THESIS REPORT

Ph.D.

Bayesian Change Detection

by D.C. MacEnany
Advisor: J.S. Baras

Ph.D. 91-8

INSTITUTE FOR SYSTEMS RESEARCH

*Sponsored by
the National Science Foundation
Engineering Research Center Program,
the University of Maryland,
Harvard University,
and Industry*

BAYESIAN CHANGE DETECTION

by

David C. Mac Enany

Thesis submitted to the Faculty of the Graduate School of the University of Maryland in partial fulfillment of the degree of Doctor of Philosophy

1991

ABSTRACT

In this thesis we consider certain problems of optimal change detection in which the task is to decide in a sequential manner which of two probabilistic system descriptions account for given, observed data. Optimal decisions are defined according to an average cost criterion which has a penalty which increases with time and a penalty for incorrect decisions. We consider observation processes of both the diffusion and point process kind. A main result is a verification-type theorem which permits one to prove the optimality of candidate decision policies provided one can find a certain function and interval. The form of the theorem suggests how to go about looking for such a pair. As applications we consider the sequential detection and disruption problems involving diffusion observations and give new proofs of the existence of the optimal thresholds as well as a new, simple algorithm for their computation. In the case of sequential detection between Poisson processes we solve the so-called *overshoot problem* exactly for the first time using the same algorithm.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

In this thesis we consider certain problems of optimal change detection. The quintessential problem of change detection was first considered by A. Wald in his generalization of the Neyman–Pearson approach to binary hypothesis testing based on observations of a stochastic process. In the widely known Neyman-Pearson problem, one observes a stochastic process which has one of two distinct probabilistic descriptions and the task is to collect data for a fixed, finite amount of time, let's say T, and then decide which description models the data with the highest probability of detection for a given maximum probability of false alarm. The solution to this problem takes the form: at the fixed time T make a binary decision according to whether the likelihood ratio is above or below a threshold which is precomputed to satisfy the error constraints.

In the modification made by Wald to this basic problem, the task is to decide in a random, finite amount of time which description accounts for the data with a fixed probability of detection and a given maximum probability of false alarm [W]. His solution to this problem, the so-called sequential probability ratio test, was shown to not only satisfy the constraints on the detection and false alarm probabilities but also to yield an expected data collection time (averaging with respect to either hypothesis!) which is not longer than the expected time spent observing the data using any rival test which achieves the same probabilities of both types. Thus in the Wald formulation T becomes a nonnegative random variable and the solution takes the form: observe the likelihood ratio process in an interval with endpoints precomputed to satisfy the error constraints, stop collecting data at the random time T when the likelihood ratio exits this *continuation* interval and then make a binary decision according to whether the likelihood exits to the right or left.

Since Wald's highly original contribution, much research has focused on certain difficulties associated with it. One such difficulty is that oftentimes in practice, particularly in radar applications, one does not have the luxury of waiting a random amount of time prior to making a decision. Usually, there is a maximum length of time allowed for data collection and that's it. This leads to the consideration of truncated sequential tests. Another practical difficulty not fully appreciated in Wald's day is that by careful choice of a fixed T, the simpler Neyman–Pearson solution leads to performance which is often as good as the more complex *optimal* Wald solution. This situation is for the most part due to the statistical modeling uncertainties which are neglected in the optimization.

Another difficulty which has nagged researchers and engineers alike from the very beginning has been the overshoot problem. When the underlying likelihood ratio has continuous sample paths it can only exit the continuation interval in a continuous manner. Of course, the sample-path continuity of the likelihood derives from the continuity of the trajectories of the observed data. If the data come from a jump process then of course the likelihood ratio will also possess discontinuous sample paths and as a result it may exit the continuation interval by jumping across the boundary. This discontinuous behavior manifests itself in the optimization for the continuation interval in an exceedingly complicated way. Dealing with this complication has been given a name, the *overshoot problem.* Typical approaches to the overshoot problem involve special cases which are well approximated by diffusions. The favorite "solution" is to *neglect the excess over the boundary*, but of course neglecting the excess is no real solution. Other solutions seek to redefine the problem so that approximate solutions are optimal [K&P-K]. Nevertheless, it is safe to say that the overshoot problem has never been dealt with head-on and that prior to the results contained herein, a useful method to compute the optimal thresholds has never been given.

The Wald problem has also generated alot of research devoted to generalizations which it suggests. One such generalization is the notion of optimal stopping. As stated, the Wald formulation requires a solution consisting of two parts: a random stopping time and a random binary decision. His solution however, in essence reduces the problem to only finding the optimal stopping time in terms of the optimal exit interval, the actual decision therefore is a by-product of this recipe. Viewed as such, sequential detection is a problem of optimal stopping with a solution in the form of a *first exit policy.*

Much of the theory of optimal stopping concerns itself with such problems involving Markov processes for which optimal first exit policies exist. A particular example of a problem of this type is the disorder or *disruption problem* first considered by A. Shiryayev [S]. The disruption problem differs from the sequential detection problem in that the observed data are initially generated according to one probabilistic description up until some random time, and generated thereafter according to a second probabilistic model; the random time itself comes equipped with its own probabilistic description. In the sequential detection problem by contrast, the observed data are generated for all time according to one of the known statistical models. We see therefore that the disruption problem is inherently sequential, there is no rival Neyman–Pearson approach involving some fixed time T. Indeed, Shiryayev argues just this way in defense of the importance of sequential analysis techniques.

We shall consider a modest generalization to the disruption problem and call it the **change detection problem**. The generalization we make is to allow the disruption time to take values in $\Re \cup \{+\infty\}$ instead of just $\Re$. This modesty allows us to give sequential detection and disruption a unified treatment and to consider each as resulting from a specific choice of probabilistic model for the time of change. In particular, the probabilistic model for the time of change in the case of sequential detection is that of a random variable taking values in $\{0, +\infty\}$; in the disruption problem, a random variable with values in $[0, \infty)$. Viewed in this manner, one can argue that the problem of sequential detection is a kind of degenerate change detection problem. With this interpretation perhaps it is less unexpected that in practice the fixed-time Neyman–Pearson approach works as well as it does by comparison with sequential methods for the detection problem.

We have chosen the Bayesian approach in our consideration of these problems of change detection. The rationale behind this choice is not so much a manifestation of a personal philosophy regarding the proper modeling of prior knowledge, but merely a reflection of the fact that general results obtained within the Bayesian framework imply the analogous results in the Wald case by exploiting the extra degree of freedom in the costs to match the error probabilities [L]. Within this framework we model the observation processes using semimartingale methods in order to take advantage of powerful filtering results for semimartingales. Moreover, by using martingale methods we are able to treat diffusion-type processes and jump processes in a unified manner. Indeed, one of the triumphs of modern filtering theory, as pioneered by Snyder [Sn] and others, is that martingale methods permit one to handle filtering problems involving both types of noise processes in an abstract way using a *single* set of mathematical tools. With respect to problems of optimal stopping involving jump semimartingales however, this benefit has never been fully realized. Unfortunately, the triumphant abstraction obtained in filtering theory with the help of the abstract Girsanov theorem in no way dealt with the overshoot problem. As a result, when it came down to computing the optimal continuation interval or even proving its existence and uniqueness, one was left with *two* different sets of techniques and the same set of difficulties involving the *excess over the boundary*. In this thesis we outline a new approach which complements the abstractions obtained in filtering theory and allow one to give a unified treatment of the change detection problem for both types of semimartingales. We employ this approach to give an exact solution to the overshoot problem in the case of sequential detection between Poisson processes. We emphasize that the uniformity of viewpoint afforded by our methods extends all the way from the definition of the filtering problem through to the computation of the optimal thresholds.

A brief outline of the thesis is as follows. In Chapter 2 we consider the general theory of Bayesian change detection using martingale methods. The main result is the Verification Theorem which permits one to prove optimality results provided one can find a certain function and interval. The form of this theorem suggests how one should go about looking for such a pair. In Chapter 3 we consider the change detection problem for diffusion-type observations. By considering this more familiar case first we are able to point up the differences with respect to the less familiar point-process situation. In Chapter 4 we consider the point process case and solve the overshoot problem exactly for the first time in the case of sequential detection between Poisson processes. The thesis concludes with a consideration of future research directions.

Chapter 2

Bayesian Change Detection: General Theory

2.1 Introduction

In this chapter we consider some general probability theory underlying certain problems of optimal change detection. In this introduction we impart the feel of such problems by way of examples. The kinds of changes that we are concerned with detecting are ones in which a partially observed system undergoes a fundamental alteration in its dynamics. Loosely speaking, the prototypical problem can be described as follows. One is given the complete probabilistic descriptions of two stochastic dynamical systems. For instance, these two system models may be of the "signal+noise" or "noise-only" type found in classical Wald sequential detection. An essential ingredient is that each distinct system, say S_0 and S_1, comes equipped with one of two distinct probability measures, say P_0 and P_1. For $i = 0, 1$, P_i allows us to assess the chances of dynamical events involving S_i, possibly conditional upon knowledge of some of the 'output' of S_i. For 'i' given, computing probabilities associated with S_i amounts to simple conditional estimation theory. A wrinkle appears because although we are indeed given output data, we are not told that it is simply the output of S_0 or S_1—in other words we are not given 'i'. Instead we are given *another* probabilistic model describing how the output can change between S_0 and S_1. For instance, in Wald sequential detection we are told that the output for all time is definitely due to one and only one of S_0 or S_1. In the Bayesian approach to the same basic situation, we are also given a *prior* probability for, say, the second possibility. Supposing this prior probability to be $\pi \in [0, 1]$, we then construct a new probability measure, called Bayes' probability measure, as a convex combination of P_0 and P_1,

$$P_\pi\{\cdot\} := \pi\, P_1\{\cdot\} + (1 - \pi)\, P_0\{\cdot\}. \tag{2.1}$$

Equivalently, we may suppose that $P_\pi\{\cdot \mid S_i\} := P_i\{\cdot\}$ for $i = 0, 1$. For obvious reasons, the overall situation in the Bayes' sequential detection problem is usually modeled with a two-part formulation: first, a π-biased coin is flipped and then the flip determines which of S_0 or S_1 has its output gated to the observer.

As another example, consider the typical problem of disruption. We are given a nonnegative random variable modeling a random time and told that up until this random time the observed output is due to S_0 and thereafter due to S_1. Clearly, a coin-flipping formulation does not suffice to capture the basic ingredients of the disruption problem. The converse however is true. Indeed, if we suppose that the nonnegative random variable representing the disruption time takes values on $[0, \infty]$ and employ a measure which gives zero probability to $(0, \infty)$, then the disruption time formulation essentially mimics a coin-flipping situation in which heads is $\{0\}$ and tails is $\{+\infty\}$. In this thesis we shall exploit this converse to give a uniform formulation for both types of change detection problems. In other words, we shall view the sequential detection problem as a special case of a slightly more general form of disruption problem which we call **change detection**.

So far we have outlined in broad strokes the probabilistic set-up of the observed and unobserved dynamics in the basic problem of simple, binary Bayesian change detection. The remainder of the problem set-up concerns how the output data can be used to make decisions concerning which of S_0 or S_1 is responsible for the output at the time the decision is made, and also, how the data can be used to determine when to stop and make this decision. Not surprisingly, this two-part decision-making process "stopping/deciding" leads us to consider a two-part cost function called Bayes' cost for which a minimum is sought. In the type of problems we consider the solution to this minimization problem yields an on-line *decision policy* which is simple to implement and achieves the minimum average penalty amongst all admissible policies.

The remainder of the chapter has the following outline:

Section 1. PROBABILISTIC FRAMEWORK

In this section we describe the basic set-up for the prior probability measures and system dynamics filtrations. In this set-up we impose only two conditions on the family of prior measures—the deeper structure of this family is not further elaborated upon until Section 7. We also make preliminary assumptions concerning the nature of the random disruption time, we define Bayes' cost function and characterize the set of admissible decision policies. We finish with a definition of Bayes' optimality and an aside to show that levying penalties for correct decisions is a complication resulting in no greater generality.

Section 2. OPTIMAL STOPPING

The main point of this section is to show how the two-parameter optimization of Section 1 can be replaced by a single parameter optimization over the admissible stopping times alone. In so doing we introduce Π, the *a posteriori* probability process.

Section 3. CONCAVE RUNNING COST

The purpose here is to impose an additional, modest assumption upon the structure of Bayes' cost. This results in a simple proof that the optimal Bayes' cost is a concave function of the prior probability, an extremely useful fact which is fully exploited in Section 5 to help demonstrate the Bayes' optimality of a policy of the type described in the Section 4.

Section 4. FIRST EXIT POLICIES

In this section we define a special class of decision policies which are based on the time at which the *a posteriori* probability process first exits an interval. Such policies are useful because they are easy to implement in practice and also yield a tractable analysis in the abstract.

Section 5. SUFFICIENCY AND VERIFICATION

This section is concerned with posing a set of conditions involving a function and an interval which are sufficient to demonstrate the Bayes' optimality of the first exit policy based on this interval and also to uniquely characterize Bayes' optimal cost computationally.

Section 6. A LIKELIHOOD RATIO

In this section we further specialize our probabilistic models. We also define a Radon-Nikodym derivative, consider its conditioning upon a certain stopped σ–algebra, and deal with its reciprocal.

Section 7. THE PRIOR MEASURES: A MODEL

Here we fill in the remainder of the general description of the family of Bayes' measures, $\{P_\pi : 0 \leq \pi \leq 1\}$. Indeed, we make assumptions concerning the properties of P_0 and P_1 which are consistent with our modeling intentions and then define the entire family in terms of P_0 and P_1 alone.

Section 8. MARTINGALE DYNAMICS OF Π

This final section deals with the consequences of the assumptions posed in Section 7 and their bearing upon the semimartingale description of the *a posteriori* probability process.

2.2 Probabilistic Framework

On a measurable space $(\Omega, \mathcal{A})$ we are given a family of probability measures $\{P_\pi : 0 \leq \pi \leq 1\}$ and two right-continuous filtrations of $\mathcal{A}$, $\{\mathcal{A}_t\}_{t\geq 0}$ and $\{\mathcal{O}_t\}_{t\geq 0}$, both $(\mathcal{A}, P_\pi)$–completed for all $t \geq 0$ and all $\pi \in [0,1]$. With $\mathcal{A}_0$ denoting the $(\mathcal{A}, P_\pi)$–completed trivial σ–algebra on $\mathcal{A}$, we define $\mathcal{O}_0 := \mathcal{A}_0$ and assume that,

$$\mathcal{O}_t \subset \mathcal{A}_t \qquad \forall\, t \geq 0. \tag{2.2}$$

The smaller filtration will serve to model the observable consequences of unobservable events whose unfolding will in turn be modeled within the abstraction of a larger filtration. For reasons that will become clear later P_π is called Bayes' measure on $(\Omega, \mathcal{A})$ with prior π in $[0,1]$. Let $\mathcal{O} = \bigvee_{t\geq 0} \mathcal{O}_t$. We make the following assumption concerning the family of Bayes' measures,

$$\text{(A1):} \qquad P_\pi\{O\} = \lambda\, P_{\pi'}\{O\} + (1-\lambda)\, P_{\pi''}\{O\} \qquad \forall\, O \in \mathcal{O},$$

whenever λ, π, π', π'' *in* $[0,1]$ *are related by,*

$$\pi = \lambda\, \pi' + (1-\lambda)\, \pi''.$$

This linearity of the family of Bayes' measures in the prior index will be central to our investigations in this chapter.

With this partial characterization of the structure of Bayes' probability measures in hand, we also suppose we are given some $\mathcal{A}_t$–stopping time υ (*upsilon*) called the **disruption time** and assume that,

$$(A2): \quad P_\pi\{\upsilon = 0\} = \pi \qquad \forall\, \pi \in [0,1].$$

We define a point process $\Upsilon = \{\Upsilon_t\}_{t\geq 0}$ (cap-*upsilon*) in terms of the disruption time as,

$$\Upsilon_t := 1\{\upsilon \leq t\} \quad \forall t \geq 0, \tag{2.3}$$

so that Υ is a binary point process which is nondecreasing, right-continuous, and has a single jump from $\{0\}$ to $\{1\}$ at the random time υ. In the context of the general sequential binary decision problem, it is assumed that the single-jump point process Υ is not directly observable but that one can observe some $\mathcal{O}_t$–adapted semimartingale with a compensator depending on Υ as follows. For all times less than υ this observable semimartingale is modeled as having an $\mathcal{A}_t$–predictable compensator whose statistics are described by the probability measure P_0, and thereafter by P_1. The role of Υ in this set-up is to act as a 'switch process', switching (only once) between the two alternate compensator models. One of our first goals will be to formulate a reasonable method to decide for each instant of time if the switch has occurred, the decision based solely upon information in $\mathcal{O}_t$ and made according to some reasonable cost criterion. Within this formulation, Υ is said to be $\mathcal{O}_t$-partially observable. The following proposition concerns the semimartingale nature of Υ.

Proposition 2.1 *The binary single-jump $\mathcal{A}_t$–adapted point process Υ has the semimartingale representation*

$$\Upsilon_t = \Upsilon_0 + K_t + M_t \qquad \forall t \geq 0,$$

where K is its $(\mathcal{A}_t, P_\pi)$–predictable, integrable compensator having sample paths of locally finite variation which are increasing with P_π–probability one and M is an $(\mathcal{A}_t, P_\pi)$–martingale. Moreover, K is P_π–integrable.

Proof: We first show that $\Upsilon - \Upsilon_0$ is a class D, corlol $(\mathcal{A}_t, P_\pi)$–submartingale for then all but the P_π–integrability of K will follow from a celebrated theorem due to P.A. Meyer [W&H, p225]. The corlol property follows immediately from the definition of Υ. Next note,

$$E_\pi[\Upsilon_t | \mathcal{A}_s] \geq E_\pi[\Upsilon_s | \mathcal{A}_s] = \Upsilon_s \qquad 0 \leq s \leq t, \tag{2.4}$$

because $s \leq t$ implies $\Upsilon_t \geq \Upsilon_s$, P_π–a.s. Hence $\Upsilon - \Upsilon_0$ is an $(\mathcal{A}_t, P_\pi)$–submartingale. To show that $\Upsilon - \Upsilon_0$ is class D we need to show that the collection of random variables $\mathcal{L}$,

$$\mathcal{L} := \{\Upsilon_\tau - \Upsilon_0 \,:\, \tau \text{ is an } \mathcal{A}_t\text{–stopping time}\}, \tag{2.5}$$

is uniformly P_π–integrable. This follows trivially because for all $X \in \mathcal{L}$ $|X|$ is bounded, but for completeness, recall, $\mathcal{L}$ is a uniformly P_π-integrable collection of random variables if,

$$\lim_{c\to\infty} \sup_{X\in\mathcal{L}} E_\pi[|X| 1\{|X| \geq c\}] = 0. \tag{2.6}$$

For any $X \in \mathcal{L}$ note that $|X| \leq 1$ so that,

$$E_\pi[|X| 1\{|X| \geq c\}] \leq E_\pi[1\{|X| \geq c\}] \leq 1\{1 \geq c\} \quad \forall c \geq 0. \tag{2.7}$$

Hence,

$$0 \leq \lim_{c\to\infty} \sup_{X\in\mathcal{L}} E_\pi[|X| 1\{|X| \geq c\}] \leq \lim_{c\to\infty} 1\{1 \geq c\} = 0. \tag{2.8}$$

Thus, $\Upsilon - \Upsilon_0$ is class D.

To show that $K = \{K_t\}_{t\geq 0}$ is P_π–integrable note that,

$$E_\pi K_t = E_\pi[\Upsilon_t - \Upsilon_0] \leq 1 \quad \forall t \geq 0. \tag{2.9}$$

Because $0 = K_0 \leq K_n \leq K_{n+1}$ we see that there exists a random variable K_∞ such that,

$$K_n \longrightarrow K_\infty \quad P_\pi\text{–a.s.} \tag{2.10}$$

Therefore using the Monotone Convergence Theorem we have,

$$E_\pi K_n \longrightarrow E_\pi K_\infty \tag{2.11}$$

so that $E_\pi K_\infty \leq 1$, i.e., K is P_π–integrable. □

In the task of attempting to determine if Υ has jumped—which is equivalent to deciding if the P_1–governed compensator is currently modulating the observed data—consider the set of **decision policies**: all pairs of random variables of the form (τ, δ), where τ is some $\mathcal{O}_t$–stopping time, written $\tau \in T(\mathcal{O})$, and δ is some $\mathcal{O}_\tau$–measurable binary random variable, written $\delta \in B(\mathcal{O}_\tau)$. One may think of δ as a guess or decision made at time τ about the value of Υ_τ based upon the information content of $\mathcal{O}_\tau$. The decision time τ itself may be considered as a kind of estimate of the jump time v of the Υ process.

Over the set of decision policies define Bayes' cost function

$$\rho_\pi(\tau, \delta) := E_\pi[\int_0^\tau \mathcal{C}_s\, ds + \mathcal{E}(\Upsilon_\tau, \delta)], \tag{2.12}$$

where $\mathcal{C} = \{\mathcal{C}_t\}_{t \geq 0}$ is some nonnegative $\mathcal{O}_t$–adapted process satisfying,

$$\begin{aligned} &(\text{C1}): \quad E_\pi \int_0^t \mathcal{C}_s\, ds < \infty \quad \forall t \geq 0, \quad \forall \pi \in [0,1]; \\ &(\text{C2}): \quad P_\pi\{\int_0^\infty \mathcal{C}_s\, ds = \infty\} = 1, \quad \forall \pi \in [0,1], \end{aligned}$$

and where,

$$\mathcal{E}(x,y) := \begin{cases} c^0 & \text{if } x = 1 \;\&\; y = 0; \\ 0 & \text{if } x = y; \\ c^1 & \text{if } x = 0 \;\&\; y = 1, \end{cases} \tag{2.13}$$

with $0 < c^0, c^1 \leq \infty$ not both infinite. For a given decision policy (τ, δ), we call the integral portion of Bayes' cost, $E_\pi[\int_0^\tau C_s\, ds]$, the expected running cost of the policy and interpret it loosely as the average cost of not making a decision until the random time τ. Condition (C1) requiring the running cost to be integrable for all possible priors works to ensure that there is no fixed, deterministic time before which any policy must make a decision. Condition (C2), on the other hand, essentially guarantees that decision policies having stopping times which are not P_π–a.s. finite have unbounded cost. The portion $E_\pi[\mathcal{E}(\Upsilon_\tau, \delta)]$ is called the expected decision cost of the policy with the obvious interpretation.

At this point we can succinctly state that the main goal of this chapter is to give conditions which are sufficient to guarantee the existence and uniqueness of a decision policy (τ_*, δ_*) which minimizes Bayes' cost over the set of all decision policies. From the definition of Bayes' cost we see that if we define the simple, deterministic policy $(\tilde{\tau}, \tilde{\delta})$ as, $\tilde{\tau} \equiv 0$ and $\tilde{\delta} = 1\{c^1 \leq c^0\}$, we get

$$\rho_\pi(\tilde{\tau}, \tilde{\delta}) = \min\{c^0, c^1\} < \infty. \tag{2.14}$$

Because our goal is to minimize Bayes' cost, we obviously desire a policy which does no worse than $(\tilde{\tau}, \tilde{\delta})$. With this motivation, we make the following definition.

Definition 2.1 *Let $\mathcal{T}_{ad}$ denote the subset of $\mathcal{O}_t$–stopping times,*

$$\mathcal{T}_{ad} := \{\, \tau \in T(\mathcal{O}) \; : \; E_\pi \int_0^\tau \mathcal{C}_s \, ds < \infty \;\; \forall \pi \in [0,1] \,\}.$$

We call $\mathcal{T}_{ad}$ the set of **admissible** *stopping times. Let Δ_τ denote the subset of $\mathcal{O}_\tau$–measurable binary random variables,*

$$\Delta_\tau := \{\, \delta \in B(\mathcal{O}_\tau) \; : \; E_\pi[\, \mathcal{E}(\Upsilon_\tau, \delta)\,] < \infty \;\; \forall \pi \in [0,1] \,\}.$$

Any decision policy, (τ, δ), is said to be admissible if $\tau \in \mathcal{T}_{ad}$ and δ is in Δ_τ. We also define the set of P_π–a.s. finite stopping times,

$$\mathcal{T} := \{\, \tau \in T(\mathcal{O}) \; : \; P_\pi\{\, \tau < \infty\} = 1 \;\; \forall \pi \in [0,1] \,\}.$$

□

With $\mathcal{T}$ and $\mathcal{T}_{ad}$ so defined, we see condition (C1) guarantees that $\mathcal{T}_{ad}$ is not empty while (C2) implies that $\mathcal{T}_{ad} \subset \mathcal{T}$. It is also not too difficult to see that $\mathcal{T}_{ad}$ is a convex set. We can now give a definition of optimality.

Definition 2.2 *An admissible policy, (τ_*, δ_*), is said to be* **Bayesian optimal** *if*

$$\rho_\pi(\tau_*, \delta_*) = \inf_{(\tau,\delta)} \rho_\pi(\tau, \delta) \qquad \forall \pi \in [0,1],$$

where the infimum is over all admissible policies. We define

$$\rho(\pi) := \inf_{(\tau,\delta)} \rho_\pi(\tau, \delta),$$

and call it the **Bayes' optimal cost.** □

From this definition it becomes clear that in assuming c^0 and c^1 are both nonzero we have removed a trivial case. For if $c^0\,c^1 = 0$ then it follows easily that $\rho \equiv 0$ since the degenerate policy (τ_*, δ_*) with $\tau_* \equiv 0$ and $\delta_* := 1\{c^1 = 0\}$ is seen to yield the optimal Bayes' cost. The definition also makes clear that it is without loss of generality that no cost is levied for correct decisions in 2.13, i.e., as long as we agree that the cost of a correct decision is less than the cost of either type of incorrect decision. The reason for this is that any such problem can be converted into one having a cost of the form 2.13. To see this write $\mathcal{E}(\Upsilon_\tau, \delta) = \mathcal{E}(\Upsilon_\tau, \delta; c^0, c^1)$, emphasizing the dependence of $\mathcal{E}(\Upsilon_\tau, \delta)$ on c^0 and c^1 and then define $\overline{\mathcal{E}}(c^0, c^1) := \mathcal{E}(\Upsilon_\tau, \delta; c^0, c^1)$, deemphasizing the dependence on Υ_τ and δ. Let $\tilde{c} > 0$ denote a cost for correct decisions and compute,

$$\begin{aligned} E_\pi[\tilde{c}\,1\{\delta = \Upsilon_\tau\}] &= E_\pi[\tilde{c}\,\delta\,\Upsilon_\tau + \tilde{c}\,(1-\delta)\,(1-\Upsilon_\tau)] \\ &= \tilde{c} - E_\pi[\tilde{c}\,(1-\delta)\,\Upsilon_\tau + \tilde{c}\,\delta\,(1-\Upsilon_\tau)] \\ &= \tilde{c} - E_\pi[\overline{\mathcal{E}}(\tilde{c}, \tilde{c})], \end{aligned} \tag{2.15}$$

and this gives us,

$$\begin{aligned} E_\pi[\mathcal{E}(\Upsilon_\tau, \delta; c^0, c^1) + \tilde{c}\,1\{\delta = \Upsilon_\tau\}] &= \tilde{c} + E_\pi[\overline{\mathcal{E}}(c^0, c^1) - \overline{\mathcal{E}}(\tilde{c}, \tilde{c})] \\ &= \tilde{c} + E_\pi[\overline{\mathcal{E}}(\tilde{c}^0, \tilde{c}^1)] \\ &= \tilde{c} + E_\pi[\mathcal{E}(\Upsilon_\tau, \delta; \tilde{c}^0, \tilde{c}^1)], \end{aligned} \tag{2.16}$$

where we have defined $\tilde{c}^i = c^i - \tilde{c}$ for $i = 0, 1$. Now, since we indeed assume in this thesis that the penalty for correct decisions is less than the penalty for incorrect decisions—a not unreasonable assumption—we have $\tilde{c} < \min\{c^0, c^1\}$ and therefore $0 < \tilde{c}^i$ for $i = 0, 1$. As a result, the final expectation in 2.16 has the exact same form as $E_\pi[\mathcal{E}(\Upsilon_\tau, \delta; c^0, c^1)]$; in other words, a minimization of 2.16 over (τ, δ) admissible does not explicitly involve $\tilde{c}$ and is of the form which our methods are intended to address.

2.3 Optimal Stopping

In our effort to simplify the minimization of Bayes' cost, we begin by defining the $(\mathcal{O}_t, P_\pi)$-conditional probability of the event $\{\Upsilon_t = 1\}$ as

$$\Pi_t := P_\pi\{\Upsilon_t = 1 \,|\, \mathcal{O}_t\} \qquad \forall\, t \geq 0. \tag{2.17}$$

Note that $P_\pi\{\Pi_0 = \pi\} = 1$ in view of assumption (A2); for all later times, a probabilistic description of the dynamical structure of Π is in part characterized by the following.

Proposition 2.2 *The conditional probability process, $\Pi = \{\Pi_t\}_{t\geq 0}$, has the semimartingale representation*

$$\Pi_t = \Pi_0 + \hat{K}_t + \overline{M}_t \qquad \forall\, t \geq 0,$$

where $\hat{K}$ denotes its unique $(\mathcal{O}_t, P_\pi)$-predictable compensator having sample paths of locally finite variation which are increasing with P_π-probability one and $\overline{M}$ is an $(\mathcal{O}_t, P_\pi)$-martingale. Moreover, $\hat{K}$ is P_π-integrable.

Proof: The proof of this proposition is the same as the proof of Proposition 2.1 with the obvious notational changes. □

From the definition of Π note that we may also write

$$\Pi_t = E_\pi[\Upsilon_t | \mathcal{O}_t] \qquad \forall\, t \geq 0, \tag{2.18}$$

so that Π_t is the projection of Υ_t onto $\mathcal{O}_t$ and therefore we have $\hat{K}_t = E_\pi[K_t | \mathcal{O}_t]$ using the Projection Theorem [W&H, P7.1.3].

We next give a theorem which greatly simplifies the minimization used in Definition 2.2 to define Bayes' optimal cost. The theorem makes explicit use of the Π process to reduce the two-dimensional minimization over all decision policies (τ, δ) as posed in Definition 2.2 to a one-dimensional minimization over all τ admissible, a considerable and welcome simplification.

Theorem 2.1 (Optimal Stopping) *For $\tau \in \mathcal{T}_{ad}$ define,*

$$\rho_\pi(\tau) := E_\pi\left[\int_0^\tau \mathcal{C}_s\, ds + e(\Pi_\tau)\right],$$

with

$$e(\pi) := \min\{c^0\pi, c^1(1-\pi)\}.$$

Then,

$$\inf_{(\tau,\delta)} \rho_\pi(\tau, \delta) = \inf_{\tau \in \mathcal{T}_{ad}} \rho_\pi(\tau),$$

where the infimum on the left is over all admissible policies.

Proof: Let (τ, δ) denote some arbitrary admissible policy. Consider the policy (τ, δ_*) obtained from (τ, δ) by replacing δ with $\delta_* = \delta_*(\Pi_\tau)$ given by

$$\delta_* := \begin{cases} 1 & \text{if } \Pi_\tau \geq \pi_e; \\ 0 & \text{if } \Pi_\tau \leq \pi_e, \end{cases} \tag{2.19}$$

where $\pi_e := \frac{c^1}{c^0+c^1}$; note that $\pi_e = \arg\max\{e(\pi) : 0 \leq \pi \leq 1\}$. We will show that Bayes' cost for (τ, δ) is not less than that for (τ, δ_*). Using 2.13 we have,

$$\mathcal{E}(\Upsilon_\tau, \delta) = c^0\, 1\{\delta = 0\} 1\{\Upsilon_\tau = 1\} + c^1\, 1\{\delta = 1\} 1\{\Upsilon_\tau = 0\} \qquad P_\pi\text{–a.s.} \tag{2.20}$$

By the definition of admissibility we know $\{\delta = i\} \in \mathcal{O}_\tau$ for $i = 0, 1$ and hence,

$$\begin{aligned} E_\pi[1\{\delta = 0\}\, 1\{\Upsilon_\tau = 1\}] &= E_\pi[1\{\delta = 0\}\, E_\pi[1\{\Upsilon_\tau = 1\} | \mathcal{O}_\tau]] \\ &= E_\pi[1\{\delta = 0\}\, \Pi_\tau] \\ &= E_\pi[(1 - \delta)\, \Pi_\tau], \end{aligned} \tag{2.21}$$

and similarly,

$$E_\pi[1\{\delta = 1\}\,1\{\Upsilon_\tau = 0\}] = E_\pi[\delta(1 - \Pi_\tau)]. \tag{2.22}$$

Thus,

$$\begin{aligned} E_\pi[\mathcal{E}(\Upsilon_\tau, \delta)] &= E_\pi\left[c^0(1-\delta)\Pi_\tau + c^1\delta(1-\Pi_\tau)\right] \\ &\geq E_\pi\left[\min\{c^0\Pi_\tau, c^1(1-\Pi_\tau)\}\right] \\ &= E_\pi[e(\Pi_\tau)], \end{aligned} \tag{2.23}$$

so that for all admissible pairs (τ, δ),

$$\rho_\pi(\tau, \delta) \geq E_\pi\left[\int_0^\tau \mathcal{C}_s\, ds + e(\Pi_\tau)\right] = \rho_\pi(\tau). \tag{2.24}$$

On the other hand, δ_* is $\mathcal{O}_\tau$–measurable as is clear from 2.19 and so,

$$\begin{aligned} \rho_\pi(\tau, \delta_*) &= E_\pi\left[\int_0^\tau \mathcal{C}_s\, ds + \mathcal{E}(\Upsilon_\tau, \delta_*)\right] \\ &= E_\pi\left[\int_0^\tau \mathcal{C}_s\, ds + c^0\,\Pi_\tau\,(1-\delta_*) + c^1\,(1-\Pi_\tau)\,\delta_*\right] \\ &= E_\pi\left[\int_0^\tau \mathcal{C}_s\, ds + e(\Pi_\tau)\right] \\ &= \rho_\pi(\tau). \end{aligned} \tag{2.25}$$

Hence,

$$\rho_\pi(\tau, \delta) \geq \rho_\pi(\tau, \delta_*), \tag{2.26}$$

for all admissible pairs (τ, δ). This implies

$$\inf_{\tau \in \mathcal{T}_{ad}} \rho_\pi(\tau) = \inf_{(\tau,\delta)} \rho_\pi(\tau, \delta), \tag{2.27}$$

where the infimum on the right is over all admissible policies. This completes the proof. □

As a result, the search for an optimal policy is reduced to a search for an optimal stopping time. Note that according to Definition 2.2 we have shown

$$\rho(\pi) = \inf_{\tau \in \mathcal{T}_{ad}} \rho_\pi(\tau), \tag{2.28}$$

and so we shall henceforth call $\rho_\pi(\tau)$ Bayes' cost corresponding to $\tau \in \mathcal{T}_{ad}$. Another consequence of the Optimal Stopping Theorem is that it shifts the focus of our attention from the point process Υ to the conditional probability process Π and it suggests that the properties of the Π process deserve closer investigation.

2.4 Concave Running Cost

The following proposition describes a useful property enjoyed by Π which it inherits directly from condition (A1) imposed on P_π, or, more honestly, this desirable property strongly inspires our definition of the P_π measure. This property immediately suggests that an additional assumption be imposed on the running cost.

Proposition 2.3 *Suppose π, π', $\pi'' \in [0,1)$ are in the relation*

$$\pi = \lambda\,\pi' + (1-\lambda)\,\pi'' \quad \textit{for some } \lambda \in [0,1]. \tag{2.29}$$

Define $\Pi'_\tau := E_{\pi'}[\Upsilon_\tau \,|\, \mathcal{O}_\tau]$ and by analogy Π''_τ for $\tau \in \mathcal{T}_{ad}$. Let Λ' denote the Radon-Nikodym derivative of $P_{\pi'}$ with respect to P_π when both are restricted to $\mathcal{O}$ and let Λ'' denote the analogous derivative. Then for all $\tau \in \mathcal{T}_{ad}$ we have,

$$\Pi_\tau = \lambda\,\Lambda'\,\Pi'_\tau + (1-\lambda)\,\Lambda''\,\Pi''_\tau \qquad P_\pi\text{-}a.s.$$

Proof: From the assumption (A1) regarding the family of Bayes measures we know that the restriction of P_π to $\mathcal{O}$ satisfies,

$$P_\pi\{O\} = \lambda\,P_{\pi'}\{O\} + (1-\lambda)\,P_{\pi''}\{O\} \qquad \forall\, O \in \mathcal{O}. \tag{2.30}$$

and clearly $P_{\pi'}$ and $P_{\pi''}$ are absolutely continuous with respect to P_π on $\mathcal{O}$. Let $O \in \mathcal{O}_\tau$ and compute,

$$\int_O E_\pi[\Upsilon_\tau \,|\, \mathcal{O}_\tau]\,dP_\pi \;=\; \int_O \Upsilon_\tau\,dP_\pi$$

$$
\begin{aligned}
&= \lambda \int_O \Upsilon_\tau \, dP_{\pi'} + (1-\lambda) \int_O \Upsilon_\tau \, dP_{\pi''} \\
&= \lambda \int_O E_{\pi'}[\Upsilon_\tau | \mathcal{O}_\tau] \, dP_{\pi'} + (1-\lambda) \int_O E_{\pi''}[\Upsilon_\tau | \mathcal{O}_\tau] \, dP_{\pi''} \\
&= \lambda \int_O \Pi'_\tau \, dP_{\pi'} + (1-\lambda) \int_O \Pi''_\tau \, dP_{\pi''} \\
&= \int_O [\lambda \, \Lambda' \, \Pi'_\tau + (1-\lambda) \, \Lambda'' \Pi''_\tau] \, dP_\pi, \qquad (2.31)
\end{aligned}
$$

i.e., for any $\tau \in \mathcal{T}_{ad}$,

$$
\Pi_\tau = E_\pi[\Upsilon_\tau | \mathcal{O}_\tau] = \lambda \, \Lambda' \, \Pi'_\tau + (1-\lambda) \, \Lambda'' \, \Pi''_\tau, \qquad (2.32)
$$

except possibly on $\mathcal{O}_\tau$–sets of P_π–measure zero. This gives us the result. □

This property of the *a posteriori* probability suggests that by placing an additional, modest assumption on the running cost we can still further simplify the nature of the infimum involved in the search for the optimal stopping time beyond that which we have achieved so far. Moreover, the suggested assumption by no means trivializes the kinds of running costs which can be considered. Thus, before we address further the computation of the infimum in 2.28, we impose the following assumption upon the running cost. This assumption will be sufficient to imply that ρ is a concave function on $(0,1)$. In particular we assume that the $\mathcal{O}_t$–adapted process, $\mathcal{C} = \{\mathcal{C}_t\}_{t \geq 0}$, does not depend on the prior $\pi \in [0,1]$ and is of the form,

$$
\mathcal{C}_t := \mathcal{C}_t(\Pi_t) \qquad \forall \, t \geq 0. \qquad (2.33)
$$

With this understanding we place a third restriction on $\mathcal{C}$:

$$
(\mathrm{C3}): \quad \mathcal{C}(\Pi) \text{ is concave in } \Pi,
$$

i.e., given any $\pi, \pi', \pi'', \lambda \in [0,1]$ such that $\pi = \lambda\,\pi' + (1-\lambda)\,\pi''$, (C3) is taken to mean,

$$\mathcal{C}_t(\pi) \geq \lambda\,\mathcal{C}_t(\pi') + (1-\lambda)\,\mathcal{C}_t(\pi'') \qquad \forall\, t \geq 0,\ P_\pi\text{–a.s.} \tag{2.34}$$

The next proposition demonstrates that the concavity of both $\mathcal{C}$ and e implies the concavity of ρ.

Proposition 2.4 *Under condition* (C3) *the Bayes' optimal cost, ρ, is concave on* $(0,1)$.

Proof: Suppose that π, π', $\pi'' \in [0,1)$ are in the relation $\pi = \lambda\,\pi' + (1-\lambda)\,\pi''$ for some $\lambda \in [0,1]$. In Proposition 2.3 we showed,

$$\Pi_\tau = \lambda\,\Lambda'\,\Pi'_\tau + (1-\lambda)\,\Lambda''\,\Pi''_\tau \qquad P_\pi\text{–a.s.},\ \forall\,\tau \in \mathcal{T}_{ad}, \tag{2.35}$$

which says that Π is a convex combination of $\Lambda'\,\Pi'$ and $\Lambda''\,\Pi''$. However, from assumption (A1) regarding P_π we find,

$$\lambda\,\Lambda' + (1-\lambda)\,\Lambda'' = 1 \qquad P_\pi\text{–a.s.}, \tag{2.36}$$

so that Π is in fact *also* a (stochastic) convex combination of Π' and Π''. Thus, because the mapping e is concave we get,

$$e(\Pi_t) \geq \lambda\,\Lambda'\,e(\Pi'_t) + (1-\lambda)\,\Lambda''\,e(\Pi''_t) \quad \forall\, t \geq 0, \quad P_\pi\text{–a.s.} \tag{2.37}$$

Hence,

$$\begin{aligned} E_\pi[e(\Pi_t)] &\geq \lambda\,E_\pi[\Lambda'\,e(\Pi'_t)] + (1-\lambda)\,E_\pi[\Lambda''\,e(\Pi''_t)] \\ &= \lambda\,E_{\pi'}[e(\Pi'_t)] + (1-\lambda)\,E_{\pi''}[e(\Pi''_t)]. \end{aligned} \tag{2.38}$$

Likewise, the concavity of $\mathcal{C}(\Pi)$ as postulated in (C3) implies,

$$E_\pi \int_0^t \mathcal{C}_s(\Pi_s)\, ds \geq \lambda\, E_{\pi'} \int_0^t \mathcal{C}_s(\Pi'_s)\, ds + (1-\lambda)\, E_{\pi''} \int_0^t \mathcal{C}_s(\Pi''_s)\, ds. \qquad (2.39)$$

Combining these results we have,

$$\rho_\pi(t) \geq \lambda\, \rho_{\pi'}(t) + (1-\lambda)\, \rho_{\pi''}(t) \qquad \forall\, t \geq 0, \qquad (2.40)$$

and therefore,

$$\rho_\pi(\tau) \geq \lambda\, \rho_{\pi'}(\tau) + (1-\lambda)\, \rho_{\pi''}(\tau) \qquad \forall\, \tau \in \mathcal{T}_{ad}. \qquad (2.41)$$

Finally, we see that,

$$\begin{aligned} \rho(\pi) = \inf_{\tau\in\mathcal{T}_{ad}} \rho_\pi(\tau) &\geq \lambda \inf_{\tau\in\mathcal{T}_{ad}} \rho_{\pi'}(\tau) + (1-\lambda) \inf_{\tau\in\mathcal{T}_{ad}} \rho_{\pi''}(\tau) \\ &= \lambda\, \rho(\pi') + (1-\lambda)\, \rho(\pi''), \end{aligned} \qquad (2.42)$$

in other words, ρ, being an infimum of a family of concave functions is necessarily concave. Of course, a simple corollary to the concavity of ρ on $(0,1)$ is its continuity there. □

For the remainder of this thesis the general problem with which we shall concern ourselves can now be stated rigorously and concisely as:

$$(\mathcal{P}): \qquad \text{Find } \tau_* \in \mathcal{T}_{ad} \text{ such that } \rho(\pi) = \inf_{\tau\in\mathcal{T}_{ad}} \rho_\pi(\tau) = \rho_\pi(\tau_*).$$

Although we know that ρ exists, is nonnegative and finite, and is concave on $(0,1)$, we do not know if a solution to $(\mathcal{P})$ exists or if it possesses a unique solution. We know a fair amount about the conditional probability process Π and have used it to simplify the minimization in the definition of ρ. In the next section we continue to exploit the properties of Π to still further simplify the minimization in $(\mathcal{P})$ when certain sufficient conditions are met. In particular, we introduce the notion of *first exit policy* based on the notion of *first exit time*, a random time at which Π first exits some open interval. This leads us to consider the whereabouts of Π at the time of first exit. Using these notions, we can give a set of conditions involving a particular interval, I_*, which if it exists and satisfies these conditions is sufficient to imply that the first exit policy based on I_* is optimal in that it solves problem $(\mathcal{P})$. These constraints also serve to characterize a mapping $r_* : [0,1] \to \Re$ which permits one to compute Bayes' optimal cost provided the mapping exists and it too meets the set of constraints.

2.5 First Exit Policies

The next step towards our goal is to transform the minimization problem described by $(\mathcal{P})$ into one still more manageable. In attempting to do this, it becomes clear that a most important subclass of admissible stopping times are those which are first exit times of the Π process from an interval. This is due both to their simple specification and remarkable optimality properties.

Definition 2.3 *Define the collections of* **continuation intervals**,

$$\begin{aligned} \mathcal{I}^0 \quad &:= \{\,[0,b) \,:\, 0 < b < 1\,\}; \\ \mathcal{I}^+ \quad &:= \{\,(a,b) \,:\, 0 < a < b < 1\,\}, \end{aligned}$$

and $\mathcal{I} := \mathcal{I}^0 \cup \mathcal{I}^+$. *The* **first exit time** *of* Π *from* $\mathrm{I} \in \mathcal{I}$ *is an* $\mathcal{O}_t$*-stopping time defined as*

$$\tau^{\mathrm{I}} = \inf\{t \geq 0 : \; \Pi_t \notin \mathrm{I}\}.$$

We denote by $\overline{\mathcal{T}}$ *the collection of all such first exit times; we observe that* τ^{I} *is not necessarily an admissible stopping time. For each* $\mathrm{I} \in \mathcal{I}$ *we call* $(\tau^{\mathrm{I}}, \delta_*(\Pi_{\tau^{\mathrm{I}}}))$ *the* **first exit policy** *based on* I. □

We emphasize the fact the the collection of continuation intervals $\mathcal{I}$ lacks symmetry because $\pi = 1$ is always an absorbing boundary for the Π process whereas $\pi = 0$ may or may not be. Indeed,

$$\Pi_t \,|_{\pi=1} = P_1\{\, v \leq t \,|\, \mathcal{O}_t\,\} = P_1\{\, 0 \leq t \,|\, \mathcal{O}_t\,\} = 1 \quad \forall\, t \geq 0. \tag{2.43}$$

Let $\mathrm{I} \in \mathcal{I}$ and $\pi \in \mathrm{I}$. If $\tau^{\mathrm{I}} \in \mathcal{T}$ then from the definition of $\mathcal{T}$,

$$P_\pi\{\tau^{\mathrm{I}} < \infty\} = 1 \qquad \forall\, \pi \in \mathrm{I}, \tag{2.44}$$

and this means that Π eventually exits I, P_π–a.s. For instance, if the Π process possesses continuous sample paths then $P_\pi\{\,\Pi_{\tau^{\rm I}} \in \partial\,{\rm I}\,\} = 1$ where $\partial{\rm I}$ denotes the boundary of I, i.e., its endpoints. In general however, the Π process may possess jump discontinuities and as a result there is little or no chance that Π lies on the boundary of I at the time of escape. Nevertheless, when Π exits an interval, questions naturally arise as to its whereabouts at the time of escape. The following definition provides a means to phrase such questions.

Definition 2.4 *Let* ${\rm I} \in \mathcal{I}$ *and* $\Omega_{\rm I} := \{\Pi_0 \in {\rm I}\}$. *The* Π–**boundary** *of* I *is defined as,*

$$\partial_\Pi\,{\rm I} := \bigcup_{\omega\in\Omega_{\rm I}} \{\,\Pi_{\tau^{\rm I}}\,\}.$$

In words, $\partial_\Pi\,{\rm I}$ *comprises all the point values that* Π *may take on when it exits* I. *The* Π–**closure** *of* I *is defined as,*

$$[\,{\rm I}\,]_\Pi := {\rm I} \cup \partial\,{\rm I} \cup \partial_\Pi\,{\rm I}.$$

It will also be convenient to define the upper and lower Π*-boundaries of* I *as* $\partial^+_\Pi{\rm I}$ *and* $\partial^-_\Pi{\rm I}$*, respectively, i.e.,* $\partial_\Pi{\rm I} = \partial^-_\Pi{\rm I} \cup \partial^+_\Pi{\rm I}$, $\partial^-_\Pi{\rm I} \cap \partial^+_\Pi{\rm I} = \emptyset$ *and* $\pi^+ \in \partial^+_\Pi{\rm I}$ *implies* $\pi^+ > \pi$ *for all* $\pi \in \partial^-_\Pi{\rm I}$*, and vice versa.* □

Our plan for solving the minimization over $\mathcal{T}_{ad}$ in 2.28 is to characterize an interval ${\rm I}_* \in \mathcal{I}$ such that for all $\pi \in [0,1]$,

$$\rho_\pi(\tau^{{\rm I}_*}) \le \rho_\pi(\tau) \qquad \forall\,\tau \in \mathcal{T}_{ad}. \tag{2.45}$$

At first glance, 2.45 does not say that $\tau^{{\rm I}_*}$ solves problem $\mathcal{P}$ because we also need to show that $\tau^{{\rm I}_*}$ is admissible. However, given any $\tau \in \mathcal{T}_{ad}$ a second glance at 2.45 shows,

$$E_\pi \int_0^{\tau^{\mathrm{I}*}} \mathcal{C}_s \, ds \le \rho_\pi(\tau^{\mathrm{I}*}) \le \rho_\pi(\tau) < \infty \quad \forall \pi \in [0,1], \tag{2.46}$$

and therefore $\tau^{\mathrm{I}*} \in \mathcal{T}_{ad}$.

The following simple proposition characterizes Bayes' cost for a first exit policy based on a continuation interval I in $\mathcal{I}$ for any prior not in I.

Proposition 2.5 *Let* $\mathrm{I} \in \mathcal{I}$. *Then,*

$$\rho_\pi(\tau^{\mathrm{I}}) = e(\pi) \quad \forall \pi \notin \mathrm{I}.$$

Proof: If π is not in I then $\tau^{\mathrm{I}} = 0$, P_π–a.s. Hence,

$$\begin{aligned} \rho_\pi(\tau^{\mathrm{I}}) &= E_\pi\Big[\int_0^{\tau^{\mathrm{I}}} \mathcal{C}_s \, ds + e(\Pi_{\tau^{\mathrm{I}}}) \Big] \\ &= E_\pi[\, 0 + e(\Pi_0) \,] = e(\pi), \end{aligned} \tag{2.47}$$

since $P_\pi\{\, \Pi_0 = \pi \,\} = 1$. □

We are also obviously interested in dealing with the quantity $\rho_\pi(\tau^{\mathrm{I}})$ for $\mathrm{I} \in \mathcal{I}$ when $\pi \in \mathrm{I}$. In order to do this in some generality we need to impose the following **escape** condition on the Π process:

(E)

Either: (E^0) $\quad P_\pi\{\tau^{[0,b)} < \infty\} = 1 \quad \forall \pi \in [0,b),\ b < 1;$

Or: (E^+) $\quad P_\pi\{\tau^{(a,b)} < \infty\} = 1 \quad \forall \pi \in (a,b),\ 0 < a < b < 1.$

With economy of notation in mind define I_∞ via,

$$\mathrm{I}_\infty := \begin{cases} [0,1) & \text{if } (\mathrm{E}^0) \text{ holds;} \\ (0,1) & \text{if } (\mathrm{E}^+) \text{ holds,} \end{cases} \tag{2.48}$$

and let $\mathcal{I}_\infty$ denote

$$\mathcal{I}_\infty := \begin{cases} \mathcal{I}^0 & \text{if } (\mathrm{E}^0) \text{ holds;} \\ \mathcal{I}^+ & \text{if } (\mathrm{E}^+) \text{ holds.} \end{cases} \tag{2.49}$$

Note that $\mathrm{I}_\infty \notin \mathcal{I}_\infty$ since $\mathrm{I}_\infty \notin \mathcal{I}$, however, I_∞ can be expressed as a limit of intervals all in $\mathcal{I}_\infty$. Furthermore, given any increasing sequence of intervals $\{\mathrm{I}_n\}_{n\geq 1}$ in $\mathcal{I}_\infty$ such that $\cup_{n\geq 1}\mathrm{I}_n = \mathrm{I}_\infty$, then $\tau^{\mathrm{I}_n} \nearrow \tau^{\mathrm{I}_\infty}$, P_π–a.s. for all $\pi \in \mathrm{I}_\infty$. Using this definition of $\mathcal{I}_\infty$ the escape condition can be rewritten more compactly as,

$$\text{(E)} \qquad P_\pi\{\tau^{\mathrm{I}} < \infty\} = 1 \quad \forall\, \pi \in \mathrm{I},\, \forall\, \mathrm{I} \in \mathcal{I}_\infty.$$

In terms of our earlier notation the escape condition says that $\tau^{\mathrm{I}} \in \mathcal{T}$ for all $\mathrm{I} \in \mathcal{I}_\infty$, i.e, the Π process is guaranteed to escape any admissible interval. In other words, the *absorbing* points of the Π process are *attracting*. Under certain conditions which we shall address later these absorbing points, namely $\partial\mathrm{I}_\infty$, are not only attracting but also *unattainable* in finite time.

Another consequence of (E) concerning the sample paths of Π is that when Π exits any $\mathrm{I} \in \mathcal{I}_\infty$ there is always another $\mathrm{J} \in \mathcal{I}_\infty$ which contains both I and the exit point. We can can state this more technically by saying that $[\,\mathrm{I}\,]_\Pi$ is a proper subset of I_∞ for all I in $\mathcal{I}_\infty$. An important implication of this fact which we will exploit later on is that,

$$\sup[\,\mathrm{I}\,]_\Pi < 1 \qquad \forall\, \mathrm{I} \in \mathcal{I}_\infty. \tag{2.50}$$

2.6 Sufficiency and Verification

We are now ready to state a set of conditions which are sufficient to imply that there exists a continuation interval $\mathrm{I}_* \in \mathcal{I}$ such that

$$\rho(\pi) = \rho_\pi(\tau^{\mathrm{I}_*}), \tag{2.51}$$

which says that the optimal Bayes' cost is achieved by the first exit policy based on I_*. Define the class of $\mathcal{O}_t$–stopping times $\mathcal{T}_{\mathrm{m}}$ by,

$$\mathcal{T}_{\mathrm{m}} := \{\tau \in \mathcal{T}_{ad} : \tau \leq \tau^{\mathrm{I}}\ P_\pi\text{–a.s. } \forall\, \pi \in [0,1] \text{ for some } \mathrm{I} \in \mathcal{I}_\infty\,\}, \tag{2.52}$$

i.e., $\mathcal{T}_{\mathrm{m}}$ consists of all those $\mathcal{O}_t$–stopping times which are *majorized* by some stopping time $\tau^{\mathrm{I}} \in \overline{\mathcal{T}}$ with $\mathrm{I} \in \mathcal{I}_\infty$.

Let $\mathrm{I}_* \in \mathcal{I}_\infty$ and let $r_* : [0,1] \to \Re^e$ denote a corlol function. Consider the following conditions on the pair (r_*, I_*) :

$$\begin{array}{llcll}
(\mathrm{V1a}): & \textit{For all } \tau \in \mathcal{T}_{\mathrm{m}}, & & & \\
& E_\pi[r_*(\Pi_\tau) - r_*(\Pi_0)] & \geq & -E_\pi \displaystyle\int_0^\tau \mathcal{C}_s\, ds & \forall\, \pi \in [\,\mathrm{I}_*\,]_\Pi; \\
(\mathrm{V1b}): & E_\pi[r_*(\Pi_{\tau^{\mathrm{I}_*}}) - r_*(\Pi_0)] & = & -E_\pi \displaystyle\int_0^{\tau^{\mathrm{I}_*}} \mathcal{C}_s\, ds & \forall\, \pi \in [\,\mathrm{I}_*\,]_\Pi; \\
(\mathrm{V2}): & r_*(\pi) & = & e(\pi) & \forall\, \pi \in \partial_\Pi\, \mathrm{I}_*; \\
(\mathrm{V3}): & r_*(\pi) & < & e(\pi) & \forall\, \pi \notin \partial_\Pi\, \mathrm{I}_*;
\end{array}$$

and also,

(V4): *r_* is bounded and continuous on* $[\,\mathrm{I}_*\,]_\Pi$.

The consequences of these conditions are considered in the following series of simple lemmas.

Lemma 2.1 *Suppose there exists a pair* (r_*, I_*) *satisfying* (V1b) *and* (V4). *Then,*

$$P_\pi\{\,\Pi_{\tau^{\mathrm{I}_*}} \in \partial_\Pi \mathrm{I}_*\,\} = 1 \quad \forall \pi \in [\,\mathrm{I}_*\,]_\Pi.$$

Proof: Let $\pi \in [\,\mathrm{I}_*\,]_\Pi$; condition (V4) implies that there exists B_* finite such that,

$$|r_*(\pi)| \le B_* < \infty. \tag{2.53}$$

From the definition of $[\,\mathrm{I}_*\,]_\Pi$ we know $P_\pi\{\,\Pi_t \in [\,\mathrm{I}_*\,]_\Pi,\ 0 \le t \le \tau^{\mathrm{I}_*}\,\} = 1$ so that,

$$\begin{aligned} |\,E_\pi[r_*(\Pi_{\tau^{\mathrm{I}_*}}) - r_*(\Pi_0)]\,| &\le E_\pi[\,|\,r_*(\Pi_{\tau^{\mathrm{I}_*}}) - r_*(\Pi_0)\,|\,] \\ &\le 2B_*. \end{aligned} \tag{2.54}$$

Hence, (V1b) yields,

$$E_\pi \int_0^{\tau^{\mathrm{I}_*}} \mathcal{C}_s\, ds \le 2B_* < \infty, \tag{2.55}$$

and therefore,

$$\infty > E_\pi \int_0^{\tau^{\mathrm{I}_*}} \mathcal{C}_s\, ds \ge E_\pi[\,1\{\tau^{\mathrm{I}_*} = \infty\} \int_0^{\infty} \mathcal{C}_s\, ds\,]. \tag{2.56}$$

In view of condition (C2) on the running cost, this last line leads to a contradiction unless $P_\pi\{\,\tau^{\mathrm{I}_*} < \infty\,\} = 1$. Since π in $[\,\mathrm{I}_*\,]_\Pi$ was chosen arbitrarily this proves the assertion. □

Corollary 2.1 *The stopping time* τ^{I_*} *is admissible, i.e.,* $\tau^{\mathrm{I}_*} \in \mathcal{T}_{ad}$.

Proof: Use expression 2.55 above and Proposition 2.5. □

Lemma 2.2 *Suppose there exists a pair* (r_*, I_*) *satisfying* (V1b), (V2), *and* (V4). *Then,*

$$\rho_\pi(\tau^{\mathrm{I}_*}) = r_*(\pi) \quad \forall\, \pi \in [\,\mathrm{I}_*\,]_\Pi.$$

Proof: Let $\pi \in [\,\mathrm{I}_*\,]_\Pi$. Using Lemma 2.1 and (V2),

$$\begin{aligned} E_\pi[\, r_*(\Pi_{\tau^{\mathrm{I}_*}})\,] &= E_\pi[\, 1\{\, \Pi_{\tau^{\mathrm{I}_*}} \in \partial_\Pi \mathrm{I}_*\,\}\, r_*(\Pi_{\tau^{\mathrm{I}_*}})\,] \\ &= E_\pi[\, 1\{\, \Pi_{\tau^{\mathrm{I}_*}} \in \partial_\Pi \mathrm{I}_*\,\}\, e(\Pi_{\tau^{\mathrm{I}_*}})\,] \\ &= E_\pi[\, e(\Pi_{\tau^{\mathrm{I}_*}})\,]. \end{aligned} \tag{2.57}$$

From (V1b) it therefore follows that,

$$\begin{aligned} \rho_\pi(\tau^{\mathrm{I}_*}) &= E_\pi[\int_0^{\tau^{\mathrm{I}_*}} \mathcal{C}_s\, ds + e(\Pi_{\tau^{\mathrm{I}_*}})\,] \\ &= r_*(\pi) + E_\pi[\, e(\Pi_{\tau^{\mathrm{I}_*}}) - r_*(\Pi_{\tau^{\mathrm{I}_*}})\,] \\ &= r_*(\pi), \end{aligned} \tag{2.58}$$

and this proves the lemma. □

Lemma 2.3 *Suppose there exists a pair* (r_*, I_*) *satisfying* (V1)–(V4). *Then,*

$$\rho(\pi) = \inf_{\tau \in \mathcal{T}_{ad}} \rho_\pi(\tau) = r_*(\pi) \quad \forall\, \pi \in [\,\mathrm{I}_*\,]_\Pi.$$

Proof: Fix $\pi \in [\,\mathrm{I}_*\,]_\Pi$. Define the sequence of intervals I_n for all $n \geq 1$ as,

$$\mathrm{I}_n := \begin{cases} [\,0, \frac{n-1}{n}) & \text{if } (\mathrm{E}^0) \text{ holds;} \\ (\frac{1}{n}, \frac{n-1}{n}) & \text{if } (\mathrm{E}^+) \text{ holds,} \end{cases} \tag{2.59}$$

and note that $\lim \mathrm{I}_n = \mathrm{I}_\infty$. For any $\tau \in \mathcal{T}_{ad}$ we have,

$$\begin{aligned}
\rho_\pi(\tau \wedge \tau^{\mathrm{I}_n}) &= E_\pi[\int_0^{\tau\wedge\tau^{\mathrm{I}_n}} \mathcal{C}_s\, ds + e(\Pi_{\tau\wedge\tau^{\mathrm{I}_n}})] \\
&\leq E_\pi[\int_0^{\tau} \mathcal{C}_s\, ds + e(\Pi_{\tau\wedge\tau^{\mathrm{I}_n}})] \\
&= E_\pi[\int_0^{\tau} \mathcal{C}_s\, ds + e(\Pi_\tau) + e(\Pi_{\tau\wedge\tau^{\mathrm{I}_n}}) - e(\Pi_\tau)] \\
&= \rho_\pi(\tau) + E_\pi[\, e(\Pi_{\tau\wedge\tau^{\mathrm{I}_n}}) - e(\Pi_\tau)\,]. \qquad (2.60)
\end{aligned}$$

Since $\tau \wedge \tau^{\mathrm{I}_n}$ is in $\mathcal{T}_{\mathrm{m}}$ for all $n \geq 1$ from (V1a) we obtain,

$$E_\pi \int_0^{\tau\wedge\tau^{\mathrm{I}_n}} \mathcal{C}_s\, ds \geq E_\pi[\, r_*(\Pi_0) - r_*(\Pi_{\tau\wedge\tau^{\mathrm{I}_n}})\,], \qquad (2.61)$$

and therefore,

$$\rho_\pi(\tau \wedge \tau^{\mathrm{I}_n}) \geq r_*(\pi) + E_\pi[\, e(\Pi_{\tau\wedge\tau^{\mathrm{I}_n}}) - r_*(\Pi_{\tau\wedge\tau^{\mathrm{I}_n}})]. \qquad (2.62)$$

The pair (V2)–(V3) yield $e(\pi) - r_*(\pi) \geq 0$ for all $\pi \in [0,1]$ so this last line implies,

$$\rho_\pi(\tau \wedge \tau^{\mathrm{I}_n}) \geq r_*(\pi). \qquad (2.63)$$

Combining 2.60 and 2.63 yields,

$$r_*(\pi) \leq \rho_\pi(\tau) + E_\pi[\, e(\Pi_{\tau\wedge\tau^{\mathrm{I}_n}}) - e(\Pi_\tau)\,]. \qquad (2.64)$$

We know the escape condition (E) implies $\tau^{\mathrm{I}_n} \nearrow \tau^{\mathrm{I}_\infty}$, P_π–a.s. Therefore from the continuity of e on $[0,1]$ and the Bounded Convergence Theorem we get,

$$\lim_{n\to\infty} E_\pi[\, e(\Pi_{\tau\wedge\tau_n})\,] = E_\pi[\, e(\Pi_{\tau\wedge\tau^{\mathrm{I}_\infty}})\,]. \qquad (2.65)$$

Using 2.65 we pass to the limit in 2.64 and obtain,

$$r_*(\pi) \leq \rho_\pi(\tau) + E_\pi[\, e(\Pi_{\tau\wedge\tau^{\mathrm{I}_\infty}}) - e(\Pi_\tau)\,]. \qquad (2.66)$$

Since $\tau \in \mathcal{T}_{ad}$ the same argument as in Lemma 2.1 shows that also $\tau \in \mathcal{T}$, i.e., $P_\pi\{\tau < \infty\} = 1$. Thus, it follows directly from the definition of I_∞ that,

$$
\begin{aligned}
E_\pi[\, e(\Pi_{\tau \wedge \tau^{\mathrm{I}_\infty}}) - e(\Pi_\tau)\,] &= E_\pi[\,(\, e(\Pi_{\tau^{\mathrm{I}_\infty}}) - e(\Pi_\tau)\,)\, 1\{\tau^{\mathrm{I}_\infty} < \tau\}\,] + 0 \\
&= E_\pi[\,(\, e(\Pi_{\tau^{\mathrm{I}_\infty}}) - e(\Pi_\tau)\,)\, 1\{\tau^{\mathrm{I}_\infty} < \tau < \infty\}\,] \\
&= 0. \qquad (2.67)
\end{aligned}
$$

As a result, the expression 2.66 gives,

$$r_*(\pi) \leq \rho_\pi(\tau). \qquad (2.68)$$

Because we chose $\pi \in [\,\mathrm{I}_*\,]_\Pi$ arbitrarily 2.68 implies,

$$\inf_{\tau \in \mathcal{T}_{ad}} \rho_\pi(\tau) \geq r_*(\pi) \quad \forall \pi \in [\,\mathrm{I}_*\,]_\Pi. \qquad (2.69)$$

Finally, applying Lemma 2.2 and the corollary to Lemma 2.1 we obtain the result. □

Corollary 2.2 *The mapping* r_* *is nonnegative and continuous on* $\mathrm{int}([\,\mathrm{I}_*\,]_\Pi)$.

Proof: Apply Proposition 2.4. □

The next lemma is a further characterization of I_*.

Lemma 2.4 *Suppose there exists a pair* (r_*, I_*) *satisfying* (V1)–(V4). *Then the point at which the terminal cost function* e *attains its maximum value, namely* $\pi_e = c^1/(c^0 + c^1)$, *is contained in* I_*.

Proof: Pick $\pi \in \mathrm{int}(\mathrm{I}_*)$ and choose $\lambda \in (0,1)$ to satisfy $\pi = \lambda\, a_* + (1-\lambda)\, b_*$ where a_* and b_* denote the endpoints of I_*. From the corollary to Lemma 2.3

r_* is concave on $\text{int}(\mathrm{I}_*)$ (*a fortiori*) and so,

$$\begin{aligned} r_*(\pi) &\geq \lambda\, r_*(a_*) + (1-\lambda)\, r_*(b_*) \\ &= \lambda\, e(a_*) + (1-\lambda)\, e(b_*), \end{aligned} \tag{2.70}$$

using condition (V2). Now suppose that $\pi_e < a_*$; we will derive a contradiction. From this supposition, it is obvious that $\pi_e \notin \mathrm{I}_*$ irrespective of whether I_* is open or closed at a_*. From this and the definition of the mapping e we can conclude that e is strictly linear on I_*. The linearity of e on I_* therefore yields,

$$e(\pi) = \lambda\, e(a_*) + (1-\lambda)\, e(b_*), \tag{2.71}$$

so that combining 2.70 and 2.71 we obtain $r_*(\pi) \geq e(\pi)$. Since $\pi \in \text{int}(\mathrm{I}_*)$ and therefore $\pi \notin \partial_\Pi \mathrm{I}_*$, this yields a clear contradiction to (V3) and therefore it must be true that $a_* \leq \pi_e$. Similarly, $b_* \leq \pi_e$ implies $\pi_e \notin \mathrm{I}_*$ since I_* is open at b_* and again, we derive a contradiction to (V3). Hence, we insist that $\pi_e \in \mathrm{I}_*$. □

The importance of Lemma 2.4 is that it implies that e is strictly linear outside of I_*. The importance of the next and final lemma is that it characterizes ρ for all priors not in $[\,\mathrm{I}_*\,]_\Pi$.

Lemma 2.5 *Suppose there exists a pair* (r_*, I_*) *satisfying* (V1)–(V4). *Then,*

$$\rho(\pi) = e(\pi) \qquad \forall \pi \notin [\,\mathrm{I}_*]_\Pi.$$

Proof: Let $\tau_0 \in \mathcal{T}_{ad}$ denote the stopping time which is identically zero, $\tau_0 \equiv 0$. Obviously,

$$0 \leq \rho(\pi) \leq \rho_\pi(\tau_0) = e(\pi) \qquad \forall\, \pi \in [0,1]. \tag{2.72}$$

Since $e(1) = 0$ it then follows that $\rho(1) = 0$. Now suppose we choose $\pi^+ \in \partial_\Pi^+ \mathrm{I}_*$ so that π^+ is in the Π–closure of I_* and therefore Lemma 2.3 yields $\rho(\pi^+) = e(\pi^+)$. From Lemma 2.4 it follows that e is strictly linear on the interval $[\pi^+, 1]$. It is therefore geometrically obvious that ρ, bounded above by e and concave as shown in Proposition 2.4, is also strictly linear on $[\pi^+, 1]$ and in fact equal to e on this interval. This gives us what we want for all points in the unit interval to the right of $[\,\mathrm{I}_*\,]_\Pi$.

For points in the unit interval to the left of $[\,\mathrm{I}_*\,]_\Pi$ a slightly more delicate argument works. Either $\pi_e = 0$ or not. If $\pi_e > 0$ then $e(0) = 0$ and from 2.72 it follows that $\rho(0) = 0$. In this case we can argue as above employing some $\pi^- \in \partial_\Pi^- \mathrm{I}_*$. On the other hand, if $\pi_e = 0$ then $\mathrm{I}_* \in \mathcal{I}_0$ and we obtain the result vacuously since there are no points in the unit interval to the right of $[\,\mathrm{I}_*\,]_\Pi$. Hence, in either case we get what we want. □

Using the results of the previous lemmas we arrive at last at the main result of this chapter.

Theorem 2.2 (Verification) *Suppose there exists a pair* (r_*, I_*) *satisfying* (V1)–(V4). *Then,*

$$\rho(\pi) = \rho_\pi(\tau^{\mathrm{I}_*}) \qquad \forall\, \pi \in [0, 1].$$

Proof:

From Lemma 2.2 and Proposition 2.5 we have,

$$\rho_\pi(\tau^{\mathrm{I}_*}) = \begin{cases} r_*(\pi) & \text{if } \pi \in [\,\mathrm{I}_*\,]_\Pi; \\ e(\pi) & \text{if } \pi \notin [\,\mathrm{I}_*\,]_\Pi. \end{cases} \tag{2.73}$$

Combining Lemma 2.3 and Lemma 2.5 we have,

$$\rho(\pi) = \begin{cases} r_*(\pi) & \text{if } \pi \in [\mathrm{I}_*]_\Pi; \\ e(\pi) & \text{if } \pi \notin [\mathrm{I}_*]_\Pi, \end{cases} \tag{2.74}$$

so that,

$$\rho(\pi) = \rho_\pi(\tau^{\mathrm{I}_*}) \qquad \forall\, \pi \in [0,1]. \tag{2.75}$$

This gives us the result. □

From (V2) and 2.74 a corollary to the theorem is the following simpler expression for ρ,

$$\rho(\pi) = \begin{cases} r_*(\pi) & \text{if } \pi \in \mathrm{I}_*; \\ e(\pi) & \text{if } \pi \notin \mathrm{I}_*. \end{cases} \tag{2.76}$$

We see that Theorem 2.2 *verifies* that the continuation interval I_* characterizes a Bayes' optimal stopping policy, namely $(\tau^{\mathrm{I}_*}, \delta_*)$, and also *verifies* that the function r_* together with I_* characterizes Bayes' risk via 2.76.

The next result states that the pair (r_*, I_*) is **essentially unique**, i.e., if there exists another pair, say (s_*, J_*), satisfying (V1)–(V4) then $\mathrm{I}_* \equiv \mathrm{J}_*$ and

$$r_*(\pi) = s_*(\pi) \qquad \forall\, \pi \in [\mathrm{I}_*]_\Pi. \tag{2.77}$$

We point out that the possible lack of uniqueness of r_* outside of I_* is irrelevant to any questions concerning the optimal Bayes' cost or the optimal first exit policy. For our purposes, only 'uniqueness' as above is 'essential'. We can now prove the last result of this section.

Theorem 2.3 (Essential Uniqueness) *If a pair (r_*, I_*) exists satisfying* (V1)–(V4) *then it is essentially unique.*

Proof: Suppose there exists another pair satisfying these conditions, say (s_*, J_*). Let $\mathrm{I}_* = (a_*, b_*)$, and $\mathrm{J}_* = (c_*, d_*)$. From Lemma 2.4 we know that $\mathrm{I}_* \cap \mathrm{J}_* \neq \emptyset$. Hence, if $a_* \neq c_*$ then either $c_* \in (a_*, b_*)$ or $a_* \in (c_*, d_*)$. Both possibilities lead to a contradiction. For instance, if $c_* \in (a_*, b_*)$ then according to Lemma 2.3 and condition (V3) there holds,

$$\rho(c_*) = r_*(c_*) < e(c_*). \tag{2.78}$$

But Lemma 2.3 applied to s_* implies,

$$\rho(c_*) = s_*(c_*) = e(c_*). \tag{2.79}$$

A similar contradiction is obtained if $a_* \in (c_*, d_*)$ and hence, $a_* = c_*$. By an analogous argument, $b_* \neq d_*$ is untenable and therefore $\mathrm{I}_* \equiv \mathrm{J}_*$ which is half of essential uniqueness. Applying Lemma 2.3 again now yields,

$$r_*(\pi) = \rho(\pi) = s_*(\pi) \qquad \forall\, \pi \in [\mathrm{I}_*]_\Pi, \tag{2.80}$$

which is the other half. □

2.7 A Likelihood Ratio

In this section we fill in still more detail to the probabilistic framework established so far which we will need in the chapters to follow. On the measurable space $(\Omega, \mathcal{A})$ we assume that in addition to the filtrations $\{\mathcal{A}_t\}_{t\geq 0}$ and $\{\mathcal{O}_t\}_{t\geq 0}$ on $\mathcal{A}$ we are also given a third filtration $\{\mathcal{G}_t\}_{t\geq 0}$, such that $\mathcal{O}_t \subset \mathcal{G}_t \subset \mathcal{A}_t$ for all $t \geq 0$. We take $\mathcal{G}_0 = \mathcal{A}_0$ and let $\mathcal{G} = \bigvee_{t\geq 0} \mathcal{G}_t$; we assume that $\mathcal{G}_t$ is completed with respect to P_0 for all $t \geq 0$. The intermediary filtration $\mathcal{G}_t$ will serve to model those system dynamics which still remain only partially observable even with full knowledge of Υ.

Let $P_0^{\mathcal{G}}$ and $P_1^{\mathcal{G}}$ denote the restrictions of P_0 and P_1 to the measurable space $(\Omega, \mathcal{G})$ and assume that $P_0^{\mathcal{G}}$ and $P_1^{\mathcal{G}}$ are mutually absolutely continuous probability measures; this we indicate by $P_0^{\mathcal{G}} \equiv P_1^{\mathcal{G}}$. Define L_∞ to be the Radon-Nikodym derivative of $P_1^{\mathcal{G}}$ with respect to $P_0^{\mathcal{G}}$ and then define $L = \{L_t\}_{t\geq 0}$ according to $L_t := E_0[L_\infty | \mathcal{G}_t]$. Note that $E_0 L_t = E_0 L_\infty = 1$ and then because $\mathcal{G}_0 = \mathcal{A}_0$ we have $L_0 = E_0[L_\infty | \mathcal{A}_0] = E_0 L_\infty = 1$. Hence the mapping $t \mapsto E_0 L_t$ is right-continuous. Also, L is a $\mathcal{G}_t$–martingale and $\mathcal{G}_t$ is an $(\mathcal{A}, P_0)$–completed, right-continuous filtration of $\mathcal{G}$. Thus, L has a corlol modification which is the one we agree that L represents. Of course it is also true that $\{L_t\}_{t\geq 0}$ is a family of uniformly P_0–integrable $\mathcal{G}_t$–martingales.

For any $\mathcal{G}_t$–stopping time ν (nu) the properties of the $\mathcal{G}$–measurable random variable L_ν is understood in terms of the well-defined sub–σ–algebra $\mathcal{G}_\nu$ of $\mathcal{A}$ and we write $L_\nu = E_0[L_\infty \,|\, \mathcal{G}_\nu]$. When ν is more generally some $\mathcal{A}_t$–stopping time, the meaning of the formal symbol L_ν again rests on the meaning of the formal symbol $\mathcal{G}_\nu$ which we now define as,

$$\mathcal{G}_\nu := \{ A \in \mathcal{G} \bigvee \sigma(\nu) \,|\, A \cap \{\nu \le t\} \in \mathcal{G}_t \bigvee \mathcal{U}_t,\, \forall t \ge 0 \}, \tag{2.81}$$

where $\mathcal{U}_t$ conveniently denotes $\bigvee_{s \le t} \sigma(\Upsilon_s)$ for all $t \ge 0$. It is straightforward to show that $\mathcal{G}_\nu$ so defined is closed with respect to complementation and countable intersection and therefore is a σ–algebra; it is obvious therefore that $\mathcal{G}_\nu$ is a sub–σ–algebra of $\mathcal{A}$. Note that if the $\mathcal{A}_t$–stopping time ν is in fact a $\mathcal{G}_t$–stopping time then $\mathcal{G} \bigvee \sigma(\nu) = \mathcal{G}$ and $\mathcal{G}_t \bigvee \mathcal{U}_t = \mathcal{G}_t$ so that the more general definition of $\mathcal{G}_\nu$ above reduces to the usual one. The argument that ν is $\mathcal{G}_\nu$–measurable for ν an $\mathcal{A}_t$–stopping time parallels the same argument when ν is a $\mathcal{G}_t$–stopping time; for X some $\mathcal{G}_t$–progressive process, the proof that X_ν is $\mathcal{G}_\nu$–measurable for ν an $\mathcal{A}_t$–stopping time is the analog of the standard proof that X_ν is $\mathcal{G}_\nu$–measurable for ν a $\mathcal{G}_t$–stopping time. Given this definition for $\mathcal{G}_\nu$ we can now define $L_\nu := E_0[L_\infty \,|\, \mathcal{G}_\nu\,]$ which agrees with our earlier definition when $\nu = t$, a deterministic "$\mathcal{A}_t$–stopping time".

In the development below we will need to consider the process $L^{-1} := \{L_t^{-1}\}_{t \ge 0}$ and thus we present the following proposition.

Proposition 2.6 *The right-continuous $\mathcal{G}_t$–adapted random process L^{-1} is locally bounded, i.e., there exists an increasing sequence of $\mathcal{G}_t$–stopping times $\{\sigma_n\}_{n \ge 1}$ with unbounded limit such that the random process $\{L_{t \wedge \sigma_n}^{-1}\}_{t \ge 0}$ is bounded.*

Proof: Begin by defining,

$$\sigma_n := \inf\{t \ge 0 \,|\, L_t \le \frac{1}{n}\}, \tag{2.82}$$

and note that the sequence $\{\sigma_n(\omega)\}_{n\geq 1}$ is increasing for each $\omega \in \Omega$ so that the limit $\sigma_\infty := \lim_{n\to\infty} \sigma_n$ exists. We want to show that $\sigma_\infty = \infty$, P_0–a.s. To this end note that on the set $\{\sigma_n < n\}$ that $L_{\sigma_n} \leq \frac{1}{n}$: if $\{\sigma_n < n\}$ is empty it holds vacuously; if it is not empty then for all ω in $\{\sigma_n < n\}$ it follows that $L_{\sigma_n(\omega)} \leq \frac{1}{n}$ because $\sigma(\omega) < \infty$ and L is corlol and therefore right-continuous *a fortiori*. Now we take advantage of the fact that L is a uniformly P_0–integrable corlol $\mathcal{G}_t$–martingale and Doob's Optional Sampling theorem to compute,

$$\begin{aligned} E_0[\, L_\infty 1\{\sigma_\infty < n\}\,] &\leq E_0[\, L_\infty 1\{\sigma_n < n\}\,] \\ &= E_0[\, E_0[\, L_\infty 1\{\sigma_n < n\} \,|\, \mathcal{G}_{\sigma_n}\,]\,] \\ &= E_0[\, L_{\sigma_n} 1\{\sigma_n < n\}\,] \\ &\leq \frac{1}{n}. \end{aligned} \tag{2.83}$$

Hence,

$$\lim_{n\to\infty} E_0[\, L_\infty 1\{\sigma_\infty < n\}\,] = 0, \tag{2.84}$$

so that employing the Lebesgue Dominated Convergence theorem we obtain,

$$\begin{aligned} \lim_{n\to\infty} E_0[\, L_\infty 1\{\sigma_\infty < n\}\,] &= E_0[\, L_\infty \lim_{n\to\infty} 1\{\sigma_\infty < n\}\,] \\ &= E_0[\, L_\infty 1\{\sigma_\infty < \infty\}\,], \end{aligned} \tag{2.85}$$

and therefore $E_0[\, L_\infty 1\{\sigma_\infty < \infty\}\,] = 0$. But if we can show that $P_0\{L_\infty > 0\} = 1$ then this implies that $P_0\{\,\sigma_\infty < \infty\,\} = 0$ and as a result we get what we want namely, $P_0\{\,\sigma_\infty = \infty\,\} = 1$. Hence, we compute,

$$P_1\{\, L_\infty = 0\} = \int_{\{L_\infty = 0\}} L_\infty \, dP_0 = 0, \tag{2.86}$$

which yields $P_0\{L_\infty = 0\} = 0$ because $P_0^{\mathcal{G}} \equiv P_1^{\mathcal{G}}$ and so indeed $L_\infty > 0$, P_0–a.s.

To complete the proof, observe that the $\mathcal{A}_t$–adapted process $\{L_{t\wedge\sigma_n}\}_{t\geq 0}$ is bounded away from zero for all $n \geq 1$ and hence $\{L_{t\wedge\sigma_n}^{-1}\}_{t\geq 0}$ is a bounded random process for all $n \geq 1$. Finally, since $\{\sigma_n\}_{n\geq 1}$ is an increasing sequence of $\mathcal{G}_t$–stopping times with limit $\sigma_\infty = \infty$, P_0–a.s. we see that L^{-1} is a locally bounded random process. □

2.8 The Prior Measures: A Model

In this section we focus on the family of Bayes' probability measures, i.e., the prior measures $\{P_\pi : 0 \le \pi \le 1\}$, and distinguish P_0 and P_1 as characterizing the entire family via certain structural assumptions.

Keeping Proposition 2.6 in mind, for any $\mathcal{A}_t$–stopping time ν and any $\mathcal{A}$–measurable random variable X we implicitly define the random measure Q^ν on $(\Omega, \mathcal{A})$ via,

$$\int_J E_{Q^\nu} X \, dP_0 = \int_J E_0[\, L_\nu^{-1} \, X L_\infty \,|\, \sigma(\nu) \,] \, dP_0 \qquad \forall\, J \in \sigma(\nu), \tag{2.87}$$

so that,

$$E_{Q^\nu} X = E_0[\, L_\nu^{-1} \, X L_\infty \,|\, \sigma(\nu) \,] \qquad P_0\text{–a.s.} \tag{2.88}$$

For $\nu = u$ with u in $[0, \infty]$ deterministic we have $\sigma(\nu) = \mathcal{A}_0$ and therefore,

$$E_{Q^u} X = E_0 \, L_u^{-1} \, X \, L_\infty, \tag{2.89}$$

and as a result,

$$Q^u\{A\} = \int_A L_u^{-1} \, L_\infty \, dP_0 \qquad \forall\, A \in \mathcal{A}, \tag{2.90}$$

and also,

$$Q^0\{A\} = P_1\{A\} \qquad \forall\, A \in \mathcal{A}. \tag{2.91}$$

By the first half of Tonelli's theorem we see for all A in $\mathcal{A}$ that the mapping $u \mapsto Q^u\{A\}$ is Borel measurable and thus $\{Q^u : 0 \le u \le 1\}$ defines a transition measure from $([0, \infty], \mathcal{B}([0, \infty]))$ to $(\Omega, \mathcal{A})$, where $\mathcal{B}([0, \infty])$ denotes the Borel σ–algebra on $[0, \infty]$.

Because $P_0^{\mathcal{G}} \equiv P_1^{\mathcal{G}}$ the expressions 2.90 and 2.91 combine to give,

$$Q^u\{G\} = \int_G L_u^{-1}\, dP_1 \qquad \forall\, G \in \mathcal{G},\ 0 \le u \le \infty. \tag{2.92}$$

It is then straightforward to show that,

$$Q^u\{G\} = \left\{ \begin{array}{ll} P_1\{G\} & \text{if } u = 0 \\ \int_G L_u^{-1}\, dP_1 & \text{if } 0 < u < t \\ P_0\{G\} & \text{if } t \le u \le \infty \end{array} \right\} \quad \forall\, G \in \mathcal{G}_t,\ t \ge 0. \tag{2.93}$$

We now turn our attention to the modeling of the $\mathcal{A}_t$–stopping time v from a conditional distribution standpoint with respect to the probability measures P_0 and P_1. Let $F = \{F_t\}_{t\ge 0}$ denote an $\mathcal{O}_t$–predictable, right-continuous, nondecreasing process satisfying,

$$F_0 \equiv 0 \qquad \text{and} \qquad F_\infty = 1, \quad P_0\text{–a.s.}, \tag{2.94}$$

and impose the following technical condition on F:

$$(\mathrm{F}): \quad E_0 \int_0^\tau (1 - F_s)^{-1}\, dF_s < \infty \quad \forall\, \tau \in \mathcal{T}.$$

We describe the $(\mathcal{A}, P_0)$–completed internal history $\{\mathcal{F}_t\}_{t\ge 0}$ of F according to,

$$\mathcal{F}_0 = \mathcal{O}_0; \quad \mathcal{F}_t := \sigma\{\, F_s : 0 \le s \le t \,\}; \quad \mathcal{F} = \bigvee_{t \ge 0} \mathcal{F}_t. \tag{2.95}$$

Recalling that v denotes an $\mathcal{A}_t$–stopping time taking values in $[0, \infty]$, we make the following assumptions concerning the distributional description of v under P_0 and P_1,

$$(\mathrm{D0}) \qquad P_0\{\, v \le t \,|\, \mathcal{F}_t \,\} = F_t \qquad \forall\, t \ge 0,$$

and,

$$\text{(D1)} \qquad P_1\{v = 0\} = 1.$$

From (D0) we see that $P_0\{v = 0\} = 0$. Thus the assumptions (D0) and (D1) are consistent with the assumption (A2). However, it is our goal in this section to *replace* the assumptions (A1) and (A2) pertaining to P_π with the assumptions (D0) and (D1) concerning only P_0 and P_1 and then *obtain* (A1) and (A2) from (D0) and (D1) by properly defining P_π in terms of P_0 and P_1 for all π in $[0, 1]$. Moreover, we shall obtain a more detailed characterization of the family of Bayes' measures which we shall exploit in the section to follow. With our stated goal in mind define,

$$E_\pi X := \pi\, E_1 X + (1 - \pi)\, E_0\, L_v^{-1}\, X\, L_\infty \qquad \forall\, \pi \in [0, 1], \tag{2.96}$$

for all $\mathcal{A}$–measurable random variables X. For notational convenience it makes sense to define an auxiliary measure P via,

$$P\{A\} := \int_A L_v^{-1}\, L_\infty\, dP_0 \qquad \forall\, A \in \mathcal{A}. \tag{2.97}$$

With this convenient notation we see for all π in $[0, 1]$ that,

$$P_\pi\{A\} = \pi\, P_1\{A\} + (1 - \pi)\, P\{A\} \qquad \forall\, A \in \mathcal{A}. \tag{2.98}$$

Observe,

$$\begin{aligned}
P\{\Omega\} &= \int_\Omega L_v^{-1}\, L_\infty\, dP_0 \\
&= \int_\Omega E_0[\, L_v^{-1}\, L_\infty \,|\, \mathcal{G}_v\,]\, dP_0 \\
&= \int_\Omega L_v^{-1}\, E_0[\, L_\infty \,|\, \mathcal{G}_v\,]\, dP_0 \\
&= \int_\Omega L_v^{-1}\, L_v\, dP_0
\end{aligned}$$

$$
\begin{aligned}
&= P_0\{\Omega\} \\
&= 1,
\end{aligned} \tag{2.99}
$$

so that $P\{\Omega\} = 1$ and $P_\pi\{\Omega\} = 1$. Thus both P and P_π are indeed probability measures on $(\Omega, \mathcal{A})$.

We now demonstrate that (A1) and (A2) follow from (D0), (D1) and the definition of P_π in terms of P_0 and P_1. Clearly, whenever λ, π, π', π'' in $[0,1]$ are related by,

$$
\pi = \lambda\,\pi' + (1-\lambda)\,\pi'', \tag{2.100}
$$

our definition of P_π yields for any A in $\mathcal{A}$ that,

$$
\begin{aligned}
P_\pi\{A\} &= [\lambda\,\pi' + (1-\lambda)\,\pi'']\,P_1\{A\} \\
&\qquad + [\lambda\,(1-\pi') + (1-\lambda)\,(1-\pi'')]\,P\{A\} \\
&= \lambda\,[\pi'\,P_1\{A\} + (1-\pi')\,P\{A\}] \\
&\qquad + (1-\lambda)\,[\pi''\,P_1\{A\} + (1-\pi'')\,P\{A\}], \\
&= \lambda\,P_{\pi'}\{A\} + (1-\lambda)\,P_{\pi''}\{A\},
\end{aligned} \tag{2.101}
$$

so that (A1) follows *a fortiori*. In addition, by employing (D0) and (D1) we have for all π in $[0,1]$ that,

$$
\begin{aligned}
P_\pi\{\,v=0\,\} &= \pi\,P_1\{\,v=0\,\} + (1-\pi)\,E_0[\,L_0^{-1}\,1\{v=0\}\,L_\infty\,] \\
&= \pi\cdot 1 + (1-\pi)\,E_0[\,1\cdot 1\{v=0\}\,L_\infty\,] \\
&= \pi + (1-\pi)\cdot 0 \\
&= \pi,
\end{aligned} \tag{2.102}
$$

so that (A2) is also obtained.

We proceed to demonstrate how the definition of the family of Bayes' measures in terms of P_0 and P_1 in concert with the distributional assumptions

(D0) and (D1) and the definition of the random measure Q^v lends desirable properties to the Bayes' family. To begin with we note that completion by P_0 and P_1 of any sub–σ–algebra of $\mathcal{A}$ is sufficient to guarantee its completion by the family of Bayes' measures. Moreover, the collection of suppositions which we have made so far imparts to the family $\{P_\pi : 0 \leq \pi \leq 1\}$ the following properties. For any set J in $\sigma(v)$ and any $\mathcal{G}$–measurable, P_0–integrable random variable X,

$$\begin{aligned} E_0[\,1_J\, L_v^{-1}\, X\, L_\infty\,] &= E_0[\,1_J\, E_0\,[\,L_v^{-1}\, X\, L_\infty |\, \sigma(v)\,]\,] \\ &= E_0[1_J\, E_{Q^v}\,[\,X\,]\,] \\ &= E_0[\,1\{\,v \in v(J)\,\}\int_\Omega X\, dQ^v\,] \\ &= E_0[\,E_0[1\{\,v \in v(J)\,\}\int_\Omega X\, dQ^v\,|\mathcal{F}\,]\,] \\ &= E_0 \int_0^\infty 1\{\,u \in v(J)\}\int_\Omega X\, dQ^u\, dF_u, \end{aligned} \tag{2.103}$$

and similarly,

$$\begin{aligned} E_1[\,1_J\, X\,] &= E_1[1\{\,v \in v(J)\}\,X\,] \\ &= 1\{\,0 \in v(J)\}E_1[\,X\,1\{v = 0\}\,] \\ &= 1\{0 \in v(J)\}\,E_1 X. \end{aligned} \tag{2.104}$$

As a result it follows for all π in $[0,1]$ that,

$$\begin{aligned} E_\pi[1_J\, X] &= (1-\pi)\,E_0\int_0^\infty 1\{\,u \in v(J)\}\int_\Omega X\, dQ^u\, dF_u \\ &\quad + \pi\, 1\{0 \in v(J)\}\, E_1 X. \end{aligned} \tag{2.105}$$

If $J = \{v \leq t\,\}$ then $v(J) = [\,0, t\,]$ and the preceding expression yields,

$$E_\pi[\,1\{v \leq t\,\}\,X] = \pi\, E_1 X + (1-\pi)\,E_0 \int_0^t \int_\Omega X\, dQ^u\, dF_u, \tag{2.106}$$

for all $\mathcal{G}$-measurable X. For $J = \Omega$ set $t = \infty$ in the above to obtain,

$$E_\pi X = \pi\, E_1 X + (1 - \pi)\, E_0 \int_0^\infty \int_\Omega X\, dQ^u\, dF_u. \tag{2.107}$$

Comparing this with the definition of P we see that,

$$P\{G\} = E_0 \int_0^\infty Q^u\{G\}\, dF_u \qquad \forall\, G \in \mathcal{G}. \tag{2.108}$$

2.9 Martingale Dynamics of Π

Under the newly imposed assumptions (A1) and (A2), the results of last section allow us to give a more detailed semimartingale representation for the *a posteriori* probability process. We proceed by first obtaining a more explicit semimartingale representation for the $\mathcal{A}_t$–adapted point process Υ. We then project Υ_t onto the observation filtration and so obtain the desired, more precise semimartingale representation for Π.

Recall that with respect to P_0 the $\mathcal{A}_t$–stopping time v at which Υ jumps from zero to one has the $\mathcal{F}_t$–conditional cumulative distribution function F. Employing a point process representation theorem [B, III.T7] we find that,

$$\Upsilon_t - \int_0^{t\wedge v} (1-F_s)^{-1}\, dF_s \quad \text{is an } (\mathcal{A}_t, P_0)\text{–martingale,} \tag{2.109}$$

so that we may write,

$$\tilde{M}_t = \Upsilon_t - \int_0^{t\wedge v} (1-F_s)^{-1}\, dF_s, \tag{2.110}$$

for some $(\mathcal{A}_t, P_0)$–martingale $\tilde{M}$. Computing with $r \le t$ and A in $\mathcal{A}_r$ yields,

$$\begin{aligned} E_\pi[1_A(\tilde{M}_t - \Upsilon_0)] &= \pi\, E_1[1_A(\tilde{M}_t - \Upsilon_0)] + (1-\pi)\, E_0[1_A(\tilde{M}_t - \Upsilon_0)] \\ &= 0 + (1-\pi) E_0[1_A(\tilde{M}_t - 0)] \\ &= (1-\pi) E_0[1_A E_0[\tilde{M}_t | \mathcal{A}_r]] \\ &= (1-\pi) E_0[1_A \tilde{M}_r] \\ &= \pi\, E_1[1_A(\tilde{M}_r - \Upsilon_0)] + (1-\pi)\, E_0[1_A(\tilde{M}_r - \Upsilon_0)] \\ &= E_\pi[1_A(\tilde{M}_r - \Upsilon_0)], \end{aligned} \tag{2.111}$$

and we conclude that,

$$\int_A E_\pi[\tilde{M}_t - \Upsilon_0 | \, \mathcal{A}_r]\, dP_\pi = \int_A (\tilde{M}_r - \Upsilon_0)\, dP_\pi \qquad \forall\, A \in \mathcal{A}, \tag{2.112}$$

and therefore $\tilde{M} - \Upsilon_0$ is an $(\mathcal{A}_t, P_\pi)$–martingale.

We next exploit the uniqueness of predictable compensators and Proposition 2.1 to make the identifications,

$$K_t = \int_0^{t\wedge v} (1 - F_s)^{-1} \, dF_s \qquad \forall\, t \geq 0, \; P_\pi\text{–a.s.}, \tag{2.113}$$

and thence $M = \tilde{M} - \Upsilon_0$, P_π-a.s., so that we may therefore write,

$$\Upsilon_t = \Upsilon_0 + \int_0^{t\wedge v} (1 - F_s)^{-1} \, dF_s + M_t, \tag{2.114}$$

for the $(\mathcal{A}_t, P_\pi)$–martingale M. Because $\Upsilon_t = 1\{v \leq t\}$ we can in turn rewrite this as,

$$\Upsilon_t = \Upsilon_0 + \int_0^t (1 - \Upsilon_s)\,(1 - F_s)^{-1} \, dF_s + M_t, \tag{2.115}$$

and then,

$$\Upsilon_t = \Upsilon_0 + \int_0^t (1 - \Upsilon_{s-})\,(1 - F_s)^{-1} \, dF_s + M_t, \tag{2.116}$$

since F has no jumps in common with Υ. This provides us with the desired semimartingale representation for Υ.

Note that in view of the technical assumption (F) we have,

$$\begin{aligned} E_\pi \int_0^t (1 - \Upsilon_s)\,(1 - F_s)^{-1} \, dF_s &= (1 - \pi)\, E_0 \int_0^t (1 - \Upsilon_s)\,(1 - F_s)^{-1} \, dF_s \\ &\leq E_0 \int_0^t (1 - F_s)^{-1} \, dF_s < \infty, \end{aligned} \tag{2.117}$$

and Υ indeed has the P_π–integrable compensator that Proposition 2.1 demands.

We end this section by projecting Υ_t onto the observations. Observe that,

$$P_\pi\{\, \Upsilon_0 = 0 \,|\, \mathcal{O}_0 \,\} = P_\pi\{\, \Upsilon_0 = 0 \,\} = \pi, \tag{2.118}$$

and recall $P_\pi\{\Pi_0 = \pi\} = 1$ for all π in $[0,1]$ so that using Proposition 2.1 once again and making a modest modification to the Projection Theorem [W&H, P7.1.3] we obtain for $\Pi_t = E_\pi[\,\Upsilon_t\,|\,\mathcal{O}_t\,]$,

$$\Pi_t = \Pi_0 + \int_0^t (1-\Pi_s)\,(1-F_s)^{-1}\,dF_s + \overline{M}_t \quad \forall t \geq 0,\ P_\pi\text{–a.s.}, \tag{2.119}$$

making use of the $\mathcal{O}_t$–predictability of F. We recall from Proposition 2.2 that $\overline{M}$ denotes an $(\mathcal{O}_t, P_\pi)$–martingale. In the applications in the chapters to follow, the precise nature of the observation filtration will be characterized and this in turn will characterize the structure of $\overline{M}$ as a stochastic integral.

Chapter 3

Change Detection: Diffusion Data

3.1 Introduction

In this chapter we consider the problem of Bayesian optimal change detection when the observed data are modeled by generalized diffusions. The chapter has the following outline:

Section 1. SYSTEM DYNAMICS

In this section we make decisions concerning the general dynamics models for the two underlying systems. One system is modeled as a purely noisy Brownian motion, the other system is modeled as a general drift process with the Brownian motion superimposed. We impose technical conditions on the drift process which will be needed to apply filtering results and obtain escape properties. We derive a representation for the likelihood ratio process using Girsanov's Theorem and employ this representation to obtain a martingale description of the observation process with respect to the prior measure.

Section 2. OBSERVABLE DYNAMICS

The main point of this section is to obtain a martingale description of the observation process with respect to the prior measure when the drift is conditioned upon the observation filtration. To do this we work with the observation process when the drift is smoothed with respect to this filtration and derive the associated likelihood ratio in which only the smoothed drift appears. Thus, this section is largely the analog of the previous but with the drift conditioned upon the observation history.

Section 3. THE CONDITIONAL PROBABILITY

The purpose here is to derive an explicit semimartingale representation for the *a posteriori* probability process by estimating the jump process for the time of change conditioned with respect to the observations. This smoothed representation gives us a filter for the underlying jump state.

Section 4. PRELUDE TO VERIFICATION

In this section we anticipate application of the Verification Theorem of Chapter 2 and compute a more explicit version of the first verification condition taking advantage of our specialization in this chapter to diffusion data.

Section 5. SEQUENTIAL DETECTION

This section is concerned with the classical Bayesian problem of sequential detection based on observations arising from one of two generalized diffusions. We show how this detection problem can be recast as a problem of change detection by properly modeling the jump time within the framework developed and by proper choice of the Bayes cost. We also consider the escape properties of the *a posteriori* probability process prior to setting up the Stefan problem implied by the Verification Theorem. Using a novel approach involving convexity notions and existence and uniqueness for ODE's we solve the Stefan problem and arrive at one of the main results of the chapter: There exists a Bayes' optimal first exit policy for the problem of sequential drift detection with energy cost and terminal error penalties. We conclude the section with an example involving a constant drift diffusion.

Section 6. DISRUPTION

In this section we formulate a general problem of disruption in the case of diffusion-type data. We reformulate this problem directly as a generic change detection problem, employing the jump process to model the disrupted drift. We demonstrate how Bayes' cost as defined here is general enough to capture both the panic cost for deciding too soon that disruption has occurred and also the elapsed energy cost. We examine the escape properties of the conditional probability process inherited from the properties of the smoothed drift and set up the Stefan problem obtained by recourse to the Verification Theorem. The solution to this

Stefan problem is found using the same method as employed in the case of sequential detection. The section concludes with a concrete example involving a constant drift diffusion and an exponentially distributed disruption time.

3.2 System Dynamics

As in the previous chapter we are working on the measurable space $(\Omega, \mathcal{A})$ equipped with two probability measures P_0 and P_1 which induce the family of Bayes' measures $\{P_\pi : 0 \leq \pi \leq 1\}$. We shall draw freely upon our previous results and notation with only the briefest reminders. In a particular application, for instance the one here in which we consider diffusion data, it remains to specify the nature of $\mathcal{O}_t$, $\mathcal{G}_t$, and $\mathcal{A}_t$, to calculate $L = \{L_t\}_{t \geq 0}$, to make choices specializing the $\mathcal{F}_t$–conditional distribution of the disruption time v, and lastly to specify the cost functions. Once these things are done we can make use of the results of the last chapter to solve for the optimal first exit policy.

We begin by supposing that we are given $H = \{H_t\}_{t \geq 0}$, a corlol random process on $(\Omega, \mathcal{A})$ called the **drift**. Define the drift filtration on $\mathcal{A}$ via,

$$\mathcal{H}_t := \bigvee_{0 \leq s \leq t} \sigma(H_s) \qquad \forall\, t \geq 0, \tag{3.1}$$

and take this filtration as $(\mathcal{A}, P_0)$–completed. We are also given $Y = \{Y_t\}_{t \geq 0}$, $Y_0 \equiv 0$, another corlol random process on $(\Omega, \mathcal{A})$ called the **observation**. Define the **observation filtration** on $\mathcal{A}$ via,

$$\mathcal{O}_t := \bigvee_{0 \leq s \leq t} \sigma(Y_s) \qquad \forall\, t \geq 0, \tag{3.2}$$

and take this filtration as $(\mathcal{A}, P_0)$–completed. For any $\mathcal{A}_t$–adapted process $X = \{X_t\}_{t \geq 0}$ we define the $\mathcal{O}_t$–adapted process $\hat{X} = \{\hat{X}_t\}_{t \geq 0}$ via,

$$\hat{X}_t := E_1[\, X_t \,|\, \mathcal{O}_t \,] \qquad \forall\, t \geq 0. \tag{3.3}$$

We collect the following technical conditions involving the drift process H and its $(\mathcal{O}_t, P_1)$–smoothing $\hat{H}$:

$$\begin{array}{llll}
\text{(H0):} & E_i \int_0^t H_s^2\, ds < \infty & \forall\, t \geq 0, & i = 0, 1; \\
\text{(H1):} & E_i \int_0^\tau \hat{H}_s^2\, ds < \infty & \forall\, \tau \in \mathcal{T}, & i = 0, 1; \\
\text{(H2):} & P_i \{ \int_0^\infty \hat{H}_s^2\, ds = \infty \} = 1 & & i = 0, 1.
\end{array}$$

Next we define $\mathcal{G}_t := \mathcal{O}_t \bigvee \mathcal{H}_t$ and make the following assumptions concerning the dynamics model for the observation process under P_0 and P_1:

$$\begin{array}{lll}
\text{(DM0)}: & Y_t & \text{is a } (\mathcal{G}_t, P_0)\text{–Wiener martingale;} \\
\text{(DM1)}: & Y_t - \int_0^t H_s\, ds & \text{is a } (\mathcal{G}_t, P_1)\text{–Wiener martingale.}
\end{array}$$

Recalling that υ denotes a measurable mapping from Ω to $[0, \infty]$, that $\Upsilon_t := 1\{\upsilon \leq t\}$, and that $\mathcal{U}_t = \bigvee_{s \leq t} \sigma(\Upsilon_t)$ we define,

$$\mathcal{A}_t := \mathcal{G}_t \bigvee \mathcal{U}_t \bigvee \mathcal{A}_0 \qquad \forall\, t > 0, \tag{3.4}$$

where $\mathcal{A}_0$ denotes the smallest σ–algebra containing both the P_0–null sets of $\mathcal{A}$ and the P_1–null sets of $\mathcal{A}$. We have come to our first proposition.

Proposition 3.1 *The $\mathcal{G}_t$–adapted process L is given by,*

$$L_t = \exp\{ \int_0^t H_s\, dY_s - \frac{1}{2} \int_0^t H_s^2\, ds \} \qquad \forall\, t \geq 0.$$

Proof: The proof of this representation is well-known and standard but we include it for completeness and because some of its ingredients are reused in succeeding propositions. Applying the Itô rule to the natural logarithm of the $(\mathcal{G}_t, P_0)$–martingale $L = \{L_t\}_{t\geq 0}$ with $L_0 \equiv 0$ we obtain,

$$\log\, L_t = \int_0^t L_s^{-1}\, dL_s - \frac{1}{2}\int_0^t L_s^{-2}\, d[L,L]_s^c \quad \forall\, t \geq 0,\ P_0\text{–a.s.} \tag{3.5}$$

Define $X = \{X_t\}_{t\geq 0}$ via $X := L^{-1} \bullet L$ where,

$$L^{-1} \bullet L_t := \int_0^t L_{s-}^{-1}\, dL_s = \int_0^t L_s^{-1}\, dL_s \qquad \forall\, t \geq 0. \tag{3.6}$$

Because L is a $(\mathcal{G}_t, P_0)$–martingale and L^{-1} is a $(\mathcal{G}_t, P_0)$–progressive process which according to Proposition 2.6 is locally bounded, we see that X is a $(\mathcal{G}_t, P_0)$–local martingale. Next note that,

$$\begin{aligned} L^{-2} \bullet [L,L]^c &= L^{-1} \bullet [L^{-1} \bullet L, L]^c \\ &= L^{-1} \bullet [L, L^{-1} \bullet L]^c \\ &= [L^{-1} \bullet L, L^{-1} \bullet L]^c = [X,X]^c, \end{aligned} \tag{3.7}$$

i.e.,

$$\int_0^t L_s^{-2}\, d[L,L]_s^c = [X,X]_t^c \qquad \forall\, t \geq 0,\ P_0\text{–a.s.} \tag{3.8}$$

As a result we see that,

$$\log\, L_t = X_t - \frac{1}{2}\,[X,X]_t^c, \tag{3.9}$$

and therefore,

$$L_t = \exp\{X_t - \frac{1}{2}\,[X,X]_t^c\}. \tag{3.10}$$

Since X is a $(\mathcal{G}_t, P_0)$–local martingale with $X_0 \equiv 0$ we can employ the Martingale Representation Theorem [W&H, P6.7.3] to conclude that X has the stochastic integral representation,

$$X_t = \int_0^t \xi_s\, dY_s \qquad \forall\, t \geq 0,\ P_0\text{–a.s.}, \tag{3.11}$$

for some $\mathcal{G}_t$-progressive process $\xi = \{\xi_t\}_{t\geq 0}$ satisfying $\int_0^t \xi_s^2\, ds < \infty$, P_0–a.s., for all $t \geq 0$. Since $< Y, X >_t = [Y, X]_t = \int_0^t \xi_s\, ds$ and because both X and Y are $(\mathcal{G}_t, P_0)$–local martingales we may apply the abstract Girsanov Theorem [W&H, P6.7.2] and so conclude that $Y- < Y, X >$ is a $(\mathcal{G}_t, P_1)$–local martingale. But then employing (DM1) and the fact that predictable compensators are unique we conclude $\xi = H$, P_0–a.s., so that,

$$X_t = \int_0^t L_s^{-1}\, dL_s = \int_0^t H_s\, dY_s, \tag{3.12}$$

and therefore,

$$L_t = \exp\{\int_0^t H_s\, dY_s - \frac{1}{2}\int_0^t H_s^2\, ds\}, \tag{3.13}$$

which completes the proof. □

The next proposition which we shall give begins to reveal the role of the random measure Q^v on $(\Omega, \mathcal{A})$, but first we need to prove the following lemma involving the Q^u measure on $(\Omega, \mathcal{G})$ for any u in $[0, \infty]$.

Lemma 3.1 *For each $u \in \overline{\Re}_+$,*

$$Y_t - \int_0^t U_s\, H_s\, ds \quad \textit{is a } (\mathcal{G}_t, Q^u)\textit{–martingale,}$$

where U denotes the deterministic indicator process $U_t := 1\{u \leq t\}$.

Proof: For $u \in [0, \infty]$, define the auxiliary $(\mathcal{G}_t, P_0)$–local martingale M^u,

$$M_t^u := \int_0^t U_s\, H_s\, dY_s \qquad \forall\, t \geq 0, \tag{3.14}$$

and the $(\mathcal{G}_t, P_0)$–adapted process L^u,

$$L_t^u := \exp\{M_t^u - \frac{1}{2}[M^u, M^u]_t\} \qquad \forall\, t \geq 0. \tag{3.15}$$

A simple calculation shows that,

$$L_t^u = \begin{cases} 1 & 0 \le t < u \le \infty; \\ L_u^{-1} L_t & 0 \le u \le t \le \infty, \end{cases} \tag{3.16}$$

and from this there follows,

$$L_\infty^u := \lim_{n \to \infty} L_n^u = L_u^{-1} L_\infty \qquad \forall u \in [0, \infty], \tag{3.17}$$

using [W&H, P6.1.4]. Recalling the definition of Q^u we see that,

$$Q^u\{G\} = \int_G L_\infty^u \, dP_0 \qquad \forall G \in \mathcal{G}. \tag{3.18}$$

Computing under P_0 we obtain,

$$< Y, M^u >_t = [Y, M^u]_t = \int_0^t U_s \, H_s \, ds. \tag{3.19}$$

Thus, applying Girsanov's Theorem we conclude that $Y- < Y, M^u >$ is a $(\mathcal{G}_t, Q^u)$–local martingale and this gives us the result. □

This brings us to the following proposition which describes the unobservable dynamics of the Y process on $(\Omega, \mathcal{A})$ with respect to the P_π measure for any π in $[0, 1]$ fixed.

Proposition 3.2

$$Y_t - \int_0^t \Upsilon_s \, H_s \, ds \ \textit{ is an } (\mathcal{A}_t, P_\pi)\textit{–martingale} \quad \forall \pi \in [0, 1].$$

Proof: Fix π in $[0, 1]$. Let $r \le t$ and pick A in $\mathcal{A}_r$ satisfying $A = J \cap G$ with $J \in \mathcal{U}_r$ and $G \in \mathcal{G}_r$. Using Fubini's Theorem we obtain,

$$E_\pi \, 1_A \int_0^t \Upsilon_s \, H_s \, ds = \int_0^t E_\pi \Upsilon_s \, 1_A \, H_s \, ds. \tag{3.20}$$

Next employing expression 2.105 and then Fubini's Theorem twice it follows that the right-hand side of 3.20 is equal to,

$$\pi\, E_1 1_A \int_0^t H_s\, ds + (1-\pi)\, E_0 \int_0^t \int_0^\infty 1\{u \in v(J_s)\} \int_G H_s\, dQ^u\, dF_u\, ds, \quad (3.21)$$

where J_s denotes the $\sigma(v)$–measurable set $J \cap \{v \leq s\}$. Clearly, $v(J_s) = [0,s] \cap v(J)$ and therefore $1\{u \in v(J_s)\} = U_s\, 1\{u \in v(J)\}$ where again, U_t is the indicator function $1\{u \leq t\}$. Thus, the expectation in the second term equals,

$$E_0 \int_0^t \int_{v(J)} \int_G U_s\, H_s\, dQ^u\, dF_u\, ds, \quad (3.22)$$

and with two more applications of Fubini's Theorem this expectation can be rewritten as,

$$E_0 \int_0^t \int_{v(J)} \int_G U_s\, H_s\, dQ^u\, dF_u\, ds = E_0 \int_{v(J)} \int_G \int_0^t U_s\, H_s\, ds\, dQ^u\, dF_u. \quad (3.23)$$

Thus far our calculations imply that,

$$E_\pi 1_A \int_0^t \Upsilon_s H_s ds = \pi\, E_1 1_A \int_0^t H_s ds + (1-\pi)\, E_0 \int_{v(J)} \int_G \int_0^t U_s H_s\, ds\, dQ^u dF_u. \quad (3.24)$$

Employing 2.105 once again, similar appeals to Fubini yield,

$$E_\pi 1_A\, Y_t = \pi\, E_1 1_A\, Y_t + (1-\pi)\, E_0 \int_{v(J)} \int_G Y_t\, dQ^u\, dF_u. \quad (3.25)$$

The next step is to to compute $E_\pi 1_A[Y_t - \int_0^t \Upsilon_s\, H_s\, ds]$, the difference of 3.24 and 3.25; for notational economy define the $\mathcal{A}_t$–adapted process $W^v = \{W_t^v\}_{t\geq 0}$ via,

$$W_t^v := Y_t - \int_0^t \Upsilon_s\, H_s\, ds \qquad \forall\, t \geq 0, \quad (3.26)$$

and for u in $[0, \infty]$ define the $\mathcal{G}_t$–adapted process $W^u = \{W_t^u\}_{t \geq 0}$ by,

$$W_t^u := Y_t - \int_0^t U_s \, H_s \, ds \qquad \forall \, t \geq 0. \tag{3.27}$$

In this notation the goal of the present proposition is to prove that W^v is an $(\mathcal{A}_t, P_\pi)$–martingale; note also in this notation Lemma 3.2 says that W^u is a $(\mathcal{G}_t, Q^u)$–martingale for all u in $[0, \infty]$. Subtract 3.24 from 3.25 to obtain,

$$E_\pi 1_A \, W_t^v = \pi \, E_1 1_A W_t^0 + (1-\pi) \, E_0 \int_{v(J)} \int_G W_t^u \, dQ^u \, dF_u. \tag{3.28}$$

Working on the second expectation in 3.28 we compute,

$$\begin{aligned} E_0 \int_{v(J)} \int_G W_t^u \, dQ^u \, dF_u &= E_0 \int_{v(J)} E_{Q^u}[\, 1_G \, W_t^u \,] \, dF_u \\ &= E_0 \int_{v(J)} E_{Q^u}[\, 1_G \, E_{Q^u}[W_t^u \, | \mathcal{G}_r \,] \,] \, dF_u \\ &= E_0 \int_{v(J)} E_{Q^u}[\, 1_G \, W_r^u \,] \, dF_u \\ &= E_0 \int_{v(J)} \int_G W_r^u \, dQ^u \, dF_u, \end{aligned} \tag{3.29}$$

again, because Lemma 3.2 says that W^u is a $(\mathcal{G}_t, Q^u)$–martingale. Of course W^0 is a $(\mathcal{G}_t, P_1)$–martingale so that,

$$\begin{aligned} E_1 1_A W_t^0 &= E_1 \, 1\{v \in v(J)\} 1_G W_t^0 \\ &= 1\{0 \in v(J)\} \, E_1 1_G W_t^0 \\ &= 1\{0 \in v(J)\} \, E_1[\, 1_G \, E_1[W_t^0 \, | \, \mathcal{G}_r \,] \,] \\ &= 1\{0 \in v(J)\} \, E_1 1_G \, W_r^0 \\ &= E_1 \, 1\{v \in v(J)\} 1_G W_r^0 \\ &= E_1 1_A W_r^0. \end{aligned} \tag{3.30}$$

Plugging the last two results into 3.28 yields,

$$E_\pi 1_A \, W_t^v = \pi \, E_1 1_A W_r^0 + (1-\pi) \, E_0 \int_{v(J)} \int_G W_r^u \, dQ^u \, dF_u. \tag{3.31}$$

But according to 3.28 the right-hand side of 3.31 is precisely $E_\pi 1_A W_r^v$. Thus we have shown,

$$E_\pi 1_A W_t^v = E_\pi 1_A W_r^v. \tag{3.32}$$

The conclusion: For any r in $[0, t]$ and any $t \geq 0$,

$$E_\pi 1_{J\cap G} W_t^v = E_\pi 1_{J\cap G} W_r^v \quad \forall J \in \mathcal{U}_r \text{ and } G \in \mathcal{G}_r. \tag{3.33}$$

We now use the fact that the σ–algebra $\mathcal{A}_r = \mathcal{U}_r \bigvee \mathcal{G}_r$ is generated by the sets of the form $J \cap G$ as in 3.33 to show that 3.33 holds more generally on all of $\mathcal{A}_r$. To be precise, define the π-system,

$$\mathcal{M} := \{A \in \mathcal{A}_r : A = J \cap G \text{ with } J \in \mathcal{U}_r \text{ and } G \in \mathcal{G}_r\}, \tag{3.34}$$

and define two measures μ_r and μ_t on $(\Omega, \mathcal{A}_r)$ as,

$$\mu_r\{A\} := \int_A W_r^v \, dP_\pi \quad \forall A \in \mathcal{A}_r, \tag{3.35}$$

and

$$\mu_t\{A\} := \int_A W_t^v \, dP_\pi \quad \forall A \in \mathcal{A}_r. \tag{3.36}$$

In this notation 3.33 says,

$$\mu_r\{M\} = \mu_t\{M\} \qquad \forall M \in \mathcal{M}. \tag{3.37}$$

Because W_r^v and W_t^v are P_π–integrable and $\Omega \in \mathcal{M}$, we see that μ_r and μ_t are *a fortiori* σ–finite on the π–system $\mathcal{M}$. These properties of μ_r and μ_t plus the fact that $\sigma(\mathcal{M})$ equals $\mathcal{A}_r$ allows us to apply an extension to Dynkin's π-λ Theorem [BILL, T10.3] to conclude that,

$$\mu_r\{A\} = \mu_t\{A\} \qquad \forall A \in \mathcal{A}_r. \tag{3.38}$$

As a result we therefore obtain,

$$\int_A W_t^v \, dP_\pi = \int_A E_\pi[\, W_t^v \,|\, \mathcal{A}_r \,] \, dP_\pi = \int_A W_r^v \, dP_\pi \quad \forall \, A \in \mathcal{A}_r, \tag{3.39}$$

so that,

$$E_\pi[\, W_t^v \,|\, \mathcal{A}_r \,] = W_r^v \qquad P_\pi\text{–a.s.}, \tag{3.40}$$

and in words, W^v is an $(\mathcal{A}_t, P_\pi)$–martingale. □

The proposition tells us that the observation process has the semimartingale dynamics

$$Y_t = \int_0^t \Upsilon_s \, H_s \, ds + W_t^v \qquad \forall \, t \geq 0, \tag{3.41}$$

for some $(\mathcal{A}_t, P_\pi)$–martingale $W^v = \{W_t^v\}_{t \geq 0}$.

3.3 Observable Dynamics

Given that the main goal of this section is to show that,

$$Y_t - \int_0^t \Pi_s \hat{H}_s \, ds \text{ is an } (\mathcal{O}_t, P_\pi)\text{–martingale,} \tag{3.42}$$

it comes as no surprise that this section parallels the previous. We begin with some new notation. Define the $\mathcal{O}_t$–adapted process $\check{X} = \{\check{X}_t\}_{t\geq 0}$ via,

$$\check{X}_t := E_0[\, X_t \,|\, \mathcal{O}_t \,] \qquad \forall\, t \geq 0. \tag{3.43}$$

This notation complements our earlier definition of $\hat{X}_t$ as $E_1[\, X_t \,|\, \mathcal{O}_t \,]$: the superscribed accent indicating the $\mathcal{O}_t$–smoothing and its up/down directionality reminding us that this smoothing is with respect to P_0 or P_1, respectively, a mnemonic suggested by the relative positions of these two measures in the symbol for the Radon–Nikodym derivative L_∞ as we defined it in Chapter 2. The following proposition gives us a representation for $\check{L}$.

Proposition 3.3 *The $\mathcal{O}_t$–adapted process $\check{L}$ is given by,*

$$\check{L}_t = \exp\{\int_0^t \hat{H}_s \, dY_s - \frac{1}{2}\int_0^t \hat{H}_s^2 \, ds\} \qquad \forall\, t \geq 0.$$

Proof: The proof of this proposition is the analog of the proof of Proposition 3.1. With X redefined here as $X := \check{L}^{-1} \bullet \check{L}$ we obtain by a similar argument that,

$$\check{L}_t = \exp\{X_t - \frac{1}{2}[X, X]_t\}. \tag{3.44}$$

A slight modification to Proposition 2.6 will show that $\check{L}^{-1}$ is a locally bounded $(\mathcal{O}_t, P_0)$–progressive process; it is obvious that $\check{L}$ is a $(\mathcal{O}_t, P_0)$–martingale. Thus, from its definition we see that X is a $(\mathcal{O}_t, P_0)$–local martingale with $X_0 \equiv 0$. From the Martingale Representation Theorem [W&H, P6.7.3] we know that X can be expressed as,

$$X_t = \int_0^t \xi_s \, dY_s \qquad \forall\, t \geq 0,\ P_0\text{–a.s.}, \tag{3.45}$$

for some $\mathcal{O}_t$–progressive process $\xi = \{\xi_t\}_{t \geq 0}$ satisfying $\int_0^t \xi_s^2 \, ds < \infty$, P_0–a.s., for all $t \geq 0$. Applying Girsanov's Theorem [W&H, P6.7.2] we can conclude that $Y - < Y, X >$ is an $(\mathcal{O}_t, P_1)$–local martingale. Then by the uniqueness of predictable compensators we obtain $\xi = \hat{H}$, P_0–a.s., since,

$$Y_t - \int_0^t \hat{H}_s \, ds \quad \text{is an } (\mathcal{O}_t, P_1)\text{–martingale}, \tag{3.46}$$

where 3.46 follows by an easy application of the Projection Theorem [W&H P7.1.3]. Therefore,

$$X_t = \int_0^t \check{L}_s^{-1} \, d\check{L}_s = \int_0^t \hat{H}_s \, dY_s, \tag{3.47}$$

and finally,

$$\check{L}_t = \exp\{\int_0^t \hat{H}_s \, dY_s - \frac{1}{2} \int_0^t \hat{H}_s^2 \, ds\}, \tag{3.48}$$

and the proof is complete. □

Next, we prove the following lemma involving the Q^u measure on $(\Omega, \mathcal{O})$ for any u in $[0, \infty]$.

Lemma 3.2 *For each $u \in \overline{\Re}_+$,*

$$Y_t - \int_0^t U_s \hat{H}_s \, ds \ \ \textit{is an } (\mathcal{O}_t, Q^u)\textit{–martingale},$$

where U denotes the deterministic indicator process $U_t := 1\{u \leq t\}$.

Proof: From Lemma 3.2 we know that,

$$Y_t - \int_0^t U_s\, H_s\, ds \text{ is a } (\mathcal{G}_t, Q^u)\text{–martingale.} \tag{3.49}$$

Thus, a routine application of the Projection Theorem yields the result. □

This brings us to the following proposition which describes the dynamics of Y on $(\Omega, \mathcal{O})$ with respect to the P_π measure for any fixed π in $[0,1]$.

Proposition 3.4

$$Y_t - \int_0^t \Pi_s\, \hat{H}_s\, ds \;\; \textit{is an } (\mathcal{O}_t, P_\pi)\textit{–martingale} \quad \forall\, \pi \in [0,1].$$

Proof: For $0 \le r \le t$ and with $\mathcal{O}_r$ in the role of $\mathcal{G}_r$, the same argument as in the proof of Proposition 3.2 leads to the implication that,

$$\int_A E_\pi[\, \tilde{W}_t^v \,|\, \mathcal{U}_r \bigvee \mathcal{O}_r \,]\, dP_\pi = \int_A \tilde{W}_r^v\, dP_\pi \quad \forall\, A \in \mathcal{U}_r \bigvee \mathcal{O}_r, \tag{3.50}$$

where for the purposes of this proposition we define,

$$\tilde{W}_t^v := Y_t - \int_0^t \Upsilon_s\, \hat{H}_s\, ds \quad \forall\, t \ge 0, \tag{3.51}$$

and let $\tilde{W}^v$ take on the role of W^v in Proposition 3.2. Hence the analogous conclusion is reached, i.e., $\tilde{W}^v$ is an $(\mathcal{U}_t \bigvee \mathcal{O}_t, P_\pi)$–martingale. Using this fact, the quite obvious fact that $\mathcal{O}_t$ is a sub–σ–algebra of $\mathcal{U}_t \bigvee \mathcal{O}_t$, and the Projection Theorem [W&H, P7.1.3] yields,

$$\begin{aligned} E_\pi[\, \tilde{W}_t^v \,|\, \mathcal{O}_t \,] &= Y_t - \int_0^t E_\pi[\, \Upsilon_s \,|\, \mathcal{O}_s \,]\, \hat{H}_s\, ds \\ &= Y_t - \int_0^t \Pi_s\, \hat{H}_s\, ds, \end{aligned} \tag{3.52}$$

for all $t \geq 0$. Using this result and two applications of the smoothing property of conditional expectation we compute,

$$\begin{aligned} E_\pi[\,Y_t - \int_0^t \Pi_s \hat{H}_s\, ds \,|\, \mathcal{O}_r\,] &= E_\pi[\,E_\pi[\,\tilde{W}_t^v \,|\, \mathcal{O}_t\,] \,|\, \mathcal{O}_r\,] \\ &= E_\pi[\,\tilde{W}_t^v \,|\, \mathcal{O}_r\,] \\ &= E_\pi[\,E_\pi[\,\tilde{W}_t^v \,|\mathcal{U}_r \bigvee \mathcal{O}_r\,] \,|\, \mathcal{O}_r\,] \\ &= E_\pi[\,\tilde{W}_r^v \,|\, \mathcal{O}_r\,] \\ &= Y_r - \int_0^r \Pi_s \hat{H}_s\, ds, \end{aligned} \tag{3.53}$$

and we see that our claim is indeed true. □

Let's take this opportunity to summarize the results of this section and the last. We have given two models for the total dynamics of an observable process Y under two different probability measures:

$$Y_t \quad \text{is a } (\mathcal{G}_t, P_0)\text{–Wiener martingale,} \tag{3.54}$$

and,

$$Y_t - \int_0^t H_s\, ds \quad \text{is a } (\mathcal{G}_t, P_1)\text{–Wiener martingale.} \tag{3.55}$$

We interpret each measure as modeling a distinct mode of operation of some underlying dynamical system on $(\Omega, \mathcal{G})$ which we observe through Y. Under P_0 the hidden dynamics are modeled as merely a Brownian motion and they influence the observations directly. Under P_1 the hidden dynamics are modeled via $H = \{H_t\}_{t \geq 0}$ and these dynamics influence the observations through the typical 'signal plus noise' set-up. The $\mathcal{G}_t$–progressive, $\mathcal{O}_t$–partially observable dynamics themselves, H, may arise according to any number of models, for instance, a memoryless nonlinear transformation of a process with a linear stochastic differential description.

For all G in $\mathcal{G}$ note that $Q^0\{G\} = P_1\{G\}$ while $Q^\infty\{G\} = P_0\{G\}$. We see that the transition measure Q^u 'links' the two distinct modes of system dynamics. Indeed, if the time of mode change were deterministic then $\{Q^u : 0 \leq u \leq \infty\}$ provides us with a tool to answer questions concerning the (still stochastic) behavior of the underlying system. However, we are interested in problems where the overall system has observable dynamics which can change at a random time v from say mode-0 modeled according 3.54 to mode-1, a different mode in which the observable dynamics are properly modeled using 3.55 above. Moreover, in this problem we assume there is π-probability that the system is initially in mode-1 and $(1-\pi)$–probability that the jump time v is positive and distributed according to the $\mathcal{O}_t$–conditional cumulative distribution function F. On the probability space $(\Omega, \mathcal{A}, P_\pi)$ which we have constructed to model this situation, the observable dynamics have a single representation—they behave according to,

$$Y_t - \int_0^t \Upsilon_s H_s \, ds \quad \text{is an } (\mathcal{A}_t, P_\pi)\text{–martingale}, \tag{3.56}$$

which is the conclusion to Proposition 3.2.

To complete our summary, we note that we also have two models for the projections of the dynamics of Y onto the observations under P_0 and P_1:

$$Y_t \quad \text{is an } (\mathcal{O}_t, P_0)\text{–Wiener martingale}, \tag{3.57}$$

and,

$$Y_t - \int_0^t \hat{H}_s \, ds \quad \text{is an } (\mathcal{O}_t, P_1)\text{–Wiener martingale}. \tag{3.58}$$

Using these projections, in Proposition 3.4 we obtained the analogous paritally observable representation for the observation dynamics on $(\Omega, \mathcal{A}, P_\pi)$

as,

$$Y_t - \int_0^t \Pi_s \hat{H}_s \, ds \quad \text{is an } (\mathcal{O}_t, P_\pi)\text{–martingale.} \tag{3.59}$$

This last martingale is the one we need to obtain explicit filtering results involving the *a posteriori* probability. We take up this task in the section to follow.

We end the current section with a result concerning an escape property of the likelihood ratio process $\check{L} = \{\check{L}_t\}_{t \geq 0}$ which follows directly from the condition (H1) imposed on the drift process at the outset of this chapter.

Proposition 3.5 *Under assumption* (H1),

$$P_\pi\{ \sup_{0 \leq t \leq n} |\log \check{L}_t| = \infty \} = 0 \quad \forall n \geq 1, \; \forall \pi \in \mathrm{I}_\infty.$$

Proof: Fix $\pi \in \mathrm{I}_\infty$. From Proposition 3.3 we know that the $\mathcal{O}_t$–adapted process $\log \check{L}$ is given by,

$$\log \check{L}_t = \int_0^t \hat{H}_s \, dY_s - \frac{1}{2} \int_0^t \hat{H}_s^2 \, ds \quad \forall t \geq 0, \tag{3.60}$$

so that computing under P_π we obtain (see Proposition 3.4),

$$\begin{aligned} |\log \check{L}_t| &= |\int_0^t (\Pi_s - \frac{1}{2}) \hat{H}_s^2 \, ds + \int_0^t \hat{H}_s \, d\overline{W}_s^{\upsilon}| \\ &\leq |\frac{1}{2} \int_0^t \hat{H}_s^2 \, ds| + |\int_0^t \hat{H}_s \, d\overline{W}_s^{\upsilon}|. \end{aligned} \tag{3.61}$$

Hence for all $n \geq 1$,

$$P_\pi\{ \sup_{0 \leq t \leq n} |\log \check{L}_t| = \infty \} \;\leq\; P_\pi\{ \int_0^n \hat{H}_s^2 \, ds + \sup_{0 \leq t \leq n} |\int_0^t \hat{H}_s \, d\overline{W}_s^{\upsilon}| = \infty \}$$

$$= P_\pi\{\sup_{0\le t\le n} |\int_0^t \hat{H}_s\, d\overline{W}^v_s| = \infty\}, \tag{3.62}$$

where the last line follows *a foritiori* from (H1).

Applying Kolmogorov's inequality we find that,

$$P_\pi\{\sup_{0\le t\le n} |\int_0^t \hat{H}_s\, d\overline{W}^v_s| \ge m\} \le \frac{E_\pi\left[\int_0^n \hat{H}_s\, d\overline{W}^v_s\right]^2}{m^2} \qquad \forall m \ge 1, \tag{3.63}$$

while the Itô product rule and familiar localization arguments imply,

$$E_\pi\left[\int_0^n \hat{H}_s\, d\overline{W}^v_s\right]^2 = E_\pi[\hat{H}\bullet\overline{W}^v, \hat{H}\bullet\overline{W}^v]_n = E_\pi \int_0^n \hat{H}^2_s\, ds. \tag{3.64}$$

Combining the last two expressions and employing (H1) again we obtain,

$$\begin{aligned} P_\pi\{\sup_{0\le t\le n} |\int_0^t \hat{H}_s\, d\overline{W}^v_s| = \infty\} &\le P_\pi\{\sup_{0\le t\le n} |\int_0^t \hat{H}_s\, d\overline{W}^v_s| \ge m\} \\ &\le \frac{E_\pi \int_0^n \hat{H}^2_s\, ds}{m^2} \qquad \forall m \ge 1. \end{aligned} \tag{3.65}$$

Passing to the limit above and taking expression 3.62 into account yields,

$$P_\pi\{\sup_{0\le t\le n} |\log \check{L}_t| = \infty\} = 0 \quad \forall n \ge 1, \tag{3.66}$$

and the proof is complete. □

3.4 The Conditional Probability

Fix $\pi \in [0,1]$. We conveniently collect some of our results so far to compute a final, explicit representation for the projection of Υ_t onto $\mathcal{O}_t$.

1. From 2.116 we see that the single-jump, binary point process Υ has the $(\mathcal{A}_t, P_\pi)$–semimartingale representation,

$$\Upsilon_t = \Upsilon_0 + \int_0^t (1 - \Upsilon_{s-})\,(1 - F_s)^{-1}\, dF_s + M_t \quad \forall\, t \geq 0,\ P_\pi\text{–a.s.,} \tag{3.67}$$

   with Υ_0 an $\mathcal{A}_0$–measurable binary random variable satisfying $E_\pi \Upsilon_0 = \pi$, with $(1 - \Upsilon_{s-})$ an $\mathcal{A}_t$–predictable process, and with M an $(\mathcal{A}_t, P_\pi)$–martingale.

2. From Proposition 3.2 the observation process satisfies,

$$Y_t = \int_0^t \Upsilon_{s-}\, H_s\, ds + W_t^v \quad \forall\, t \geq 0,\ P_\pi\text{–a.s.} \tag{3.68}$$

   for H some $\mathcal{G}_t$–progressive process such that,

$$E_\pi \int_0^t H_s^2\, ds < \infty \quad \forall\, t \geq 0, \tag{3.69}$$

   where W^v denotes an $(\mathcal{A}_t, P_\pi)$–Wiener martingale. The bounding in 3.69 follows from assumption (H0) and expression 2.96.

We frame the filter for the projection of Υ_t onto $\mathcal{O}_t$ in the following proposition.

Proposition 3.6 *The filter for* $\Pi_t = E_\pi[\,\Upsilon_t \,|\, \mathcal{O}_t\,]$ *is given by,*

$$\Pi_t = \Pi_0 + \int_0^t (1 - \Pi_s)\,(1 - F_s)^{-1}\, dF_s + \int_0^t \Pi_s\,(1 - \Pi_s)\, \hat{H}_s\, d\overline{W}_s^v \quad ,$$

for all $t \geq 0$*,* P_π*-a.s., where* $\overline{W}^v$ *is an* $(\mathcal{O}_t, P_\pi)$*-martingale; we recall that by definition* $\hat{H}_t = E_1[H_t|\,\mathcal{O}_t\,]$.

Proof: Applying a filtering theorem [W&H, P7.4.1] we obtain,

$$\Pi_t = E_\pi[\Upsilon_0|\mathcal{O}_0] + \int_0^t E_\pi[\,1 - \Upsilon_{s-}\,|\,\mathcal{O}_s\,]\,(1 - F_s)^{-1}\,dF_s + \int_0^t \Phi_s\,d\overline{W}_s^v \qquad (3.70)$$

for all $t \geq 0$, P_π–a.s., where $\overline{W}^v$ is some $(\mathcal{O}_t, P_\pi)$–martingale and where Φ is an $\mathcal{O}_t$–progressive process given by

$$\Phi_t := E_\pi[\phi_t|\mathcal{O}_t] + E_\pi[\Upsilon_t(\Upsilon_t\,H_t - E_\pi[\Upsilon_t\,H_t|\,\mathcal{O}_t])|\,\mathcal{O}_t] \qquad \forall\, t \geq 0, \qquad (3.71)$$

for some $\mathcal{A}_t$–predictable process ϕ satisfying,

$$\langle M, W^v\rangle_t = \int_0^t \phi_s\,ds \qquad \forall\, t \geq 0,\ P_\pi\text{–a.s.} \qquad (3.72)$$

Let's compute Φ starting with $E_\pi[\phi_t|\mathcal{O}_t]$. Recall that $\langle M, W^v\rangle$ denotes the predictable compensator of $[M, W^v]$, the co-quadratic variation of MW^v. Observe that the additive noise W^v in the observation Y has no jumps and thus it *a fortiori* has no jumps in common with M, the zero-mean martingale driving Υ which has sample paths of locally finite variation, indeed, each a single jump from zero to one. Therefore, $[M, W^v] \equiv \langle M, W^v\rangle \equiv 0$. Hence we can take $\phi \equiv 0$ so that $E_\pi[\phi_t|\mathcal{O}_t] \equiv 0$.

Next, we use the fact that $\Upsilon^2 \equiv \Upsilon$ and find,

$$E_\pi[\Upsilon_t(\Upsilon_t\,H_t - E_\pi[\Upsilon_t\,H_t|\,\mathcal{O}_t])|\,\mathcal{O}_t] = E_\pi[\Upsilon_t\,H_t|\,\mathcal{O}_t]\,(1 - \Pi_t). \qquad (3.73)$$

As for $E_\pi[\Upsilon_t\,H_t|\,\mathcal{O}_t]$, pick any O in $\mathcal{O}_t$ and using expression 2.106 compute,

$$E_\pi[\Upsilon_t\,1_O\,H_t] \;=\; \pi\,E_1 1_O\,H_t + (1-\pi)\,E_0 \int_0^t \int_O H_t\,dQ^u\,dF_u$$

$$
\begin{aligned}
&= \pi\, E_1[1_O\, E_1[\, H_t \,|\, \mathcal{O}_t]\,] + (1-\pi)\, E_0 \int_0^t \int_O L_u^{-1}\, H_t\, dP_1\, dF_u \\
&= \pi\, E_1 1_O\, \hat{H}_t + (1-\pi)\, E_0 \int_0^t \int_O L_u^{-1}\, \hat{H}_t\, dP_1\, dF_u \\
&= \pi\, E_1 1_O\, \hat{H}_t + (1-\pi)\, E_0 \int_0^t \int_O \hat{H}_t\, dQ^u\, dF_u \\
&= E_\pi[\Upsilon_t\, 1_O\, \hat{H}_t] \\
&= E_\pi[\Pi_t\, 1_O\, \hat{H}_t] \qquad (3.74)
\end{aligned}
$$

and therefore,

$$E_\pi[\Upsilon_t\, H_t |\, \mathcal{O}_t] = \Pi_t\, \hat{H}_t \qquad \forall\, t \geq 0,\ P_\pi\text{–a.s.} \qquad (3.75)$$

Combining these results we obtain,

$$\Phi_t = \Pi_t\, (1 - \Pi_t)\, \hat{H}_t \qquad \forall\, t \geq 0,\ P_\pi\text{–a.s.} \qquad (3.76)$$

Next we observe,

$$E_\pi[1 - \Upsilon_{t-} |\, \mathcal{O}_t] = 1 - \Pi_{t-} = 1 - \Pi_t \quad \forall\, t \geq 0,\ P_\pi\text{–a.s.}, \qquad (3.77)$$

and finally,

$$E_\pi[\Upsilon_0 | \mathcal{O}_0] = P_\pi\{\Upsilon_0 = 1 |\, \mathcal{O}_0\} = \Pi_0 \quad P_\pi\text{–a.s.} \qquad (3.78)$$

Collecting these results and plugging them into 3.70 yields the desired expression. □

3.5 Prelude to Verification

Letting $\mathcal{BC}^2(\mathrm{I})$ denote the class of all functions which are twice continuously differentiable on $\mathrm{I} \in \mathcal{I}_\infty$ as well as bounded there, define

$$\mathcal{BC}^2(\mathcal{I}_\infty) := \bigcap_{\mathrm{I} \in \mathcal{I}_\infty} \mathcal{BC}^2(\mathrm{I}). \tag{3.79}$$

An example of a function in this class is the mapping $\pi \mapsto \pi^2$ for $\pi \in [0,1]$. Using the filter for Π developed in the last section, we wish to compute the expectation,

$$E_\pi[\, r(\Pi_\tau) - r(\Pi_0)\,], \tag{3.80}$$

for all $\tau \in \mathcal{T}_{\mathrm{m}}$ and all mappings $r \in \mathcal{BC}^2(\mathcal{I}_\infty)$ in anticipation of applying the Verification Theorem. The main result of this section rests on the next two lemmas.

Lemma 3.3 *Let $\tau \in \mathcal{T}_{\mathrm{m}}$. Then,*

$$E_\pi \int_0^\tau \Pi_s^2 \, \hat{H}_s^2 \, ds < \infty \qquad \forall \, \pi \in \mathrm{I}_\infty.$$

Proof: Fix $\pi \in \mathrm{I}_\infty$ and $\tau \in \mathcal{T}_{\mathrm{m}}$; hence $\tau \leq \tau^{\mathrm{I}}$ P_π–a.s. for some I in $\mathcal{I}_\infty$ and moreover $\tau \in \mathcal{T}$ since $\mathcal{T}_{\mathrm{m}} \subset \mathcal{T}_{ad} \subset \mathcal{T}$. Using the Itô stochastic integration formula [W&H, P6.6.2] and Proposition 3.6 yields,

$$\Pi_t^2 - \Pi_0^2 = \int_0^t (1-\Pi_s)^2 \, \Pi_s^2 \, \hat{H}_s^2 \, ds + 2\int_0^t (1-\Pi_s)\, \Pi_s \, (1-F_s)^{-1} \, dF_s + \tilde{M}_t, \tag{3.81}$$

for all $t \geq 0$, P_π–a.s. where we have defined $\tilde{M} = \{\tilde{M}_t\}_{t \geq 0}$ as,

$$\tilde{M}_t := 2\int_0^t (1-\Pi_s)\, \Pi_s^2 \, \hat{H}_s \, d\overline{W}_s^v \quad \forall \, t \geq 0. \tag{3.82}$$

We note that the integrand in the definition of the stochastic integral $\tilde{M}$ is an $\mathcal{O}_t$–progressive and locally bounded process while the integrator, $\overline{W}^{\upsilon}$, is an $(\mathcal{O}_t, P_\pi)$–martingale. Hence, $\tilde{M}$ is an $(\mathcal{O}_t, P_\pi)$–local martingale and there exists an increasing sequence of $\mathcal{O}_t$–stopping times, $\{\mu_n\}_{n\geq 1}$, converging to infinity P_π–a.s. and such that $\{\tilde{M}_{t\wedge\mu_n}\}_{t\geq 0}$ is an $(\mathcal{O}_t, P_\pi)$–martingale. As a result, we can employ a strong version of Doob's Optional Sampling Theorem [E, T4.12] to find,

$$E_\pi[\tilde{M}_{\tau\wedge\mu_n}] = 0 \qquad \forall\, n \geq 1. \tag{3.83}$$

In addition, the representation in 3.81 in fact holds for any $\tau \in \mathcal{T}$, i.e., for any $\mathcal{O}_t$–stopping time which is P_π–a.s. finite [L&S1, p121]. Therefore we may write,

$$\begin{aligned} \Pi^2_{\tau\wedge\mu_n} - \Pi^2_0 \;&=\; 2\int_0^{\tau\wedge\mu_n} (1-\Pi_s)\,\Pi_s\,(1-F_s)^{-1}\,dF_s \\ &\quad + \int_0^{\tau\wedge\mu_n} (1-\Pi_s)^2\,\Pi_s^2\,\hat{H}_s^2\,ds + \tilde{M}_{\tau\wedge\mu_n}. \end{aligned} \tag{3.84}$$

Taking expectations on both sides gives,

$$\begin{aligned} E_\pi[\,\Pi^2_{\tau\wedge\mu_n} - \Pi^2_0\,] \;&=\; 2\,E_\pi\int_0^{\tau\wedge\mu_n} (1-\Pi_s)\,\Pi_s\,(1-F_s)^{-1}\,dF_s \\ &\quad + E_\pi\int_0^{\tau\wedge\mu_n} (1-\Pi_s)^2\,\Pi_s^2\,\hat{H}_s^2\,ds, \end{aligned} \tag{3.85}$$

and from this there follows,

$$1 \geq E_\pi\int_0^{\tau\wedge\mu_n} (1-\Pi_s)^2\,\Pi_s^2\,\hat{H}_s^2\,ds \geq (1-\sup[\,\mathrm{I}\,]_\Pi)^2 E_\pi\int_0^{\tau\wedge\mu_n} \Pi_s^2\,\hat{H}_s^2\,ds. \tag{3.86}$$

As a result,

$$E_\pi\int_0^{\tau\wedge\mu_n} \Pi_s^2\,\hat{H}_s^2\,ds \leq (1-\sup[\,\mathrm{I}\,]_\Pi)^{-2} < \infty, \tag{3.87}$$

since $\sup[\mathrm{I}]_{\Pi} < 1$ as guaranteed by (E). Hence, applying the Monotone Convergence Theorem yields,

$$E_{\pi} \int_0^{\tau} \Pi_s^2 \, \hat{H}_s^2 \, ds \le (1 - \sup[\mathrm{I}]_{\Pi})^{-2} < \infty, \tag{3.88}$$

and the lemma is shown. □

To simplify the notation in the next lemma, for each $t \ge 0$ define the $\mathcal{O}_t$–measurable linear stochastic differential operator $D_{\mathcal{O}_t}$ for all functions $r \in \mathcal{C}^2(\mathrm{I}_{\infty})$ via,

$$D_{\mathcal{O}_t} \, r(\pi) := \frac{1}{2} \, (1 - \pi)^2 \, \pi^2 \, r''(\pi) \, \hat{H}_t^2 \, dt + (1 - \pi) \, r'(\pi) \, (1 - F_t)^{-1} \, dF_t \tag{3.89}$$

Lemma 3.4 *Let $\{\sigma_n\}$ denote an increasing sequence of $\mathcal{O}_t$–stopping times which converge to infinity P_{π}–a.s. and suppose that $r \in \mathcal{BC}^2(\mathcal{I}_{\infty})$. Then,*

$$E_{\pi} \int_0^{\tau \wedge \sigma_n} D_{\mathcal{O}_s} \, r(\Pi_s) \longrightarrow E_{\pi} \int_0^{\tau} D_{\mathcal{O}_s} \, r(\Pi_s) \qquad \forall \, \tau \in \mathcal{T}_{\mathrm{m}}, \; \pi \in \mathrm{I}_{\infty}.$$

Proof: Fix $\pi \in \mathrm{I}_{\infty}$ and τ in $\mathcal{T}_{\mathrm{m}}$; hence $\tau \le \tau^{\mathrm{I}}$ P_{π}–a.s. for some I in $\mathcal{I}_{\infty}$. Define,

$$X_t := \int_0^t (1 - \Pi_s)^2 \, \Pi_s^2 \, r''(\Pi_s) \, \hat{H}_s^2 \, ds, \tag{3.90}$$

and

$$Z_t := \int_0^t (1 - \Pi_s) \, r'(\Pi_s) \, (1 - F_s)^{-1} \, dF_s, \tag{3.91}$$

and note that,

$$\int_0^t D_{\mathcal{O}_s} \, r(\Pi_s) = \frac{1}{2} X_t + Z_t \qquad \forall \, t \ge 0, \; P_{\pi}\text{–a.s.} \tag{3.92}$$

We will show that $E_\pi X_{\tau\wedge\sigma_n} \to E_\pi X_\tau$ and $E_\pi Z_{\tau\wedge\sigma_n} \to E_\pi Z_\tau$. Then, because

$$E_\pi \int_0^{\tau\wedge\sigma_n} D_{\mathcal{O}_s} r(\Pi_s) = \frac{1}{2} E_\pi X_{\tau\wedge\sigma_n} + E_\pi Z_{\tau\wedge\sigma_n} \qquad P_\pi\text{–a.s.,} \tag{3.93}$$

the theorem will follow. We rule out a trivial case: if $\pi \notin \mathrm{I}$ then $\tau^{\mathrm{I}} \equiv 0$, and $\tau \wedge \sigma_n = 0$ for all $n \geq 1$. Thus, both sides of the above expression are zero for all $n \geq 1$ as are the limits.

Part 1.

Let $\pi \in \mathrm{I}$; define $V := (1-\Pi)^2\,\Pi^2\,\hat{H}^2\,r''(\Pi)$ and note,

$$X_{\tau\wedge\sigma_n} = \int_0^\infty 1\{s \leq \tau \wedge \sigma_n\}\, V_s\, ds \quad P_\pi\text{–a.s.} \tag{3.94}$$

Since $\mathrm{I} \in \mathcal{I}_\infty$ the condition (E) implies that $[\,\mathrm{I}\,]_\Pi$ is a proper subset of I_∞. From this and the fact that $r \in \mathcal{BC}^2(\mathcal{I}_\infty)$ it follows that,

$$r''(\pi) \leq B''_{\mathrm{I}} \quad \forall\, \pi \in [\,\mathrm{I}\,]_\Pi, \tag{3.95}$$

for some bound B''_{I} which is finite. From the definition of $[\,\mathrm{I}\,]_\Pi$ and the assumption that $\pi \in [\,\mathrm{I}\,]_\Pi$ we know that $\Pi_t \in [\,\mathrm{I}\,]_\Pi$, P_π–a.s., for all t in the stochastic interval $[0, \tau^{\mathrm{I}}]$. Then since $\tau \leq \tau^{\mathrm{I}}$, P_π–a.s., we can conclude,

$$r''(\Pi_{\tau\wedge\sigma_n}) \leq B''_{\mathrm{I}} < \infty \quad P_\pi\text{–a.s.,} \quad \forall\, n \geq 1. \tag{3.96}$$

With this in mind note that

$$\begin{aligned} |\,1\{s \leq \tau \wedge \sigma_n\}\, V_s\,| &\leq 1\{s \leq \tau\}\,|\,V_s\,| \\ &\leq 1\{s \leq \tau\}\, B''_{\mathrm{I}}\, \Pi_s^2\, \hat{H}_s^2, \end{aligned} \tag{3.97}$$

and also,

$$\int_0^\infty 1\{s \leq \tau\}\, B''_{\mathrm{I}}\, \Pi_s^2\, \hat{H}_s^2\, ds = B''_{\mathrm{I}} \int_0^\tau \Pi_s^2\, \hat{H}_s^2\, ds < \infty \quad P_\pi\text{–a.s.,} \tag{3.98}$$

where the last bound follows *a fortiori* from Lemma 3.3. Thus, 3.97 and 3.98 taken together imply that the $\mathcal{O}_t$–progressive process $\{\, 1\{t \le \tau\}\, V_t \,\}_{t\ge 0}$ is P_π–a.s. Lebesgue integrable on $[0,\infty]$. Hence, employing the Lebesgue Dominated Convergence Theorem with respect to Lebesgue measure we have,

$$\int_0^\infty 1\{s \le \tau \wedge \sigma_n\}\, V_s\, ds \longrightarrow \int_0^\infty 1\{s \le \tau\}\, V_s\, ds \quad P_\pi\text{–a.s.} \tag{3.99}$$

We can also express this as,

$$\int_0^{\tau\wedge\sigma_n} V_s\, ds \longrightarrow \int_0^\tau V_s\, ds \quad P_\pi\text{–a.s.} \tag{3.100}$$

Likewise,

$$|\int_0^{\tau\wedge\sigma_n} V_s\, ds\,| \le \int_0^\tau |\, V_s\,|\, ds \quad P_\pi\text{–a.s.}, \tag{3.101}$$

and similarly Lemma 3.3 yields,

$$E_\pi \int_0^\tau |\, V_s\,|\, ds \le B_{\mathrm{I}}''\, E_\pi \int_0^\tau \Pi_s^2\, \hat{H}_s^2\, ds < \infty. \tag{3.102}$$

Employing the Lebesgue Dominated Convergence Theorem now with respect to the P_π–measure gives,

$$E_\pi \int_0^{\tau\wedge\sigma_n} V_s\, ds \longrightarrow E_\pi \int_0^\tau V_s\, ds, \tag{3.103}$$

and we conclude,

$$E_\pi X_{\tau\wedge\sigma_n} \longrightarrow E_\pi X_\tau. \tag{3.104}$$

This gives us the first half of what we want for all $\tau \in \mathcal{T}_{\mathrm{m}}$ and $\pi \in [0,1]$.

Part 2.

Let $\pi \in \mathrm{I}$; redefine $V := (1-\Pi)\, r'(\Pi)\,(1-F)^{-1}$ and note,

$$Z_{\tau\wedge\sigma_n} = \int_0^\infty 1\{s \le \tau \wedge \sigma_n\}\, V_s\, dF_s \quad P_\pi\text{–a.s.} \tag{3.105}$$

Arguing as in Part 1 for the second derivative we conclude here for the first derivative that,

$$r'(\Pi_{\tau\wedge\sigma_n}) < B'_{\mathrm{I}} \qquad P_\pi\text{–a.s.}, \quad \forall\, n \geq 1. \tag{3.106}$$

for some bound B'_{I} which is finite. With this in mind we obtain,

$$E_\pi \int_0^\tau V_s\, dF_s \leq B'_{\mathrm{I}}\, E_\pi \int_0^\tau (1 - F_s)^{-1}\, dF_s < \infty \quad P_\pi\text{–a.s.}, \tag{3.107}$$

where the finiteness follows from technical assumption (F) together with expression 2.96 and the fact that $P_0^{\mathcal{O}} \equiv P_1^{\mathcal{O}}$. Continuing the analog of the argument in Part 1 yields,

$$\int_0^{\tau\wedge\sigma_n} V_s\, dF_s \longrightarrow \int_0^\tau V_s\, dF_s \quad P_\pi\text{–a.s.}, \tag{3.108}$$

and then an appeal to the Lebesgue Dominated Convergence Theorem gives,

$$E_\pi \int_0^{\tau\wedge\sigma_n} V_s\, dF_s \longrightarrow E_\pi \int_0^\tau V_s\, dF_s, \tag{3.109}$$

and we conclude,

$$E_\pi Z_{\tau\wedge\sigma_n} \longrightarrow E_\pi Z_\tau. \tag{3.110}$$

Combining the results of Part 1 and Part 2 in the obvious way proves the lemma. □

We have the following proposition.

Proposition 3.7 *Let $r \in \mathcal{BC}^2(\mathcal{I}_\infty)$. Then,*

$$E_\pi[r(\Pi_\tau) - r(\Pi_0)] = E_\pi \int_0^\tau D_{\mathcal{O}_s}\, r(\Pi_s) \qquad \forall\, \tau \in \mathcal{T}_{\mathrm{m}},\ \pi \in \mathrm{I}_\infty.$$

Proof: Fix $\pi \in \mathrm{I}_\infty$. For $r \in \mathcal{BC}^2(\mathcal{I}_\infty)$ the Itô formula yields *a fortiori*,

$$r(\Pi_t) - r(\Pi_0) = \int_0^t r'(\Pi_s)\, d\Pi_s + \frac{1}{2}\int_0^t r''(\Pi_s)\, d[\Pi,\Pi]_s, \qquad (3.111)$$

for all $t \geq 0$, P_π–a.s. Using Proposition 3.6 we compute the quadratic variation,

$$\begin{aligned}
[\Pi,\Pi]_t &= \int_0^t (1-\Pi_s)\,\Pi_s\,\hat{H}_s\, d[\overline{W}^v, ((1-\Pi)\,\Pi\,\hat{H}) \bullet \overline{W}^v]_s \\
&= \int_0^t (1-\Pi_s)^2\,\Pi_s^2\,\hat{H}_s^2\, d[\overline{W}^v, \overline{W}^v]_s \\
&= \int_0^t (1-\Pi_s)^2\,\Pi_s^2\,\hat{H}_s^2\, ds \quad \forall\, t \geq 0,\ P_\pi\text{–a.s.}, \qquad (3.112)
\end{aligned}$$

and the stochastic differential,

$$d\Pi_t = (1-\Pi_t)(1-F_t)^{-1}\, dF_t + \Pi_t\,(1-\Pi_t)\,\hat{H}_t\, d\overline{W}_t^v. \qquad (3.113)$$

Substituting 3.112 and 3.113 into 3.111 we obtain,

$$\begin{aligned}
r(\Pi_t) - r(\Pi_0) &= \int_0^t (1-\Pi_s)\,(1-F_s)^{-1}\, r'(\Pi_s)\, dF_s \\
&\quad + \int_0^t (1-\Pi_s)\,\Pi_s\,\hat{H}_s\, r'(\Pi_s)\, d\overline{W}_s^v \\
&\quad + \frac{1}{2}\int_0^t (1-\Pi_s)^2\,\Pi_s^2\,\hat{H}_s^2\, r''(\Pi_s)\, ds, \qquad (3.114)
\end{aligned}$$

for all $t \geq 0$, P_π–a.s. With the $(\mathcal{O}_t, P_\pi)$–local martingale $\tilde{M} = \{\tilde{M}_t\}_{t\geq 0}$ defined here as,

$$\tilde{M}_t := \int_0^t (1-\Pi_s)\,\Pi_s\,\hat{H}_s\, r'(\Pi_s)\, d\overline{W}_s^v \quad \forall\, t \geq 0, \qquad (3.115)$$

and using the definition of $D_{\mathcal{O}_t}$ given in 3.89 we can rewrite 3.114 more compactly as,

$$r(\Pi_t) - r(\Pi_0) = \int_0^t D_{\mathcal{O}_s}\, r(\Pi_s) + \tilde{M}_t \quad \forall\, t \geq 0,\ P_\pi\text{–a.s.} \qquad (3.116)$$

From Lemma 3.4 we obtain,

$$\lim_{n\to\infty} E_\pi[\, r(\Pi_{\tau\wedge\mu_n}) - r(\Pi_0)\,] = E_\pi \int_0^\tau D_{\mathcal{O}_s} r(\Pi_s), \tag{3.117}$$

where again $\{\mu_n\}_{n\geq 1}$ denotes a localization sequence for $\tilde{M}$. Since $\tau \in \mathcal{T}_m$ there exists $\mathrm{I} \subset \mathrm{I}_\infty$ such that $\tau \leq \tau^{\mathrm{I}}$, P_π–a.s. and so,

$$|\, r(\Pi_{\tau\wedge\mu_n})\,| \leq |\, r(\Pi_\tau)\,| \leq B_{\mathrm{I}} < \infty, \tag{3.118}$$

for some bound B_{I} known to exist since $r \in \mathcal{BC}(\mathrm{I}_\infty)$ *a fortiori*. Finally, employing the Bounded Convergence theorem and an argument paralleling the one used in Lemma 2.1 yields,

$$\lim_{n\to\infty} E_\pi[\, r(\Pi_{\tau\wedge\mu_n})\,] = E_\pi[\, r(\Pi_\tau)\,], \tag{3.119}$$

and we get what we want in 3.117 on the left-hand side. □

3.6 Sequential Detection

In this section we will formulate a classical Bayesian sequential binary hypothesis testing or *sequential detection* problem on the drift of an observed diffusion-type stochastic process. We will show how it can be recast as a *change detection* problem within our framework and then tackled via the Verification Theorem. Recast in this way, the classical problem of sequential detection is given a fresh interpretation as a problem of *lack-of-change* detection. The Verification Theorem will lead us to consider a type of free-boundary value or *Stefan* problem whose solution is addressed using techniques from ODE theory and convex analysis. We will end the section with an example involving a diffusion with constant drift.

3.6.1 Problem Statement

On a measurable space $(\Omega, \mathcal{G})$ equipped with two mutually absolutely continuous probability measures P_0 and P_1 we observe a stochastic process $Y = \{Y_t\}_{t \geq 0}$ for which one of the following hypotheses is true:

$$\begin{aligned} \text{(Noise Only)}: \quad & Y_t = W_t & 0 \leq t < \infty; \\ \text{(Signal Plus Noise)}: \quad & Y_t = \int_0^t H_s \, ds + W_t & 0 \leq t < \infty, \end{aligned}$$

where W is a $(\mathcal{G}_t, P_i)$–standard Wiener martingale for $i = 0, 1$, and $H = \{H_t\}_{t \geq 0}$ is a $\mathcal{G}_t$–progressive process satisfying,

$$E_i \int_0^t H_s^2 \, ds < \infty \qquad \forall\, t \geq 0,\ i = 0, 1. \tag{3.120}$$

Following the Bayesian philosophy, we are told in advance that the (Signal Plus Noise) hypothesis occurs with prior probability $\pi \in [0, 1]$. We are tasked with deciding in a sequential manner which hypothesis is indeed responsible for what is observed. Decision policies are defined as a pair, (τ, δ), where τ is a stopping time with respect to the P_0–completed observation filtration $\mathcal{O}_t = \bigvee_{s \leq t} \sigma(Y_s)$ and δ is a binary random variable representing our decision and therefore measurable with respect to $\mathcal{O}_\tau$. We are told that the goodness of any sequential decision policy, (τ, δ), is judged according to the following 'elapsed-energy + incorrect decision' cost criterion,

$$\begin{aligned}\bar{\rho}_\pi(\tau, \delta) &= \pi\, E_1[\int_0^\tau c\, \hat{H}_s^2\, ds + c^0\, 1\{\delta = 0\}\,] \\ &\quad + (1-\pi)\, E_0[\int_0^\tau c\, \hat{H}_s^2\, ds + c^1\, 1\{\delta = 1\}\,], \qquad (3.121)\end{aligned}$$

where c, c^0, and c^1 are strictly positive and finite. We are asked to minimize $\bar{\rho}_\pi(\tau, \delta)$ over all decision pairs. We are told to restrict our attention to those $\mathcal{O}_t$–stopping times which satisfy,

$$E_i \int_0^\tau \hat{H}_s^2\, ds < \infty, \quad i = 0, 1, \qquad (3.122)$$

and in addition we are given the following technical condition on the running cost:

$$P_i\,\{\int_0^\infty \hat{H}_s^2\, ds = \infty\,\} = 1, \quad i = 0, 1. \qquad (3.123)$$

3.6.2 Problem Reformulation

We now show how the sequential detection problem can be recast as a problem of change detection on the probability space $(\Omega, \mathcal{A}, P_\pi)$ with π the same prior as in the previous subsection. Drawing freely upon our earlier results and notation, the first step is to notice that the (Noise Only) and (Signal Plus

Noise) hypotheses are captured by the model dynamics (DM0) and (DM1), respectively. A little consideration shows that the *sequential* aspect of sequential detection can be recovered within the change detection format by properly specifying the random disruption time υ previously characterized only up to the nature of $F = \{F_t\}_{t\geq 0}$, its $\mathcal{O}_t$–adapted, conditional cumulative distribution function with respect to the P_0–measure. To this end make the following sequential detection modeling assumption:

$$(\mathrm{F_{SD}}) : \qquad F_t := 1\{t = \infty\} \quad \forall\, t \in [0, \infty].$$

We see that F is a legitimate $\mathcal{O}_t$–adapted, conditional cumulative distribution function on $[0, \infty] = \Re_+ \cup \{\infty\}$ satisfying the requirements set down in Chapter 2: $F_0 = 0$, F is an increasing function on $[0, \infty]$ which is deterministic and therefore trivially $\mathcal{O}_t$–adapted, $F_\infty = 1$, and lastly, F satisfies condition (F). Under assumption ($\mathrm{F_{SD}}$) we have therefore,

$$P_0\{\,\upsilon \leq t \,|\, \mathcal{O}_t\,\} = P_0\{\,\upsilon \leq t\,\} = 0 \quad \forall\, t \in [0, \infty), \tag{3.124}$$

and $P_0\{\,\upsilon = \infty\,\} = 1$. With this choice of F, the P_0–measure gives all its probability to the event $\{\,\upsilon = \infty\,\}$; the P_1–measure of course still gives all its probability to the event $\{\,\upsilon = 0\,\}$. An immediate consequence of this is that the P–measure defined in 2.97 is precisely the P_0–measure and as a result the P_π-measure simplifies to,

$$\begin{aligned} P_\pi\{A\} &= \pi\, P_1\{A\} + (1-\pi)\, P\{A\} \\ &= \pi\, P_1\{A\} + (1-\pi)\, P_0\{A\}, \end{aligned} \tag{3.125}$$

for all π in $[0, 1]$ and A in $\mathcal{A}$. From this we see,

$$P_\pi\{0 < \upsilon < \infty\} = \pi\, P_1\{0 < \upsilon < \infty\} + (1-\pi)\, P_0\{0 < \upsilon < \infty\} = 0, \tag{3.126}$$

or, what is equivalent $P_\pi\{\,v = 0\,\} = \pi$ and $P_\pi\{\,v = \infty\,\} = 1 - \pi$.

The ($\mathrm{F_{SD}}$) assumption also has the consequence of simplifying the semimartingale representation for Υ which is obviously now reduced to $\Upsilon_t = \Upsilon_0$, P_π–a.s., for all $t \geq 0$. Using this fact we have $E_\pi\mathcal{E}(\Upsilon_\tau, \delta) = E_\pi\mathcal{E}(\Upsilon_0, \delta)$ for all τ in $\mathcal{T}_{ad}$ an $\mathcal{T}$. From this we compute,

$$
\begin{aligned}
E_\pi\mathcal{E}(\Upsilon_\tau, \delta) &= E_\pi\mathcal{E}(\Upsilon_0, \delta) \\
&= \pi\, E_1\mathcal{E}(\Upsilon_0, \delta) + (1-\pi)\, E_0\mathcal{E}(\Upsilon_0, \delta) \\
&= \pi\, E_1\mathcal{E}(1, \delta) + (1-\pi)\, E_0\mathcal{E}(0, \delta) \\
&= \pi\, E_1[\, c^0\, 1\{\delta = 0\} \cdot 1 + 0\,] + (1-\pi)\, E_0[\, 0 + c^1\, 1\{\delta = 1\} \cdot 1\,] \\
&= \pi\, E_1[\, c^0\, 1\{\delta = 0\}\,] + (1-\pi)\, E_0[\, c^1\, 1\{\delta = 1\}\,]. \qquad (3.127)
\end{aligned}
$$

Next, we choose the running cost $\mathcal{C}$ according to,

$$
\mathcal{C}_t := c\,\hat{H}_t^2 \quad \forall\, t \geq 0, \qquad (3.128)
$$

and note that $\mathcal{C}$ satisfies the cost assumptions (C1), (C2) and (C3): (C1) and (C2) follow directly from (H1), (H2) and 3.125 while (C3) follows since $\mathcal{C}$ is trivially concave in Π. Now compute Bayes' cost,

$$
\begin{aligned}
\rho_\pi(\tau, \delta) &= E_\pi[\int_0^\tau \mathcal{C}_s\, ds + \mathcal{E}(\Upsilon_\tau, \delta)\,] \\
&= \pi\, E_1[\int_0^\tau \mathcal{C}_s\, ds + \mathcal{E}(\Upsilon_0, \delta)\,] \\
&\qquad + (1-\pi)\, E_0[\int_0^\tau \mathcal{C}_s\, ds + \mathcal{E}(\Upsilon_0, \delta)\,] \\
&= \pi\, E_1[\int_0^\tau c\,\hat{H}_s^2\, ds + c^0\, 1\{\delta = 0\}\,] \\
&\qquad + (1-\pi)\, E_0[\int_0^\tau c\,\hat{H}_s^2\, ds + c^1\, 1\{\delta = 1\}\,] \\
&= \bar{\rho}_\pi(\tau, \delta). \qquad (3.129)
\end{aligned}
$$

Hence, this choice of running cost and assumption (F_{SD}) reduce the change detection problem as defined in Chapter 2 to a classical Bayes sequential detection problem. Indeed, when viewed as a change detection problem, a sequential detection problem is seen to be the extremal case: the *lack-of-change* change detection problem. We can therefore solve the sequential detection problem by solving the optimal stopping problem ($\mathcal{P}$) of Chapter 2 under the conditions prevailing in the current section. The plan of course is to solve the optimal stopping problem using a first exit policy found by recourse to the Verification Theorem. We end this subsection with a convenient and complete reformulation of the sequential detection problem, taking into account the nature of the observations, the sequential detection assumption, and the choice of running cost.

On the probability space $(\Omega, \mathcal{A}, P_\pi)$ with $\pi \in [0,1]$ given, choose $F_t = 1\{v = \infty\}$ so that the jump time v obeys,

$$P_\pi\{v = 0\} = \pi = 1 - P_\pi\{v = \infty\}. \tag{3.130}$$

Since $\Upsilon_t = 1\{v \leq t\}$ we note that $\Upsilon_t = \Upsilon_0$ for all $t \geq 0$, P_π–a.s. Thus, there exist only two possible sample paths upon which P_π is concentrated. The observations are given by,

$$Y_t = \int_0^t \Upsilon_s H_s \, ds + W_t^v \qquad \forall\, t \geq 0, \tag{3.131}$$

where the drift H satisfies 3.120, 3.122 and 3.123 in view of the technical assumptions (H0), (H1), and (H2) imposed in the last section. Due to the essentially two-valued nature of v we see that,

$$Y_t = \Upsilon_0 \int_0^t H_s \, ds + W_t^v \qquad \forall\, t \geq 0, \tag{3.132}$$

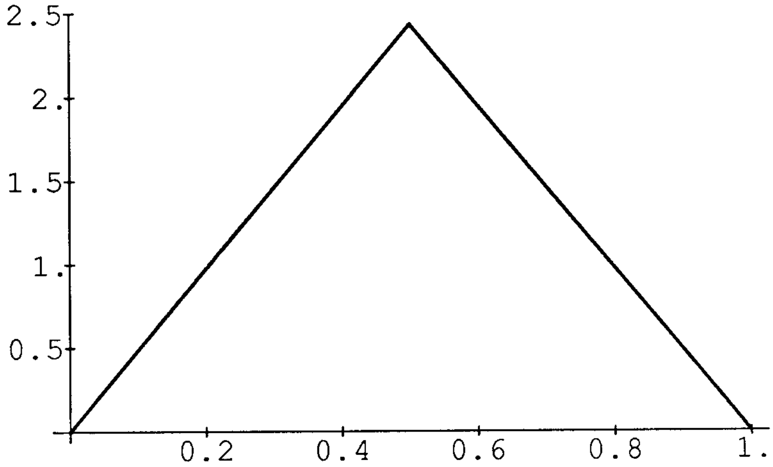

Figure 3.1: Terminal cost function. $c^0 = c^1 = 3 - \frac{1}{3} + 2\log 3$, $\pi_e = \frac{1}{2}$.

and therefore,

$$Y_t = \begin{cases} W_t^v & 0 \le t < v = \infty; \qquad \text{(Noise Only)} \\ \\ \int_0^t H_s\, ds + W_t^v & 0 = v \le t < \infty. \quad \text{(Signal Plus Noise)} \end{cases} \tag{3.133}$$

Hence, choosing F as in ($\mathrm{F_{SD}}$) captures the classical sequential detection set-up within the change detection set-up. Bayes' cost is taken to be,

$$\rho_\pi(\tau) = E_\pi[\int_0^\tau c\hat{H}_s^2\, ds + e(\Pi_\tau)], \tag{3.134}$$

with $c > 0$ and where the terminal cost, e, is given by,

$$e(\pi) = \min\{c^0\pi, c^1(1-\pi)\} \qquad \forall \pi \in [0,1], \tag{3.135}$$

with $0 < c^0, c^1 < \infty$ (see Figure 3.1). For this choice of running cost an admissible stopping time τ is any $\mathcal{O}_t$–stopping time satisfying,

$$E_\pi \int_0^\tau \mathcal{C}_s\, ds = E_\pi \int_0^\tau c\hat{H}_s^2\, ds < \infty. \tag{3.136}$$

Thus, the change detection problem to solve which is equivalent to the sequential detection problem is,

$$(\mathcal{P}_{\text{SD}}): \qquad \text{Find } \tau_* \in \mathcal{T}_{ad} \text{ such that } \rho_\pi(\tau_*) = \inf_{\tau \in \mathcal{T}_{ad}} \rho_\pi(\tau). \tag{3.137}$$

3.6.3 Escape Properties

The next question to answer is whether problem $(\mathcal{P}_{\text{SD}})$ is sufficiently well-posed that the escape condition (E) is satisfied and then which of (E^+) or (E^0) holds. We note that (F_{SD}) implies that $dF_t = 0$ for all $t \geq 0$, P_π–a.s., so that $\Upsilon \equiv \Upsilon_0$ and the stochastic differential description of the sample paths of Π reduces to,

$$\Pi_t = \Pi_0 + \int_0^t \Pi_s (1 - \Pi_s) \hat{H}_s \, d\overline{W}^v_s. \tag{3.138}$$

In the Proposition below we show that condition (H2) implies that (E^+) holds so that $\mathcal{I}_\infty = \mathcal{I}^+$ and $\text{I}_\infty = (0,1)$ and thus the trajectories of Π are guaranteed to escape any proper subinterval of $(0,1)$.

Proposition 3.8 *Under assumption* (H2) *the escape condition* (E^+) *holds, i.e., given*

$$P_i\{ \int_0^\infty \hat{H}_s^2 \, ds = \infty \} = 1 \qquad i = 0, 1,$$

then,

$$P_\pi\{\tau^{\text{I}} < \infty\} = 1 \qquad \forall \pi \in \text{I}, \, \forall \text{I} \in \mathcal{I}^+.$$

Proof: Choose any I in $\mathcal{I}^+$ so that $\text{I} = (a, b)$ with $0 < a < b < 1$ and suppose $\pi \in \text{I}$. From the Itô product formula and 3.138 we obtain,

$$\Pi_t^2 - \Pi_0^2 = 2 \int_0^t \Pi_s^2 (1 - \Pi_s) \hat{H}_s \, d\overline{W}^v_s + [\Pi, \Pi]_t, \tag{3.139}$$

and we see that the first term on the right is an $(\mathcal{O}_t, P_\pi)$–local martingale with a localization sequence, say $\{\mu_n\}_{n\geq 1}$. Using by now familiar arguments we obtain,

$$E_\pi[\,\Pi^2_{\tau^{\rm I}\wedge\mu_n} - \Pi^2_0\,] = E_\pi[\Pi,\Pi]_{\tau^{\rm I}\wedge\mu_n} = E_\pi \int_0^{\tau^{\rm I}\wedge\mu_n} \Pi^2_s\,(1-\Pi_s)^2\,\hat{H}^2_s\,ds. \quad (3.140)$$

Obviously,

$$E_\pi[\,\Pi^2_{\tau^{\rm I}\wedge\mu_n} - \Pi^2_0\,] \leq 1. \quad (3.141)$$

Moreover, since π is in $\mathrm{I} = (a,b)$ we know that $\Pi_{\tau^{\rm I}\wedge\mu_n}$ is in $[\,\mathrm{I}\,]_\Pi$, P_π–a.s., and therefore,

$$E_\pi \int_0^{\tau^{\rm I}\wedge\mu_n} \Pi^2_s\,(1-\Pi_s)^2\,\hat{H}^2_s\,ds \geq a^2\,(1-b)^2 E_\pi \int_0^{\tau^{\rm I}\wedge\mu_n} \hat{H}^2_s\,ds. \quad (3.142)$$

Combining these results yields,

$$E_\pi \int_0^{\tau^{\rm I}\wedge\mu_n} \hat{H}^2_s\,ds \leq \frac{1}{a^2\,(1-b)^2} < \infty. \quad (3.143)$$

Using additional familiar arguments we pass to the limit and obtain,

$$E_\pi \int_0^{\tau^{\rm I}} \hat{H}^2_s\,ds \leq \frac{1}{a^2\,(1-b)^2} < \infty, \quad (3.144)$$

and thus,

$$\infty > E_\pi \int_0^{\tau^{\rm I}} \hat{H}^2_s\,ds \geq E_\pi[\,1\{\tau^{\rm I} = \infty\} \int_0^{\infty} \hat{H}^2_s\,ds\,]. \quad (3.145)$$

This last line yields a contradiction unless $P_\pi\{\tau^{\rm I} = \infty\} = 0$ because it is obvious from (H2) that $P_\pi\{\,\int_0^\infty \hat{H}^2_s\,ds = \infty\,\} = 1$ for all π in I. □

Motivated by this proposition we are justified in searching for $\mathrm{I}_* \in \mathcal{I}^+$ so that if we write $\mathrm{I}_* = (a_*, b_*)$ then $0 < a_* < b_* < 1$ and moreover, in view of Lemma 2.4, we can expect that $a_* \leq \pi_e \leq b_*$. Before embarking on our search for I_* however, we consider one more result which is companion to the last. Indeed, while the last proposition shows (H2) implies that the Π process is guaranteed to escape any interval in $\mathcal{I}_\infty$, in the following proposition we will show (H1) guarantees that Π can only escape $I_\infty = (0,1)$ in infinite time.

Proposition 3.9 *Under assumption* (H1),

$$P_\pi\{\tau^{I_\infty} = \infty\} = 1 \qquad \forall \pi \in I_\infty.$$

Proof: Fix $\pi \in I_\infty = (0,1)$. From Proposition 3.5 we have

$$P_\pi\{\sup_{0 \leq t \leq n} |\log \check{L}_t| = \infty\} = 0 \quad \forall n \geq 1, \ \forall \pi \in I_\infty. \tag{3.146}$$

This fact is sufficient to imply that $P_\pi\{\tau^{I_\infty} < \infty\} = 0$. Employing the Itô stochastic integration formula one can show for all $\pi \in (0,1)$ that (see [MS, App I]),

$$\check{L} \equiv \frac{1-\pi}{\pi} \frac{\Pi}{1-\Pi}. \tag{3.147}$$

From this it is obvious that,

$$\{|\Pi_{\tau^{I_\infty}} - \frac{1}{2}| = \frac{1}{2}\} = \{|\log \check{L}_{\tau^{I_\infty}}| = \infty\}, \tag{3.148}$$

and therefore,

$$\{\tau^{I_\infty} < \infty\} \subset \{|\log \check{L}_{\tau^{I_\infty}}| = \infty\}. \tag{3.149}$$

Hence,

$$P_\pi\{\tau^{I_\infty} < \infty\} \;=\; P_\pi\{|\log \check{L}_{\tau^{I_\infty}}| = \infty, \tau^{I_\infty} < \infty\}$$

$$
\begin{aligned}
&= P_\pi\{|\log \check{L}_{\tau^{\mathrm{I}_\infty}}| = \infty, \cup_{n\geq 1}\tau^{\mathrm{I}_\infty} \leq n\} \\
&\leq \sum_{n=1}^{\infty} P_\pi\{|\log \check{L}_{\tau^{\mathrm{I}_\infty}}| = \infty, \tau^{\mathrm{I}_\infty} \leq n\} \\
&\leq \sum_{n=1}^{\infty} P_\pi\{\sup_{0\leq t\leq n} |\log \check{L}_t| = \infty, \tau^{\mathrm{I}_\infty} \leq n\} \\
&\leq \sum_{n=1}^{\infty} P_\pi\{\sup_{0\leq t\leq n} |\log \check{L}_t| = \infty\} \\
&= \lim_{m\to\infty} \sum_{n=1}^{m} 0 = 0,
\end{aligned}
\tag{3.150}
$$

i.e., Π has no chance of escaping from $\mathrm{I}_\infty = (0,1)$ in finite time. □

3.6.4 Verification: A Stefan Problem

For easy reference, we state the verification conditions (V1)–(V4) for problem $(\mathcal{P}_{\mathrm{SD}})$ given in the previous subsection with $\mathrm{I}_\infty = (0,1)$, $\mathcal{I}_\infty = \mathcal{I}^+$, and $\mathrm{I}_* \in \mathcal{I}^+$. We also incorporate the proposed running cost given in 3.128:

(V1a) : *For all* $\tau \in \mathcal{T}_{\mathrm{m}}$,

$$
\begin{aligned}
&& E_\pi[r_*(\Pi_\tau) - r_*(\Pi_0)] &\geq -E_\pi \int_0^\tau c\hat{H}_s^2\, ds && \forall \pi \in [\,\mathrm{I}_*\,]_\Pi; \\
&(\mathrm{V1b}): & E_\pi[r_*(\Pi_{\tau^{\mathrm{I}_*}}) - r_*(\Pi_0)] &= -E_\pi \int_0^{\tau^{\mathrm{I}_*}} c\hat{H}_s^2\, ds && \forall \pi \in [\,\mathrm{I}_*\,]_\Pi; \\
&(\mathrm{V2}): & r_*(\pi) &= e(\pi) && \forall \pi \in \partial_\Pi \mathrm{I}_*; \\
&(\mathrm{V3}): & r_*(\pi) &< e(\pi) && \forall \pi \notin \partial_\Pi \mathrm{I}_*;
\end{aligned}
$$

and,

(V4) : r_* is bounded and continuous on $[\,\mathrm{I}_*\,]_\Pi$.

Let us attempt to find $r \in \mathcal{C}^2(0,1)$, an attempt facilitated by the availability of the Itô stochastic integration formula for twice smooth functions and suggested by the fact that $\mathcal{C}^2(0,1) = \mathcal{BC}^2(\mathcal{I}^+)$ as is easy to show. Taking the detection assumption ($\mathrm{F_{SD}}$) into account, the $\mathcal{O}_t$–measurable linear stochastic differential operator $D_{\mathcal{O}_t}$ defined in 3.89 for all mappings $r \in \mathcal{C}^2(0,1)$ simplifies to,

$$D_{\mathcal{O}_t}\, r(\pi) = \frac{1}{2}\,(1-\pi)^2\,\pi^2\,\hat{H}_t^2\, r''(\pi)\,dt. \tag{3.151}$$

With this in mind, define the deterministic linear differential operator D_Π via

$$D_\Pi\, r(\pi) := \frac{1}{2}\,(1-\pi)^2\pi^2 r''(\pi), \tag{3.152}$$

so that Proposition 3.7 states,

$$E_\pi[\,r(\Pi_\tau) - r(\Pi_0)\,] = E_\pi \int_0^\tau \hat{H}_s^2\, D_\Pi r(\Pi_s)\,ds \quad \forall\, \tau \in \mathcal{T}_{\mathrm{m}}. \tag{3.153}$$

Hence, we can rewrite (V1a) as,

$$E_\pi \int_0^\tau \hat{H}_s^2\, D_\Pi\, r_*(\Pi_s)\,ds \geq E_\pi \int_0^\tau \hat{H}_s^2(-c)\,ds \quad \forall\, \tau \in \mathcal{T}_{\mathrm{m}}, \tag{3.154}$$

and (V1b) as,

$$E_\pi \int_0^{\tau^{\mathrm{I}_*}} \hat{H}_s^2\, D_\Pi\, r_*(\Pi_s)\,ds = E_\pi \int_0^{\tau^{\mathrm{I}_*}} \hat{H}_s^2(-c)\,ds. \tag{3.155}$$

Thus, to arrange for (V1) it is sufficient that r_* satisfy,

$$D_\Pi\, r_*(\pi) = -c \qquad \forall\, \pi \in (0,1), \tag{3.156}$$

since this will give both (V1b) and (V1a) (with equality). As for conditions (V2)–(V4) we need to compute $\partial_\Pi\,\mathrm{I}_*$. We know that since I_* satisfies the escape condition (E) with respect to Π then $\partial_\Pi\,\mathrm{I}_* \neq \emptyset$. In this case then, because the sample paths of the *martingale* Π are P_π–a.s. continuous we see that $\partial_\Pi\,\mathrm{I}_* = \partial\mathrm{I}_* = \{a_*, b_*\}$, and moreover, $[\,\mathrm{I}_*\,]_\Pi = [a_*, b_*]$. Let's restate the verification conditions with these changes. Thus, we seek $r_* \in \mathcal{C}^2(0,1)$ and $\mathrm{I}_* \in \mathcal{I}^+$ such that,

$$
\begin{array}{llcll}
(\mathrm{S1}): & D_\Pi\, r_*(\pi) & = & -c & \forall\, \pi \in (0,1); \\
(\mathrm{S2}): & r_*(\pi) & = & e(\pi) & \forall\, \pi \in \{a_*, b_*\}; \\
(\mathrm{S3}): & r_*(\pi) & < & e(\pi) & \forall\, \pi \notin \{a_*, b_*\};
\end{array}
$$

(S4): r_* is bounded and continuous on $[a_*, b_*]$.

Conditions (S1) and (S2) give us a free-boundary-value problem, a so-called *Stefan* problem. The functional constraint (S3) serves to uniquely determine a solution amongst all the solutions to (S1)–(S2). In the next subsection we will show that (S1)–(S3) has a unique solution pair (r_*, I_*) which satisfies (S4) and therefore an appeal to the Verification Theorem will imply that $\tau^{\mathrm{I}_*} \in \mathcal{T}$ solves problem $(\mathcal{P}_{\mathrm{SD}})$.

3.6.5 Convexity Analysis of the Stefan Problem

We now proceed to solve the problem posed by (S1)-(S3). From 3.152 we know that $D_\Pi\, r$ for $r \in \mathcal{C}^2(0,1)$ is given by,

$$D_\Pi\, r(\pi) = \frac{1}{2}\,\pi^2\,(1-\pi)^2\, r''(\pi) \quad \forall\, \pi \in (0,1). \tag{3.157}$$

Fix any $a,\ b \in (0,1)$ satisfying $0 < a \leq \pi_e \leq b < 1$ and $a < b$, define $\mathrm{I} = (a,b)$, and consider the following ODE for $r \in \mathcal{C}^2(0,1)$,

$$
\begin{aligned}
\frac{1}{2}\,\pi^2\,(1-\pi)^2\, r''(\pi) &= -c \quad \forall\, \pi \in (0,1); \\
r(a) &= e(a); \\
r(b) &= e(b).
\end{aligned} \tag{3.158}
$$

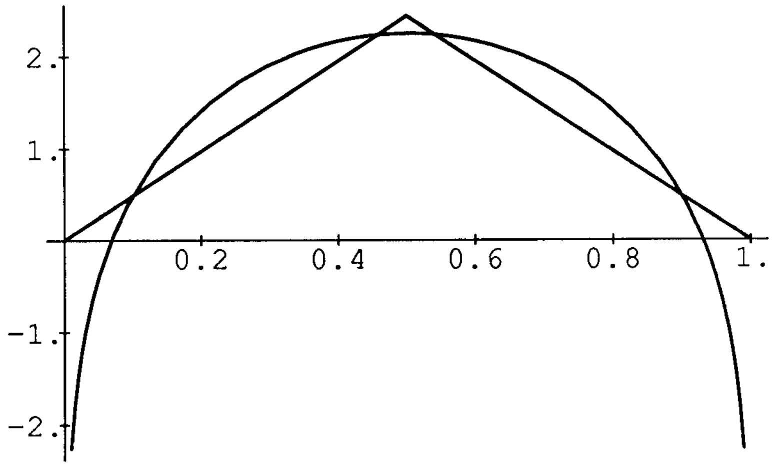

Figure 3.2: Graph of e and r_{I} with $\text{I} = (\frac{1}{10}, \frac{9}{10})$, $c = 1$.

We emphasize that the choice of a and b satisfying $0 < a \leq \pi_e \leq b < 1$ is motivated by Lemma 2.4. The connection between the above ODE and (S1)–(S2) should be clear. Under the assumption that c, c^0, and c^1 are positive and finite, by elementary ODE theory we know that a unique, nontrivial solution to (S1)–(S2) exists; call it r_{I}, with $r_{\text{I}} \in \mathcal{C}^2(0,1)$ (see Figure 3.2). From 3.158 it is clear that $r_{\text{I}}'' < 0$ and so r_{I} is (strictly) concave. Define the auxiliary function, $s_{\text{I}} := r_{\text{I}} - e$ (see Figure 3.3). Let S denote the hypograph of s_{I} and let $\overline{S} = Co(S)$, its convex hull. Since we have assumed that c^0 and c^1 are positive and finite we see that there is always a 'kink' in s_{I} at $\pi_e \in (0,1)$. Therefore $\overline{S} \backslash S \neq \emptyset$; in particular we can always choose that point $z_e \in \partial(\overline{S} \backslash S)$ with π–coordinate π_e. Construct the hyperplane at z_e supported by $\overline{S}$ (see Figure 3.4). Let $\ell_* : [0,1] \to \Re$ denote the line whose graph is this hyperplane. Since $z_e \notin \partial S$, there exist $\lambda \in (0,1)$ and x_*, $y_* \in \partial S$, with $x_* < \pi_e$ and $y_* > \pi_e$ such that $z_e = \lambda x_* + (1-\lambda) y_*$. We have

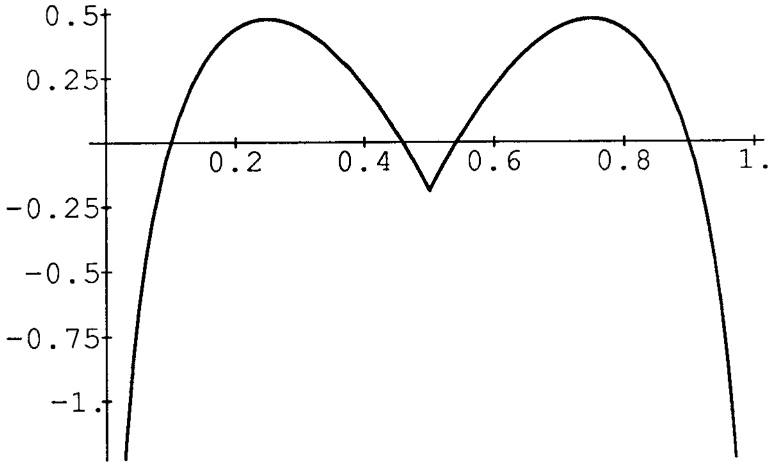

Figure 3.3: Graph of $s_{\mathrm{I}} = r_{\mathrm{I}} - e$.

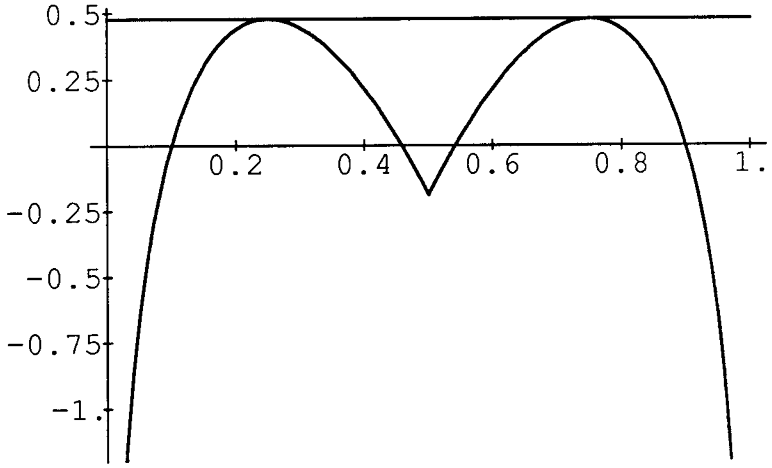

Figure 3.4: Graph of ℓ_* "supported" by s_{I}.

$x_* = (a_*, \ell_*(a_*)), y_* = (b_*, \ell_*(b_*))$ for some a_*, b_* with $0 < a_* < \pi_e < b_* < 1$ for which $s_{\rm I}(a_*) = \ell_*(a_*)$ and $s_{\rm I}(b_*) = \ell_*(b_*)$. We note that,

$$\ell_*(\pi) \geq s_{\rm I}(\pi) = r_{\rm I}(\pi) - e(\pi) \qquad \forall \pi \in [0, 1], \tag{3.159}$$

and if we define,

$$r_*(\pi) := r_{\rm I}(\pi) - \ell_*(\pi), \tag{3.160}$$

then obviously,

$$r_*(\pi) \leq e(\pi) \qquad \forall \pi \in [0, 1]. \tag{3.161}$$

Moreover,

$$\begin{aligned} r_*(a_*) &= r_{\rm I}(a_*) - \ell_*(a_*) \\ &= e(a_*) + (r_{\rm I}(a_*) - e(a_*)) - \ell_*(a_*) \\ &= e(a_*) + s_{\rm I}(a_*) - \ell_*(a_*) = e(a_*), \end{aligned} \tag{3.162}$$

and similarly,

$$r_*(b_*) = e(b_*). \tag{3.163}$$

Hence, $(r_*, {\rm I}_*)$ satisfies (S2). In addition note that,

$$D_\Pi r_*(\pi) = D_\Pi r_{\rm I}(\pi) - D_\Pi \ell_*(\pi) = -c - 0 = -c, \tag{3.164}$$

and therefore r_* satisfies (S1) also. Moreover, in view of the strict concavity of $r_{\rm I}$, it follows from the definition of r_* that these are the only two points for which the inequality in 3.161 is not strict, i.e.,

$$r_*(\pi) < e(\pi) \qquad \forall \pi \notin \{a_*, b_*\}, \tag{3.165}$$

which gives (S3). Thus, r_* as defined in 3.160 and ${\rm I}_* := (a_*, b_*)$ solves the Stefan problem posed by (S1)–(S2) with the functional constraint (S3). The convexity analysis also shows that $\mathrm{I}_* \in \mathcal{I}^+$, i.e., I_* is a proper subinterval of $\mathrm{I}_\infty = (0,1)$. As pointed out earlier, given (E) the fact that I_* is a proper subinterval of I_∞ implies that $[\,\mathrm{I}_*\,]_\Pi$ is also a proper subinterval of I_∞. Thus (S4) follows automatically because $r_* \in \mathcal{C}^2(0,1)$ implies *a fortiori* that $r_* \in \mathcal{BC}([\,\mathrm{I}_*\,]_\Pi)$. In addition, the convexity analysis shows that $a_* \leq \pi_e \leq b_*$. It only remains to show that the running cost conditions are satisfied. We consider this point and present the theorem for sequential detection in the next subsection.

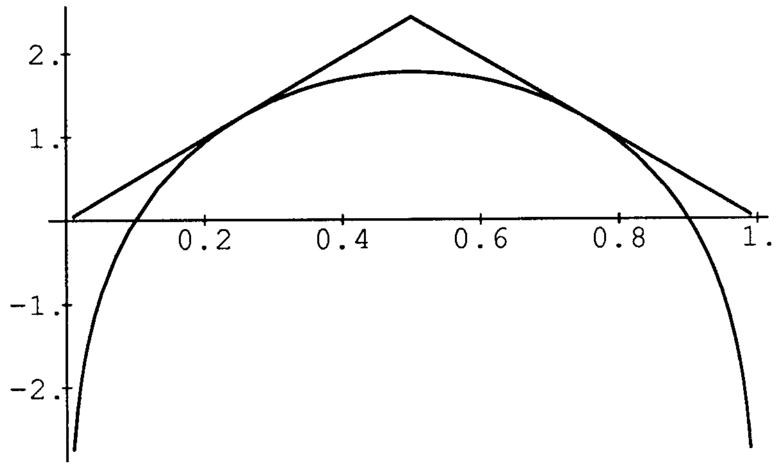

Figure 3.5: Graph of e and r_* with $a_* = \frac{1}{4}$, $b_* = \frac{3}{4}$.

3.6.6 Main Result

For ease of reference, we redisplay the technical conditions involving the drift process and its P_1–projection onto the observations:

$$\begin{aligned}
&\text{(H0):}\quad E_i \int_0^t H_s^2\, ds < \infty && \forall\, t \geq 0, && i = 0, 1; \\
&\text{(H1):}\quad E_i \int_0^\tau \hat{H}_s^2\, ds < \infty && \forall\, \tau \in \mathcal{T}, && i = 0, 1; \\
&\text{(H2):}\quad P_i \{ \int_0^\infty \hat{H}_s^2\, ds = \infty \} = 1 && && i = 0, 1.
\end{aligned}$$

We have come to the main result of this section.

Theorem 3.1 *Assume that the conditions* (H0), (H1),*and* (H2) *hold. In the problem of sequential detection based on observations of the process,*

$$Y_t = \begin{cases} W_t & t \geq 0 \quad \text{if } v = \infty; \\ \int_0^t H_s\, ds + W_t & t \geq 0 \quad \text{if } v = 0, \end{cases}$$

with average running cost,

$$E_\pi \int_0^\tau \mathcal{C}_s\, ds = E_\pi \int_0^\tau c\, \hat{H}_s^2\, ds \quad c > 0,$$

and average decision cost,

$$E_\pi[\, \mathcal{E}(\Upsilon_\tau, \delta)] = E_\pi[c^0\, (1-\delta)\, \Upsilon_\tau + c^1\, \delta\, (1 - \Upsilon_\tau)]$$

with $0 < c^0,\ c^1 < \infty$*, there exist* a_*, b_* *unique with* $0 < a_* < \pi_e < b_* < 1$*, such that the first exit policy* (τ^{I_*}, δ_*) *based on the continuation interval* $I_* = (a_*, b_*)$ *achieves Bayes' optimal cost, i.e.,*

$$\rho_\pi(\tau^{I_*}) = \inf_{\tau \in \mathcal{T}_{ad}} \rho_\pi(\tau) \qquad \forall\, \pi \in [0, 1],$$

where,

$$\rho_\pi(\tau) = E_\pi[\int_0^\tau \mathcal{C}_s\, ds + e(\Pi_\tau)] \quad \forall\, \pi \in [0,1] \text{ and } \tau \in \mathcal{T}_{ad}.$$

In addition, there exists $r_* \in \mathcal{C}^2(0,1)$*, the solution to* (S1)–(S4) *above, such that*

$$\rho_\pi(\tau^{\mathrm{I}_*}) = \begin{cases} r_*(\pi) & \text{if } \pi \in \mathrm{I}_*; \\ e(\pi) & \text{if } \pi \notin \mathrm{I}_*, \end{cases}$$

where,

$$e(\pi) = \min\{c^0\,\pi, c^1\,(1-\pi)\}.$$

Proof:

In the previous subsection we solved the associated Stefan problem and showed that there exists a pair (r_*, I_*) satisfying (S1)–(S3) for which the conditions on the cost coefficients guarantee that $0 < a_* < \pi_e < b_* < 1$. Therefore we have in fact exhibited a pair (r_*, I_*), $\mathrm{I}_* \in \mathcal{I}^+$ and $r_* \in \mathcal{BC}(\mathcal{I}^+)$, which satisfy (V1)–(V3). Moreover, since $\mathrm{I}_* \in \mathcal{I}^+$ we know that $[\mathrm{I}_*]_\Pi = [a_*, b_*] \subset (0,1)$ and therefore (r_*, I_*) also satisfies (V4) *a fortiori* in view of the fact that $r_* \in \mathcal{C}^2(0,1)$. Thus, we have found a pair satisfying (V1)–(V4).

To employ the Verification Theorem and therefore prove the theorem at hand it remains only to show that (C1), (C2), and (C3) hold since (E) follows from Proposition 3.8. With the above choice of running cost we see that (C1) follows *a fortiori* from (H1) since,

$$E_\pi \int_0^\tau \mathcal{C}_s\, ds = \pi\, E_1 \int_0^\tau c\, \hat{H}_s^2\, ds + (1-\pi)\, E_0 \int_0^\tau c\, \hat{H}_s^2\, ds < \infty. \qquad (3.166)$$

Next, it is obvious that (C2) follows from (H2) since $c > 0$ and thus,

$$\begin{aligned} P_\pi\{\int_0^\infty \mathcal{C}_s\, ds = \infty\} &= \pi\, P_1\{\int_0^\infty c\, \hat{H}_s^2\, ds\} + (1-\pi)\, P_0\{\int_0^\infty c\, \hat{H}_s^2\, ds\} \\ &= \pi \cdot 1 + (1-\pi)\cdot 1 = 1. \end{aligned} \quad (3.167)$$

Finally, condition (C3) follows since the running cost $\mathcal{C}$ is trivially concave in Π. Thus τ^{I_*} is the optimal stopping time for this problem and r_* characterizes Bayes' cost. □

3.6.7 Example

We end this section of the chapter with a concrete example of a sequential detection problem involving a diffusion with constant drift. We observe a stochastic process $Y = \{Y_t\}_{t\geq 0}$ for which one of the following hypotheses is true:

$$\begin{aligned} \text{(Noise Only)}: &\quad Y_t = W_t \qquad t \geq 0; \\ \text{(Signal Plus Noise)}: &\quad Y_t = t + W_t \qquad t \geq 0, \end{aligned}$$

where W is a $(\mathcal{G}_t, P_i)$–standard Wiener martingale for $i = 0,\ 1$. It is given that the (Signal Plus Noise) hypothesis occurs with prior probability $\pi \in [0, 1]$. Define the Bayes' cost,

$$\bar{\rho}_\pi(\tau, \delta) = \pi\, E_1[c\,\tau + c^0\, 1\{\delta = 0\}\,] + (1-\pi)\, E_0[c\,\tau + c^1\, 1\{\delta = 1\}\,], \quad (3.168)$$

where c, c^0, and c^1 are strictly positive and finite. We are asked to minimize $\bar{\rho}_\pi(\tau, \delta)$ over all decision pairs for which $P_i\{\tau < \infty\} = 1$ for $i = 0, 1$.

We see that this is precisely the form we are equipped to handle as long as we make the identification $H \equiv 1$ for then our choice of running cost collapses down to

$$\int_0^{\tau} \mathcal{C}_s\, ds = \int_0^{\tau} c\,\hat{H}_s\, ds = \int_0^{\tau} c\, ds = c\tau. \qquad (3.169)$$

We point out that each of (H0), (H1) and (H2) are trivially satisfied. Thus we can apply Theorem 3.1 to deduce that an optimal first exit policy exists. A formula for r_* can be found in [S, C4] and [MS, C4]. One can solve for the optimal interval I_* using the convexity approach given here or see [S, C4] and [MS, C4] for a system of equations to solve. In the symmetric case when $c^0 = c^1$ one can show that $a_* = 1 - b_*$; in the special symmetric case when $c^0 = c^1 = (3 - \frac{1}{3} + 2\log 3) \cdot c$ one can show that $a_* = \frac{1}{4}$ and $b_* = \frac{3}{4}$ (see [MS, C4]). The graphs in this case for $c = 1$ are depicted in Figures 3.1 through 3.5. The graph of ρ, the Bayes' optimal risk is given in Figure 3.6 along with the worst case risk. Figure 3.6 neatly depicts the so-called *smooth pasting property* of ρ and e as discussed in [S, MS]. The convexity analysis given here provides an elegant alternative to this approach to finding the optimal risk.

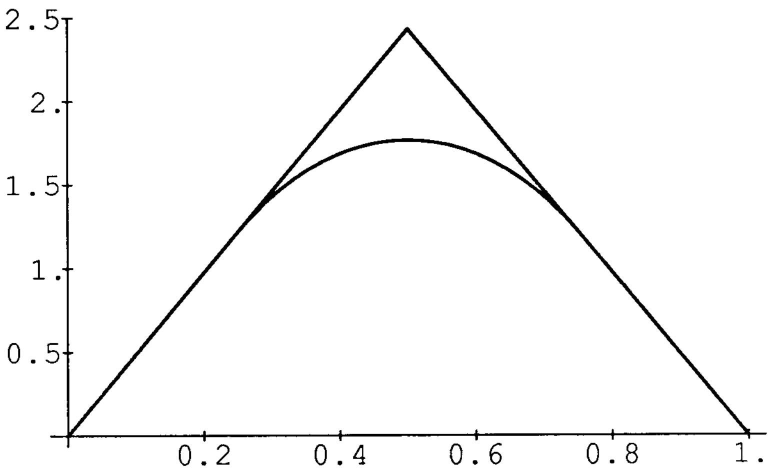

Figure 3.6: Graph of terminal cost and Bayes' optimal risk.

3.7 Disruption

In this section we will formulate the Bayesian *disruption* problem on the drift of an observed diffusion-type stochastic process. We will reformulate it as a *change detection* problem and solve it by recourse to the Verification Theorem. As in the last section this will lead us to consider a *Stefan* problem with its solution addressed using techniques from ODE theory and convex analysis. We will conclude the section with an example involving a system which jumps from a zero-drift diffusion to a positive-drift diffusion with a random jump time which is exponentially distributed.

3.7.1 Problem Statement

On a probability space $(\Omega, \mathcal{A}, P_\pi)$ we observe a stochastic process $Y = \{Y_t\}_{t\geq 0}$ for which,

$$Y_t = \begin{cases} W_t & 0 \leq t < v; \qquad \text{(Noise Only)} \\ \\ \int_v^t H_s\, ds + W_t & 0 \leq v \leq t, \quad \text{(Signal Plus Noise)} \end{cases} \tag{3.170}$$

where W is an $(\mathcal{A}_t, P_\pi)$–standard Wiener martingale, $H = \{H_t\}_{t\geq 0}$ is an $\mathcal{A}_t$–progressive process satisfying,

$$E_\pi \int_0^t H_s^2\, ds < \infty \quad \forall\, t \geq 0, \tag{3.171}$$

and v is an $\mathcal{A}_t$–stopping time for which $P_\pi\{v = 0\} = \pi$ for some number π in $[0, 1]$. In the cases when $\pi \neq 1$, $P_\pi\{v > 0\} > 0$, and we are given that

$$P\{\, v \leq t \,|\, \mathcal{F}_t \,\} = F_t \quad \forall\, t \geq 0, \tag{3.172}$$

where $P\{A\} := P_\pi\{\, A \,|\, v > 0 \,\}$ and with $F = \{F_t\}_{t \geq 0}$ modeled according to,

$$\text{(D)} \qquad F_t := 1 - \exp\{-\int_0^t \alpha\, \hat{H}_s^2\, ds\} \quad t \geq 0,\ \alpha > 0. \tag{3.173}$$

This defines an $\mathcal{O}_t$–measurable conditional cumulative distribution function for v under the P–measure; of course, $\mathcal{O}_t$ is the P_π–completed observation filtration $\mathcal{O}_t = \bigvee_{s \leq t} \sigma(Y_s)$. We are tasked with guessing as time progresses whether disruption has occured. A decision policy is defined as any P_π–a.s. finite $\mathcal{O}_t$–stopping time τ. The goodness of any such stopping time is to be judged according to the 'elapsed energy excess + panic cost' penalty criterion,

$$\bar{\rho}_\pi(\tau) = E_\pi[\max\{0, \int_v^\tau c\, \hat{H}_s^2\, ds\}\,] + P_\pi\{\tau < v\}, \tag{3.174}$$

where c is strictly positive. We are asked to minimize $\bar{\rho}_\pi(\tau)$ over all P_π–a.s. finite stopping times.

3.7.2 Problem Reformulation

Quite clearly, the disruption problem is *the* generic change detection problem with observations,

$$Y_t = \int_0^t \Upsilon_s\, H_s\, ds + W_t^v \quad \forall\, t \geq 0, \tag{3.175}$$

and disruption time v having the conditional cumulative distribution function F as defined in the previous subsection. Note that $\hat{H}_t = E_1[H_t|\mathcal{O}_t]$ and $P_1\{A\} = P_\pi\{A|v = 0\}$ if $\pi < 1$ and $P_1\{A\} = P_\pi\{A\}$ if $\pi = 1$ for all A in $\mathcal{A}$; of course we assume that the technical conditions (H0)–(H2) are in force. To solve the disruption problem we need only account for the classic two-part cost given above within the cost structure for problems of change detection as defined in Chapter 2. Choosing $c^0 = \infty$ and $c^1 = 1$ gives us,

$$\mathcal{E}(\Upsilon_\tau, \delta) = \begin{cases} \infty & \text{if } \Upsilon_\tau = 1 \text{ and } \delta = 0; \\ 1 & \text{if } \Upsilon_\tau = 0 \text{ and } \delta = 1. \end{cases} \tag{3.176}$$

We see therefore that this choice of terminal decision cost forces us to choose $\delta \equiv 1$ or accept infinite cost, thus,

$$\begin{aligned} E_\pi \mathcal{E}(\Upsilon_\tau, \delta) &= E_\pi \mathcal{E}(\Upsilon_\tau, 1) \\ &= E_\pi[\,0 + c^1\, 1\{\Upsilon_\tau = 0\} \cdot 1\,] \\ &= E_\pi[\,1 - 1\{\Upsilon_\tau = 1\}\,] \\ &= 1 - E_\pi[\,P_\pi\{\Upsilon_\tau = 1 \,|\, \mathcal{O}_\tau\,\}\,] \\ &= E_\pi[1 - \Pi_\tau] \\ &= E_\pi[\,e(\Pi_\tau)\,]. \end{aligned} \tag{3.177}$$

On the other hand, the panic cost satisfies

$$\begin{aligned} P_\pi\{\tau < \upsilon\} &= 1 - P_\pi\{\upsilon \leq \tau\} \\ &= 1 - E_\pi[\,\Upsilon_\tau\,] \\ &= 1 - E_\pi[E_\pi[\Upsilon_\tau | \mathcal{O}_\tau]] \\ &= E_\pi[1 - \Pi_\tau] \\ &= E_\pi[\,e(\Pi_\tau)\,]. \end{aligned} \tag{3.178}$$

Hence, the choices $c^0 = \infty$ and $c^1 = 1$ capture the panic cost model within the Bayes' terminal cost set-up. Next we consider the running cost portion of the penalty criterion for P_π–a.s. finite $\mathcal{O}_t$–stopping times $\tau \in \mathcal{T}$. We compute,

$$E_\pi[\,\max\{0, \int_\upsilon^\tau c\,\hat{H}_s^2\, ds\}\,] = E_\pi[\,1\{\upsilon \leq \tau\} \int_\upsilon^\tau c\,\hat{H}_s^2\, ds\,]$$

$$
\begin{aligned}
&= E_\pi[\int_v^\tau c\,1\{v \le s\}\,\hat{H}_s^2\,ds\,] \\
&= E_\pi[\int_v^\tau c\,\Upsilon_s\,\hat{H}_s^2\,ds\,].
\end{aligned} \tag{3.179}
$$

With the intention of simplifying this last expression define $X = \{X_t\}_{t \ge 0}$ via,

$$
X_t := E_\pi[\int_0^t (\Upsilon_s - \Pi_s)\,\hat{H}_s^2\,ds \,|\, \mathcal{O}_t\,] \quad \forall t \ge 0. \tag{3.180}
$$

Compute with $t \ge 0$ and $r \le t$,

$$
\begin{aligned}
E_\pi[X_t \,|\, \mathcal{O}_r\,] &= E_\pi[\int_0^t (\Upsilon_s - \Pi_s)\,\hat{H}_s^2\,ds \,|\, \mathcal{O}_r\,] \\
&= X_r + E_\pi[\int_r^t (\Upsilon_s - \Pi_s)\,\hat{H}_s^2\,ds \,|\, \mathcal{O}_r\,] \\
&= X_r + \int_r^t E_\pi[(\Upsilon_s - \Pi_s)\,\hat{H}_s^2 \,|\, \mathcal{O}_r\,]\,ds,
\end{aligned} \tag{3.181}
$$

where the last line follows from a routine Fubini argument. Next observe,

$$
E_\pi[(\Upsilon_s - \Pi_s)\,\hat{H}_s^2 \,|\, \mathcal{O}_r\,] = E_\pi[\,E_\pi[\,(\Upsilon_s - \Pi_s) \,|\, \mathcal{O}_s\,]\,\hat{H}_s^2 \,|\, \mathcal{O}_r\,] = 0, \tag{3.182}
$$

for all $s \ge r \ge 0$. Therefore,

$$
E_\pi[X_t \,|\, \mathcal{O}_r\,] = X_r \quad 0 \le r \le t, \tag{3.183}
$$

or in words, X is an $(\mathcal{O}_t, P_\pi)$–martingale. From this it follows using the strong form of Doob's Optional Sampling Theorem [E, T4.12] that,

$$
E_\pi X_\tau = E_\pi[\,X_\tau \,|\, \mathcal{O}_0\,] = X_0 = 0. \tag{3.184}
$$

We conclude,

$$
E_\pi \int_0^\tau \Upsilon_s\,\hat{H}_s^2\,ds = E_\pi \int_0^\tau \Pi_s\,\hat{H}_s^2\,ds \quad \forall \tau \in \mathcal{T}. \tag{3.185}
$$

With this computation in hand we see that if we choose the nonnegative running cost $\mathcal{C}$ according to,

$$\mathcal{C}_t := c\,\Pi_t\,\hat{H}_t^2 \quad \forall\, t \geq 0,\; c > 0. \tag{3.186}$$

then taking 3.185 and 3.178 into account Bayes' cost is given by,

$$\begin{aligned} \rho_\pi(\tau,\delta) &= E_\pi[\int_0^\tau \mathcal{C}_s\,ds + \mathcal{E}(\Upsilon_\tau,\delta)\,] \\ &= E_\pi[\int_0^\tau c\,\Pi_s\,\hat{H}_s^2\,ds + \mathcal{E}(\Upsilon_\tau,1)\,] \\ &= E_\pi[\int_0^\tau c\,\Upsilon_s\,\hat{H}_s^2\,ds + e(\Pi_\tau)\,] \\ &= E_\pi[\max\{0,\int_\upsilon^\tau c\,\hat{H}_s^2\,ds\}\,] + P_\pi\{\tau < \upsilon\} \\ &= \bar{\rho}_\pi(\tau). \end{aligned} \tag{3.187}$$

Hence, our choices of running cost and terminal cost together with assumption (D) yield a classical disruption problem involving the drift of a generalized diffusion. To solve this disruption problem it is therefore sufficient to solve the optimal stopping problem ($\mathcal{P}$) of Chapter 2 under the conditions prevailing in the current section. Again, the plan is the same as in the sequential detection case: employ the Verification Theorem to characterize the associated Stefan problem and then solve the Stefan problem via ODE theory and a similar convexity analysis. Bayes' cost for the disruption problem is,

$$\rho_\pi(\tau) = E_\pi[\int_0^\tau c\,\Pi_s\,\hat{H}_s^2\,ds + e(\Pi_\tau)], \tag{3.188}$$

with $c > 0$ and with the terminal cost given by,

$$e(\pi) = \min\{c^0\pi, c^1(1-\pi)\} = \min\{\infty, 1-\pi\} = 1-\pi, \tag{3.189}$$

for all π in $[0,1]$ as shown in Figure 3.7. For the choice of running cost $\mathcal{C}$

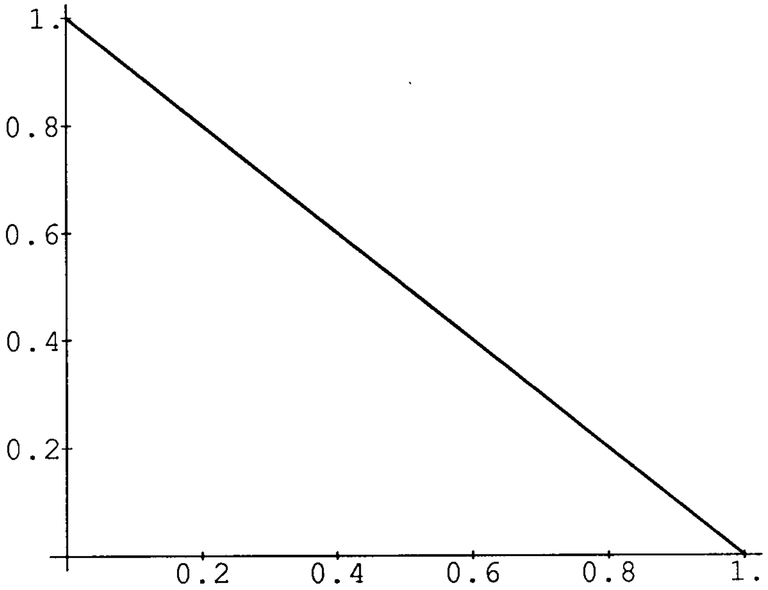

Figure 3.7: Terminal "panic" cost. $c^0 = \infty$, $c^1 = 1$, $\pi_e = 0$.

made in 3.186, condition (C1) translates to: τ admissible is any $\mathcal{O}_t$–stopping time satisfying,

$$E_\pi \int_0^\tau \mathcal{C}_s\, ds = E_\pi \int_0^\tau c\,\Pi_s\, \hat{H}_s^2\, ds < \infty. \tag{3.190}$$

To summarize, the change detection problem to solve equivalent to the disruption problem is:

$$\mathcal{P}_{\mathrm{D}}: \qquad \text{Find } \tau_* \in \mathcal{T}_{ad} \text{ such that } \rho_\pi(\tau_*) = \inf_{\tau \in \mathcal{T}_{ad}} \rho_\pi(\tau). \tag{3.191}$$

3.7.3 Escape Properties

The next question to answer is whether problem ($\mathcal{P}_{\mathrm{D}}$) is sufficiently well-posed that the escape condition (E) is satisfied and then which of (E^+) or (E^0) holds. We note that (F_{D}) implies that the stochastic differential dF is given by,

$$dF_t = \exp\{-\int_0^t \alpha\,\hat{H}_s^2\,ds\}\,\alpha\,\hat{H}_t^2\,dt = (1 - F_t)\,\alpha\,\hat{H}_t^2\,dt. \tag{3.192}$$

From this it follows that,

$$(1 - F_t)^{-1}\,dF_t = \alpha\,\hat{H}_t^2\,dt, \tag{3.193}$$

and this specializes the semimartingale representation for Υ to,

$$\Upsilon_t = \Upsilon_0 + \int_0^t \alpha\,(1 - \Upsilon_s)\,\hat{H}_s^2\,ds + M_t \quad \forall\, t \geq 0,\; P_\pi\text{–a.s.} \tag{3.194}$$

The filter for Π similarly reduces to,

$$\Pi_t = \Pi_0 + \int_0^t \alpha\,(1 - \Pi_s)\,\hat{H}_s^2\,ds + \int_0^t \Pi_s\,(1 - \Pi_s)\,\hat{H}_s\,d\overline{W}_s^{\upsilon}. \tag{3.195}$$

In the Proposition below we show that in this case condition (H2) implies that (E^0) holds so that $\mathcal{I}_\infty = \mathcal{I}^0$ and $\mathrm{I}_\infty = [0,1)$, and therefore the trajectories of Π are guaranteed to escape any subinterval of $[0,1)$ possessing a right-hand endpoint which is strictly less than one.

Proposition 3.10 *Under assumption* (H2) *the escape condition* (E^0) *holds, i.e., given*

$$P_i\{\int_0^\infty \hat{H}_s^2\,ds = \infty\} = 1 \qquad i = 0, 1,$$

then,

$$P_\pi\{\tau^{\mathrm{I}} < \infty\} = 1 \qquad \forall\, \pi \in \mathrm{I},\; \forall\, \mathrm{I} \in \mathcal{I}^0.$$

Proof: Choose any I in $\mathcal{I}^0$ so that $\mathrm{I} = [0, b)$ with $0 < b < 1$ and suppose $\pi \in \mathrm{I}$. From 3.195 we obtain,

$$\Pi_t - \Pi_0 = \alpha \int_0^t (1 - \Pi_s)\, \hat{H}_s^2\, ds + \overline{M}_t, \tag{3.196}$$

where $\overline{M}$ is an $(\mathcal{O}_t, P_\pi)$–local martingale with a localization sequence, say $\{\mu_n\}_{n \geq 1}$. Using by now familiar arguments we obtain,

$$E_\pi[\, \Pi_{\tau^{\mathrm{I}} \wedge \mu_n} - \Pi_0\,] = \alpha E_\pi \int_0^{\tau^{\mathrm{I}} \wedge \mu_n} (1 - \Pi_s)\, \hat{H}_s^2\, ds. \tag{3.197}$$

Obviously,

$$E_\pi[\, \Pi_{\tau^{\mathrm{I}} \wedge \mu_n} - \Pi_0\,] \leq 1. \tag{3.198}$$

Because $\pi \in (0, b)$, we know that $\Pi_{\tau^{\mathrm{I}} \wedge \mu_n} \leq b$, P_π–a.s., and thus,

$$E_\pi \int_0^{\tau^{\mathrm{I}} \wedge \mu_n} (1 - \Pi_s)\, \hat{H}_s^2\, ds \geq (1 - b) E_\pi \int_0^{\tau^{\mathrm{I}} \wedge \mu_n} \hat{H}_s^2\, ds. \tag{3.199}$$

Combining these results yields,

$$E_\pi \int_0^{\tau^{\mathrm{I}} \wedge \mu_n} \hat{H}_s^2\, ds \leq \frac{\alpha^{-1}}{(1 - b)} < \infty. \tag{3.200}$$

Using additional familiar arguments we pass to the limit and obtain,

$$E_\pi \int_0^{\tau^{\mathrm{I}}} \hat{H}_s^2\, ds \leq \frac{\alpha^{-1}}{(1 - b)} < \infty. \tag{3.201}$$

and so,

$$\infty > E_\pi \int_0^{\tau^{\mathrm{I}}} \hat{H}_s^2\, ds \geq E_\pi[\, 1\{\tau^{\mathrm{I}} = \infty\} \int_0^{\infty} \hat{H}_s^2\, ds\,]. \tag{3.202}$$

This last line yields a contradiction unless $P_\pi\{\tau^{\mathrm{I}} = \infty\} = 0$ since it is obvious from (H2) that $P_\pi\{\, \int_0^\infty \hat{H}_s^2\, ds = \infty\,\} = 1$ for all π in I. If we now define $\tau_b := \inf\{t \geq 0 : \Pi_t \geq b\}$ it is easy to see that the same argument works to show that $P_\pi\{\tau_b < \infty\} = 1$. Hence we conclude that Π exits $\mathrm{I} \in \mathcal{I}^0$ only to the right, P_π–a.s. for all π in I. $\square$

Motivated by this proposition we are justified in searching for $\mathrm{I}_* \in \mathcal{I}^0$ so that if we write $\mathrm{I}_* = [0, b_*)$ then $0 < b_* < 1$. Note how this is in harmony with Lemma 2.4 since $0 = a_* \leq \pi_e = 0 \leq b_*$. Before embarking on our search for I_* we consider one more result which is companion to the last. Indeed, while the last proposition shows (H2) implies that the Π process is guaranteed to escape any interval in $\mathcal{I}_\infty$, in the following proposition we will show (H1) guarantees that Π can only escape $\mathrm{I}_\infty = [0,1)$ in infinite time.

Proposition 3.11 *Under assumption* (H1),

$$P_\pi\{\tau^{\mathrm{I}_\infty} = \infty\} = 1 \qquad \forall \pi \in \mathrm{I}_\infty.$$

Proof: Fix $\pi \in \mathrm{I}_\infty = [0,1)$. With $t \geq 0$ and $O \in \mathcal{O}_t$ we have,

$$\begin{aligned} \int_O \Pi_t \, dP_\pi &= \int_O P_\pi\{v \leq t \,|\, \mathcal{O}_t\} \, dP_\pi = E_\pi[1\{v \leq t\} \, 1_O] \\ &= \pi \, P_1\{O\} + (1-\pi) \, E_0 \int_0^t Q^u\{O\} \, dF_u, \end{aligned} \tag{3.203}$$

where the last line follows from 2.106. Working on the integral in the second term of 3.203 gives,

$$\begin{aligned} E_0 \int_0^t Q^u\{O\} \, dF_u &= E_0 \int_0^t \int_O \frac{dQ^u}{dP_0} \, dP_0 \, dF_u \\ &= E_0 \int_0^t \int_O E_0\left[\frac{dQ^u}{dP_0} \,|\, \mathcal{O}_t \right] \, dP_0^{\mathcal{O}_t} \, dF_u \end{aligned}$$

$$= E_0 \int_O \int_0^t E_0\left[\frac{dQ^u}{dP_0} \,|\, \mathcal{O}_t\right] dF_u \, dP_0^{\mathcal{O}_t}$$
$$= \int_O \int_0^t E_0\left[\frac{dQ^u}{dP_0} \,|\, \mathcal{O}_t\right] dF_u \, dP_0, \qquad (3.204)$$

where the second to last line follows from Fubini's Theorem. Also we see that,

$$P_1\{\, O \,\} = \int_O E_0\left[\frac{dP_1}{dP_0} \,|\, \mathcal{O}_t\right] dP_0. \qquad (3.205)$$

By previous results we can write,

$$\check{L}_t = E_0\left[\frac{dP_1}{dP_0} \,|\, \mathcal{O}_t\right] \quad t \geq 0, \qquad (3.206)$$

and

$$\check{L}_u^{-1} \, \check{L}_t = E_0\left[\frac{dQ^u}{dP_0} \,|\, \mathcal{O}_t\right] \quad 0 \leq u \leq t. \qquad (3.207)$$

Hence combining expressions 3.203 through 3.207 yields,

$$\int_O \Pi_t \, dP_\pi = \int_O \left(\pi \check{L}_t + (1-\pi) \int_0^t \check{L}_u^{-1} \, \check{L}_t \, dF_u\right) \frac{dP_0}{dP_\pi} \, dP_\pi$$
$$= \int_O \left(\pi \check{L}_t + (1-\pi) \int_0^t \check{L}_u^{-1} \, \check{L}_t \, dF_u\right) E_\pi\left[\frac{dP_0}{dP_\pi} \,|\, \mathcal{O}_t\right] dP_\pi$$
$$= \int_O \frac{\left(\pi \check{L}_t + (1-\pi) \int_0^t \check{L}_u^{-1} \, \check{L}_t \, dF_u\right)}{E_0\left[\frac{dP_\pi}{dP_0} \,|\, \mathcal{O}_t\right]} \, dP_\pi, \qquad (3.208)$$

where the simplification in the last line results from using Bayes' Formula [B, VI.3.L5] which lets us write,

$$E_\pi\left[\frac{dP_0}{dP_\pi} \,|\, \mathcal{O}_t\right] = \left(E_0\left[\frac{dP_\pi}{dP_0} \,|\, \mathcal{O}_t\right]\right)^{-1} \quad P_\pi\text{–a.s.} \qquad (3.209)$$

It is straightforward to show,

$$E_0\left[\frac{dP_\pi}{dP_0} \,|\, \mathcal{O}_t\right] = \pi \, \check{L}_t + (1-\pi) \int_0^t \check{L}_u^{-1} \, \check{L}_t \, dF_u + (1-\pi) \, P_\pi\{t < v\}, \quad (3.210)$$

and we conclude for all $O \in \mathcal{O}_t$ that,

$$\int_O \Pi_t \, dP_\pi = \int_O \frac{\pi \check{L}_t + (1-\pi) \int_0^t \check{L}_u^{-1} \check{L}_t \, dF_u}{\pi \check{L}_t + (1-\pi) \int_0^t \check{L}_u^{-1} \check{L}_t \, dF_u + (1-\pi) P_\pi\{t < \upsilon\}} \, dP_\pi, \tag{3.211}$$

and therefore,

$$\Pi_t = \frac{\pi \check{L}_t + (1-\pi) \int_0^t \check{L}_u^{-1} \check{L}_t \, dF_u}{\pi \check{L}_t + (1-\pi) \int_0^t \check{L}_u^{-1} \check{L}_t \, dF_u + (1-\pi) P_\pi\{t < \upsilon\}} \quad P_\pi\text{-a.s.} \tag{3.212}$$

From this it is clear that,

$$\{|\Pi_{\tau^{\mathrm{I}_\infty}} - \frac{1}{2}| = \frac{1}{2}\} = \{|\log \check{L}_{\tau^{\mathrm{I}_\infty}}| = \infty\}, \tag{3.213}$$

because $P_\pi\{t < \upsilon\} > 0$ for all $t \geq 0$, and therefore,

$$\{\tau^{\mathrm{I}_\infty} < \infty\} \subset \{|\log \check{L}_{\tau^{\mathrm{I}_\infty}}| = \infty\}. \tag{3.214}$$

Hence,

$$\begin{aligned}
P_\pi\{\tau^{\mathrm{I}_\infty} < \infty\} &= P_\pi\{|\log \check{L}_{\tau^{\mathrm{I}_\infty}}| = \infty, \tau^{\mathrm{I}_\infty} < \infty\} \\
&= P_\pi\{|\log \check{L}_{\tau^{\mathrm{I}_\infty}}| = \infty, \cup_{n\geq 1} \tau^{\mathrm{I}_\infty} \leq n\} \\
&\leq \sum_{n=1}^{\infty} P_\pi\{|\log \check{L}_{\tau^{\mathrm{I}_\infty}}| = \infty, \tau^{\mathrm{I}_\infty} \leq n\} \\
&\leq \sum_{n=1}^{\infty} P_\pi\{\sup_{0\leq t\leq n} |\log \check{L}_t| = \infty, \tau^{\mathrm{I}_\infty} \leq n\} \\
&\leq \sum_{n=1}^{\infty} P_\pi\{\sup_{0\leq t\leq n} |\log \check{L}_t| = \infty\} \\
&= \lim_{m\to\infty} \sum_{n=1}^{m} 0 = 0,
\end{aligned} \tag{3.215}$$

where the last equality follows from Proposition 3.5. Thus, Π has no chance of escaping from $\mathrm{I}_\infty = [0,1)$ in finite time. □

3.7.4 Verification: A Stefan Problem

For ease of reference, we state the verification conditions (V1)–(V4) for problem $(\mathcal{P}_{\mathrm{D}})$ given in the previous subsection with $\mathrm{I}_\infty = [0,1)$, $\mathcal{I}_\infty = \mathcal{I}^0$, and $\mathrm{I}_* \in \mathcal{I}^0$. We also incorporate the proposed running cost given in 3.186:

$$\begin{array}{llcll}
(\mathrm{V1a}): & \textit{For all } \tau \in \mathcal{T}_{\mathrm{m}}, & & & \\
& E_\pi[r_*(\Pi_\tau) - r_*(\Pi_0)] & \geq & -E_\pi \int_0^\tau c\,\Pi_s\,\hat{H}_s^2\,ds & \forall \pi \in [\,\mathrm{I}_*]_\Pi; \\
(\mathrm{V1b}): & E_\pi[r_*(\Pi_{\tau^{\mathrm{I}_*}}) - r_*(\Pi_0)] & = & -E_\pi \int_0^{\tau^{\mathrm{I}_*}} c\,\Pi_s\,\hat{H}_s^2\,ds & \forall \pi \in [\,\mathrm{I}_*]_\Pi; \\
(\mathrm{V2}): & r_*(\pi) & = & e(\pi) & \forall \pi \in \partial_\Pi \mathrm{I}_*; \\
(\mathrm{V3}): & r_*(\pi) & < & e(\pi) & \forall \pi \notin \partial_\Pi \mathrm{I}_*;
\end{array}$$

and,

$$(\mathrm{V4}): \;\; r_* \text{ is bounded and continuous on } [\,\mathrm{I}_*]_\Pi.$$

As in the sequential detection case we confine our search for r_* to $\mathcal{BC}^2(\mathcal{I}^0)$ and it is not too difficult to see that $\mathcal{BC}^2(\mathcal{I}^0) = \mathcal{C}^2[0,1)$, a subset of $\mathcal{C}^2(0,1)$. Taking the disruption assumption $(\mathrm{F_D})$ into account, the $\mathcal{O}_t$–measurable linear stochastic differential operator $D_{\mathcal{O}_t}$ defined in 3.89 for all mappings $r \in \mathcal{C}^2(0,1)$ simplifies to,

$$\begin{aligned}
D_{\mathcal{O}_t}\, r(\pi) &= \alpha\,(1-\pi)\,r'(\pi)\,(1-F_t)^{-1}\,dF_t + \frac{1}{2}\,(1-\pi)^2\,\pi^2\,\hat{H}_t^2\,r''(\pi)\,dt \\
&= \hat{H}_t^2\,[\,\alpha\,(1-\pi)\,r'(\pi) + \frac{1}{2}\,(1-\pi)^2\,\pi^2\,r''(\pi)\,]\,dt. \qquad (3.216)
\end{aligned}$$

If we conveniently define,

$$D_\Pi\, r(\pi) := \alpha\,(1-\pi)\,r'(\pi) + \frac{1}{2}\,(1-\pi)^2\,\pi^2\,r''(\pi), \qquad (3.217)$$

then Proposition 3.7 can be rewritten as,

$$E_\pi[\,r(\Pi_\tau) - r(\Pi_0)\,] = E_\pi \int_0^\tau \hat{H}_s^2\, D_\Pi\, r(\Pi_s)\,ds \quad \forall\, \tau \in \mathcal{T}_{\mathrm{m}}. \qquad (3.218)$$

Using this we can rewrite (V1a) as,

$$E_\pi \int_0^\tau \hat{H}_s^2 \, D_\Pi \, r_*(\Pi_s) \, ds \geq E_\pi \int_0^\tau \hat{H}_s^2(-c\,\Pi_s) \, ds \quad \forall \, \tau \in \mathcal{T}_{\mathrm{m}}, \tag{3.219}$$

and (V1b) as,

$$E_\pi \int_0^{\tau^{\mathrm{I}_*}} \hat{H}_s^2 \, D_\Pi \, r_*(\Pi_s) \, ds = E_\pi \int_0^{\tau^{\mathrm{I}_*}} \hat{H}_s^2(-c\,\Pi_s) \, ds. \tag{3.220}$$

Thus, to arrange for both (V1a) (with equality) and (V1b) it is sufficient that r_* satisfy,

$$D_\Pi \, r_*(\pi) = -c\,\pi \qquad \forall \, \pi \in (0,1). \tag{3.221}$$

To restate conditions (V2), (V3), and (V4) for the specific problem at hand, we need to compute the Π–boundary $\partial_\Pi \, \mathrm{I}_*$ and the Π–closure $[\,\mathrm{I}_*\,]_\Pi$. Since the sample paths of the submartingale Π are P_π–a.s. continuous we know that,

$$\partial_\Pi \, \mathrm{I}_* \subseteq \partial \mathrm{I}_* = \{a_*, b_*\} = \{0, b_*\}, \tag{3.222}$$

and this gives,

$$[\,\mathrm{I}_*\,]_\Pi = [a_*, b_*] = [0, b_*]. \tag{3.223}$$

Since $\mathrm{I}_* \in \mathcal{I}^0$, Proposition 3.10 implies that $P_\pi\{\Pi_{\tau^{\mathrm{I}_*}} = 0\} = 0$ and therefore $P_\pi\{\Pi_{\tau^{\mathrm{I}_*}} = b\} = 1$, i.e., Π always exits I_* to the right. From this it follows that,

$$\partial_\Pi \, \mathrm{I}_* = \{b_*\}. \tag{3.224}$$

We are now in a position to restate the verification conditions which take into account the specifics of the disruption problem. Thus, we seek $r_* \in \mathcal{C}^2[0,1)$

and $b_* \in (0,1)$ such that,

$$\begin{array}{lcccl} \text{(S1)}: & D_\Pi\, r_*(\pi) & = & -c\,\pi & \forall\, \pi \in [0,1); \\ \text{(S2)}: & r_*(\pi) & = & e(\pi) & \text{if } \pi = b_*; \\ \text{(S3)}: & r_*(\pi) & < & e(\pi) & \text{if } \pi \neq b_*; \end{array}$$

(S4): r_* is bounded and continuous on $[0, b_*]$.

Conditions (S1) and (S2) give us another *Stefan* problem. The functional constraints (S3)–(S4) serve to uniquely determine a solution amongst all the solutions to (S1)–(S2). In the next subsection we will show that (S1)–(S4) has a unique solution pair (r_*, I_*) and therefore by the Verification Theorem $\tau^{\mathrm{I}_*} \in \mathcal{T}_{ad}$ solves problem $(\mathcal{P}_{\mathrm{D}})$.

3.7.5 Convexity Analysis of the Stefan Problem

We now proceed to solve the problem defined by (S1)–(S4). From 3.217 we know that $D_\Pi\, r$ for $r \in \mathcal{C}^2(0,1)$ is given by,

$$D_\Pi\, r(\pi) = \alpha\,(1-\pi)\,r'(\pi) + \frac{1}{2}\,(1-\pi)^2\,\pi^2\,r''(\pi) \quad \forall\, \pi \in (0,1). \tag{3.225}$$

Fix any $b \in (0,1)$, define $\mathrm{I} = (0,b)$, and consider the following ODE for $r \in \mathcal{C}^2[0,1)$,

$$\begin{aligned} \alpha\,(1-\pi)\,r'(\pi) + \frac{1}{2}\,(1-\pi)^2\,\pi^2\,r''(\pi) &= -c\pi \quad \forall\, \pi \in [0,1); \\ r(b) &= e(b) \quad 0 < b < 1. \end{aligned} \tag{3.226}$$

Thus (S1)–(S2) yield a linear, second-order ODE which is reducible: let $q(\pi) := r'(\pi)$. This yields the family of integral curves satisfying,

$$\alpha\,(1-\pi)\,q(\pi) + \frac{1}{2}\,\pi^2(1-\pi)^2\,q'(\pi) = -c\,\pi \quad \forall\, \pi \in (0,1). \tag{3.227}$$

Conditions (S1) and (S2) offer no help in uniquely specifying a member of this family; the ODE in 3.226 does not posses a unique solution. However from (S4) we know that the solution which we seek should also be bounded and continuous on $[0, b)$. Since,

$$q'(\pi) = \frac{-c\,\pi - \alpha\,(1-\pi)\,q(\pi)}{\frac{1}{2}\,\pi^2\,(1-\pi)^2} \qquad \forall\,\pi \in (0,1), \tag{3.228}$$

we see that there is a vector field separatrix at $\pi = 0$. Hence (S4) requires that we choose this separatrix as the desired solution of the reduced equation since it is the only member of the family of integral curves which is bounded in a neighborhood of the origin. Thus, we seek a solution to,

$$\begin{aligned} \alpha\,(1-\pi)\,q(\pi) + \frac{1}{2}\,\pi^2(1-\pi)^2\,q'(\pi) &= -c\,\pi \qquad \forall\,\pi \in [0,1); \\ q(0) &= 0. \end{aligned} \tag{3.229}$$

By elementary ODE theory we know that a unique nontrivial solution exists, call it q_0. This gives us a solution to the original ODE, call it r_{I}, with $r_{\mathrm{I}} \in C^2\,[0,1)$, $r_{\mathrm{I}}'(0) = 0$ and $\mathrm{I} = [0, b)$; indeed (see Figure 3.8),

$$r_{\mathrm{I}}(\pi) = 1 - b + \int_b^{\pi} q_0(p)\,dp \qquad \forall\,\pi \in [0,1). \tag{3.230}$$

In anticipation of a convexity analysis similar to that in the last chapter, consider the following proposition.

Proposition 3.12 *For* $\mathrm{I} = [0, b)$ *with* $0 < b < 1$, *the solution* r_{I} *defined above is strictly concave.*

Proof: From 3.228 we see that,

$$r_{\mathrm{I}}''(\pi) = \frac{-2c\,\pi - 2\alpha\,(1-\pi)\,q_0(\pi)}{\pi^2\,(1-\pi)^2}$$

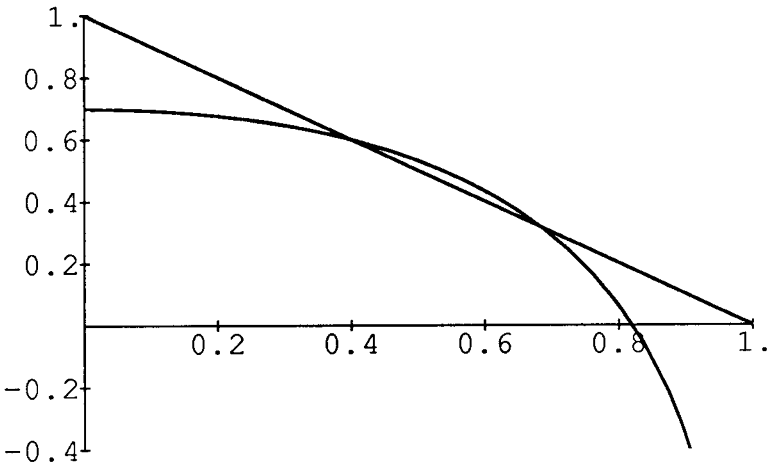

Figure 3.8: Graph of e and r_{I} with I = [0, .4) and $\alpha = 1$, $c = 1$.

$$
\begin{aligned}
&= \frac{2c}{\pi^2(1-\pi)}\left[-\frac{\alpha}{c}\,q_0(\pi) - \frac{\pi}{(1-\pi)}\right] \\
&= \frac{2c}{\pi^2(1-\pi)}\,\exp\{-2\alpha\, s(\pi)\}\,[\,J(\pi) - K(\pi)\,], \qquad (3.231)
\end{aligned}
$$

where $s(\pi) := \log(\frac{p}{1-p}) - 1/p$,

$$J(\pi) := -\frac{\alpha}{c}\,\exp\{2\alpha\, s(\pi)\}\,q_0(\pi), \qquad (3.232)$$

and

$$K(\pi) := \frac{\pi}{1-\pi}\,\exp\{2\alpha\, s(\pi)\}. \qquad (3.233)$$

It is straightforward to show,

$$q_0(\pi) = -2c\int_0^{\pi} \frac{\exp\{-2\alpha[\,s(\pi) - s(p)\,]\}}{p\,(1-p)^2}\,dp. \qquad (3.234)$$

Likewise, it is easy to show that,

$$J'(\pi) - K'(\pi) = \frac{-\pi}{2\alpha}\,J'(\pi) = -[\frac{\exp\{2\alpha\, s(\pi)\}}{(1-\pi)^2}] < 0 \quad \forall \pi \in (0,1) \qquad (3.235)$$

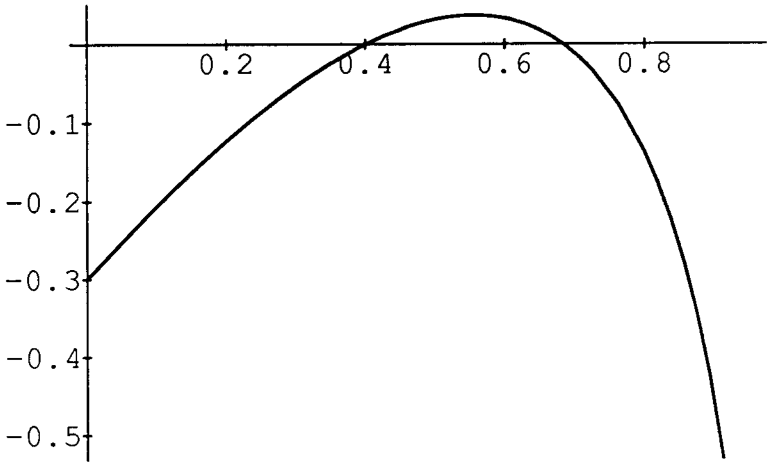

Figure 3.9: Graph of $s_{\mathrm{I}} = r_{\mathrm{I}} - e$.

and therefore since $J(0) = 0 = K(0)$ we obtain,

$$J(\pi) - K(\pi) < 0 \quad \forall \pi \in (0, 1). \tag{3.236}$$

As a result from 3.231 we obtain $r''_{\mathrm{I}}(\pi) < 0$ for all π in $(0, 1)$. In other words r_{I} is strictly concave. □

If we now define $s_{\mathrm{I}} := r_{\mathrm{I}} - e$, then from this proposition it follows that s_{I} is also strictly concave since e is just a line (see Figure 3.9). Thus the hypograph S of s_{I} is a convex set identical to its own convex hull, $\overline{S}$. In keeping with our discussion following the convexity analysis of the Stefan problem arising in the sequential detection case, we know that we seek a support function for $\overline{S}$ in $\ker D_{\Pi}$. From the definition of D_{Π} we see that its kernel consists of precisely the constant functions. Letting ℓ_* denote the graph of the unique line of support for $\overline{S}$ in $\ker D_{\Pi}$ which has zero slope we

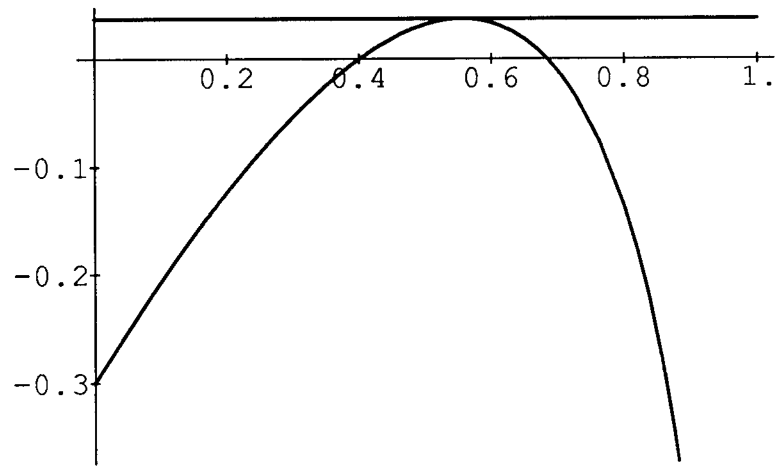

Figure 3.10: Graph of ℓ_* "supported" by s_{I}.

define

$$r_* := r_{\text{I}} - \ell_*. \tag{3.237}$$

Obviously then,

$$r_*(\pi) \leq e(\pi) \quad \forall \pi \in [0,1]. \tag{3.238}$$

Moreover, if we let b_* denote the π-coordinate of the point at which ℓ_* is tangent to $\overline{S}$ (see Figure 3.10) then,

$$\begin{aligned} r_*(b_*) &= r_{\text{I}}(b_*) - \ell_*(b_*) \\ &= e(b_*) + [\, r_{\text{I}}(b_*) - e(b_*)\,] - \ell_*(b_*) \\ &= e(b_*) + s_{\text{I}}(b_*) - \ell_*(b_*) \\ &= e(b_*), \end{aligned} \tag{3.239}$$

and (S2) is satisfied. Moreover, in view of the strict concavity of r_{I} it follows that r_* is also strictly concave and therefore b_* is the only point at which the

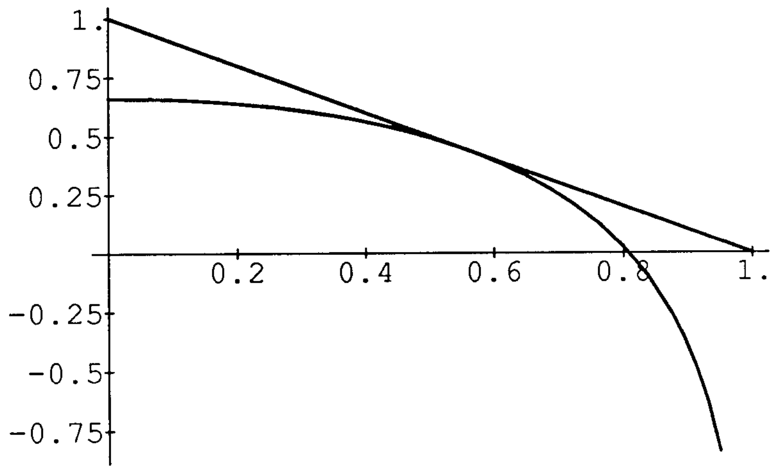

Figure 3.11: Graph of e and r_* with $b_* \approx 0.55607$.

inequality in 3.238 is not strict, i.e.,

$$r_*(\pi) < e(\pi) \qquad \forall \pi \neq b_*, \tag{3.240}$$

and this of course is (S3). By construction we know that r_* satifies (S4) and finally since $\ell_* \in \ker D_\Pi$ we obtain,

$$D_\Pi\, r_*(\pi) = D_\Pi\, r_{\text{I}}(\pi) - D_\Pi\, \ell_*(\pi) = -c\,\pi - 0 = -c \quad \forall \pi \in [0,1), \tag{3.241}$$

and hence we also have satisfied (S1). Thus, we can conclude that the pair (r_*, I_*), r_* as defined in 3.237 and $\text{I}_* = [0, b_*)$, satisfies (S1)–(S4) and therefore solves the problem posed by (V1)–(V4). As a result we are ready to apply the Verification Theorem.

3.7.6 Main Result

For easy reference, we redisplay the technical conditions involving the drift process and its P_1–projection onto the observations:

$$
\begin{aligned}
&\text{(H0):} \quad E_i \int_0^t H_s^2\, ds < \infty && \forall\, t \geq 0, \quad i = 0, 1; \\
&\text{(H1):} \quad E_i \int_0^\tau \hat{H}_s^2\, ds < \infty && \forall\, \tau \in \mathcal{T}, \quad i = 0, 1; \\
&\text{(H2):} \quad P_i \{ \int_0^\infty \hat{H}_s^2\, ds = \infty \} = 1 && i = 0, 1.
\end{aligned}
$$

We have come to the main result of this section.

Theorem 3.2 *Assume that the conditions* (H0), (H1),*and* (H2) *hold. In the problem of disruption based on observations of the process,*

$$Y_t = \int_0^t \Upsilon_s\, H_s\, ds + W_t^v \qquad t \geq 0,$$

with average running cost,

$$E_\pi \int_0^\tau \mathcal{C}_s\, ds = E_\pi \int_0^\tau c\, \Pi_s\, \hat{H}_s^2\, ds \quad c > 0,$$

and panic cost,

$$P_\pi\{\tau < v\} = E_\pi[\, 1 - \Pi_\tau\,],$$

there exists b_* unique, $0 < b_* < 1$, such that the first exit policy $(\tau^{\mathrm{I}_*}, \delta_*)$ based on the continuation interval $\mathrm{I}_* = [0, b_*)$ achieves Bayes' optimal cost, i.e.,

$$\rho_\pi(\tau^{\mathrm{I}_*}) = \inf_{\tau \in \mathcal{T}_{ad}} \rho_\pi(\tau) \qquad \forall\, \pi \in [0, 1],$$

where,

$$\rho_\pi(\tau) = E_\pi[\int_0^\tau \mathcal{C}_s\, ds + e(\Pi_\tau)\,] \quad \forall\, \pi \in [0,1] \text{ and } \tau \in \mathcal{T}_{ad}.$$

In addition, there exists $r_* \in \mathcal{C}^2[0,1)$, *the solution to* (S1)–(S4) *above, such that*

$$\rho_\pi(\tau^{\mathrm{I}_*}) = \begin{cases} r_*(\pi) & \textit{if } \pi \in \mathrm{I}_*; \\ e(\pi) & \textit{if } \pi \notin \mathrm{I}_*, \end{cases}$$

where,

$$e(\pi) = 1 - \pi \qquad \forall\, \pi \in [0,1]. \tag{3.242}$$

Proof: In the previous subsection we solved the associated Stefan problem and showed that there exists a pair (r_*, I_*) satisfying (S1)–(S4) for which the conditions on the cost coefficients guarantee that $0 < b_* < 1$. Therefore we have in fact exhibited a pair (r_*, I_*), $\mathrm{I}_* \in \mathcal{I}^0$ and $r_* \in \mathcal{C}^2[0,1)$, which satisfy (V1)–(V4).

To employ the Verification Theorem and therefore prove the theorem at hand it remains only to show that (C1), (C2), and (C3) hold since (E) follows from Proposition 3.10. With the above choice of running cost we see that (C1) follows from (H1) since,

$$\begin{aligned} E_\pi \int_0^\tau \mathcal{C}_s\, ds &= E_\pi \int_0^\tau c\, \Pi_s\, \hat{H}_s^2\, ds \\ &\leq E_\pi \int_0^\tau c\, \hat{H}_s^2\, ds \\ &= \pi\, E_1 \int_0^\tau c\, \hat{H}_s^2\, ds + (1-\pi)\, E_0 \int_0^\tau c\, \hat{H}_s^2\, ds < \infty. \end{aligned} \tag{3.243}$$

Since (H1) implies $P_\pi\{\int_0^t \hat{H}_s^2\, ds < \infty\} = 1$ for all $t \geq 0$ while (H2) impies $P_\pi\{\int_0^\infty \hat{H}_s^2\, ds = \infty\} = 1$, it is obvious that

$$P_\pi\{\int_t^\infty \hat{H}_s^2\, ds = \infty\} = 1 \quad \forall\, t \geq 0. \tag{3.244}$$

From this for all $n \geq 1$ we obtain,

$$\begin{aligned} P_\pi\{\int_0^\infty \mathcal{C}_s\, ds < \infty\} &= P_\pi\{\int_0^\infty c\,\Pi_s\, \hat{H}_s^2\, ds < \infty\} \\ &\leq P_\pi\{\int_n^\infty \Pi_s\, \hat{H}_s^2\, ds < \infty\} \\ &\leq P_\pi\{\inf_{n \leq s \leq \infty} \Pi_s \int_t^\infty \hat{H}_s^2\, ds < \infty\} \\ &= P_\pi\{\inf_{n \leq s \leq \infty} \Pi_s = 0\}. \end{aligned} \tag{3.245}$$

Since $\lim_{n\to\infty} P_\pi\{\inf_{n\leq s\leq\infty} \Pi_s = 0\} = 0$ we see that,

$$P_\pi\{\int_0^\infty \mathcal{C}_s\, ds = \infty\} = 1, \tag{3.246}$$

which is to say that (C2) follows. Finally, condition (C3) follows since the running cost $\mathcal{C}$ is linear in Π and hence concave in Π. □

3.7.7 Example

We end this section of the chapter with a concrete example of a sequential detection problem involving a diffusion with constant drift. We observe a stochastic process $Y = \{Y_t\}_{t\geq 0}$ which prior to a random jump time υ is a zero-drift diffusion and after the jump is a diffusion with unity drift:

$$\begin{aligned} &(\text{Before Jump}): \quad Y_t = W_t \qquad && t \geq 0; \\ &(\text{After Jump}): \quad Y_t = t + W_t \qquad && t \geq 0, \end{aligned}$$

where W is a $(\mathcal{G}_t, P_\pi)$–standard Wiener martingale. It is given that the jump occurs at time zero with prior probability $\pi \in [0,1]$ and conditional upon $\upsilon > 0$ the jump time is exponentially distributed with parameter $\lambda > 0$. Define the Bayes' cost,

$$\bar{\rho}_\pi(\tau) = E_\pi[\max\{0, \int_\upsilon^\tau c\,\hat{H}_s^2\, ds\}] + P_\pi\{\tau < \upsilon\}, \tag{3.247}$$

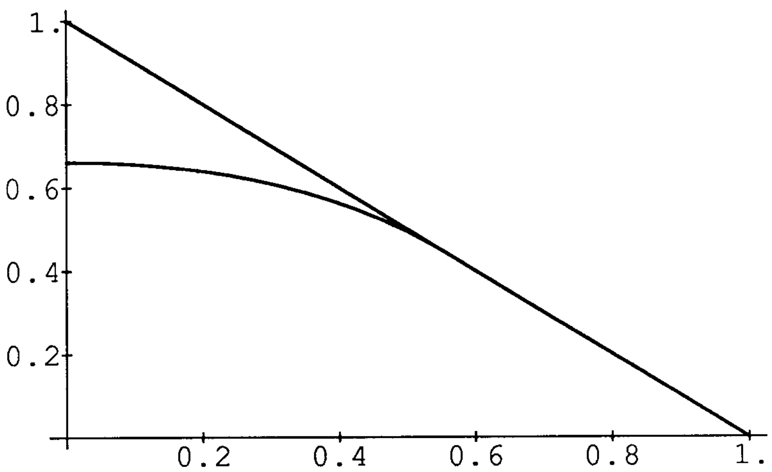

Figure 3.12: Graph of panic cost and Bayes' optimal risk.

where c is strictly positive. We are asked to minimize $\tilde{\rho}_\pi(\tau)$ over all P_π–a.s. finite stopping times.

We see that this is precisely the form we are equipped to handle as long as we make the identification $H \equiv 1$ for then our choice of running cost collapses down to

$$\int_0^\tau C_s\, ds = \int_0^\tau c\, \Pi_s \hat{H}_s\, ds = c \int_0^\tau \Pi_s\, ds. \tag{3.248}$$

We point out that each of (H0), (H1) and (H2) are trivially satisfied. Thus we can apply Theorem 3.2 to deduce that an optimal first exit policy exists. One can solve for the optimal interval I_* using the convexity approach given here or see [S, C4] for an integral equation to invert. The graphs for $\alpha = 1$ and $c = 1$ are depicted in Figures 3.7 through 3.11. The graph of ρ, the Bayes' optimal risk is given in Figure 3.12 along with the worst case risk.

Chapter 4

Change Detection: Point Process Data

4.1 Introduction

In this chapter we consider the problem of Bayesian optimal change detection when the observed data are modeled by a point process. The chapter has the following outline:

Section 1. SYSTEM DYNAMICS

In this section we make decisions concerning the general dynamics models for the two underlying systems. One system is modeled as a standard Poisson process, the other system is modeled as a general intensity process driving a point process. We impose technical growth conditions on the intensity process in order to apply filtering results and obtain the necessary escape properties on the *a posteriori* probability. We derive a representation for the likelihood ratio process using Girsanov's Theorem and then employ this representation to obtain a martingale description of the observation process with respect to the prior measure.

Section 2. OBSERVABLE DYNAMICS

The main point of this section is to obtain a martingale description of the observation process with respect to the prior measure when the intensity is conditioned upon the observation filtration. To do this we work with the observation process when the intensity is smoothed with respect to this filtration and derive the associated likelihood ratio in which only the smoothed intensity appears. Thus, this section is largely the analog of the previous but with the intensity conditioned upon the observation history. Likewise, it parallels Section 3.2 with the intensity in the role of the drift.

Section 3. THE CONDITIONAL PROBABILITY

The purpose here is to derive an explicit semimartingale representation for the *a posteriori* probability process by estimating the jump process for the time of change conditioned with respect to the observations. This smoothed representation gives us a filter for the state of the underlying jump.

Section 4. PRELUDE TO VERIFICATION

This section is in anticipation of the application of the Verification Theorem of Chapter 2 and thus its purpose is to compute a more explicit version of the first verification condition taking advantage of our specialization in this chapter to point process data.

Section 5. SEQUENTIAL DETECTION

This section is concerned with the classical Bayesian problem of sequential detection when the observations arise from one of two point processes. We show how this detection problem can be recast as a problem of change detection by properly modeling the jump time within the framework developed and by proper choice of the Bayes cost. We also consider the escape properties of the *a posteriori* probability process prior to setting up the Stefan problem implied by the Verification Theorem. Using a novel approach suggested by the convexity notions employed in Chapter 3 and the theory of functional–differential equations, we solve the Stefan problem and arrive at the one of the main results of the chapter: There exists Bayes' optimal first exit policy for the problem of sequential detection with elapsed energy cost and terminal error penalties. We conclude the section with an example involving a constant intensity Poisson process.

4.2 System Dynamics

As in Chapter 2 we are working on the measurable space $(\Omega, \mathcal{A})$ equipped with two probability measures P_0 and P_1 which induce the family of Bayes' measures $\{P_\pi : 0 \leq \pi \leq 1\}$. In this chapter we consider the general problem of change detection when the underlying dynamics give rise to point process data. Mathematically speaking, the methods of modern martingale analysis make this chapter largely a rewrite of the last. We shall draw freely upon the results and notations of Chapter 2 and Chapter 3 with only the briefest reminders. As in Chapter 3 it remains to specify the nature of $\mathcal{O}_t$, $\mathcal{G}_t$, and $\mathcal{A}_t$, to calculate $L = \{L_t\}_{t\geq 0}$, to make choices specializing the $\mathcal{F}_t$–conditional distribution F of the disruption time υ, and lastly to specify the running cost and terminal cost functions. Once these things are done we can make use of the Verification Theorem to solve for the optimal first exit policy.

We begin by supposing that we are given $J = \{J_t\}_{t\geq 0}$, a continuous random process on $(\Omega, \mathcal{A})$ called the **jump intensity** (or simply the **intensity** so as not to confuse it with the compensator of the jump process Υ) for which $\log J$ is a nonnegative process. Define the intensity filtration on $\mathcal{A}$ via,

$$\mathcal{J}_t := \bigvee_{0\leq s\leq t} \sigma(J_s) \qquad \forall\, t \geq 0, \tag{4.1}$$

and take this filtration as $(\mathcal{A}, P_0)$–completed. We are also given $N = \{N_t\}_{t\geq 0}$, $N_0 \equiv 0$, a P_π–nonexplosive point process on $(\Omega, \mathcal{A})$ called the **observation**. Define the **observation filtration** on $\mathcal{A}$ via,

$$\mathcal{O}_t := \bigvee_{0\leq s\leq t} \sigma(N_s) \qquad \forall\, t \geq 0, \tag{4.2}$$

and take this filtration as $(\mathcal{A}, P_0)$–completed. For any $\mathcal{A}_t$–adapted process $X = \{X_t\}_{t\geq 0}$ we recall that the $\mathcal{O}_t$–adapted process $\hat{X} = \{\hat{X}_t\}_{t\geq 0}$ is defined via,

$$\hat{X}_t := E_1[\,X_t \,|\, \mathcal{O}_t\,] \qquad \forall t \geq 0. \tag{4.3}$$

We collect the following technical conditions involving the intensity process J and its $(\mathcal{O}_t, P_1)$–smoothing $\hat{J}$:

$$\begin{aligned}
&\text{(J0):}\quad E_i \int_0^t J_s\, ds < \infty && \forall t \geq 0,\quad i = 0,1;\\
&\text{(J1):}\quad E_i \int_0^t \hat{J}_s^2\, ds < \infty && \forall t \geq 0,\quad i = 0,1;\\
&\text{(J2):}\quad P_i\{\int_0^\infty (\hat{J}_s - 1)^2\, ds = \infty\} = 1 && i = 0,1.
\end{aligned}$$

Next we define $\mathcal{G}_t := \mathcal{O}_t \bigvee \mathcal{J}_t$ and note that J being both $\mathcal{G}_t$–adapted and (left) continuous is therefore also $\mathcal{G}_t$–predictable. We make the following assumptions concerning the dynamics model for the observation process under P_0 and P_1:

$$\begin{aligned}
&\text{(DM0)}: && N_t \quad \text{is a } (\mathcal{G}_t, P_0)\text{–Poisson counting process;}\\
&\text{(DM1)}: && N_t - \int_0^t J_s\, ds \quad \text{is a } (\mathcal{G}_t, P_1)\text{–martingale.}
\end{aligned}$$

Recalling that v denotes a measurable mapping from Ω to $[0,\infty]$, that $\Upsilon_t := 1\{v \leq t\}$, and that $\mathcal{U}_t = \bigvee_{s\leq t} \sigma(\Upsilon_t)$ we define,

$$\mathcal{A}_t := \mathcal{G}_t \bigvee \mathcal{U}_t \bigvee \mathcal{A}_0 \qquad \forall t > 0, \tag{4.4}$$

where $\mathcal{A}_0$ denotes the smallest σ–algebra containing both the P_0–null sets of $\mathcal{A}$ and the P_1–null sets of $\mathcal{A}$. We have come to our first proposition.

Proposition 4.1 *The $\mathcal{G}_t$–adapted process L is given by,*

$$L_t = \exp\{\int_0^t \log J_s\, dN_s - \int_0^t (J_s - 1)\, ds\} \quad \forall t \geq 0,\ P_0\text{–}a.s.$$

Proof: The proof of this representation is well-known and standard but as in Chapter 3 we include it for completeness and because some of its ingredients are reused in succeeding propositions. Applying the generalized Itô rule [W&H, P6.6.2] to the natural logarithm of the $(\mathcal{G}_t, P_0)$–martingale $L = \{L_t\}_{t\geq 0}$ with $L_0 \equiv 0$ we obtain,

$$\log\, L_t = \int_0^t L_{s-}^{-1}\, dL_s - \frac{1}{2}\int_0^t L_s^{-2}\, d[L,L]_s^c + \sum_{s\leq t}\left[\log\left(1+\frac{\Delta L_s}{L_{s-}}\right) - \frac{\Delta L_s}{L_{s-}}\right] \tag{4.5}$$

for all $t \geq 0$, P_0–a.s. Define $X = \{X_t\}_{t\geq 0}$ via,

$$X_t := L^{-1} \bullet L_t = \int_0^t L_{s-}^{-1}\, dL_s \qquad \forall\, t \geq 0, \tag{4.6}$$

and note that,

$$\Delta X_t = \frac{\Delta L_t}{L_{t-}} = L_{t-}^{-1}\, \Delta L_t \quad \forall\, t \geq 0. \tag{4.7}$$

Because L is a $(\mathcal{G}_t, P_0)$–martingale and $\{L_{t-}^{-1}\}_{t>0}$ is a $(\mathcal{G}_t, P_0)$–predictable process which according to Proposition 2.6 is locally bounded, we see that X is a $(\mathcal{G}_t, P_0)$–local martingale with $X_0 \equiv 0$ and therefore we may employ the Martingale Representation Theorem [W&H, P6.7.3] to conclude that X has the stochastic integral representation (let $\eta_t := N_t - t$),

$$X_t = \int_0^t \xi_s\, d\eta_s \qquad \forall\, t \geq 0,\ P_0\text{–a.s.}, \tag{4.8}$$

for some $\mathcal{G}_t$–predictable process $\xi = \{\xi_t\}_{t\geq 0}$ satisfying $\int_0^t |\xi_s|\, ds < \infty$, P_0–a.s., for all $t \geq 0$. From this it follows that,

$$L^{-2} \bullet [L,L]^c = [L^{-1} \bullet L, L^{-1} \bullet L]^c = [X,X]^c \equiv 0, \tag{4.9}$$

i.e.,

$$\int_0^t L_s^{-2}\, d[L,L]_s^c = 0 \qquad \forall\, t \geq 0,\ P_0\text{–a.s.} \tag{4.10}$$

Combining 4.5, 4.7 and 4.10 gives,

$$\log L_t = X_t + \sum_{s \leq t} \left[\log\left(1 + \Delta X_s\right) - \Delta X_s\right], \tag{4.11}$$

and therefore,

$$L_t = \exp\{X_t + \sum_{s \leq t} [\log\left(1 + \Delta X_s\right) - \Delta X_s]\}. \tag{4.12}$$

It only remains to compute an $(\mathcal{O}_t, P_0)$–representation for ξ to obtain the result. Since $< \eta, X >= [\eta, X] = [X, \eta] = \xi \bullet [\eta, \eta]$ we see that,

$$< \eta, X >_t = \xi \bullet [\eta, \eta]_t = \xi \bullet [N, N]_t = \int_0^t \xi_s\, ds. \tag{4.13}$$

Because both X and η are $(\mathcal{G}_t, P_0)$–local martingales we may apply the abstract Girsanov Theorem [W&H, P6.7.2] and so conclude that $\eta - < \eta, X >$ is a $(\mathcal{G}_t, P_1)$–local martingale. But then notice,

$$\begin{aligned} \eta_t - < \eta, X >_t &= N_t - t - \int_0^t \xi_s\, ds \\ &= N_t - \int_0^t (\xi_s + 1)\, ds, \end{aligned} \tag{4.14}$$

so that employing (DM1) and the fact that predictable compensators are unique implies $J_t = \xi_t + 1$ for all $t \geq 0$, P_0–a.s. Hence,

$$X_t = \int_0^t \xi_s\, d\eta_s = \int_0^t (J_s - 1)\, d\eta_s, \tag{4.15}$$

and from this we see,

$$\Delta X_t = (J_t - 1)\Delta \eta_t = (J_t - 1)\Delta N_t \quad \forall\, t \geq 0. \tag{4.16}$$

Plugging these expressions into 4.5 yields,

$$L_t = \exp\{\int_0^t (J_s - 1)\, d\eta_s + \sum_{s\leq t} [\log(1 + (J_s - 1)\Delta N_s) - (J_s - 1)\,\Delta N_s]\}. \tag{4.17}$$

To simplify this last expression note that,

$$\sum_{s\leq t}(J_s - 1)\,\Delta N_s = \sum_{\substack{s\leq t\\ \Delta N_s = 1}} (J_s - 1) = \int_0^t (J_s - 1)\, dN_s, \tag{4.18}$$

and therefore,

$$\sum_{s\leq t}(J_s - 1)\,\Delta N_s = \int_0^t (J_s - 1)\, d\eta_s + \int_0^t (J_s - 1)\, ds. \tag{4.19}$$

Also observe that,

$$\sum_{s\leq t} \log(1 + (J_s - 1)\Delta N_s) = \sum_{\substack{s\leq t\\ \Delta N_s = 1}} \log J_s = \int_0^t \log J_s\, dN_s. \tag{4.20}$$

From these expressions we obtain,

$$L_t = \exp\{\int_0^t \log J_s\, dN_s - \int_0^t (J_s - 1)\, ds\}, \tag{4.21}$$

which completes the proof. □

The next proposition which we shall give mirrors Proposition 3.2 in that it exploits the random measure Q^v as a device to obtain a martingale description for the system dynamics with respect to the P_π–measure on $(\Omega, \mathcal{A})$. We begin with the following lemma, the logical companion to Lemma 3.2.

Lemma 4.1 *For each $u \in [0, \infty]$,*

$$\eta_t - \int_0^t U_s\,(J_s - 1)\, ds \;\; \textit{is a } (\mathcal{G}_t, Q^u)\textit{–martingale,}$$

where U denotes the deterministic indicator process $U_t := 1\{u \leq t\}$.

Proof: For $u \in [0, \infty]$ define the auxiliary $(\mathcal{G}_t, P_0)$–local martingale M^u,

$$M_t^u := \int_0^t U_s\,(J_s - 1)\,d\eta_s \qquad \forall\, t \geq 0, \tag{4.22}$$

and the $(\mathcal{G}_t, P_0)$–adapted process L^u,

$$L_t^u := \exp\{M_t^u + \sum_{s \leq t}[\log(1 + \Delta M_s^u) - \Delta M_s^u]\} \qquad \forall\, t \geq 0. \tag{4.23}$$

Computing under P_0 we obtain,

$$< \eta, M^u >_t = [\eta, M^u]_t = \int_0^t U_s\,(J_s - 1)\,ds. \tag{4.24}$$

From Girsanov's Theorem we know that $\eta - < \eta, M^u >$ is a $(\mathcal{G}_t, Q^u)$–local martingale and this gives us the result. □

We now give the proposition describing the unobservable dynamics of the N process on $(\Omega, \mathcal{A})$ with respect to the P_π measure for any π in $[0, 1]$ fixed.

Proposition 4.2

$N_t - \int_0^t [\Upsilon_s \cdot J_s + (1 - \Upsilon_s) \cdot 1]\; ds$ *is an* $(\mathcal{A}_t, P_\pi)$*–martingale* $\quad \forall\, \pi \in [0, 1]$.

Proof: Fix π in $[0, 1]$. Define the $\mathcal{A}_t$–adapted process $\eta^v = \{\eta_t^v\}_{t \geq 0}$ via,

$$\eta_t^v := \eta_t - \int_0^t \Upsilon_s\,(J_s - 1)\,ds \qquad \forall\, t \geq 0, \tag{4.25}$$

and for u in $[0, \infty]$ define the $\mathcal{G}_t$–adapted process $\eta^u = \{\eta_t^u\}_{t \geq 0}$ by,

$$\eta_t^u := \eta_t - \int_0^t U_s\,(J_s - 1)\,ds \qquad \forall\, t \geq 0. \tag{4.26}$$

Our immediate goal is to prove that η^v is an $(\mathcal{A}_t, P_\pi)$–martingale; note in this notation that Lemma 4.2 says η^u is a $(\mathcal{G}_t, Q^u)$–martingale for all u in $[0, \infty]$.

With η here in the role of Y in Proposition 3.2 and similarly $(J-1)$, η^v, η^u in the roles of H, W^v, W^u, respectively, the same logical argument employed in Proposition 3.2 to show that W^v is an $(\mathcal{A}_t, P_\pi)$–martingale works here to show that,

$$E_\pi[\,\eta_t^v \,|\, \mathcal{A}_r\,] = \eta_r^v \quad \forall\, r \in [0,t],\ t \geq 0,\ P_\pi\text{–a.s.,} \tag{4.27}$$

and in words, η^v is an $(\mathcal{A}_t, P_\pi)$–martingale. This tells us that η has the semimartingale dynamics

$$\eta_t = \int_0^t \Upsilon_s\,(J_s - 1)\,ds + \eta_t^v \qquad \forall\, t \geq 0, \tag{4.28}$$

for the $(\mathcal{A}_t, P_\pi)$–martingale $\eta^v = \{\eta_t^v\}_{t\geq 0}$. Since $\eta_t = N_t - t$ for all $t \geq 0$ this gives us the result. □

4.3 Observable Dynamics

The main goal of this section is to show that,

$$N_t - \int_0^t \Big[\Pi_s \cdot \hat{J}_s + (1-\Pi_s)\cdot 1\Big]\, ds \text{ is an } (\mathcal{O}_t, P_\pi)\text{–martingale.} \qquad (4.29)$$

Recall for all $t \geq 0$ that the $\mathcal{O}_t$–adapted process $\check{X} = \{\check{X}_t\}_{t\geq 0}$ is given by, $\check{X}_t = E_0[\, X_t \,|\, \mathcal{O}_t\,]$. The following proposition gives us a representation for $\check{L}$.

Proposition 4.3 *The $\mathcal{O}_t$–adapted process $\check{L}$ is given by,*

$$\check{L}_t = \exp\{\int_0^t \log \hat{J}_{s-}\, dN_s - \int_0^t (\hat{J}_s - 1)\, ds\} \qquad \forall\, t \geq 0.$$

Proof: The proof of this proposition is the analog of the proofs of both Proposition 4.1 and Proposition 3.3. With X defined here as $X := \check{L}^{-1} \bullet \check{L}$ we obtain by similar arguments that,

$$\check{L}_t = \exp\{X_t + \sum_{s\leq t} [\log(1+\Delta X_s) - \Delta X_s]\}, \qquad (4.30)$$

with X expressible as,

$$X_t = \int_0^t \xi_s\, d\eta_s \qquad \forall\, t \geq 0,\ P_0\text{–a.s.}, \qquad (4.31)$$

for some $\mathcal{O}_t$–predictable process $\xi = \{\xi_t\}_{t\geq 0}$ satisfying $\int_0^t |\xi_s|\, ds < \infty$, P_0–a.s., for all $t \geq 0$. Applying Girsanov's Theorem [W&H, P6.7.2] we likewise conclude that $\eta - < \eta, X >$ is an $(\mathcal{O}_t, P_1)$–local martingale and therefore by the uniqueness of predictable compensators we obtain $\xi_t = \hat{J}_{t-} - 1$, for all $t \geq 0$, P_0–a.s. Therefore,

$$X_t = \int_0^t \check{L}_{s-}^{-1}\, d\check{L}_s = \int_0^t (\hat{J}_{s-} - 1)\, d\eta_s, \qquad (4.32)$$

and this completes the proof. □

Next we prove the following logical analog to Lemma 3.2 and Lemma 4.2.

Lemma 4.2 *For each $u \in [0, \infty]$,*

$$\eta_t - \int_0^t U_s (\hat{J}_s - 1)\, ds \ \ \textit{is an } (\mathcal{O}_t, Q^u)\textit{-martingale,}$$

where U denotes the deterministic indicator process $U_t := 1\{u \leq t\}$.

Proof: From Lemma 4.2 we know that,

$$\eta_t - \int_0^t U_s (J_s - 1)\, ds \text{ is a } (\mathcal{G}_t, Q^u)\text{-martingale.} \tag{4.33}$$

Thus, a routine application of the Projection Theorem yields the result. □

This brings us to the following proposition which describes the dynamics of N on $(\Omega, \mathcal{O})$ with respect to the P_π measure for any fixed π in $[0, 1]$.

Proposition 4.4

$$N_t - \int_0^t \Big[\Pi_s \cdot \hat{J}_s + (1 - \Pi_s) \cdot 1 \Big]\, ds \ \ \textit{is an } (\mathcal{O}_t, P_\pi)\textit{-martingale} \quad \forall \pi \in [0, 1].$$

Proof: If we define,

$$\tilde{\eta}_t^v := \eta_t - \int_0^t \Upsilon_s (\hat{J}_s - 1)\, ds \quad \forall t \geq 0, \tag{4.34}$$

then proof of this proposition follows the same argument as Proposition 3.4 with $\tilde{\eta}^v$ in the role of $\tilde{W}^v$. □

Let's take this opportunity to summarize the results of this section and the last. We have given two models for the total dynamics of the observable process N under two different probability measures:

$$N_t \quad \text{is a } (\mathcal{G}_t, P_0)\text{–Poisson counting process,} \tag{4.35}$$

and,

$$N_t - \int_0^t J_s\, ds \quad \text{is a } (\mathcal{G}_t, P_1)\text{-martingale.} \tag{4.36}$$

As in Chapter 3 we interpret each measure as modeling a distinct mode of operation of some underlying dynamical system on $(\Omega, \mathcal{G})$—the difference in this chapter is that we partially observe these underlying dynamics through a counting process rather than a diffusion. Under P_0 the hidden dynamics are modeled as merely a Poisson process which influence the observations directly. Under P_1 the hidden dynamics are modeled via $J = \{J_t\}_{t \geq 0}$ and these dynamics influence the observations through the 'signal plus noise' set-up. The $\mathcal{G}_t$–progressive, $\mathcal{O}_t$–partially observable dynamics themselves, J, may arise according to any number of models, for instance, a memoryless nonlinear transformation of a process with a linear stochastic differential description.

For all G in $\mathcal{G}$ note that $Q^0\{G\} = P_1\{G\}$ while $Q^\infty\{G\} = P_0\{G\}$ and once again we see that the transition measure Q^u 'links' the two distinct modes of system dynamics. Indeed, if the time of mode change were deterministic then $\{Q^u : 0 \leq u \leq \infty\}$ provides us with a tool to answer questions concerning the (still stochastic) behavior of the underlying system. However, we are interested in problems where the overall system has observable dynamics which can change at a random time v from say mode-0 modeled according 4.35 to mode-1, a different mode in which the observable dynamics are properly modeled using 4.36 above. Moreover, in this problem we assume there is π-probability that the system is initially in mode-1 and $(1-\pi)$–probability that the jump time v is positive and distributed according to the

$\mathcal{O}_t$–conditional cumulative distribution function F. On the probability space $(\Omega, \mathcal{A}, P_\pi)$ which we have constructed in Chapter 2 to model this situation, the observable dynamics have a single representation—they behave according to,

$$N_t - \int_0^t [\Upsilon_s \cdot J_s + (1 - \Upsilon_s) \cdot 1] \, ds \quad \text{is an } (\mathcal{A}_t, P_\pi)\text{–martingale,} \tag{4.37}$$

which is the conclusion to Proposition 4.2.

To complete our summary, we note that we also have two models for the partially–observed dynamics of N under both P_0 and P_1:

$$N_t \quad \text{is an } (\mathcal{O}_t, P_0)\text{–Poisson process,} \tag{4.38}$$

and,

$$N_t - \int_0^t \hat{J}_s \, ds \quad \text{is an } (\mathcal{O}_t, P_1)\text{–martingale.} \tag{4.39}$$

Using these representations, in Proposition 4.4 we obtained the analogous partially–observable representation for the dynamics on $(\Omega, \mathcal{A}, P_\pi)$ as,

$$N_t - \int_0^t \left[\Pi_s \cdot \hat{J}_s + (1 - \Pi_s) \cdot 1\right] \, ds \quad \text{is an } (\mathcal{O}_t, P_\pi)\text{–martingale.} \tag{4.40}$$

This last martingale is the one we need in this chapter to obtain explicit filtering results involving the *a posteriori* probability. We take up this task in the section to follow.

We end the current section with a result concerning an escape property of the likelihood ratio process $\check{L} = \{\check{L}_t\}_{t \geq 0}$ which follows directly from the condition (J1) imposed on the drift process at the outset of this chapter.

Proposition 4.5 *Under assumption* (J1),

$$P_\pi\{\sup_{0 \leq t \leq n} |\log \check{L}_t| = \infty\} = 0 \quad \forall\, n \geq 1, \ \forall\, \pi \in \mathrm{I}_\infty.$$

Proof: Fix $\pi \in \mathrm{I}_\infty$. From Proposition 4.3 we know that the $\mathcal{O}_t$–adapted process $\log \check{L}$ is given by,

$$\log \check{L}_t = \int_0^t \log \hat{J}_{s-}\, dN_s - \int_0^t (\hat{J}_s - 1)\, ds \quad \forall\, t \geq 0, \tag{4.41}$$

so that computing under P_π we obtain (see Proposition 4.4),

$$|\log \check{L}_t| \leq |\int_0^t (\hat{J}_{s-} - 1)\, dN_s + \int_0^t [1 + (\hat{J}_s - 1)^2]\, ds, \tag{4.42}$$

using the simple inequalities $\log x \leq x - 1 \leq x^2$ and the fact that dN is a positive (random) measure. Hence for all $n \geq 1$,

$$\begin{aligned}
P_\pi\{\sup_{0\leq t\leq n} |\log \check{L}_t| = \infty\} &\leq P_\pi\{\int_0^n (\hat{J}_{s-} - 1)\, dN_s \\
&\qquad + \int_0^n [1 + (\hat{J}_s - 1)^2]\, ds = \infty\} \\
&= P_\pi\{\int_0^n (\hat{J}_{s-} - 1)\, dN_s + \int_0^n (\hat{J}_s - 1)^2\, ds = \infty\} \\
&= P_\pi\{\int_0^n (\hat{J}_{s-} - 1)\, dN_s = \infty\} \\
&\leq P_\pi\{\int_0^n (\hat{J}_{s-} - 1)\, dN_s \geq m\} \quad \forall\, m \geq 1,
\end{aligned} \tag{4.43}$$

where the second to last line follows *a fortiori* from (J1). It is not difficult to show that $(\hat{J} - 1) \bullet N_t$ is a right-continuous $(\mathcal{O}_t, P_\pi)$–submartingale and therefore we may apply a basic submartingale inequality [L&S1, T3.2] to the right-hand side of 4.43 and find,

$$P_\pi\{\sup_{0\leq t\leq n} |\log \check{L}_t| = \infty\} \leq \frac{E_\pi \int_0^n (\hat{J}_{s-} - 1)\, dN_s}{m} \quad \forall\, m \geq 1. \tag{4.44}$$

Employing Proposition 4.4 we obtain

$$E_\pi \int_0^n (\hat{J}_{s-} - 1)\, dN_s \;=\; E_\pi \int_0^n [1 + \Pi_s\, (\hat{J}_s - 1)\,](\hat{J}_s - 1)\, ds$$

$$
\begin{aligned}
&= E_\pi \int_0^n (\hat{J}_s - 1)\, ds + E_\pi \int_0^n \Pi_s\, (\hat{J}_s - 1)^2\, ds \\
&\leq n + 2E_\pi \int_0^n (\hat{J}_s - 1)^2\, ds. \qquad (4.45)
\end{aligned}
$$

Combining the last two expressions and then (J1) again yields,

$$
P_\pi\{ \sup_{0 \leq t \leq n} |\log \check{L}_t| = \infty\} \leq \lim_{m \to \infty} \frac{n + 2E_\pi \int_0^n (\hat{J}_s - 1)^2\, ds}{m} = 0, \qquad (4.46)
$$

and therefore,

$$
P_\pi\{ \sup_{0 \leq t \leq n} |\log \check{L}_t| = \infty\} = 0 \quad \forall n \geq 1, \qquad (4.47)
$$

which completes the proof. □

4.4 The Conditional Probability

Fix $\pi \in [0,1]$. We conveniently collect some of the results obtained so far to compute a more explicit representation for the projection of Υ_t onto $\mathcal{O}_t$, taking the point process nature of the observations into account.

1. From 2.116 we know that the single-jump, binary point process Υ has the $(\mathcal{A}_t, P_\pi)$–semimartingale representation,

$$\Upsilon_t = \Upsilon_0 + \int_0^t (1-\Upsilon_{s-})\,(1-F_s)^{-1}\,dF_s + M_t \quad \forall\, t \geq 0,\; P_\pi\text{–a.s.}, \tag{4.48}$$

   with Υ_0 an $\mathcal{A}_0$–measurable binary random variable satisfying $E_\pi \Upsilon_0 = \pi$, with $(1-\Upsilon_{s-})$ an $\mathcal{A}_t$–predictable process, and with M an $(\mathcal{A}_t, P_\pi)$–martingale.

2. From Proposition 4.2 the observation process satisfies,

$$N_t = \int_0^t [\Upsilon_s \cdot J_s + (1-\Upsilon_s)\cdot 1]\; ds + \eta_t^v \quad \forall\, t \geq 0,\; P_\pi\text{–a.s.} \tag{4.49}$$

   for J some $\mathcal{G}_t$–predictable process such that,

$$E_\pi \int_0^t J_s\, ds < \infty \quad \forall\, t \geq 0, \tag{4.50}$$

   where η^v denotes an $(\mathcal{A}_t, P_\pi)$–martingale. The bounding in 4.50 follows from assumption (J0) and expression 2.96.

The following proposition gives the filter for the projection of Υ_t onto $\mathcal{O}_t$.

Proposition 4.6 *The filter for* $\Pi_t = E_\pi[\,\Upsilon_t \,|\, \mathcal{O}_t\,]$ *is given by,*

$$\Pi_t = \Pi_0 + \int_0^t (1-\Pi_s)\,(1-F_s)^{-1}\,dF_s + \int_0^t \frac{(1-\Pi_{s-})\,\Pi_{s-}\,(\hat{J}_{s-}-1)}{1+\Pi_{s-}(\hat{J}_{s-}-1)}\,d\overline{\eta}_s^v,$$

for all $t \geq 0$, P_π–a.s., where $\overline{\eta}^v$ is an $(\mathcal{O}_t, P_\pi)$–martingale; we recall that by definition $\hat{J}_t = E_1[J_t | \mathcal{O}_t]$.

Proof: Define,

$$\lambda_t := \Upsilon_t \cdot J_t + (1 - \Upsilon_t) \cdot 1 \quad \forall t \geq 0, \tag{4.51}$$

so that we may write,

$$N_t = \int_0^t \lambda_s \, ds + \eta_t^v \quad \forall t \geq 0. \tag{4.52}$$

Applying a filtering theorem [W&H, P7.4.1] to compute $\Pi = E_\pi[\Upsilon | \mathcal{O}]$ we obtain,

$$\Pi_t = \Pi_0 + \int_0^t (1 - \Pi_s)(1 - F_s)^{-1} \, dF_s + \int_0^t \Phi_{s-} \, d\overline{\eta}_s^v \tag{4.53}$$

for all $t \geq 0$, P_π–a.s., where $\overline{\eta}^v$ is some $(\mathcal{O}_t, P_\pi)$–martingale and where Φ is an $\mathcal{O}_t$–progressive process given by

$$\Phi_t := E_\pi[\phi_t | \mathcal{O}_t] + \frac{E_\pi[\Upsilon_t(\lambda_t - E_\pi[\lambda_t | \mathcal{O}_t]) | \mathcal{O}_t]}{E_\pi[\lambda_t | \mathcal{O}_t]}, \tag{4.54}$$

for all $t \geq 0$ and for some $\mathcal{A}_t$–predictable process ϕ satisfying,

$$\langle M, \eta^v \rangle_t = \int_0^t \phi_s \, ds \qquad \forall t \geq 0, \, P_\pi\text{–a.s.} \tag{4.55}$$

Let's compute Φ starting with $E_\pi[\phi_t | \mathcal{O}_t]$. Recall that $\langle M, \eta^v \rangle$ denotes the predictable compensator of $[M, \eta^v]$, the co-quadratic variation of $M \eta^v$. Observe that the additive noise η^v in the observation N has no jumps in common with M, the zero-mean martingale driving Υ which quite obviously has sample paths of locally finite variation. Therefore, $[M, \eta^v] \equiv \langle M, \eta^v \rangle \equiv 0$. Hence we can take $\phi \equiv 0$ so that $E_\pi[\phi_t | \mathcal{O}_t] \equiv 0$.

Next, use the fact that $\Upsilon^2 \equiv \Upsilon$ and $\Upsilon(1-\Upsilon) \equiv 0$ to find,

$$\begin{aligned} E_\pi[\,\Upsilon_t\,\lambda_t\,|\,\mathcal{O}_t] &= E_\pi[\,\Upsilon_t(\Upsilon_t \cdot J_t + (1-\Upsilon_t)\cdot 1)\,|\,\mathcal{O}_t] \\ &= E_\pi[\,\Upsilon_t\,J_t\,|\,\mathcal{O}_t]. \end{aligned} \tag{4.56}$$

To compute $E_\pi[\,\Upsilon_t\,J_t\,|\,\mathcal{O}_t]$, pick any O in $\mathcal{O}_t$ and use expression 2.106 to obtain,

$$\begin{aligned} E_\pi[\Upsilon_t\,1_O\,J_t] &= \pi\,E_1 1_O\,J_t + (1-\pi)\,E_0 \int_0^t \int_O J_t\,dQ^u\,dF_u \\ &= \pi\,E_1[1_O\,E_1[\,J_t\,|\,\mathcal{O}_t]\,] + (1-\pi)\,E_0 \int_0^t \int_O L_u^{-1}\,J_t\,dP_1\,dF_u \\ &= \pi\,E_1 1_O\,\hat{J}_t + (1-\pi)\,E_0 \int_0^t \int_O L_u^{-1}\,\hat{J}_t\,dP_1\,dF_u \\ &= \pi\,E_1 1_O\,\hat{J}_t + (1-\pi)\,E_0 \int_0^t \int_O \hat{J}_t\,dQ^u\,dF_u \\ &= E_\pi[\Upsilon_t\,1_O\,\hat{J}_t] \\ &= E_\pi[\Pi_t\,1_O\,\hat{J}_t], \end{aligned} \tag{4.57}$$

and therefore,

$$E_\pi[\Upsilon_t\,J_t\,|\,\mathcal{O}_t] = \Pi_t\,\hat{J}_t \qquad \forall\, t \geq 0,\ P_\pi\text{–a.s.} \tag{4.58}$$

From this we see that,

$$E_\pi[\,\Upsilon_t\,\lambda_t\,|\,\mathcal{O}_t] = \Pi_t\,\hat{J}_t \qquad \forall\, t \geq 0,\ P_\pi\text{–a.s.}, \tag{4.59}$$

and also,

$$\begin{aligned} E_\pi[\,\lambda_t\,|\,\mathcal{O}_t] &= E_\pi[\,\Upsilon_t \cdot J_t + (1-\Upsilon_t)\cdot 1\,|\,\mathcal{O}_t] \\ &= \Pi_t \cdot \hat{J}_t + (1-\Pi_t)\cdot 1 \\ &= 1 + \Pi_t(\hat{J}_t - 1) \quad \forall\, t \geq 0,\ P_\pi\text{–a.s.} \end{aligned} \tag{4.60}$$

Combining these results yields,

$$\Phi_t = \frac{(1-\Pi_t)\,\Pi_t\,(\hat{J}_t - 1)}{1+\Pi_t(\hat{J}_t - 1)} \quad \forall\, t \geq 0,\ P_\pi\text{-a.s.}, \tag{4.61}$$

and plugging this into 4.53 yields the desired expression. □

An easy consequence of this proposition is the following equivalent representation for the *a posteriori* probability process which cleanly separates out the jumps,

$$\begin{aligned} \Pi_t - \Pi_0 \;&=\; \int_0^t (1-\Pi_s)\,(1-F_s)^{-1}\,dF_s - \int_0^t (1-\Pi_s)\,\Pi_s\,(\hat{J}_s - 1)\,ds \\ &\quad + \sum_{\substack{s<t \\ \Delta N_s = 1}} \frac{(1-\Pi_{s-})\,\Pi_{s-}\,(\hat{J}_{s-} - 1)}{1+\Pi_{s-}\,(\hat{J}_{s-} - 1)}. \end{aligned} \tag{4.62}$$

4.5 Prelude to Verification

For $\mathrm{I} \in \mathcal{I}_\infty$ define $\mathcal{PC}^+(\mathrm{I})$ to be the class of piecewise continuous functions which are right-continuous on I except at most a countable number of isolated points and which lose their right-continuity at these isolated points only in the following way: if $r(\pi+) \neq r(\pi-)$ then,

$$r(\pi) = r(\pi+) < r(\pi-). \tag{4.63}$$

Let $\mathcal{BC}^{1+}(\mathrm{I})$ denote the subclass of functions in $\mathcal{PC}^+(\mathrm{I})$ which are bounded on I as well as once right-continuously differentiable there except at a countable number of isolated points. Now define,

$$\mathcal{BC}^{1+}(\mathcal{I}_\infty) := \bigcap_{\mathrm{I} \in \mathcal{I}_\infty} \mathcal{BC}^{1+}(\mathrm{I}). \tag{4.64}$$

Simple examples of functions in $\mathcal{BC}^{1+}(\mathcal{I}_\infty)$ are,

$$r(\pi) := \begin{cases} 1 & \text{for } 0 \leq \pi < \frac{1}{2}; \\ 0 & \text{for } \frac{1}{2} \leq \pi \leq 1, \end{cases} \tag{4.65}$$

and,

$$r(\pi) := \begin{cases} \pi & \text{for } 0 \leq \pi < \frac{1}{2}; \\ 1 - \pi & \text{for } \frac{1}{2} \leq \pi \leq 1, \end{cases} \tag{4.66}$$

and also the product of these two mappings. In the next section we shall be interested in a particular function $r_* \in \mathcal{BC}^{1+}(\mathcal{I}_\infty)$. Using the filter for Π developed in the last section our goal is to compute the expectation, $E_\pi[\, r(\Pi_\tau) - r(\Pi_0)\,]$, for all $\tau \in \mathcal{T}_\mathrm{m}$ and all mappings $r \in \mathcal{BC}^{1+}(\mathcal{I}_\infty)$ in anticipation of applying the Verification Theorem. To simplify the notation, for each $t \geq 0$ define the $\mathcal{O}_t$–measurable linear stochastic differential operator $D_{\mathcal{O}_t}$ for all functions $r \in \mathcal{C}^{1+}(\mathrm{I}_\infty)$ via,

$$\begin{aligned} D_{\mathcal{O}_t} r(\pi) \;:=\; & \Big[-(1-\pi)\,\pi\, r'(\pi) \\ & \qquad + (\Psi_t + \pi)\Big[r\left(\frac{\Psi_t + 1}{\Psi_t + \pi}\,\pi\right) - r(\pi)\Big]\Big](\hat{J}_t - 1)\,dt \\ & \qquad + (1-\pi)\, r'(\pi)\,(1 - F_t)^{-1}\, dF_t \qquad (4.67) \end{aligned}$$

where,

$$\Psi_t := (\hat{J}_{t-} - 1)^{-1} \quad \forall\, t \geq 0. \qquad (4.68)$$

Observe that,

$$0 \leq \frac{\Psi_t + 1}{\Psi_t + \pi}\,\pi \leq 1 \qquad \forall\, \pi \in [0,1],\; t \geq 0, \qquad (4.69)$$

and therefore this ratio is always in the domain of $r \in \mathcal{C}^{1+}(\mathrm{I}_\infty)$.

For $\mathrm{I} \in \mathcal{I}_\infty$, $r \in \mathcal{PC}^{+}(\mathrm{I})$ and $t \geq 0$ we define $J^r(\Pi)$ to be the sum of the jumps in the $r(\Pi)$ process which are due solely to the discontinuities of r, i.e.,

$$J_t^r(\Pi) := \sum_{s \leq t} [r(\Pi_s^-) - r(\Pi_s)] \qquad \forall\, t \geq 0. \qquad (4.70)$$

The following proposition is the analog of Proposition 3.7 but with the partially–observed drift structure of Π in the diffusion case of Chapter 3 replaced here by the partially–observed intensity structure of Π in the point process case.

Proposition 4.7 *Let $r \in \mathcal{BC}^{1+}(\mathcal{I}_\infty)$ and suppose that $D_{\mathcal{O}_t} r(\pi) < 0$, P_π–a.s. for all $\pi \in \mathrm{I}_\infty$ and $t \geq 0$. Then,*

$$E_\pi[r(\Pi_\tau) - r(\Pi_0)] = E_\pi \int_0^\tau D_{\mathcal{O}_s} r(\Pi_s) + J_\tau^r(\Pi) \qquad \forall\, \tau \in \mathcal{T}_{\mathrm{m}},\; \pi \in \mathrm{I}_\infty.$$

Proof: Fix $\pi \in \mathrm{I}_\infty$. For $r \in \mathcal{BC}^{1+}(\mathcal{I}_\infty)$ the generalized Itô formula yields (see Appendix A),

$$r(\Pi_t) - r(\Pi_0) = \int_0^t r'(\Pi_s)\, d\Pi_s^c + \sum_{s\le t}[\, r(\Pi_s^-) - r(\Pi_{s-})\,], \tag{4.71}$$

which we can rewrite as,

$$r(\Pi_t) - r(\Pi_0) = \int_0^t r'(\Pi_s)\, d\Pi_s + \sum_{s\le t}[\, r(\Pi_s^-) - r(\Pi_{s-}) - r'(\Pi_{s-})\Delta\Pi_s\,], \tag{4.72}$$

for all $t \ge 0$, P_π–a.s. Using Proposition 4.6 we compute the stochastic differential,

$$\begin{aligned} d\Pi_t &= (1-\Pi_t)\,(1-F_t)^{-1}\, dF_t + \frac{\Pi_{t-}\,(1-\Pi_{t-})\,(\hat{J}_{t-}-1)}{1+\Pi_{t-}(\hat{J}_{t-}-1)}\, d\overline{\eta}_t^{\upsilon} \\ &= (1-\Pi_t)\,(1-F_t)^{-1}\, dF_t \\ &\qquad + \Pi_{t-}\,(1-\Pi_{t-})\,(\Psi_t+\Pi_{t-})^{-1}\, d\overline{\eta}_t^{\upsilon} \end{aligned} \tag{4.73}$$

and the stochastic difference,

$$\begin{aligned} \Delta\Pi_t &= \Pi_{t-}\,(1-\Pi_{t-})\,(\Psi_t+\Pi_{t-})^{-1}\, \Delta\overline{\eta}_t^{\upsilon} \\ &= \Pi_{t-}\,(1-\Pi_{t-})\,(\Psi_t+\Pi_{t-})^{-1}\, \Delta N_t. \end{aligned} \tag{4.74}$$

We can now employ 4.74 to compute the last term of the sum in 4.72,

$$\begin{aligned} \sum_{s\le t} r'(\Pi_{s-})\,\Delta\Pi_s &= \sum_{s\le t} r'(\Pi_{s-})\,\Pi_{s-}\,(1-\Pi_{s-})\,(\Psi_s+\Pi_{s-})^{-1}\,\Delta N_s \\ &= \sum_{\substack{s\le t \\ \Delta N_s = 1}} \frac{\Pi_{s-}\,(1-\Pi_{s-})}{\Psi_s+\Pi_{s-}}\, r'(\Pi_{s-}) \\ &= \int_0^t \frac{\Pi_{s-}\,(1-\Pi_{s-})}{\Psi_s+\Pi_{s-}}\, r'(\Pi_{s-})\, dN_s, \end{aligned} \tag{4.75}$$

and using Proposition 4.4 we can rewrite 4.75 as,

$$\sum_{s\le t} r'(\Pi_{s-})\,\Delta\Pi_s = \int_0^t \frac{\Pi_{s-}\,(1-\Pi_{s-})}{\Psi_s+\Pi_{s-}}\, r'(\Pi_{s-})\, d\overline{\eta}_s^{\upsilon}$$

$$
\begin{aligned}
& - \int_0^t \frac{\Pi_s\,(1-\Pi_s)}{\Psi_s + \Pi_s}\, r'(\Pi_s)\,[1 + \Pi_s(\hat{J}_s - 1)]\, ds \\
= & \int_0^t r'(\Pi_{s-})\,\Pi_{s-}\,(1-\Pi_{s-})\,(\Psi_s + \Pi_{s-})^{-1}\, d\overline{\eta}_s^{\upsilon} \\
& - \int_0^t r'(\Pi_s)\,\Pi_s\,(1-\Pi_s)\,\Psi_s^{-1}\, ds. \qquad (4.76)
\end{aligned}
$$

Substituting 4.73 and 4.76 into 4.72 and then simplifying the result we obtain,

$$
\begin{aligned}
r(\Pi_t) - r(\Pi_0) = & \int_0^t (1-\Pi_s)\,(1-F_s)^{-1}\, r'(\Pi_s)\, dF_s \\
& - \int_0^t \Pi_s\,(1-\Pi_s)\, r'(\Pi_s)\,\Psi_s^{-1}\, ds \\
& + \sum_{s \leq t} [\, r(\Pi_s^-) - r(\Pi_{s-})\,]. \qquad (4.77)
\end{aligned}
$$

It only remains to compute the sum in 4.77. The total stochastic difference of $r(\Pi)$ can be decomposed as,

$$
r(\Pi_t^-) - r(\Pi_{t-}) = r(\Pi_t) - r(\Pi_{t-}) + r(\Pi_t^-) - r(\Pi_t) \qquad P_\pi\text{-a.s.} \qquad (4.78)
$$

From 4.74 we find,

$$
\begin{aligned}
\Pi_t & = \Pi_{t-} + \Delta\Pi_t \\
& = [\Psi_t + \Pi_{t-} + (1-\Pi_{t-})\,\Delta N_t]\,(\Psi_t + \Pi_{t-})^{-1}\,\Pi_{t-} \\
& = \left(\frac{\Psi_t + \Delta N_t + \Pi_{t-}(1-\Delta N_t)}{\Psi_t + \Pi_{t-}} \right) \Pi_{t-} \qquad (4.79)
\end{aligned}
$$

and therefore using 4.79 we can compute,

$$
\begin{aligned}
\sum_{s \leq t} [r(\Pi_s) - r(\Pi_{s-})] & = \sum_{s \leq t} [r(\Pi_s) - r(\Pi_{s-})]\,\Delta N_s \\
& = \sum_{\substack{s \leq t \\ \Delta N_s = 1}} [r(\Pi_{s-} + \Delta\Pi_s) - r(\Pi_{s-})] \\
& = \sum_{\substack{s \leq t \\ \Delta N_s = 1}} [r\left(\frac{\Psi_s + 1}{\Psi_s + \Pi_{s-}}\,\Pi_{s-} \right) - r(\Pi_{s-})]
\end{aligned}
$$

$$= \int_0^t [r\left(\frac{\Psi_s+1}{\Psi_s+\Pi_{s-}}\,\Pi_{s-}\right) - r(\Pi_{s-})]\,dN_s. \quad (4.80)$$

Plugging 4.80 into 4.77 gives,

$$\begin{aligned} r(\Pi_t) - r(\Pi_0) &= \int_0^t (1-\Pi_s)\,(1-F_s)^{-1}\,r'(\Pi_s)\,dF_s \\ &\quad - \int_0^t \Pi_s\,(1-\Pi_s)\,r'(\Pi_s)\,\Psi_s^{-1}\,ds \\ &\quad + \int_0^t [r\left(\frac{\Psi_s+1}{\Psi_s+\Pi_{s-}}\,\Pi_{s-}\right) - r(\Pi_{s-})]\,dN_s \\ &\quad + \sum_{s\le t}[\,r(\Pi_s^-) - r(\Pi_s)\,], \end{aligned} \quad (4.81)$$

for all $t \geq 0$, P_π–a.s. Making use of the definition of $D_{\mathcal{O}_t}$ given in 4.67 and the definition of $J^r(\Pi)$ we can use Proposition 4.4 once again to rewrite 4.81 more compactly as,

$$r(\Pi_t) - r(\Pi_0) = \int_0^t D_{\mathcal{O}_s}\,r(\Pi_s) + J_t^r(\Pi) + \tilde{M}_t \quad \forall t \geq 0,\ P_\pi\text{–a.s.}, \quad (4.82)$$

where $\tilde{M} = \{\tilde{M}_t\}_{t\geq 0}$ is defined here as,

$$\tilde{M}_t := \int_0^t [r\left(\frac{\Psi_s+1}{\Psi_s+\Pi_{s-}}\,\Pi_{s-}\right) - r(\Pi_{s-})]\,d\tilde{\eta}_s^v \quad \forall t \geq 0. \quad (4.83)$$

Next we will show that $\{\tilde{M}_{t\wedge\tau}\}_{t\geq 0}$ defined by,

$$\tilde{M}_{t\wedge\tau} := \int_0^t 1\{s \leq \tau\}[r\left(\frac{\Psi_s+1}{\Psi_s+\Pi_{s-}}\,\Pi_{s-}\right) - r(\Pi_{s-})]\,d\tilde{\eta}_s^v \quad \forall t \geq 0, \quad (4.84)$$

is an $(\mathcal{O}_t, P_\pi)$–local martingale. To do this it is only necessary to show that the integrand in 4.84 is $(\mathcal{O}_t, P_\pi)$–predictable and locally bounded. The predictability is clear. Since $\tau \leq \tau^{\mathrm{I}}$ we see that $\Pi_{\tau^{\mathrm{I}}} \in [\,\mathrm{I}\,]_\Pi$ and therefore,

$$r(\Pi_{\tau-}) \leq B_{\mathrm{I}} < \infty \quad P_\pi\text{–a.s.},\ \forall \pi \in \mathrm{I}_\infty, \quad (4.85)$$

for some B_{I} known to exist since $r \in \mathcal{BC}^{1+}(\mathcal{I}_\infty)$. Similarly,

$$r\left(\frac{\Psi_\tau + 1}{\Psi_\tau + \Pi_{\tau-}} \Pi_{\tau-}\right) \leq B_{\mathrm{I}} \quad P_\pi\text{–a.s.}, \; \forall \pi \in \mathrm{I}_\infty. \tag{4.86}$$

This last bound follows from 4.79 if we keep in mind that even when $\Delta N_\tau = 1$ so that Π jumps at τ we have,

$$\Pi_\tau = \left(\frac{\Psi_\tau + 1}{\Psi_\tau + \Pi_{\tau-}}\right) \Pi_{\tau-}, \tag{4.87}$$

and therefore,

$$\left(\frac{\Psi_\tau + 1}{\Psi_\tau + \Pi_{\tau-}}\right) \Pi_{\tau-} \in [\,\mathrm{I}\,]_\Pi \qquad P_\pi\text{–a.s.} \tag{4.88}$$

Thus, the integrand in expression 4.84 is in fact P_π–a.s. bounded. As a result $\tilde{M}_{t\wedge\tau}$ is an $(\mathcal{O}_t, P_\pi)$–local martingale with a localization sequence, say $\{\mu_n\}_{n\geq 1}$ so that,

$$E_\pi \tilde{M}_{\mu_n\wedge\tau} = 0 \qquad \forall n \geq 1. \tag{4.89}$$

In view of expression 4.82 this gives for all $n \geq 1$,

$$E_\pi[\, r(\Pi_{\mu_n\wedge\tau}) - r(\Pi_0)\,] = E_\pi \int_0^{\mu_n\wedge\tau} D_{\mathcal{O}_s} r(\Pi_s) + E_\pi\, J^r_{\mu_n\wedge\tau}(\Pi), \tag{4.90}$$

which if we take the definition of $J^r(\Pi)$ into account can be rewritten,

$$E_\pi[\, r(\Pi_{\mu_n\wedge\tau-}) - r(\Pi_0)\,] = E_\pi \int_0^{\mu_n\wedge\tau} D_{\mathcal{O}_s} r(\Pi_s) + E_\pi\, J^r_{\mu_n\wedge\tau-}(\Pi), \tag{4.91}$$

Making use of the Monotone Convergence Theorem we obtain,

$$\lim_{n\to\infty} E_\pi \int_0^{\mu_n\wedge\tau} D_{\mathcal{O}_s} r(\Pi_s) = E_\pi \int_0^\tau D_{\mathcal{O}_s} r(\Pi_s), \tag{4.92}$$

and also,

$$\lim_{n\to\infty} E_\pi\, J^r_{\mu_n\wedge\tau-}(\Pi) = E_\pi\, J^r_{\tau-}(\Pi). \tag{4.93}$$

As a result,

$$\lim_{n\to\infty} E_\pi[\, r(\Pi_{\mu_n\wedge\tau-}) - r(\Pi_0)\,] = E_\pi \int_0^\tau D_{\mathcal{O}_s}\, r(\Pi_s) + E_\pi\, J^r_{\tau-}(\Pi). \qquad (4.94)$$

Employing the Bounded Convergence Theorem and an argument paralleling the one used in Lemma 2.1 we find,

$$\lim_{n\to\infty} E_\pi[\, r(\Pi_{\mu_n\wedge\tau-})\,] = E_\pi[\, r(\Pi_{\tau-})\,], \qquad (4.95)$$

so that 4.94 yields,

$$E_\pi[\, r(\Pi_{\tau-}) - r(\Pi_0)\,] = E_\pi \int_0^\tau D_{\mathcal{O}_s}\, r(\Pi_s) + E_\pi\, J^r_{\tau-}(\Pi). \qquad (4.96)$$

Finally, from this last expression we obtain,

$$E_\pi[\, r(\Pi_\tau) - r(\Pi_0)\,] = E_\pi \int_0^\tau D_{\mathcal{O}_s}\, r(\Pi_s) + E_\pi\, J^r_\tau(\Pi), \qquad (4.97)$$

as promised. □

4.6 Sequential Detection

In this section we will formulate a classical Bayesian sequential binary hypothesis testing or *sequential detection* problem on the intensity of an observed point process process. We will show how it can be recast as a *change detection* problem within our framework and then tackled via the Verification Theorem. Recast in this way, the classical problem of sequential detection is given a fresh interpretation as a problem of *lack-of-change* detection. The Verification Theorem will lead us to consider a type of free-boundary value or *Stefan* problem whose solution is addressed using techniques from the theory of functional differential equations and a generalized convex analysis. We will end the section with an example involving a Poisson process.

4.6.1 Problem Statement

On a measurable space $(\Omega, \mathcal{G})$ equipped with two mutually absolutely continuous probability measures P_0 and P_1 we observe a counting process $N = \{N_t\}_{t\geq 0}$ for which one of the following hypotheses is true:

$$\begin{aligned} \text{(Noise Only)}: \quad & N_t = \int_0^t 1\, ds + \eta_t & \qquad 0 \leq t < \infty; \\ \text{(Signal Plus Noise)}: \quad & N_t = \int_0^t J_s\, ds + \eta_t & \qquad 0 \leq t < \infty, \end{aligned}$$

where η is a $(\mathcal{G}_t, P_i)$–martingale for $i = 0, 1$, and $J = \{J_t\}_{t\geq 0}$ is a $\mathcal{G}_t$–progressive process which is strictly greater than unity and satisfies,

$$E_i \int_0^t J_s\, ds < \infty \qquad \forall\, t \geq 0,\ i = 0, 1. \tag{4.98}$$

Following the Bayesian philosophy, we are told in advance that the (Signal Plus Noise) hypothesis occurs with prior probability $\pi \in [0, 1]$. We are tasked with deciding in a sequential manner which hypothesis is indeed responsible for what is observed. Decision policies are defined as a pair, (τ, δ), where τ is a stopping time with respect to the P_0–completed observation filtration $\mathcal{O}_t = \bigvee_{s \leq t} \sigma(N_s)$ and δ is a binary random variable representing our decision and therefore measurable with respect to $\mathcal{O}_\tau$. We are told that the goodness of any sequential decision policy, (τ, δ), is judged according to the following 'elapsed-energy + incorrect decision' cost criterion,

$$\begin{aligned} \bar{\rho}_\pi(\tau, \delta) &= \pi\, E_1[\int_0^\tau c\,(\hat{J}_s - 1)\, ds + c^0\, 1\{\delta = 0\}\,] \\ &\quad + (1 - \pi)\, E_0[\int_0^\tau c\,(\hat{J}_s - 1)\, ds + c^1\, 1\{\delta = 1\}\,], \end{aligned} \tag{4.99}$$

where c, c^0, and c^1 are strictly positive and finite. We are asked to minimize $\bar{\rho}_\pi(\tau, \delta)$ over all decision pairs. We are told to restrict our attention to those $\mathcal{O}_t$–stopping times which satisfy,

$$E_i \int_0^\tau (\hat{J}_s - 1)^2\, ds < \infty, \quad i = 0, 1, \tag{4.100}$$

and in addition we are given the following technical condition on the running cost:

$$P_i\,\{\int_0^\infty (\hat{J}_s - 1)^2\, ds = \infty\,\} = 1, \quad i = 0, 1. \tag{4.101}$$

4.6.2 Problem Reformulation

We now show how the sequential detection problem can be recast as a problem of change detection on the probability space $(\Omega, \mathcal{A}, P_\pi)$ with π the same prior as in the previous subsection. Drawing freely upon our earlier results and notation, the first step is to notice that the (Noise Only) and (Signal

Plus Noise) hypotheses are captured by the model dynamics (DM0) and (DM1), respectively, just as in the diffusion case. As expected then, a little consideration shows that the *sequential* aspect of sequential detection can be recovered within the change detection format by properly specifying the random disruption time v previously characterized only up to the nature of $F = \{F_t\}_{t\geq 0}$, its $\mathcal{O}_t$–adapted, conditional cumulative distribution function with respect to the P_0–measure. To this end make the following sequential detection modeling assumption:

$$(\mathrm{F_{SD}}): \qquad F_t := 1\{t = \infty\} \quad \forall\, t \in [0, \infty].$$

We see that F is a legitimate $\mathcal{O}_t$–adapted, conditional cumulative distribution function on $[0, \infty] = \Re_+ \cup \{\infty\}$ satisfying the requirements set down in Chapter 2. Under assumption ($\mathrm{F_{SD}}$) we have therefore,

$$P_0\{\, v \leq t \,|\, \mathcal{O}_t \,\} = P_0\{\, v \leq t \,\} = 0 \quad \forall\, t \in [0, \infty), \tag{4.102}$$

and $P_0\{\, v = \infty \,\} = 1$. With this choice of F, the P_0–measure gives all its probability to the event $\{\, v = \infty \,\}$; the P_1–measure of course still gives all its probability to the event $\{\, v = 0 \,\}$. An immediate consequence of this is that the P–measure defined in 2.97 is precisely the P_0–measure and as a result the P_π-measure simplifies to,

$$\begin{aligned} P_\pi\{A\} &= \pi\, P_1\{A\} + (1 - \pi)\, P\{A\} \\ &= \pi\, P_1\{A\} + (1 - \pi)\, P_0\{A\}, \end{aligned} \tag{4.103}$$

for all π in $[0, 1]$ and A in $\mathcal{A}$. From this we see,

$$P_\pi\{0 < v < \infty\} = \pi\, P_1\{0 < v < \infty\} + (1 - \pi)\, P_0\{0 < v < \infty\} = 0, \tag{4.104}$$

or, what is equivalent $P_\pi\{ v = 0 \} = \pi$ and $P_\pi\{ v = \infty \} = 1 - \pi$.

The ($\mathrm{F_{SD}}$) assumption also has the consequence of simplifying the semimartingale representation for Υ which is obviously now reduced to $\Upsilon_t = \Upsilon_0$, P_π–a.s., for all $t \geq 0$. Using this fact we have $E_\pi \mathcal{E}(\Upsilon_\tau, \delta) = E_\pi \mathcal{E}(\Upsilon_0, \delta)$ for all τ in $\mathcal{T}_{ad}$. From this we compute,

$$\begin{aligned} E_\pi \mathcal{E}(\Upsilon_\tau, \delta) &= E_\pi \mathcal{E}(\Upsilon_0, \delta) \\ &= \pi\, E_1 \mathcal{E}(\Upsilon_0, \delta) + (1-\pi)\, E_0 \mathcal{E}(\Upsilon_0, \delta) \\ &= \pi\, E_1 \mathcal{E}(1, \delta) + (1-\pi)\, E_0 \mathcal{E}(0, \delta) \\ &= \pi\, E_1[\, c^0\, 1\{\delta = 0\} \cdot 1 + 0\,] + (1-\pi)\, E_0[\, 0 + c^1\, 1\{\delta = 1\} \cdot 1\,] \\ &= \pi\, E_1[\, c^0\, 1\{\delta = 0\}\,] + (1-\pi)\, E_0[\, c^1\, 1\{\delta = 1\}\,]. \end{aligned} \tag{4.105}$$

Next, we choose the running cost $\mathcal{C}$ according to,

$$\mathcal{C}_t := c\,(\hat{J}_t - 1) \quad \forall\, t \geq 0, \tag{4.106}$$

and note that $\mathcal{C}$ satisfies the cost assumptions (C1), (C2) and (C3): (C1) and (C2) follow directly from (J1), (J2) and 4.103 while (C3) follows since $\mathcal{C}$ is trivially concave in Π. Now compute Bayes' cost,

$$\begin{aligned} \rho_\pi(\tau, \delta) &= E_\pi[\int_0^\tau \mathcal{C}_s\, ds + \mathcal{E}(\Upsilon_\tau, \delta)\,] \\ &= \pi\, E_1[\int_0^\tau \mathcal{C}_s\, ds + \mathcal{E}(\Upsilon_0, \delta)\,] \\ &\qquad + (1-\pi)\, E_0[\int_0^\tau \mathcal{C}_s\, ds + \mathcal{E}(\Upsilon_0, \delta)\,] \\ &= \pi\, E_1[\int_0^\tau c\,(\hat{J}_s - 1)\, ds + c^0\, 1\{\delta = 0\}\,] \\ &\qquad + (1-\pi)\, E_0[\int_0^\tau c\,(\hat{J}_s - 1)\, ds + c^1\, 1\{\delta = 1\}\,] \\ &= \bar{\rho}_\pi(\tau, \delta). \end{aligned} \tag{4.107}$$

Hence, this choice of running cost and assumption ($\mathrm{F_{SD}}$) reduce the change detection problem as defined in Chapter 2 to a classical Bayes sequential detection problem. Indeed, when viewed as a change detection problem, a sequential detection problem is seen to be the extremal case: the *lack-of-change* change detection problem. We can therefore solve the sequential detection problem by solving the optimal stopping problem ($\mathcal{P}$) of Chapter 2 under the conditions prevailing in the current section. The plan of course is to solve the optimal stopping problem using a first exit policy found by recourse to the Verification Theorem. We end this subsection with a convenient and complete reformulation of the sequential detection problem, taking into account the nature of the observations, the sequential detection assumption, and the choice of running cost.

On the probability space $(\Omega, \mathcal{A}, P_\pi)$ with $\pi \in [0,1]$ given, choose $F_t = 1\{v = \infty\}$ so that the jump time v obeys,

$$P_\pi\{v = 0\} = \pi = 1 - P_\pi\{v = \infty\}. \tag{4.108}$$

Since $\Upsilon_t = 1\{v \leq t\}$ we note that $\Upsilon_t = \Upsilon_0$ for all $t \geq 0$, P_π–a.s. Thus, there exist only two possible sample paths upon which P_π is concentrated. The observations are given by,

$$N_t = \int_0^t [\,(1 - \Upsilon_s) \cdot 1 + \Upsilon_s \cdot J_s\,]\, ds + \eta_t^v \qquad \forall\, t \geq 0, \tag{4.109}$$

where the intensity J satisfies 4.98, 4.100 and 4.101 in view of the technical assumptions (J0), (J1), and (J2) imposed in the last section. Due to the essentially two-valued nature of v we see that,

$$N_t = (1 - \Upsilon_0)\, t + \Upsilon_0 \int_0^t J_s\, ds + \eta_t^v \qquad \forall\, t \geq 0, \tag{4.110}$$

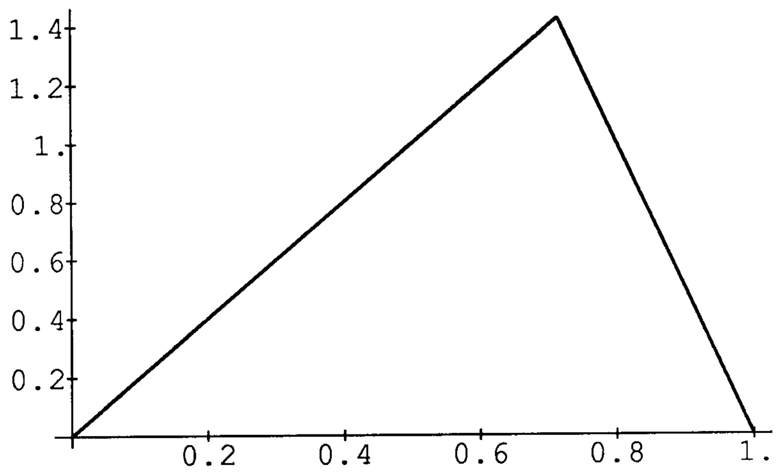

Figure 4.1: Terminal cost function. $c^0 = 2$, $c^1 = 5$, $\pi_e = \frac{5}{7}$.

and therefore,

$$N_t = \begin{cases} t + \eta_t^v & 0 \leq t < v = \infty; \qquad \text{(Noise Only)} \\ \\ \displaystyle\int_0^t J_s\, ds + \eta_t^v & 0 = v \leq t < \infty. \quad \text{(Signal Plus Noise)} \end{cases} \tag{4.111}$$

Hence, choosing F as in (F_{SD}) captures the classical sequential detection set-up within the change detection set-up. Bayes' cost is taken to be,

$$\rho_\pi(\tau) = E_\pi[\int_0^\tau c\,(\hat{J}_s - 1)\, ds + e(\Pi_\tau)\,], \tag{4.112}$$

with $c > 0$ and where the terminal cost, e, is given by,

$$e(\pi) = \min\{c^0\pi, c^1(1-\pi)\} \qquad \forall\, \pi \in [0,1], \tag{4.113}$$

with $0 < c^0, c^1 < \infty$ (see Figure 4.1). For this choice of running cost an admissible stopping time τ is any $\mathcal{O}_t$–stopping time satisfying,

$$E_\pi \int_0^\tau \mathcal{C}_s\, ds = E_\pi \int_0^\tau c\,(\hat{J}_s - 1)\, ds < \infty. \tag{4.114}$$

Thus, the change detection problem to solve which is equivalent to the sequential detection problem is,

$$(\mathcal{P}_{\rm SD}): \qquad \text{Find } \tau_* \in \mathcal{T}_{ad} \text{ such that } \rho_\pi(\tau_*) = \inf_{\tau \in \mathcal{T}_{ad}} \rho_\pi(\tau). \tag{4.115}$$

4.6.3 Escape Properties

The next question to answer is whether problem $(\mathcal{P}_{\rm SD})$ is sufficiently well-posed that the escape condition (E) is satisfied and then which of (E^+) or (E^0) holds. We note that $(\mathrm{F}_{\rm SD})$ implies that $dF_t = 0$ for all $t \geq 0$, P_π–a.s., so that $\Upsilon \equiv \Upsilon_0$ and the stochastic differential description of the sample paths of Π reduces to,

$$\Pi_t = \Pi_0 + \int_0^t \frac{(1 - \Pi_{s-})\,\Pi_{s-}\,(\hat{J}_{s-} - 1)}{1 + \Pi_{s-}\,(\hat{J}_{s-} - 1)}\, d\overline{\eta}^v_s. \tag{4.116}$$

Recalling that Ψ is defined to be the predictable version of $(\hat{J} - 1)^{-1}$ we can rewrite this as,

$$\Pi_t = \Pi_0 + \int_0^t \frac{(1 - \Pi_{s-})\,\Pi_{s-}}{\Psi_s + \Pi_{s-}}\, d\overline{\eta}^v_s. \tag{4.117}$$

In Proposition 4.8 below we show that condition (J2) implies that (E^+) holds so that $\mathcal{I}_\infty = \mathcal{I}^+$ and $\mathrm{I}_\infty = (0,1)$ and thus the trajectories of Π are guaranteed to escape any proper subinterval of $(0,1)$.

Proposition 4.8 *Under assumption* (J2) *the escape condition* (E^+) *holds, i.e., given*

$$P_i\{ \int_0^\infty (\hat{J}_s - 1)^2\, ds = \infty \} = 1 \qquad i = 0, 1,$$

then,

$$P_\pi\{\tau^{\rm I} < \infty\} = 1 \qquad \forall\, \pi \in \mathrm{I},\ \forall\, \mathrm{I} \in \mathcal{I}^+.$$

Proof: Choose any I in $\mathcal{I}^+$ so that $\mathrm{I} = (a, b)$ with $0 < a < b < 1$ and suppose $\pi \in \mathrm{I}$. Define the mapping R according to,

$$R(\pi) := \frac{\pi}{1-\pi} \quad \forall\, \pi \in [0,1]. \tag{4.118}$$

From Proposition 4.7 and 4.116 we obtain,

$$E_\pi[\, R(\Pi_{\tau^{\mathrm{I}}}) - R(\Pi_0)\,] = E_\pi \int_0^{\tau^{\mathrm{I}}} \frac{\Pi_s^2\,(\hat{J}_s - 1)^2}{(1-\Pi_s)}\, ds. \tag{4.119}$$

Obviously,

$$\frac{b}{1-b} \geq E_\pi[\, R(\Pi_{\tau^{\mathrm{I}}}) - R(\Pi_0)\,], \tag{4.120}$$

and likewise,

$$E_\pi \int_0^{\tau^{\mathrm{I}}} \frac{\Pi_s^2\,(\hat{J}_s - 1)^2}{(1-\Pi_s)}\, ds \geq \frac{a^2}{1-a} E_\pi \int_0^{\tau^{\mathrm{I}}} (\hat{J}_s - 1)^2\, ds. \tag{4.121}$$

Moreover it is clear that,

$$E_\pi \int_0^{\tau^{\mathrm{I}}} (\hat{J}_s - 1)^2\, ds \geq E_\pi[\, 1\{\tau^{\mathrm{I}} = \infty\} \int_0^{\tau^{\mathrm{I}}} (\hat{J}_s - 1)^2\, ds\,]. \tag{4.122}$$

Combining these results yields,

$$E_\pi[\, 1\{\tau^{\mathrm{I}} = \infty\} \int_0^{\infty} (\hat{J}_s - 1)^2\, ds\,] \leq \frac{(1-a)\, b}{(1-b)\, a^2} < \infty. \tag{4.123}$$

This last line yields a contradiction unless $P_\pi\{\tau^{\mathrm{I}} = \infty\} = 0$ because it is obvious from (J2) that $P_\pi\{\, \int_0^\infty (\hat{J}_s - 1)^2\, ds = \infty\,\} = 1$ for all π in I. □

Motivated by this proposition we are justified in searching for $\mathrm{I}_* \in \mathcal{I}^+$ so that if we write $\mathrm{I}_* = (a_*, b_*)$ then $0 < a_* < b_* < 1$ and moreover, in view of Lemma 2.4, we expect that $a_* \leq \pi_e \leq b_*$. Before embarking on our search for I_* however, we consider one more result which is companion to the last. Indeed, while the last proposition shows (J2) implies that the Π process is guaranteed to escape any interval in $\mathcal{I}_\infty$ in finite time, in the following proposition we will show (J1) guarantees that Π can only escape $I_\infty = (0,1)$ in infinite time.

Proposition 4.9 *Under assumption* (J1),

$$P_\pi\{\tau^{I_\infty} = \infty\} = 1 \qquad \forall\, \pi \in I_\infty.$$

Proof: Fix $\pi \in I_\infty = (0,1)$. From Proposition 4.5 we have

$$P_\pi\{\sup_{0\le t\le n} |\log \check{L}_t| = \infty\} = 0 \quad \forall\, n \ge 1,\ \forall\, \pi \in I_\infty. \tag{4.124}$$

This fact is sufficient to imply that $P_\pi\{\tau^{I_\infty} < \infty\} = 0$. Employing the Itô stochastic integration formula one can show for all $\pi \in (0,1)$ that (see [MS, App I]),

$$\check{L} \equiv \frac{1-\pi}{\pi}\,\frac{\Pi}{1-\Pi}. \tag{4.125}$$

From this it is obvious that,

$$\{|\Pi_{\tau^{I_\infty}} - \frac{1}{2}| = \frac{1}{2}\} = \{|\log \check{L}_{\tau^{I_\infty}}| = \infty\}, \tag{4.126}$$

and therefore,

$$\{\tau^{I_\infty} < \infty\} \subset \{|\log \check{L}_{\tau^{I_\infty}}| = \infty\}. \tag{4.127}$$

Hence,

$$P_\pi\{\tau^{I_\infty} < \infty\} \;=\; P_\pi\{|\log \check{L}_{\tau^{I_\infty}}| = \infty, \tau^{I_\infty} < \infty\}$$

$$
\begin{aligned}
&= P_\pi\{|\log \check{L}_{\tau^{\mathrm{I}_\infty}}| = \infty, \cup_{n\geq 1}\{\tau^{\mathrm{I}_\infty} \leq n\}\} \\
&\leq \sum_{n=1}^{\infty} P_\pi\{|\log \check{L}_{\tau^{\mathrm{I}_\infty}}| = \infty, \tau^{\mathrm{I}_\infty} \leq n\} \\
&\leq \sum_{n=1}^{\infty} P_\pi\{\sup_{0\leq t\leq n} |\log \check{L}_t| = \infty, \tau^{\mathrm{I}_\infty} \leq n\} \\
&\leq \sum_{n=1}^{\infty} P_\pi\{\sup_{0\leq t\leq n} |\log \check{L}_t| = \infty\} \\
&= \lim_{m\to\infty} \sum_{n=1}^{m} 0 = 0,
\end{aligned} \tag{4.128}
$$

i.e., Π has no chance of escaping from $\mathrm{I}_\infty = (0,1)$ in finite time. □

4.6.4 Verification: A Stefan Problem

For easy reference, we state the verification conditions (V1)–(V4) for problem $(\mathcal{P}_{\mathrm{SD}})$ given in the previous subsection with $\mathrm{I}_\infty = (0,1)$, $\mathcal{I}_\infty = \mathcal{I}^+$, and $\mathrm{I}_* \in \mathcal{I}^+$. We also incorporate the proposed running cost given in 4.106:

(V1a) : *For all* $\tau \in \mathcal{T}_{\mathrm{m}}$,

$$E_\pi[r_*(\Pi_\tau) - r_*(\Pi_0)] \geq -E_\pi \int_0^\tau c(\hat{J}_s - 1)\, ds \quad \forall \pi \in [\mathrm{I}_*]_\Pi;$$

(V1b) : $$E_\pi[r_*(\Pi_{\tau^{\mathrm{I}_*}}) - r_*(\Pi_0)] = -E_\pi \int_0^{\tau^{\mathrm{I}_*}} c(\hat{J}_s - 1)\, ds \quad \forall \pi \in [\mathrm{I}_*]_\Pi;$$

(V2) : $$r_*(\pi) = e(\pi) \quad \forall \pi \in \partial_\Pi \mathrm{I}_*;$$

(V3) : $$r_*(\pi) < e(\pi) \quad \forall \pi \notin \partial_\Pi \mathrm{I}_*;$$

and,

(V4) : r_* is bounded and continuous on $[\mathrm{I}_*]_\Pi$.

Let us attempt to find $r \in \mathcal{BC}^{1+}(0,1)$, an attempt facilitated by the availability of the generalized Itô stochastic integration formula given in Appendix A. Taking the detection assumption $(\mathrm{F}_{\mathrm{SD}})$ into account, the $\mathcal{O}_t$–measurable linear stochastic differential operator $D_{\mathcal{O}_t}$ defined in 4.67 for all mappings $r \in \mathcal{BC}^{1+}(0,1)$ simplifies to,

$$D_{\mathcal{O}_t}\, r(\pi) \;:=\; \Big[-(1-\pi)\,\pi\, r'(\pi) + (\Psi_t+\pi)[\, r\left(\frac{\Psi_t+1}{\Psi_t+\pi}\,\pi\right) - r(\pi)\,]\Big]\Psi_t^{-1}\, dt. \quad (4.129)$$

With this in mind, for each $t \geq 0$ define another stochastic linear differential operator D_{Ψ_t} via

$$D_{\Psi_t}\, r(\pi) := -(1-\pi)\,\pi\, r'(\pi) + (\Psi_t+\pi)[\, r\left(\frac{\Psi_t+1}{\Psi_t+\pi}\,\pi\right) - r(\pi)\,], \quad (4.130)$$

so that for all τ in $\mathcal{T}_{\mathrm{m}}$ Proposition 4.7 states,

$$E_\pi[\, r(\Pi_\tau) - r(\Pi_0)\,] = E_\pi \int_0^\tau \Psi_s^{-1}\, D_{\Psi_s}\, r(\Pi_s)\, ds + J_\tau^r(\Pi). \quad (4.131)$$

Hence, we can rewrite (V1a) as,

$$E_\pi \int_0^\tau (\hat{J}_s - 1)\, D_{\Psi_s}\, r(\Pi_s)\, ds + J_\tau^r(\Pi) \geq E_\pi \int_0^\tau \Big(-c\,(\hat{J}_s - 1)\Big)\; ds, \quad (4.132)$$

and (V1b) as,

$$E_\pi \int_0^{\tau^{\mathrm{I}_*}} (\hat{J}_s - 1)\, D_{\Psi_s}\, r(\Pi_s)\, ds + J_{\tau^{\mathrm{I}_*}}^r(\Pi) = E_\pi \int_0^{\tau^{\mathrm{I}_*}} \Big(-c\,(\hat{J}_s - 1)\Big)\; ds. \quad (4.133)$$

We note that

$$J^r(\Pi) \geq 0 \qquad \forall\, r \in \mathcal{BC}^{1+}(0,1), \quad (4.134)$$

and if $r_* \in \mathcal{BC}^{1+}(0,1)$ exists satisfying (V4) then,

$$J_{\tau^{\mathrm{I}_*}}^{r_*}(\Pi) = 0. \quad (4.135)$$

Therefore, to arrange for (V1) it is sufficient that r_* satisfy,

$$D_{\Psi_t}\, r_*(\pi) = -c \qquad \forall\, t \geq 0,\;\; P_\pi\text{–a.s.}, \forall\, \pi \in (0,1), \quad (4.136)$$

provided (V4) is also satisfied.

Here we see that we have arrived at a major difference between the point process case and the diffusion case. In the last expression the left-hand side is stochastic while the right-hand side is deterministic—simply factoring out the drift term does not lead to a deterministic Stefan problem as it did in the diffusion case. The only easy way around this state of affairs is less than satisfactory, but not exactly trivial either, and that is to assume that J and therefore $\hat{J}$ is a constant rate. With this choice the conditions on J, namely (J0)–(J2), reduce to $J > 1$. We therefore make the following assumption,

$$J = u^{-1} + 1, \tag{4.137}$$

for some $u > 0$ so that $J > 1$ and $\Psi_t = u$ for all $t \geq 0$. With this simplification the stochastic differential operator D_{Ψ_t} becomes deterministic. As in the diffusion case we will let D_Π denote the deterministic differential operator, namely,

$$D_\Pi\, r(\pi) := -(1-\pi)\,\pi\, r'(\pi) + (u+\pi)[\, r\left(\frac{u+1}{u+\pi}\,\pi\right) - r(\pi)\,]. \tag{4.138}$$

As for conditions (V2)–(V4) we need to compute $\partial_\Pi \mathrm{I}_*$. We know that since I_* satisfies the escape condition (E) with respect to Π then $\partial_\Pi \mathrm{I}_* \neq \emptyset$. In this case then, because the sample paths of the *martingale* Π are P_π–a.s. only right-continuous we see that,

$$\partial_\Pi^- \mathrm{I}_* = \{a_*\}, \tag{4.139}$$

while due to the possibility of jumping across b_*,

$$\partial_\Pi^+ \mathrm{I}_* = [b_*, \frac{u+1}{u+b_*}\, b_*). \tag{4.140}$$

Hence,

$$\partial_\Pi \mathrm{I}_* = \{a_*, [b_*, \frac{u+1}{u+b_*}\, b_*)\}, \tag{4.141}$$

and,

$$[\mathrm{I}_*]_\Pi = [a_*, \frac{u+1}{u+b_*}\, b_*). \tag{4.142}$$

Let's restate the verification conditions with these changes. Thus, we seek $r_* \in \mathcal{BC}^{1+}(0,1)$ and $\mathrm{I}_* \in \mathcal{I}^+$,

$$\begin{array}{lcccl}
(\mathrm{S1}): & D_\Pi\, r_*(\pi) & = & -c & \forall\, \pi \in (0,1); \\
(\mathrm{S2}): & r_*(\pi) & = & e(\pi) & \forall\, \pi \in \{a_*, [b_*, \frac{u+1}{u+b_*}\, b_*)\}; \\
(\mathrm{S3}): & r_*(\pi) & < & e(\pi) & \forall\, \pi \notin \{a_*, [b_*, \frac{u+1}{u+b_*}\, b_*)\};
\end{array}$$

(S4): r_* is bounded and continuous on $[a_*, \frac{u+1}{u+b_*}\, b_*)$.

Conditions (S1) and (S2) give us another free-boundary-value problem, a functional–differential variant of the ordinary differential *Stefan* problems (2nd order) considered in the last chapter. The functional constraints (S3) and (S4) serve to uniquely determine a solution amongst all the solutions to (S1)–(S2). In the next subsection we will show that (S1)–(S2) has a unique solution pair (r_*, I_*) which satisfies (S3)–(S4) and therefore an appeal to the Verification Theorem will imply that $\tau^{\mathrm{I}_*} \in \mathcal{T}$ solves problem $(\mathcal{P}_{\mathrm{SD}})$. The proof of the existence and uniqueness is carried out with convexity-like notions having the same flavor as those employed in the diffusion case.

4.6.5 Analysis of the Stefan Problem

We now proceed to solve the problem posed by (S1)-(S4). From 3.152 we know that $D_\Pi r$ for $r \in \mathcal{C}^{1+}(0,1)$ is given by,

$$D_\Pi r(\pi) = -\pi(1-\pi)r'(\pi) + (u+\pi)\left[r\left(\frac{u+1}{u+\pi}\pi\right) - r(\pi)\right] \quad \forall \pi \in (0,1), \tag{4.143}$$

where the derivative r' is of course computed from the right. For notational convenience define the *advance* operator $\mathcal{A}_u$ via,

$$\mathcal{A}_u \pi := \frac{u+1}{u+\pi}\pi \qquad 0 < \pi < 1. \tag{4.144}$$

Fix any a, $b \in (0,1)$ satisfying $0 < a \le \pi_e \le b < 1$ and $a < b$, define $\mathrm{I} = (a,b)$, and consider the following functional–differential equation (FDE) for $r \in \mathcal{BC}^{1+}(0,1)$,

$$\begin{aligned} -\pi(1-\pi)r'(\pi) + (u+\pi)\left[r(\mathcal{A}_u\pi) - r(\pi)\right] &= -c \quad \forall \pi \in (0,1); \\ r(\pi) &= e(\pi) \quad \forall \pi \in [b, \mathcal{A}_u b); \\ r(a) &= e(a). \end{aligned} \tag{4.145}$$

We emphasize that the choice of a and b satisfying $0 < a \le \pi_e \le b < 1$ is motivated by Lemma 2.4. The connection between the above FDE and (S1)–(S2) should be clear. Under the assumption that c, c^0, and c^1 are positive and finite, it follows from elementary FDE theory (see [MS]) that a unique, nontrivial solution to (S1)–(S2) exists; call it r_{I}, with $r_{\mathrm{I}} \in \mathcal{BC}^{1+}(0,1)$ (see Figure 4.2). Define the auxiliary function, $s_{\mathrm{I}} := r_{\mathrm{I}} - e$ (see Figure 4.3). To maintain the analogy with the diffusion problems we define the *generalized hyperplane* ℓ_* "supported" by s_{I} to be that function in $\ker \mathcal{D}_\Pi$,

$$\ker \mathcal{D}_\Pi := \{\, r \in \mathcal{BC}^{1+}(0,1) : \mathcal{D}_\Pi r \equiv 0 \,\}, \tag{4.146}$$

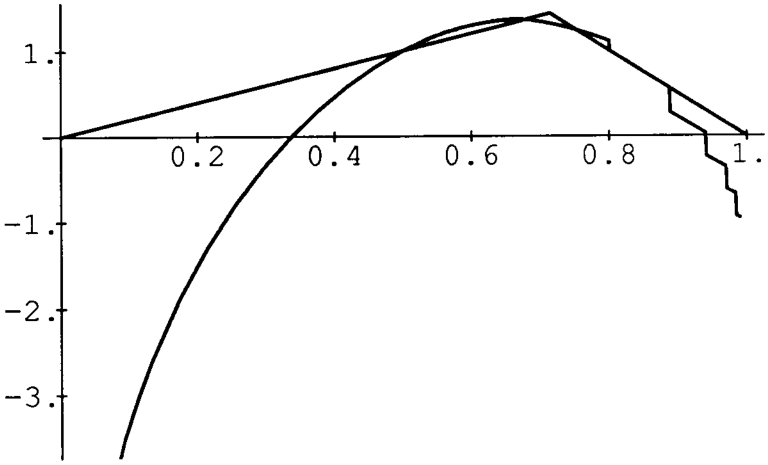

Figure 4.2: Graph of e and r_{I} with $\text{I} = (\frac{5}{10}, \frac{8}{10})$, $c = \frac{1}{2}$, $u = 1$.

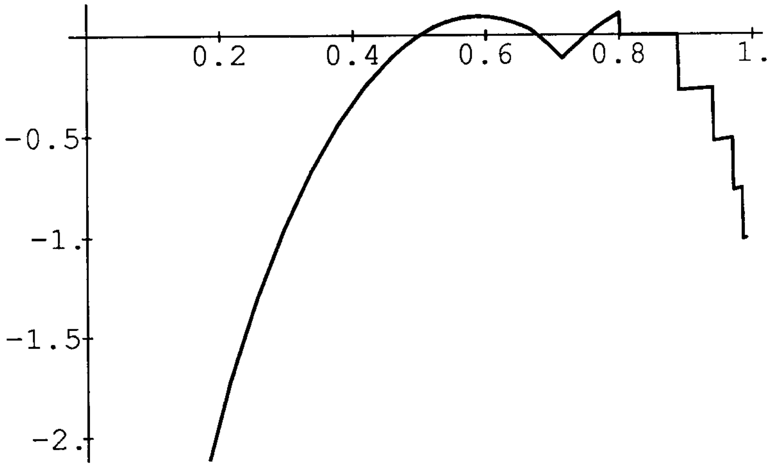

Figure 4.3: Graph of $s_{\text{I}} = r_{\text{I}} - e$.

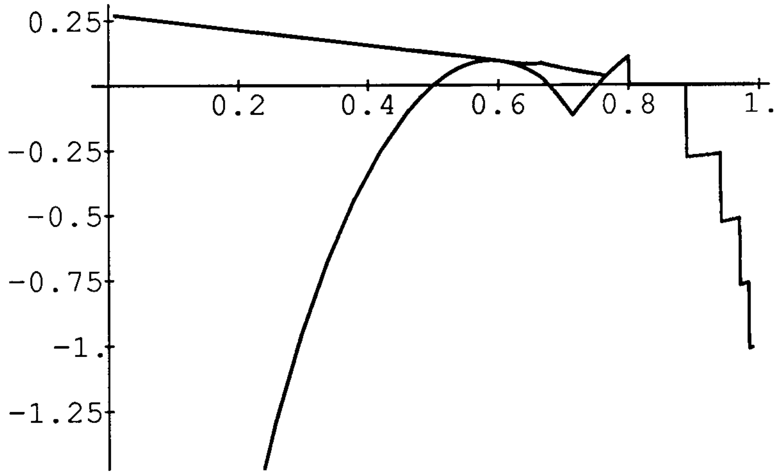

Figure 4.4: Graph of ℓ_* "supported" by s_{I}.

for which,

$$\ell_*(\pi) \geq s_{\text{I}}(\pi) \qquad \forall\, \pi \in [0,1], \tag{4.147}$$

and

$$\ell_*(\pi) = s_{\text{I}}(\pi) \qquad \forall\, \pi \in \partial_{\Pi}\, \text{I}. \tag{4.148}$$

The construction of this "hyperplane" is carried out in Appendix B and is depicted in Figure 4.4 and Figure 4.5. Now we can define the two numbers a_* and b_* via,

$$a_* := \sup\{\, \pi \leq \pi_e : \ell_*(\pi) = s_{\text{I}}(\pi)\,\}, \tag{4.149}$$

and

$$b_* := \inf\{\, \pi \geq \pi_e : \ell_*(\pi) = s_{\text{I}}(\pi)\,\}. \tag{4.150}$$

Since these sets are compact while s_{I} and ℓ_* are continuous on $(0, \mathcal{A}_u\, b)$, we see that $s_{\text{I}}(a_*) = \ell_*(a_*)$ and $s_{\text{I}}(b_*) = \ell_*(b_*)$ and by definition $0 < a_* < \pi_e <$

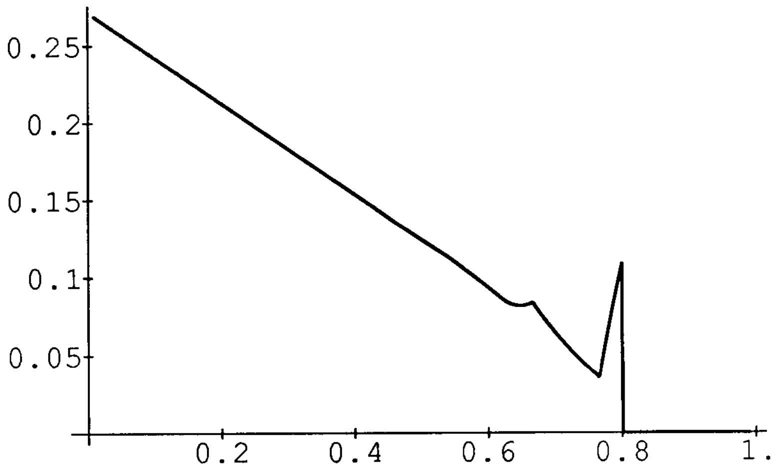

Figure 4.5: Graph of ℓ_*.

$b_* < 1$. Hence, if we define,

$$r_*(\pi) := r_{\mathrm{I}}(\pi) - \ell_*(\pi), \tag{4.151}$$

then obviously,

$$r_*(\pi) \leq e(\pi) \qquad \forall\, \pi \in [0,1], \tag{4.152}$$

and moreover,

$$\begin{aligned} r_*(a_*) &= r_{\mathrm{I}}(a_*) - \ell_*(a_*) \\ &= e(a_*) + (r_{\mathrm{I}}(a_*) - e(a_*)) - \ell_*(a_*) \\ &= e(a_*) + s_{\mathrm{I}}(a_*) - \ell_*(a_*) = e(a_*). \end{aligned} \tag{4.153}$$

Similarly, one can show,

$$r_*(b_*) = e(b_*), \tag{4.154}$$

and in fact,

$$r_*(\pi) = e(\pi) \qquad \forall\, \pi \in [b_*, \mathcal{A}_u\, b_*). \tag{4.155}$$

Hence, (r_*, I_*) satisfies (S2). In addition note that,

$$D_\Pi r_*(\pi) = D_\Pi r_{\mathrm{I}}(\pi) - D_\Pi \ell_*(\pi) = -c - 0 = -c, \qquad (4.156)$$

and therefore r_* satisfies (S1) also. Moreover, in view of the strict concavity of r_* on $(0, b_*)$, and its behavior on $[\mathcal{A}_u\, b_*, 1)$ (see Appendix C), it follows that the inequality in 4.152 is strict off of $\partial_\Pi\, \mathrm{I}_*$, i.e.,

$$r_*(\pi) < e(\pi) \qquad\qquad \forall\, \pi \notin \{a_*, [b_*, \mathcal{A}_u\, b_*)\}, \qquad (4.157)$$

and this gives (S3). Thus, r_* as defined in 4.151 and $\mathrm{I}_* := (a_*, b_*)$ solves the Stefan problem posed by (S1)–(S2) with the functional constraint (S3). The "convexity analysis" above also shows that $\mathrm{I}_* \in \mathcal{I}^+$, i.e., I_* is a proper subinterval of $\mathrm{I}_\infty = (0,1)$. As pointed out earlier, given (E) the fact that I_* is a proper subinterval of I_∞ implies that $[\,\mathrm{I}_*\,]_\Pi$ is also a proper subinterval of I_∞ and indeed $[a_*, \mathcal{A}_u b_*) \subset (0,1)$. Condition (S4) is obtained since $r_* \in \mathcal{BC}([\,\mathrm{I}_*\,]_\Pi)$. In addition, our analysis shows that $a_* \le \pi_e \le b_*$. It only remains to show that the running cost conditions are satisfied. We consider this point and present the theorem for sequential detection in the next subsection.

4.6.6 Main Result

Consider the following technical conditions involving the intensity process and its P_1–projection onto the observations:

$$\begin{aligned}
&\text{(J0)}:\quad E_i \int_0^t J_s\, ds < \infty && \forall\, t \ge 0, \quad i = 0,1;\\
&\text{(J1)}:\quad E_i \int_0^\tau \hat{J}_s^2\, ds < \infty && \forall\, \tau \in \mathcal{T}, \quad i = 0,1;\\
&\text{(J2)}:\quad P_i \{ \int_0^\infty \hat{J}_s^2\, ds = \infty \} = 1 && i = 0,1.
\end{aligned}$$

We have come to the main result of this section.

Theorem 4.1 *Assume that the conditions* (J0), (J1),*and* (J2) *hold. In the problem of sequential detection based on observations of the process,*

$$N_t = \begin{cases} \lambda^0 \int_0^t J_s\, ds + \eta_t & t \geq 0 \quad \text{if } \; v = \infty; \\ \lambda^1 \int_0^t J_s\, ds + \eta_t & t \geq 0 \quad \text{if } \; v = 0, \end{cases}$$

with $\lambda^1 > \lambda^0 > 0$ *and with average running cost,*

$$E_\pi \int_0^\tau \mathcal{C}_s\, ds = E_\pi \int_0^\tau c\, \hat{J}_s\, ds \qquad c > 0,$$

and average decision cost,

$$E_\pi[\,\mathcal{E}(\Upsilon_\tau, \delta)] = E_\pi[c^0\,(1-\delta)\,\Upsilon_\tau + c^1\,\delta\,(1-\Upsilon_\tau)]$$

with $0 < c^0,\ c^1 < \infty$, *there exist* a_*, b_* *unique with* $0 < a_* < \pi_e < b_* < 1$, *such that the first exit policy* $(\tau^{\mathrm{I}_*}, \delta_*)$ *based on the continuation interval* $\mathrm{I}_* = (a_*, b_*)$ *achieves Bayes' optimal cost, i.e.,*

$$\rho_\pi(\tau^{\mathrm{I}_*}) = \inf_{\tau \in \mathcal{T}_{ad}} \rho_\pi(\tau) \qquad \forall\, \pi \in [0,1],$$

where,

$$\rho_\pi(\tau) = E_\pi[\int_0^\tau \mathcal{C}_s\, ds + e(\Pi_\tau)] \quad \forall\, \pi \in [0,1] \text{ and } \tau \in \mathcal{T}_{ad}.$$

In addition, there exists $r_* \in \mathcal{C}^{1+}(0,1)$, *the solution to* (S1)–(S4) *above, such that*

$$\rho_\pi(\tau^{\mathrm{I}_*}) = \begin{cases} r_*(\pi) & \text{if } \pi \in \mathrm{I}_*; \\ e(\pi) & \text{if } \pi \notin \mathrm{I}_*, \end{cases}$$

where,

$$e(\pi) = \min\{c^0\,\pi, c^1\,(1-\pi)\}.$$

Proof:

In the previous subsection we solved an associated Stefan problem for which it was shown that there exists a pair (r_*, I_*) satisfying (S1)–(S4). This Stefan problem grew out of a binary hypothesis testing problem comparing an intensity process J against a unity intensity. The reader may check that it is only the ratios of the intensity that matter and here that ratio is $\frac{\lambda^1}{\lambda^0}$. Thus the same results apply if we interpret u as

$$u = \frac{\lambda^0}{\lambda^1 - \lambda^0}. \tag{4.158}$$

With these changes, the results of the last subsection serve to demonstrate the existence of a pair (r_*, I_*), $\mathrm{I}_* \in \mathcal{I}^+$ and $r_* \in \mathcal{BC}(\mathcal{I}^+)$, which satisfy (V1)–(V4). Moreover, $[\mathrm{I}_*]_\Pi = [a_*, \mathcal{A}_u b_*)$ is a strict subset of $(0,1)$ and therefore (V4) is also obtained.

To employ the Verification Theorem and therefore prove the theorem at hand it remains only to show that (C1), (C2), and (C3) hold since (E) follows from Proposition 4.8. With the above choice of running cost we see that (C1) follows *a fortiori* from (J1) since,

$$E_\pi \int_0^\tau \mathcal{C}_s\, ds = \pi\, E_1 \int_0^\tau c\,\hat{J}_s\, ds + (1-\pi)\, E_0 \int_0^\tau c\,\hat{J}_s\, ds < \infty. \tag{4.159}$$

Next, it is obvious that (C2) follows from (J2) since $c > 0$ and thus,

$$\begin{aligned} P_\pi\{\int_0^\infty \mathcal{C}_s\, ds = \infty\} &= \pi\, P_1\{\int_0^\infty c\,\hat{J}_s\, ds\} + (1-\pi)\, P_0\{\int_0^\infty c\,\hat{J}_s\, ds\} \\ &= \pi \cdot 1 + (1-\pi)\cdot 1 = 1. \end{aligned} \tag{4.160}$$

Finally, condition (C3) follows since the running cost $\mathcal{C}$ is trivially concave in Π. Thus τ^{I_*} is the optimal stopping time for this problem and r_* characterizes Bayes' cost. □

4.6.7 Example

We end this section of the chapter with a concrete example of a sequential detection problem involving a Poisson process with one of two constant rates. We observe a Poisson counting process $N = \{N_t\}_{t\geq 0}$ for which one of the following hypotheses is true:

$$\begin{aligned} &(\text{Unit Rate}): \quad N_t = t + \eta_t \qquad && t \geq 0; \\ &(\text{Higher Rate}): \quad N_t = \lambda t + \eta_t \qquad && t \geq 0, \end{aligned}$$

where η is a $(\mathcal{G}_t, P_i)$–martingale for $i = 0$, 1 and $\lambda > 1$. It is given that the (Higher Rate) hypothesis occurs with prior probability $\pi \in [0, 1]$. Define the Bayes' cost,

$$\bar{\rho}_\pi(\tau, \delta) = \pi\, E_1[c\,\tau + c^0\, 1\{\delta = 0\}\,] + (1 - \pi)\, E_0[c\,\tau + c^1\, 1\{\delta = 1\}\,], \quad (4.161)$$

where c, c^0, and c^1 are strictly positive and finite. We are asked to minimize $\bar{\rho}_\pi(\tau, \delta)$ over all decision pairs for which $P_i\{\tau < \infty\} = 1$ for $i = 0, 1$.

We see that this is precisely the form we are equipped to handle and our choice of running cost collapses down to

$$\int_0^\tau \mathcal{C}_s\, ds = c\tau. \quad (4.162)$$

We point out that each of (J0), (J1) and (J2) are trivially satisfied. Thus we can apply Theorem 3.1 to deduce that an optimal first exit policy exists. One can solve for the optimal pair (r_*, I_*) using the "convexity" approach given above. The graphs in this case for $c = 1$ and $\lambda = 2$ are depicted in Figures 3.1 through 3.5. The graph of ρ, the Bayes' optimal risk is given in Figure 3.6 along with the worst case risk. Figure 3.6 neatly shows quite clearly that the so-called *smooth pasting property* of ρ and e as discussed in the literature is not obtained for the Poisson process case.

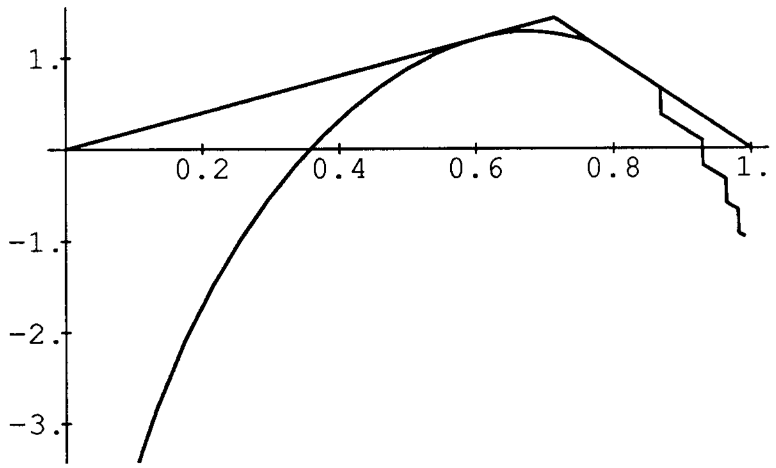

Figure 4.6: Graph of e and r_* with $a_* \approx 0.60445$, $b_* \approx 0.76627$.

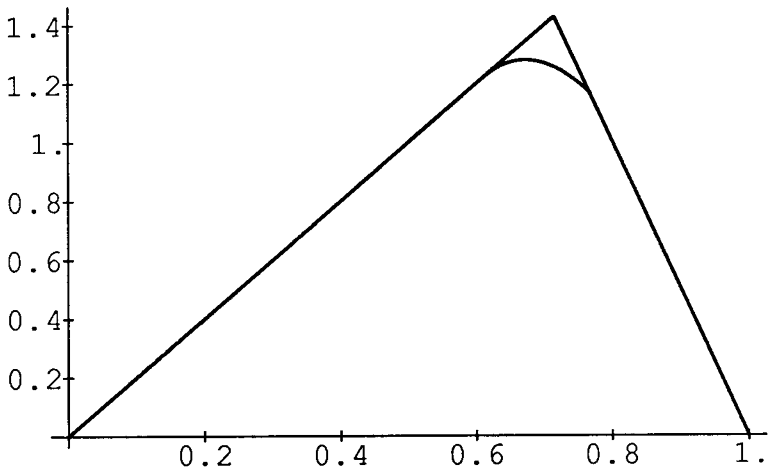

Figure 4.7: Graph of terminal cost and Bayes' optimal risk.

Chapter 5

Conclusion

In this thesis we considered problems of change detection under Bayesian assumptions for the costs and prior information. Our data streams were modeled using semimartingales so that we could employ modern martingale filtering results. We considered sequential detection problems and disruption problems in a unified framework using the simple device of allowing the time of change to take on possibly infinite values and defining its general probability model on this larger set.

We circumscribed a set of conditions, called the *verification* conditions which permit one to verify the optimality of a proposed threshold policy. We showed that under certain modeling assumptions made with respect to the observed data and performance costs that the verification conditions reduce the problem of searching for an optimal first exit problem to solving a kind of free boundary value problem. We showed how to solve this kind of Stefan problem in three problems of change detection using the same abstract approach. This approach was likened to a kind of convex analysis involving the notion of generalized hyperplane. Whether the validity of this viewpoint will be borne out by future investigations remains to be seen, but the authors feel strongly that the viewpoint deserves merit. The major difficulty to achieving a complete notion of generalized convexity along the lines of [BECK] turned out to be the lack of continuity generic to the solutions of the FDEs which have been called the *bugbear of the subject* [MHAD]. This notwithstanding, using the results contained herein, the actual solutions for the optimal thresholds are now easily obtained with roughly the same amount of work in the jump process case as in the diffusion case. Computationally at least therefore, our work goes a long way towards taming the animal. Moreover, the methods employed here are sufficiently general to allow one to compute the optimal thresholds in examples which we did not consider. This is made possible by the discovery of the algorithm for computing the optimal continuation interval which relies on the sequence of approximating generalized hyperplanes. This algorithm, which has the same specification irrespective of whether the noise is impulsive or not, amounts to an efficient method to compute the optimal thresholds, a task which has always been considered impractical and next to impossible. Much of the thesis was devoted to showing that it also provides the basis for an abstract theory powerful enough to prove that optimal first exit policies exist and can be characterized.

Appendix A

Some Integration Formulas

In this appendix we give some generalizations to Lebesgue-Stieltjes integration formulas. In the lemma below the formula given is the same as the usual one under a modest relaxation of continuity of the derivative of the composing function. We will denote by X^c the continuous part of X, namely,

$$X_t^c := X_t - \sum_{s \le t} [X_s - X_{s-}] \quad \forall\, t \ge 0, \tag{A.1}$$

where,

$$X_{t-} := \lim_{s \uparrow t} X_s \quad \forall\, t \ge 0. \tag{A.2}$$

Lemma A.1 *Let $\{X_t\}_{t \ge 0}$ denote a real-valued corlol function of bounded variation. Let $F : \Re \to \Re$ be continuous with a piecewise-continuous derivative having at most a finite number of discontinuities*[1]*. Then,*

$$F(X_t) = F(X_0) + \int_0^t F'(X_s)\, dX_s^c + \sum_{s \le t} [F(X_s) - F(X_{s-})] \quad \forall\, t \ge 0.$$

[1] Or, a countable set of isolated discontinuities, i.e., with no accumulation point.

Proof: Suppose F has a single jump discontinuity in its derivative at $x \in \Re$; for convenience we can take $x = 0$. We know that we can approximate F with a sequence of continuously differentiable functions $\{F_n\}_{n\geq 1}$ such that F_n converges uniformly to F on $\Re$ and F_n' converges in L^1 to F'. Since F_n is continuously differentiable for all $n \geq 1$ we may apply the usual Lebesgue-Stieltjes formula [E, L13.2] to obtain,

$$F_n(X_t) - F_n(X_0) = \int_0^t F_n'(X_s)\, dX_s^c + \sum_{s\leq t}[F_n(X_s) - F_n(X_{s-})] \quad \forall\, t \geq 0. \tag{A.3}$$

The limit on the left-hand side is obviously $F(X_t) - F(X_0)$. As for the right-hand side, begin by writing,

$$\int_0^t F_n'(X_s)\, dX_s^c = \int_0^t F_n'(X_s)\, dX_{1,s}^c - \int_0^t F_n'(X_s)\, dX_{2,s}^c, \tag{A.4}$$

where X_1^c, X_2^c denote the positive and negative variation of X^c, respectively. Using the Bounded Convergence theorem we have for each $i = 1, 2$,

$$\lim_{n\to\infty} \int_0^t F_n'(X_s)\, dX_{i,s}^c = \int_0^t F'(X_s)\, dX_{i,s}^c, \tag{A.5}$$

and hence there follows,

$$\lim_{n\to\infty} \int_0^t F_n'(X_s)\, dX_s^c = \int_0^t F'(X_s)\, dX_s^c. \tag{A.6}$$

As for the summation, define,

$$B(F, X, t) := \sup\{\, |F'(x)| : 0 < |x| \leq \sup_{s\leq t} |X_s| \,\}. \tag{A.7}$$

Letting $[x]_\pm$ denote $\max\{0, \pm x\}$ we see that,

$$\sum_{s\leq t}[F_n(X_s) - F_n(X_{s-})]_\pm \leq B(F_n, X, t) \sum_{s\leq t}[X_s - X_{s-}]_\pm, \tag{A.8}$$

and clearly,

$$B(F_n, X, t) \leq B(F_n - F, X, t) + B(F, X, t) \quad \forall t \geq 0. \qquad \text{(A.9)}$$

Also, since the set $\{ x \in \Re : 0 < |x| \leq \sup_{s \leq t} |X_s| \}$ is bounded for all $t \geq 0$ and since the convergence of F_n' to F is uniform almost everywhere, for any $\epsilon > 0$ we can find n large enough such that,

$$B(F_n - F, X, t) = \sup\{ |F_n'(x) - F'(x)| : 0 < |x| \leq \sup_{s \leq t} |X_s| \} \leq 1. \quad \text{(A.10)}$$

Therefore, letting $B_t := 1 + B(F, X, t)$ we see that for large enough n,

$$\sum_{s \leq t} [F_n(X_s) - F_n(X_{s-})]_{\pm} \leq B_t \sum_{s \leq t} |X_s - X_{s-}| < \infty, \qquad \text{(A.11)}$$

and so by the Monotone Convergence theorem,

$$\lim_{n \to \infty} \sum_{s \leq t} [F_n(X_s) - F_n(X_{s-})]_{\pm} = \sum_{s \leq t} [F(X_s) - F(X_{s-})]_{\pm} < \infty. \qquad \text{(A.12)}$$

As a result,

$$\begin{aligned} \lim_{n \to \infty} \sum_{s \leq t} [F_n(X_s) - F_n(X_{s-})] &= \sum_{s \leq t} [F(X_s) - F(X_{s-})]_{+} \\ &\qquad - \sum_{s \leq t} [F(X_s) - F(X_{s-})]_{-} \\ &= \sum_{s \leq t} [F(X_s) - F(X_{s-})]. \qquad \text{(A.13)} \end{aligned}$$

Thus, taking limits in A.3 gives us the result for F having a single discontinuity in its derivative. The general case simply follows by linearity. □

In the next lemma and its corollary we consider a degenerate Lebesgue-Stieltjes integration formula for the case where the derivative of the composing function is zero almost everywhere.

Lemma A.2 *Let* $\{X_t\}_{t\geq 0}$ *denote a real-valued corlol function and let* D_X *denote the set of points at which* X *is not continuous. Assume that* X *is decreasing on* D_X*, nondecreasing off of* D_X*, and that* D_X *is finite. Suppose that* F *is given by,*

$$F(x) := \begin{cases} 0 & x < 0; \\ 1 & x \geq 0. \end{cases}$$

Then $\{F(X_t)\}_{t\geq 0}$ *is corlol and obeys the formula,*

$$F(X_t) = F(X_0) + \sum_{s\leq t}[F(X_s) - F(X_{s-}^-)] \quad \forall\, t \geq 0,$$

where,

$$F(x^-) := \lim_{y\uparrow x} F(y).$$

Proof: It is helpful to consider Table A.1. The first two rows of the table identify the zero-crossings of X which affect $F(X)$. According to these two entries we may write for all $t \geq 0$,

$$\sum_{s\leq t}[F(X_s) - F(X_{s-}^-)] = \sum_{s\leq t}[1\{X_s = 0 = X_{s-}\} - 1\{X_s < 0 < X_{s-}\}]. \quad \text{(A.14)}$$

The third and fourth rows of the table delineate the "not-quite" zero-crossings of X which have no effect on $F(X)$. The last two rows are those discontinuities of X which quite obviously have no effect on the trajectories of $F(X)$.

Let's check the formula. Assume that $F(X_0) = 0$; the argument for $F(X_0) = 1$ is entirely symmetric. Observe, $F(X_0) = 0$ implies $X_0 < 0$. If $X_s < 0$ for all $s \leq t$ then A.14 is identically zero and $F(X_t) = 0 = F(X_0)$

$t \geq 0$	$F(X_t)$	$F(X_{t-}^-)$	$F(X_t) - F(X_{t-}^-)$
$X_t < 0 < X_{t-}$	0	1	-1
$X_t = 0 = X_{t-}$	1	0	1
$X_t < 0 = X_{t-}$	0	0	0
$X_t = 0 < X_{t-}$	1	1	0
$X_t < X_{t-} < 0$	0	0	0
$0 < X_t < X_{t-}$	1	1	0

Table A.1: Zero-crossing behavior of F

verifies the formula. So suppose instead that $X_s \geq 0$ for some $s \leq t$. Either $X_t \geq 0$ or not. Consider $X_t < 0$. Then $X_t < 0 \leq X_s$ and $F(X_t) = 0$ which says that X must have zero-crossed an even number of times: an odd number of times, say $n \geq 1$, in a continuous manner and the same odd number of times discontinuously. Thus we see that $F(X)$ again obeys the formula since,

$$\begin{aligned} F(X_t) &= F(X_0) + \sum_{s \leq t}[F(X_s) - F(X_{s-}^-)] \\ &= 0 + \sum_{s \leq t} 1\{X_s = 0 = X_{s-}\} - \sum_{s \leq t} 1\{X_s < 0 < X_{s-}\} \\ &= n - n = 0. \end{aligned} \tag{A.15}$$

Now consider the remaining case $X_t \geq 0$ for which $F(X_t) = 1$. Here, X has zero-crossed an odd number of times: n times continuously and $n-1$ times by jumping. Once again $F(X)$ obeys the formula since,

$$\begin{aligned} F(X_t) &= F(X_0) + \sum_{s \leq t} 1\{X_s = 0 = X_{s-}\} - \sum_{s \leq t} 1\{X_s < 0 < X_{s-}\} \\ &= 0 + n - (n-1) = 1. \end{aligned} \tag{A.16}$$

Consequently, it follows that $F(X)$ satisfies the formula in general and we can now employ it to show that $F(X)$ is corlol. That $F(X)$ has limits on the left is clear. To see that it is also continuous on the right note that,

$$\begin{aligned}
\lim_{u\downarrow t} F(X_u) &= F(X_0) + \lim_{u\downarrow t} \sum_{s\leq u} [F(X_s) - F(X_{s-}^-)] \\
&= F(X_t) + \lim_{u\downarrow t} \sum_{t<s\leq u} [F(X_s) - F(X_{s-}^-)] \\
&= F(X_t) + \lim_{u\downarrow t} \sum_{t<s\leq u} [1\{X_s = 0 = X_{s-}\} - 1\{X_s < 0 < X_{s-}\}] \\
&= F(X_t), \qquad (A.17)
\end{aligned}$$

and this completes the proof.

Corollary A.1 *Suppose $F : \Re \to \Re$ is right-continuous and piecewise constant. Let D_F denote the points at which F is discontinuous. If D_F is countable and has no accumulation point then,*

$$F(X_t) = F(X_0) + \sum_{s\leq t} [F(X_s) - F(X_{s-}^-)] \quad \forall\, t \geq 0.$$

Proof: Write $D_F = \{x_1, x_2, \ldots\}$ and denote the jump sizes via,

$$j_n := F(x_n) - F(x_n^-) \quad \forall\, n \geq 1. \qquad (A.18)$$

This implies that,

$$F(x) = \sum_{n=1}^{\infty} j_n\, 1\{x_n \leq x\} \quad \forall\, x \in \Re, \qquad (A.19)$$

and of course there are no convergence technicalities since the summation is in fact finite due to our assumptions concerning D_F. Now define,

$$F_n(x) := 1\{x_n \leq x\} \quad \forall\, x \in \Re, \qquad (A.20)$$

so that from the lemma we obtain (by a simple translation of the origin),

$$F_n(X_t) = F_n(X_0) + \sum_{s \le t} [F_n(X_s) - F_n(X_{s-}^-)]. \tag{A.21}$$

Plugging this into the above yields,

$$\begin{aligned}
F(X_t) &= \sum_{n=1}^{\infty} j_n \, F_n(X_t) \\
&= \sum_{n=1}^{\infty} j_n \left[F_n(X_0) + \sum_{s \le t} [F_n(X_s) - F_n(X_{s-}^-)] \right] \\
&= \sum_{n=1}^{\infty} j_n \, F_n(X_0) + \sum_{n=1}^{\infty} \sum_{s \le t} j_n \, [F_n(X_s) - F_n(X_{s-}^-)] \\
&= F(X_0) + \sum_{s \le t} \left[\sum_{n=1}^{\infty} j_n \, F_n(X_s) - \sum_{n=1}^{\infty} j_n \, F_n(X_{s-}^-) \right] \\
&= F(X_0) + \sum_{s \le t} [F(X_s) - F(X_{s-}^-)],
\end{aligned} \tag{A.22}$$

i.e., we get what we want by exploiting linearity. □

In the next proposition we combine the previous lemmas to give an integration formula which is a modest generalization of the usual Lebesgue-Stieltjes formula.

Proposition A.1 *Let $\{X_t\}_{t \ge 0}$ denote a real-valued corlol function as in Lemma A.2 and let $F : \Re \to \Re$ be a right-continuous function with a piecewise-constant (right) derivative F', both F and F' possessing a finite number of discontinuities. Then,*

$$F(X_t) = F(X_0) + \int_0^t F'(X_s) \, dX_s^c + \sum_{s \le t} [F(X_s) - F(X_{s-}^-)] \quad \forall t \ge 0.$$

Proof: Define the function,

$$F_1(x) := \sum_{w \le x} [F(w) - F(w^-)] \quad \forall x \in \Re, \tag{A.23}$$

which is seen to satisfy the hypotheses of Lemma A.2 (right-continuous, finite number of positive jumps). Now if we define $F_2 := F - F_1$, then F_2 satisfies the hypotheses of Lemma A.1 (corlol, continuous, F' piecewise continuous). Therefore, applying Lemma A.1 to F_2 gives,

$$F_2(X_t) = F_2(X_0) + \int_0^t F_2(X_s)\, dX_s^c + \sum_{s \le t} [F_2(X_s) - F_2(X_{s-}^-)] \quad \forall t \ge 0, \tag{A.24}$$

and applying Lemma A.2 to F_1 yields,

$$F_1(X_t) = F_1(X_0) + \sum_{s \le t} [F_1(X_s) - F_1(X_{s-}^-)]. \tag{A.25}$$

Hence, because $F = F_1 + F_2$ these results give,

$$F(X_t) = F(X_0) + \int_0^t F_2'(X_s)\, dX_s^c + \sum_{s \le t} [F(X_s) - F_2(X_{s-}) - F_1(X_{s-}^-)]. \tag{A.26}$$

Clearly,

$$\int_0^t F_1'(X_s)\, dX_s^c = 0 \quad \forall t \ge 0, \tag{A.27}$$

and also

$$F_2(X_{t-}) + F_1(X_{t-}^-) = F_2(X_{t-}^-) + F_1(X_{t-}^-) = F(X_{t-}^-) \quad \forall t \ge 0. \tag{A.28}$$

Substitute these two expressions into the previous and simplify to obtain,

$$F(X_t) = F(X_0) + \int_0^t F'(X_s)\, dX_s^c + \sum_{s \le t} [F(X_s) - F(X_{s-}^-)] \quad \forall t \ge 0, \tag{A.29}$$

and this is the result. □

Before ending this appendix we include another proposition which is companion to the previous. It considers the case where F is still right-continuous but X instead has positive jumps and is otherwise decreasing. In this case the formula can be changed slightly in order to work, but it fails to hold on a set of measure zero. Moreover, due to the "incompatible" continuity handedness of F and X we cannot conclude that $F(X)$ is corlol, indeed, in general it is neither right-continuous nor left-continuous although it does possess limits on both the left and right. We start as before with a lemma.

Lemma A.3 *Let $\{X_t\}_{t\geq 0}$ denote a real-valued corlol function and let D_X denote the set of points at which X is not continuous. Assume that X is increasing on D_X, nonincreasing off of D_X and D_X is finite. Suppose F is right-continuous and piecewise constant. Let D_F denote the points at which F is discontinuous. If D_F is countable and has no accumulation point then,*

$$F(X_t) = F(X_0) + \sum_{s\leq t}[F(X_s) - F(X_{s-}^{-})] \quad \forall\, t \geq 0,$$

except on the finite set $\{t \geq 0 : X_{t-} < 0 = X_t\}$.

Proof: For simplicity we can argue with $F(x) = 1\{0 \leq x\}$ as in Lemma A.2. Consider the analogous table, Table A.2. From this table it follows that,

$$\sum_{s\leq t}[F(x_s^-) - F(X_s)] = \sum_{s\leq t}[1\{\, X_{t-} < 0 < X_t\} - 1\{\, X_{t-} = 0 = X_t\,\}]. \quad \text{(A.30)}$$

We can now repeat the arguments of Lemma A.2 and its corollary to prove the claim. □

$t \geq 0$	$F(X_t^-)$	$F(X_{t-})$	$F(X_t^-) - F(X_{t-})$
$X_{t-} = 0 = X_t$	0	1	-1
$X_{t-} < 0 < X_t$	1	0	1
$X_{t-} < X_t < 0$	0	0	0
$X_{t-} = 0 < X_t$	1	1	0
$X_{t-} < 0 = X_t$	0	0	0
$0 < X_{t-} < X_t$	1	1	0

Table A.2: Zero-crossing behavior of F

As for $F(X)$ not being corlol in the lemma above note that when $X_{t-} = 0 = X_t$, i.e., when X crosses the origin continuously (downward) we have,

$$F(X_t) = \lim_{s \uparrow t} F(X_s) = 1 \neq 0 = \lim_{s \downarrow t} F(X_s). \tag{A.31}$$

So, $F(X)$ cannot be right-continuous. On the other hand, if $X_{t-} < 0 < X_t$ then,

$$F(X_t) = \lim_{s \downarrow t} F(X_s) = 1 \neq 0 = \lim_{s \uparrow t} F(X_s). \tag{A.32}$$

So, $F(X)$ is not left-continuous either. Even more pathologically, if $X_{t-} < 0 = X_t$, then $F(X_t) = 1$ but,

$$\lim_{s \uparrow t} F(X_s) = 0 = \lim_{s \downarrow t} F(X_s). \tag{A.33}$$

Note that from a stochastic viewpoint, for instance when X is some counting process, things are typically arranged so that the event $\{X_{t-} < 0 = X_t\}$ has zero probability. Such events can be handled so that $F(X)$ can usually be replaced by a stochastically equivalent left- or right-continuous *version* by employing a standard process *modification* argument. Therefore when we compose F with a stochastic process of locally finite variation and bounded mean, the deterministic requirement that X possess a finite number of discontinuities is replaced by an assumption that X is *non-explosive.* In this way the results carry over via simple pathwise arguments. We can now state the final result of the appendix.

Proposition A.2 *Let* $\{X_t\}_{t\geq 0}$ *denote a real-valued corlol function as in Lemma A.3 and let* $F : \Re \rightarrow \Re$ *be a right-continuous function with a piecewise-constant derivative* F'*, both* F *and* F' *possessing a finite number of discontinuities. Then,*

$$F(X_t) = F(X_0) + \int_0^t F'(X_s)\,dX_s^c + \sum_{s\leq t}[F(X_s^-) - F(X_{s-})] \quad \forall\, t \geq 0,$$

except on the finite set $\{t \geq 0 : X_{t-} < 0 = X_t\}$.

Proof: Use the same arguments as in Proposition A.1. □

Appendix B

Construction of the Generalized Hyperplane

The purpose of this appendix is to prove the existence and uniqueness of the generalized hyperplane associated with the computation of the optimal thresholds in the detection problem for the jump process case. The proof consists of a limiting argument involving a sequence of (generalized) approximating hyperplanes. It is somewhat constructive in the sense that it suggests a practical algorithm for computing the optimal thresholds.

Recall the problem posed by (S1)–(S2) (see Chapter 4, Section 5). Define the linear functional-differential operator D_u for all $r \in \mathcal{C}^{1+}(0,1)$ via,

$$D_u\, r(\pi) := -\pi(1-\pi)r'(\pi) + (u+\pi)\left[r\left(\mathcal{A}_u\pi\right) - r(\pi)\right]. \tag{B.1}$$

The derivative r' is understood as taken from the right and,

$$\mathcal{A}_u(\pi) := \frac{u+1}{u+\pi}\,\pi \qquad 0 < \pi < 1,\, u > 0. \tag{B.2}$$

With a, $b \in (0,1)$ satisfying $0 < a \leq \pi_e \leq b < 1$ and $a < b$, the problem involves the following functional-differential equation (FDE) for $r \in \mathcal{C}^{1+}(0,1)$,

$$D_u\, r(\pi) = -c \qquad \forall\, \pi \in (0,1), \tag{B.3}$$

together with the equality constraints,

$$\begin{aligned} r(a) &= e(a); \\ r(\pi) &= e(\pi) \qquad \forall\, \pi \in [b, \mathcal{A}_u b). \end{aligned} \tag{B.4}$$

The unique solution to this FDE is given explicitly in Appendix C which for convenience we repeat here,

$$r(\pi; a, b) = e_1(\pi) + d(\pi; b) + \overline{D}(\pi; b) + K(a,b)\,\overline{H}(\pi; b) \quad \forall\, \pi \in (0,1), \tag{B.5}$$

with u, e_1, d, $\overline{D}$, K and $\overline{H}$ given also in Appendix C. The reader can verify that B.5 satisfies B.3, in fact for any $0 < a < b < 1$. In what follows we shall give a recipe for choosing a sequence of continuation intervals (a_n, b_n), $n = 0, 1, \ldots$, with endpoints satisfying $0 < a_n \leq \pi_e \leq b_n < 1$ and $a_n < b_n$, which converge to the optimal continuation interval (a_*, b_*). Equivalently, this will define a sequence of functions r_n which converge pointwise to r_*, the solution to (S1)-(S4) for the jump case.

An important ingredient in our recipe is the **gap function** defined for $a \in (0,1)$ via,

$$G(\pi; a) := \lim_{x \uparrow \pi} r(x; a, \pi) - e_1(\pi) \qquad 0 < a < \pi < 1. \tag{B.6}$$

The gap function with respect to any $a \in (0,1)$ gives us for each $\pi \in (a,1)$ the height of the discontinuity in $r(\cdot\,; a, \pi)$ at π. It is shown in Appendix C

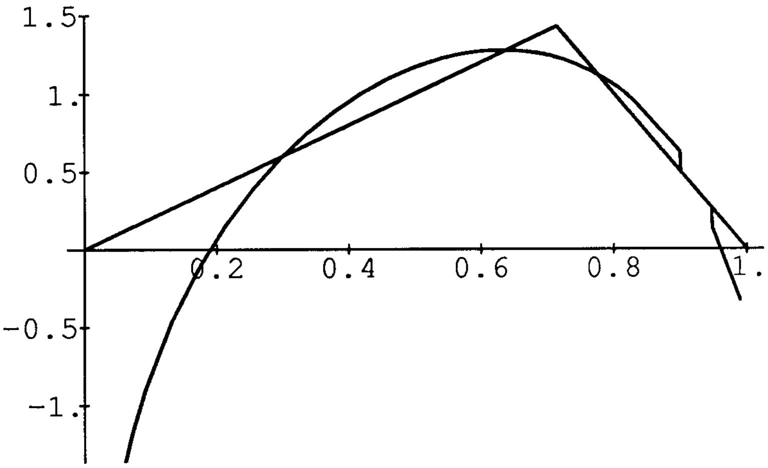

Figure B.1: Graph of e and r_0 with $c^0 = 2$, $c^1 = 5$; $\lambda^0 = 1$, $\lambda^1 = 2$, $c = \frac{1}{4}$ and $a_0 = \frac{3}{10}$, $b_0 = \frac{9}{10}$.

that $G(\cdot\,; a)$ is a continuous function for all $a \in (0,1)$ and that,

$$\lim_{\pi \downarrow a} G(\pi; a) < 0, \tag{B.7}$$

while,

$$\lim_{\pi \uparrow 1} G(\pi; a) > 0. \tag{B.8}$$

Hence, for any $a \in (0,1)$ there exists $b = b(a) \in (a,1)$ such that,

$$G(b\,; a) = 0. \tag{B.9}$$

This property of the gap function will be exploited at key steps in the convergence argument below.

STEP 1. Pick $a_0 \in (0, \pi_e)$ and $b_0 \in (\pi_e, 1)$. Let r_0 denote $r(\cdot\,; a_0, b_0)$; for a graph of r_0 see Figure B.1.

By definition r_0 crosses the terminal cost function 'triangle' e at a_0. In general there will be two distinct crossover points to the left of π_e, $\underline{a}_0$ and $\overline{a}_0$, at which r_0 crosses e; obviously $a_0 \in \{\underline{a}_0, \overline{a}_0\}$. If there is only one crossover point, i.e., $\underline{a}_0 = \overline{a}_0$ then skip to the last step of the argument. Otherwise assume $\underline{a}_0 < \overline{a}_0$. Before proceeding we need the following lemma.

Lemma B.1 *Fix $b \in (0,1)$ and consider $L \in \ker D_u$ such that,*

$$L(\pi) = 0 \qquad \forall \pi \geq b. \tag{B.10}$$

Suppose L has a continuous derivative on $(0, \mathcal{A}_u^{-1} b)$ and is positive at $\mathcal{A}_u^{-1} b$. Then L is positive and decreasing on $(0, b)$, i.e.,

$$L'(\pi) < 0 < L(\pi) \qquad \forall \pi < b. \tag{B.11}$$

Proof: Let B denote the point to the left of b for which $\mathcal{A}_u B = b$; therefore we may write $B = \mathcal{A}_u^{-1} b$ for simpler notation in the proof. Since $L \in \ker D_u$ we have,

$$L'(\pi) = \frac{u + \pi}{\pi(1-\pi)}[L(\mathcal{A}_u \pi) - L(\pi)] \qquad \forall \pi \in (0,1), \tag{B.12}$$

so that the assumption $L \equiv 0$ on $[b, 1)$ leads to,

$$L'(\pi) = -\left[\frac{u + \pi}{\pi(1-\pi)}\right] L(\pi) \qquad \forall \pi \in [B, b). \tag{B.13}$$

Integrating this ODE yields,

$$L(\pi) = L(B) \exp\{\int_B^\pi \frac{u + s}{s(1-s)} ds\} \qquad \forall \pi \in [B, b). \tag{B.14}$$

Hence $L(B) > 0$ implies L is positive on $[B, b)$. As a result,

$$L'(\pi) < 0 < L(\pi) \qquad \forall \pi \in [\mathcal{A}_u^{-1} b, b). \tag{B.15}$$

The remainder of the proof is an induction argument. Pick $n \geq 1$ and suppose,

$$L'(\pi) < 0 < L(\pi) \quad \forall \pi \in [\mathcal{A}_u^{-n} b, b). \tag{B.16}$$

Thus,

$$L'(\mathcal{A}_u^{-n} b) < 0, \tag{B.17}$$

so that by the continuity of L' on $(0, \mathcal{A}_u^{-n} b)$ there exists a smallest $\pi_0 \in [\mathcal{A}_u^{-n-1} b, \mathcal{A}_u^{-n} b)$ such that,

$$L'(\pi) < 0 \qquad \forall \pi \in (\pi_0, \mathcal{A}_u^{-n} b). \tag{B.18}$$

From expression B.12 we obtain,

$$L(\pi) > L(\mathcal{A}_u \pi_0) \qquad \forall \pi \in (\pi_0, \mathcal{A}_u^{-n} b), \tag{B.19}$$

and thence by continuity,

$$L(\pi_0) \geq L(\mathcal{A}_u \pi_0). \tag{B.20}$$

In fact, since L is strictly decreasing on $(\pi_0, \mathcal{A}_u^{-n} b)$ we have,

$$L(\pi_0) > L(\mathcal{A}_u \pi_0). \tag{B.21}$$

Now by the upper inequality B.16 of the induction hypothesis this leads to,

$$L(\pi_0) > 0, \tag{B.22}$$

and using B.12 again,

$$L'(\pi_0) < 0. \tag{B.23}$$

As a result we must conclude that $\pi_0 = \mathcal{A}_u^{-n-1}b$ and then,

$$L'(\pi) < 0 < L(\pi) \qquad \forall\, \pi \in [\mathcal{A}_u^{-n-1}b, \mathcal{A}_u^{-n}b). \tag{B.24}$$

A final appeal to B.16 gives,

$$L'(\pi) < 0 < L(\pi) \qquad \forall\, \pi \in [\mathcal{A}_u^{-n-1}b, b). \tag{B.25}$$

Since any $\pi \in (0, b)$ is contained in a semi-open interval of this form for some $n \geq 1$ the lemma is established. □

With this lemma in hand we may proceed.

STEP 2. Pick $a_1 \in (\underline{a}_0, \overline{a}_0)$ and define,

$$L_1(\pi) := \left[\frac{r_0(a_1) - e(a_1)}{\overline{H}(a_1; b_0)}\right] \overline{H}(\pi; b_0) \qquad 0 < \pi < 1. \tag{B.26}$$

It is not too difficult to show for any $b \in (0,1)$ that $\overline{H}(\cdot\,, b)$ is in $\ker D_u$ and also that it satisfies the hypotheses of the lemma above. As a result L_1 as defined above also satisfies these hypotheses. Now define R_1 via $R_1 := r_0 - L_1$. The graph of L_1 together with $r_0 - e$ is given in Figure B.2. For a graph of r_0 and R_1 together with e see Figure B.3.

Observe that R_1 lies strictly below r_0 on $(0, b_0)$ because Lemma B.1 ensures that L_1 is positive. Moreover, given the fact that L_1 is in $\ker D_u$ it follows from the existence and uniqueness properties of the FDE that,

$$R_1(\pi) = r(\pi; \underline{a}_1, b_0) \qquad \forall\, \pi \in (0, 1). \tag{B.27}$$

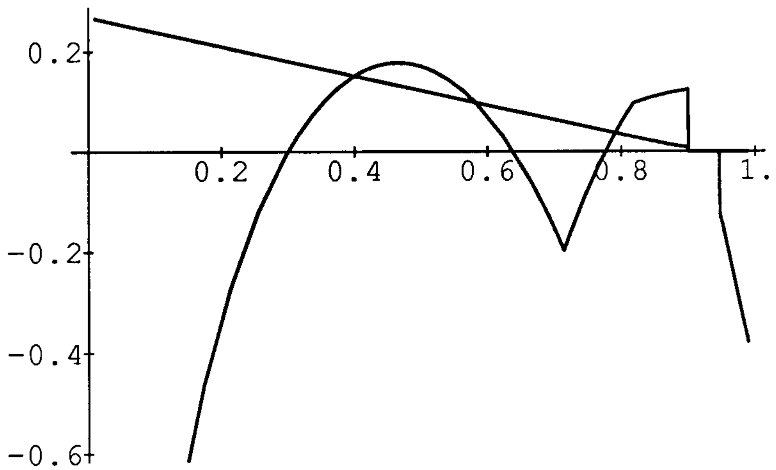

Figure B.2: Graph of $r_0 - e$ and L_1 with $a_1 = \frac{4}{10}$.

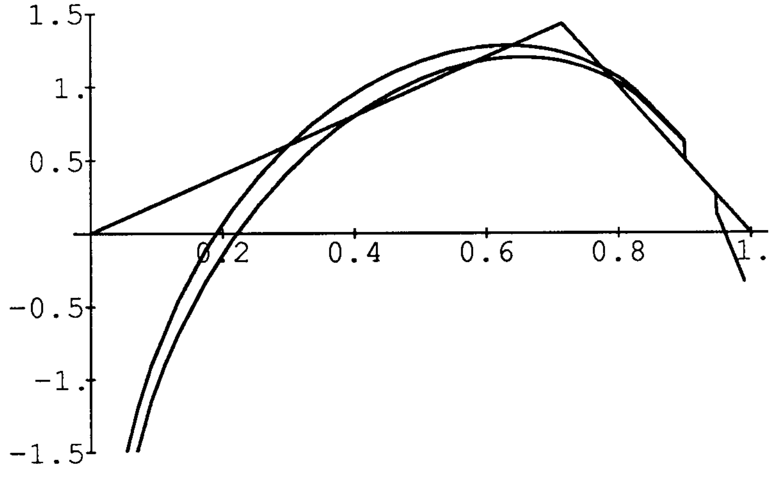

Figure B.3: Graph of r_0 and R_1.

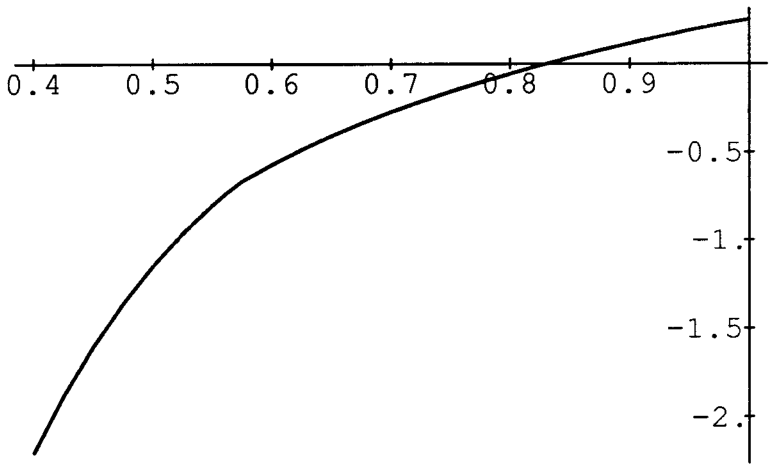

Figure B.4: Graph of the gap function for $a_1 = \frac{4}{10}$, $b_1 = b(a_1) \approx 0.8355$.

Again, in general R_1 will cross e in two places say $\underline{A}$ and $\overline{A}$; see Figure B.3. Without loss of generality we may supose that $a_1 = \underline{A}$. It follows obviously from the positivity of L_1 that

$$0 \leq \overline{A} - \underline{A} < \overline{a}_0 - \underline{a}_0. \tag{B.28}$$

In the next step we employ the gap function G which was defined in B.6.

STEP 3. Define $b_1 \in (a_1, 1)$ to be the solution to $G(\pi; a_1) = 0$ (see Figure B.4) and let r_1 denote $r(\cdot\,; a_1, b_1)$.

By definition r_1 is guaranteed to cross e at a_1. For a graph of r_1 together with R_1 see Figure B.5; a graph of the initial approximation r_0 is displayed with r_1 in Figure B.6. What is crucial to the success of the present argument is that r_1 lies entirely beneath r_0 on at least $(\underline{A}, \overline{A})$. This is because it is our intent to show that $(\underline{a}_1, \overline{a}_1) \subset (\underline{a}_0, \overline{a}_0)$ where $\underline{a}_1$ and $\overline{a}_1$ denote the crossover points for r_1 which lie to the left of π_e. In fact, what appears to be true is

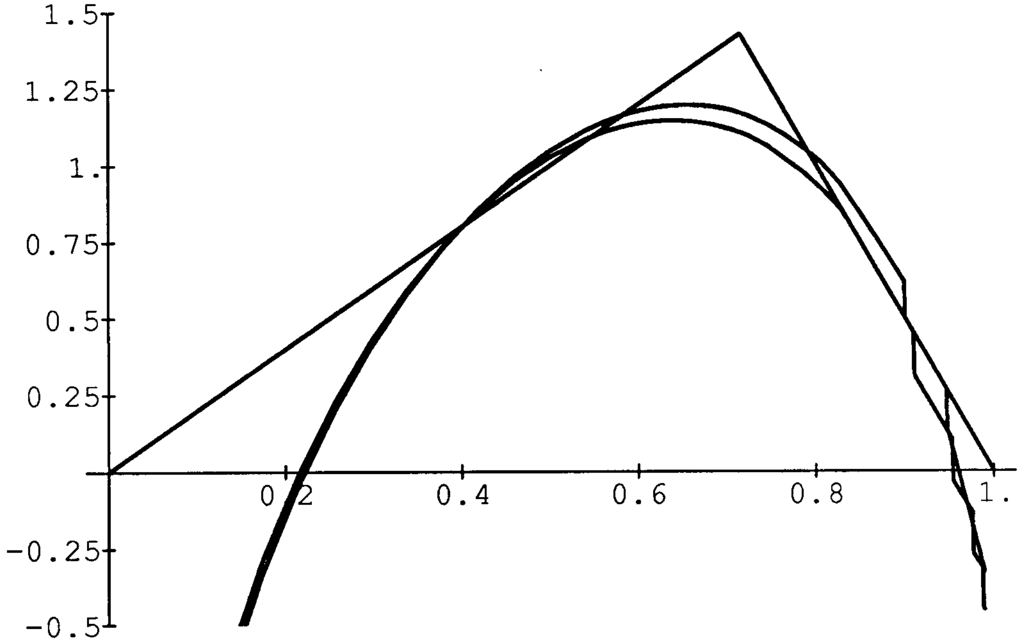

Figure B.5: Graph of R_1 and r_1.

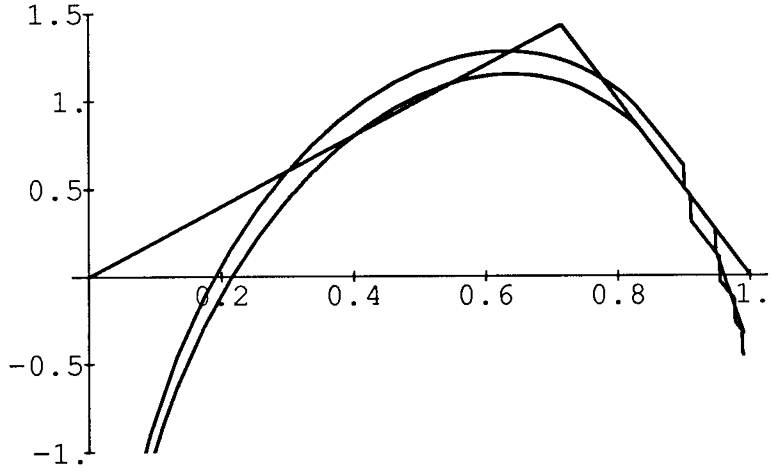

Figure B.6: Graph of r_0 and r_1.

the stronger fact that r_1 lies entirely beneath R_1 on $(\underline{A}, \overline{A})$. However, a proof of either of these hypotheses has not yet been obtained. Instead, we content ourselves with a more complicated argument, namely choose b_1 as close as possible to the solution to $G(\pi; a_1) = 0$ so that the upper crossover point of r_1 remains less that or equal to $\overline{a}_0$. Since the crossover points of the solution to the FDE depend continuously on the initial data (see Appendix C) this can always be done. In doing so we see that we have at least reduced the gap by a positive amount and are ensured that,

$$0 \leq \overline{a}_1 - \underline{a}_1 < \overline{a}_0 - \underline{a}_0. \tag{B.29}$$

We can now iterate the entire process defining $\underline{a}_n$ and $\overline{a}_n$, R_n, b_n, and r_n in the obvious way. In doing so we are guaranteed that,

$$0 \leq \overline{a}_n - \underline{a}_n < \overline{a}_{n-1} - \underline{a}_{n-1} \quad \forall n \geq 1. \tag{B.30}$$

Hence we may define,

$$a_* := \lim_{n \to \infty} \overline{a}_n = \lim_{n \to \infty} \underline{a}_n, \tag{B.31}$$

and then take b_* as the solution to $G(\pi; a_*)$. Lastly, define r_* via

$$r_*(\pi) := r(\pi; a_*, b_*) \qquad \forall \pi \in (0, 1). \tag{B.32}$$

The function r_* satisfies the FDE by definition but also satisfies,

$$r_*(\pi) < e(\pi) \qquad \forall \pi \in (0, b) \setminus \{a_*\}, \tag{B.33}$$

as is clear from the construction of a_*. Of course r_* equals e at a_* and on $[b_*, \mathcal{A}_u b_*)$ since it is a solution of the FDE. For values above $\mathcal{A}_u b_*$, r_* is also

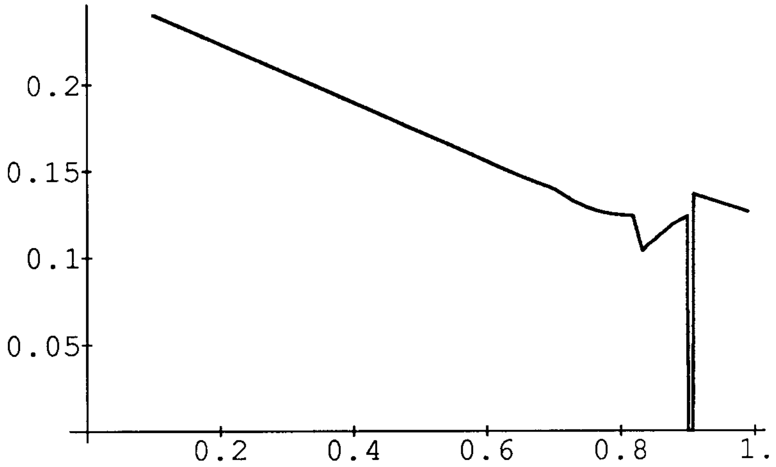

Figure B.7: Graph of ℓ_*.

strictly beneath e since this is a general property of every solution to the FDE as shown in Appendix C. Finally, we define ℓ_* uniquely according to $\ell_* := r_0 - r_*$ and this is precisely the generalization to the hyperplane in the diffusion case which we seek. The graph of ℓ_* for particular values of a_0 and b_0 is displayed by itself in Figure B.7 and with $r_0 - e$ in Figure B.8. For completeness the graph of r_* is depicted in Figure B.9.

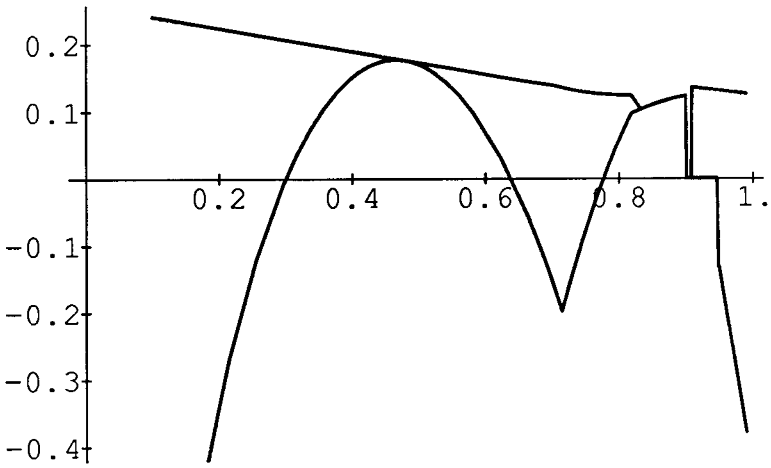

Figure B.8: Graph of ℓ_* and $r_0 - e$.

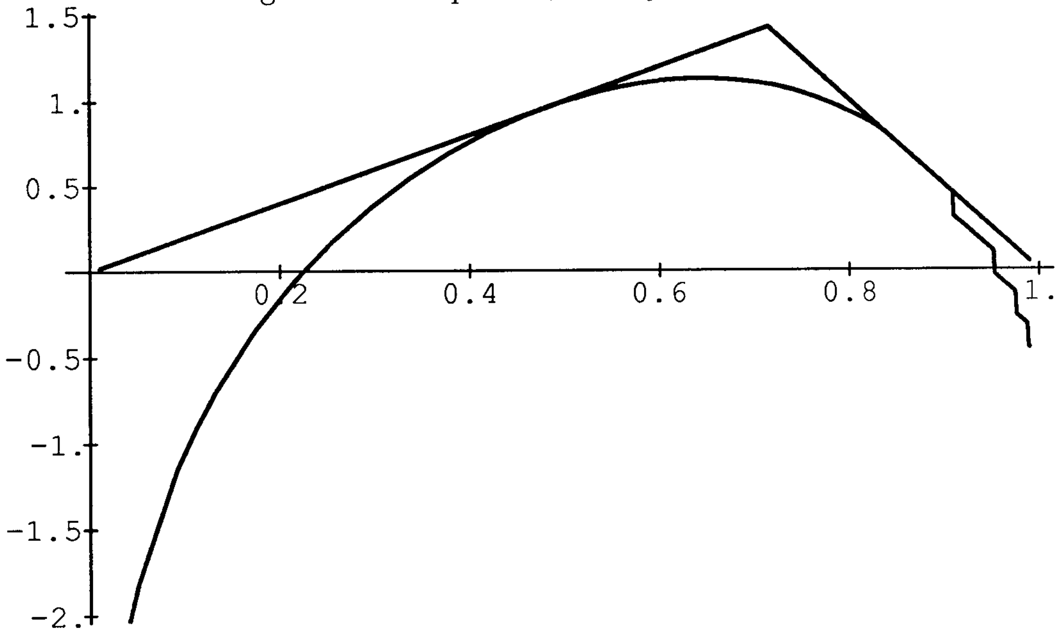

Figure B.9: Graph of e and r_* with $a_* \approx 0.4750$ and $b_* \approx 0.8325$.

Appendix C

Auxiliary Results

The purpose of this appendix is to set down some of the formulas used in Chapter 4 and Appendix B and to give arguments for the various existence, uniqueness and continuity properties of which we have had occasion to make use. To get things off the ground suppose that $X : (0,1) \mapsto \Re$ is given both continuous and monotone such that

$$X(\pi) = -X(1-\pi) \qquad \forall\, \pi \in (0,1). \tag{C.1}$$

Then if we define γ via,

$$\gamma(a,b) := X(b) - X(a) \qquad 0 < a, b < 1, \tag{C.2}$$

this implies that γ is a continuous mapping of the (open) unit square such that $\gamma(a,\cdot) : (0,1) \mapsto \Re$ is nondecreasing for $a \in (0,1)$ fixed, and for which $\gamma(\cdot,b) : (0,1) \mapsto \Re$ is nonincreasing for $b \in (0,1)$ fixed. In addition it implies that,

$$\gamma(a,b) = \gamma(1-b, 1-a) \qquad \forall\, a, b \in (0,1). \tag{C.3}$$

Next consider the *trick* floor function $\underline{\lfloor} \cdot \underline{\rfloor} : \Re \mapsto \mathcal{Z}$ defined as,

$$\underline{\lfloor} x \underline{\rfloor} := \begin{cases} \lfloor x \rfloor & \text{if } \lfloor x \rfloor < x; \\ \lfloor x \rfloor - 1 & \text{if } \lfloor x \rfloor = x, \end{cases} \tag{C.4}$$

where $\lfloor \cdot \rfloor$ denotes the usual floor function. The adjective *trick* is suggestive of an obvious extension to Knuth's picturesque appellation *floor function* namely, a *floor* function with a '*trapdoor*'. Another way to view $\underline{\lfloor} \cdot \underline{\rfloor}$ is as a left-continuous version of the right-continuous floor function. Having defined these objects we go on to define,

$$\mathcal{H}(a, b\,; \alpha) := \sum_{n=0}^{\underline{\lfloor}\gamma(a,b)\underline{\rfloor}} \frac{(-1)^n}{n!} \alpha^n \left[\gamma(a, b) - n\right]^n \qquad 0 < a, b < 1, \tag{C.5}$$

for $\alpha > 0$ and,

$$\mathcal{D}(a, b\,; \alpha, \beta) := \sum_{n=0}^{\underline{\lfloor}\gamma(a,b)\underline{\rfloor}-1} \alpha^n \sum_{m=0}^{n} \frac{(-1)^m}{m!} \beta^m \left[\gamma(a, b) - n - 1\right]^n \qquad 0 < a, b < 1, \tag{C.6}$$

for both α and β positive.

With these quantities in hand we can now give a complete and concise definition of the solution family $\{ r(\cdot\,; a, b) : 0 < a < b < 1 \}$ for the Stefan problem arising in the Poisson process case. This description can be found with minor notation changes in [MS] and [B&M]. In particular for all a, b such that $0 < a < b < 1$ define,

$$r(\pi; a, b) := e_1(\pi) + d(\pi; b) + \overline{D}(\pi; b) + K(a, b)\,\overline{H}(\pi; b) \quad 0 < \pi < 1, \tag{C.7}$$

with e_1, d, $\overline{D}$, K, and $\overline{H}$ defined as follows. The function d is given by,

$$d(\pi; b) := C\,\left(\lambda^1(1 - \pi) + \lambda^0\right)(N_b(\pi) + 1) \quad 0 < \pi, b < 1, \tag{C.8}$$

where $N_b(\pi) := \lfloor \gamma(\pi, b) \rfloor$ and (see C.1 and C.2),

$$X(\pi) := \log \left[\frac{\pi}{1-\pi}\right] / \log \frac{\lambda^1}{\lambda^0}, \quad C = c\,\frac{\lambda^1 - \lambda^0}{\lambda^0 \lambda^1}, \tag{C.9}$$

for $c > 0$ and $0 < \lambda^0 < \lambda^1$. The function e_1 is given by

$$e_1(\pi) = c^1\,(1-\pi) \qquad \forall\,\pi \in [0,1],\; c^1 > 0. \tag{C.10}$$

As for $\overline{H}(\pi; b)$, it equals,

$$\overline{H}(\pi; b) := \lambda^1\,(1-\pi)\,H_0(\pi; b) + \lambda^0\,\pi\,H_1(\pi; b) \quad 0 < \pi, b < 1, \tag{C.11}$$

where refering to C.5 we write,

$$H_i(\pi; b) := e^{-\nu_i\,X(\pi)} \mathcal{H}(\pi, b\,; \nu_i\, e^{-\nu_i}) \quad i = 0, 1, \tag{C.12}$$

with,

$$\nu_i := \frac{\lambda^i \log \frac{\lambda^1}{\lambda^0}}{\lambda^1 - \lambda^0} \quad i = 0, 1. \tag{C.13}$$

Refering to C.6 we define,

$$D_i(\pi; b) := -C\, e^{\nu_i\,(X(b) - X(\pi) - 1)} \mathcal{D}(\pi, b\,; e^{-\nu_i}, \nu_i) \quad i = 0, 1, \tag{C.14}$$

and then $\overline{D}$ is expressible as,

$$\overline{D}(\pi; b) := \lambda^1\,(1-\pi)\,D_0(\pi; b) + \lambda^0\,\pi\,D_1(\pi; b) \quad 0 < \pi, b < 1. \tag{C.15}$$

Finally, K is defined via,

$$K(a, b) := \frac{e_0(\pi) - [e_1(\pi) + d(a; b) + \overline{D}(a; b)]}{\overline{H}(a; b)} \quad 0 < a < b < 1, \tag{C.16}$$

with e_0 given by,

$$e_0(\pi) = c^0\,\pi \qquad \forall\,\pi \in [0,1],\; c^0 > 0. \tag{C.17}$$

Note that the risk 'triangle' e equals $e = e^0 \wedge e^1$, the minimum of e^0 and e^1. In all of these expressions we employ the conventions that an empty sum is zero and $0^0 = 1$.

Writing $r(\pi; a, b)$ more simply as $r(\pi)$ and letting r' denote the derivative taken from the right, the reader may show that,

$$r(a) = e_0(a) \qquad \forall\, a \in (0, b),\ 0 < b < 1. \tag{C.18}$$

If one defines the *advance* operator $\mathcal{A}_u$ by

$$\mathcal{A}_u\, \pi := \frac{u+1}{u+\pi}\, \pi \qquad 0 \leq \pi \leq 1,\ u > 0, \tag{C.19}$$

then one can also show,

$$r(\pi) = e_1(\pi) \qquad \forall\, \pi \in [b, \mathcal{A}_u b),\ 0 < b < 1. \tag{C.20}$$

It is called an advance operator since $\mathcal{A}_u\, b > b$ for all $b \in (0,1)$. For our purposes we shall take u as,

$$u = \frac{\lambda^0}{\lambda^1 - \lambda^0}. \tag{C.21}$$

Finally one can show that,

$$-\,\pi(1-\pi)r'(\pi) + (u+\pi)\,[r(\mathcal{A}_u \pi) - r(\pi)] = -c \quad \forall\, \pi \in (0,1), \tag{C.22}$$

i.e., r as defined above satisfies a functional-differential equation of the advance type. In fact it is shown in [MS] that r is the unique solution to this problem which is continuous on $(0, b)$ for any $b \in (0,1)$ and $c \geq 0$. The continuity of this solution family is obtained by its construction but it can also be demonstrated directly by assessing the continuity of its constituent parts.

Indeed, one can show directly that the homogeneous part of the solution $\overline{H}(\cdot\,;b)$ is continuous on $(0,b)$ for $0 < b < 1$. Essentially such a demonstration reduces to the consideration of the continuity of $\mathcal{H}(\cdot\,,b\,;\alpha)$ on $(0,b)$ for $b \in (0,1)$ and $\alpha > 0$. It is not difficult to see that any loss of continuity of $\mathcal{H}$ can only occur at those a, b for which $\gamma(a,b)$ is an integer. It is precisely at these values however that the next term in the sum starts off at zero. Hence the homogeneous part is continuous. Analogously, demonstrating the continuity of the particular solution reduces to the consideration of $d(\cdot\,;b)$ and $\overline{D}(\cdot\,;b)$. Likewise, any discontinuous behavior is confined to those points at which $\gamma = 0$. Examination of those points shows that both $d(\cdot\,;b)$ and $\overline{D}(\cdot\,;b)$ are discontinuous there but have simple jumps of equal magnitude and opposite sign. Hence, the sum of the two is continuous. To summarize the argument, we are guaranteed that $\overline{H}(\cdot\,;b)$ and $d(\cdot\,;b)+\overline{D}(\cdot\,;b)$ are continuous on $(0,b)$ for $0 < b < 1$. Now, by exploiting the symmetry inherent in the definition of γ as is manifest by expression C.3 it should come as no surprise that $\overline{H}(a;\cdot)$ and $d(a;\cdot) + \overline{D}(a;\cdot)$ are continuous on $(a,1)$ for $0 < a < 1$.

Next we turn our attention to the gap function G which we defined in Appendix B as,

$$G(\pi;a) := \lim_{x \to \pi^-} r(x;a,\pi) - e_1(\pi) \qquad 0 < a < \pi < 1. \tag{C.23}$$

From this definition it is not difficult to show that,

$$G(\pi;a) = \lambda(\pi) + K(a,\pi)H(\pi) \qquad 0 < a < \pi < 1, \tag{C.24}$$

where K is defined above, λ is given by,

$$\lambda(\pi) := \lambda^1(1-\pi) + \lambda^0\pi \qquad 0 < \pi < 1, \tag{C.25}$$

and H is given by

$$H(\pi) := \lambda^1 (1-\pi)\, e^{-\nu_0 X(\pi)} + \lambda^0\, \pi\, e^{-\nu_1 X(\pi)} \qquad 0 < \pi < 1. \tag{C.26}$$

From expression C.24 we see that the continuity of $G(\cdot\,,a)$ on $(a,1)$ for $a \in (0,1)$ given rests on the continuity of K under the same conditions. From our argument above we see that this continuity is guaranteed. Another implication which follows straightforwardly from C.24 is that

$$\lim_{\pi \to a^+} G(\pi; a) = e_0(a) - e_1(a), \tag{C.27}$$

so that $G(a^+; a) < 0$ whenever $a < \pi_e$; we remind the reader that π_e is the abscissa of the apex of the risk triangle. A third easily verified implication is that

$$\lim_{\pi \to 1} G(\pi; a) = \lambda^0 > 0, \tag{C.28}$$

for any $a \in (0,1)$. Combining these results we see that given any $a \in (0, \pi_e)$ there exists $b = b(a)$ in $(a,1)$ such that $G(b; a) = 0$.

A fact which we employed in Appendix B is that the crossover points $\underline{a}$ and $\overline{a}$ depend continuously on b. This can be shown in a general way by making the usual Lipshitz continuity arguments as in ODE theory which in our case reduces to the $C^1(0,1)$ properties of the mapping defining the derivative in the FDE (see [MS]). A direct proof follows easily from our continuity arguments above since for $\epsilon < b - a$ and $0 < \pi < b < 1$ we can write,

$$r(\pi; a, b\pm\epsilon) = e_1(\pi) + [\, d(\pi; b\pm\epsilon) + \overline{D}(\pi; b\pm\epsilon)\,] + K(a, b\pm\epsilon)\, \overline{H}(\pi; b\pm\epsilon). \tag{C.29}$$

We also claimed in Appendix B that r is strictly beneath e for all $\pi \geq b$. This follows quite easily since

$$r(\pi) = e_1(\pi) + \lambda(\pi)\,(N_b(\pi) + 1) \qquad \forall\, \pi \in [b, 1), \tag{C.30}$$

and as is evident from its definition,

$$N_b(\pi) \leq -1 \qquad \forall\, pi \in [b, 1). \tag{C.31}$$

The last bit of tidying up concerns the concavity properties of a given solution r to the FDE given at the begining of this appendix. This housekeeping is contained in the following proposition.

Proposition C.1 *Suppose that r is piecewise twice continuously differentiable and,*

$$D_u\, r(\pi) = -c \qquad \forall\, \pi \in (0, 1). \tag{C.32}$$

Then r is piecewise concave, the pieces defined by the breakpoints of its derivative.

Proof:

Let I denote an interval such that $r \in \mathcal{C}^2(\mathrm{I})$. The proof will proceed by a contrapositive argument. Since $r \in \mathcal{C}^2(\mathrm{I})$ then if r is not concave on I it must be true that there is some small open interval, say U, upon which r'' is strictly positive. With Π as in Section 4.5 define,

$$\tau^{\mathrm{U}} := \inf\{t > 0 : \Pi_t \notin \mathrm{U}\}, \tag{C.33}$$

and suppose $\pi \in \mathrm{U}$ so that $P_\pi\{\Pi_0 \in \mathrm{U}\} = 1$. Since r is strictly convex on U we have,

$$E_\pi[\, r\,(\Pi_{\tau^{\mathrm{U}}})\,] \geq r\,(E_\pi[\,\Pi_{\tau^{\mathrm{U}}}\,]) = r(\pi) \tag{C.34}$$

where the inequality is an application of Jensen's Inequality and the equality follows from the fact that Π is a uniformly integrable martingale. On the other hand, since r is supposed to satisfy the FDE condition for $c > 0$ we have,

$$D_u\, r(\pi) < 0. \tag{C.35}$$

From our previous results we have,

$$\begin{aligned} E_\pi[\, r\,(\Pi_{\tau^{\mathrm{U}}}) - r(\Pi_0)\,] &= E_\pi \int_0^{\tau^{\mathrm{U}}} D_u\, r(\Pi_s)\, ds + E_\pi\, J^r_{\tau^{\mathrm{U}}}(\Pi) \\ &= E_\pi \int_0^{\tau^{\mathrm{U}}} D_u\, r(\Pi_s)\, ds, \end{aligned} \tag{C.36}$$

so that C.35 and C.36 combine to yield,

$$E_\pi[\, r\,(\Pi_{\tau^{\mathrm{U}}}) - r(\Pi_0)\,] < 0. \tag{C.37}$$

From this last expression we obtain therefore,

$$E_\pi[\, r\,(\Pi_{\tau^{\mathrm{U}}})\,] < E_\pi[\, r(\Pi_0)\,] = r(\pi), \tag{C.38}$$

which is a direct contradiction to expression C.34. Hence the hypothesized interval U does not exist and r must be concave on I. Since this is true for any I upon which r is $\mathcal{C}^2$ the result is shown. □

The proposition gives us what we need since it was shown in [MS] that every solution to the FDE is piecewise twice continuously differentiable *a fortiori*. Moreover, employing the Corollary to Theorem 1.2.1 in [MS] implies that r_* is twice continuously differentiable on $(0, \mathcal{A}_u^{-1}\, b)$, once continuously differentiable on $(0, b)$ and continuous on $(0, \mathcal{A}_u\, b)$. These facts taken together yield r_* concave on $(0, \mathcal{A}_u\, b)$.

References

[B&M] Baras, J.S., and D.C. MacEnany (1987). 'Sequential Testing on the Rate of a Counting Process', Proc 26th IEEE Conf. on Dec. and Contr. 1592-1593.

[BECK] Beckenbach, E.F. (1937) 'Generalized Convex Functions', Bull. Amer. Math. Soc. 43

[B] Bremaud, P. (1981). *Point Processes and Queues. Martingale Dynamics.* Springer-Verlag, New York-Heidelberg-Berlin-Toyko.

[BILL] Billingsley, P. (1979) *Probability and Measure.* New York: Wiley.

[E] Elliot, R.J. (1982). *Stochastic Calculus and Applications* Springer-Verlag, New York-Heidelberg-Berlin-Toyko.

[L&S1] Liptser, R.S., and A.N. Shiryayev (1978). *Statistics of Random Processes I: Applications.* Springer-Verlag, New York-Heidelberg-Berlin-Toyko.

[MS] MacEnany, D.C. (1987). 'Bayesian Sequential Hypothesis Testing', Tech. Rep. SRC-TR87-107, Systems Research Center, Univ. of Maryland. Master's Thesis.

[K&P-K] Kazakos, D., and P. Papantoni-Kazakos (1980). Sequential Detection Between Poisson Processes. IEEE Transactions on Information Theory, IT-26, No.1 January, 1980.

[L] Lehmann, E.L. (1958). *Testing Statistical Hypotheses.* John Wiley and Sons, New York.

[MHAD] Davis, M.H.A. (1984). 'Piecewise-deterministic Markov Processes: A General Class of Non-diffusion Stochastic Models', Jour. Royal Stat. Soc. B 46, No. 3 353–388.

[S] Shiryayev, A.N. (1977). *Optimal Stopping Rules.* Springer-Verlag, New York-Heidelberg-Berlin-Toyko.

[Sn] Snyder, D.L. (1975). *Random Point Processes.* Wiley-Interscience, New York.

[W] Wald, A. (1947). *Sequential Analysis.* Dover Publications, Inc. New York.

[W&H] Wong, E., and B. Hajek (1985). *Stochastic Processes in Engineering Systems.* Springer-Verlag, New York-Heidelberg-Berlin.