let $a(x) > 0$ and $b(x)$ be bounded continuous functions and let $x_0 \in \mathbb{R}$. Let X be the diffusion with drift $b(x)$ and variance $a(x)$, and $X_0 = x_0$. Let $(X^n, \mathbb{F}^n)$ be a sequence of processes. Suppose that for all $L > 0$*

(a) $X_0^n \Rightarrow X_0$

(b) $E\left[\sup_{t \le L} \left(X_t^n - X_{t-}^n\right)^2\right] \to 0$ *as* $n \to \infty$.

Suppose also that for each n, there exist N_t^n and $\mathcal{N}_t^n$ adapted to $\mathbb{F}^n$ such that for all $L > 0$

(c) $(M^n, \mathbb{F}^n)$ *is a martingale, where* $M_t^n = X_t^n - \int_0^t b(X_s^n)ds - N_t^n$

(d) $(S^n, \mathbb{F}^n)$ *is a martingale, where* $S_t^n = (M_t^n)^2 - \int_0^t a(X_s^n)ds - \mathcal{N}_t^n$

(e) $\sup_{T \in \mathcal{T}_L^n} E\left[(N_T^n)^2\right] \to 0$ *as* $n \to \infty$

(f) $\sup_{T \in \mathcal{T}_L^n} E\left[|\mathcal{N}_T^n|\right] \to 0$ *as* $n \to \infty$,

where $\mathcal{T}_L^n$ is the set of $\mathbb{F}^n$-stopping times bounded by L. Then $X^n \Rightarrow X$ (i.e., weak convergence).

Proposition IV.14 (Proposition 21.17, (Aldous))**.** *Let $(Y^n, \mathbb{F}^n)$ be a sequence of processes, and X the diffusion with drift $b(x)$ and variance $a(x)$. In order that $Y^n \Rrightarrow X$ (i.e. extended weak convergence), it is necessary and sufficient that there exist X^n adapted to $\mathbb{F}^n$ such that*

(i) $\sup_{t \le L} |X_t^n - Y_t^n| \to 0$ *in probability*

(ii) $(X^n, \mathbb{F}^n)$ *satisfies the hypotheses of Proposition IV.13.*

Proposition IV.15. *As* $n \to \infty$, $\widetilde{\Phi}_t^n \Rrightarrow \Phi_t^c$

Proof. The proof consists of checking the six conditions in Proposition IV.13, which necessitates establishing some moment inequalities on $\widetilde{\Phi}^n$. We refer the reader to Section IV.9. $\square$

Corollary IV.16. *As* $n \to \infty$, $\widetilde{\Phi}^{\mathfrak{D},n} \Rrightarrow \Phi^c$.

Proof. Let $g_n(\phi) = e^{\frac{\lambda}{n}}(\phi+1) - 1 - \phi = (\phi+1) \cdot O\left(\frac{1}{n}\right)$.

Note that

$$\sup_{0 \le t \le L} |\widetilde{\Phi}_t^n - \widetilde{\Phi}_t^{\mathfrak{D},n}| = g_n\left(\max_{0 \le k \le k_{max}} \widetilde{\Phi}_{\frac{k}{n}}^n\right).$$

As in the proof of Proposition IV.15, $\widetilde{\Phi}^n$ is a submartingale, and so by Doob's L^2 Inequality,

$$E\left[\sup_{0 \le t \le L} |\widetilde{\Phi}_t^n - \widetilde{\Phi}_t^{\mathfrak{D},n}|^2\right] = O\left(\frac{1}{n^2}\right) E\left[\left(1 + \widetilde{\Phi}_T^n\right)^2\right].$$

Using the moment bounds on $\widetilde{\Phi}^n$ established in Section IV.9, we see that this last quantity above is $O\left(\frac{1}{n^2}\right)$. Now, we can see that Condition (i) in Proposition IV.14 is satisfied. Applying it with Proposition IV.15, we deduce the Corollary. $\square$

Corollary IV.17. *As* $n \to \infty$, $\widetilde{v}_n^{\mathfrak{D}}(\phi) \to v_c(\phi)$ *and* $v_n(\phi) \to v_c(\phi)$.

Proof. First, note that for any $\epsilon > 0$, there exists a $L = L(\epsilon)$ such that for all n,

$$\widetilde{v}_n^{\mathfrak{D}}(\phi) > \inf_{\tau \in \widetilde{\mathcal{T}}^n, \tau \le L} E\left[\int_0^\tau e^{-\lambda s}\left(\widetilde{\Phi}_s^{\mathfrak{D},n} - \frac{\lambda}{c}\right) ds\right] - \epsilon,$$

and the same type of inequality holds true for $v_c(\phi)$. This is because the running reward function at time s is greater than $-\frac{\lambda}{c}e^{-\lambda s}$, as $\widetilde{\Phi}^{\mathfrak{D},n}$ and Φ^c are nonnegative. Therefore, the value functions $\widetilde{v}_n^{\mathfrak{D}}$ and v_c are uniformly approximated by problems where the allowed stopping times are uniformly bounded. Therefore, to show that

$\widetilde{v}_n^{\mathfrak{D}}$ converges to v_c, we may assume that all stopping times are bounded by some constant L.

Now, we cannot apply Proposition IV.12 directly, since the value functions $\widetilde{v}_n^{\mathfrak{D}}$ and v_c are optimal stopping problems, not on $\widetilde{\Phi}^{\mathfrak{D},n}$ and Φ^c, but on their time integrals. Fortunately, there is a simple way to work around this technical difficulty, using one last result from [1].

Lemma IV.18. *Let $H : D(\mathbb{R}) \to D(\mathbb{R})$ be a continuous mapping such that if $f(u) = g(u)$ for $u \leq t$ then $(Hf)(t) = (Hg)(t)$. Then if $(X^n, \mathbb{F}^n) \Rrightarrow (X^\infty, \mathbb{F}^\infty)$ and $Y^n = H(X^n)$, $(Y^n, \mathbb{F}^n) \Rrightarrow (Y^\infty, \mathbb{F}^\infty)$.*

For H defined by $(Hf)(t) = \int_0^t f(s)ds$, it is clear that the conditions of Lemma IV.18 are satisfied, at the very least when H is restricted to continuous paths. Therefore, $\int_0^\cdot \widetilde{\Phi}_s^{\mathfrak{D},n} ds \Rrightarrow \int_0^\cdot \Phi_s^c ds$. Therefore, by Proposition IV.12, we have $\widetilde{v}_n^{\mathfrak{D}} \to v_c$. In computing $v_n^{\mathfrak{D}}$, we take $\lfloor Ln \rfloor$ observations, so $v_n^{\mathfrak{D}} \geq v_{\lfloor Ln \rfloor} \geq v_c$. By the monotonicity of v_n with respect to n, it follows that $v_n \to v_c$. □

IV.4 The Stochastic Arrival Rate n-Observation Problem: Setup, Existence of Optimal Strategies

We will consider two subcases of this problem. First, we assume that a total of n observation rights arrive via a Poisson process. Second, we assume that the rates arrive indefinitely from a Poisson process. The second case will be addressed as a limiting case of the former. Suppose that, in addition to supporting a Wiener process X and the random variable Θ, the space (Ω, P) supports an independent, completely observable Poisson process $\{N_t\}_{t \geq 0}$ with arrival rate $\mu > 0$. Let $\eta_1 \leq \eta_2 \leq \cdots$ denote the increasing sequence of jumps times of N. For convenience, take $\eta_0 = 0$. We will define the set of allowed observation strategies for both the "n total observation rights" problem and for the infinite observation rights problem. When we consider the "n total observation rights" problem, we will stop N after n arrivals, and assume that $\eta_{n+1} = \infty$. As before, we will first define the set of admissible observation strategies.

Definition IV.19. For a sequence of random variables $\psi_1 \leq \psi_2 \leq \cdots \leq \psi_n$, we say $\Psi = \{\psi_1, \ldots, \psi_n\}$ is an admissible observation strategy in the stochastic arrival rate n-observation problem, written $\Psi \in \boldsymbol{\mathfrak{O}}^n$, if

$$\psi_j \in m\ \sigma(X_{\psi_1}, \ldots, X_{\psi_{j-1}}, \psi_1, \ldots, \psi_{j-1}, \eta_1, \ldots, \eta_j) \text{ and } \psi_j \geq \eta_j$$

for each $1 \leq j \leq n$. For convenience, we will always set $\psi_0 = 0$ for any Ψ.

Definition IV.20. For $\Psi = \{\psi_1, \psi_2, \ldots\}$, we say that Ψ is an admissible observation strategy in the stochastic arrival rate infinite observation problem, written $\Psi \in \boldsymbol{\mathfrak{O}}^\infty$, if for each n, $\{\psi_1, \ldots, \psi_n\} \in \boldsymbol{\mathfrak{O}}^n$.

Using the same construction as in the previous section, each $\Psi \in \boldsymbol{\mathfrak{O}}^n$, $1 \leq n \leq \infty$, induces a continuous time filtration $\widetilde{\mathbb{F}}^\Psi = (\widetilde{\mathcal{F}}_t^\Psi)_{t \geq 0}$ which is built up from the discrete observations made at times ψ_i. We take $\mathbb{F}^\Psi = \widetilde{\mathbb{F}}^\Psi \vee \mathbb{F}^N$, $\mathbb{F}^N$ being the filtration generated by the Poisson process N. Each Ψ induces the set $\mathcal{T}^\Psi$ of $\mathbb{F}^\Psi$-stopping times and the observed posterior process Φ^Ψ defined by (4.1).

Take $1 \leq n \leq \infty$. As before, according to Lemma 3.1 of [9], the minimum Bayes risk equals $\mathbf{R}_n(p) = 1 - p + (1-p)c\mathbf{V}_n(p/(1-p))$, where

$$\mathbf{V}_n(\phi) \triangleq \inf_{\Psi \in \boldsymbol{\mathfrak{O}}^n} \inf_{\tau \in \mathcal{T}^\Psi} E^\phi \left[\int_0^\tau e^{-\lambda t} \left(\Phi_t^\Psi - \frac{\lambda}{c} \right) dt \right], \tag{4.7}$$

and the expectation $E^\phi[\cdot]$ is with respect to a probability measure P under which X is a standard Weiner process and $\Phi_0^\Psi = \phi$. Hence, we will focus on solving (4.7).

We will first specialize to the case where $n < \infty$. As in the previous section, in considering the problem $\mathbf{V}_n$, we can optimize over a smaller set of stopping times than $\mathcal{T}^\Psi$.

Definition IV.21. Let $\Psi \in \mathfrak{O}^n$, and let $\tau \in \mathcal{T}^\Psi$. We say that $\tau \in \mathcal{T}_s^\Psi$ if $\{\psi_i < \tau < \psi_{i+1}\} \cap \{\psi_i \geq \eta_{i+1}\} = \emptyset$, for each $0 \leq i \leq n-1$.

Note that $\{\psi_i \geq \eta_{i+1}\}$ represents the scenarios when, after making observation i, the agent has additional observation rights stockpiled. Therefore, a stopping time $\tau \in \mathcal{T}_s^\Psi$ is one that does not stop while there are unused observation rights. As in Section IV.2, we have

Proposition IV.22. $\boldsymbol{V}_n(\phi) = \inf_{\Psi \in \mathfrak{O}^n} \inf_{\tau \in \mathcal{T}_s^\Psi} E^\phi \left[\int_0^\tau e^{-\lambda t} \left(\Phi_t^\Psi - \frac{\lambda}{c} \right) dt \right].$

Proof. Let $\Psi \in \mathfrak{O}^n$, and $\tau \in \mathcal{T}^\Psi$. First, note that every stopping time $\tau \in \mathcal{T}^\Psi$ satisfies $\{\psi_0 < \tau < \psi_1\} \cap \{\psi_0 \geq \eta_1\} = \emptyset$, simply because $\{\psi_0 \geq \eta_1\} = \emptyset$, η_1 being a strictly positive random variable. The proof now is essentially identical to that of Proposition IV.2. As before, if we were to stop the game while having unused observation rights, we could construct a new observation strategy which adds in an additional observation at that stopping time, without changing the value received. $\square$

Now, we will define some operators, which have analogs in the lump sum n-observation problem. Let $\Lambda^F \subset \mathbb{R}_+^2$ denote the set of feasible values of the state process (t, Φ_t). Precisely:

$$\Lambda^F = \left\{ (y, \phi) : y \geq 0, \phi \geq e^{\lambda y} - 1 \right\}.$$

Here y represents the time since the last observation. In the absence of observations, the trajectory of the state process follows the path $(t, e^{\lambda t}(\phi + 1) - 1)$, starting from ϕ at time zero. Since $\phi \geq 0$ at time zero, all trajectories must lie in Λ^F.

Recall the operator K from (4.3). We will extend it as follows: for $w : \Lambda^F \to \mathbb{R}$ bounded, define

$$\boldsymbol{K} w(t,\phi) \triangleq \int_{-\infty}^{\infty} w(0, j(t,\phi,z)) \frac{\exp(-z^2/2)}{\sqrt{2\pi}} dz. \tag{4.8}$$

In the next two operators, the "0" superscript stands for "no observations stockpiled". We define, for $w : \Lambda^F \to \mathbb{R}$ bounded,

$$J^0 w(t,y,\phi)$$

(4.9)
$$\triangleq \int_0^\infty \mu e^{-\mu u} \left(\int_0^{u \wedge t} e^{-\lambda r} \left(\varphi(r,\phi) - \frac{\lambda}{c} \right) dr + e^{-\lambda u} 1_{\{t>u\}} w(y+u, \varphi(u,\phi)) \right) du,$$

$$J_0^0 w(y,\phi) \triangleq \inf_{t \geq 0} J^0 w(t,y,\phi). \tag{4.10}$$

Let us explain the operator J^0. It describes the situation in which the agent has no observations stockpiled, the posterior is ϕ, and y units of time have passed since the last observation was made. Faced with this scenario, he stops at time t, which may be prior to the arrival time u of the next observation right, or after it. An agent will be left with no observations stockpiled only if he has just used an observation, so for these operators y will effectively be zero. For subsequent operators, we will consider scenarios where y is positive, and so for this reason we keep the notation consistent.

Next we define jump operators J^+ and J_0^+, corresponding to the scenario when the agent has stockpiled observation rights after he has either just made an observation or received an observation right. We define, for $w^1, w^2 : \Lambda^F \to \mathbb{R}$ bounded,

(4.11) $J^+(w^1, w^2)(t,y,\phi)$
$$\triangleq \int_0^\infty \mu e^{-\mu u} \Bigg(\int_0^{u \wedge t} e^{-\lambda r} \left(\varphi(r,\phi) - \frac{\lambda}{c} \right) dr$$
$$+ e^{-\lambda t} 1_{\{t<u\}} \boldsymbol{K} w^1 \left(y+t, \varphi(-y,\phi) \right) + e^{-\lambda u} 1_{\{u \leq t\}} w^2(y+u, \varphi(u,\phi)) \Bigg) du,$$

$$J_0^+(w^1,w^2)(y,\phi) \triangleq \inf_{t\geq 0} J^+(w^1,w^2)(t,y,\phi). \tag{4.12}$$

From Proposition IV.22, we have seen that it is never optimal for an agent to stop while he has unused observation rights. Therefore, if he has observation rights stock-piled, the agent either observes immediately ($t=0$), which is equivalent to stopping, or chooses his next observation time $t>0$. If u is the next arrival time of an additional observation right, then his next observation time t may be either prior to or after the arrival of the next observation right. Here of the two continuation functions w^1 and w^2, w^1 corresponds to this former scenario, and w^2 to the latter. The variable y denotes the amount of time that has passed since the agent has last made an observation, which may be nonzero if an observation right has arrived more recently than the last time an observation was made.

We will need one more pair of operators, corresponding to the times when all n observation rights have been received. Note that this scenario explains why the "lump sum n observation rights" problem is essentially embedded in this one. Therefore, note the similarity between J^e, J_0^e, defined below, and J, J_0, defined in (4.5). The main difference consists in allowing y to be nonzero, allowing for the possibility that time has elapsed since the last observation. We define, for $w:\Lambda^F\to\mathbb{R}$ bounded,

$$J^e w(t,y,\phi) \triangleq \int_0^t e^{-\lambda r}\left(\varphi(r,\phi)-\frac{\lambda}{c}\right)dr + e^{-\lambda t}\boldsymbol{K}w(y+t,\varphi(-y,\phi)), \tag{4.13}$$

$$J_0^e w(y,\phi) \triangleq \inf_{t\geq 0} J^e w(t,y,\phi). \tag{4.14}$$

Fix $1\leq n<\infty$. Set $\boldsymbol{v}^n_{n,n+1}(y,\phi)\triangleq 0$. For $0\leq k\leq n$, set $\boldsymbol{v}^n_{n,k}(y,\phi)\triangleq J_0^e\boldsymbol{v}^n_{n,k+1}(y,\phi)$. The superscript "n" corresponds to n total observation rights, while the subscript "n,k" corresponds to n observation rights received and k observations used. Note that when there are n observation rights arriving stochastically, it is the case that once all of these n observations have arrived, we essentially revert to the lump sum problem. We now define $\boldsymbol{v}^n_{n-1,n-1}(y,\phi) \triangleq J^0_0 \boldsymbol{v}^n_{n,n-1}(y,\phi)$ and, for $0 \le k < n-1$, $\boldsymbol{v}^n_{n-1,k}(y,\phi) \triangleq J^+_0(\boldsymbol{v}^n_{n-1,k+1}, \boldsymbol{v}^n_{n,k})(y,\phi)$. Proceeding inductively in this way, we define

$$\boldsymbol{v}^n_{j,j}(y,\phi) \triangleq J^0_0(\boldsymbol{v}^n_{j+1,j})(y,\phi), 0 \le j \le n,$$

$$\boldsymbol{v}^n_{j,k}(y,\phi) \triangleq J^+_0(\boldsymbol{v}^n_{j,k+1}, \boldsymbol{v}^n_{j+1,k})(y,\phi), 0 \le k < j \le n.$$

The function $\boldsymbol{v}^n_{j,k}(y,\phi)$ is a value function, representing the value when there are n total observation rights, of which j have been received and $k \le j$ spent, the current posterior level is ϕ, and y units of time have elapsed since the last observation was made. Note that by definition of J^0_0, $\boldsymbol{v}^n_{j,j} \le 0$, $0 \le j \le n$. From this and the definition of J^+_0, it also follows that $\boldsymbol{v}^n_{j,k} \le 0$ for all j and k. We illustrate the relationship between the value functions and the jump operators through the following figures.

Figure 4.1: Schematic of Stochastic Arrival Problem, $n = 4$

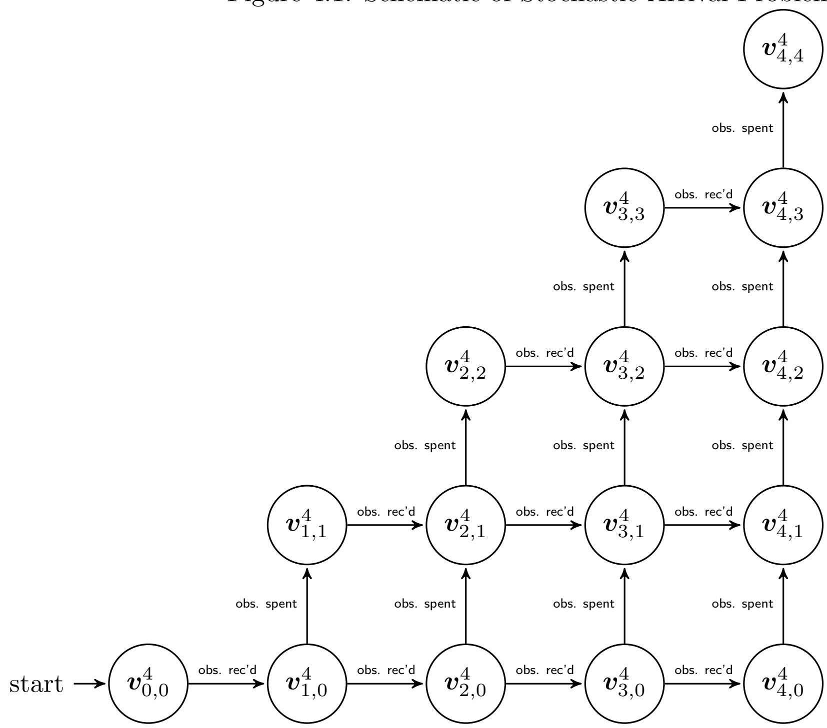
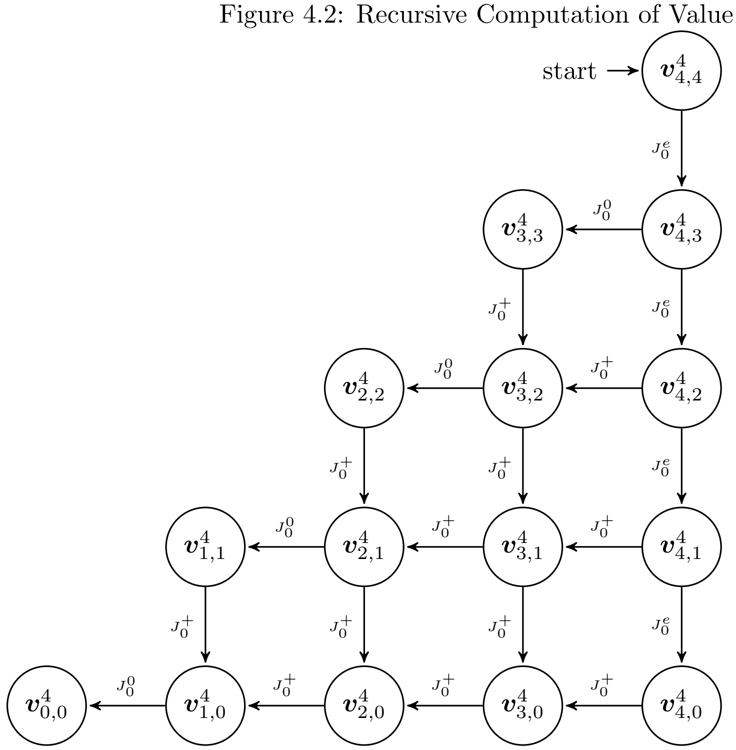


Figure 4.2: Recursive Computation of Value Functions, $n = 4$



Fix $0 \le k \le n$, and let $\Psi^k = \{\psi_1^k, \dots, \psi_k^k\} \in \mathfrak{O}^k$, from Definition IV.19. We set, for $k \le j \le n$,

$$\mathfrak{O}_{j,k}^n(\Psi^k) \triangleq \left\{\Psi = \{\psi_1, \dots, \psi_n\} \in \mathfrak{O}^n : \psi_i = \psi_i^k, 1 \le i \le k \text{ and } \psi_{k+1} \ge \eta_j\right\}.$$

Intuitively, $\mathfrak{O}_{j,k}^n(\Psi^k)$ consists of the observation strategies one can pursue after observing at $\psi_1^k, \dots, \psi_k^k$, and refraining from observing next until η_j. Note that the last requirement $\psi_{k+1} \ge \eta_j$ is vacuous when $j = k, k+1$. We let, for $0 \le k \le n$ and $k \le j \le n$,

$$\boldsymbol{\gamma}_{j,k}^n(\Psi^k) \triangleq \operatorname*{ess\,inf}_{\Psi \in \mathfrak{O}_{j,k}^n(\Psi^k)} \operatorname*{ess\,inf}_{\tau \in \mathcal{T}_s^\Psi, \tau \ge \psi_k^k \vee \eta_j} E\left[\int_{\psi_k^k \vee \eta_j}^{\tau} e^{-\lambda(t - \psi_k^k \vee \eta_j)} \left(\Phi_t^\Psi - \frac{\lambda}{c}\right) dt \Big| \mathcal{F}^{\Psi^k}_{\psi_k^k \vee \eta_j}\right].$$

Note that the "reference" time above is $\psi_k^k \vee \eta_j$. We are in a scenario where j observation rights have been received and k spent; if $\psi_k^k > \eta_j$, we arrived at this state from "j observation rights received, $k-1$ observations spent", and if $\eta_j > \psi_k^k$, we arrived at this state from "$j-1$ observations received, k observations spent".

Proposition IV.23. *For any n, $0 \le k \le j \le n$ and $\Psi^k = \{\psi_1^k, \dots, \psi_k^k\} \in \mathfrak{O}^k$,*

$$\boldsymbol{\gamma}_{j,k}^n(\Psi^k) \ge \boldsymbol{v}_{j,k}^n\left(\psi_k^k \vee \eta_j - \psi_k^k, \Phi^{\Psi^k}_{\psi_k^k \vee \eta_j}\right) \tag{4.15}$$

on the set $\{\psi_k^k < \eta_{j+1}\}$.

Proof. See Section IV.10. □

For the proof of the other inequality, we will need to construct some optimal stopping times, describing when one should either observe the process or stop and accept the change hypothesis. We will do this inductively, with the help of some auxiliary functions. Set $s_{n,n}^n(y, \phi) = s_{n,n}^n(\phi) \triangleq t_0^*(\phi)$, defined by

$$t_0^*(\phi) = \frac{1}{\lambda} \log\left(\frac{c + \lambda}{c(\phi + 1)}\right) \vee 0. \tag{4.16}$$

For $0 \le k < n$, define $o^n_{n,k}(y,\phi) \triangleq \inf\left\{s \ge 0 : J^e \boldsymbol{v}^n_{n,k+1}(s,y,\phi) = J^e_0 \boldsymbol{v}^n_{n,k+1}(y,\phi)\right\}$. We define, for $0 \le j < n$,

$$s^n_{j,j}(y,\phi) \triangleq \inf\left\{s \ge 0 : J^0 \boldsymbol{v}^n_{j+1,j}(s,y,\phi) = J^0_0 \boldsymbol{v}^n_{j+1,j}(y,\phi)\right\}$$

and, for $0 \le j < n$, $0 \le k < j$,

$$o^n_{j,k}(y,\phi) \triangleq \inf\left\{s \ge 0 : J^+\left(\boldsymbol{v}^n_{j,k+1}, \boldsymbol{v}^n_{j+1,k}\right)(s,y,\phi) = J^+_0\left(\boldsymbol{v}^n_{j,k+1}, \boldsymbol{v}^n_{j+1,k}\right)(y,\phi)\right\}.$$

The notation "s" and "o" stands for, respectively, stop, and observe. This is in line with the reasoning that one should stop only when there are no available observation rights, i.e. $j = k$.

For $\Psi^k = \{\psi^k_1, \ldots, \psi^k_k\} \in \boldsymbol{\mathfrak{O}}^k$, we define the "action times" (either stopping or making an observation) $\widehat{\tau}^n_{j,k}$, $k \le j \le n$. Set

$$\begin{aligned}\widehat{\tau}^n_{n,k} &= \widehat{\tau}^n_{n,k}(\Psi^k)\\ &\triangleq \psi^k_k \vee \eta_n + o^n_{n,k}\left(\psi^k_k \vee \eta_n - \psi^k_k, \Phi^{\Psi^k}_{\psi^k_k \vee \eta_n}\right),\end{aligned}$$

and we will inductively define, on the set $\{\psi^k_k < \eta_{j+1}\}$, $k < j < n$,

$$\begin{aligned}&\widehat{\tau}^n_{j,k} = \widehat{\tau}^n_{j,k}(\Psi^k)\\ &\quad\triangleq \begin{cases}\psi^k_k \vee \eta_j + o^n_{j,k}\left(\psi^k_k \vee \eta_j - \psi^k_k, \Phi^{\Psi^k}_{\psi^k_k \vee \eta_j}\right) & \text{if } \psi^k_k \vee \eta_j + o^n_{j,k}\left(\psi^k_k \vee \eta_j - \psi^k_k, \Phi^{\Psi^k}_{\psi^k_k \vee \eta_j}\right) < \eta_{j+1}\\ \widehat{\tau}^n_{j+1,k}(\Psi^k) & \text{if } \psi^k_k \vee \eta_j + o^n_{j,k}\left(\psi^k_k \vee \eta_j - \psi^k_k, \Phi^{\Psi^k}_{\psi^k_k \vee \eta_j}\right) \ge \eta_{j+1}\end{cases}\end{aligned}$$

and

$$\begin{aligned}\widehat{\tau}^n_{k,k} &= \widehat{\tau}^n_{k,k}(\Psi^k)\\ &\triangleq \begin{cases}\psi^k_k + s^n_{k,k}\left(\Phi^{\Psi^k}_{\psi^k_k}\right) & \text{if } \psi^k_k + s^n_{j,k}\left(\Phi^{\Psi^k}_{\psi^k_k}\right) < \eta_{k+1}\\ \widehat{\tau}^n_{k+1,k}(\Psi^k) & \text{if } \psi^k_k + s^n_{k,k}\left(\Phi^{\Psi^k}_{\psi^k_k}\right) \ge \eta_{k+1}.\end{cases}\end{aligned}$$

Proposition IV.24. *For any n, $0 \le k \le j \le n$ and $\Psi^k = \{\psi^k_1, \ldots, \psi^k_k\} \in \boldsymbol{\mathfrak{O}}^k$, on the set $\{\psi^k_k < \eta_{j+1}\}$,*

$$\boldsymbol{\gamma}^n_{j,k}(\Psi^k) \le \boldsymbol{v}^n_{j,k}\left(\psi^k_k \vee \eta_j - \psi^k_k, \Phi^{\Psi^k}_{\psi^k_k \vee \eta_j}\right). \tag{4.17}$$

Hence, $\boldsymbol{\gamma}^n_{j,k}(\Psi^k) = \boldsymbol{v}^n_{j,k}\left(\psi^k_k \vee \eta_j - \psi^k_k, \Phi^{\Psi^k}_{\psi^k_k \vee \eta_j}\right)$. *Furthermore, on the set* $\{\psi^j_j < \eta_{j+1}\}$,

$$\boldsymbol{v}^n_{j,j}\left(0, \Phi^{\Psi^j}_{\psi^j_j}\right) = E\left[\int_{\psi^j_j}^{\widehat{\tau}^n_{j,j}\wedge\eta_{j+1}} e^{-\lambda(s-\psi^j_j)}\left(\varphi\left(s-\psi^j_j, \Phi^{\Psi^j}_{\psi^j_j}\right) - \frac{\lambda}{c}\right)ds + e^{-\lambda(\eta_{j+1}-\psi^j_j)} 1_{\{\widehat{\tau}^n_{j,j} > \eta_{j+1}\}} \boldsymbol{v}^n_{j+1,j}\left(\eta_{j+1}-\psi^j_j, \varphi\left(\eta_{j+1}-\psi^j_j, \Phi^{\Psi^j}_{\psi^j_j}\right)\right) | \mathcal{F}^{\Psi^j}_{\psi^j_j}\right], \tag{4.18}$$

and for $k < j$, *on the set* $\{\psi^k_k < \eta_{j+1}\}$,

$$\begin{aligned}\boldsymbol{v}^n_{j,k}&\left(\psi^k_k \vee \eta_j - \psi^k_k, \Phi^{\Psi^k}_{\psi^k_k\vee\eta_j}\right)\\ &= E\Bigg[\int_{\psi^k_k\vee\eta_j}^{\widehat{\tau}^n_{j,k}\wedge\eta_{j+1}} e^{-\lambda(s-\psi^k_k\vee\eta_j)}\left(\varphi\left(s-\psi^k_k\vee\eta_j, \Phi^{\Psi^k}_{\psi^k_k\vee\eta_j}\right) - \frac{\lambda}{c}\right)ds\\ &\quad + e^{-\lambda(\widehat{\tau}^n_{j,k}-\psi^k_k\vee\eta_j)} 1_{\{\widehat{\tau}^n_{j,k} < \eta_{j+1}\}} \boldsymbol{K}\boldsymbol{v}^n_{j,k+1}\left(\widehat{\tau}^n_{j,k} - \psi^k_k, \varphi\left(-(\psi^k_k\vee\eta_j - \psi^k_k), \Phi^{\Psi^k}_{\psi^k_k\vee\eta_j}\right)\right)\\ &\quad + e^{-\lambda(\eta_{j+1}-\psi^k_k\vee\eta_j)} 1_{\{\eta_{j+1}\le\widehat{\tau}^n_{j,k}\}} \boldsymbol{v}^n_{j+1,k}\left(\eta_{j+1}-\psi^k_k, \varphi\left(\eta_{j+1}-\psi^k_k\vee\eta_j, \Phi^{\Psi^k}_{\psi^k_k\vee\eta_j}\right)\right) | \mathcal{F}^{\Psi^k}_{\psi^k_k\vee\eta_j}\Bigg].\end{aligned} \tag{4.19}$$

In particular, $\boldsymbol{V}_n(\phi) = \boldsymbol{v}^n_{0,0}(\phi)$.

Proof. See Section IV.10. $\square$

As a consequence of Proposition IV.24, we may inductively describe the optimal observation strategies and stopping times, which are as follows. Consider a given instant of time, when an observation has just been spent or an observation right has just been received. Let j be the number of observation rights received, $k \le j$ be the number of observation rights used, let ϕ be the current value of the posterior process, and let y be the amount of time elapsed since the last time an observation was made.

Corollary IV.25. *The following observation/stopping strategy is optimal. Suppose that an observation has just been made or an observation right has just been received.*

Observation Strategy

(1) If $k = j$, *there are no available observation rights. Wait until time* η_{j+1}, *increment* j *by* 1, *and proceed to (2) with the appropriate changes to* ϕ, y.

(2) If $k < j$, calculate $\widehat{\tau}^n_{j,k}$, which is a function of j, k, ϕ, y, and the current time. If $\widehat{\tau}^n_{j,k} < \eta_{j+1}$, spend an observation right at time $\widehat{\tau}^n_{j,k}$, and increment k by 1. If $k+1 < j$, proceed to (2) and if $k+1 = j$, proceed to (1), making the appropriate changes to ϕ and y. Otherwise, if $\widehat{\tau}^n_{j,k} \geq \eta_{j+1}$, increment j by 1, and proceed to (2) with the appropriate changes to ϕ and y.

Stopping Strategy *If $k = j$ and $\widehat{\tau}^n_{j,j} < \eta_{j+1}$, stop the game at time $\widehat{\tau}^n_{j,j}$. If $\widehat{\tau}^n_{j,j} \geq \eta_{j+1}$, increment j by 1 and proceed to Observation Strategy (2). The game is never stopped when $k < j$.*

Remark IV.26. In the corollary above, we say that the agent optimally stops the game only when he has no spare observation rights. This is essentially a formalism. One can envision a scenario in which the agent makes an observation, and notices that the posterior is at a very high level, indicating that it is very likely that the disorder has occurred. The agent will want to stop the game immediately. In our setup, the agent, if he has spare observation rights, will exercise them all instantaneously to get to the point where he has no observation rights remaining, after which he will stop the game.

IV.4.1 The Infinite Horizon Problem

We consider now the subcase of the stochastic arrival problem in which observation rights continue to arrive indefinitely. The following proposition says that the value function in the infinite arrival problem is approximated uniformly by the value function in the n arrival problem, as n goes to infinity. Since the strategy space for the n arrival problem is contained within the strategy space for the infinite arrival problem, it therefore follows that we may use optimal strategies in the n arrival problem to find near optimal strategies in the infinite arrival problem.

Proposition IV.27. *The value functions $\boldsymbol{V}_n(\phi)$ converge to $\boldsymbol{V}_\infty(\phi)$ as $n \to \infty$, uniformly over ϕ. More precisely,* $0 \leq \boldsymbol{V}_\infty - \boldsymbol{V}_n \leq \frac{1}{c}\left(\frac{\mu}{\mu+\lambda}\right)^{n+1}$.

Proof. The inequality $\boldsymbol{V}_\infty \leq \boldsymbol{V}_n$ is an immediate consequence of the fact that $\mathfrak{O}^n$ is naturally included in $\mathfrak{O}^\infty$, i.e. for any element Ψ of $\mathfrak{O}^n$, there is an element $\widetilde{\Psi}$ of $\mathfrak{O}^\infty$ such that the first n observation times of $\widetilde{\Psi}$ coincide with those of Ψ.

For the second inequality, let $\Psi = \{\psi_1, \psi_2, \ldots\}$ be an arbitrary element of $\mathfrak{O}^\infty$. By definition, it must be the case that $\psi_{n+1} \geq \eta_{n+1}$, and that $\{\psi_1, \ldots, \psi_n\}$ is an element of $\mathfrak{O}^n$. Noting that for all Ψ, the posterior process Φ^Ψ is positive, it follows that

$$\boldsymbol{V}_\infty - \boldsymbol{V}_n \geq E\left[\int_{\eta_{n+1}}^\infty e^{-\lambda s}\frac{-\lambda}{c}ds\right] = \frac{1}{c}E\left[e^{-\lambda\eta_{n+1}}\right]. \tag{4.20}$$

Now, η_{n+1}, being the sum of $n+1$ independent exponential random variables with parameter μ, has the Erlang distribution $\eta_{n+1} \sim \text{Erlang}(n+1, \mu)$, which has Laplace transform $f^*(s) = \left(\frac{\mu}{\mu+s}\right)^{n+1}$. It therefore follows that the right hand side of (4.20) above is equal to $\frac{1}{c}\left(\frac{\mu}{\mu+\lambda}\right)^{n+1}$, which tends to zero as $n \to \infty$. □

Remark IV.28. A similar argument may be used to show the convergence of $\boldsymbol{v}^n_{0,0}(0,\phi)$ (n total observations arriving stochastically, of which none have yet arrived) to $v_n(\phi)$, (n total observations, all of which are available), as the *arrival rate* $\mu \to \infty$. Suppose that for some ϕ, $t^*_n(\phi)$, the optimal time to make the first observation in the lump sum n-observation problem, is strictly positive. Using the cumulative distribution of an Erlang random variable, it is easily calculated that the probability that all n observation rights arrive before time $t^*_n(\phi)$ is

$$1 - \sum_{k=0}^{n-1}\frac{1}{k!}e^{-\mu t^*_n(\phi)}(\mu t^*_n(\phi))^k.$$

When n and ϕ are fixed, so that $t^*_n(\phi)$ is fixed, this expression converges to 1 almost exponentially fast as $\mu \to \infty$. If all observation rights arrive before time

$t_n^*(\phi)$, then the stochastic arrival of the observation rights imposes no restriction on observation strategies vis-a-vis the scenario in which all n-observation rights are available all along, because the agent has received all observation rights by the time he wishes to make even a single observation. Therefore, with probability at least $1 - \sum_{k=0}^{n-1} \frac{1}{k!} e^{-\mu t_n^*(\phi)} (\mu t_n^*(\phi))^k$ the strategy that one would pursue in the Lump Sum n-observation problem is also feasible in the stochastic arrival rate problem. This implies that for fixed n and ϕ, $\boldsymbol{v}_{0,0}^n(0, \phi)$ should converge at least almost exponentially fast to $v_n(\phi)$ as $\mu \to \infty$. Note that a uniform rate of convergence over all ϕ is not guaranteed. When $t_n^*(\phi)$ is very close to zero, it becomes increasingly important to have observation rights immediately available. Additionally, this argument does not hold uniformly over all n as $n \to \infty$. In fact, the convergence rate of $\boldsymbol{v}_{0,0}^\infty(0, \phi)$ to $v_c(\phi)$ as a function of μ will be comparable to the convergence of $v_n(\phi)$ to $v_c(\phi)$ as a function of n, which is rather slow (see Table 4.1). This is because, in any finite time interval, the expected number of received observation rights will be proportional to the arrival rate μ.

IV.5 An Algorithm for the Lump Sum n-Observation Problem

In this section, we explicitly describe an algorithm for computing the value functions $v_0, v_1, \ldots, v_N$, as well as the boundaries which determine when observations should be made. We give a rigorous construction which shows how solutions may be constructed up to any specified error tolerance. Actual code, in which certain heuristics are used to speed up implementation, will be given in Appendix B. We have the following main result in this section, giving worst case error bounds:

Proposition IV.29. *Fix a positive integer N. Then in $O(\frac{N^6}{\epsilon^3})$ function evaluations, we may uniformly approximate $v_0(\phi), v_1(\phi), \ldots, v_N(\phi)$ to within ϵ.*

We note that in the process of calculating the value functions, we also determine the boundaries which determine the optimal observation behavior. We outline the steps of the algorithm in Subsection IV.5.1. In Subsection IV.5.2, we justify the error bounds of Step 2 of the algorithm. In Subsection IV.5.3, we explain the error bounds of Step 3, as well as explaining how an upper bound $\overline{\phi}$ may be constructed. Finally, in Subsection IV.5.4, we give error bounds for iterating Steps 2 and 3 multiple times.

IV.5.1 Pseudo-code for the Algorithm

Here we outline the steps of the algorithm. Subsequent parts of this section will explain why such an algorithm works to uniformly approximate the value functions.

(1) Fix N, the total number of observations. Discretize the ϕ variable into $\phi_0 = 0, \phi_1, \ldots, \phi_J = \overline{\phi}$, and set all value functions equal to zero for $\phi \geq \overline{\phi}$.

(2) The function $v_0(\phi)$ can be analytically computed. Fix ϕ_j, and approximately minimize $t \mapsto Jv_0(t, \phi_j)$ by computing $Jw(t_i, \phi_j)$, for $t_0 = 0, t_1, \ldots, t_K = T$ a discretization of t, and T an upper bound on the size of optimal t, established in Lemma IV.32. Let the minimizer be $\widehat{t}_n^*(\phi_j)$.

(3) Having computed above an approximation to $v_1(\phi_j)$, interpolate these values in a piecewise constant fashion to obtain a function $\widehat{v}_1(\phi)$ which approximates $v_1(\phi)$.

(4) Let the collection of points $(\widehat{t}_n^*(\phi_j), \varphi(\widehat{t}_n^*(\phi_j), \phi_j))$ define the observation barrier.

(5) Repeat Steps 2, 3, except now minimizing $t \mapsto J\widehat{v}_1(t, \phi)$, to obtain an approximation $\widehat{v}_2(\phi)$ to $v_2(\phi)$.

(6) Continue this procedure until $\widehat{v}_N(\phi)$ is computed.

IV.5.2 Minimizing $t \mapsto Jw(t,\phi)$ for ϕ Fixed

Lemma IV.30. *For each $n \geq 0$, $\frac{-1}{c} \leq V_n(\phi) \leq 0$ for all ϕ.*

Proof. According to Proposition IV.5, $V_n = v_n$, where $v_{-1} \equiv 0$, and for $n \geq 0$, $v_n = J_0 v_{n-1}$. Here J_0 is the jump operator defined in Section IV.2. Note that 0 clearly satisfies the conclusion of the lemma. Therefore, it suffices to establish the inductive step: if $\frac{-1}{c} \leq v_n \leq 0$, then $-\frac{1}{c} \leq J_0 v_n \leq 0$. The upper bound follows from $J_0 v_n(\phi) \leq J v_n(0,\phi) = 0$. For the lower bound, calculate that for any t,

$$\begin{aligned}
&Jv_n(t,\phi) \\
&\quad = \int_0^t e^{-\lambda u}\left(\varphi(u,\phi) - \frac{\lambda}{c}\right) du + e^{-\lambda t} K v_n(t,\phi) \\
&\quad \geq \int_0^t e^{-\lambda u}\left(\varphi(u,\phi) - \frac{\lambda}{c}\right) du + e^{-\lambda t}\left(\frac{-1}{c}\right) \\
&\quad \geq \int_0^t e^{-\lambda u}\left(-\frac{\lambda}{c}\right) du + e^{-\lambda t}\left(\frac{-1}{c}\right) \\
&\quad = \frac{-1}{c}.
\end{aligned}$$

Taking the infimum over all t yields $J_0 v_n(\phi) \geq \frac{-1}{c}$. □

Lemma IV.31. *For each $n \geq 0$, $v_n(\phi)$ is concave and increasing in ϕ.*

Proof. Follows by definition, and the fact that an infimum of concave functions is again concave. □

Lemma IV.32. *For $T \triangleq \frac{1}{\lambda}\left(1 + \frac{\lambda}{c}\right) + \frac{1}{c}$ and each n,*

$$\begin{aligned}
v_n(\phi) &= J_0 v_{n-1}(\phi) \\
&= \inf_{0 \leq t \leq T} \int_0^t e^{-\lambda u}\left(\varphi(u,\phi) - \frac{\lambda}{c}\right) du + e^{-\lambda t} K v_{n-1}(t,\phi).
\end{aligned}$$

In other words, the optimal time t^* can be assumed to be less than or equal to T.

Proof. Note that

$$\int_0^t e^{-\lambda u}\left(\varphi(u,\phi)-\frac{\lambda}{c}\right)du = \int_0^t \left[\phi+1-e^{-\lambda u}\left(1+\frac{\lambda}{c}\right)\right]du$$
$$= (\phi+1)t+\frac{1}{\lambda}\left(1+\frac{\lambda}{c}\right)(e^{-\lambda t}-1)$$
$$\geq t-\frac{1}{\lambda}\left(1+\frac{\lambda}{c}\right).$$

It follows therefore, that for $t \geq \frac{1}{\lambda}\left(1+\frac{\lambda}{c}\right)+\frac{1}{c}=T$, $Jv_n(t,\phi)\geq 0$ for any n. Here we have used the uniform lower bound for v_n established in Lemma IV.30. By the upper bound in that Lemma, $v_n \leq 0$, so it is sufficient to minimize $Jv_n(t,\phi)$ over $t\in[0,T]$. □

Lemma IV.33. *Let* $||\cdot||_{Lip}$ *denote the Lipschitz norm. For each* $n\geq 0$, $||v_n||_{Lip}\leq T+||v_{n-1}||_{Lip}$.

Proof. Take $\phi_1<\phi_2$. We have, using Lemma IV.32 for the first inequality,

$$|v_n(\phi_1)-v_n(\phi_2)| \leq \sup_{0\leq t\leq T}\left|\int_0^t \left[e^{-\lambda u}\left(\varphi(u,\phi_1)-\frac{\lambda}{c}\right)-e^{-\lambda u}\left(\varphi(u,\phi_2)-\frac{\lambda}{c}\right)\right]du\right|$$
$$+\sup_{0\leq t\leq T}\left|e^{-\lambda t}\left(Kv_{n-1}(t,\phi_1)-Kv_{n-1}(t,\phi_2)\right)\right|.$$

We treat these two terms on the right hand side separately. We calculate that the first term is actually equal to

$$\sup_{0\leq t\leq T}\left|\int_0^t(\phi_1-\phi_2)du\right| = T|\phi_1-\phi_2|.$$

To calculate the second term, fix $t\in[0,T]$. Then

$$\left|e^{-\lambda t}\int_{-\infty}^{\infty}\Big(v_{n-1}(j(t,\phi_1,z))-v_{n-1}(j(t,\phi_2,z))\Big)\frac{e^{-z^2/2}}{\sqrt{2\pi}}dz\right|$$
$$\leq ||v_{n-1}||_{Lip}\ e^{-\lambda t}\int_{-\infty}^{\infty}|j(t,\phi_1,z)-j(t,\phi_2,z)|\frac{e^{-z^2/2}}{\sqrt{2\pi}}dz$$
$$=||v_{n-1}||_{Lip}e^{-\lambda t}|\phi_1-\phi_2|\int_{-\infty}^{\infty}e^{\alpha z\sqrt{t}+(\lambda-\alpha^2/2)t}\frac{e^{-z^2/2}}{\sqrt{2\pi}}dz$$
$$=||v_{n-1}||_{Lip}e^{-\lambda t}|\phi_1-\phi_2|e^{\lambda t}$$
$$=||v_{n-1}||_{Lip}|\phi_1-\phi_2|,$$

where the first equality uses the definition of $j(t, \phi, z)$, in Section IV.2. It now follows that $|v_n(\phi_1) - v_n(\phi_2)| \leq (T + ||v_{n-1}||_{Lip}) \, |\phi_1 - \phi_2|$. $\square$

Lemma IV.34. *The mapping $t \mapsto Jv_n(t, \phi)$ is $\frac{1}{2}$-Hölder continuous. In particular, $|\frac{d}{dt} Jv_n(t, \phi)| \leq \phi + a + b||v_n||_{Lip} t^{-1/2}$, for constants*

$$a = \left(1 + \frac{\lambda}{c} + \frac{1}{c}\right) e^{-\lambda t},$$
$$b = \left(\phi\left(\frac{1}{2}\alpha C_1 + \lambda\right) + \lambda + \frac{1}{2\lambda} C_1\right).$$

Proof. We calculate that

$$\frac{d}{dt} Jv_n(t, \phi) = \phi + 1 - e^{-\lambda t}\left(1 + \frac{\lambda}{c}\right) + e^{-\lambda t}\frac{d}{dt} Kv_n(t, \phi) - \lambda e^{-\lambda t} Kv_n(t, \phi).$$

Using $|v_n| \leq \frac{1}{c}$, this implies that

$$\left|\frac{d}{dt} Jv_n(t, \phi)\right| \leq \phi + e^{-\lambda t}\left(1 + \frac{\lambda}{c} + \frac{1}{c}\right) + \left|e^{-\lambda t}\frac{d}{dt} Kv_n(t, \phi)\right|. \tag{4.21}$$

Taking $a = \left(1 + \frac{\lambda}{c}\right) e^{-\lambda t} + \frac{1}{c}$, it therefore suffices to bound the last term on the right hand side above. So,

$$\begin{aligned} \left|e^{-\lambda t}\frac{d}{dt} Kv_n(t, \phi)\right| &= \left|e^{-\lambda t}\frac{d}{dt}\int_{-\infty}^{\infty} v_n(j(t, \phi, z))\frac{e^{-z^2/2}}{\sqrt{2\pi}} dz\right| \\ &= \left|e^{-\lambda t}\int_{-\infty}^{\infty} v_n'(j(t, \phi, z))\frac{dj}{dt}\frac{e^{-z^2/2}}{\sqrt{2\pi}} dz\right| \\ &\leq e^{-\lambda t}||v_n||_{Lip}\int_{-\infty}^{\infty}\left|\frac{dj}{dt}\frac{e^{-z^2/2}}{\sqrt{2\pi}}\right| dz, \end{aligned} \tag{4.22}$$

with the exchange of derivatives and integrals in the second equality justified by the fact that v_n' is bounded, established in Lemma IV.33. We now examine more closely the integrand in the last line above. We calculate that

$$\frac{dj}{dt} = \left(\frac{1}{2}\alpha z t^{-1/2} + (\lambda - \alpha^2/2)\right) \exp\left\{\alpha z\sqrt{t} + (\lambda-\alpha^2/2)t\right\}\phi \tag{4.23}$$

$$+\lambda \exp\left\{\alpha z\sqrt{t} + (\lambda-\alpha^2/2)t\right\} \tag{4.24}$$

$$+\int_0^t \lambda\left(\frac{-1}{2}\alpha z u t^{-3/2} + \frac{1}{2}\alpha^2 u^2 t^{-2}\right)\exp\left\{\left(\lambda + \frac{\alpha z}{\sqrt{t}}\right)u - \frac{\alpha^2 u^2}{2t}\right\} du, \tag{4.25}$$

and here the first term above came from differentiating the first term of j, and the second and third terms came from differentiating the second term of j. We label these terms in (4.23) $(A_1), (A_2), (A_3)$. Then

$$\begin{aligned}
\int_{-\infty}^{\infty} |(A_1)|\, e^{-z^2/2}/\sqrt{2\pi} &= e^{\lambda t}\phi \int_{-\infty}^{\infty} \left|\frac{1}{2}\alpha z t^{-1/2} + (\lambda - \alpha^2/2)\right| e^{-(z-\alpha\sqrt{t})^2/2}/\sqrt{2\pi}dz \\
&\le e^{\lambda t}\phi \int_{-\infty}^{\infty} \left(\frac{1}{2}\alpha t^{-1/2}\left|z - \alpha\sqrt{t}\right| + \lambda\right) e^{-(z-\alpha\sqrt{t})^2/2}/\sqrt{2\pi}dz \\
&= e^{\lambda t}\phi\left(\frac{1}{2}\alpha t^{-1/2}C_1 + \lambda\right),
\end{aligned}$$

with C_1 a universal constant equal arising from the expectation of the absolute value of a standard normal r.v. The second term can be treated similarly, yielding

$$\int_{-\infty}^{\infty} |(A_2)| e^{-z^2/2}/\sqrt{2\pi}dz \le \lambda e^{\lambda t}.$$

For the third term, we have, using Fubini's Theorem for the first inequality,

$$\begin{aligned}
\int_{-\infty}^{\infty} |(A_3)|\frac{e^{-z^2}}{\sqrt{2\pi}}\, dz &= \lambda\int_0^t \int_{-\infty}^{\infty} \frac{1}{2}\alpha u t^{-3/2} e^{\lambda u}\left|z - \frac{\alpha u}{\sqrt{t}}\right| \frac{e^{-(z-\alpha u/\sqrt{t})^2/2}}{\sqrt{2\pi}}\, dz du \\
&= \lambda\frac{1}{2}\alpha t^{-3/2}\int_0^t u e^{\lambda u}\int_{-\infty}^{\infty}\left|z - \frac{\alpha u}{\sqrt{t}}\right| \frac{e^{-(z-\alpha u/\sqrt{t})^2/2}}{\sqrt{2\pi}}\, dz du \\
&= \lambda\frac{1}{2}\alpha t^{-3/2} C_1 \int_0^t u e^{\lambda u} du \\
&\le \lambda\frac{1}{2}\alpha t^{-3/2} C_1 \frac{e^{\lambda t}(\lambda t - 1) + 1}{\lambda^2}.
\end{aligned}$$

This has absolute value less than or equal to $\lambda\frac{1}{2}\alpha t^{-3/2}C_1\frac{e^{\lambda t}t}{\lambda} = \lambda\frac{1}{2}\alpha t^{-1/2}C_1\frac{e^{\lambda t}}{\lambda}$.

Plugging these three estimates into (4.22), we obtain:

$$\begin{aligned}
&\left| e^{-\lambda t} \frac{d}{dt} K v_n(t, \phi) \right| \\
&\qquad \leq e^{-\lambda t} ||v_n||_{Lip} \int_{-\infty}^{\infty} \left| \frac{dj}{dt} \frac{e^{-z^2/2}}{\sqrt{2\pi}} \right| dz \\
&\qquad \leq e^{-\lambda t} ||v_n||_{Lip} \left(e^{\lambda t} \phi \left(\frac{1}{2} \alpha t^{-1/2} C_1 + \lambda \right) + \lambda e^{\lambda t} + \lambda \frac{1}{2} \alpha t^{-1/2} C_1 \frac{e^{\lambda t}}{\lambda} \right) \\
&\qquad = b ||v_n||_{Lip} t^{-1/2}, \qquad (4.26)
\end{aligned}$$

with $b = \left(\phi \left(\frac{1}{2} \alpha C_1 + \lambda \right) + \lambda + \frac{1}{2} \alpha C_1 \right)$. □

Using the Hölder continuity established above, we can do a trivial discretization to find ϵ-optimal times.

Corollary IV.35. *Fix $\overline{\phi} > 0$. For $0 \leq \phi \leq \overline{\phi}$, one may find $t^*(\phi, \epsilon)$ such that $Jv_n(t^*(\phi, \epsilon), \phi) < \min_{0 \leq t \leq T} Jv_n(t, \phi) + \epsilon$ by making $||v_n||_{Lip} \cdot O(\frac{1}{\epsilon^2})$ evaluations of $Jv_n(\cdot, \phi)$.*

Proof. Discretize $[0, T]$ into N equally spaced points $t_1, \ldots, t_N$, where $N = \lceil \frac{M}{\epsilon^2} \rceil$, and M is derived from the Hölder constant established in Lemma IV.34; for example, we may take $M \triangleq \overline{\phi} + a + b||v_n||_{Lip}$, with $a = 1 + \frac{\lambda}{c} + \frac{1}{c}$ and $b = \left(\overline{\phi} \left(\frac{1}{2} \alpha C_1 \right) + \lambda + \frac{1}{2\lambda} C_1 \right)$. Then, choose $t^*(\phi, \epsilon) \in \arg\min_{1 \leq i \leq N} Jv_n(t_i, \phi)$. By Lemma IV.34, $\left| \min_{1 \leq i \leq n} Jv_n(t_i, \phi) - \min_{0 \leq t \leq T} Jv_n(t, \phi) \right| \leq \epsilon$, so $t^*(\epsilon, \phi)$ must be ϵ-optimal. □

We will uniformly approximate v_n by a function $\widehat{v}_n$, but we do not know a priori what Lipschitz properties the approximation $\widehat{v}_{n-1}$ has, only that it is close to v_{n-1}. Therefore we need Corollary IV.37 and Lemma IV.36 to estimate $J_0 \widehat{v}_n(\phi)$.

Lemma IV.36. *Suppose that $||w_1 - w_2||_{L^\infty} < \epsilon$. Then $||J_0 w_1 - J_0 w_2||_{L^\infty} < \epsilon$ and $|Jw_1(t, \phi) - Jw_2(t, \phi)| < \epsilon$ for all $t, \phi \geq 0$.*

Proof. The proof follows by noticing that

$$Jw_1(t, \phi) - Jw_2(t, \phi) = e^{-\lambda t} \int_{-\infty}^{\infty} \left[w_1(j(t, z, \phi)) - w_2(j(t, z, \phi)) \right] e^{-z^2/2} / \sqrt{2\pi} dz,$$

which is bounded in size by ϵ for all t and ϕ. $\square$

Corollary IV.37. *Suppose that* $w : \mathbb{R}_+ \to \mathbb{R}$ *is a function such that* $||w - v_n||_{L^\infty} < \epsilon_1$*. Fix* $\overline{\phi} > 0$*. For* $0 \le \phi \le \overline{\phi}$*, one may find* $t^{**}(\phi, \epsilon)$ *such that*

$$\left| Jw(t^{**}(\phi, \epsilon), \phi) - \min_{0 \le t \le T} Jv_n(t, \phi) \right| < \epsilon + 3\epsilon_1$$

by making $||v_n||_{Lip} \cdot O(\frac{1}{\epsilon^2})$ *evaluations of* $Jw(\cdot, \phi)$*.*

Proof. Perform the same discretization as in Corollary IV.37, and let $t^{**}(\phi, \epsilon) \in \underset{1 \le i \le N}{\arg\min}\ Jw(t_i, \phi)$. Since $||w - v_n||_{L^\infty} < \epsilon_1$, Lemma IV.36 implies that $|Jw(t, \phi) - Jv_n(t, \phi)| < \epsilon_1$ for any $t, \phi \ge 0$. then Lemma IV.34 implies that

$$\left| Jv_n(t^{**}(\phi, \epsilon)) - \inf_{0 \le t \le T} Jv_n(t, \phi) \right| < \epsilon + 2\epsilon_1.$$

Using $||w - v_n||_{L^\infty} < \epsilon_1$ again,

$$|Jv_n(t^{**}(\phi, \epsilon)) - Jw(t^{**}(\phi, \epsilon))| < \epsilon_1.$$

Therefore,

$$\left| Jw(t^{**}(\phi, \epsilon)) - \inf_{0 \le t \le T} Jv_n(t, \phi) \right| < \epsilon + 3\epsilon_1.$$

$\square$

IV.5.3 Approximating $J_0 v_n(\phi)$ over all ϕ, for Fixed n

Lemma IV.38. *Fix* $n \ge 0$*, and suppose that* $w : \mathbb{R}_+ \to \mathbb{R}$ *satisfies* $||w - v_n||_{L^\infty} < \epsilon_1$*. Then using* $||v_n||_{Lip}||v_{n+1}||_{Lip} O\left(\frac{1}{\epsilon^3}\right)$ *evaluations of* $Jw(t, \phi)$*, we can construct a function* $\widehat{J}_0 w$ *such that* $||\widehat{J}_0 w(\phi) - J_0 v_n(\phi)||_{L^\infty} < \epsilon + 3\epsilon_1$*.*

Proof. Following Section 4.4 of [53], the function $v_c(\phi)$ can be explicitly computed, and in fact we can construct $\overline{\phi}$ such that $v_c(\overline{\phi}) = 0$. By Proposition IV.8, $v_n(\phi) \ge v_c(\phi)$, and v_n is increasing and nonnegative, so it follows that $v_n(\phi) = 0$ for $\phi \ge \overline{\phi}$

for all $n \geq 0$. Therefore, we set $\widehat{J_0}w(\phi) = 0$ for $\phi \geq \overline{\phi}$. So from now on, we assume that $0 \leq \phi \leq \overline{\phi}$. Discretize $[0, \overline{\phi}]$ into $R = \left\lceil \frac{||v_{n+1}||_{Lip}}{\epsilon} \right\rceil$ points $\phi_1, \ldots, \phi_R$, with $\phi_1 = 0$ and $\phi_R = \overline{\phi}$. Using Corollary IV.37, for each i, we may, given w, in $||v_n||_{Lip} O\left(\frac{1}{\epsilon^2}\right)$ function evaluations calculate $\widehat{J_0}w(\phi_i)$ such that $|\widehat{J_0}w(\phi_i) - J_0 v_n(\phi_i)| < \epsilon + 3\epsilon_1$ for $1 \leq i \leq R$. For $\phi_i \leq \phi < \phi_{i+1}$, $1 \leq i \leq R-1$, set $\widehat{J_0}w(\phi) = \widehat{J_0}w(\phi_i)$. We have, for $\phi_i \leq \phi < \phi_{i+1}$,

$$
\begin{aligned}
|\widehat{J_0}w(\phi) - J_0 v_n(\phi)| &= |\widehat{J_0}w(\phi_i) - J_0 v_n(\phi)| \\
&\leq |\widehat{J_0}w(\phi_i) - J_0 v_n(\phi_i)| + |J_0 v_n(\phi_i) - J_0 v_n(\phi)| \\
&\leq (\epsilon + 3\epsilon_1) + \epsilon,
\end{aligned}
$$

where the second ϵ term above is derived from the Lipschitzness of $J_0 v_n = v_{n+1}$, established in Lemma IV.33. Since each point i requires $||v_n||_{Lip} O\left(\frac{1}{\epsilon^2}\right)$ function evaluations, computing the approximations for all $R \approx \frac{||v_{n+1}||_{Lip}}{\epsilon}$ points requires $||v_n||_{Lip} ||v_{n+1}||_{Lip} O\left(\frac{1}{\epsilon^3}\right)$ function evaluations. □

IV.5.4 Approximating $v_n(\phi)$, for all $0 \leq n \leq N$

Proof of Proposition IV.29. The function $v_0(\phi)$ may be computed analytically. According to Lemma IV.38, we may compute a function $\widehat{v_1}(\phi)$ such that

$$
||\widehat{v_1}(\phi) - v_1(\phi)||_{L^\infty} < \frac{\epsilon}{N}
$$

in $||v_0||_{Lip} ||v_1||_{Lip} O\left(\frac{N^3}{\epsilon^3}\right)$ function evaluations. Applying Lemma IV.38 again, we construct $\widehat{v_2}(\phi)$ satisfying

$$
\left|\left| \widehat{v_2}(\phi) - \underbrace{J_0 v_1(\phi)}_{v_2(\phi)} \right|\right|_{L^\infty} < \frac{\epsilon}{N} + \frac{\epsilon}{N}
$$

in $||v_1||_{Lip} ||v_2||_{Lip} O\left(\frac{N^3}{\epsilon^3}\right)$ function evaluations. Arguing inductively in this way, we see that we may compute $\widehat{v_N}(\phi)$ satisfying

$$
||\widehat{v_N}(\phi) - v_N(\phi)||_{L^\infty} < \frac{N\epsilon}{N} = \epsilon
$$

in

$$O\left(\frac{N^3}{\epsilon^3}\right)\sum_{i=0}^{N-1}||v_i||_{Lip}||v_{i+1}||_{Lip}$$

function evaluations. By Lemma IV.33, $||v_i||_{Lip} \leq iT$. Therefore

$$\begin{aligned}
O\left(\frac{N^3}{\epsilon^3}\right)\sum_{i=0}^{N-1}||v_i||_{Lip}||v_{i+1}||_{Lip} &= O\left(\frac{N^3}{\epsilon^3}\right)\sum_{i=0}^{N-1} i^2T \\
&\leq O\left(\frac{N^3}{\epsilon^3}\right)N^3T \\
&= O\left(\frac{N^6}{\epsilon^3}\right).
\end{aligned}$$

$\square$

IV.6 Numerical Results for the Lump Sum n-Observation Problem

IV.6.1 Comparison to the Continuous Value Function

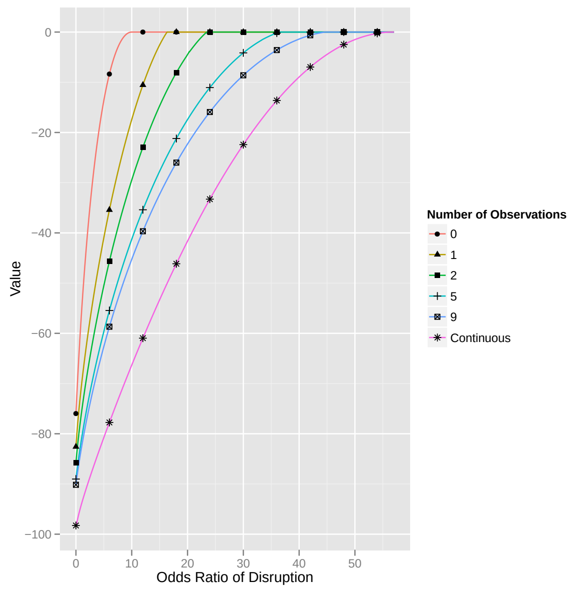


Figure 4.3: Value function $v_n(\phi)$ for $0 \leq n \leq 9$, and continuous observation value function $v_c(\phi)$, $\lambda = .1$, $c = .01$, $\alpha = 1$.

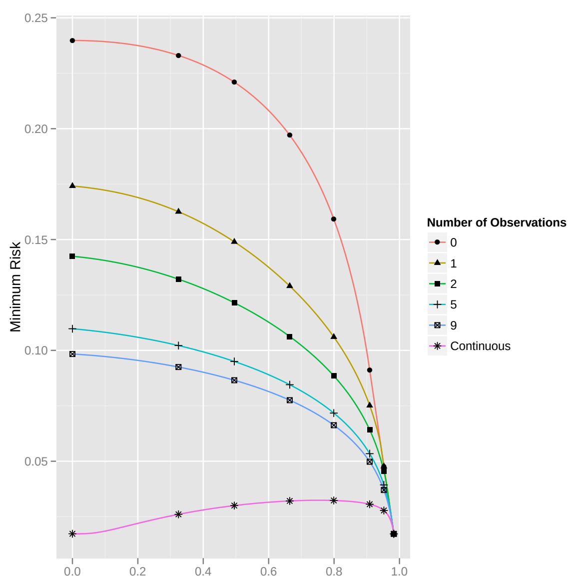


Figure 4.4: Bayesian Risk Associated with Total Number of Observations, $\lambda = .1$, $c = .01$, $\alpha = 1$.

Table 4.1: Effect of Observation Size on Value Functions at $\phi = 0$

Observations (n)	$v_n(0)$	$v_c(0)$	$v_n(0) - v_c(0)$	$\frac{\log(v_n(0)-v_c(0))}{\log(n+1)} - \log(v_0(0) - v_c(0)), n \geq 1$
0	-76.021	-98.237	22.216	
1	-82.586	-98.237	15.651	-.505
2	-85.755	-98.237	12.482	-.525
3	-87.410	-98.237	10.827	-.518
4	-88.392	-98.237	9.845	-.506
5	-89.024	-98.237	9.213	-.491
6	-89.455	-98.237	8.782	-.477
7	-89.762	-98.237	8.475	-.463
8	-89.990	-98.237	8.247	-.451
9	-90.163	-98.237	8.074	-.440
10	-90.299	-98.237	7.938	-.429

As expected, the value functions $v_n(\phi)$ are all concave and increasing, and between $\frac{-1}{c} = 100$ and 0. Furthermore, as n increases, the value functions decrease. From Figure 4.3, it is not immediately obvious whether $\lim_{n\to\infty} v_n = v_c$, although the results of Section IV.3 prove that this is the case. In any case, the convergence rate is quite slow, as demonstrated by Table 4.1.

IV.6.2 Comparison to Dayanik's Discrete Observation Model

In this subsection, we compare the value functions of the lump sum n-observation problem with those found in Dayanik's model of discrete observation, [15]. More precisely, we consider models of one or five total observation rights, and specify fixed time intervals at which observations will be made. In Figures 4.5 and 4.6, it is not surprising that the value function from our n observation problem is smallest, but the efficiency gap can be quite large, especially for higher values of ϕ when it can be crucial to make an observation quickly. Furthermore, we can see that the value functions associated to fixed observation schedules have widely varying performance on different levels of ϕ, and one which performs well for one value of ϕ may do quite poorly at another. Therefore, it is hard to achieve good performance using fixed observation strategies. This should not be surprising: our value function is the concave hull of the value functions corresponding to deterministic observation schedules. The difference is magnified with more observations, as flexibility becomes more important.

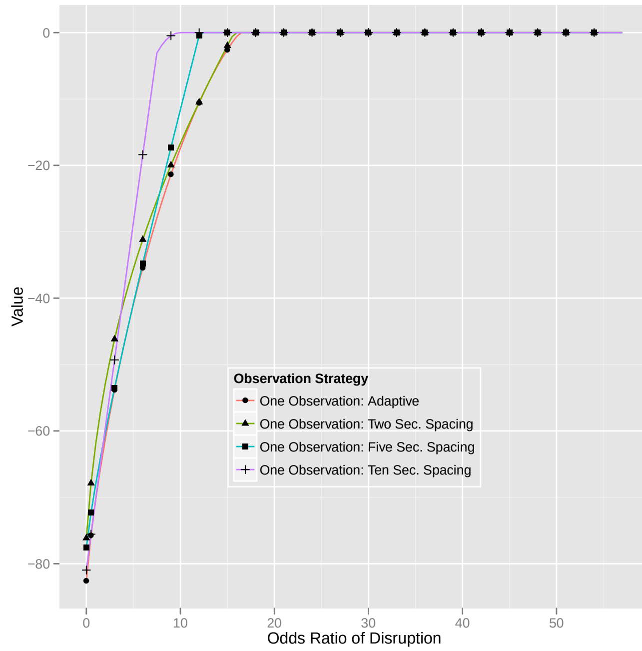


Figure 4.5: Value Functions for One Observation: At Two, Five, and Ten Seconds, and Chosen Adaptively, $\lambda = .1$, $c = .01$, $\alpha = 1$.

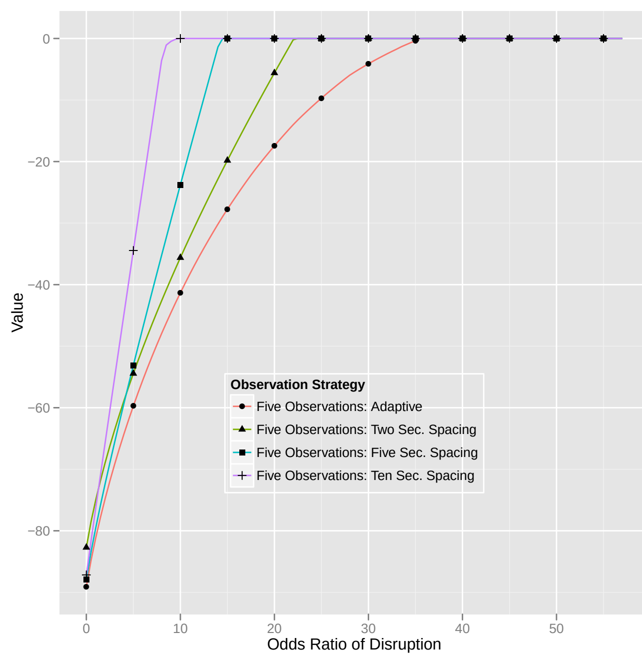


Figure 4.6: Value Functions for Five Observations: At Two, Five, and Ten Seconds Intervals, and Chosen Adaptively, $\lambda = .1$, $c = .01$, $\alpha = 1$.

IV.6.3 Depiction of Observation Boundaries

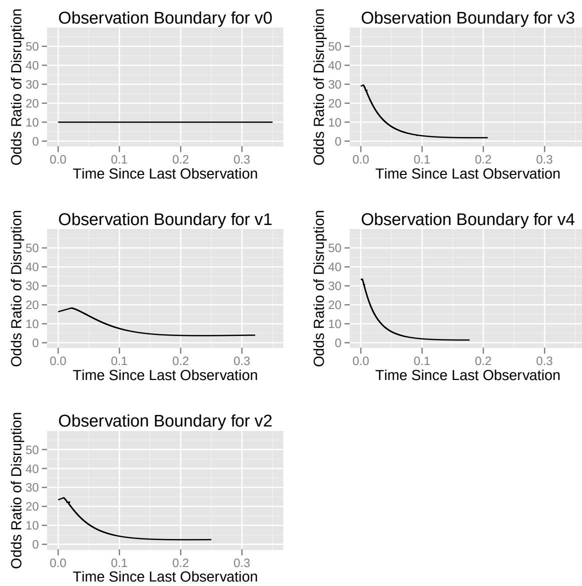


Figure 4.7: Observation Boundaries, $\lambda = .1$, $c = .01$, $\alpha = 1$

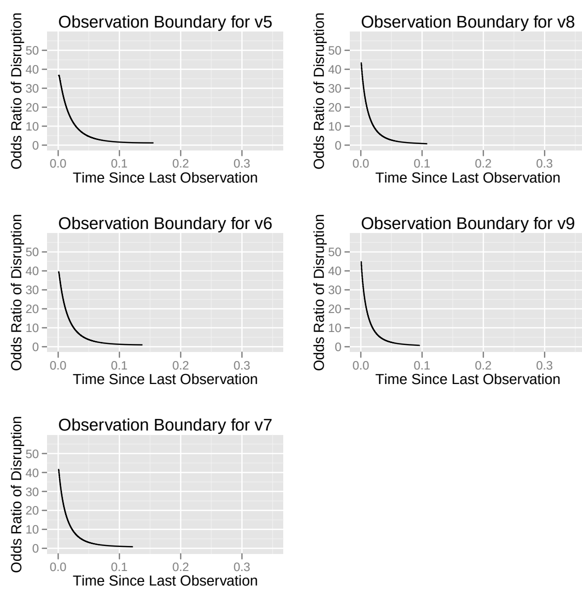


Figure 4.8: Observation Boundaries, $\lambda = .1$, $c = .01$, $\alpha = 1$.

The observation boundary for v_0 is of course identically equal to a horizontal line at $10 = \frac{\lambda}{c}$. Without any observations, the posterior process is perpetually increasing, and so the observation boundary (which is really a stopping boundary here) should always stop when ϕ is equal to $\frac{\lambda}{c}$. Note furthermore that the boundary does not depend on the time since the last observation: since there will never be any more observations made, there is time homogeneity.

In the rest of the observation boundaries, we notice two general trends: first, the curves are decreasing in n for large values of time, and second, they are increasing in n for small values of time. The former phenomenon reflects the idea that if an agent has more observations, than he should be more willing to use them, which corresponds to the barrier being easier to get to, and hence lower. At small time values, however, the barriers are increasing. This reflects the fact that as one has more observations, one is less willing to "give up" and stop. For example, when one has only a single observation and the odds process is above 20, an optimally acting agent knows that he only hurts himself by waiting, and so will observe immediately at time zero (which is equivalent to stopping the game). With more observations, however, the agent is willing to wait a little bit and see how things will go, and for this reason, the curves increase at small times.

A natural question is whether, as the observations increase, do the curves tend to infinity for very small values of time? In fact, the observation boundaries are uniformly bounded for all n and t. One may deduce this fact by comparing the discrete observation value functions with the continuous observation value function. In Figure 4.3, one sees that the continuous value function is zero for $\phi \geq \overline{\phi}$ (here $\overline{\phi} \approx 55$, and $\overline{\phi}$ can be explicitly computed: it is the optimal threshold level for stopping in the continuous observation problem,). This implies that every discrete observation value function is also zero for $\phi \geq \overline{\phi}$. Therefore, one always wants to observe immediately at such ϕ values. It follows then that all observation boundaries from Figures 4.7 and 4.8 will be bounded from above by $\overline{\phi}$.

IV.7 Numerics for the Stochastic Arrival Rate Problem

IV.7.1 A Heuristic Algorithm for the Stochastic Arrival Rate n-Observation Problem

Here we outline a computational algorithm for solving the stochastic arrival rate problem. The infinite horizon problem is a limiting case of this one. As discussed before, the lump sum n-observation problem is essentially embedded into this problem, so it should be of no surprise that this problem must be solved first.

(1) Fix the total number of observations N. Discretize ϕ into $\phi_0 = 0, \phi_1, \ldots, \phi_J = \overline{\phi}$ as before, and set all value functions equal to zero for $\phi \geq \overline{\phi}$. Compute the (approximations to) value functions $\widehat{\boldsymbol{v}}^N_{N,j}(0, \phi), 0 \leq j \leq N$, as in the previous section, as well as the optimal times $\widehat{t}^{*,N}_{N,j}(\phi)$.

(2) We have computed $\widehat{\boldsymbol{v}}^N_{N,j}(0, \phi)$, the approximation to $\boldsymbol{v}^N_{N,j}(0, \phi)$. For $y, t \geq 0$ such that $(y + t, \varphi(t, \phi)) \in \Lambda^F$, define $\widehat{\boldsymbol{v}}^N_{N,j}(y + t, \varphi(t, \phi))$ by

$$\text{(4.27)} \quad e^{-\lambda(y+t)}\widehat{\boldsymbol{v}}^N_{N,j}(y+t, \varphi(t, \phi)) = \begin{cases} \widehat{\boldsymbol{v}}^N_{N,j}(0, \varphi(-y, \phi)) - \int_0^{y+t} e^{-\lambda r} \left(\varphi(r - y, \phi) - \frac{\lambda}{c}\right) dr & \text{if } y + t < \widehat{t}^{*,N}_{N,j}(\varphi(-y, \phi)) \\ e^{-\lambda(y+t)} \boldsymbol{K}\widehat{\boldsymbol{v}}^N_{N,j+1}(y + t, \varphi(-y, \phi)) & \text{if } y + t \geq \widehat{t}^{*,N}_{N,j}(\varphi(-y, \phi)) \end{cases}$$

(3) Fix ϕ_j. Discretize time into $t_0 = 0, t_1, \ldots, t_K = T$, for an appropriately chosen upper bound T, as in Lemma IV.32. Compute $\widehat{\boldsymbol{v}}^N_{N-1,N-1}(0, \phi_j)$ by minimizing $J^0\widehat{\boldsymbol{v}}^N_{N,N-1}(t_i, 0, \phi_j)$ over the t_i. Let $\widehat{t}^{*,N}_{N-1,N-1}(\phi_j)$ be the minimizing t_i.

(4) Interpolate to find a function $\widehat{\boldsymbol{v}}^N_{N-1,N-1}(0, \phi)$ which approximates $\boldsymbol{v}^N_{N-1,N-1}(0, \phi)$, and a stopping boundary $\widehat{t}^{*,N}_{N-1,N-1}(\phi)$.

(5) Fix ϕ_j. As in Step 3, compute $\widehat{\boldsymbol{v}}^N_{N-1,N-2}(0,\phi_j)$ by minimizing $J^+(\widehat{\boldsymbol{v}}^N_{N-1,N-1}, \widehat{\boldsymbol{v}}^N_{N,N-2})(t_i,0,\phi_j)$ over the t_i. Let $\widehat{t}^{*,N}_{N-1,N-2}(\phi_j)$ be the minimizing t_i.

(6) Interpolate to find a function $\widehat{\boldsymbol{v}}^N_{N-1,N-2}(0,\phi)$ which approximates $\boldsymbol{v}^N_{N-1,N-2}(0,\phi)$, and an observation boundary $\widehat{t}^{*,N}_{N-1,N-2}(\phi)$.

(7) Repeat Steps 5 and 6 to compute $\widehat{\boldsymbol{v}}^N_{N-1,j}(0,\phi)$ and $\widehat{t}^{*,N}_{N-1,j}(\phi)$ for $0 \le j \le N-2$.

(8) We now need to repeat the analog of Step 2. For $y, t \ge 0$ such that $(y+t, \varphi(t,\phi)) \in \Lambda^F$, inductively define $\widehat{\boldsymbol{v}}^N_{N-1,j}(y+t,\varphi(t,\phi)$, $0 \le j \le N-2$, by

$$e^{-\lambda(y+t)}e^{-\mu(y+t)}\widehat{\boldsymbol{v}}^N_{N-1,j}(y+t,\varphi(t,\phi)) =$$

$$\begin{cases} \widehat{\boldsymbol{v}}^N_{N-1,j}(0,\varphi(-y,\phi)) & \text{if } y+t < \widehat{t}^{*,N}_{N-1,j}(\varphi(-y,\phi)) \\ \quad -\int_0^\infty \mu e^{-\mu u}\Bigg(\int_0^{u\wedge(y+t)} e^{-\lambda r}\left(\varphi(r-y,\phi) - \frac{\lambda}{c}\right) dr & \\ \quad -1_{\{u \le y+t\}} e^{-\lambda u}\widehat{\boldsymbol{v}}^N_{N-1,j+1}(y+u,\varphi(u,\phi))\Bigg) du & \\ e^{-\lambda(y+t)}e^{-\mu(y+t)}\boldsymbol{K}\widehat{\boldsymbol{v}}^N_{N-1,j+1}(y+t,\varphi(-y,\phi)) & \text{if } y+t \ge \widehat{t}^{*,N}_{N-1,j}(\varphi(-y,\phi)) \end{cases}$$

(9) Repeat steps 3 through 8 for each $0 \le n \le N-2$, computing $\widehat{v}^N_{n,j}$, $0 \le j \le n$ and their associated optimal times $\widehat{t}^{*,n}_{n,j}(\phi)$.

IV.7.2 Discussion of the Heuristic

1. The formula in Step (2) comes from a dynamic programming principle. In a simplified version (with $y = 0$), the dynamic programming principle says that for $t \le \widehat{t}^{*,N}_{N,j}(\phi)$,

$$\boldsymbol{v}^N_{N,j}(0,\phi) = \int_0^t e^{-\lambda r}\left(\varphi(r,\phi) - \frac{\lambda}{c}\right) dr + e^{-\lambda t}\boldsymbol{v}^N_{N,j}(t,\varphi(t,\phi)) :$$

in other words, if it is not optimal to make an observation before time t, then by waiting until time t no utility is lost, and the only difference between the value functions at the two times is the exponential discounting and the running cost lost between them. On the other hand, for all $t > \widehat{t}^{*,N}_{N,j}(\phi)$, we assume that it is optimal to immediately observe, hence the term $e^{-\lambda t}\boldsymbol{K}\boldsymbol{v}^N_{N,j+1}(t,\phi)$ in (4.27), describing the expected value after an observation is made. This step is a heuristic because we have not shown that the optimal observation behavior has this simple strategy. It is in theory possible, although unlikely in practice, that, when starting at $t=0$ the optimal observation time is $\widehat{t}^{*,N}_{N,j}(\phi)$, but when starting at some $t_1 > \widehat{t}^{*,N}_{N,j}(\phi)$, the optimal observation time is not t_1, but some other $t_2 > t_1$. Numerical evidence suggests that this is not the case, but we do not have a proof of this fact.

2. In Step 3, The J_0^0 operator is applied in the case when the agent has no spare observation rights. If he must wait to receive an observation, then even in that first instant when he receives this right, a positive amount of time has passed since the last observation was made. Therefore, we need information about the value function, i.e. $\boldsymbol{v}^N_{N,N-1}(\cdot,\cdot)$, when its first argument is positive: this explains the necessity of Step 2. Similar considerations apply to the calculation in Step 5.

3. The derivation of Step 8 is similar to that of Step 2, except that whereas in Step 2, we took the dynamic programming principle with all observation rights received, here we use the dynamic programming principle when there are still observation rights receiving. For example, dynamic programming implies that for $t \le t^{*,N}_{N-1,j}(\phi)$,

$$\begin{aligned}\boldsymbol{v}^N_{N-1,j}(0,\phi) = \int_0^\infty \mu e^{-\mu}\Bigg(&\int_0^{u\wedge t} e^{-\lambda r}\left(\varphi(r,\phi)-\frac{\lambda}{c}\right)dr \\ &+ 1_{\{t<u\}}e^{-\lambda t}\boldsymbol{v}^N_{N-1,j}(t,\varphi(t,\phi)) + 1_{\{u\le t\}}e^{-\lambda u}\boldsymbol{v}^N_{N,j}(u,\varphi(u,\phi))\Bigg)du.\end{aligned}$$

From this equation, we can solve for $\boldsymbol{v}^N_{N-1,j}(t,\varphi(t,\phi))$ to obtain the formula used in Step 8.

4. In Step 3, the minimum of a function is calculated by an exhaustive search on grid points. Numerical evidence suggests that $J^0\widehat{\boldsymbol{v}}^N_{N,N-1}(t,0,\phi)$ is actually convex as a function of t, which would allow for much more efficient ways of finding its minimum. The same holds true for the minimization in Step 5. We currently do not have analytic proofs of these facts. We note that this reasoning can additionally be applied to the minimization of $Jv_n(\cdot,\phi)$ in the lump sum n-observation problem.

IV.7.3 Numerical Results for the Stochastic Arrival Rate n-Observation Problem

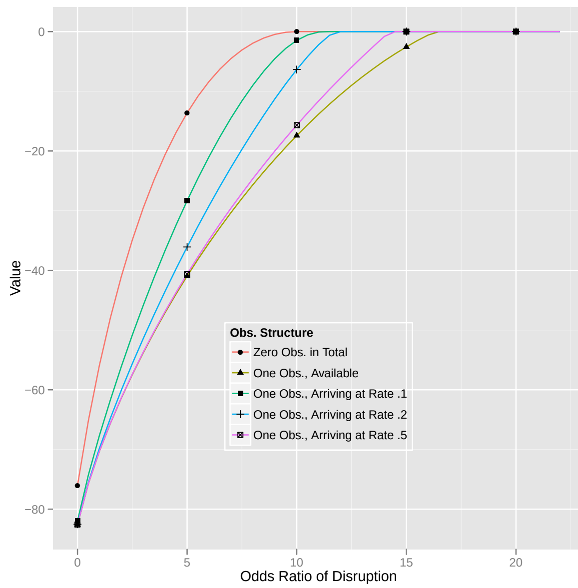


Figure 4.9: $\boldsymbol{v}^1_{1,1}(\phi), \boldsymbol{v}^1_{1,0}(\phi)$, and $\boldsymbol{v}^1_{0,0}(\phi)$ for different arrival rates μ; $\lambda = .1$, $c = .01$, $\alpha = 1$.

Concerning Figure 4.9, we make a few basic observations. First, it should be clear that $v^1_{1,1}$, corresponding to the case where the single observation right has been used, is the worst-performing value function, and $v^1_{1,0}$, corresponding to the case where the observation right has been received but not used, is the best-performing. We also expect $v^1_{0,0}$, as the arrival parameter μ varies, to interpolate between these two extreme curves, so that $v^1_{0,0}$ resembles $v^1_{1,0}$ when μ is large, and $v^1_{1,1}$ when μ is small. Furthermore, the gap between $v^1_{0,0}$ and $v^1_{1,0}$ is smallest when ϕ is small. This reflects the fact that when ϕ is small, an optimally acting agent, even if he had an observation right in hand, would wait to exercise it. As a consequence, having to wait to receive such a right is a less stringent constraint than when ϕ is large, in which case it is important to make an observation relatively quickly.

IV.8 Proofs from Section IV.2

Proof of Proposition IV.3. Let $\Psi = \{\psi_1, \ldots, \psi_n\} \in \mathfrak{O}^n$, and let $\tau \in \mathcal{T}^{\Psi}$. Supposing that $\tau \geq \psi_k$, we will show how to modify Ψ to yield an observation strategy $\widetilde{\Psi}$ under which τ is a $\mathbb{F}^{\widetilde{\Psi}}$-stopping time that does not stop between ψ_k and ψ_{k+1}. By inductively following the same procedure, this allows us to construct an observation strategy Ψ' such that $\tau \in \mathcal{T}_o^{\Psi'}$ and $E^{\phi}\left[\int_0^{\tau} e^{-\lambda t}\left(\Phi_t^{\Psi} - \frac{\lambda}{c}\right)dt\right] = E^{\phi}\left[\int_0^{\tau} e^{-\lambda t}\left(\Phi_t^{\Psi'} - \frac{\lambda}{c}\right)dt\right]$. This will establish the lemma.

The basic method is to just add in an observation whenever τ stops between observations; if a stop is made between observations, there are always "spare observations". Let $A = \{\psi_k < \tau < \psi_{k+1}\}$. According to Proposition IV.2, $A \in \mathcal{F}^{\Psi}_{\psi_k}$, or $A \in \sigma\left(\{X_{\psi_i}, \psi_i\}_{1\leq i\leq k}\right)$. Define $\widetilde{\psi}_1 = \psi_1, \ldots, \widetilde{\psi}_k = \psi_k$, $\widetilde{\psi}_{k+1} = 1_A\tau + 1_{A^c}\psi_{k+1}$, and for $k+2 \leq j \leq n$, $\widetilde{\psi}_j = 1_A\psi_{j-1} + 1_{A^c}\psi_j$. With $A \in \sigma\left(\{X_{\psi_i}, \psi_i\}_{1\leq i\leq k}\right) = \sigma\left(\left\{X_{\widetilde{\psi}_i}, \widetilde{\psi}_i\right\}_{1\leq i\leq k}\right)$, the following claim will imply that for each $k+1 \leq j \leq n$,

$$\widetilde{\psi}_j \in m\ \sigma(X_{\widetilde{\psi}_1}, \ldots, X_{\widetilde{\psi}_{j-1}}, \widetilde{\psi}_1, \ldots, \widetilde{\psi}_{j-1}).$$

This fact is obvious for $j = 1, \ldots, k$, as $\widetilde{\psi}_j = \psi_j$ for $j \leq k$. Set $\widetilde{\Psi} \triangleq \{\widetilde{\psi}_1, \ldots, \widetilde{\psi}_n\}$.

Claim IV.39. Let Let $k+1 \leq j \leq n$. Let $X \in \mathcal{F}^{\Psi}_{\psi_{j-1}}$ and let $Y \in \mathcal{F}^{\Psi}_{\psi_j}$. Then $[A \cap X] \cup [A^c \cap Y] \in \mathcal{F}^{\widetilde{\Psi}}_{\widetilde{\psi}_j}$.

Proof Of Claim. Write $1_X = x\left(\{X_{\psi_i}, \psi_i : 1 \leq i \leq j-1\}\right)$,
$1_Y = y\left(\{X_{\psi_i}, \psi_i : 1 \leq i \leq j\}\right)$, $1_A = a\left(\{X_{\psi_i}, \psi_i\}_{1\leq i\leq k}\right)$, where x, y, and a are all Borel functions with respective domains $\mathbb{R}^{2(j-1)}, \mathbb{R}^{2j}$, and $\mathbb{R}^{2k}$.

Then

$$
\begin{aligned}
&1_A 1_X + 1_{A^c} 1_Y \\
&\quad = 1_A x\left(\{X_{\psi_i}, \psi_i : 1 \le i \le j-1\}\right) \\
&\qquad + 1_{A^c} y\left(\{X_{\psi_i}, \psi_i : 1 \le i \le j\}\right) \\
&\quad = 1_A x\left(X_{\psi_1}, \psi_1, \{1_A X_{\psi_{i-1}} + 1_{A^c} X_{\psi_i}, 1_A \psi_{i-1} + 1_{A^c}\psi_i : 3 \le i \le j\}\right) \\
&\qquad + 1_{A^c} y\Big(X_{\psi_1}, \psi_1, 1_A X_\tau + 1_{A^c} X_{\psi_2}, 1_A \tau + 1_{A^c}\psi_2, \\
&\qquad\qquad \{1_A X_{\psi_{i-1}} + 1_{A^c} X_{\psi_i}, 1_A \psi_{i-1} + 1_{A^c}\psi_i : 3 \le i \le j\}\Big) \\
&\quad = a\left(\{X_{\psi_i}, \psi_i\}_{1 \le i \le k}\right) x\left(X_{\psi_1}, \psi_1, \{1_A X_{\psi_{i-1}} + 1_A X_{\psi_i}, 1_A \psi_{i-1} + 1_{A^c}\psi_i : 3 \le i \le j\}\right) \\
&\qquad + \left(1 - a\left(\{X_{\psi_i}, \psi_i\}_{1 \le i \le k}\right)\right) y\Big(X_{\psi_1}, \psi_1, 1_A X_\tau + 1_{A^c} X_{\psi_2}, 1_A \tau + 1_{A^c}\psi_2, \\
&\qquad\qquad \{1_A X_{\psi_{i-1}} + 1_{A^c} X_{\psi_i}, 1_A \psi_{i-1} + 1_{A^c}\psi_i : 3 \le i \le j\}\Big) \in m\mathcal{F}^{\widetilde{\Psi}}_{\widetilde{\psi}_j}.
\end{aligned}
$$

$\square$

Having concluded the proof of the claim, we have shown that the observation strategy $\widetilde{\Psi}$ is admissible, or $\widetilde{\Psi} \in \mathfrak{O}^n$. By construction, τ does not stop between $\widetilde{\psi}_k$ and $\widetilde{\psi}_{k+1}$.

We have to check that τ is a $\mathbb{F}^{\widetilde{\Psi}}$-stopping time. Let $t > 0$. Then $\{\tau \le t\} \in \mathcal{F}^{\widetilde{\Psi}}_t$ if and only if $\{\tau \le t\} \cap \{\widetilde{\psi}_j \le t\} \in \mathcal{F}^{\widetilde{\Psi}}_{\widetilde{\psi}_j}$ for each $1 \le j \le n$. This is clear for $1 \le j \le k$ as $\widetilde{\psi}_j = \psi_j$ for $1 \le j \le k$ and τ is a $\mathbb{F}^{\Psi}$-stopping time. For $j = k+1$, we work on A and A^c separately. We must show

$$\{\tau \le t\} \cap \{\widetilde{\psi}_{k+1} \le t\} \in \mathcal{F}^{\widetilde{\Psi}}_{\widetilde{\psi}_{k+1}}.$$

Invoking the claim, it is sufficient to show that

$$\{\tau \le t\} \cap \{\widetilde{\psi}_{k+1} \le t\} \cap A \in \mathcal{F}^{\Psi}_{\psi_k} \tag{4.28}$$

and

$$\{\tau \leq t\} \cap \{\widetilde{\psi}_{k+1} \leq t\} \cap A^c \in \mathcal{F}^{\Psi}_{\psi_{k+1}}. \tag{4.29}$$

To show (4.28), note that on the set A, $\widetilde{\psi}_{k+1} = \tau$, and so $\{\tau \leq t\} \cap \{\widetilde{\psi}_{k+1} \leq t\} \cap A = \{\tau \leq t\} \cap A$. Since τ is a $\mathbb{F}^{\Psi}$-stopping time, greater than or equal to ψ_k, it follows that $\{\tau \leq t\} = \{\tau \leq t\} \cap \{\psi_k \leq t\} \in \mathcal{F}^{\Psi}_{\psi_k}$. To show (4.29), note that on the set A^c, $\widetilde{\psi}_{k+1} = \psi_{k+1}$, so $\{\tau \leq t\} \cap \{\widetilde{\psi}_{k+1} \leq t\} \cap A^c = \{\tau \leq t\} \cap \{\psi_{k+1} \leq t\} \cap A^c \in \mathcal{F}^{\Psi}_{\psi_{k+1}}$, noting that τ is a $\mathbb{F}^{\Psi}$ stopping time and $A \in \mathcal{F}^{\Psi}_{\psi_k}$.

Now, fix $j \geq k+2$. We may write

$$\{\tau \leq t\} \cap \{\widetilde{\psi}_j \leq t\} = \Big[A \cap \{\tau \leq t\} \cap \{\psi_{j-1} \leq t\}\Big] \cup \Big[A^c \cap \{\tau \leq t\} \cap \{\psi_j \leq t\}\Big].$$

Since τ is a $\mathbb{F}^{\Psi}$-stopping time, we know that $\{\tau \leq t\} \cap \{\psi_{j-1} \leq t\} \in \mathcal{F}^{\Psi}_{\psi_{j-1}}$ and $\{\tau \leq t\} \cap \{\psi_j \leq t\} \in \mathcal{F}^{\Psi}_{\psi_j}$. Therefore, $\{\tau \leq t\} \cap \{\widetilde{\psi}_j \leq t\} \in \mathcal{F}^{\widetilde{\Psi}}_{\widetilde{\psi}_j}$, again by Claim IV.39.

Finally, note that $E^{\phi}\left[\int_0^{\tau} e^{-\lambda t}\left(\Phi^{\Psi}_t - \frac{\lambda}{c}\right) dt\right] = E^{\phi}\left[\int_0^{\tau} e^{-\lambda t}\left(\Phi^{\widetilde{\Psi}}_t - \frac{\lambda}{c}\right) dt\right]$ because $\Phi^{\Psi} = \Phi^{\widetilde{\Psi}}$ a.s. on the random time interval $[0, \tau)$.

$\square$

Proof of Proposition IV.4. Fix $n \geq 0$. We proceed by backwards induction, so consider the base case $k = n$, and take $\Psi^n = \{\psi^n_1, \ldots, \psi^n_n\} \in \mathfrak{O}^n$. We will prove equality here. We must show that

$$\begin{aligned} v_0\left(\Phi^{\Psi^n}_{\psi^n_n}\right) &= \operatorname*{ess\,inf}_{\Psi \in \mathfrak{O}^n(\Psi^n)} \operatorname*{ess\,inf}_{\tau \in \mathcal{T}^{\Psi}_o, \tau \geq \psi^n_n} E\left[\int_{\psi^n_n}^{\tau} e^{-\lambda(t-\psi^n_n)}\left(\Phi^{\Psi}_t - \frac{\lambda}{c}\right) dt \middle| \mathcal{F}^{\Psi^n}_{\psi^n_n}\right] \\ &= \operatorname*{ess\,inf}_{\tau \in \mathcal{T}^{\Psi^n}, \tau \geq \psi^n_n} E\left[\int_{\psi^n_n}^{\tau} e^{-\lambda(t-\psi^n_n)}\left(\Phi^{\Psi^n}_t - \frac{\lambda}{c}\right) dt \middle| \mathcal{F}^{\Psi^n}_{\psi^n_n}\right], \end{aligned}$$

the second inequality following from the fact that we can observe a total of n times, so that $\mathfrak{O}^n(\Psi^n)$ is a singleton, and equal to Ψ^n.

Recall

$$t_0^*(\phi) = \frac{1}{\lambda}\log\left(\frac{c+\lambda}{c(\phi+1)}\right) \vee 0. \tag{4.30}$$

A simple calculation confirms that $J_00(\phi) = J0(t^*(\phi), \phi)$.

According to Proposition IV.2, there is a one-to-one correspondence between $\{\tau \in \mathcal{T}^{\Psi^n} : \tau \geq \psi_n^n\}$ and the set $\{\psi_n^n + R_n : R_n \geq 0, R_n \in m\mathcal{F}^{\Psi^n}_{\psi_n^n}\}$. So, take $\tau^* = \psi_n^n + t_0^*(\Phi^{\Psi^n}_{\psi_n^n}) \in \{\tau \in \mathcal{T}^{\Psi^n} : \tau \geq \psi_n^n\}$. Thus,

$$\begin{aligned}
&\operatorname*{ess\,inf}_{\tau\in\mathcal{T}^{\Psi^n},\tau\geq\psi_n^n} E\left[\int_{\psi_n^n}^{\tau} e^{-\lambda(t-\psi_n^n)}\left(\Phi_t^{\Psi^n} - \frac{\lambda}{c}\right)dt \Big| \mathcal{F}^{\Psi^n}_{\psi_n^n}\right] \\
&\quad\leq E\left[\int_{\psi_n^n}^{\tau^*} e^{-\lambda(t-\psi_n^n)}\left(\Phi_t^{\Psi^n} - \frac{\lambda}{c}\right)dt \Big| \mathcal{F}^{\Psi^n}_{\psi_n^n}\right].
\end{aligned}$$

According to (4.1), the process $\Phi_t^{\Psi^n}$ is increasing for $t \geq \psi_n^n$, and if $t_0^*(\Phi_t^{\Psi^n}) > 0$, then $\psi_n^n + t_0^*(\Phi^{\Psi^n}_{\psi_n^n})$ is the unique time $t \geq \psi_n^n$ when $\Phi_t^{\Psi^n} - \frac{\lambda}{c}$ changes sign from negative to positive. If $t_0^*(\Phi_t^{\Psi^n}) = 0$, then $\Phi_t^{\Psi^n}$ is always greater than λ/c. Therefore,

$$\begin{aligned}
&\operatorname*{ess\,inf}_{\tau\in\mathcal{T}^{\Psi^n},\tau\geq\psi_n^n} E\left[\int_{\psi_n^n}^{\tau} e^{-\lambda(t-\psi_n^n)}\left(\Phi_t^{\Psi^n} - \frac{\lambda}{c}\right)dt \Big| \mathcal{F}^{\Psi^n}_{\psi_n^n}\right] \\
&\quad\geq E\left[\int_{\psi_n^n}^{\tau^*} e^{-\lambda(t-\psi_n^n)}\left(\Phi_t^{\Psi^n} - \frac{\lambda}{c}\right)dt \Big| \mathcal{F}^{\Psi^n}_{\psi_n^n}\right],
\end{aligned}$$

the inequality holding at the pointwise level. It therefore follows that $\gamma_n^n(\Psi^n) = v_0(\Phi^{\Psi^n}_{\psi_n^n})$.

Now, for the inductive step, suppose that $\gamma_{k+1}^n(\Psi^{k+1}) \geq v_{n-k-1}\left(\Phi^{\Psi^{k+1}}_{\psi^{k+1}_{k+1}}\right)$ for all $\Psi^{k+1} = \{\psi_i^{k+1}\}_{1\leq i\leq k+1} \in \mathfrak{O}_{k+1}$. Let $\Psi^k = \{\psi_i^k\}_{1\leq i\leq k} \in \mathfrak{O}^k$. We wish to show that

$$\gamma_k^n(\Psi^k) \geq v_{n-k}\left(\Psi^k_{\psi_k^k}\right).$$

Let $\widetilde{\Psi} \in \mathfrak{O}^n(\Psi^k)$, and write $\widetilde{\Psi} = \{\psi_1^k, \ldots, \psi_k^k, \widetilde{\psi}_{k+1}, \ldots, \widetilde{\psi}_n\}$. Without loss of generality, we assume that $\widetilde{\psi}_{k+1} > \psi_k^k$.

Then

$$
\begin{aligned}
&\operatorname*{ess\,inf}_{\tau\in\mathcal{T}_o^{\widetilde{\Psi}},\tau\geq\psi_k^k} E\left[\int_{\psi_k^k}^{\tau} e^{-\lambda(t-\psi_k^k)}\left(\Phi_t^{\widetilde{\Psi}}-\frac{\lambda}{c}\right)dt\big|\mathcal{F}^{\Psi^k}_{\psi_k^k}\right]\\
&=\min\left\{0,\operatorname*{ess\,inf}_{\tau\in\mathcal{T}_o^{\widetilde{\Psi}},\tau\geq\widetilde{\psi}_{k+1}} E\left[\int_{\psi_k^k}^{\widetilde{\psi}_{k+1}} e^{-\lambda(t-\psi_k^k)}\left(\Phi_t^{\widetilde{\Psi}}-\frac{\lambda}{c}\right)dt+\int_{\widetilde{\psi}_{k+1}}^{\tau} e^{-\lambda(t-\psi_k^k)}\left(\Phi_t^{\widetilde{\Psi}}-\frac{\lambda}{c}\right)dt\big|\mathcal{F}^{\Psi^k}_{\psi_k^k}\right]\right\}\\
&=\min\left\{0,E\left[\int_{\psi_k^k}^{\widetilde{\psi}_{k+1}} e^{-\lambda(t-\psi_k^k)}\left(\Phi_t^{\widetilde{\Psi}}-\frac{\lambda}{c}\right)dt\big|\mathcal{F}^{\Psi^k}_{\psi_k^k}\right]\right.\\
&\qquad\left.+\operatorname*{ess\,inf}_{\tau\in\mathcal{T}_o^{\widetilde{\Psi}},\tau\geq\widetilde{\psi}_{k+1}} E\left[\int_{\widetilde{\psi}_{k+1}}^{\tau} e^{-\lambda(t-\psi_k^k)}\left(\Phi_t^{\widetilde{\Psi}}-\frac{\lambda}{c}\right)dt\big|\mathcal{F}^{\Psi^k}_{\psi_k^k}\right]\right\}\\
&\geq\min\left\{0,\int_0^{\widetilde{\psi}_{k+1}-\psi_k^k} e^{-\lambda t}\left(\varphi\left(t,\Phi^{\Psi^k}_{\psi_k^k}\right)-\frac{\lambda}{c}\right)dt+e^{-(\widetilde{\psi}_{k+1}-\psi_k^k)}E\left[v_{n-k-1}\left(\Phi^{\widetilde{\Psi}}_{\widetilde{\psi}_{k+1}}\right)|\mathcal{F}^{\Psi^k}_{\psi_k^k}\right]\right\}
\end{aligned}
$$

by the inductive hypothesis, where we have used the fact that $\widetilde{\psi}_{k+1}$ is $\mathcal{F}^{\Psi^k}_{\psi_k^k}$-measurable as well as the deterministic dynamics of $\Phi^{\widetilde{\Psi}}$ in between jumps. Next,

$$
\begin{aligned}
&\min\left\{0,\int_0^{\widetilde{\psi}_{k+1}-\psi_k^k} e^{-\lambda t}\left(\varphi\left(t,\Phi^{\Psi^k}_{\psi_k^k}\right)-\frac{\lambda}{c}\right)dt+e^{-(\widetilde{\psi}_{k+1}-\psi_k^k)}E\left[v_{n-k-1}\left(\widetilde{\Psi}_{\widetilde{\psi}_{k+1}}\right)|\mathcal{F}^{\Psi^k}_{\psi_k^k}\right]\right\}\\
&\quad=\min\left\{0,\int_0^{\widetilde{\psi}_{k+1}-\psi_k^k} e^{-\lambda t}\left(\varphi\left(t,\Phi^{\Psi^k}_{\psi_k^k}\right)-\frac{\lambda}{c}\right)dt\right.\\
&\qquad\left.+e^{-(\widetilde{\psi}_{k+1}-\psi_k^k)}E\left[v_{n-k-1}\left(j\left(\widetilde{\psi}_{k+1}-\psi_k^k,\Phi^{\Psi^k}_{\psi_k^k},\frac{X_{\widetilde{\psi}_{k+1}}-X_{\psi_k^k}}{\sqrt{\widetilde{\psi}_{k+1}-\psi_k^k}}\right)\right)|\mathcal{F}^{\Psi^k}_{\psi_k^k}\right]\right\}\\
&\quad=\min\left\{0,\int_0^{\widetilde{\psi}_{k+1}-\psi_k^k} e^{-\lambda t}\left(\varphi\left(t,\Phi^{\Psi^k}_{\psi_k^k}\right)-\frac{\lambda}{c}\right)dt+e^{-(\widetilde{\psi}_{k+1}-\psi_k^k)}Kv_{n-k-1}\left(\widetilde{\psi}_{k+1}-\psi_k^k,\Phi^{\Psi^k}_{\psi_k^k}\right)\right\}\\
&\quad=\min\left\{0,Jv_{n-k-1}\left(\widetilde{\psi}_{k+1}-\psi_k^k,\Phi^{\Psi^k}_{\psi_k^k}\right)\right\},
\end{aligned}
$$

noting that $\widetilde{\psi}_{k+1}-\psi_k^k>0$ and that $\frac{X_{\widetilde{\psi}_{k+1}}-X_{\psi_k^k}}{\sqrt{\widetilde{\psi}_{k+1}-\psi_k^k}}$ has standard normal distribution, independent of $\mathcal{F}^{\Psi^k}_{\psi_k^k}$. Finally, $\min\left\{0,Jv_{n-k-1}\left(\widetilde{\psi}_{k+1}-\psi_k^k,\Phi^{\Psi^k}_{\psi_k^k}\right)\right\}\geq J_0v_{n-k-1}\left(\Phi^{\Psi^k}_{\psi_k^k}\right)=v_{n-k}\left(\Phi^{\Psi^k}_{\psi_k^k}\right)$.

This establishes that

$$
\operatorname*{ess\,inf}_{\tau\in\mathcal{T}_o^{\widetilde{\Psi}},\tau\geq\psi_k^k} E\left[\int_{\psi_k^k}^{\tau} e^{-\lambda(t-\psi_k^k)}\left(\Phi_t^{\widetilde{\Psi}}-\frac{\lambda}{c}\right)dt\big|\mathcal{F}^{\Psi^k}_{\psi_k^k}\right]\geq v_{n-k}\left(\Phi^{\Psi^k}_{\psi_k^k}\right).
$$

Taking the infimum over all $\widetilde{\Psi}\in\mathfrak{O}^n(\Psi^k)$, we obtain that $\gamma_k^n(\Psi^k)\geq v_{n-k}\left(\Psi^k_{\psi_k^k}\right)$. $\square$

Proof. The " $\geq$ " inequality has been established in Proposition IV.4, and it suffices to show that $\gamma_k^n(\Psi^k)\leq v_{n-k}(\Phi^{\Psi^k}_{\psi_k})$. As in Proposition IV.4, we proceed by reverse induction, and note that the base case has already been established in Proposition

IV.4. Therefore, we assume that $\gamma^n_{k+1}(\Psi^{k+1}) = v_{n-k-1}\left(\Phi^{\Psi^{k+1}}_{\psi^{k+1}_{k+1}}\right)$ for all $\Psi^{k+1} \in \mathfrak{O}^{k+1}$. Let $\Psi^k \in \mathfrak{O}^k$. We wish to show that $\gamma^n_k(\Psi^k) \leq v_{n-k}\left(\Phi^{\Psi^k}_{\psi^k_k}\right)$. Recall the functions

$$h^0_{n-k}(\phi) = \min\{s \geq 0 : Jv_{n-k}(s,\phi) \leq J_0 v_{n-k}(\phi)\},$$

and the stopping times $\widehat{\psi}^{k+1} \triangleq \psi^k_k + h^0_{n-k}\left(\Phi^{\Psi^k}_{\psi^k_k}\right)$. Note that we can use a minimum above instead of an infimum because $Jv_{n-k}(s,\phi)$ is lower semi-continuous in s, as v_{n-k} is nonpositive. Write $\widehat{\Psi}^{k+1} = \left\{\psi^k_1, \ldots, \psi^k_k, \widehat{\psi}_{k+1}\right\}$.

Then

$$\begin{aligned}
\gamma^n_k(\Psi^k) &= \operatorname*{ess\,inf}_{\Psi \in \mathfrak{O}^n(\Psi^k)} \operatorname*{ess\,inf}_{\tau \in \mathcal{T}^\Psi_o, \tau \geq \psi^k_k} E\left[\int_{\psi^k_k}^{\tau} e^{-\lambda(t-\psi^k_k)}\left(\Phi^\Psi_t - \frac{\lambda}{c}\right) dt \middle| \mathcal{F}^{\Psi^k}_{\psi^k_k}\right] \\
&\leq \operatorname*{ess\,inf}_{\Psi \in \mathfrak{O}^n(\widehat{\Psi}^{k+1})} \operatorname*{ess\,inf}_{\tau \in \mathcal{T}^\Psi_o, \tau \geq \psi^k_k} E\left[\int_{\psi^k_k}^{\tau} e^{-\lambda(t-\psi^k_k)}\left(\Phi^\Psi_t - \frac{\lambda}{c}\right) dt \middle| \mathcal{F}^{\Psi^k}_{\psi^k_k}\right] \\
&= \min\Bigg\{0, E\Bigg[\int_{\psi^k_k}^{\widehat{\psi}_{k+1}} e^{-\lambda(t-\psi^k_k)}\left(\Phi^\Psi_t - \frac{\lambda}{c}\right) dt \\
&\qquad + \operatorname*{ess\,inf}_{\Psi \in \mathfrak{O}^n(\widehat{\Psi}^{k+1})} \operatorname*{ess\,inf}_{\tau \in \mathcal{T}^\Psi_o, \tau \geq \widehat{\psi}_{k+1}} \int_{\widehat{\psi}_{k+1}}^{\tau} e^{-\lambda(t-\psi^k_k)}\left(\Phi^\Psi_t - \frac{\lambda}{c}\right) dt \Big| \mathcal{F}^{\Psi^k}_{\psi^k_k}\Bigg]\Bigg\} \\
&= \min\left\{0, E\left[\int_{\psi^k_k}^{\widehat{\psi}_{k+1}} e^{-\lambda(t-\psi^k_k)}\left(\Phi^{\widehat{\Psi}^{k+1}}_t - \frac{\lambda}{c}\right) dt + e^{-(\widehat{\psi}_{k+1}-\psi^k_k)}\gamma^n_{k+1}\left(\Phi^{\widehat{\Psi}^{k+1}}_{\widehat{\psi}_{k+1}}\right) \Big| \mathcal{F}^{\Psi^k}_{\psi^k_k}\right]\right\} \\
&= \min\left\{0, E\left[\int_0^{\widehat{\psi}_{k+1}-\psi^k_k} e^{-\lambda t}\left(\varphi(t, \Phi^{\Psi^k}_{\psi^k_k}) - \frac{\lambda}{c}\right) dt + e^{-(\widehat{\psi}_{k+1}-\psi^k_k)} v_{n-k-1}\left(\Phi^{\widehat{\Psi}^{k+1}}_{\widehat{\psi}_{k+1}}\right) \Big| \mathcal{F}^{\Psi^k}_{\psi^k_k}\right]\right\},
\end{aligned}$$

with the last equality following from the inductive hypothesis and the deterministic evolution of Φ^{Ψ^k} in between jumps. It now follows, using an argument similar to that at the end of the proof of Proposition IV.4, that

$$\begin{aligned}
&\min\left\{0, E\left[\int_0^{\widehat{\psi}^{k+1}-\psi^k_k} e^{-\lambda t}\left(\varphi(t, \Phi^{\Psi^k}_{\psi^k_k}) - \frac{\lambda}{c}\right) dt + e^{-(\widehat{\psi}^{k+1}-\psi^k_k)} v_{n-k-1}\left(\Phi^{\widehat{\Psi}^{k+1}}_{\widehat{\psi}^{k+1}}\right) \Big| \mathcal{F}^{\Psi^k}_{\psi^k_k}\right]\right\} \\
&= J_0 v_{n-k-1}\left(\Phi^{\Psi^k}_{\psi^k_k}\right) \\
&= v_{n-k}\left(\Phi^{\Psi^k}_{\psi^k_k}\right),
\end{aligned}$$

with the first equality above by definition of $\widehat{\psi}_{k+1}$ and h^0_{n-k}. It follows then that

$$\gamma^n_k(\Psi^k) \leq v_{n-k}(\Phi^{\Psi^k}_{\psi_k}),$$

and the equality now follows. □

Proof of Lemma IV.6. We claim that $h^{\epsilon}_{n-k}(\phi)$ is lower semi-continuous. This would imply that $h^{\epsilon}_{n-k}(\phi)$ is Borel-measurable, which will in turn imply that $\widehat{\psi}^{k+1}$ is a stopping time.

Let $\phi_i \to \phi_\infty$ in $\mathbb{R}_+$, and let $s_i = h^{\epsilon}_{n-k}(\phi_i)$. Note that $J_0 v_{n-k}(\cdot)$ is bounded, and that

$$\lim_{s\to\infty} \inf_{\phi\geq 0} J v_{n-k}(s,\phi) = +\infty.$$

Therefore, we may assume that the sequence $\{s_i\}_{i\geq 0}$ is bounded. It follows then that $\liminf_{i\to\infty}\ s_i = s_\infty < \infty$. It is straightforward to see that $Jv_{n-k}(\cdot,\cdot)$ and $J_0 v_{n-k}(\cdot)$ are continuous in their arguments. Therefore,

$$\begin{aligned} Jv_{n-k}(s_\infty,\phi_\infty) &\leq \liminf_{i\to\infty}\ Jv_{n-k}(s_i,\phi_i) \\ &\leq \lim_{i\to\infty}\ J_0 v_{n-k}(\phi_i) + \epsilon \\ &= J_0 v_{n-k}(\phi_\infty) + \epsilon. \end{aligned}$$

Thus,

$$\begin{aligned} h^{\epsilon}_{n-k}(\phi_\infty) &\leq s_\infty \\ &= \liminf_{i\to\infty}\ s_i \\ &= \liminf_{i\to\infty}\ h^{\epsilon}_{n-k}(\phi_i), \end{aligned}$$

establishing lower semi-continuity. □

IV.9 Proofs from Section IV.3

IV.9.1 Estimates for the Second Moment of $\widetilde{\Phi}^n_t$

Lemma IV.40. *Let Z be a standard normal random variable. Then for $t > 0$ and $\phi \geq 0$, $E[j(t,Z,\phi)] = e^{\lambda t}(\phi+1) - 1$.*

Proof. We may write

$$E[j(t,Z,\phi)] = \int_{-\infty}^{\infty} \left(e^{\alpha z\sqrt{t} + (\lambda - \alpha^2/2)t}\phi + \int_0^t \lambda e^{\lambda u + \frac{\alpha z}{\sqrt{t}}u - \frac{\alpha^2}{2t}u^2} \right) e^{-z^2/2}/\sqrt{(2\pi)}dz.$$

The integral of the first term is calculated by completing the square and equals $e^{\lambda t}\phi$. The integral of the second term is calculated by switching the order of integration and completing the square, yielding $e^{\lambda t} - 1$. $\square$

Corollary IV.41. *Let n and k be positive integers. Then*

$$E\left[\widetilde{\Phi}^n_{\frac{k}{n}} \middle| \mathcal{F}^n_{\frac{k-1}{n}}\right] = e^{\frac{\lambda}{n}}\left(\widetilde{\Phi}^n_{\frac{k-1}{n}} + 1\right) - 1.$$

Proof. Apply Lemma IV.40 with $t = \frac{1}{n}$, using the fact that $\widetilde{\Phi}^n_{\frac{k}{n}} = j\left(\frac{1}{n}, Z_k, \widetilde{\Phi}^n_{\frac{k-1}{n}}\right)$, with Z_k independent of $\mathcal{F}^n_{\frac{k-1}{n}}$. $\square$

Lemma IV.42. *Let Z be a standard normal random variable. Then for $\phi \geq 0$ and $t > 0$,*

$$E[j(t,Z,\phi)^2] \leq \phi^2 e^{2\lambda t + \alpha^2 t} + 2\phi e^{\lambda t}\frac{\lambda}{\lambda + \alpha^2}\left(e^{\lambda t + \alpha^2 t} - 1\right) + e^{\alpha^2 t}\left(e^{\lambda t} - 1\right)^2.$$

Proof. We start by expanding $E\left[j(t,Z,\phi)^2\right]$ into three terms:

$$\begin{aligned}
&E\left[j(t,Z,\phi)^2\right] \\
&\quad = \left.\int_{-\infty}^{\infty} \phi^2 e^{2\alpha z\sqrt{t} + 2(\lambda - \alpha^2/2)t} e^{-z^2/2}/\sqrt{(2\pi)}dz\right\} (1) \\
&\qquad + \left. 2\phi\int_{-\infty}^{\infty} e^{\alpha z\sqrt{t} + (\lambda - \alpha^2/2)t}\left(\int_0^t \lambda e^{\lambda u + \frac{\alpha z}{\sqrt{t}}u - \frac{\alpha^2}{2t}u^2}du\right)e^{-z^2/2}/\sqrt{2\pi}dz\right\} (2) \\
&\qquad + \left.\int_{-\infty}^{\infty}\left(\int_0^t \lambda e^{\lambda u + \frac{\alpha z}{\sqrt{t}}u - \frac{\alpha^2}{2t}u^2}du\right)\left(\int_0^t \lambda e^{\lambda w + \frac{\alpha z}{\sqrt{t}}w - \frac{\alpha^2}{2t}w^2}dw\right)e^{-z^2/2}/\sqrt{2\pi}dz\right\} (3)
\end{aligned}$$

We calculate each of these terms separately. (1) is the simplest, and by completing the square, we can calculate its value to be $\phi^2 e^{2\lambda t}e^{\alpha^2 t}$. We have

$$\begin{aligned}
(2) &= 2\phi\int_0^t \lambda e^{\lambda t}e^{\lambda u}e^{\alpha^2 u}\left(\int_{-\infty}^{\infty} e^{-\left(z - \frac{\alpha u}{\sqrt{t}} - \alpha\sqrt{t}\right)^2/2}/\sqrt{2\pi}dz\right)du \\
&= 2\phi e^{\lambda t}\int_0^t \lambda e^{\lambda u}e^{\alpha^2 u}du \\
&= 2\phi e^{\lambda t}\frac{\lambda}{\lambda + \alpha^2}\left(e^{\lambda t + \alpha^2 t} - 1\right).
\end{aligned}$$

Now, the third term we cannot calculate exactly, but we give an upper bound for it which will be good enough:

$$\begin{aligned}
(3) &= \int_0^t \int_0^t \lambda e^{\lambda u} e^{\lambda w} e^{\frac{\alpha^2 uw}{t}} \left(\int_{-\infty}^{\infty} e^{-\left(z - \frac{\alpha u}{\sqrt{t}} - \frac{\alpha w}{\sqrt{t}}\right)^2/2} / \sqrt{2\pi} dz \right) dudw \\
&= \int_0^t \int_0^t \lambda e^{\lambda u} e^{\lambda w} e^{\frac{\alpha^2 uw}{t}} dudw \\
&\leq \int_0^t \int_0^t \lambda e^{\lambda u} e^{\lambda w} e^{\alpha^2 t} dudw \\
&= e^{\alpha^2 t} \left(e^{\lambda t} - 1\right)^2,
\end{aligned}$$

where in the inequality above, we have the used the fact that $uw \leq t^2$ for $u, w \in [0, t]$. Combining $(1), (2)$, and (3), we deduce the lemma. □

Corollary IV.43. *Let n and k be positive integers. Then*

$$E\left[\left(\widetilde{\Phi}^n_{\frac{k}{n}}\right)^2 |\mathcal{F}^n_{\frac{k-1}{n}}\right] \leq \left(\widetilde{\Phi}^n_{\frac{k-1}{n}}\right)^2 e^{\frac{1}{n}(2\lambda+\alpha^2)} + 2\widetilde{\Phi}^n_{\frac{k-1}{n}} e^{\frac{\lambda}{n}} \frac{\lambda}{\lambda+\alpha^2} \left(e^{\frac{1}{n}(\lambda+\alpha^2)} - 1\right) + e^{\frac{\alpha^2}{n}} \left(e^{\frac{\lambda}{n}} - 1\right)^2.$$

Proof. Apply Lemma IV.42, as Lemma IV.40 is used in the proof of Corollary IV.41.

□

Since we are interested in the limit as $n \to \infty$, we we also set down asymptotic versions of Lemma IV.42 and Corollary IV.43.

Lemma IV.44. *Let $\phi \geq 0$, and Z a standard normal random variable. Then as $t \downarrow 0$,*

$$E\left[j(t, Z, \phi)^2\right] = \phi^2(1 + 2\lambda t + \alpha^2 t + O(t^2)) + \phi(2\lambda t + O(t^2)) + O(t^2).$$

Proof. By examining the proof of Lemma IV.42, we can see that $E\left[j(t, Z, \phi)^2\right] = \phi^2 e^{2\lambda t} e^{\alpha^2 t} + 2\phi e^{\lambda t} \frac{\lambda}{\lambda+\alpha^2}\left(e^{\lambda t + \alpha^2 t} - 1\right)$, plus a positive third term, which can be bounded from above by $e^{\alpha^2 t}\left(e^{\lambda t} - 1\right)^2$. Note that this term

$$\begin{aligned}
e^{\alpha^2 t}\left(e^{\lambda t} - 1\right)^2 &= (1 + \alpha^2 t + O(t^2))(\lambda t + O(t^2))^2 \\
&= O(t^2).
\end{aligned}$$

So, we are left with the first two terms. We have

$$\begin{aligned}\phi^2 e^{2\lambda t} e^{\alpha^2 t} &= \phi^2(1+2\lambda t+O(t^2))(1+\alpha^2 t+O(t^2))\\ &= \phi^2(1+2\lambda t+\alpha^2 t+O(t^2)),\end{aligned}$$

and

$$\begin{aligned}2\phi e^{\lambda t}\frac{\lambda}{\lambda+\alpha^2}\left(e^{\lambda t+\alpha^2 t}-1\right) &= 2\phi(1+\lambda t+O(t^2))\frac{\lambda}{\lambda+\alpha^2}(\lambda t+\alpha^2 t+O(t^2))\\ &= 2\phi(t+O(t^2)).\end{aligned}$$

$\square$

Corollary IV.45. *Let k be a positive integer. As $n \to \infty$,*

$$E\left[\left(\widetilde{\Phi}^n_{\frac{k}{n}}\right)^2 |\mathcal{F}^n_{\frac{k-1}{n}}\right] = \left(\widetilde{\Phi}^n_{\frac{k-1}{n}}\right)^2\left(1+\frac{2\lambda}{n}+\frac{\alpha^2}{n}+O\left(\frac{1}{n^2}\right)\right)+\widetilde{\Phi}^n_{\frac{k-1}{n}}\left(\frac{2\lambda}{n}+O\left(\frac{1}{n^2}\right)\right)+O\left(\frac{1}{n^2}\right).$$

IV.9.2 Proving Proposition IV.15 by Establishing the Conditions of Proposition IV.13

We verify the six conditions of Proposition IV.13 separately. First note that Condition (a) is satisfied by construction. We will defer (b) until last.

Conditions $(c),(e)$. First, we will construct the process N^n_t. This is essentially done by Doob Decomposition. We set $N^n_0 = 0$. First, we will define N^n at grid points $\{\frac{1}{n},\frac{2}{n},\ldots\}$, and then extend to all of $\mathbb{R}_+$. We construct N^n on the grid points inductively: for $k \geq 1$,

$$\begin{aligned}N^n_{\frac{k}{n}} - N^n_{\frac{k-1}{n}} &\triangleq E\left[\widetilde{\Phi}^n_{\frac{k}{n}} - \widetilde{\Phi}^n_{\frac{k-1}{n}} - \int_{\frac{k-1}{n}}^{\frac{k}{n}} b(\widetilde{\Phi}^n_s)ds|\mathcal{F}^n_{\frac{k-1}{n}}\right]\\ &= E\left[\widetilde{\Phi}^n_{\frac{k}{n}} - \widetilde{\Phi}^n_{\frac{k-1}{n}} - \frac{1}{n}\lambda\left(1+\widetilde{\Phi}^n_{\frac{k-1}{n}}\right)|\mathcal{F}^n_{\frac{k-1}{n}}\right],\end{aligned} \tag{4.31}$$

using $b(x) = \lambda(1+x)$, as well as the fact that $\widetilde{\Phi}^n_s = \widetilde{\Phi}^n_{\frac{k-1}{n}}$ for $\frac{k-1}{n} \leq s < \frac{k}{n}$ by construction. Now by Corollary IV.43,

$$E\left[\widetilde{\Phi}^n_{\frac{k}{n}}\middle|\mathcal{F}^n_{\frac{k-1}{n}}\right] = e^{\frac{\lambda}{n}}\left(\widetilde{\Phi}^n_{\frac{k-1}{n}}+1\right)-1$$

$$(4.32)\qquad = \left(1+\frac{\lambda}{n}+\frac{\lambda^2}{n^2}+\cdots\right)\left(\widetilde{\Phi}^n_{\frac{k-1}{n}}+1\right)-1$$

$$= \left(1+\frac{\lambda}{n}+O\left(\frac{1}{n^2}\right)\right)\left(\widetilde{\Phi}^n_{\frac{k-1}{n}}+1\right)-1.$$

Therefore, plugging (4.32) in (4.31),

(4.33)

$$N^n_{\frac{k}{n}} - N^n_{\frac{k-1}{n}} = \left(1+\frac{\lambda}{n}+O\left(\frac{1}{n^2}\right)\right)\left(\widetilde{\Phi}^n_{\frac{k-1}{n}}+1\right)-1-\widetilde{\Phi}^n_{\frac{k-1}{n}}-\frac{1}{n}\lambda\left(1+\widetilde{\Phi}^n_{\frac{k-1}{n}}\right)$$

$$= \left(\frac{\lambda^2}{n^2}+\frac{\lambda^3}{n^3}+\cdots\right)\left(\widetilde{\Phi}^n_{\frac{k-1}{n}}+1\right)$$

$$= O\left(\frac{1}{n^2}\right)\left(\widetilde{\Phi}^n_{\frac{k-1}{n}}+1\right).$$

This defines N^n on all grid points. Next, for $\frac{k-1}{n} \le t < \frac{k}{n}$,

$$N^n_t - N^n_{\frac{k-1}{n}} \triangleq E\left[-\int_{\frac{k-1}{n}}^t b(\widetilde{\Phi}^n_s)ds\middle|\mathcal{F}^n_{\frac{k-1}{n}}\right]$$

$$= \left(\frac{k-1}{n}-t\right)\lambda\left(1+\widetilde{\Phi}^n_{\frac{k-1}{n}}\right),$$

using as before the forms of $b(x)$ and $\widetilde{\Phi}^n_t$ in between grid points.

We have now defined N^n_t for all $t \ge 0$, and by construction, (b) is satisfied. The goal now is to show that Condition (e) is satisfied by the N^n's as $n \to \infty$. We start with some observations about the process N^n. Recall the fixed $L > 0$ from Proposition IV.13. We let $k_{max} = k_{max}(n) \triangleq Ln$.

(1) For all n, the sequence $\overline{N}^n \triangleq \{N^n_0, N^n_{\frac{1}{n}}, N^n_{\frac{2}{n}}, \ldots\}$, is increasing: Note that from (4.33)

$$N^n_{\frac{k}{n}} - N^n_{\frac{k-1}{n}} = \left(\frac{\lambda^2}{n^2}+\frac{\lambda^3}{n^3}+\cdots\right)\left(\widetilde{\Phi}^n_{\frac{k-1}{n}}+1\right) \ge 0.$$

As a consequence of this fact, the maximum of $E\left[\overline{N}^n_T\right]$ over all of its stopping times T is equal to $E\left[\overline{N}^n_{\frac{k_{max}}{n}}\right]$.

(2) For any k, $N^n_t - N^n_{\frac{k-1}{n}} = \left(\frac{k-1}{n} - t\right) \lambda \left(1 + \widetilde{\Phi}^n_{\frac{k-1}{n}}\right)$ for $\frac{k-1}{n} \leq t < \frac{k}{n}$. This is a negative term whose magnitude is maximized when t approaches $\frac{k}{n}$. Furthermore, $N^n_{\frac{k}{n}-} - N^n_{\frac{k-1}{n}} = -\frac{\lambda}{n}\left(1 + \widetilde{\Phi}^n_{\frac{k-1}{n}}\right)$. Since $N^n_{\frac{k}{n}}$ is always nonnegative, this implies that

$$\inf_{0 \leq t \leq L} N^n_t \geq \min_{1 \leq k \leq k_{max}} -\frac{\lambda}{n}\left(1 + \widetilde{\Phi}^n_{\frac{k-1}{n}}\right). \tag{4.34}$$

Note that $\left\{\widetilde{\Phi}^n_{\frac{k}{n}} : 0 \leq k \leq k_{max}\right\}$ is a submartingale, as seen in (4.32), and so its negative is a supermartingale. Therefore, by Doob's inequality,

$$E\left[\left(\min_{1 \leq k \leq k_{max}} -\frac{\lambda}{n}\left(1 + \widetilde{\Phi}^n_{\frac{k-1}{n}}\right)\right)^2\right] \leq 2\frac{\lambda^2}{n^2} E\left[\left(1 + \widetilde{\Phi}^n_{\frac{k_{max}}{n}}\right)^2\right]. \tag{4.35}$$

We iteratively apply Corollary IV.43 to estimate the size of $\widetilde{\Phi}^n_{\frac{k}{n}}$, deducing that

$$E\left[\left(\widetilde{\Phi}^n_{\frac{k}{n}}\right)^2\right] \ll (\phi^2 + \phi) e^{\frac{k}{n}(2\lambda + \alpha^2)}. \tag{4.36}$$

In particular, for all n and for k less than or equal to $k_{max}(n) = Ln$ and fixed ϕ, the above quantity is uniformly bounded over all k. It follows from (4.34), (4.35), and (4.36), therefore, that $|| \inf_{0 \leq t \leq L} N^n_t ||_{L^2} = O\left(\frac{1}{n}\right)$. We find, therefore, that

$$\sup_{T \in \mathcal{T}^n_L} E\left[(N^n_T)^2\right] = E\left[\left(N^n_{\frac{k_{max}}{n}}\right)^2\right] + O\left(\frac{1}{n^2}\right).$$

It remains to control the size of this last term. Recall from (4.33) that $N^n_{\frac{k}{n}} - N^n_{\frac{k-1}{n}} = O\left(\frac{1}{n^2}\right)\left(\widetilde{\Phi}^n_{\frac{k-1}{n}} + 1\right)$. Iterating this formula over $1 \leq k \leq k_{max}$, we have

$$\begin{aligned} N^n_{\frac{k_{max}}{n}} &= O\left(\frac{1}{n^2}\right) \sum_{k=1}^{k_{max}} \left(\widetilde{\Phi}^n_{\frac{k-1}{n}} + 1\right) \\ &= O\left(\frac{1}{n^2}\right) \sum_{k=1}^{k_{max}} \left(\widetilde{\Phi}^n_{\frac{k-1}{n}}\right) + O\left(\frac{1}{n}\right), \end{aligned}$$

using $k_{max} = Ln$. Squaring both sides of this equation and taking expectations, we deduce by (4.36) and Hölder's Inequality that

$$
\begin{aligned}
&E\left[\left(N^n_{\frac{k_{max}}{n}}\right)^2\right] \\
&\quad = O\left(\frac{1}{n^4}\right)(k_{max}(n))^2 e^{2L(2\lambda+\alpha^2)} + O\left(\frac{1}{n^3}\right) k_{max}(n) e^{L(2\lambda+\alpha^2)} + O\left(\frac{1}{n^2}\right) \\
&\quad = O\left(\frac{1}{n^4}\right)(Ln)^2 e^{2L(2\lambda+\alpha^2)} + O\left(\frac{1}{n^3}\right)(Ln) e^{L(2\lambda+\alpha^2)} + O\left(\frac{1}{n^2}\right) \\
&\quad = O\left(\frac{1}{n^2}\right).
\end{aligned}
$$

This establishes Condition (e). $\square$

We next address Conditions (d) and (f).

Conditions (d) *and* (f). First, we will construct the process $\mathcal{N}^n_t$. The procedure mimics the one in the construction of N^n_t. We write

$$
\begin{aligned}
S^n_{\frac{k}{n}} - S^n_{\frac{k-1}{n}} &= \left(M^n_{\frac{k}{n}}\right)^2 - \left(M^n_{\frac{k}{n}}\right)^2 - \int_{\frac{k-1}{n}}^{\frac{k}{n}} a(\widetilde{\Phi}^n_s)ds \\
&= \left(M^n_{\frac{k}{n}}\right)^2 - \left(M^n_{\frac{k}{n}}\right)^2 - \frac{\alpha^2}{n}\left(\widetilde{\Phi}^n_{\frac{k-1}{n}}\right)^2,
\end{aligned} \tag{4.37}
$$

using the fact that $a(x) = \alpha^2 x^2$, and that $\widetilde{\Phi}^n_s = \widetilde{\Phi}^n_{\frac{k-1}{n}}$ for $\frac{k-1}{n} \leq s < \frac{k}{n}$. We set $\mathcal{N}^n_0 = 0$. We first define $\mathcal{N}^n$ on the grid points $\{\frac{1}{n}, \frac{2}{n}, \ldots\}$. Define

$$
\begin{aligned}
\mathcal{N}^n_{\frac{k}{n}} - \mathcal{N}^n_{\frac{k-1}{n}} &\triangleq E\left[S^n_{\frac{k}{n}} - S^n_{\frac{k-1}{n}} \middle| \mathcal{F}^n_{\frac{k-1}{n}}\right] \\
&= E\left[\left(\widetilde{\Phi}^n_{\frac{k}{n}} - \int_0^{\frac{k}{n}} b(\widetilde{\Phi}^n_s)ds - N^n_{\frac{k}{n}}\right)^2 \middle| \mathcal{F}^n_{\frac{k-1}{n}}\right] \\
&\quad - E\left[\left(\widetilde{\Phi}^n_{\frac{k-1}{n}} - \int_0^{\frac{k-1}{n}} b(\widetilde{\Phi}^n_s)ds - N^n_{\frac{k-1}{n}}\right)^2 \middle| \mathcal{F}^n_{\frac{k-1}{n}}\right] \\
&\quad - \frac{\alpha^2}{n}\left(\widetilde{\Phi}^n_{\frac{k-1}{n}}\right)^2.
\end{aligned} \tag{4.38}
$$

We can expand the first term in (4.38) as

$$
E\left[\left(\widetilde{\Phi}^n_{\frac{k-1}{n}} + \left(\widetilde{\Phi}^n_{\frac{k}{n}} - \widetilde{\Phi}^n_{\frac{k-1}{n}}\right) - \int_0^{\frac{k-1}{n}} b(\widetilde{\Phi}^n_s)ds - \int_{\frac{k-1}{n}}^{\frac{k}{n}} b(\widetilde{\Phi}^n_s)ds - N^n_{\frac{k-1}{n}} - \left(N^n_{\frac{k}{n}} - N^n_{\frac{k-1}{n}}\right)\right)^2 \middle| \mathcal{F}^n_{\frac{k-1}{n}}\right],
$$

or

$$E\left[\left(\left(\widetilde{\Phi}^n_{\frac{k-1}{n}} - \int_0^{\frac{k-1}{n}} b(\widetilde{\Phi}^n_s)ds - N^n_{\frac{k-1}{n}}\right) + \left(\widetilde{\Phi}^n_{\frac{k}{n}} - \widetilde{\Phi}^n_{\frac{k-1}{n}}\right) - \int_{\frac{k-1}{n}}^{\frac{k}{n}} b(\widetilde{\Phi}^n_s)ds - \left(N^n_{\frac{k}{n}} - N^n_{\frac{k-1}{n}}\right)\right)^2 \middle| \mathcal{F}^n_{\frac{k-1}{n}}\right].$$

This is designed to cancel with the second term in (4.38). After some algebraic manipulation, we arrive at

$$\begin{aligned}
&\mathcal{N}^n_{\frac{k}{n}} - \mathcal{N}^n_{\frac{k-1}{n}} \\
&= 2\left(\widetilde{\Phi}^n_{\frac{k-1}{n}} - \int_0^{\frac{k-1}{n}} b(\widetilde{\Phi}^n_s)ds - N^n_{\frac{k-1}{n}}\right) E\left[\left(\widetilde{\Phi}^n_{\frac{k}{n}} - \widetilde{\Phi}^n_{\frac{k-1}{n}}\right) - \int_{\frac{k-1}{n}}^{\frac{k}{n}} b(\widetilde{\Phi}^n_s)ds - \left(N^n_{\frac{k}{n}} - N^n_{\frac{k-1}{n}}\right) \middle| \mathcal{F}^n_{\frac{k-1}{n}}\right] \\
&\quad + E\left[\left(\left(\widetilde{\Phi}^n_{\frac{k}{n}} - \widetilde{\Phi}^n_{\frac{k-1}{n}}\right) - \int_{\frac{k-1}{n}}^{\frac{k}{n}} b(\widetilde{\Phi}^n_s)ds - \left(N^n_{\frac{k}{n}} - N^n_{\frac{k-1}{n}}\right)\right)^2 \middle| \mathcal{F}^n_{\frac{k-1}{n}}\right] \\
&\quad - \frac{\alpha^2}{n}\left(\widetilde{\Phi}^n_{\frac{k-1}{n}}\right)^2.
\end{aligned}$$

In the first term of the right hand side above, the conditional expectation is zero, by definition of N^n. So we have

$$\begin{aligned}
\mathcal{N}^n_{\frac{k}{n}} - \mathcal{N}^n_{\frac{k-1}{n}} &= E\left[\left(\left(\widetilde{\Phi}^n_{\frac{k}{n}} - \widetilde{\Phi}^n_{\frac{k-1}{n}}\right) - \int_{\frac{k-1}{n}}^{\frac{k}{n}} b(\widetilde{\Phi}^n_s)ds - \left(N^n_{\frac{k}{n}} - N^n_{\frac{k-1}{n}}\right)\right)^2 \middle| \mathcal{F}^n_{\frac{k-1}{n}}\right] - \frac{\alpha^2}{n}\left(\widetilde{\Phi}^n_{\frac{k-1}{n}}\right)^2 \\
&= E\left[\left(\widetilde{\Phi}^n_{\frac{k}{n}} - \left(\widetilde{\Phi}^n_{\frac{k-1}{n}} + \frac{\lambda}{n}\left(1+\widetilde{\Phi}^n_{\frac{k-1}{n}}\right) + \left(N^n_{\frac{k}{n}} - N^n_{\frac{k-1}{n}}\right)\right)\right)^2 \middle| \mathcal{F}^n_{\frac{k-1}{n}}\right] - \frac{\alpha^2}{n}\left(\widetilde{\Phi}^n_{\frac{k-1}{n}}\right)^2.
\end{aligned}$$

We will now expand this term above. Recall the useful facts that $E\left[\widetilde{\Phi}^n_{\frac{k}{n}} \middle| \mathcal{F}^n_{\frac{k-1}{n}}\right] = e^{\frac{\lambda}{n}}\left(\widetilde{\Phi}^n_{\frac{k-1}{n}} + 1\right) - 1$, and that $N^n_{\frac{k}{n}}$ is $\mathcal{F}^n_{\frac{k-1}{n}}$-measurable. We have

$$\begin{aligned}
\mathcal{N}^n_{\frac{k}{n}} - \mathcal{N}^n_{\frac{k-1}{n}} &= E\left[\left(\widetilde{\Phi}^n_{\frac{k}{n}}\right)^2 \middle| \mathcal{F}^n_{\frac{k-1}{n}}\right] - 2\left(\widetilde{\Phi}^n_{\frac{k-1}{n}} + \frac{\lambda}{n}\left(1+\widetilde{\Phi}^n_{\frac{k-1}{n}}\right) + \left(N^n_{\frac{k}{n}} - N^n_{\frac{k-1}{n}}\right)\right) E\left[\widetilde{\Phi}^n_{\frac{k}{n}} \middle| \mathcal{F}^n_{\frac{k-1}{n}}\right] \\
&\quad + \left(\widetilde{\Phi}^n_{\frac{k-1}{n}} + \frac{\lambda}{n}\left(1+\widetilde{\Phi}^n_{\frac{k-1}{n}}\right) + \left(N^n_{\frac{k}{n}} - N^n_{\frac{k-1}{n}}\right)\right)^2 - \frac{\alpha^2}{n}\left(\widetilde{\Phi}^n_{\frac{k-1}{n}}\right)^2.
\end{aligned} \tag{4.39}$$

The second term on the right hand side of (4.39) is equal to

$$-2\left(\widetilde{\Phi}^n_{\frac{k-1}{n}} + \frac{\lambda}{n}\left(1+\widetilde{\Phi}^n_{\frac{k-1}{n}}\right) + \left(N^n_{\frac{k}{n}} - N^n_{\frac{k-1}{n}}\right)\right)\left(e^{\frac{\lambda}{n}}\left(\widetilde{\Phi}^n_{\frac{k-1}{n}} + 1\right) - 1\right). \tag{4.40}$$

According to Equations (4.33) and (4.36), $\left(N^n_{\frac{k}{n}} - N^n_{\frac{k-1}{n}}\right) = O\left(\frac{1}{n^2}\right)\left(\widetilde{\Phi}^n_{\frac{k-1}{n}} + 1\right)$, uniformly over all $k \le k_{max}(n)$. Then (4.40) becomes

$$-2\left[\widetilde{\Phi}^n_{\frac{k-1}{n}} + \frac{\lambda}{n}\left(1+\widetilde{\Phi}^n_{\frac{k-1}{n}}\right) + O\left(\frac{1}{n^2}\right)\left(\widetilde{\Phi}^n_{\frac{k-1}{n}} + 1\right)\right]\left[e^{\frac{\lambda}{n}}\left(\widetilde{\Phi}^n_{\frac{k-1}{n}} + 1\right) - 1\right]$$

$$= -2\left[\widetilde{\Phi}^n_{\frac{k-1}{n}} + \frac{\lambda}{n}\left(1+\widetilde{\Phi}^n_{\frac{k-1}{n}}\right) + O\left(\frac{1}{n^2}\right)\left(\widetilde{\Phi}^n_{\frac{k-1}{n}}+1\right)\right]\left[\widetilde{\Phi}^n_{\frac{k-1}{n}} + \frac{\lambda}{n}\left(\widetilde{\Phi}^n_{\frac{k-1}{n}}+1\right) + O\left(\frac{1}{n^2}\right)\left(\widetilde{\Phi}^n_{\frac{k-1}{n}}+1\right)\right].$$

As we can see, this term partially cancels with the third term in Equation (4.39). In doing so, we obtain

(4.41)

$$\begin{aligned}\mathcal{N}^n_{\frac{k}{n}} - \mathcal{N}^n_{\frac{k-1}{n}} &= E\left[\left(\widetilde{\Phi}^n_{\frac{k}{n}}\right)^2 |\mathcal{F}^n_{\frac{k-1}{n}}\right] - \left(\widetilde{\Phi}^n_{\frac{k-1}{n}} + \frac{\lambda}{n}\left(1+\widetilde{\Phi}^n_{\frac{k-1}{n}}\right)\right)^2 - \frac{\alpha^2}{n}\left(\widetilde{\Phi}^n_{\frac{k-1}{n}}\right)^2 + O\left(\frac{1}{n^2}\right)\left(\widetilde{\Phi}^n_{\frac{k-1}{n}}\right)^2 \\ &= E\left[\left(\widetilde{\Phi}^n_{\frac{k}{n}}\right)^2 |\mathcal{F}^n_{\frac{k-1}{n}}\right] - \left(\widetilde{\Phi}^n_{\frac{k-1}{n}}\right)^2\left(1+\frac{2\lambda}{n}\right) - \frac{2\lambda}{n}\widetilde{\Phi}^n_{\frac{k-1}{n}} + O\left(\frac{1}{n^2}\right)\left(\left(\widetilde{\Phi}^n_{\frac{k-1}{n}}\right)^2 + \widetilde{\Phi}^n_{\frac{k-1}{n}} + 1\right).\end{aligned}$$

From Corollary IV.45, we have

$$E\left[\left(\widetilde{\Phi}^n_{\frac{k}{n}}\right)^2 |\mathcal{F}^n_{\frac{k-1}{n}}\right] = \left(\widetilde{\Phi}^n_{\frac{k-1}{n}}\right)^2\left(1+\frac{2\lambda}{n}+\frac{\alpha^2}{n}+O\left(\frac{1}{n^2}\right)\right) + \widetilde{\Phi}^n_{\frac{k-1}{n}}\left(\frac{2\lambda}{n}+O\left(\frac{1}{n^2}\right)\right) + O\left(\frac{1}{n^2}\right),$$

and this perfectly cancels out with the second and third terms in (4.41). We are ultimately left with

$$\mathcal{N}^n_{\frac{k}{n}} - \mathcal{N}^n_{\frac{k-1}{n}} = O\left(\frac{1}{n^2}\right)\left(\left(\widetilde{\Phi}^n_{\frac{k-1}{n}}\right)^2 + \widetilde{\Phi}^n_{\frac{k-1}{n}} + 1\right). \tag{4.42}$$

Before pursuing this line of reasoning further, let us define $\mathcal{N}^n_t$ between grid points. For $\frac{k-1}{n} \le t < \frac{k}{n}$,

$$\begin{aligned}\mathcal{N}^n_t - \mathcal{N}^n_{\frac{k-1}{n}} &\triangleq E\left[-\int_{\frac{k-1}{n}}^t a(\widetilde{\Phi}^n_s)ds|\mathcal{F}^n_{\frac{k-1}{n}}\right] \\ &= \left(\frac{k-1}{n} - t\right)\alpha^2\left(\widetilde{\Phi}^n_{\frac{k-1}{n}}\right)^2,\end{aligned}$$

using $a(x) = \alpha^2 x^2$ as well as the the fact that $\widetilde{\Phi}^n_s = \widetilde{\Phi}^n_{\frac{k-1}{n}}$ for $\frac{k-1}{n} \le s < \frac{k}{n}$. As in the proof of Parts (c) and (e), we may show, using the submartingality of $\left\{\widetilde{\Phi}^n_{\frac{k}{n}} : k = 0, 1, \ldots\right\}$, that the L^1 norm $\left|\left|\inf_{0\le k\le k_{max}(n)} -\frac{\alpha^2}{n}\left(\widetilde{\Phi}^n_{\frac{k-1}{n}}\right)^2\right|\right|_{L^1} \ll \frac{1}{n}||\widetilde{\Phi}^n_{\frac{k_{max}}{n}}||_{L^2}$, which is $O\left(\frac{1}{n}\right)$. Therefore, in attempting to establish (e), we may ignore any possible times in between grid points $\{0, \frac{1}{n}, \ldots\}$. Consequently, we consider the discrete-time process

$$\overline{\mathcal{N}}^N \triangleq \{\mathcal{N}^n_0, \mathcal{N}^n_{\frac{1}{n}}, \ldots, \mathcal{N}^n_{\frac{k_{max}}{n}}\},$$

with associated bounded stopping times $\overline{\mathcal{T}}_L^N$, and we must show

$$\sup_{\tau \in \overline{\mathcal{T}}_L^N} E[|\mathcal{N}_T^N|] \to 0 \text{ as } n \to \infty.$$

According to (4.42), for any k,

$$\left|\mathcal{N}_{\frac{k}{n}}^n\right| = O\left(\frac{1}{n^2}\right) \sum_{j=1}^{k} \left(\left(\widetilde{\Phi}_{\frac{k-1}{n}}^n\right)^2 + \widetilde{\Phi}_{\frac{k-1}{n}}^n + 1\right).$$

Therefore, noting that $\widetilde{\Phi}_{\frac{k}{n}}^n$ is always nonnegative

$$\sup_{0 \le k \le k_{max}} \left|\mathcal{N}_{\frac{k}{n}}^n\right| = O\left(\frac{1}{n^2}\right) \sum_{j=1}^{k_{max}} \left(\left(\widetilde{\Phi}_{\frac{k-1}{n}}^n\right)^2 + \widetilde{\Phi}_{\frac{k-1}{n}}^n + 1\right).$$

Now, since $\left\{\widetilde{\Phi}_{\frac{k}{n}}^n\right\}_{k \ge 0}$ is also a submartingale and $k_{max} = Ln$, it follows that

$$\begin{aligned} E\left[\sup_{0 \le k \le k_{max}} \left|\mathcal{N}_{\frac{k}{n}}^n\right|\right] &= O\left(\frac{1}{n^2}\right) E\left[\sum_{j=1}^{k_{max}} \left(\left(\widetilde{\Phi}_{\frac{k-1}{n}}^n\right)^2 + \widetilde{\Phi}_{\frac{k-1}{n}}^n + 1\right)\right] \\ &\le O\left(\frac{1}{n^2}\right) (Ln) E\left[\left(\widetilde{\Phi}_{\frac{k_{max}}{n}}^n\right)^2 + \widetilde{\Phi}_{\frac{k_{max}}{n}}^n + 1\right] \\ &= O\left(\frac{1}{n}\right). \end{aligned}$$

□

Condition (*b*). First, we claim that $E\left[(j(t, Z, \phi) - \phi)^2\right] = O(t)\phi^2$ as $t \to 0$, for Z a standard normal random variable. Supposing that this claim is established, then applied conditionally, it entails that

$$E\left[\left(\widetilde{\Phi}_{\frac{k}{n}}^n - \widetilde{\Phi}_{\frac{k-1}{n}}^n\right)^2 |\mathcal{F}_{\frac{k-1}{n}}^n\right] = O\left(\frac{1}{n}\right) \left(\widetilde{\Phi}_{\frac{k-1}{n}}^n\right)^2,$$

and this will establish (*b*), given our control over the L^2 norm of $\sup_{0 \le k \le Ln} \left|\widetilde{\Phi}_{\frac{k}{n}}^n\right|$. So, let us address the claim.

We have

$$\begin{aligned} E\left[(j(t, Z, \phi) - \phi)^2\right] &= E[j(t, Z, \phi)^2] - 2\phi E[j(t, Z, \phi)] + \phi^2 \\ &= \phi^2 (1 + O(t)) - 2\phi(\phi + O(t)) + \phi^2 \\ &= O(t)\phi^2; \end{aligned}$$

here, we have used Lemma IV.44 for the first term, and Lemma IV.40 for the second term.

□

IV.10 Proofs from Section IV.4

Proof of Proposition IV.23. The proof will be by backwards induction on j. For reference, the reader should see Figures 4.1 and 4.2: we will use backwards induction on the columns in Figure 4.1, and in each column, we will use backwards induction on the elements of the column. First, the base case, where we will prove equality. Suppose that $j = n$ and $0 \le k \le n$. Let $\Psi^k = \{\psi_1^k, \ldots, \psi_k^k\} \in \boldsymbol{\mathfrak{O}}^k$. As mentioned before, the Poisson process N is assumed to be stopped at η_n, so that $\eta_{n+1} = \infty$. Since all observation rights have arrived by the time η_j when $j = n$, we have remaining a total of $n - k$ observation rights, with no restrictions on when they may be used. Therefore, following the proof of Proposition IV.4 with slight modifications, we may establish that for each $0 \le k \le n$, $\boldsymbol{\gamma}_{n,k}^n(\Psi^k) = \boldsymbol{v}_{n,k}^n \left(\psi_k^k \vee \eta_n - \psi_k^k, \Phi_{\psi_k^k \vee \eta_n}^{\Psi^k}\right)$.

We next tackle the inductive step. Suppose then that

$$\boldsymbol{\gamma}_{j+1,k}^n(\Psi'^k) \ge \boldsymbol{v}_{j+1,k}^n \left(\psi_k'^k \vee \eta_{j+1} - \psi_k'^k, \Phi_{\psi_k'^k \vee \eta_{j+1}}^{\Psi'^k}\right)$$

holds for all $0 \le k \le j+1$ and any $\Psi'^k \in \boldsymbol{\mathfrak{O}}^k$, on the set $\{\psi_k'^k < \eta_{(j+1)+1}\}$. We wish to show that $\boldsymbol{\gamma}_{j,k}^n(\Psi^k) \ge \boldsymbol{v}_{j,k}^n \left(\psi_k^k \vee \eta_j - \psi_k^k, \Phi_{\psi_k^k \vee \eta_j}^{\Psi^k}\right)$ holds for all $0 \le k \le j$ and any $\Psi^k \in \boldsymbol{\mathfrak{O}}^k$, on the set $\{\psi_k^k < \eta_{j+1}\}$. To establish this, we will proceed with another round of backwards induction, this time on k, starting from $k = j$. So, we fix $\Psi^j = \{\psi_1^j, \ldots, \psi_j^j\} \in \boldsymbol{\mathfrak{O}}^j$. Let $\widetilde{\Psi} = \{\psi_1^j, \ldots, \psi_j^j, \widetilde{\psi}_{j+1}, \ldots, \widetilde{\psi}_n\} \in \boldsymbol{\mathfrak{O}}_{j,j}^n(\Psi^j)$. Note that since $\widetilde{\Psi} \in \boldsymbol{\mathfrak{O}}^n$, it is the case that $\psi_j^j \ge \eta_j$, so that $\psi_j^j \vee \eta_j = \psi_j^j$. We have, with all arguments taking place on the set $\{\psi_j^j < \eta_{j+1}\}$,

$$\text{(4.43)} \qquad \operatorname*{ess\,inf}_{\tau \in \mathcal{T}_s^{\widetilde{\Psi}}, \tau \geq \psi_j^j} E\left[\int_{\psi_j^j}^{\tau} e^{-\lambda(t-\psi_j^j)}\left(\Phi_t^{\widetilde{\Psi}} - \frac{\lambda}{c}\right) dt \big| \mathcal{F}_{\psi_j^j}^{\Psi^j}\right]$$

$$= \operatorname*{ess\,inf}_{\tau \in \mathcal{T}_s^{\widetilde{\Psi}}, \tau \geq \psi_j^j} E\Bigg[\int_{\psi_j^j}^{\eta_{j+1}\wedge\tau} e^{-\lambda(t-\psi_j^j)}\left(\Phi_t^{\widetilde{\Psi}} - \frac{\lambda}{c}\right) dt$$

$$+1_{\{\tau>\eta_{j+1}\}}\int_{\eta_{j+1}}^{\tau\vee\eta_{j+1}} e^{-\lambda(t-\psi_j^j)}\left(\Phi_t^{\widetilde{\Psi}} - \frac{\lambda}{c}\right) dt \big| \mathcal{F}_{\psi_j^j}^{\Psi^j}\Bigg].$$

For each such τ above, $\tau \vee \eta_{j+1}$ is an element of $\mathcal{T}^{\widetilde{\Psi}}$ which is greater than $\eta_{j+1} = \psi_j^j \vee \eta_{j+1}$. Therefore, the right hand side of (4.43) above is greater than or equal to

$$\operatorname*{ess\,inf}_{\tau \in \mathcal{T}_s^{\widetilde{\Psi}}, \tau \geq \psi_j^j} E\left[\int_{\psi_j^j}^{\eta_{j+1}\wedge\tau} e^{-\lambda(t-\psi_j^j)}\left(\Phi_t^{\widetilde{\Psi}} - \frac{\lambda}{c}\right) dt + 1_{\{\tau>\eta_{j+1}\}} e^{-\lambda(\eta_{j+1}-\psi_j^j)} \boldsymbol{\gamma}_{j+1,j}^n(\Psi^j) \big| \mathcal{F}_{\psi_j^j}^{\Psi^j}\right],$$

which by the (initial) induction hypothesis is greater than or equal to

$$\text{(4.44)} \quad \operatorname*{ess\,inf}_{\tau \in \mathcal{T}_s^{\widetilde{\Psi}}, \tau \geq \psi_j^j} E\Bigg[\int_{\psi_j^j}^{\eta_{j+1}\wedge\tau} e^{-\lambda(t-\psi_j^j)}\left(\Phi_t^{\widetilde{\Psi}} - \frac{\lambda}{c}\right) dt$$

$$+1_{\{\tau>\eta_{j+1}\}} e^{-\lambda(\eta_{j+1}-\psi_j^j)} \boldsymbol{v}_{j+1,j}^n\left(\eta_{j+1}-\psi_j^j, \varphi\left(\eta_{j+1}-\psi_j^j, \Phi_{\psi_j^j}^{\Psi^j}\right)\right) \big| \mathcal{F}_{\psi_j^j}^{\Psi^j}\Bigg].$$

Now, recall that $\eta_{j+1} - \psi_j^j$ is independent of $\mathcal{F}_{\psi_j^j}^{\Psi^k}$ and exponentially distributed with parameter μ, on the set $\{\psi_j^j < \eta_{j+1}\}$. Additionally, from Theorem 3.2 of [15], it can be seen that for some nonnegative random random variable $R_j \in m\mathcal{F}_{\psi_j^j}^{\Psi^j}$, $1_{\{\tau>\eta_{j+1}\}} = 1_{\{R_j>\eta_{j+1}-\psi_j^j\}}$, and $\tau 1_{\{\tau\leq\eta_{j+1}\}} = (\psi_j^j + R_j)1_{\{R_j\leq\eta_{j+1}-\psi_j^j\}}$. Using the deterministic dynamics of Φ^{Ψ^j} in between observations, (4.44) becomes

$$\operatorname*{ess\,inf}_{R_j \in m\mathcal{F}_{\psi_j^j}^{\Psi^j}, R_j\geq 0} \int_0^\infty \mu e^{-\mu u}\left[\int_0^{R_j\wedge u} e^{-\lambda t}\left(\varphi\left(t, \Phi_{\psi_j^j}^{\Psi^j}\right) - \frac{\lambda}{c}\right) dt + 1_{\{R_j>u\}} e^{-\lambda u} \boldsymbol{v}_{j+1,j}^n\left(u, \varphi\left(u, \Phi_{\psi_j^j}^{\Psi^j}\right)\right)\right] du,$$

which is equal to

$$\operatorname*{ess\,inf}_{R_j \in m\mathcal{F}_{\psi_j^j}^{\Psi^j}, R_j\geq 0} J^0 \boldsymbol{v}_{j+1,j}^n\left(R_j, 0, \Phi_{\psi_j^j}^{\Psi^j}\right) \geq J_0^0 \boldsymbol{v}_{j+1,j}^n\left(0, \Phi_{\psi_j^j}^{\Psi^j}\right)$$

$$= \boldsymbol{v}_{j,j}^n\left(0, \Phi_{\psi_j^j}^{\Psi^j}\right).$$

This implies that

$$\operatorname*{ess\,inf}_{\tau \in \mathcal{T}_s^{\widetilde{\Psi}}, \tau \geq \psi_j^j} E\left[\int_{\psi_j^j}^{\tau} e^{-\lambda(t-\psi_j^j)}\left(\Phi_t^{\widetilde{\Psi}} - \frac{\lambda}{c}\right) dt \big| \mathcal{F}_{\psi_j^j}^{\Psi^k}\right] \geq \boldsymbol{v}_{j,j}^n\left(0, \Phi_{\psi_j^j}^{\Psi^j}\right)$$

on the set $\{\psi_j^j < \eta_{j+1}\}$. Taking the infimum over all $\widetilde{\Psi} \in \mathfrak{O}_{j,j}^n(\Psi^j)$, we obtain that $\boldsymbol{\gamma}_{j,j}^n(\Psi^j) \geq \boldsymbol{v}_{j,j}^n\left(0, \Phi_{\psi_j^j}^{\Psi^j}\right)$ on $\{\psi_j^j < \eta_{j+1}\}$, which establishes the base case in the second induction.

We now proceed to the inductive step in the second induction. Our hypothesis is that $\boldsymbol{\gamma}_{j,k+1}^n\left(\Psi^{k+1}\right) \geq \boldsymbol{v}_{j,k+1}^n\left(\psi_{k+1}^{k+1} \vee \eta_j - \psi_{k+1}^{k+1}, \Phi_{\psi_{k+1}^{k+1}\vee\eta_j}^{\Psi^{k+1}}\right)$ holds for all $\Psi^{k+1} \in \mathfrak{O}^{k+1}$ on the set $\left\{\psi_{k+1}^{k+1} < \eta_{j+1}\right\}$. We wish to show that for any $\Psi^k = \{\psi_1^k, \ldots, \psi_k^k\} \in \mathfrak{O}^k$, it is the case that

$$\boldsymbol{\gamma}_{j,k}^n(\Psi^k) \geq \boldsymbol{v}_{j,k}^n\left(\psi_k^k \vee \eta_j - \psi_k^k, \Phi_{\psi_k^k\vee\eta_j}^{\Psi^k}\right) \text{ on } \{\psi_k^k < \eta_{j+1}\}.$$

So, let $\widetilde{\Psi} = \{\psi_1^k, \ldots, \psi_k^k, \widetilde{\psi}_{k+1}, \ldots, \widetilde{\psi}_n\} \in \mathfrak{O}_{j,k}^n(\Psi^k)$. We write, on the set $\{\psi_k^k < \eta_{j+1}\}$,

$$\begin{aligned}
&\operatorname*{ess\,inf}_{\tau\in\mathcal{T}_s^{\widetilde{\Psi}},\tau\geq\psi_k^k\vee\eta_j} E\left[\int_{\psi_k^k\vee\eta_j}^{\tau} e^{-\lambda(t-\psi_k^k\vee\eta_j)}\left(\Phi_t^{\widetilde{\Psi}} - \frac{\lambda}{c}\right)\big|\mathcal{F}_{\psi_k^k\vee\eta_j}^{\widetilde{\Psi}}\right] \\
&\quad = \min\Bigg\{0, \operatorname*{ess\,inf}_{\tau\in\mathcal{T}_s^{\widetilde{\Psi}},\tau\geq\widetilde{\psi}_{k+1}\wedge\eta_{j+1}} E\Bigg[\int_{\psi_k^k\vee\eta_j}^{\widetilde{\psi}_{k+1}\wedge\eta_{j+1}} e^{-\lambda(t-\psi_k^k\vee\eta_j)}\left(\Phi_t^{\widetilde{\Psi}} - \frac{\lambda}{c}\right)dt \\
&\qquad + \int_{\widetilde{\psi}_{k+1}\wedge\eta_{j+1}}^{\tau} e^{-\lambda(t-\psi_k^k\vee\eta_j)}\left(\Phi_t^{\widetilde{\Psi}} - \frac{\lambda}{c}\right)dt\big|\mathcal{F}_{\psi_k^k\vee\eta_j}^{\widetilde{\Psi}}\Bigg]\Bigg\},
\end{aligned}$$

using the fact that $\tau \in \mathcal{T}_s^{\Psi}$, i.e. it does not stop the game between observations while there are remaining observation rights. This then equals

$$\begin{aligned}
&\min\Bigg\{0, E\left[\int_{\psi_k^k\vee\eta_j}^{\widetilde{\psi}_{k+1}\wedge\eta_{j+1}} e^{-\lambda(t-\psi_k^k\vee\eta_j)}\left(\Phi_t^{\widetilde{\Psi}} - \frac{\lambda}{c}\right)dt\big|\mathcal{F}_{\psi_k^k\vee\eta_j}^{\widetilde{\Psi}}\right] \\
&\qquad + \operatorname*{ess\,inf}_{\tau\in\mathcal{T}_s^{\widetilde{\Psi}},\tau\geq\widetilde{\psi}_{k+1}\wedge\eta_{j+1}} E\left[\int_{\widetilde{\psi}_{k+1}\wedge\eta_{j+1}}^{\tau} e^{-\lambda(t-\psi_k^k\vee\eta_j)}\left(\Phi_t^{\widetilde{\Psi}} - \frac{\lambda}{c}\right)dt\big|\mathcal{F}_{\psi_k^k\vee\eta_j}^{\widetilde{\Psi}}\right]\Bigg\} \\
&\quad = \min\Bigg\{0, E\left[\int_{\psi_k^k\vee\eta_j}^{\widetilde{\psi}_{k+1}\wedge\eta_{j+1}} e^{-\lambda(t-\psi_k^k\vee\eta_j)}\left(\Phi_t^{\widetilde{\Psi}} - \frac{\lambda}{c}\right)dt\big|\mathcal{F}_{\psi_k^k\vee\eta_j}^{\widetilde{\Psi}}\right] \\
&\qquad + \operatorname*{ess\,inf}_{\tau\in\mathcal{T}_s^{\widetilde{\Psi}},\tau\geq\widetilde{\psi}_{k+1}\wedge\eta_{j+1}} E\Bigg[1_{\{\widetilde{\psi}_{k+1}<\eta_{j+1}\}}\int_{\widetilde{\psi}_{k+1}}^{\tau} e^{-\lambda(t-\psi_k^k\vee\eta_j)}\left(\Phi_t^{\widetilde{\Psi}} - \frac{\lambda}{c}\right)dt \\
&\qquad + 1_{\{\eta_{j+1}\leq\widetilde{\psi}_{k+1}\}}\int_{\eta_{j+1}}^{\tau} e^{-\lambda(t-\psi_k^k\vee\eta_j)}\left(\Phi_t^{\widetilde{\Psi}} - \frac{\lambda}{c}\right)dt\big|\mathcal{F}_{\psi_k^k\vee\eta_j}^{\widetilde{\Psi}}\Bigg]\Bigg\}
\end{aligned}$$

$$
\begin{aligned}
(4.45)\quad &\geq \min\Bigg\{0, E\left[\int_{\psi_k^k\vee\eta_j}^{\widetilde{\psi}_{k+1}\wedge\eta_{j+1}} e^{-\lambda(t-\psi_k^k\vee\eta_j)}\left(\Phi_t^{\widetilde{\Psi}}-\frac{\lambda}{c}\right)dt\Big|\mathcal{F}^{\widetilde{\Psi}}_{\psi_k^k\vee\eta_j}\right]\\
&+\operatorname*{ess\,inf}_{\tau\in\mathcal{T}_s^{\widetilde{\Psi}},\tau\geq\widetilde{\psi}_{k+1}\wedge\eta_{j+1}} E\left[1_{\{\widetilde{\psi}_{k+1}<\eta_{j+1}\}}\int_{\widetilde{\psi}_{k+1}}^{\tau} e^{-\lambda(t-\psi_k^k\vee\eta_j)}\left(\Phi_t^{\widetilde{\Psi}}-\frac{\lambda}{c}\right)dt\Big|\mathcal{F}^{\widetilde{\Psi}}_{\psi_k^k\vee\eta_j}\right]\\
&+\operatorname*{ess\,inf}_{\tau\in\mathcal{T}_s^{\widetilde{\Psi}},\tau\geq\widetilde{\psi}_{k+1}\wedge\eta_{j+1}} E\left[1_{\{\eta_{j+1}\leq\widetilde{\psi}_{k+1}\}}\int_{\eta_{j+1}}^{\tau} e^{-\lambda(t-\psi_k^k\vee\eta_j)}\left(\Phi_t^{\widetilde{\Psi}}-\frac{\lambda}{c}\right)dt\Big|\mathcal{F}^{\widetilde{\Psi}}_{\psi_k^k\vee\eta_j}\right]\Bigg\},
\end{aligned}
$$

which we claim is equal to

$$
\begin{aligned}
(4.46)\quad &\min\Bigg\{0, E\left[\int_{\psi_k^k\vee\eta_j}^{\widetilde{\psi}_{k+1}\wedge\eta_{j+1}} e^{-\lambda(t-\psi_k^k\vee\eta_j)}\left(\Phi_t^{\widetilde{\Psi}}-\frac{\lambda}{c}\right)dt\Big|\mathcal{F}^{\widetilde{\Psi}}_{\psi_k^k\vee\eta_j}\right]\\
&+\operatorname*{ess\,inf}_{\tau\in\mathcal{T}_s^{\widetilde{\Psi}},\tau\geq\widetilde{\psi}_{k+1}} E\left[1_{\{\widetilde{\psi}_{k+1}<\eta_{j+1}\}}\int_{\widetilde{\psi}_{k+1}}^{\tau} e^{-\lambda(t-\psi_k^k\vee\eta_j)}\left(\Phi_t^{\widetilde{\Psi}}-\frac{\lambda}{c}\right)dt\Big|\mathcal{F}^{\widetilde{\Psi}}_{\psi_k^k\vee\eta_j}\right]\\
&+\operatorname*{ess\,inf}_{\tau\in\mathcal{T}_s^{\widetilde{\Psi}},\tau\geq\eta_{j+1}} E\left[1_{\{\eta_{j+1}\leq\widetilde{\psi}_{k+1}\}}\int_{\eta_{j+1}}^{\tau} e^{-\lambda(t-\psi_k^k\vee\eta_j)}\left(\Phi_t^{\widetilde{\Psi}}-\frac{\lambda}{c}\right)dt\Big|\mathcal{F}^{\widetilde{\Psi}}_{\psi_k^k\vee\eta_j}\right]\Bigg\}.
\end{aligned}
$$

It is trivially true that (4.45) $\leq$ (4.46), since its infimums are over a larger set of stopping times. We claim that the other inequality also holds. To see this, let $\tau \geq \widetilde{\psi}_{k+1}\wedge\eta_{j+1}$ be given. Consider $\tau' = \tau 1_{\{\widetilde{\psi}_{k+1}<\eta_{j+1}\}} + \infty 1_{\{\eta_{j+1}\leq\widetilde{\psi}_{k+1}\}}$, which is a stopping time greater than $\widetilde{\psi}_{k+1}$ and takes the same value as τ on the set $1_{\{\widetilde{\psi}_{k+1}<\eta_{j+1}\}}$. The existence of such a τ' implies that minimizing over $\{\tau\in\mathcal{T}_s^{\widetilde{\Psi}} : \tau\geq\widetilde{\psi}_{k+1}\wedge\eta_{j+1}\}$ is equivalent to minimizing over $\{\tau\in\mathcal{T}_s^{\widetilde{\Psi}} : \tau\geq\widetilde{\psi}_{k+1}\}$, on the set $\{\widetilde{\psi}_{k+1}<\eta_{j+1}\}$. A similar procedure may be done for stopping times on the set $1_{\{\eta_{j+1}\leq\widetilde{\psi}_{k+1}\}}$, to establish the equality of (4.45) and (4.46).

Next, conditioning the second and third terms of (4.46) on, respectively, $\mathcal{F}^{\widetilde{\Psi}}_{\widetilde{\psi}_{k+1}\vee\eta_j}$ and $\mathcal{F}^{\widetilde{\Psi}}_{\widetilde{\psi}_k\vee\eta_{j+1}}$, and using the definition of $\boldsymbol{\gamma}^n_{j,k+1}(\widetilde{\Psi})$, $\boldsymbol{\gamma}^n_{j+1,k}(\widetilde{\Psi})$, we obtain

$$
\begin{aligned}
&(4.46)\\
&\geq \min\Bigg\{0, E\left[\int_{\psi_k^k\vee\eta_j}^{\widetilde{\psi}_{k+1}\wedge\eta_{j+1}} e^{-\lambda(t-\psi_k^k\vee\eta_j)}\left(\Phi_t^{\widetilde{\Psi}}-\frac{\lambda}{c}\right)dt\Big|\mathcal{F}^{\widetilde{\Psi}}_{\psi_k^k\vee\eta_j}\right]\\
&\quad+E\left[1_{\{\widetilde{\psi}_{k+1}<\eta_{j+1}\}}e^{-\lambda(\widetilde{\psi}_{k+1}-\psi_k^k\vee\eta_j)}\boldsymbol{\gamma}^n_{j,k+1}(\widetilde{\Psi})\Big|\mathcal{F}^{\widetilde{\Psi}}_{\psi_k^k\vee\eta_j}\right]\\
&\quad+E\left[1_{\{\eta_{j+1}\leq\widetilde{\psi}_{k+1}\}}e^{-\lambda(\eta_{j+1}-\psi_k^k\vee\eta_j)}\boldsymbol{\gamma}^n_{j+1,k}(\widetilde{\Psi})\Big|\mathcal{F}^{\widetilde{\Psi}}_{\psi_k^k\vee\eta_j}\right]\Bigg\},
\end{aligned}
$$

where we have used the fact that $\widetilde{\psi}_{k+1} \geq \eta_j$ (so $\widetilde{\psi}_{k+1} \vee \eta_j = \widetilde{\psi}_{k+1}$) for the second term, and the fact that we are on the set $\{\psi_k^k < \eta_{j+1}\}$ for the third term, so that $\psi_k^k \vee \eta_{j+1} = \eta_{j+1}$. Now, applying the induction hypothesis, this is greater than or equal to

$$\min\Bigg\{0, E\left[\int_{\psi_k^k \vee \eta_j}^{\widetilde{\psi}_{k+1} \wedge \eta_{j+1}} e^{-\lambda(t-\psi_k^k \vee \eta_j)} \left(\Phi_t^{\widetilde{\Psi}} - \frac{\lambda}{c}\right) dt \Big| \mathcal{F}^{\widetilde{\Psi}}_{\psi_k^k \vee \eta_j}\right]$$
$$+E\left[1_{\{\widetilde{\psi}_{k+1} < \eta_{j+1}\}} e^{-\lambda(\widetilde{\psi}_{k+1} - \psi_k^k \vee \eta_j)} \boldsymbol{v}^n_{j,k+1}\left(\widetilde{\psi}_{k+1} - \widetilde{\psi}_{k+1}, \Phi^{\widetilde{\Psi}}_{\widetilde{\psi}_{k+1}}\right) \Big| \mathcal{F}^{\widetilde{\Psi}}_{\psi_k^k \vee \eta_j}\right]$$
$$+E\left[1_{\{\eta_{j+1} \leq \widetilde{\psi}_{k+1}\}} e^{-\lambda(\eta_{j+1} - \psi_k^k \vee \eta_j)} \boldsymbol{v}^n_{j+1,k}\left(\psi_k^k \vee \eta_{j+1} - \psi_k^k, \Phi^{\widetilde{\Psi}}_{\psi_k^k \vee \eta_{j+1}}\right) \Big| \mathcal{F}^{\widetilde{\Psi}}_{\psi_k^k \vee \eta_j}\right]\Bigg\},$$

equal to

$$\min\Bigg\{0, E\left[\int_{\psi_k^k \vee \eta_j}^{\widetilde{\psi}_{k+1} \wedge \eta_{j+1}} e^{-\lambda(t-\psi_k^k \vee \eta_j)} \left(\Phi_t^{\widetilde{\Psi}} - \frac{\lambda}{c}\right) dt \Big| \mathcal{F}^{\widetilde{\Psi}}_{\psi_k^k \vee \eta_j}\right]$$
$$+E\left[1_{\{\widetilde{\psi}_{k+1} < \eta_{j+1}\}} e^{-\lambda(\widetilde{\psi}_{k+1} - \psi_k^k \vee \eta_j)} \boldsymbol{v}^n_{j,k+1}\left(0, \Phi^{\widetilde{\Psi}}_{\widetilde{\psi}_{k+1}}\right) \Big| \mathcal{F}^{\widetilde{\Psi}}_{\psi_k^k \vee \eta_j}\right]$$
$$+E\left[1_{\{\eta_{j+1} \leq \widetilde{\psi}_{k+1}\}} e^{-\lambda(\eta_{j+1} - \psi_k^k \vee \eta_j)} \boldsymbol{v}^n_{j+1,k}\left(\eta_{j+1} - \psi_k^k, \Phi^{\widetilde{\Psi}}_{\eta_{j+1}}\right) \Big| \mathcal{F}^{\widetilde{\Psi}}_{\psi_k^k \vee \eta_j}\right]\Bigg\}. \tag{4.47}$$

Now we make some observations. First, on the set $\{\psi_k^k < \eta_{j+1}\}$, $\eta_{j+1} - \psi_k^k \vee \eta_j$ is independent of $\mathcal{F}^{\widetilde{\Psi}}_{\psi_k^k \vee \eta_j}$, and conditioned on this sigma algebra, is distributed as an exponential random variable with parameter μ. Second, between $\psi_k^k \vee \eta_j$ and $\widetilde{\psi}_{k+1} \wedge \eta_{j+1}$, $\Phi^{\widetilde{\Psi}}$ has deterministic dynamics described by (4.1). Third, for some nonnegative $\mathcal{F}^{\widetilde{\Psi}}_{\psi_k^k \vee \eta_j}$-random variable $R_{j,k}$, $\widetilde{\psi}_{k+1} = \psi_k^k \vee \eta_j + R_{j,k}$, $1_{\{\widetilde{\psi}_{k+1} < \eta_{j+1}\}} = 1_{\{R_{j,k} < \eta_{j+1} - \psi_k^k \vee \eta_j\}}$, and $1_{\{\widetilde{\psi}_{k+1} \geq \eta_{j+1}\}} = 1_{\{R_{j,k} \geq \eta_{j+1} - \psi_k^k \vee \eta_j\}}$. Fourth, we note that, based upon (4.1) and arguing as in Proposition IV.4,

$$
\begin{aligned}
&E\left[\boldsymbol{v}^n_{j,k+1}\left(0,\Phi^{\widetilde{\Psi}}_{\widetilde{\psi}_{k+1}}\right)|\mathcal{F}^{\widetilde{\Psi}}_{\psi^k_k\vee\eta_j}\right]\\
&\quad=\boldsymbol{K}\boldsymbol{v}^n_{j,k+1}\left(\widetilde{\psi}_{k+1}-\psi^k_k,\Phi^{\Psi^k}_{\psi^k_k}\right)\\
&\quad=\boldsymbol{K}\boldsymbol{v}^n_{j,k+1}\left((\widetilde{\psi}_{k+1}-\psi^k_k\vee\eta_j)+(\psi^k_k\vee\eta_j-\psi^k_k),\Phi^{\Psi^k}_{\psi^k_k}\right)\\
&\quad=\boldsymbol{K}\boldsymbol{v}^n_{j,k+1}\left((\widetilde{\psi}_{k+1}-\psi^k_k\vee\eta_j)+(\psi^k_k\vee\eta_j-\psi^k_k),\varphi\left(-(\psi^k_k\vee\eta_j-\psi^k_k),\Phi^{\widetilde{\Psi}}_{\psi^k_k\vee\eta_j}\right)\right).
\end{aligned}
$$

Therefore,

$$
\begin{aligned}
&(4.47)\\
&\quad=\min\Bigg\{0,\int_0^\infty\mu e^{-\lambda u}\Bigg[\int_0^{R_{j,k}\wedge u}e^{-\lambda t}\left(\varphi\left(t,\Phi^{\widetilde{\Psi}}_{\psi^k_k\vee\eta_j}\right)-\frac{\lambda}{c}\right)dt\\
&\qquad+1_{\{R_{j,k}<u\}}e^{-\lambda R_{j,k}}\boldsymbol{K}\boldsymbol{v}^n_{j,k+1}\left(R_{j,k}+(\psi^k_k\vee\eta_j-\psi^k_k),\varphi\left(-(\psi^k_k\vee\eta_j-\psi^k_k),\Phi^{\widetilde{\Psi}}_{\psi^k_k\vee\eta_j}\right)\right)\\
&\qquad+1_{\{u\le R_{j,k}\}}e^{-\lambda u}\boldsymbol{v}^n_{j+1,k}\left(u+(\psi^k_k\vee\eta_j-\psi^k_k),\varphi\left(u,\Phi^{\widetilde{\Psi}}_{\psi^k_k\vee\eta_j}\right)\right)\Bigg]du\Bigg\},
\end{aligned}
$$

which is equal to

$$
\begin{aligned}
&\min\left\{0,J^+\left(\boldsymbol{v}^n_{j,k+1},\boldsymbol{v}^n_{j+1,k}\right)\left(R_{j,k},\psi^k_k\vee\eta_j-\psi^k_k,\Phi^{\Psi^k}_{\psi^k_k\vee\eta_j}\right)\right\}\\
&\quad\ge J^+_0\left(\boldsymbol{v}^n_{j,k+1},\boldsymbol{v}^n_{j+1,k}\right)\left(\psi^k_k\vee\eta_j-\psi^k_k,\Phi^{\Psi^k}_{\psi^k_k\vee\eta_j}\right)\\
&\quad=\boldsymbol{v}^n_{j,k}\left(\psi^k_k\vee\eta_j-\psi^k_k,\Phi^{\Psi^k}_{\psi^k_k\vee\eta_j}\right).
\end{aligned}
$$

We have shown that

$$
\begin{aligned}
&\operatorname*{ess\,inf}_{\tau\in\mathcal{T}^{\widetilde{\Psi}}_s,\tau\ge\psi^k_k\vee\eta_j}E\left[\int_{\psi^k_k\vee\eta_j}^{\tau}e^{-\lambda(t-\psi^k_k\vee\eta_j)}\left(\Phi^{\widetilde{\Psi}}_t-\frac{\lambda}{c}\right)dt\Big|\mathcal{F}^{\Psi^k}_{\psi^k_k\vee\eta_j}\right]\\
&\quad\ge\boldsymbol{v}^n_{j,k}\left(\psi^k_k\vee\eta_j-\psi^k_k,\Phi^{\Psi^k}_{\psi^k_k\vee\eta_j}\right)\ \text{on}\ \{\psi^k_k<\eta_{j+1}\}.
\end{aligned}
$$

Thus, after taking the infimum over all $\widetilde{\Psi}\in\boldsymbol{\mathfrak{O}}^n_{j,k}(\Psi^k)$, we deduce that

$$
\boldsymbol{\gamma}^n_{j,k}(\Psi^k)\ge\boldsymbol{v}^n_{j,k}\left(\psi^k_k\vee\eta_j-\psi^k_k,\Phi^{\Psi^k}_{\psi^k_k\vee\eta_j}\right)\ \text{on}\ \{\psi^k_k<\eta_{j+1}\},
$$

as claimed. $\square$

Proof of Proposition IV.24. As in Proposition IV.23, the proof is handled by a double backwards induction, first on the number of observation rights received j, and then on the number of observation rights spent, k. The proof of the base case, $j = n$, corresponds to times when all possible observation rights have been received, and so as before, this case is handled essentially identically as in Proposition IV.5. Therefore, we move on to the inductive step. Suppose that the result has been proven for $j+1$ observation rights received; we will prove it for j. Now comes a second induction, on k, so we will take up the base case of this, and take $k = j$. So, let $\Psi^j \in \mathfrak{O}^j$. We have

$$\begin{aligned}
(4.48)\ \ &\boldsymbol{\gamma}^n_{j,j}(\Psi^j) \\
&= \operatorname*{ess\,inf}_{\Psi \in \mathfrak{O}^n_{j,j}(\Psi^j)} \ \operatorname*{ess\,inf}_{\tau \in \mathcal{T}^{\Psi}_s, \tau \geq \psi^j_j} E\left[\int_{\psi^j_j}^{\tau} e^{-\lambda(t-\psi^j_j)}\left(\Phi^{\Psi}_t - \frac{\lambda}{c}\right) dt \Big| \mathcal{F}^{\Psi^j}_{\psi^j_j}\right] \\
(4.49)\ \ &\leq E\left[\int_{\psi^j_j}^{\widehat{\tau}^n_{j,j}\wedge\eta_{j+1}} e^{-\lambda(t-\psi^j_j)}\left(\varphi\left(s-\psi^j_j, \Phi^{\Psi^j}_{\psi^j_j}\right) - \frac{\lambda}{c}\right) ds \Big| \mathcal{F}^{\Psi^j}_{\psi^j_j}\right] \\
&\quad + \operatorname*{ess\,inf}_{\Psi \in \mathfrak{O}^n_{j+1,j}(\Psi^j)} \ \operatorname*{ess\,inf}_{\tau \in \mathcal{T}^{\Psi}_s, \tau \geq \eta_{j+1}} E\left[1_{\{\widehat{\tau}^n_{j,j} \geq \eta_{j+1}\}} \int_{\eta_{j+1}}^{\tau} e^{-\lambda(t-\psi^j_j)}\left(\Phi^{\Psi}_t - \frac{\lambda}{c}\right) dt \Big| \mathcal{F}^{\Psi^j}_{\psi^j_j}\right],
\end{aligned}$$

where we have used the following facts:

(a) The $j+1^{st}$ observation can not be made prior to the arrival time η_{j+1} of the $j+1^{st}$ arrival time, so Φ^{Ψ} evolves deterministically between ψ^j_j and η_j for all $\Psi \in \mathfrak{O}^n_{j,j}(\Psi^j)$.

(b) For any $\tau \in \mathcal{T}^{\Psi}_s$ with $\tau \geq \eta_{j+1}$, we may construct the stopping time $\widetilde{\tau} = \widehat{\tau}^n_{j,j} 1_{\{\widehat{\tau}^n_{j,j} < \eta_{j+1}\}} + \tau 1_{\{\widehat{\tau}^n_{j,j} \geq \eta_{j+1}\}}$ which agrees with τ on $\{\widehat{\tau}^n_{j,j} \geq \eta_{j+1}\}$, and is an element of $\mathcal{T}^{\Psi}_s$ such that $\widetilde{\tau} \geq \psi^j_j$. Therefore, the infimum over $\tau \geq \psi^j_j$ in (4.48) can be replaced with the infimum over $\tau \geq \eta_{j+1}$ in (4.49).

After conditioning the interior of $E\left[1_{\{\widehat{\tau}^n_{j,j} \geq \eta_{j+1}\}} \int_{\eta_{j+1}}^{\tau} e^{-\lambda(t-\psi^j_j)}\left(\Phi^{\Psi}_t - \frac{\lambda}{c}\right) dt \Big| \mathcal{F}^{\Psi^j}_{\psi^j_j}\right]$ on

$\mathcal{F}^{\Psi^j}_{\eta_{j+1}}$, we see that (4.49) is equal, by definition, to

$$E\left[\int_{\psi_j^j}^{\widehat{\tau}^n_{j,j}\wedge\eta_{j+1}} e^{-\lambda(t-\psi_j^j)}\left(\varphi\left(s-\psi_j^j,\Phi^{\Psi^j}_{\psi_j^j}\right)-\frac{\lambda}{c}\right)ds + 1_{\{\widehat{\tau}^n_{j,j}\geq\eta_{j+1}\}}\boldsymbol{\gamma}^n_{j+1,j}(\Psi^j)\big|\mathcal{F}^{\Psi^j}_{\psi_j^j}\right],$$

which by the induction hypothesis is equal to

$$E\Bigg[\int_{\psi_j^j}^{\widehat{\tau}^n_{j,j}\wedge\eta_{j+1}} e^{-\lambda(t-\psi_j^j)}\left(\varphi\left(s-\psi_j^j,\Phi^{\Psi^j}_{\psi_j^j}\right)-\frac{\lambda}{c}\right)ds$$
$$+1_{\{\widehat{\tau}^n_{j,j}\geq\eta_{j+1}\}}\boldsymbol{v}^n_{j+1,j}\left(\eta_{j+1}-\psi_j^j,\varphi\left(\eta_{j+1}-\psi_j^j,\Phi^{\Psi^j}_{\psi_j^j}\right)\right)\big|\mathcal{F}^{\Psi^j}_{\psi_j^j}\Bigg],$$

which is equal to

$$\int_0^\infty \mu e^{-\mu u}\Bigg[\int_0^{s^n_{j,j}\left(\Phi^{\Psi^j}_{\psi_j^j}\right)\wedge u} e^{-\lambda s}\left(\varphi\left(s,\Phi^{\Psi^j}_{\psi_j^j}\right)-\frac{\lambda}{c}\right)ds$$
$$+1_{\left\{s^n_{j,j}\left(\Phi^{\Psi^j}_{\psi_j^j}\right)\geq u\right\}}\boldsymbol{v}^n_{j+1,j}\left(u,\varphi\left(u,\Phi^{\Psi^j}_{\psi_j^j}\right)\right)\Bigg]du, \tag{4.50}$$

as $\widehat{\tau}^n_{j,j}=\psi_j^j+s^n_{j,j}\left(\Phi^{\Psi^j}_{\psi_j^j}\right)$, and using the fact that on the set $\{\psi_j^j<\eta_{j+1}\}$, $\eta_{j+1}-\psi_j^j$ is independent of $\mathcal{F}^{\Psi^j}_{\psi_j^j}$, and distributed as an exponential random variable with parameter μ.

Now, (4.50) is equal to $J^0\boldsymbol{v}^n_{j+1,j}\left(s^n_{j,j}\left(\Phi^{\Psi^j}_{\psi_j^j}\right),0,\Phi^{\Psi^j}_{\psi_j^j}\right)$, which by construction of $s^n_{j,j}$, is equal to $J^0_0\boldsymbol{v}^n_{j+1,j}\left(0,\Phi^{\Psi^j}_{\psi_j^j}\right)=\boldsymbol{v}^n_{j,j}\left(0,\Phi^{\Psi^j}_{\psi_j^j}\right)$.

Thus, we have established that $\boldsymbol{\gamma}^n_{j,j}(\Psi^j)\leq\boldsymbol{v}^n_{j,j}\left(0,\Phi^{\Psi^j}_{\psi_j^j}\right)$ on the set $\{\psi_j^j<\eta_{j+1}\}$. In light of Proposition IV.23, these quantities are actually equal. Furthermore, since $\boldsymbol{v}^n_{j,j}\left(0,\Phi^{\Psi^j}_{\psi_j^j}\right)\leq\boldsymbol{\gamma}^n_{j,j}(\Psi^j)\leq(4.48)=\boldsymbol{v}^n_{j,j}\left(0,\Phi^{\Psi^j}_{\psi_j^j}\right)$, (4.18) is established.

We now proceed to the second inductive step. Supposing that the result has been proven for j observation rights received and $k+1\leq j$ observation rights spent, we will prove the result for j observation rights received and k observation rights spent. So, let $\Psi^k\in\boldsymbol{\mathfrak{O}}^k$, and let $\widehat{\Psi}^{k+1}=\{\psi_1^k,\ldots,\psi_k^k,\widehat{\tau}^n_{j,k}\}\in\boldsymbol{\mathfrak{O}}^{k+1}_{j,k}(\Psi^k)$, with $\widehat{\tau}^n_{j,k}=\widehat{\tau}^n_{j,k}(\Psi^k)$. We have, on the set $\{\psi_k^k<\eta_{j+1}\}$,

$$
\begin{aligned}
(4.51)\quad & \boldsymbol{\gamma}^n_{j,k}(\Psi^k) \\
&= \operatorname*{ess\,inf}_{\Psi\in\mathfrak{O}^n_{j,k}(\Psi^k)} \operatorname*{ess\,inf}_{\tau\in\mathcal{T}^\Psi_s,\tau\geq\psi^k_k\vee\eta_j} E\left[\int_{\psi^k_k\vee\eta_j}^{\tau} e^{-\lambda(t-\psi^k_k\vee\eta_j)}\left(\Phi^\Psi_t-\frac{\lambda}{c}\right)dt\big|\mathcal{F}^{\Psi^k}_{\psi^k_k\vee\eta_j}\right] \\
&\leq E\left[\int_{\psi^k_k\vee\eta_j}^{\widehat{\tau}^n_{j,k}\wedge\eta_{j+1}} e^{-\lambda(s-\psi^k_k\vee\eta_j)}\left(\varphi\left(s-\psi^k_k\vee\eta_j,\Phi^{\Psi^k}_{\psi^k_k\vee\eta_j}\right)-\frac{\lambda}{c}\right)ds\big|\mathcal{F}^{\Psi^k}_{\psi^k_k\vee\eta_j}\right] \\
&\quad+\operatorname*{ess\,inf}_{\Psi\in\mathfrak{O}^n_{j,k+1}(\widehat{\Psi}^{k+1})} \operatorname*{ess\,inf}_{\tau\in\mathcal{T}^\Psi_s,\tau\geq\widehat{\tau}^n_{j,k}} E\left[e^{-\lambda(\widehat{\tau}^n_{j,k}-\psi^k_k\vee\eta_j)}1_{\{\widehat{\tau}^n_{j,k}<\eta_{j+1}\}}\int_{\widehat{\tau}^n_{j,k}}^{\tau} e^{-\lambda(t-\widehat{\tau}^n_{j,k})}\left(\Phi^\Psi_t-\frac{\lambda}{c}\right)dt\big|\mathcal{F}^{\Psi^k}_{\psi^k_k\vee\eta_j}\right] \\
&\quad+\operatorname*{ess\,inf}_{\Psi\in\mathfrak{O}^n_{j+1,k}(\Psi^k)} \operatorname*{ess\,inf}_{\tau\in\mathcal{T}^\Psi_s,\tau\geq\eta_{j+1}} E\left[e^{-\lambda(\eta_{j+1}-\psi^k_k\vee\eta_j)}1_{\{\widehat{\tau}^n_{j,k}\geq\eta_{j+1}\}}\int_{\eta_{j+1}}^{\tau} e^{-\lambda(t-\eta_{j+1})}\left(\Phi^\Psi_t-\frac{\lambda}{c}\right)dt\big|\mathcal{F}^{\Psi^k}_{\psi^k_k\vee\eta_j}\right],
\end{aligned}
$$

where we used in the first line above the deterministic evolution of Φ^{Ψ^k} in between observations, and for the second and third lines, the structure of stopping times in this problem, arguing as in (b) above. Now, (4.51) above is equal to

$$
\begin{aligned}
E\Bigg[&\int_{\psi^k_k\vee\eta_j}^{\widehat{\tau}^n_{j,k}\wedge\eta_{j+1}} e^{-\lambda(s-\psi^k_k\vee\eta_j)}\left(\varphi\left(s-\psi^k_k\vee\eta_j,\Phi^{\Psi^k}_{\psi^k_k\vee\eta_j}\right)-\frac{\lambda}{c}\right)ds \\
&+e^{-\lambda(\widehat{\tau}^n_{j,k}-\psi^k_k\vee\eta_j)}1_{\{\widehat{\tau}^n_{j,k}<\eta_{j+1}\}}\boldsymbol{\gamma}^n_{j,k+1}(\widehat{\Psi}^{k+1})+e^{-\lambda(\eta_{j+1}-\psi^k_k\vee\eta_j)}1_{\{\widehat{\tau}^n_{j,k}\geq\eta_{j+1}\}}\boldsymbol{\gamma}^n_{j+1,k}(\Psi^k)\big|\mathcal{F}^{\Psi^k}_{\psi^k_k\vee\eta_j}\Bigg],
\end{aligned}
$$

which by the induction hypotheses (first induction for the third term and second induction for the second term) is equal to

$$
\begin{aligned}
(4.52)\quad E\Bigg[&\int_{\psi^k_k\vee\eta_j}^{\widehat{\tau}^n_{j,k}\wedge\eta_{j+1}} e^{-\lambda(s-\psi^k_k\vee\eta_j)}\left(\varphi\left(s-\psi^k_k\vee\eta_j,\Phi^{\Psi^k}_{\psi^k_k\vee\eta_j}\right)-\frac{\lambda}{c}\right)ds \\
&+e^{-\lambda(\widehat{\tau}^n_{j,k}-\psi^k_k\vee\eta_j)}1_{\{\widehat{\tau}^n_{j,k}<\eta_{j+1}\}}\boldsymbol{v}^n_{j,k+1}\left(0,\Phi^{\widehat{\Psi}^{k+1}}_{\widehat{\tau}^n_{j,k}}\right) \\
&+e^{-\lambda(\eta_{j+1}-\psi^k_k\vee\eta_j)}1_{\{\widehat{\tau}^n_{j,k}\geq\eta_{j+1}\}}\boldsymbol{v}^n_{j+1,k}\Big((\eta_{j+1}-\psi^k_k\vee\eta_j) \\
&\quad+(\psi^k_k\vee\eta_j-\psi^k_k),\varphi\left(\eta_{j+1}-\psi^k_k\vee\eta_j,\Phi^{\Psi^k}_{\psi^k_k\vee\eta_j}\right)\Big)\big|\mathcal{F}^{\Psi^k}_{\psi^k_k\vee\eta_j}\Bigg],
\end{aligned}
$$

We will now argue as in (4.47) and the discussion following it. Using the fact that $\eta_{j+1}-\psi^k_k\vee\eta_j$ on the set $\{\psi^k_k<\eta_{j+1}\}$ is independent of $\mathcal{F}^{\Psi^k}_{\psi^k_k\vee\eta_j}$ and exponentially distributed with parameter μ, and noting $\widehat{\tau}^n_{j,k}-\psi^k_k\vee\eta_j=o^n_{j,k}=o^n_{j,k}\left(\psi^k_k\vee\eta_j-\psi^k_k,\Phi^{\Psi^k}_{\psi^k_k\vee\eta_j}\right)$, (4.52) becomes

$$\int_0^\infty \mu e^{-\mu u}\Bigg[\int_0^{o^n_{j,k}\wedge u} e^{-\lambda s}\left(\varphi\left(s,\Phi^{\Psi^k}_{\psi^k_k\vee\eta_j}\right)-\frac{\lambda}{c}\right)ds$$
$$+e^{-\lambda o^n_{j,k}}1_{\{o^n_{j,k}<u\}}\boldsymbol{K}\boldsymbol{v}^n_{j,k+1}\left(o^n_{j,k}+(\psi^k_k\vee\eta_j-\psi^k_k),\varphi\left(-(\psi^k_k\vee\eta_j-\psi^k_k),\Phi^{\Psi^k}_{\psi^k_k\vee\eta_j}\right)\right)$$
$$+e^{-\lambda u}1_{\{o^n_{j,k}\ge u\}}\boldsymbol{v}^n_{j+1,k}\left(u+(\psi^k_k\vee\eta_j-\psi^k_k),\varphi\left(u,\Phi^{\Psi^k}_{\psi^k\vee\eta_j}\right)\right)\Bigg]du,$$

which is equal to $J^+\left(\boldsymbol{v}^n_{j+1,k},\boldsymbol{v}^n_{j,k+1}\right)\left(o^n_{j,k}+\psi^k_k\vee\eta_j-\psi^k_k,\Phi^{\Psi^k}_{\psi^k_k\vee\eta_j}\right)$, which by construction of $o^n_{j,k}$, is equal to

$$J^+_0\left(\boldsymbol{v}^n_{j+1,k},\boldsymbol{v}^n_{j,k+1}\right)\left(\psi^k_k\vee\eta_j-\psi^k_k,\Phi^{\Psi^k}_{\psi^k_k\vee\eta_j}\right)=\boldsymbol{v}^n_{j,k}\left(\psi^k_k\vee\eta_j-\psi^k_k,\Phi^{\Psi^k}_{\psi^k_k\vee\eta_j}\right).$$

Thus, we have established that

$$\boldsymbol{\gamma}^n_{j,k}(\Psi^k)\le\boldsymbol{v}^n_{j,k}\left(\psi^k_k\vee\eta_j-\psi^k_k,\Phi^{\Psi^k}_{\psi^k_k\vee\eta_j}\right)$$

on the set $\{\psi^k_k<\eta_{j+1}\}$. Equality is now a consequence of Proposition IV.23. Examining the proof, we immediately deduce (4.19) as well. □

Appendix A Posterior Dyamics

In this Appendix, we derive the recursive formula (4.1). It will be convenient to assume that all observations occur at deterministic times, following [15]. The reduction to this case from observations at stopping times follows immediately from iteratively taking conditional expectations. All of the material in this appendix may be found in [15].

On some probability space $(\widetilde{\Omega},\widetilde{\mathcal{F}},P)$, suppose that $\widetilde{X}$ is a standard Brownian Motion which gains drift α at time $\widetilde{\Theta}$, where $P(\widetilde{\Theta}=0)=\pi$ and $P(\widetilde{\Theta}\in dt|\Theta>0)=\lambda e^{-\lambda t}dt$. Let $0=t_0<t_1<\cdots$ be a fixed infinite sequence of numbers, describing the times at which $\widetilde{X}$ is observed. Let

$$L_t(u,x_0,x_1,\ldots)\triangleq\prod_{l\ge1,t_l\le t}\frac{1}{\sqrt{2\pi(t_l-t_{l-1})}}t\exp\left\{\frac{[x_l-x_{l-1}-\alpha(t_l-t_{l-1}\vee u)^+]^2}{2(t_l-t_{l-1})}\right\}.$$

Then

$$P(\widetilde{X}_{t_l} \in dx_l \text{ for all } l \geq 1 \text{ s.t. } t_l \leq t) = L_t(\widetilde{\Theta}, x_0, x_1, \ldots) \prod_{l \geq 1, t_l \leq t} dx_l.$$

Therefore, the conditional likelihood of the observations $\widetilde{X}_{t_0}, \widetilde{X}_{t_1}, \ldots$, given $\widetilde{\Theta} = u$, is

$$\begin{aligned}
L_t(u) &\triangleq L_t(u, \widetilde{X}_{t_0}, \widetilde{X}_{t_1}, \ldots) \\
&= \prod_{l \geq 1, t_l \leq t} \frac{1}{\sqrt{2\pi(t_l - t_{l-1})}} t \exp\left\{ \frac{[\widetilde{X}_{t_l} - \widetilde{X}_{t_{l-1}} - \alpha(t_l - t_{l-1} \vee u)^+]^2}{2(t_l - t_{l-1})} \right\}.
\end{aligned}$$

To this point, we have already assumed that such a process $\widetilde{X}$ exists. The actual construction starts from a standard Brownian Motion X, and is then achieved via a change of measure. To that end, let $(\Omega, \mathcal{F}, P_\infty)$ support a standard Brownian Motion X and an independent random variable Θ with $P_\infty(\Theta = 0) = \pi$, and $P_\infty(\Theta \in dt | \Theta > 0) = \lambda e^{-\lambda t} dt$ for $t > 0$.

Then, $P_\infty \{X_{t_l} \in dx_l \text{ for all } l \geq 1, t_l \leq t\}$ is

$$L_t(\infty, x_0, x_1, \ldots) \prod_{l \geq 1, t_l \leq t} dx_l = \prod_{l \geq 1, t_l \leq t} \frac{1}{\sqrt{2\pi(t_l - t_{l-1})}} \exp\left\{ \frac{(x_l - x_{l-1})^2}{2(t_l - t_{l-1})} \right\} dx_l$$

for all $t \geq 0$. Let $\mathbb{F}$ be the filtration obtained by observing X at fixed times $0 = t_0 < t_1 < \cdots$, and let $\mathbb{G} = (\mathcal{G}_t)_{t \geq 0}$ be the augmentation of $\mathbb{F}$ by $\sigma(\Theta)$, so that $\mathcal{G}_t = \mathcal{F}_t \wedge \sigma(\Theta)$.

Define P on $\mathcal{G}_\infty$, locally along the filtration, by

$$\begin{aligned}
\frac{dP}{dP_\infty} &= Z_t(\Theta) \\
&\triangleq \frac{L_t(\Theta)}{L_t(\infty)} \\
&= \exp\left\{ \sum_{l=1}^{\infty} 1_{\{t_l \leq t\}} \left[\frac{(X_{t_l} - X_{t_{l-1}})\, \alpha\, (t_l - (\Theta \vee t_{l-1}))^+}{t_l - t_{l-1}} - \frac{\alpha^2 \left((t_l - (\Theta \vee t_{l-1}))^+\right)^2}{2(t_l - t_{l-1})} \right] \right\}.
\end{aligned}$$

Under P, the random variables $X_{t_l} - X_{t_{l-1}}$, $l \geq 1$, conditionally on Θ, are independent and Gaussian with mean $\alpha\,(t_l - (\Theta \vee t_{l-1}))^+$ and variance $t_l - t_{l-1}$.

Since $Z_0(\Theta) = 1$, P and P_∞ are identical on $\mathcal{G}_0 = \sigma(\Theta)$, so that Θ has the same distribution under P and P_∞. Under P, X has the distribution of a Brownian Motion which gains drift α at time Θ. We will work under P for the remainder of this Appendix.

Define the conditional odds process

$$\Phi_t \triangleq \frac{P\left(\Theta \le t|\mathcal{F}_t\right)}{P\left(\Theta > t|\mathcal{F}_t\right)} = \frac{E_\infty\left[Z_t(\Theta)1_{\{\Theta\le t\}}|\mathcal{F}_t\right]}{E_\infty\left[Z_t(\Theta)1_{\{\Theta> t\}}|\mathcal{F}_t\right]}, \tag{4.53}$$

with the second equality following from Bayes' Theorem. On the set $\{\Theta > t\}$, $(t_l - (\Theta \vee t_{l-1}))^+ = (t_l - \Theta)^+ = 0$ for all $l \ge 1, t_l \le t$. Therefore, $Z_t(\Theta)1_{\{\Theta>t\}} = 1_{\{\Theta>t\}}$.

Thus,

$$\begin{aligned} E_\infty\left[Z_t(\Theta)1_{\{\Theta>t\}}|\mathcal{F}_t\right] &= P_\infty\left(\Theta > t|\mathcal{F}_t\right) \\ &= P_\infty(\Theta > t) \\ &= (1-\pi)e^{-\lambda t}. \end{aligned}$$

So, (4.53) becomes

$$\frac{E_\infty\left[Z_t(\Theta)1_{\{\Theta\le t\}}|\mathcal{F}_t\right]}{(1-\pi)e^{-\lambda t}} = \frac{e^{\lambda t}}{1-\pi}E_\infty\left[Z_t(\Theta)1_{\{\Theta\le t\}}|\mathcal{F}_t\right]. \tag{4.54}$$

We will now focus on this last term in (4.54). Write

$$E_\infty\left[Z_t(\Theta)1_{\{\Theta\le t\}}|\mathcal{F}_t\right] = \pi Z_t(0) + (1-\pi)\int_0^t \lambda e^{-\lambda t} Z_t(u)du. \tag{4.55}$$

Suppose that $t_{n-1} \le t < t_n$ for some $n \ge 1$. By definition, $Z_t(u) = Z_{t_{n-1}}(u)$ for every $u \ge 0$, and $Z_{t_{n-1}}(u) = 1$ for every $t_{n-1} \le u < t_n$. This implies that (4.55) is

$$\begin{aligned} &\pi Z_{t_{n-1}}(0) + (1-\pi)\int_0^t \lambda e^{-\lambda u} Z_{t_{n-1}}(u)du \\ &\quad = \pi Z_{t_{n-1}}(0) + (1-\pi)\int_0^{t_{n-1}} \lambda e^{-\lambda u} Z_{t_{n-1}}(u)du + (1-\pi)\int_{t_{n-1}}^t \lambda e^{-\lambda u} Z_{t_{n-1}}(u)du \\ &\quad = \frac{1-\pi}{e^{\lambda t_{n-1}}}\Phi_{t_{n-1}} + (1-\pi)\left(e^{-\lambda t_{n-1}} - e^{-\lambda t}\right). \end{aligned}$$

From this, it follows that, for $t_{n-1} \leq t < t_n$, we have that (4.54) is equal to

$$e^{\lambda(t-t_{n-1})}\Phi_{t_{n-1}} + e^{\lambda(t-t_{n-1})} - 1 = \varphi\left(t - t_{n-1}, \Phi_{t_{n-1}}\right),$$

and this establishes the first part of (4.1). We now derive the form of Φ_{t_n}, conditionally on $\Phi_{t_{n-1}}$. Since $Z_{t_{n-1}}(u) = 1$ for $u \geq t_{n-1}$, we have

$$\begin{aligned} &Z_{t_n}(u) \\ &\quad = Z_{t_{n-1}}(u) \exp\left\{ \frac{\left(X_{t_n} - X_{t_{n-1}}\right)\alpha\left(t_n - (u \vee t_{n-1})\right)^+}{t_n - t_{n-1}} - \frac{\alpha^2\left(\left(t_n - (u \vee t_{n-1})\right)^+\right)^2}{2(t_n - t_{n-1})} \right\}, u \geq 0. \end{aligned}$$

So, (4.54) becomes

$$\begin{aligned} \frac{e^{\lambda t_n}}{1-\pi}\Bigg[&\left(\pi Z_{t_{n-1}}(0) + (1-\pi)\int_0^{t_{n-1}} \lambda e^{-\lambda u} Z_{t_{n-1}}(u)du\right) \exp\left\{\left(X_{t_n} - X_{t_{n-1}}\right)\alpha - \frac{\alpha^2}{2}(t_n - t_{n-1})\right\} \\ &+ (1-\pi)\int_{t_{n-1}}^{t_n} \lambda e^{-\lambda u} \underbrace{Z_{t_{n-1}}(u)}_{=1} \exp\left\{ \frac{\left(X_{t_n} - X_{t_{n-1}}\right)\alpha\left(t_n - u\right)}{t_n - t_{n-1}} - \frac{\alpha^2\left(t_n - u\right)^2}{2(t_n - t_{n-1})} \right\} du \Bigg], \end{aligned}$$

which is equal to

$$\begin{aligned} &\exp\left\{\left(X_{t_n} - X_{t_{n-1}}\right)\alpha - \frac{\alpha^2}{2}(t_n - t_{n-1})\right\} e^{\lambda(t_n - t_{n-1})}\Phi_{t_{n-1}} \\ &\qquad + \int_{t_{n-1}}^{t_n} \lambda e^{\lambda(t_n - u)} \exp\left\{ \frac{\left(X_{t_n} - X_{t_{n-1}}\right)\alpha\left(t_n - u\right)}{t_n - t_{n-1}} - \frac{\alpha^2\left(t_n - u\right)^2}{2(t_n - t_{n-1})} \right\} du. \end{aligned}$$

After making the substitution $w = -(t_n - u)$ for the integral above, we see that this is equal to the second term in (4.1).

Appendix B Source Code

Appendix B.1 Octave Code for the Lump Sum n-Observation Problem

There are four files in this section, used to calculate the value functions and observation boundaries associated to the lump sum n-observation problem. The first three are dependencies, and are used to respectively interpolate between grid points, define the function $j(t, z, \phi)$, and define the zero observation value function $v_0(\phi)$.

The last file is the main script. Parameters of the problem and total observations can be modified inside the file.

Listing IV.1: Interpolation Between Grid Points

```
%interpolates functions, sets them to zero above an
    upperbound


function [y] = interp(invec,outvec,upperphi,x)
  y = (x <= upperphi) .* interp1(invec, outvec, x);
endfunction
```

Listing IV.2: Definition of j

```
%calculates j(t,z,phi) with parameters la = lambda, al =
    alpha


function [y] = j(t,z,phi,la,al)
b = la .+ al.*z./sqrt(t);
c = al.^2/(2 .*t);
y = exp(la.*t .+ al.*z.*sqrt(t) .- al.^2.*t./2).*phi .+ la.*
    sqrt(pi).*exp(b.^2 ./(4 .*c))./(2.*sqrt(c)).*(erf((2.*c.*t
    .-b)./(2.*sqrt(c))) .+ erf(b./(2 .*sqrt(c))));
end
```

Listing IV.3: Definition of v_0

```
%analytic definition of the 0 observation value function


function [y] = v0(phi,la,c)
```

```
    y = (1/la*((phi+1)*log((la+c)/(c*(phi+1))) + phi - la/c)
        ) .* (phi < la/c) + ...
        (0                                                    ) .* (
          phi >= la/c);
end
```

Listing IV.4: Main File for Lump Sum n-Observation Problem

```
%main file for n observation problem:


%set console to always flush, so that status output shows
page_output_immediately(1)
page_screen_output(0)


%parameters
%required files: K.m, j.m, v0.m, interp.m,
la = .1; %lambda: exponential parameter of disorder time
c = .01; %relative delay cost
al = 1; %alpha: drift rate gained after disorder
tep = .2; %step size t for computing J_0 w(t,phi)
phep = .5; %step size for approximating value function
obs = 5; %number of total observations
%grid bounds
upperphi = 5.7*la/c; %value functions set to zero for phi
   above this
tgridsize = 30/tep + 1; %compute Jw(t,phi) over t in grid
```

```
tgridstart = 2;
phigridsize = upperphi/phep + 1; %optimize Jw(t,phi) over
    each phi in grid
phigrid = 0:phep:upperphi;
%Store results here: each column corresponds to a single
    iteration
value_function_data = zeros(phigridsize, obs); %records
    value of v1(phi)
times_data = zeros(phigridsize, obs); %records value of t^*(
    phi)
%run first iteration
for n =1:phigridsize %loop over phigrid
  phi = (n-1)*phep; %set value of phi in for loop
  disp("Starting phi...")
  disp(n)
  temptvec = zeros(1,tgridsize); %for storing Jw(t,phi) for
      t in tgrid
  temptvec(1) = v0(phi,la,c); %separate rule for calculating
        at t = 0
  for time = tgridstart:tgridsize
    t = (time-1)*tep; %set value of t in loop
    v0_comp = @(z) v0(j(t, z, phi, la, al), la, c).*exp(-z
        .^2/2)./sqrt(2 .*pi); %put v0 in a form for
        numerically integrating
```

```
  temptvec(time) = (phi + 1)*t + 1/la*(1+la/c)*(exp(-la*t)
     - 1) + exp(-la*t)*quad(v0_comp, -20, 20); %calculate
     Jw(t,phi)
end
value_function_data(n,1) = min(temptvec); %set value of
   value function
if (value_function_data(n,1) == 0) %value function
   increasing: stop after it hits 0
  break;
endif
times_data(n,1) = find(temptvec == min(temptvec)); %set
   optimal time
tgridsize = times_data(n,1) + 5;
disp("Observation")
disp(1)
disp("phi is")
disp(phi)
disp("value is")
disp(value_function_data(n,1))
disp("best time is")
disp(times_data(n,1))
if ( n > 1)
  tgridstart = max(times_data(n,1) - (times_data(n-1,1) -
     times_data(n,1)) - 5, 2); %adaptively configure t's to
      search based on previous run through
```

```
    endif
  end
  %run other iterations
  for iter = 2:obs
    cur_val = @(x) interp(phigrid, value_function_data(:,iter
        -1), upperphi, x); %value function iterpolated from
        pervious iteration
    tgridsize = 30/tep + 1;
    tgridstart = 2;
    for n = 1:phigridsize
      phi = (n-1)*phep;
      disp("Starting phi...")
      disp(n)
      temptvec = zeros(1,tgridsize);
      temptvec(1) = cur_val(phi);
      for time = tgridstart:tgridsize
        t = (time-1)*tep;
        cur_val_comp = @(z) cur_val(j(t, z, phi, la, al)).*exp
            (-z.^2/2)./sqrt(2 .*pi);
        temptvec(time) = (phi + 1)*t + 1/la*(1+la/c)*(exp(-la*
            t) - 1) + exp(-la*t)*quadcc(cur_val_comp, -8, 8);
      end
    value_function_data(n,iter) = min(temptvec);
    if (value_function_data(n,iter) == 0)
      break;
```

```
    endif
    times_data(n,iter) = find(temptvec == min(temptvec));
    disp("Observation")
    disp(iter)
    disp("phi is")
    disp(phi)
    disp("value is")
    disp(value_function_data(n,iter))
    disp("best time is")
    disp(times_data(n,iter))
    tgridsize = times_data(n,iter) + 3;
    if ( n > 1)
      tgridstart = max(times_data(n,iter) - (times_data(n-1,
        iter) - times_data(n,iter)) - 3, 2);
    endif
  end
end


save value_function_data.mat value_function_data
save times_data.mat times_data
```

Appendix B.2 Octave Code for the Stochastic Arrival Rate n-Observation Problem

In this section, there is a single file, the main algorithm for the stochastic arrival rate n-observation problem. It is only implemented for one observation right, and all files from the lump sum n-observation problem are dependencies.

Listing IV.5: Main File for Stochastic Arrival Rate Problem

```
%n-observation algorithm must be run first
qd_alg4
%parameters
mu = 1; %arrival rate for observations
obs = 1; %only implemented for one observation right
%store data here
st_value_function_data = zeros(obs + 1, obs + 1, phigridsize
   );
st_times_data = zeros(obs + 1, obs + 1, phigridsize);
%set data for v^1{1,1}(0,\phi)
for i = 1:phigridsize
  st_value_function_data(1, obs + 1, i) = v0((i-1)*phep, la,
      c);
end
%get data from n-observation problem
for i = 2:obs + 1
  st_value_function_data(i, obs + 1, :) =
     value_function_data(:, i - 1);
end
for i = 2:obs + 1
  st_times_data(i, obs + 1, :) = times_data(:, i - 1);
end
%define interpolated v^1_{1,1}(0,phi)
```

```
obsval = @(phi) interp(phigrid, st_value_function_data(1,
    obs + 1, :)(1:length(st_value_function_data(1, obs + 1, :)
    )), upperphi, phi);
%define Kv^1_{1,1}(t,phi)
obsval_comp = @(t, phi, z) obsval(j(t, z, phi, la, al)).*exp
    (-z.^2/2)./sqrt(2 .*pi);
%reset time grid
tgridsize = 30/tep + 1;
tgridstart = 2;
%compute v^1_{0,0}(0,phi) and t^{*,1}_{0,0}(phi)
for n = 1:phigridsize %same phigrid as n-observation problem
   phi = (n-1)*phep;
   disp("Starting phi...")
   disp(n)
   temptvec = zeros(1, tgridsize);
   temptvec(1) = 0; %the time "t" here is a stopping time,
      and not an observation time, hence taking t = 0 yields 0
       value (compare with the n-observation problem)
   t_star = st_times_data(2, 2, n)*tep; %recalling the form
      of v^1_{1,0}(y,\phi) from Step $2$ of the stochastic
      arrival algorithm, this is t^{*,1}_{1,0}(phi)
   for time = tgridstart:tgridsize
         t = (time - 1)*tep;
         disp(t)
```

```
        temptvec(time) = st_value_function_data(2, 2, n)*(1-
            exp(-mu*min(t,t_star))) + exp(-mu*t)*((phi + 1)*t
            + (1/la + 1/c)*(exp(-la*t) - 1)) + (t > t_star)*
            quad(@(u) mu*exp(-mu*u)*((phi + 1)*u + (1/la + 1/c
            )*(exp(-la*u) - 1) + exp(-la*u)*quadcc(@(z)
            obsval_comp(u, phi, z), -8, 8, .1)), t_star, t,
            .1); %implementing Step 2 of the stochastic
            arrival algorithm
        disp(temptvec(time))
        if temptvec(time) + .001 > temptvec(time - 1) %stop
            computing times once difference between
            consecutive times becomes small
          break;
        endif
end
st_value_function_data(2, 1, n) = min(temptvec);
if st_value_function_data(2, 1, n) == 0
  break;
endif
st_times_data(2, 1, n) = min(find(temptvec == min(temptvec
    )));
tgridsize = st_times_data(2, 1, n) + 3;
%heuristic to speed up implementation: minimize only over
    t's that are "close" to where the last minimizing t was
    found
```

```
  if n > 1
    tgridstart = max(st_times_data(2, 1, n) - (st_times_data
       (2, 1, n - 1) - st_times_data(2, 1, n)) - 10, 2);
  endif
end
```

5

Appendix B.3 R Code for Plotting the n-Observation Value Functions

This script loads the data calculated in Listing IV.4, and plots it using the ggplot2 library in R. Parameters inside the script should be set identical to the ones in Listing IV.4. R libraries “ggplot2” and “reshape” are also required.

Listing IV.6: Value Function Grapher

```
###Code for graphing the value function data.  After running
    , enter p1 for value function graphs and p for Bayesian
    risk graphs.


#required libraries
library(ggplot2)
library(reshape)


#parameters: must match those from matlab code


la <- .1
al <- 1
c <- .01
```

```
phep <- .02
obs =<- 5



upperphi = 5.7*la/c
phigridsize = upperphi/phep + 1
phigrid = seq(from=0, to = upperphi, by = phep)
#value function with no observations
v0 <- function(phi) {
  if (phi < lambda/c) {
    return(1/lambda*((phi + 1)*log((lambda + c)/(c*(phi + 1)
        )) + phi - lambda/c))
  }
  else return(0)
}
#for graphing v0
V0 <- rep(0,phigridsize)
for (i in 1:phigridsize) {
  V0[i] <- v0(phigrid[i])
}
#read in discrete observation data (valdat) and continuous
    observation (contda)
valdat <- read.table("C:/cygwin/home/ross.kravitz/valfun6_
    temp.mat")
```

```
contdat <- read.table("C:/cygwin/home/ross.kravitz/contdat6.
    mat")
#formatting data
contdat2 <- contdat[,2]
contdat2 <- c(-98.271,contdat2) #-98.271 is v_C(0), had to
    be supplied separately
valdat2 <- cbind(phigrid,V0,valdat,contdat2) #combined data
#set column names
colnames(valdat2)[1] <- "Odds Ratio of Disruption"
colnames(valdat2)[2:(obs+2)] <- c(0:obs)
colnames(valdat2)[obs+3] <- "Continuous"
#format data for ggplot2
df <- melt(valdat2, id='Odds Ratio of Disruption', variable_
    name='Number of Observations')
#plot all value functions
p1 <- ggplot(df, aes(`Odds Ratio of Disruption`,value)) +
    geom_line(aes(color=`Number of Observations`, group=`
    Number of Observations`))
p1 <- p1 + ylab("Value")

#plot Bayesian Risk formulation

altrep <- valdat2

#convert grid from phi to p
```

```
for (j in 1:phigridsize) {
  altrep[j,1] <- valdat2[j,1]/(1 + valdat2[j,1])
}
#convert value functions to Bayesian risk
for (i in 1:(obs+1)) {
  for (j in 1:phigridsize) {
    altrep[j,i+1] <- (1-altrep[j,1]) + (1-altrep[j,1])*c*
       valdat2[j,i+1]
  }
}

for (j in 1:phigridsize) {
  altrep[j,obs+3] <- (1-altrep[j,1]) + (1-altrep[j,1])*c*
     valdat2[j,obs+3]
}
#set column names
colnames(altrep)[1] <- "Prior Probability of Disruption"
colnames(altrep)[2:(obs+2)] <- c(0:obs)
colnames(altrep)[obs+3] <- "Continuous"
#format for ggplot2
df <- melt(altrep, id='Prior Probability of Disruption',
   variable_name='Number of Observations')
#plot all risk functions
p <- ggplot(df, aes(`Prior Probability of Disruption`,value)
   ) + geom_line(aes(color=`Number of Observations`, group=`
```

```
    Number of Observations'))
p <- p + ylab("Minimum Risk") + xlab("Prior Probability of 
    Disruption")
```

BIBLIOGRAPHY

BIBLIOGRAPHY

[1] David Aldous. Weak convergence and the general theory of processes. *Incomplete draft of a monograph. Available at http://www.stat.berkeley.edu/ aldous/Papers/weak-gtp.pdf*, 1981.

[2] Gordon R. Antelman and I. Richard Savage. Surveillance problems: Wiener processes. *Naval Res. Logist. Quart.*, 12:35–55, 1965.

[3] David W. Balmer. On a quickest detection problem with costly information. *J. Appl. Probability*, 12:87–97, 1975.

[4] Taposh Banerjee and Venugopal V. Veeravalli. Data-efficient quickest change detection in minimax settings. *arXiv preprint arXiv:1211.3729*, 2012.

[5] Taposh Banerjee and Venugopal V. Veeravalli. Data-efficient quickest change detection with on–off observation control. *Sequential Analysis*, 31(1):40–77, 2012.

[6] John A. Bather. An optimal stopping problem with costly information. In *Proceedings of the 39th Session of the International Statistical Institute (Vienna, 1973), Vol. 3*, volume 45, pages 9–24, 1973. With discussion by J. Th. Runnenburg and Peter Mandel and with reply by the author.

[7] Erhan Bayraktar and Savas Dayanik. Poisson disorder problem with exponential penalty for delay. *Math. Oper. Res.*, 31(2):217–233, 2006.

[8] Erhan Bayraktar, Savas Dayanik, and Ioannis Karatzas. The standard Poisson disorder problem revisited. *Stochastic Process. Appl.*, 115(9):1437–1450, 2005.

[9] Erhan Bayraktar, Savas Dayanik, and Ioannis Karatzas. Adaptive Poisson disorder problem. *Ann. Appl. Probab.*, 16(3):1190–1261, 2006.

[10] Erhan Bayraktar and H. Vincent Poor. Quickest detection of a minimum of two Poisson disorder times. *SIAM J. Control Optim.*, 46(1):308–331 (electronic), 2007.

[11] Martin Beibel. Sequential change-point detection in continuous time when the post-change drift is unknown. *Bernoulli*, 3(4):457–478, 1997.

[12] Martin Beibel. A note on sequential detection with exponential penalty for the delay. *Ann. Statist.*, 28(6):1696–1701, 2000.

[13] René Carmona and Savas Dayanik. Optimal multiple stopping of linear diffusions. *Math. Oper. Res.*, 33(2):446–460, 2008.

[14] René Carmona and Nizar Touzi. Optimal multiple stopping and valuation of swing options. *Math. Finance*, 18(2):239–268, 2008.

[15] Savas Dayanik. Wiener disorder problem with observations at fixed discrete time epochs. *Math. Oper. Res.*, 35(4):756–785, 2010.

[16] Savas Dayanik, H. Vincent Poor, and Semih O. Sezer. Sequential multi-hypothesis testing for compound Poisson processes. *Stochastics*, 80(1):19–50, 2008.

[17] Freddy Delbaen, Peter Grandits, Thorsten Rheinländer, Dominick Samperi, Martin Schweizer, and Christophe Stricker. Exponential hedging and entropic penalties. *Math. Finance*, 12(2):99–123, 2002.

[18] Freddy Delbaen, Pascale Monat, Walter Schachermayer, Martin Schweizer, and Christophe Stricker. Inégalités de normes avec poids et fermeture d'un espace d'intégrales stochastiques. *C. R. Acad. Sci. Paris Sér. I Math.*, 319(10):1079–1081, 1994.

[19] Freddy Delbaen, Pascale Monat, Walter Schachermayer, Martin Schweizer, and Christophe Stricker. Weighted norm inequalities and hedging in incomplete markets. *Finance and Stochastics*, 1:181–227, 1997.

[20] Claude Dellacherie and Paul-André Meyer. *Probabilities and potential. B*, volume 72 of *North-Holland Mathematics Studies*. North-Holland Publishing Co., Amsterdam, 1982. Theory of martingales, Translated from the French by J. P. Wilson.

[21] Catherine Doléans-Dade and Paul-André Meyer. Inégalités de normes avec poids. In *Séminaire de Probabilités, XIII (Univ. Strasbourg, Strasbourg, 1977/78)*, volume 721 of *Lecture Notes in Math.*, pages 313–331. Springer, Berlin, 1979.

[22] Christoph Frei. Convergence results for the indifference value based on the stability of bsdes. To appear in Stochastics, 2011.

[23] Pavel V. Gapeev and Goran Peskir. The Wiener disorder problem with finite horizon. *Stochastic Process. Appl.*, 116(12):1770–1791, 2006.

[24] Peter Grandits and Leszek Krawczyk. Closedness of some spaces of stochastic integrals. In *Séminaire de Probabilités, XXXII*, volume 1686 of *Lecture Notes in Math.*, pages 73–85. Springer, Berlin, 1998.

[25] Peter Grandits and Thorsten Rheinländer. On the minimal entropy martingale measure. *Ann. Probab.*, 30(3):1003–1038, 2002.

[26] Ying Hu, Peter Imkeller, and Matthias Müller. Utility maximization in incomplete markets. *Ann. Appl. Probab.*, 15(3):1691–1712, 2005.

[27] Nobuyuki Ikeda and Shinzo Watanabe. *Stochastic differential equations and diffusion processes*, volume 24 of *North-Holland Mathematical Library*. North-Holland Publishing Co., Amsterdam, second edition, 1989.

[28] Aytac Ilhan, Mattias Jonsson, and Ronnie Sircar. Portfolio optimization with derivatives and indifference pricing. In *Indifference Pricing: Theory and Applications*, Princeton series in financial engineering, pages 183–197. Princeton University Press, Princeton, 2009.

[29] 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.

[30] Ioannis Karatzas and Steven E. Shreve. *Brownian motion and stochastic calculus*, volume 113 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, second edition, 1991.

[31] Constantinos Kardaras and Gordan Žitković. Stability of the utility maximization problem with random endowment in incomplete markets. *Math. Finance*, 21(2):313–333, 2011.

[32] Norihiko Kazamaki. *Continuous exponential martingales and BMO*, volume 1579 of *Lecture Notes in Mathematics*. Springer-Verlag, Berlin, 1994.

[33] Eugene M. Klimko and James Yackel. Some aspects of search strategies for Wiener processes. In *Statistical decision theory and related topics (Proc. Sympos., Purdue Univ., Lafayette, Ind., 1970)*, pages 225–237. Academic Press, New York, 1971.

[34] Dmitry Kramkov and Walter Schachermayer. The asymptotic elasticity of utility functions and optimal investment in incomplete markets. *Ann. Appl. Probab.*, 9(3):904–950, 1999.

[35] Lifeng Lai, H. Vincent Poor, Yan Xin, and Georgios Georgiadis. Quickest search over multiple sequences. *IEEE Trans. Inform. Theory*, 57(8):5375–5386, 2011.

[36] Kasper Larsen and Gordan Žitković. Stability of utility-maximization in incomplete markets. *Stochastic Process. Appl.*, 117(11):1642–1662, 2007.

[37] Robert C. Merton. Optimum consumption and portfolio rules in a continuous-time model. *J. Econom. Theory*, 3(4):373–413, 1971.

[38] Marie-Amélie Morlais. Quadratic BSDEs driven by a continuous martingale and applications to the utility maximization problem. *Finance Stoch.*, 13(1):121–150, 2009.

[39] Sabrina Mulinacci and Maurizio Pratelli. Functional convergence of Snell envelopes: applications to American options approximations. *Finance Stoch.*, 2(3):311–327, 1998.

[40] Marcel Nutz. The opportunity process for optimal consumption and investment with power utility. *Math. Financ. Econ.*, 3(3-4):139–159, 2010.

[41] Marcel Nutz. Risk aversion asymptotics for power utility maximization. *Probab. Theory Related Fields*, 152(3-4):703–749, 2012.

[42] Bernt Øksendal. *Stochastic differential equations*. Universitext. Springer-Verlag, Berlin, sixth edition, 2003. An introduction with applications.

[43] Goran Peskir and Albert Shiryaev. *Optimal stopping and free-boundary problems*. Lectures in Mathematics ETH Zürich. Birkhäuser Verlag, Basel, 2006.

[44] Goran Peskir and Albert N. Shiryaev. Solving the Poisson disorder problem. In *Advances in finance and stochastics*, pages 295–312. Springer, Berlin, 2002.

[45] Edward Posner and Howard Rumsey, Jr. Continuous sequential decision in the presence of a finite number of hypotheses. *IEEE Trans. Inf. Theor.*, 12(2):248–255, September 2006.

[46] Philip E. Protter. *Stochastic integration and differential equations*, volume 21 of *Stochastic Modelling and Applied Probability*. Springer-Verlag, Berlin, 2005. Second edition. Version 2.1, Corrected third printing.

[47] Daniel Revuz and Marc Yor. *Continuous martingales and Brownian motion*, volume 293 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer-Verlag, Berlin, second edition, 1994.

[48] Chris Rogers and David Williams. *Diffusions, Markov processes, and martingales. Vol. 2*. Cambridge Mathematical Library. Cambridge University Press, Cambridge, 2000. Itô calculus, Reprint of the second (1994) edition.

[49] Walter Schachermayer. How potential investments may change the optimal portfolio for the exponential utility, 2000.

[50] Walter Schachermayer. Optimal investment in incomplete markets when wealth may become negative. *Ann. Appl. Probab.*, 11(3):694–734, 2001.

[51] Walter Schachermayer. A super-martingale property of the optimal portfolio process. *Finance Stoch.*, 7(4):433–456, 2003.

[52] Albert N. Shiryaev. On optimum methods in quickest detection problems. *Theory of Probability & Its Applications*, 8(1):22–46, 1963.

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

[54] Daniel H. Wagner. Survey of measurable selection theorems. *SIAM J. Control Optimization*, 15(5):859–903, 1977.

[55] Abraham Wald. Sequential tests of statistical hypotheses. *Ann. Math. Statistics*, 16:117–186, 1945.

[56] Hao Xing. Stability of the exponential utility maximization problem with respect to preferences. *Pre-print. Available at arXiv.org.*

[57] Hang Yu. *Horizon dependence of utility optimizers in incomplete models.* ProQuest LLC, Ann Arbor, MI, 2011. Thesis (Ph.D.)–Carnegie Mellon University.

[58] Qing Zhao and Jia Ye. Quickest detection in multiple on-off processes. *IEEE Trans. Signal Process.*, 58(12):5994–6006, 2010.

[59] Kamil Š. Zigangirov. On a problem of optimal scanning. *Teor. Verojatnost. i Primenen*, 11:333–338, 1966.

[60] Gordan Žitković. An example of a stochastic equilibrium with incomplete markets. To appear in Finance and Stochastics, 2011.