AN IRREVERSIBLE INVESTMENT PROBLEM WITH A LEARNING-BY-DOING FEATURE

ERIK EKSTRÖM, YERKIN KITAPBAYEV, ALESSANDRO MILAZZO AND TOPIAS TOLONEN-WECKSTRÖM

Abstract. We study a model of irreversible investment for a decision-maker who has the possibility to gradually invest in a project with unknown value. In this setting, we introduce and explore a feature of "learning-by-doing", where the learning rate of the unknown project value is increasing in the decision-maker's level of investment in the project. We show that, under some conditions on the functional dependence of the learning rate on the level of investment (the "signal-to-noise ratio"), the optimal strategy is to invest gradually in the project so that a two-dimensional sufficient statistic reflects below a monotone boundary. Moreover, this boundary is characterised as the solution of a differential problem. Finally, we also formulate and solve a discrete version of the problem, which mirrors and complements the continuous version.

1. Introduction

Consider a decision-maker who aims to invest optimally in a new project, but who suffers from incomplete information and does not know the true value of the project. In the simplest Bayesian setting, we assume that the project value μ has a two-point prior distribution and can, thus, take two values: $\mu_0 < 0$ representing a "bad" project and $\mu_1 > 0$ representing a "good" project. In addition to the prior knowledge of the distribution of μ, the decision-maker has also access to a stream of noisy observations of the unknown project value μ, and can thus make further inference about its true value. Within this set-up with incomplete information, we consider a situation in which the actions of the decision-maker may affect the magnitude of the noise. More specifically, we introduce and study a notion of **learning-by-doing**: by investing more into the project (i.e., by **doing**) the decision-maker can reduce the magnitude of the noise in the observation process, thereby improving the **learning** rate of the true value of μ. The decision to increase the level of investment is irreversible, however, and there is thus a natural trade-off between early investment to increase the learning rate and a more cautious strategy to avoid investing in a potentially bad project.

We model the above situation with a learning-by-doing feature by introducing an observation process $X = (X_t)_{t\geq 0}$ of the form

$$\mathrm{d}X_t = \mu \mathrm{d}t + f(U_t)\mathrm{d}W_t, \tag{1}$$

where $W = (W_t)_{t\geq 0}$ is a standard Brownian motion, $U = (U_t)_{t\geq 0}$ is a non-decreasing control process with $0 \leq U \leq 1$ that describes the level of investment in the project, and $u \mapsto f(u)$ is a given positive and decreasing function (for a more precise description of f and the set of admissible control processes, see Sections 2-3 below). That is, the decision-maker observes the process X and may choose to increase the level of investment U at any time, thereby reducing the noise in the observation process and obtaining a

Date: June 24, 2024.

Key words and phrases. Irreversible investment; incomplete information; learning-by-doing; singular control; free-boundary problems.

Mathematics Subject Classification 2020: 93E11, 93E20, 60J60.

better estimate of the true value μ of the project. In this setting, the objective for the decision-maker is to choose a non-decreasing control U to maximize the expectation

$$\mathbb{E}\left[\int_0^\infty e^{-rt}\mu\,\mathrm{d}U_t\right] \tag{2}$$

of the true value of accumulated discounted future investments. Note that the optimization of (2) over investment strategies, subject to the learning-by-doing feature as described in (1), results in an intrinsic "cost of learning": the decision-maker naturally wants to improve her learning rate of μ by increasing the control U, but, by doing so, she may be investing in a bad project (with $\mu = \mu_0 < 0$), thereby collecting a negative pay-off. Our problem formulation of irreversible investment with learning-by-doing thus describes an instance of the classical theme of **exploration** vs. **exploitation**.

Intuitively, the set-up with the feature of learning-by-doing described above can be motivated as follows. An "outsider" (an agent who is not invested at all) may have access to noisy observations of the true value μ of a certain project, for example by observing financial statements of companies that are currently operating in a similar line of business. With no – or little – involvement in the project, however, the outsider has to rely on publicly available information, and observations of the project value are rather noisy. On the other hand, with a greater involvement in the project, as measured by the decision-maker's investment level, additional private information becomes available and more precise inference of the project value can be obtained.

For instance, an improved learning rate is a natural ingredient in situations allowing for **project expansion** (see Example 1 for a specific formulation). As an example, consider the renewable energy sector, where a firm invested in wind turbines or solar cell plants may encounter uncertainties related to factors such as weather patterns and energy output, as well as wear and tear and maintenance cost. Consequently, the firm has only access to a noisy stream of observations of the true project value. By gradual project expansion, however, some of the uncertainty factors are observed with more precision thanks to a higher experimentation rate, which enables the firm to better assess the viability of future expansion. A second example involves the launch of a new product or an existing product into a new market, where the true project value is not known due to uncertainties in, for example, production costs and demand. Again, project expansion gives rise to a higher experimentation rate, which leads to less noisy observations of the project value.

The learning-by-doing mechanism may also be associated with the level of **commitment**. For example, consider an agent who may invest in a new start-up, whose profitability is uncertain. By increasing the investment level, the agent shows commitment to the start-up and may to a larger extent gain access to board meetings and other events where more information is revealed, thereby reducing the level of noise in the observations of the project value.

1.1. **Related literature.** Problems of irreversible investment have been widely studied in the literature on stochastic control, with early references provided by [9] and [19], and more recent contributions including, among many others, [2], [6], [8] and [12]. Mathematically, the irreversible investment problem described in (1)-(2) (and formulated more precisely in Section 3 below) is a stochastic singular control problem under incomplete information, where the chosen control affects the learning rate. While stochastic control problems with incomplete information have been studied extensively (for early references, see [17] for a problem of utility maximization, and [7] for an investment timing decision), works involving control of the learning rate are more rare. Within statistics, a problem of change-point detection with a controllable learning rate has been studied in [4] (for reversible controls) and [11] (for irreversible controls), and an estimation problem with controllable learning rate was solved in [10]. In literature on operations management, related questions of the trade-off between earning and learning have been studied in the context of dynamic pricing in models with demand uncertainty, see, e.g., [15]. Within operations research, [16] studies an investment problem similar to ours, but with the main difference that investment does not affect the learning rate. More specifically, an investor may at each instant in time choose between a finite set of learning rates, where a larger learning rate comes with a larger running cost of observation, and where the unknown return has a two-point distribution. Moreover, the investor may choose an investment time, at which the optimization ends. For a related work, see also [23]. Finally, our problem is also related to classical multi-armed bandit problems, where a chosen strategy affects both learning and earning; see e.g. [14].

We also remark that, with respect to the existing literature of stochastic control problems where the control affects the learning rate (see, e.g., [4], [10] and [11]), we do not fix any specific form of the dependence of the learning rate on the control (the so-called "signal-to-noise ratio" in our problem). Instead, we develop and study a more flexible set-up by allowing for an arbitrary signal-to-noise ratio and providing sufficient conditions that guarantee the existence of a solution, which we can explicitly describe.

Since reversible controls are considered in [4], [10] and [16], the sufficient statistic in those studies consists merely of the conditional probability of one of the two possible states, and is thus one-dimensional. On the other hand, for irreversible controls (as in [11], and in the current paper), the sufficient statistic consists of the conditional probability of one of the states together with the current value of the control, and is thus two-dimensional.

The epithet "learning-by-doing" has been associated to various problems in the economics literature, mainly in settings where an experienced agent has a larger ability than a less experienced one; for a classical reference, see [1]. One may note that the notion of "learning-by-doing" as used in [1] could alternatively be described as "improving-by-doing", whereas the notion of the present paper could alternatively be referred to as "learning-faster-by-doing". Indeed, in [1] the *profitability* rate is larger for an experienced agent. In contrast, for us the investment level does not influence the project value μ, but it affects instead the rate with which the project value is revealed to the agent.

1.2. **Preview.** The remainder of the paper is organized as follows. In Section 2 we discuss a few aspects of the irreversible investment problem under consideration. Section 3 offers a precise mathematical formulation of the problem, using a weak approach. In Section 4 we provide heuristic reasoning to derive an ordinary differential equation (ODE) for a boundary, along which a candidate optimal strategy reflects the underlying sufficient statistic, and Section 5 discusses conditions under which the solution of the ODE is monotone increasing. In Section 6 we provide the Verification theorem, which guarantees that the obtained candidate strategy is indeed optimal. In Section 7 we provide some specific examples for our model with the corresponding illustrations of the optimal boundary. Finally, in Section 8 we analyze a discrete version of our problem, which corresponds to situations where the set of possible investment levels is discrete.

2. Some initial considerations

This section discusses a few aspects of the problem formulation in (1)-(2). It also serves as a bridge between the learning-by-doing problem described in Section 1 and its rigorous mathematical formulation, which is provided in Section 3.

2.1. **Investment costs.** In the general formulation (2) above, no investment costs are included. This, however, is without loss of generality. Indeed, if a constant investment cost $C > 0$ is included in the model, then the expected discounted profits would be

$$\mathbb{E}\left[\int_0^\infty e^{-rt}(\mu - C)\,\mathrm{d}U_t)\right]. \tag{3}$$

Consequently, the optimization over controls U of the expression in (3) is of the same type as in (2), but with μ replaced by $\tilde{\mu} := \mu - C$ (for a non-degenerate problem one then needs $\mu_0 < C < \mu_1$).

2.2. **Signal-to-noise ratio.** Recall the observation process X defined in (1). It is clear that, given an investment strategy U, it is equivalent to observe either X or $\tilde{X}$ where (recall that f is positive)

$$\mathrm{d}\tilde{X}_t := \frac{\mathrm{d}X_t - \mu_0 \mathrm{d}t}{f(U_t)}.$$

Then,

$$\begin{aligned}\mathrm{d}\tilde{X}_t &= \frac{\mu - \mu_0}{f(U_t)}\mathrm{d}t + \mathrm{d}W_t \\ &= \theta\rho(U_t)\mathrm{d}t + \mathrm{d}W_t, \end{aligned} \tag{4}$$

where

$$\theta := \frac{\mu - \mu_0}{\mu_1 - \mu_0} = \begin{cases} 1 & \text{if } \mu = \mu_1 \\ 0 & \text{if } \mu = \mu_0 \end{cases} \tag{5}$$

and

$$\rho(u) := \frac{\mu_1 - \mu_0}{f(u)}$$

is the *signal-to-noise ratio* of the problem. In the remainder of the article, we will often refer to the signal-to-noise ratio function $\rho(\cdot)$ rather than to the noise function $f(\cdot)$ used above.

2.3. **Admissible controls.** In our problem formulation above, it is implicitly understood that the control U should be chosen based on available observations of the process X. On the other hand, the choice of a control U affects the observation process X, cf. (1) or (4). Because of this cumbersome interplay, special care is needed when describing the set of admissible controls.

Problems of this type are well-suited for the "weak approach" based on change of measures and the Girsanov theorem. Such an approach is provided in Section 3.

3. Problem formulation

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a complete probability space, supporting a standard Brownian motion X and an independent Bernoulli random variable θ with

$$\mathbb{P}(\theta = 1) = \pi = 1 - \mathbb{P}(\theta = 0), \qquad \pi \in (0,1).$$

Let $\mathbb{F} = (\mathcal{F}_t)_{t\geq 0}$ be the smallest right-continuous filtration to which the process X is adapted, and $\mathbb{G} = (\mathcal{G}_t)_{t\geq 0}$ the smallest right-continuous filtration to which the pair (X, θ) is adapted. Denote by $\mathcal{A}$ the collection of $\mathbb{F}$-adapted, right-continuous, non-decreasing processes with values in $[0,1]$; for $u \in [0,1]$, denote by $\mathcal{A}_u$ the sub-collection of controls with initial value equal to u, i.e.,

$$\mathcal{A}_u = \{U \in \mathcal{A} : U_{0-} = u\}. \tag{6}$$

Let $\rho : [0,1] \to (0,\infty)$ be a given non-decreasing and bounded function. Then, for any $U \in \mathcal{A}$ and $t \in [0,\infty)$, we can define a measure $\mathbb{P}_t^U \sim \mathbb{P}$ on $(\Omega, \mathcal{G}_t)$ by

$$\frac{\mathrm{d}\mathbb{P}_t^U}{\mathrm{d}\mathbb{P}} := \exp\left\{\theta \int_0^t \rho(U_s)\,\mathrm{d}X_s - \frac{\theta}{2}\int_0^t \rho^2(U_s)\,\mathrm{d}s\right\} =: \eta_t^U.$$

Setting $\mathcal{G}_\infty := \sigma(\cup_{0\le t<\infty}\mathcal{G}_t)$, we may assume the existence of a probability measure $\mathbb{P}^U$ on $(\Omega, \mathcal{G}_\infty)$ that coincides with $\mathbb{P}^U_t$ on $\mathcal{G}_t$ (this can be guaranteed, e.g., by the theory of the so called Föllmer measure, cf. [13]). By the Girsanov theorem,

$$X_t = \theta \int_0^t \rho(U_s)\,\mathrm{d}s + W^U_t,$$

where W^U is a $(\mathbb{P}^U, \mathbb{G})$-Brownian motion. Note that this coincides with the dynamics of the observation process described in (4), and that ρ is the signal-to-noise ratio of the problem.

It should be noticed that the law of θ remains the same under $\mathbb{P}^U$ as under $\mathbb{P}$. Indeed, denoting by $\mathbb{E}^U$ the expectation under $\mathbb{P}^U$, we have

$$\mathbb{P}^U(\theta = 1) = \mathbb{E}^U[\mathbb{1}_{\{\theta=1\}}\eta^U_0] = \mathbb{E}[\mathbb{1}_{\{\theta=1\}}] = \mathbb{P}(\theta = 1) = \pi,$$

where the second equality follows from the fact that θ is $\mathcal{G}_0$-measurable.

For any $U \in \mathcal{A}$, define the adjusted belief process

$$\Pi^U_t := \mathbb{P}^U(\theta = 1|\mathcal{F}_t), \qquad t \in [0, \infty).$$

By the innovations approach to stochastic filtering, the so-called *innovations process*

$$\hat{W}^U_t := X_t - \int_0^t \rho(U_s)\Pi^U_s\,\mathrm{d}s$$

is a $(\mathbb{P}^U, \mathbb{F})$-Brownian motion, and (see, e.g., [18, Theorem 8.1])

$$\mathrm{d}\Pi^U_t = \rho(U_t)\Pi^U_t(1 - \Pi^U_t)\,\mathrm{d}\hat{W}^U_t. \tag{7}$$

Recall the original problem (2) that we want to solve (cf. also Section 2.2). We note that conditioning yields

$$\begin{aligned}\mathbb{E}^U\left[\int_0^\infty e^{-rt}\mu\,\mathrm{d}U_t\right] &= \mathbb{E}^U\left[\int_0^\infty e^{-rt}\mathbb{E}^U\left[\mu\middle|\mathcal{F}_t\right]\mathrm{d}U_t\right]\\ &= (\mu_1-\mu_0)\mathbb{E}^U\left[\int_0^\infty e^{-rt}(\Pi^U_t - k)\,\mathrm{d}U_t\right],\end{aligned}$$

where $k := -\mu_0/(\mu_1 - \mu_0) \in (0, 1)$ and the integral is interpreted in the Riemann-Stieltjes sense over the interval $[0, \infty)$. Therefore, in order to solve our original problem (2), we define and study the value function

$$V(u, \pi) := \sup_{U\in\mathcal{A}_u} \mathbb{E}^U_\pi\left[\int_0^\infty e^{-rt}(\Pi^U_t - k)\,\mathrm{d}U_t\right], \qquad (u, \pi) \in [0, 1] \times (0, 1), \tag{8}$$

where $k \in (0, 1)$ and the sub-index indicates the prior probability that $\theta = 1$.

4. Construction of a candidate solution

In this section we use heuristic arguments to construct a candidate value function $\hat{V}$ and a candidate optimal strategy $\hat{U}$ for the problem (8). Conditions under which $\hat{U}$ is optimal are then provided in Section 6 below, along with the equality $V = \hat{V}$.

It is intuitively clear that one should increase an optimal control $\hat{U}$ only if $\Pi^{\hat{U}}$ is large enough. Inspired by standard results in singular control, we will construct $\hat{V}$ using the Ansatz that there exists an non-decreasing boundary $\pi \mapsto h(\pi)$ and, for any initial point $(u, \pi) \in [0, 1] \times (0, 1)$, an optimal control that satisfies[1]

$$\hat{U}_t = u \vee \sup_{0\le s\le t} h(\Pi^{\hat{U}}_s). \tag{9}$$

[1] The existence of a control $\hat{U}$ that satisfies equation (9) is left for now. Note that one cannot simply see (9) as a definition, since $\hat{U}$ appears on both sides of the equation; for a formal definition of $\hat{U}$, see (30) below.

That is, we postulate that the optimal investment $\hat{U}$ is gradually increased in such a way that the two-dimensional process $(\hat{U}, \Pi^{\hat{U}})$ reflects along the boundary h, with reflection in the u-direction (see Figure 1); if the initial point (u, π) satisfies $u < h(\pi)$ then the construction results in an initial jump in the control of size $\mathrm{d}\hat{U}_0 = h(\pi) - u$.

By the dynamic programming principle, one expects the process

$$M_t := e^{-rt}\hat{V}(\hat{U}_t, \Pi_t^{\hat{U}}) + \int_0^t e^{-rs}(\Pi_s^{\hat{U}} - k)\,\mathrm{d}\hat{U}_s$$

to be a $\mathbb{P}^{\hat{U}}$-martingale. This translates into the condition

$$\frac{\rho^2(u)}{2}\pi^2(1-\pi)^2\hat{V}_{\pi\pi} - r\hat{V} = 0$$

in the *no-action region*

$$\mathcal{C} := \{(u, \pi) : u > h(\pi)\}.$$

On the boundary $\partial\mathcal{C}$, martingality of M requires that $\hat{V}_u + \pi - k = 0$; moreover, optimality of the boundary is obtained if, in addition,

$$\hat{V}_{u\pi}(h(\pi), \pi) + 1 = 0.$$

The condition on the second mixed derivative of the value function is a recurring condition for two-dimensional singular control problems (see, e.g., [20], [11],[5]).

Denoting by $b := h^{-1}$ the inverse of h, we thus formulate the following free-boundary problem: find $(\hat{V}, b)$ such that

(10)
$$\begin{cases} \frac{\rho^2(u)}{2}\pi^2(1-\pi)^2\hat{V}_{\pi\pi} - r\hat{V} = 0 & \pi < b(u) \\ \hat{V}_u = k - \pi & \pi = b(u) \\ \hat{V}_{u\pi} = -1 & \pi = b(u) \\ \hat{V}(u, 0+) = 0, & \end{cases}$$

where the last condition corresponds to no further investment in the case when the project is known to be of the bad type.

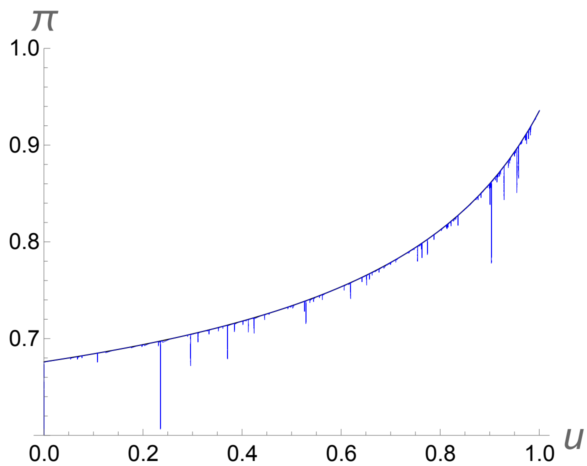

FIGURE 1. The trajectory of the pair (U, Π^U) under the reflecting strategy (9) in the case $\rho^2(u) = \frac{1}{4(1-0.9u)}$, $k = 0.5$ and $r = 0.1$.

4.1. **Deriving an ODE for the free boundary.** The general solution of the ODE in (10) is

$$\hat{V}(u,\pi) = A(u)(1-\pi)\left(\frac{\pi}{1-\pi}\right)^{\gamma(u)} + B(u)(1-\pi)\left(\frac{\pi}{1-\pi}\right)^{1-\gamma(u)},$$

where A and B are arbitrary functions and $\gamma(u) > 1$ is the unique positive solution of the quadratic equation

$$\gamma^2 - \gamma - \frac{2r}{\rho^2(u)} = 0. \tag{11}$$

More explicitly,

$$\gamma(u) = \frac{1}{2} + \sqrt{\frac{\rho^2(u) + 8r}{4\rho^2(u)}}. \tag{12}$$

Throughout Sections 4-7 we work under the following assumption.

Assumption 4.1. *The signal-to-noise ratio $\rho : [0,1] \to (0,\infty)$ is twice continuously differentiable, with $\rho'(u) > 0$ for every $u \in [0,1]$.*

Remark 4.2. It is immediate to check that Assumption 4.1 implies that $\gamma : [0,1] \to (1,\infty)$ is twice continuously differentiable, with $\gamma'(u) < 0$ for every $u \in [0,1]$.

In view of the boundary condition at $\pi = 0+$, we must have $B \equiv 0$ in the Ansatz above. Introducing the function

$$G(u,\pi) := (1-\pi)\left(\frac{\pi}{1-\pi}\right)^{\gamma(u)}, \tag{13}$$

our Ansatz then takes the form

$$\hat{V}(u,\pi) = A(u)G(u,\pi)$$

for $\pi \le b(u)$. The two conditions at the boundary (i.e., $\hat{V}_u = k - \pi$ and $\hat{V}_{u\pi} = -1$) then become

$$\begin{pmatrix} G_u(u,b(u)) & G(u,b(u)) \\ G_{u\pi}(u,b(u)) & G_\pi(u,b(u)) \end{pmatrix} \begin{pmatrix} A(u) \\ A'(u) \end{pmatrix} = \begin{pmatrix} k - b(u) \\ -1 \end{pmatrix},$$

which yields

$$\begin{pmatrix} A \\ A' \end{pmatrix} = \frac{1}{G_u G_\pi - G G_{u\pi}} \begin{pmatrix} G_\pi & -G \\ -G_{u\pi} & G_u \end{pmatrix} \begin{pmatrix} k - b \\ -1 \end{pmatrix}$$

(where the arguments of $A = A(u)$, $G = G(u, b(u))$, $b = b(u)$ and their derivatives are omitted). A straightforward calculation leads to

$$G_u G_\pi - G G_{u\pi} = \frac{-\gamma'}{b(1-b)} G^2,$$

so

$$\begin{pmatrix} A \\ A' \end{pmatrix} = \frac{b(1-b)}{\gamma' G^2} \begin{pmatrix} G_\pi & -G \\ -G_{u\pi} & G_u \end{pmatrix} \begin{pmatrix} b - k \\ 1 \end{pmatrix}. \tag{14}$$

Using

$$G_\pi = \frac{\gamma - b}{b(1-b)} G,$$

the first equation in (14) simplifies to

$$A = \frac{(\gamma + k - 1)b - \gamma k}{\gamma' G}. \tag{15}$$

Differentiation then gives

$$A' = \frac{\gamma'((\gamma+k-1)b' + \gamma'(b-k))G - ((\gamma+k-1)b - \gamma k)(\gamma'' G + \gamma' G_u + \gamma' G_\pi b')}{(\gamma')^2 G^2}.$$

Comparing the last equation with the second equation in (14), we find that

$$\begin{aligned}\gamma'((\gamma+k-1)b' + \gamma'(b-k))G - ((\gamma+k-1)b - \gamma k)(\gamma'' G + \gamma' G_u + \gamma' b' G_\pi)\\ = b(1-b)\gamma'(G_u - (b-k)G_{u\pi}),\end{aligned}$$

which simplifies to

$$b'(u) = F(u, b(u)), \tag{16}$$

where

$$F(u,b) := \frac{2(b-k)(\gamma')^2 + \big(\gamma k - (\gamma+k-1)b\big)\,\gamma''}{-\gamma(\gamma k - (\gamma+k-1)b) - (\gamma-1)(1-k)b} \cdot \frac{b(1-b)}{\gamma'}. \tag{17}$$

The ODE (16) is of first order, and we need to specify an appropriate boundary condition in order to find a unique candidate solution.

Since there is no possibility to improve learning when $u = 1$, we consider the corresponding irreversible investment problem with incomplete information *without* the learning-by-doing feature (i.e., in which the signal-to-noise ratio $u \mapsto \rho(u)$ is constant). It is intuitively clear that the two problems should coincide in the limit when $u \to 1$. We will thus use the obtained value of the boundary at $u = 1$ for the problem constant learning rates as the boundary condition for (16).

4.2. **Constant learning rates.** In this section, we study a simplistic version of our problem where the signal-to-noise ratio is constant (i.e., investing more does not provide an improvement in learning) and the decision-maker can only choose the time when to fully invest in the project. More precisely, for any fixed $u \in [0,1]$, consider the stopping problem

$$v(\pi) = v(\pi; u) := \sup_{\tau \geq 0} \mathbb{E}^u_\pi \left[e^{-r\tau}(\Pi^u_\tau - k) \right], \qquad \pi \in (0,1), \tag{18}$$

where the super-indices u denote that the signal-to-noise ratio $\rho(\cdot) \equiv \rho(u) \in (0, \infty)$ is constant, and the supremum is taken over $\mathbb{F}$-stopping times.

Remark 4.3. We have the relation $V(u,\pi) \geq (1-u)v(\pi;u)$, and the gap in the inequality represents the additional value that learning-by-doing provides in the problem formulation (8) compared to a case with a constant learning rate.

By standard methods of optimal stopping, one finds a candidate value function by solving the following free-boundary problem: construct $(\hat{v}, c)$ such that

$$\begin{cases} \frac{\rho^2(u)}{2}\pi^2(1-\pi)^2 \hat{v}_{\pi\pi} - r\hat{v} = 0, & \pi < c(u) \\ \hat{v} = \pi - k, & \pi = c(u) \\ \hat{v}_\pi = 1, & \pi = c(u) \\ \hat{v}(0+) = 0. \end{cases} \tag{19}$$

The general solution of the ODE in (19), combined with the boundary condition at $\pi = 0+$, is given by

$$\hat{v}(\pi) = D(u) G(u, \pi),$$

where G is as in (13) above and D is an arbitrary function. The two boundary conditions at $\pi = c(u)$ then yield

$$\begin{cases} D(u)G(u, c(u)) = c(u) - k \\ D(u)G_\pi(u, c(u)) = 1, \end{cases}$$

and using

$$G_\pi(u,\pi) = \frac{\gamma(u)-\pi}{\pi(1-\pi)} G(u,\pi)$$

we find that

$$c(u) = \frac{k\gamma(u)}{k+\gamma(u)-1}. \tag{20}$$

The candidate value function is thus given by

$$\hat{v}(\pi) = \begin{cases} \frac{c(u)-k}{G(u,c(u))} G(u,\pi), & \pi < c(u) \\ \pi - k, & \pi \geq c(u), \end{cases} \tag{21}$$

with c as in (20).

Since $\hat{v}$ is convex, we have $\hat{v} \geq \pi - k$. Moreover, $c \geq k$, which implies that

$$\frac{\rho^2(u)}{2}\pi^2(1-\pi)^2\hat{v}_{\pi\pi} - r\hat{v} \leq 0$$

for $\pi \neq c$. Using standard methods from optimal stopping theory (see, e.g., [21]) the verification of $\hat{v} = v$ is then straightforward and we omit the proof.

Proposition 4.4. *Let v be the value function defined as in (18), and let $\hat{v}$ be defined as in (21). Then $v = \hat{v}$, and moreover, the stopping time $\tau_c := \inf\{t \geq 0 : \Pi^u_t \geq c(u)\}$ is optimal for (18).*

5. A study of the ODE for the boundary

In this section we study the ODE (16). In particular, we first show that, when paired with its boundary condition derived in the previous section, it has a unique solution. Moreover, the heuristic derivation of (16) uses the assumption that the boundary b is monotone, so we also provide conditions under which the solution of the ODE is indeed monotone.

Consider the differential problem

$$\begin{cases} b'(u) = F(u, b(u)), & u \in (0,1) \\ b(1) = c(1), \end{cases} \tag{22}$$

where we recall from (17) that

$$F(u,b) = \frac{2(b-k)(\gamma')^2 + \big(\gamma k - (\gamma + k - 1)b\big)\,\gamma''}{-\gamma(\gamma k - (\gamma+k-1)b) - (\gamma-1)(1-k)b} \cdot \frac{b(1-b)}{\gamma'} \tag{23}$$

and where the boundary condition at $u = 1$ is given by

$$b(1) = c(1) = \frac{k\gamma(1)}{k+\gamma(1)-1} > k,$$

cf. (20).

Proposition 5.1. *The ODE (22) has a unique solution b on $[0,1]$. Moreover,*

$$0 < b(u) < c(u)$$

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

Proof. First note that the denominator of F is bounded away from 0 on the region

$$\mathcal{O} := \{(u,b) \in [0,1]^2 : b \leq c(u)\}.$$

Let $\tilde{F}$ be a modification of F which coincides with F on $\mathcal{O}$, and is Lipschitz continuous on $[0,1]\times\mathbb{R}$. It then follows from an application of the Picard-Lindelöf theorem the existence of a unique solution $\tilde{b}$ of

$$\begin{cases}\tilde{b}'(u)=\tilde{F}(u,\tilde{b}(u)), & u\in(0,1)\\ \tilde{b}(1)=c(1).\end{cases}$$

By straightforward differentiation,

$$c'(u)=\frac{-k(1-k)\gamma'}{(\gamma+k-1)^2}$$

and

$$F(u,c(u))=\frac{2(c-k)\gamma'(1-c)}{-(\gamma-1)(1-k)}=2c'(u)>c'(u).$$

Consequently, for every $u\in[0,1]$,

$$F(u,c(u))>c'(u)>0. \tag{24}$$

Therefore, $\tilde{b}(u)\le c(u)$ for all $u\in[0,1]$. Indeed, assuming that

$$u_0:=\sup\{u\in[0,1):\tilde{b}(u)=c(u)\}\ge 0,$$

we must have, by continuity,

$$F(u_0,c(u_0))=F(u_0,\tilde{b}(u_0))=\tilde{b}'(u_0)\le c'(u_0).$$

However, this contradicts (24), which proves that $\tilde{b}(u)<c(u)$ for all $u\in[0,1)$.

Similarly, $F(u,b)\le Db$ for some constant $D>0$, so by comparison we find that $\tilde{b}(u)\ge\tilde{b}(1)e^{-D(1-u)}>0$. Since $(u,\tilde{b}(u))\in\mathcal{O}$, and since $\tilde{F}\equiv F$ on $\mathcal{O}$, the result follows. □

We next study monotonicity properties of b. To do so, we need to investigate the sign of the function F in (23). Recall, from Proposition 5.1, that $0<b(u)\le c(u)$ for every $u\in[0,1]$. As a consequence,

$$-\gamma(\gamma k-(\gamma+k-1)b(u))-(\gamma-1)(1-k)b(u)<0$$

and so the sign of $b'(u)=F(u,b(u))$ coincides with the sign of the function

$$\begin{aligned}H(u,\pi)&:=2(\pi-k)(\gamma')^2+\big(\gamma k-(\gamma+k-1)\pi\big)\gamma''\\ &=\Big(2(\gamma')^2-\gamma''\gamma+(1-k)\gamma''\Big)\pi-\Big(2(\gamma')^2-\gamma''\gamma\Big)k\end{aligned}$$

evaluated at $(u,b(u))$. In particular, $b'(u)>0$ if and only if $H(u,b(u))>0$.

Note that H is affine in π, with $H(u,c(u))=2(c(u)-k)(\gamma')^2>0$ and $H(u,0)=-(2(\gamma')^2-\gamma''\gamma)k$. Consequently, if $H(u,0)\ge 0$, then b is automatically monotone increasing.

Proposition 5.2. *Assume that, for every $u\in[0,1]$,*

$$2(\gamma'(u))^2-\gamma(u)\gamma''(u)\le 0. \tag{25}$$

Then, the solution b of (22) satisfies $b'(u)>0$ for all $u\in[0,1]$.

Proof. If (25) holds, then $F(u,\pi)>0$ at all points (u,π) with $0<\pi<c(u)$. Consequently, $b'(u)>0$. □

Next, assume that $2(\gamma'(u))^2-\gamma(u)\gamma''(u)>0$ for every $u\in[0,1]$ and define

$$B(u):=\frac{2(\gamma'(u))^2-\gamma(u)\gamma''(u)}{2(\gamma'(u))^2-\gamma(u)\gamma''(u)+(1-k)\gamma''(u)}k,\qquad u\in[0,1]. \tag{26}$$

Then, $0<B(u)<c(u)$ and $H(u,B(u))=0$.

Proposition 5.3. *Assume that, for every* $u \in [0,1]$,

$$2(\gamma'(u))^2 - \gamma(u)\gamma''(u) > 0, \tag{27}$$

and $B'(u) > 0$. *Then, the solution* b *of* (22) *satisfies* $b'(u) > 0$ *and* $b(u) > B(u)$ *for all* $u \in [0,1]$.

Proof. Under the condition (27) we have $H(u,0) < 0$, and consequently $F(u,\pi) > 0$ if and only if $\pi > B(u)$. Therefore, it suffices to show that $b(u) > B(u)$.

Define

$$u_0 := \sup\{u \in [0,1] : b(u) = B(u)\}$$

and assume, to reach a contradiction, that $u_0 \geq 0$. Since $b(1) = c(1) > B(1)$, by continuity, we must have $u_0 < 1$. Moreover, by the definition of u_0, we must have $b'(u_0) \geq B'(u_0)$. However, this contradicts

$$b'(u_0) = F(u_0, b(u_0)) = 0 < B'(u_0).$$

Consequently, $b(u) > B(u)$ for all $u \in [0,1]$ and so $b'(u) > 0$. □

Remark 5.4. We note that condition (25) requires that γ is "sufficiently" convex, whereas condition (27) holds when γ is either concave or "mildly" convex. This depends, from (12), on the form of the signal-to-noise ratio.

Since it will be important to determinate whether $b(u) \geq k$ for every $u \in [0,1]$ (see Proposition 6.1 and Theorem 6.2 below), we enunciate the following corollary.

Corollary 5.5. *Assume that* γ *is concave and* $B' > 0$. *Then, the solution* b *of* (22) *satisfies* $b'(u) > 0$ *and* $b(u) > k$ *for all* $u \in [0,1]$.

Proof. The result directly follows from Proposition 5.3 and the fact that, if γ is concave, then $B(u) > k$ for every $u \in [0,1]$. □

To guarantee the monotonicity of B needed for Proposition 5.3 and Corollary 5.5, we have the following simple result.

Proposition 5.6. *Assume that* γ *is* $C^3([0,1])$, *that* $\gamma''(u) < 0$ *for* $u \in [0,1]$, *and that* γ *satisfies*

$$3(\gamma''(u))^2 < 2\gamma'(u)\gamma'''(u) \tag{28}$$

for all $u \in [0,1]$. *Then,* B *in* (26) *is strictly increasing, i.e.,* $B' > 0$.

Proof. From $H(u, B(u)) = 0$, we have

$$B'(u) = -\frac{H_u(u, B(u))}{H_\pi(u, B(u))}.$$

If γ is concave, then $H_\pi(u,\pi) > 0$ so it suffices to show that $H_u(u, B(u)) < 0$. Differentiation yields

$$\begin{aligned} H_u(u, B(u)) &= 3(B(u) - k)\gamma'\gamma'' + (\gamma k - (\gamma + k - 1)B(u))\gamma''' \\ &= 3(B(u) - k)\gamma'\gamma'' - 2(B(u) - k)\frac{(\gamma')^2\gamma'''}{\gamma''} \\ &= (B(u) - k)\frac{\gamma'}{\gamma''}(3(\gamma'')^2 - 2\gamma'\gamma''') < 0, \end{aligned}$$

where we used in the second equality that $H(u, B(u)) = 0$, and the inequality follows from (28) and the fact that $B \geq k$ since γ is concave. It follows that $B' > 0$. □

Remark 5.7. Figure 2 shows that the solution b to the ODE (22) is not always monotone.

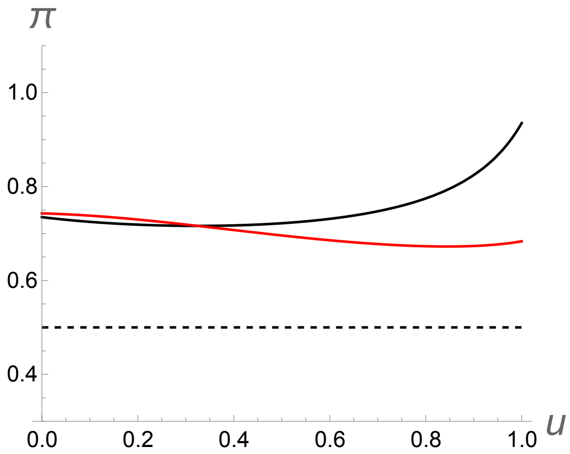

FIGURE 2. The solution b to the ODE (22) (solid black), the curve B (red) and the threshold k (dashed black), in the case $\rho^2(u) = \frac{1}{4(1-0.1u-0.8u^2)}$, $k = 0.5$ and $r = 0.1$.

6. VERIFICATION

We now formally construct the candidate optimal strategy, heuristically introduced in (9), that performs reflection along the boundary b (recall Figure 1). To do that, assume that the solution b of (22) is strictly increasing; sufficient conditions for this monotonicity were provided in Section 5 above. Recall that $h = b^{-1}$ denotes the inverse of b; it is defined on $[b(0), b(1)]$, and we extend it to $(0,1)$ so that $h(\pi) = 1$ for $\pi > b(1)$ and $h(\pi) = 0$ for $\pi < b(0)$.

For any fixed $(u,\pi) \in [0,1] \times (0,1)$, we define the candidate optimal strategy $\hat{U}$, to perform reflection along b, as follows. Denote by $C([0,\infty))$ the space of continuous functions from $[0,\infty)$ to $[0,1]$ and define the map $\tilde{U} : [0,\infty) \times C([0,\infty)) \to [0,1]$ by

$$\tilde{U}_t(\omega) := u \vee h\Big(\sup_{0\le s\le t} \omega_s\Big),$$

which will serve as the feed-back map of the optimal control. Now consider (cf. (7)) the stochastic differential equation (SDE)

$$\mathrm{d}P_t = -\rho^2(\tilde{U}_t(P))P_t^2(1-P_t)\,\mathrm{d}t + \rho(\tilde{U}_t(P))P_t(1-P_t)\,\mathrm{d}X_t, \tag{29}$$

with $P_0 = \pi$. The drift and diffusion coefficients of the SDE (29) satisfy the (locally) Lipschitz conditions of, e.g., [22, Ch. V, Th. 12.1] and thus the SDE (29) admits a unique strong solution $P = (P_t)_{t\ge 0}$. Then, define the candidate optimal control by

$$\hat{U}_{0-} = u \quad \text{and} \quad \hat{U}_t := \tilde{U}_t(P), \quad t \ge 0. \tag{30}$$

Since P is $\mathbb{F}$-adapted, we have that $\hat{U} \in \mathcal{A}_u$, as defined in (6). Recall that, by (7), we also have

$$\mathrm{d}\Pi_t^{\hat{U}} = -\rho^2(\hat{U}_t)(\Pi_t^{\hat{U}})^2(1-\Pi_t^{\hat{U}})\,\mathrm{d}t + \rho(\hat{U}_t)\Pi_t^{\hat{U}}(1-\Pi_t^{\hat{U}})\,\mathrm{d}X_t$$

i.e., $\Pi^{\hat{U}}$ satisfies (29) and so, by uniqueness, $\Pi^{\hat{U}}$ and P are indistinguishable. Notice that this also confirms our conjecture in (9).

Next, we define the candidate value function $\hat{V} : [0,1] \times (0,1) \to \mathbb{R}$ by

$$\hat{V}(u,\pi) := \begin{cases} A(u)G(u,\pi), & \pi \le b(u) \\ A(h(\pi))G(h(\pi),\pi) + (\pi - k)(h(\pi) - u), & \pi > b(u), \end{cases}$$

where

$$A(u) = \frac{(\gamma(u) + k - 1)b(u) - \gamma(u)k}{\gamma'(u)G(u, b(u))},$$

(cf. (15)) and G is as in (13). In this way, $\hat{V}$ is continuous. We now show some further properties it satisfies, which are essential to obtain the Verification theorem.

Proposition 6.1. *Assume that b is strictly increasing on $[0, 1]$. We have that*

$$\hat{V} \in C^{1,2}([0, 1] \times (0, 1)),$$

with $\hat{V}_u \le k - \pi$. Moreover, if $b(u) \ge k$ for every $u \in [0, 1]$ or if (25) holds, then

$$\frac{\rho^2}{2}\pi^2(1-\pi)^2\hat{V}_{\pi\pi} - r\hat{V} \le 0. \tag{31}$$

Proof. First, we study differentiability of $\hat{V}$. It is clear that $\hat{V}$ is of class C^1 below the boundary. Moreover, $\hat{V}_u(u, b(u)-) = k - b(u)$ by construction (recall (10)), and since $\hat{V}$ is extended linearly in u with slope $k - \pi$ for $\pi > b(u)$, it follows that $\hat{V}_u$ is continuous. More precisely, for (u, π) with $b(u) < \pi$, we have $\hat{V}(u, \pi) = \hat{V}(u_0, \pi) + (\pi - k)(u_0 - u)$ where $u_0 := h(\pi)$, and so

$$\hat{V}_\pi(u, \pi) = \hat{V}_\pi(u_0, \pi) + u_0 - u \tag{32}$$

since $\hat{V}_u(u_0, \pi) = k - \pi$. Thus, $\hat{V} \in C^1([0, 1] \times (0, 1))$.

We now check that $\hat{V}_u \le k - \pi$. We clearly have $\hat{V}_u = k - \pi$ above the boundary, so it remains to treat points below the boundary. For $\pi < b(u)$, we have

$$\hat{V}_u(u, \pi) = A'(u)G(u, \pi) + A(u)G_u(u, \pi)$$

and

$$\hat{V}_{u\pi}(u, \pi) = A'(u)G_\pi(u, \pi) + A(u)G_{u\pi}(u, \pi).$$

Since

$$G_\pi(u, \pi) = \frac{\gamma(u) - \pi}{\pi(1-\pi)}G(u, \pi) \quad \text{and} \quad G_{u\pi}(u, \pi) = \frac{\gamma(u) - \pi}{\pi(1-\pi)}G_u(u, \pi) + \frac{\gamma'(u)}{\pi(1-\pi)}G(u, \pi),$$

we obtain that

$$\hat{V}_{u\pi}(u, \pi) = \frac{\gamma(u) - \pi}{\pi(1-\pi)}\hat{V}_u(u, \pi) + \frac{\gamma'(u)}{\pi(1-\pi)}A(u)G(u, \pi). \tag{33}$$

Now assume, to reach a contradiction, that there exists $(u, \pi_0) \in [0, 1] \times (0, 1)$ with $\pi_0 < b(u)$ such that $\hat{V}_u(u, \pi_0) > k - \pi_0$. We then obtain from (33) and by (15) that

$$\begin{aligned}\hat{V}_{u\pi}(u, \pi_0) &> \frac{\gamma(u) - \pi_0}{\pi_0(1-\pi_0)}(k - \pi_0) + \frac{(\gamma(u) + k - 1)b(u) - k\gamma(u)}{\pi_0(1-\pi_0)}\frac{G(u, \pi_0)}{G(u, b(u))} \\ &\ge \frac{\gamma(u) - \pi_0}{\pi_0(1-\pi_0)}(k - \pi_0) + \frac{(\gamma(u) + k - 1)b(u) - k\gamma(u)}{\pi_0(1-\pi_0)} \\ &= \frac{(\gamma(u) + k - 1)(b(u) - \pi_0)}{\pi_0(1-\pi_0)} - 1 > -1.\end{aligned}$$

Consequently, $\pi \mapsto \hat{V}_u(u, \pi) + \pi - k$ is positive and increasing on $(\pi_0, b(u))$, which contradicts $\hat{V}_u(u, b(u)) = k - b(u)$. It follows that $\hat{V}_u \le k - \pi$ everywhere.

Differentiating (32) once more with respect to π and using $\hat{V}_{u\pi} = -1$ along the boundary (recall (10)), yields

$$\hat{V}_{\pi\pi}(u, \pi) = \hat{V}_{\pi\pi}(u_0, \pi),$$

which shows that $\hat{V}_{\pi\pi}$ is continuous.

Finally, (31) holds with equality below the boundary by construction, and above the boundary we have

$$\begin{aligned}&\frac{\rho^2(u)}{2}\pi^2(1-\pi)^2\hat{V}_{\pi\pi}(u,\pi) - r\hat{V}(u,\pi)\\&\quad = \frac{\rho^2(u)}{2}\pi^2(1-\pi)^2\hat{V}_{\pi\pi}(u_0,\pi) - r(\hat{V}(u_0,\pi) + (\pi-k)(u_0-u))\\&\quad = \left(\frac{\rho^2(u)}{\rho^2(u_0)} - 1\right) rA(u_0)G(u_0,\pi) - r(\pi-k)(u_0-u).\end{aligned}$$

Thus, if $\pi \geq b(u) \geq k$, then both terms are negative, and (31) follows.

Similarly, if (25) holds, using the expression (15) for A, we need to check that

$$\begin{aligned}&\left(\frac{\rho^2(u)}{\rho^2(u_0)} - 1\right)\frac{\gamma_0 k - (\gamma_0+k-1)\pi}{-\gamma_0'} + (\pi-k)(u-u_0)\\&= \left(\frac{\gamma_0^2-\gamma_0}{\gamma^2(u)-\gamma(u)} - 1\right)\frac{\gamma_0 k - (\gamma_0+k-1)\pi}{-\gamma_0'} + (\pi-k)(u-u_0) \leq 0,\end{aligned}$$

where $\gamma_0 := \gamma(u_0)$ and $\gamma_0' := \gamma'(u_0)$. Since $\gamma > \gamma_0$, we have

$$\frac{\gamma_0^2-\gamma_0}{\gamma^2-\gamma} \leq \frac{\gamma_0}{\gamma},$$

so it then suffices to show that

$$f(u) := \left(\frac{\gamma_0}{\gamma(u)} - 1\right)\frac{\gamma_0 k - (\gamma_0+k-1)\pi}{-\gamma_0'} + (\pi-k)(u-u_0) \leq 0.$$

However, it is clear that $f(u_0) = 0$; also,

$$f'(u_0) = \frac{(1-k)\pi}{\gamma_0} > 0$$

and f is concave by (25) and the fact that $\pi = b(u_0) \leq c(u_0)$. Consequently, $f(u) \leq 0$ for $u \leq u_0$, and (31) holds. □

The results of Proposition 6.1 lead to the Verification theorem, which we now present.

Theorem 6.2. *Let b be the solution of the differential problem (22). Assume that b is strictly increasing and either $b(u) \geq k$ for every $u \in [0,1]$ or (25) holds. Then, $V = \hat{V}$ and the strategy $\hat{U}$ is optimal in (8).*

Proof. Let $U \in \mathcal{A}_u$ be an arbitrary strategy and let

$$Y_t := Y_t^U := e^{-rt}\hat{V}(U_t,\Pi_t^U) + \int_0^t e^{-rs}\left(\Pi_s^U - k\right)\mathrm{d}U_s$$

for $t \geq 0-$. By Proposition 6.1, we can apply Itô's formula (for jump processes) to Y and obtain

$$\begin{aligned}\mathrm{d}Y_t &= e^{-rt}\left(\frac{1}{2}\rho^2(U_t)\Pi_t^2(1-\Pi_t)^2\hat{V}_{\pi\pi}(U_t,\Pi_t) - r\hat{V}(U_t,\Pi_t)\right)\mathrm{d}t + e^{-rt}(\Pi_t-k)\mathrm{d}U_t\\&\quad + e^{-rt}\hat{V}_u(U_t,\Pi_t)\,\mathrm{d}U_t^c + e^{-rt}\left(\hat{V}(U_t,\Pi_t) - \hat{V}(U_{t-},\Pi_t)\right) + e^{-rt}\hat{V}_\pi(U_t,\Pi_t)\,\mathrm{d}\Pi_t,\end{aligned}$$

where U^c denotes the continuous part of U and $\Pi := \Pi^U$. Here, the first term is non-positive by Proposition 6.1, and $\hat{V}_u \leq k-\pi$ gives that the next three ones are non-positive together. Moreover, (recall (7))

$$\int_0^t e^{-rs}\hat{V}_\pi(U_s,\Pi_s)\mathrm{d}\Pi_s = \int_0^t e^{-rs}\rho(U_s)\Pi_s(1-\Pi_s)\hat{V}_\pi(U_s,\Pi_s)\mathrm{d}\hat{W}_s^U$$

is a $(\mathbb{P}^U, \mathbb{F})$-martingale since $\hat{V}_\pi$ and ρ are bounded. Thus, the process Y is a $(\mathbb{P}^U, \mathbb{F})$-supermartingale on $[0-, \infty)$, and since Y is lower bounded it is also a $(\mathbb{P}^U, \mathbb{F})$-supermartingale on $[0-, \infty]$. It follows that

$$\hat{V}(u, \pi) = Y_{0-} \geq \mathbb{E}^U_\pi [Y_\infty] = \mathbb{E}^U_\pi \bigg[\int_0^\infty e^{-rt} (\Pi^U_t - k) \, \mathrm{d}U_t \bigg].$$

Since $U \in \mathcal{A}_u$ is arbitrary, it follows that $\hat{V} \geq V$.

To prove the reverse inequality, we consider the strategy $\hat{U}$ as in (30) (recall also (9) for an explicit form) and denote $\Pi := \Pi^{\hat{U}}$, so that $(\hat{U}, \Pi)$ always stays below the boundary b at all times t with $0 \leq t \leq \inf\{s \geq 0 : \hat{U}_s = 1\}$. Then, by construction,

$$\frac{1}{2}\rho^2(\hat{U}_t)\Pi_t^2(1 - \Pi_t)^2 \hat{V}_{\pi\pi}(\hat{U}_t, \Pi_t) - r\hat{V}(\hat{U}_t, \Pi_t) = 0$$

and

$$\hat{V}_u(\hat{U}_t, \Pi_t) \, \mathrm{d}\hat{U}^c_t = (k - \Pi_t) \, \mathrm{d}\hat{U}^c_t.$$

Moreover, at $t = 0$, if the initial point (u, π) satisfies $u < h(\pi)$, there occurs an initial and bounded jump in $\hat{U}$ of size $\mathrm{d}\hat{U}_0 = h(\pi) - u$, but no additional jumps occur. Since $\hat{V}(u, \pi) = \hat{V}(h(\pi), \pi) + (\pi - k)(h(\pi) - u)$ for $u < h(\pi)$, we have that

$$e^{-rt}(\Pi_t - k)\mathrm{d}\hat{U}_t + e^{-rt}\hat{V}_u(\hat{U}_t, \Pi_t) \, \mathrm{d}\hat{U}^c_t + e^{-rt} \Big(\hat{V}(\hat{U}_t, \Pi_t) - \hat{V}(\hat{U}_{t-}, \Pi_t) \Big) = 0.$$

Thus, by Itô's formula, the process $Y = Y^{\hat{U}}$ is a $(\mathbb{P}^{\hat{U}}, \mathbb{F})$-martingale. Since it is bounded, it is a $(\mathbb{P}^{\hat{U}}, \mathbb{F})$-martingale also on $[0-, \infty]$. It follows that

$$\hat{V}(u, \pi) = Y^{\hat{U}}_0 = \mathbb{E}^{\hat{U}}_\pi [Y_\infty] = \mathbb{E}^{\hat{U}}_\pi \bigg[\int_0^\infty e^{-rt} \, (\Pi_t - k) \, \mathrm{d}\hat{U}_t \bigg].$$

Consequently, $\hat{V} \leq V$.

Combining the two inequalities, it follows that $V \equiv \hat{V}$, and $\hat{U}$ is optimal in (8). $\square$

Remark 6.3. Theorem 6.2 shows that we can determine the solution to our problem when b is increasing and either $b(u) \geq k$ for every $u \in [0, 1]$ or (25) holds. Whether these conditions are satisfied depend on the form of the signal-to-noise ratio ρ and in Section 5 we have obtained some sufficient conditions that satisfy the hypotheses of Theorem 6.2 (recall, e.g., Corollary 5.5 and Proposition 5.6). In the next section we will provide some specific forms of ρ that fulfill the aforementioned conditions and, in particular, we will show that there are some choices of ρ under which (25) holds but $b(u) < k$ for every $u \in [0, u_0)$ and some $u_0 \in (0, 1)$ (see Figure 4).

7. Examples

In this section we provide a few examples for our model.

Example 1. (Project expansion). In this example we discuss a simplistic model for project expansion. To do that, assume that a decision-maker runs a business with unknown value μ and has access to noisy observations described by

$$\mathrm{d}X^0_t = \rho\theta \, \mathrm{d}t + \mathrm{d}W^0_t,$$

where $\rho > 0$ is a given constant, $\theta = (\mu - \mu_0)/(\mu_1 - \mu_0)$ (recall (5)) and W^0 is a Brownian motion. Moreover, assume that the decision-maker has the possibility to expand their activities by starting another identical business, but with independent noise. Thus, in addition to $\mathrm{d}X^0_t$, observations of

$$\mathrm{d}X^1_t = \rho\theta \, \mathrm{d}t + \mathrm{d}W^1_t$$

become available after expansion, where W^1 is a Brownian motion independent of W^0. Note that the drifts contain the same random factor θ, and thus the learning rate is larger after expansion. More specifically, observing $\mathrm{d}X_t^0$ and $\mathrm{d}X_t^1$ provides the same information as observing $\mathrm{d}X_t := \mathrm{d}X_t^0 + \mathrm{d}X_t^1 = 2\rho\theta\mathrm{d}t + \sqrt{2}\mathrm{d}W_t$, where W is a Brownian motion. Consequently, the signal-to-noise ratio increased from ρ to $\sqrt{2}\rho$ after expansion.

In a continuous setting, the above example generalizes to a signal-to-noise ratio $\rho(u) = C\sqrt{u}$. It is straightforward to check that (25) holds for this ρ, so Theorem 6.2 applies.

Example 2. (Linear noise function). For an illustration of the optimal reflecting boundary, we consider a special case in which $f^2(u)$ is linear, i.e., $f(u) = \sqrt{D_1 - D_2 u}$ with $0 < D_2 < D_1$. Then,

$$\rho^2(u) = \frac{C}{1 - Du}$$

for some constants $C > 0$ and $D \in (0,1)$. Differentiation of (11) yields

$$\gamma''(u) = \frac{-2(\gamma'(u))^2}{2\gamma(u) - 1} < 0.$$

Therefore, the ODE (22) becomes

$$\begin{cases} b' = \frac{(3\gamma+k-2)b-(3\gamma-1)k}{(\gamma+k-1)(\gamma-1)b-\gamma k(\gamma-b)} \cdot \frac{2\gamma' b(1-b)}{2\gamma-1}, & u \in (0,1) \\ b(1) = c(1). \end{cases}$$

We then have that the 0-level curve B of F is given by

$$B(u) = \frac{(3\gamma - 1)k}{3\gamma + k - 2} > k$$

and we note that B is monotone increasing with $B' > 0$. The existence of a monotone solution b to (22) thus follows by Corollary 5.5, and we have that $b(u) > k$ for every $u \in [0,1]$. Therefore, Theorem 6.2 applies and the optimality of the solution is verified. See Figure 3 for a plot of the optimal boundary b.

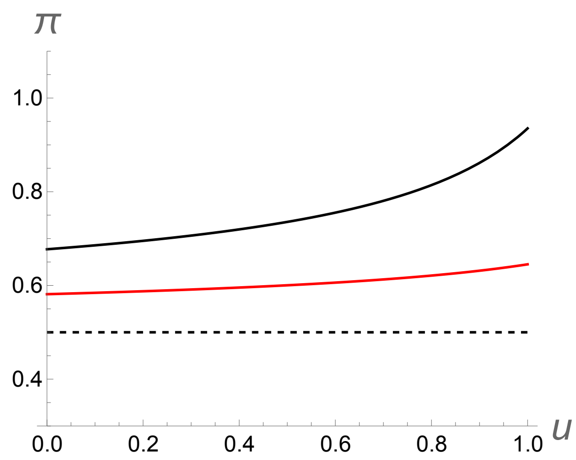

FIGURE 3. The optimal reflecting boundary b (solid black), the curve B (red) and the threshold k (dashed black) in the case $\rho^2(u) = \frac{1}{4(1-0.9u)}$, $k = 0.5$ and $r = 0.1$.

We next provide an example in which the optimal investment boundary b goes below the level k. Notice, from (8), that increasing the level of investment when Π is below k yields an instantaneous negative reward. The decision-maker should thus sometimes expand the project even though the current estimate of the project value is negative.

Example 3. $(b(u) < k)$**.** Consider the case

$$\gamma(u) = \frac{1.25}{u + 0.2}$$

for $u \in [0,1]$. It can be verified that γ satisfies (25). Hence $b'(u) > 0$ for all $u \in [0,1]$, and Theorem 6.2 applies. Figure 4 shows that the optimal boundary b goes below the threshold k for small values of u. This is remarkable since it corresponds to an instantaneous negative reward for such values (recall (8)). The explanation for this seemingly irrational behaviour is that the negative reward is compensated by a comparatively large value of future learning due to an increased learning rate.

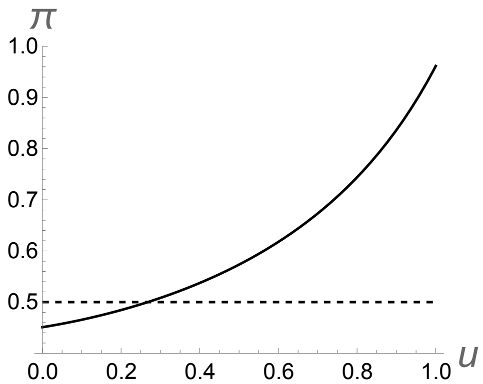

FIGURE 4. The optimal reflecting boundary b (solid black) and the threshold k (dashed black) in the case $\gamma(u) = 1.25/(u+0.2)$ and $k = 0.5$.

8. THE DISCRETE CASE

In this section we study a similar problem of irreversible investment under incomplete information and learning-by-doing, but where the control U is restricted to take values in a discrete subset of $[0,1]$. In this setting, we analyze under what conditions the related investment boundary is monotone in the number of remaining exercise rights, and we characterize the optimal strategy. The analysis of the discrete case mirrors, and complements, the continuous version of the problem of irreversible investment, as presented above.

To introduce the problem, let an integer $N \geq 0$ be given, and define $u_n = \frac{n}{N}$ for $n = 0, ..., N$. We study the following recursively defined problem:

$$\left\{ \begin{array}{l} V_N(\pi) = \sup_\tau \mathbb{E}_\pi^{u_N}\left[e^{-r\tau}\left(\Pi_\tau^{u_N} - k\right)\right], \\ V_n(\pi) = \sup_\tau \mathbb{E}_\pi^{u_n}\left[e^{-r\tau}\left(V_{n+1}(\Pi_\tau^{u_n}) + \Pi_\tau^{u_n} - k\right)\right], \qquad n = 0, ..., N-1. \end{array} \right. \tag{34}$$

Remark 8.1. The recursively defined optimization problem (34) is a discrete version of the continuous formulation (8). Indeed, if the set of admissible controls $\mathcal{A}$ is further restricted to take values only in $\{u_n\}_{n=0}^N$, then problem (8) reduces to a multiple stopping problem. By standard literature on Markovian multiple stopping problems (see, e.g., [3]), such problems can be formulated recursively as in (34).

As in the continuous case treated above, we first construct candidate solutions $\hat{V}_n(\pi)$, $n = 0, ..., N$, and we then verify that $\hat{V}_n = V_n$. The candidate solution is constructed using an Ansatz that there exists an increasing sequence $\{b_n\}_{n=0}^N$ such that

$$\tau_n := \inf\{t \geq 0 : \Pi_t^{u_n} \geq b_n\}$$

is optimal for V_n. The candidate solutions will be described using the notation

$$G_n(\pi) := G(u_n, \pi),$$

where G is as in (13) with $\gamma = \gamma_n := \gamma(u_n)$. We also let

$$c_n := c(u_n) = \frac{\gamma_n k}{\gamma_n + k - 1}.$$

8.1. **Solving the discrete problem.** First consider the last step, i.e., the stopping problem

$$V_N(\pi) = \sup_\tau \mathbb{E}^{u_N}_\pi \left[e^{-r\tau} \left(\Pi^{u_N}_\tau - k\right)\right].$$

This is a standard call option on the process Π^{u_N}, and was already treated in Subsection 4.2). In fact,

$$V_N(\pi) = \begin{cases} A_N G_N(\pi) & \pi < b_N \\ \pi - k & \pi \geq b_N, \end{cases}$$

where

$$b_N = \frac{\gamma_N k}{\gamma_N + k - 1} =: c_N$$

and A_N is a constant.

Next we treat the case $n = 0, \dots, N-1$ using induction. Assume that there are points $b_{n+1} \leq b_{n+2} \leq ... \leq b_N$ such that

$$V_m(\pi) = \begin{cases} A_m G_m(\pi) & \pi < b_m \\ \pi - k + V_{m+1}(\pi) & \pi \geq b_m, \end{cases} \tag{35}$$

for $m = n+1, ..., N$, where $V_{N+1} \equiv 0$. Also assume that

$$(\gamma_m + k - 1) b_m - \gamma_m k + (\gamma_m - \gamma_{m+1}) V_{m+1}(b_m) = 0. \tag{36}$$

Remark 8.2. Equation (36) holds for $m = N$ with $b_N = c_N$, and for $m \leq N-1$ it is a consequence of the smooth fit condition and the assumed monotonicity of the boundary. Indeed, the smooth fit condition at b_m gives the equation system

$$\begin{cases} A_m G_m(b_m) = b_m - k + V_{m+1}(b_m) \\ A_m G_m'(b_m) = 1 + V_{m+1}'(b_m), \end{cases}$$

which yields

$$(\gamma_m + k - 1) b_m - \gamma_m k = b_m(1 - b_m) V_{m+1}'(b_m) - (\gamma_m - b_m) V_{m+1}(b_m). \tag{37}$$

Now, if $b_m \leq b_{m+1}$, then

$$V_{m+1}'(b_m) = \frac{\gamma_{m+1} - b_m}{b_m(1 - b_m)} V_{m+1}(b_m),$$

and (37) reduces to (36).

We now provide conditions under which also V_n has the form

$$\hat{V}_n(\pi) = \begin{cases} A_n G_n(\pi) & \pi < b_n \\ \pi - k + V_{n+1}(\pi) & \pi \geq b_n, \end{cases} \tag{38}$$

where the boundary point b_n satisfies (39) below (which is (36) with $m = n$). To do that, first note that if there exists a boundary point b_n as in (38), then the smooth fit condition reads

$$\begin{cases} A_n G_n(b_n) = b_n - k + V_{n+1}(b_n) \\ A_n G_n'(b_n) = 1 + V_{n+1}'(b_n), \end{cases}$$

which reduces (as in Remark 8.2) to

$$(\gamma_n + k - 1) b_n - \gamma_n k + (\gamma_n - \gamma_{n+1}) V_{n+1}(b_n) = 0, \tag{39}$$

provided $b_n \le b_{n+1}$. Denote

$$f_n(b) := (\gamma_n + k - 1)b - \gamma_n k + (\gamma_n - \gamma_{n+1})V_{n+1}(b), \tag{40}$$

and let similarly

$$f_m(b) := (\gamma_m + k - 1)b - \gamma_m k + (\gamma_m - \gamma_{m+1})V_{m+1}(b)$$

for $m = n+1, ..., N$ so that $f_m(b_m) = 0$. Note that f_n is convex, with $f_n(0) = -\gamma_n k < 0$ and

$$f_n(c_n) = (\gamma_n - \gamma_{n+1})V_{n+1}(c_n) > 0.$$

Consequently, there exists a unique $b_n \in (0, c_n)$ such that $f_n(b_n) = 0$, which defines b_n.

Remark 8.3. We emphasize that the derived form of f_n (as given in (40)) uses that $b_n \le b_{n+1}$; in particular, if the solution b_n of $f_n(b_n) = 0$ satisfies $b_n > b_{n+1}$, then the smooth fit condition at b_n is not guaranteed. We also note that monotonicity of the boundary (i.e., $b_n \le b_{n+1}$) is equivalent to $f_n(b_{n+1}) \ge 0$.

Proposition 8.4. *Assume that $b_n \le b_{n+1}$. Then, $\hat{V}_n = V_n$.*

Proof. For the verification of $\hat{V}_n = V_n$, we first check that

$$\hat{V}_n(\pi) \ge V_{n+1}(\pi) + \pi - k. \tag{41}$$

Equation (41) holds automatically (with equality) for $\pi \ge b_n$. To see that it holds also below b_n, assume that $\hat{V}_n(\pi_0) < V_{n+1}(\pi_0) + \pi_0 - k$ for some $\pi_0 < b_n$. We then have that

$$\begin{aligned}\hat{V}_n'(\pi_0) - V_{n+1}'(\pi_0) - 1 &= \frac{\gamma_n - \pi_0}{\pi_0(1-\pi_0)}\hat{V}_n(\pi_0) - \frac{\gamma_{n+1} - \pi_0}{\pi_0(1-\pi_0)}V_{n+1}(\pi_0) - 1 \\ &< \frac{\gamma_n - \gamma_{n+1}}{\pi_0(1-\pi_0)}V_{n+1}(\pi_0) + \frac{(\gamma_n - \pi_0)(\pi_0 - k) - \pi_0(1-\pi_0)}{\pi_0(1-\pi_0)} \\ &= \frac{1}{\pi_0(1-\pi_0)}f_n(\pi_0) \le 0,\end{aligned}$$

where the last inequality follows from $\pi_0 \le b_n$. Thus, if $\hat{V}_n(\pi_0) < V_{n+1}(\pi_0) + \pi_0 - k$ at some point $\pi_0 < b_n$, then $\hat{V}_n(\pi) - V_{n+1}(\pi) - (\pi - k)$ is decreasing for $\pi \in [\pi_0, b_n]$, which contradicts the relation $\hat{V}_n(b_n) = V_{n+1}(b_n) + b_n - k$; consequently, (41) holds.

To complete a verification argument, we also need

$$\mathcal{L}_n\hat{V}_n := \frac{\rho_n^2}{2}\pi^2(1-\pi)^2\hat{V}_n'' - r\hat{V}_n \le 0 \tag{42}$$

for $\pi \notin \{b_n, b_{n+1}\}$. The inequality (42) holds automatically (with equality) for $\pi < b_n$, so we only need to check it above b_n. For $\pi > b_n$ we have $\hat{V}_n = V_{n+1} + \pi - k$, so

$$\begin{aligned}\mathcal{L}_n\hat{V}_n &= \tfrac{\rho_n^2}{2}\pi^2(1-\pi)^2V_{n+1}'' - rV_{n+1} - r(\pi - k) \\ &= \tfrac{\rho_n^2}{\rho_{n+1}^2}\mathcal{L}_{n+1}V_{n+1} - r\Big(1 - \tfrac{\rho_n^2}{\rho_{n+1}^2}\Big)V_{n+1} - r(\pi - k) \\ &\le -r\Big(1 - \tfrac{\rho_n^2}{\rho_{n+1}^2}\Big)V_{n+1} - r(\pi - k),\end{aligned}$$

provided $\pi \ne b_{n+1}$ (and where $\mathcal{L}_{n+1}V_{n+1} := \frac{\rho_{n+1}^2}{2}\pi^2(1-\pi)^2V_{n+1}'' - rV_{n+1} \le 0$). Since both V_{n+1} and $\pi \mapsto \pi - k$ are increasing, we note that

$$\begin{aligned}\mathcal{L}_n\hat{V}_n(\pi) &\le -r\bigg(\Big(1 - \tfrac{\rho_n^2}{\rho_{n+1}^2}\Big)V_{n+1}(b_n) + b_n - k\bigg) \\ &= -r\Big(\tfrac{\gamma_n + \gamma_{n+1} - 1}{\gamma_n^2 - \gamma_n}(\gamma_n - \gamma_{n+1})V_{n+1}(b_n) + b_n - k\Big).\end{aligned}$$

Using $f_n(b_n) = 0$, we have

$$(\gamma_n - \gamma_{n+1})V_{n+1}(b_n) = \gamma_n k - (\gamma_n + k - 1)b_n$$

and so

$$\mathcal{L}_n V_n(\pi) \le -r\left(\frac{\gamma_{n+1}(k-b_n)}{\gamma_n-1}+\frac{\gamma_n+\gamma_{n+1}-1}{\gamma_n^2-\gamma_n}(1-k)b_n\right)$$
$$\le -r\left(\frac{\gamma_{n+1}k-(\gamma_{n+1}+k-1)b_n}{\gamma_n-1}\right)\le 0,$$

since $b_n \le c_n \le c_{n+1}$. Consequently, (42) holds for all $\pi \notin \{b_n, b_{n+1}\}$. Using (41) and (42), a standard verification procedure shows that $\hat{V}_n \equiv V_n$. □

Proposition 8.4 completes the inductive construction and verification of the value function V_n. As remarked above, however, the construction depends on the assumption $b_n \le b_{n+1}$. In the next subsection we provide conditions under which the boundary is indeed monotone.

8.2. **Monotonicity of the boundary.** First note that

$$f_{N-1}(c_{N-1}) = (\gamma_{N-1}-\gamma_N)V_N(c_{N-1}) > 0,$$

so $b_{N-1}\in(0,c_{N-1})$. Since $c_{N-1}<c_N=b_N$, we automatically have $b_{N-1}\le b_N$.

Now, for $n\in\{0,...,N-2\}$, assume that $b_{n+1}\le ... \le b_N$ have been found such that (35) and (36) hold for $m=n+1,...,N$. We then have

$$V_{n+1}(b_{n+1}) = b_{n+1}-k+V_{n+2}(b_{n+1}), \tag{43}$$

and from $f_{n+1}(b_{n+1})=0$ we get

$$V_{n+2}(b_{n+1}) = \frac{\gamma_{n+1}k-(\gamma_{n+1}+k-1)b_{n+1}}{\gamma_{n+1}-\gamma_{n+2}}. \tag{44}$$

From the identities (43) and (44), we thus obtain

$$f_n(b_{n+1}) = 2(b_{n+1}-k)(\gamma_n-\gamma_{n+1})+\frac{(\gamma_{n+1}k-(\gamma_{n+1}+k-1)b_{n+1})(\gamma_n-2\gamma_{n+1}+\gamma_{n+2})}{\gamma_{n+1}-\gamma_{n+2}}.$$

Proposition 8.5. *If*

$$2(\gamma_n-\gamma_{n+1})(\gamma_{n+1}-\gamma_{n+2})-(\gamma_n-2\gamma_{n+1}+\gamma_{n+2})\gamma_{n+1}\le 0, \tag{45}$$

then $b_n\le b_{n+1}$.

Proof. Define

$$H_n(b) := 2(b-k)(\gamma_n-\gamma_{n+1})+\frac{(\gamma_{n+1}k-(\gamma_{n+1}+k-1)b)(\gamma_n-2\gamma_{n+1}+\gamma_{n+2})}{\gamma_{n+1}-\gamma_{n+2}}$$

so that $H_n(b_{n+1})=f_n(b_{n+1})$. Then

$$H_n(c_{n+1}) = 2(c_{n+1}-k)(\gamma_n-\gamma_{n+1})>0,$$

and by (45) we have

$$H_n(0) = -\frac{k}{\gamma_{n+1}-\gamma_{n+2}}\big((\gamma_n-\gamma_{n+1})(\gamma_{n+1}-\gamma_{n+2})-(\gamma_n-2\gamma_{n+1}+\gamma_{n+2})\gamma_{n+1}\big)\ge 0.$$

Since H_n is affine, it follows that $H_n(b)>0$ for all $b\in(0,c_{n+1})$. Thus $f_n(b_{n+1})\ge 0$, and the result follows. □

Remark 8.6. Note that the condition (45) is a discrete version of (25).

Figure 5 illustrates the set of the optimal boundaries b_n, $n=0,\dots,N$ for γ_n that is a discrete version of the specification used in Figure 4.

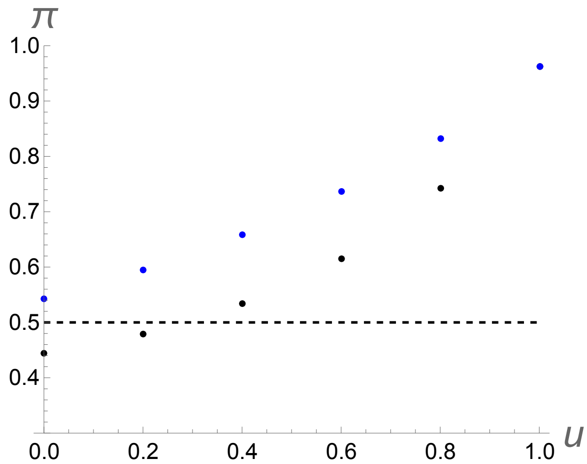

Figure 5. Optimal boundaries b_n (black dots), points c_n (blue dots), and the threshold k (black dashed) with $\gamma_n = 1.25/\left(n/5+0.2\right)$. Remaining parameters are $k = 0.5$ and $N = 5$. Note that γ_n satisfies condition (45).

References

[1] K. J. Arrow. The economic implications of learning by doing. *Rev. Econ. Stud.*, 29(3):155–173, 1962.

[2] P. Bank and D. Besslich. Modelling information flows by Meyer-σ-fields in the singular stochastic control problem of irreversible investment. *Ann. Appl. Probab.*, 30(6):2923–2962, 2020.

[3] R. Carmona and S. Dayanik. Optimal multiple stopping of linear diffusions. *Math. Oper. Res.*, 33(2):446–460, 2008.

[4] R. C. Dalang and A. N. Shiryaev. A quickest detection problem with an observation cost. *Ann. Appl. Probab.*, 25(3):1475–1512, 2015.

[5] T. De Angelis, E. Ekström, and M. Olofsson. The maximality principle in singular control with absorption and its applications to the dividend problem. *SIAM J. Control Optim.*, 62(1):91–117, 2024.

[6] T. De Angelis, S. Federico, and G. Ferrari. Optimal boundary surface for irreversible investment with stochastic costs. *Math. Oper. Res.*, 42(4):1135–1161, 2017.

[7] J.-P. Décamps, T. Mariotti, and S. Villeneuve. Investment timing under incomplete information. *Math. Oper. Res.*, 30(2):472–500, 2005.

[8] J.-P. Décamps, T. Mariotti, and S. Villeneuve. Irreversible investment in alternative projects. *Econom. Theory*, 28(2):425–448, 2006.

[9] A. K. Dixit and R. S. Pindyck. *Investment under Uncertainty.* Princeton University Press, 1994.

[10] E. Ekström and I. Karatzas. A sequential estimation problem with control and discretionary stopping. *Probab. Uncertain. Quant. Risk*, 7(3):151–168, 2022.

[11] E. Ekström and A. Milazzo. A detection problem with a monotone observation rate. *Stochastic Process. Appl.*, 172:Paper No. 104337, 19, 2024.

[12] G. Ferrari. On an integral equation for the free-boundary of stochastic, irreversible investment problems. *Ann. Appl. Probab.*, 25(1):150–176, 2015.

[13] H. Föllmer. The exit measure of a supermartingale. *Z. Wahrscheinlichkeitstheorie und Verw. Gebiete*, 21:154–166, 1972.

[14] J. Gittins, K. Glazebrook, and R. Weber. *Multi-armed bandit allocation indices.* John Wiley & Sons, Ltd., Chichester, 2011.

[15] J. M. Harrison, N. B. Keskin, and A. Zeevi. Bayesian dynamic pricing policies: Learning and earning under a binary prior distribution. *Management Science*, 58(3):570–586, 2012.

[16] J. M. Harrison and N. Sunar. Investment timing with incomplete information and multiple means of learning. *Oper. Res.*, 63(2):442–457, 2015.

[17] P. Lakner. Utility maximization with partial information. *Stochastic Process. Appl.*, 56(2):247–273, 1995.

[18] R. S. Liptser and A. N. Shiryayev. *Statistics of random processes. I.* Applications of Mathematics, Vol. 5. Springer-Verlag, New York-Heidelberg, 1977. General theory, Translated by A. B. Aries.

[19] R. McDonald and D. Siegel. The value of waiting to invest. *Quart. J. Econom.*, 101(4):707–727, 1986.

[20] A. Merhi and M. Zervos. A model for reversible investment capacity expansion. *SIAM J. Control Optim.*, 46(3):839–876, 2007.
[21] G. Peskir and A. Shiryaev. *Optimal stopping and free-boundary problems*. Lectures in Mathematics ETH Zürich. Birkhäuser Verlag, Basel, 2006.
[22] L. C. G. Rogers and D. Williams. *Diffusions, Markov processes, and martingales. Vol. 2*. Cambridge Mathematical Library. Cambridge University Press, Cambridge, 2000. Itô calculus, Reprint of the second (1994) edition.
[23] N. Sunar, S. Yu, and V. G. Kulkarni. Competitive investment with Bayesian learning: choice of business size and timing. *Oper. Res.*, 69(5):1430–1449, 2021.

E. Ekström and T. Tolonen-Weckström: Department of Mathematics, Uppsala University, Box 256, 75105 Uppsala, Sweden.
Email address: ekstrom@math.uu.se
Email address: topias.tolonen@math.uu.se

Y. Kitapbayev: Mathematics Department, Khalifa University of Science and Technology, PO Box 127788, Abu Dhabi, United Arab Emirates.
Email address: yerkin.kitapbayev@ku.ac.ae

A. Milazzo: School of Management and Economics, Dept. ESOMAS, University of Turin, C.so Unione Sovietica 218bis, 10134 Turin, Italy.
Email address: alessandro.milazzo@unito.it