Optimal Dynamic Information Acquisition

Weijie Zhong

Columbia University

Abstract

In this paper, I study a model in which an impatient decision maker (DM) acquires information prior to making her decision. I allow for fully flexible information: the DM can choose any dynamic signal process as an information source, subject to a flow cost on signal informativeness. I fully characterize the optimal learning dynamics in continuous-time: the DM seeks for an informative signal arriving as a Poisson process. The arrival of signal confirms prior belief and leads to immediate action. Learning has increasing precision and decreasing intensity over time.

Keywords: dynamic information acquisition, rational inattention, Poisson-bandits, stochastic control

JEL: D11 D81 D83

1. Introduction

When individuals make decisions, they often have imperfect information about the payoffs of different alternatives. A decision maker (DM) may acquire information about the payoffs before decision making. Examples of information acquisition activities include scientists running experiments, market analysts mining data, manufacturers building prototypes, etc. These activities share common features: 1. Information acquisition is flexible: In the aforementioned examples, researchers can design the information generating process in a flexible way. This feature has become a common practice due to the developments in data science, statistics and large scale computation. 2. Learning is costly: to gather and process information, researchers build prototype products, maintain experiment instruments, hire data analysts, etc. The cost spent is related to the information generating capacity. Moreover, the cost usually grows disproportionately when we increase the capacity. 3. Timing is critical: information takes time to acquire while delay is also costly. In financial markets investment opportunities might disappear in a wink. Even for scientific researches, considering their potential medical, commercial or national defense applications, delay might lead to high social costs.

I am indebted to Yeon-Koo Che for guidance and support. I am grateful to Mark Dean, Marina Halac, Johannes Hörner, Navin Kartik, Qingmin Liu, Konrad Mierendorf, Pietro Ortoleva and Ming Yang for insightful comments. I benefited from in-depth conversations with Sylvain Chassang, Yingni Guo, Shaowei Ke, Annie Liang, Alessandro Lizzieri, David Pearce, Debraj Ray, Ennio Stacchetti, Jakub Steiner, Philipp Strack, Glen Weyl, Michael Woodford, Siyang Xiong, as well as participants at various conferences and seminars. All remaining errors are mine.

Email address: wz2269@columbia.edu (Weijie Zhong)

First draft: June 15, 2016

This draft: April 28, 2018
costs. A question naturally arises: what is the optimal dynamic information acquisition strategy when design of information is flexible?

The theoretical answer has not yet been fully provided. In fact, this question cannot be fully answered in a canonical dynamic learning framework (Wald-type models), where information is modeled as an exogenous stochastic process, and the DM controls some parameters of the process. In these models, the modeling assumption on the stochastic process limits the types of learning the DM can choose: Gaussian learning models capture only incremental evidence, Poisson bandit models capture only certain types of breakthroughs, etc.

To address this question, I develop a dynamic information acquisition framework that endogenizes the information generating process. A DM is uncertain about a pay-off relevant state and acquires information before making a one-time choice of action. I introduce three economic assumptions on acquisition of information. First, I assume that the DM controls her belief process about the state in a non-parametric way, which equivalently models the flexible design of information. Second, the flow cost of information depends on how fast it reduces uncertainty about the unknown state. Third, the DM is impatient and the flow cost of information is strictly convex, which drives non-trivial dynamic learning behavior.

Within this framework, I try to accomplish three main goals. The first is to endogenously pin down a family of stochastic processes that model optimal information acquisition. The main finding is that it is without loss of optimality to limit the control space to the family of jump-diffusion belief processes, i.e. it is endogenously optimal to acquire a Poisson signal (which drives a jump in belief) and a Gaussian signal (which drives diffusion in belief). Within this family, the DM controls four parameters: the direction and size of belief jump, the arrival frequency of jump, and the flow variance of belief diffusion. The direction and size of a belief jump represents the type and precision of a Poisson signal. The frequency of jump represents the frequency of Poisson signal. The flow variance of belief diffusion represents the intensity/precision of Gaussian signal.

The second goal is to characterize the optimal learning dynamics. The first finding is that Poisson signal strictly dominates Gaussian signal almost surely. This finding suggests that in an environment with flexible information acquisition, dynamic learning assembles a Poisson learning model, as opposed to Gaussian learning, random sampling, or drift diffusion models etc. Then I characterize the optimal Poisson learning strategy. The DM’s optimal strategy is to search for a signal confirming her prior belief, namely observing the signal induces a belief jump towards the a priori more likely state being true. Optimal stopping time is immediately after the arrival of signal. Absence of signal is followed by continuing acquiring information. The optimal precision of the Poisson signal is increasing over time. The optimal arrival frequency of the signal and cost spent on learning is decreasing over time.

The main intuition behind the optimal strategy is mostly explained by a novel precision-frequency trade-off. A Poisson signal is characterized by two parameters: size of jump (precision) and frequency of jump. Cost of signal is increasing in both parameters. So optimal precision and frequency are pinned down by the marginal
gain from each parameter. This trade-off depends on the level of continuation value. When value is larger the DM loses more from discounting, hence prefers frequency more than signal precision. Applying this intuition pins down each of the parameters: 1. Gaussian signal is a special Poisson signal with close to zero precision and infinite frequency. As a result, Gaussian signal is strictly suboptimal except when value is very high, so high that the DM is on the edge of stopping learning. 2. If signals are confirmatory, a more extreme prior belief is closer to any given posterior belief. Hence confirmatory signal is consistent with the higher preference for frequency at more extreme beliefs (whose values are higher). 3. When signal does not arrive, belief becomes less extreme and continuation value decreases. So optimal precision increases and frequency decreases over time.

The third goal is to provide connection between dynamic information acquisition problems and static rational inattention problems. I show that in two kinds of limits, when information cost function is linear or when there is no discounting, dynamic information acquisition problems have essentially static solutions. When information cost is linear, the incentive for dynamically smoothing information diminishes, and the optimal learning strategy is to acquire all information instantaneously. When there is no discounting, the precision-frequency trade-off diminishes, so essentially all learning strategies are equivalent.

The rest of the paper is structured as follows. The related literature is reviewed in Section 2. The main continuous-time model and its Hamilton-Jacobi-Bellman (HJB) equation are introduced in Section 3. I provide foundation for this model in Section 4 by showing that it is the zero period length limit of a discrete time dynamic information acquisition model with fully flexible design of signal process. Section 5 fully characterizes the optimal strategy and illustrates the main intuition behind the result. In Section 6 I discuss assumptions used in my model. For non-crucial assumptions generalizations are provided. For the crucial posterior-separability assumption, I discuss the robustness of each of my main results.

2. Related literature

2.1. Dynamic information acquisition

My paper is closely related to a literature about designing information acquisition strategy in a dynamic way to facilitate decision making. The canonical approach is to model information flow as a family of stochastic processes. The DM controls parameters determining the information flow, and chooses when to stop learning and make decision. The earliest works focus on the duration of learning. Wald (1947) assumes information to be exogenous, and the DM has control over decision time and action choice. The problem can be formulated as an optimal stopping problem. Moscarini and Smith (2001) endogenizes the experimentation intensity by allowing the DM to control precision of a Gaussian signal. Similar learning problem is used as a learning-theoretic foundation for drift-diffusion model (DDM) by Fudenberg et al. (2015). Following a different route, Che and Mierendorff (2016), Mayskaya (2016) and
Liang et al. (2017) studies the dynamic choice of information sources when multiple signal processes are available. Other related papers on this topic include Weitzman (1979), Callander (2011), Klabjan et al. (2014), Ke and Villas-Boas (2016), Ke et al. (2016), Doval (2018), Nikandrova and Pancs (2018), etc.

In contrast to the canonical models, the key new feature of my framework is that the DM can design the information generating process non-parametrically, i.e. even the family of processes is picked endogenously. In a similar vein to this paper, two concurrent papers Steiner et al. (2016) and Hébert and Woodford (2016) also model dynamic information acquisition non-parametrically. However they assumed away the necessity to design information dynamically: in Steiner et al. (2016) the assumption of linear flow cost makes it optimal to learn instantaneously, while in Hébert and Woodford (2016) the assumption of zero discount rate makes all dynamic learning strategies essentially equivalent.\footnote{Steiner et al. (2016) assumes state to be stochastic. So all dynamics in information comes from dynamics of state directly, rather than incentive to smooth information. Hébert and Woodford (2016) picks Gaussian process exogenously to justify a neighbourhood based static information cost structure.} I analyze these two special cases in Section 6.

A major result of my framework is the endogenous optimality of Poisson signals. A stronger result is established in Section 6.1 if we only compare Poisson signal and Gaussian signal. Poisson signal dominates Gaussian signal for generic cost functions of information. This result potentially justifies Poisson learning models, which are used in a wide range of problems: e.g. Keller et al. (2005), Keller and Rady (2010), Che and Mierendorff (2016), Mayskaya (2016), see also a survey by Hörner and Skrzypacz (2016). I show that the optimal Poisson signal confirms the a priori more likely state, which is consistent with the finding of Che and Mierendorff (2016) when the DM is uncertain of the state. Che and Mierendorff (2016) predicts contradictory learning when DM is more certain of the state, because in their model signal frequency is exogenous.

2.2. Rational inattention

The intra-period information acquisition behavior in my paper is modeled based on the literature about rational inattention. A common approach in this literature is to model information as general state contingent signal distributions (Blackwell information structure). The DM can choose from the set of all information structures subject to a cost or a constraint on information. An Entropy based rational inattention framework is first introduced in Sims (2003). Matejka and McKay (2014) studied the flexible information acquisition problem using an Entropy based informativeness measure and derived a generalized Logit decision rule. Caplin and Dean (2015) takes an axiomatic approach and characterized decision rules that can be produced by an information acquisition problem.

To measure the cost of information, I borrow a posterior separable information measure from Caplin and Dean (2013). Posterior separability is a generalization of mutual information introduced in Shannon (1948), and is widely used to model cost of information in Gentzkow and Kamenica (2014), Clark (2016), Matyskova (2018), Rappoport and Somma (2017), etc. I provide an axiomatization for posterior separability based on
dynamic decomposition of information structures. Caplin et al. (2017) provides a full behavior axiomitization for (uniform) posterior separability. Morris and Strack (2017) provides a dynamic foundation for posterior separability based on implementing information structure with Gaussian learning.

2.3. Mathematical theories

**Stochastic control**

Methodologically, this paper is also closely related to the theory of continuous-time stochastic control. The early theories study control processes measurable to the natural filtration of a Brownian motion (see Fleming (1969) for a survey). The application of Bellman (1957)’s dynamic programming principle leads to the Hamilton-Jacobi-Bellman (HJB) equation characterization of value function. On the contrary, the main stochastic control problem of this paper has general martingale control process. It is a variant of the (semi)martingale models of stochastic control, studied in Davis (1979), Boel and Kohlmann (1980), Striebel (1984), etc. The existing theory provides abstract characterization and existence results, without practically tractable solving methodology provided (not even computationally tractable). This paper introduce an indirect method that establishes a tractable HJB equation. I prove that the HJB equation is identical to that of a jump-diffusion control model (see Hanson (2007)).

**Concavification**

A Blackwell information structure can be modeled using two approaches. The first is direct signal approach (e.g. Woodford (2009), Matejka and McKay (2014), Yang (2015)). Direct signal approach is not applicable here because information is acquired sequentially. I adopted the second posterior belief approach. The standard tool for this approach is a concavification method introduced in Aumann et al. (1995) (applied to Bayesian persuasion in Kamenica and Gentzkow (2009), and to information acquisition in Caplin and Dean (2013)). The concavification method maximizes the expectation of a value function by choosing a distribution of posterior beliefs. The objective function in this paper is not expected value, so I use a constrained optimalization generalization of this method developed in Zhong (2017a).

3. Model setup

The main model is a continuous-time stochastic control problem. An impatient DM chooses a one-time action at endogenous decision time. The DM controls the whole belief process before the decision time, bearing a cost on information.

**Decision problem**: Time $t \in [0, +\infty)$. The DM discounts future utility with rate $\rho > 0$. Both action space $A$ and state space $X$ are finite. Utility associated with action-state pair $(a, x) \in A \times X$ at time $t$ is $e^{-\rho t} u(a, x)$. Not knowing the true state, the DM holds a prior belief $\mu \in \Delta(X)$ about the state. Her preference under uncertainty is von Neumann-Morgenstern expected utility. Define $F(\mu) \doteq \max_{a \in A} E_{\mu}[u(a, x)]$ as the expected utility from choosing optimal action given belief $\mu$.

**Information**: I model information using a belief based approach. It is well known that a distribution of posterior beliefs is induced by an information structure according
to Bayes rule if and only if the expectation of posterior beliefs equals the prior. Hence, in a static environment the choice of an information structure can always be reformulated as the choice of a distribution of posterior beliefs (see Kamenica and Gentzkow (2009) for example).

Extending this formulation to the dynamic environment in current paper, I assume that the DM controls the posterior belief process \( \langle \mu_t \rangle \). Now Bayes rule should be satisfied at every instant of time — the expectation of \( \mu_s \) is \( \mu_t \) for any \( s > t \). Therefore I restrict \( \langle \mu_t \rangle \) to be a martingale, with \( \langle \mathcal{F}_t \rangle \) being its natural filtration. In Section 4, I formally justify this assumption by showing that in discrete-time, choosing the optimal belief martingale is equivalent to choosing the optimal dynamic information structure.

A technical remark: it is useful to define the following operator \( \mathcal{L}_t \) for \( \langle \mu_t \rangle \):

\[
\mathcal{L}_t f(\mu_t) = E \left[ \frac{df(\mu_t)}{dt} \bigg| \mathcal{F}_t \right] = \lim_{t' \to t^+} E \left[ \frac{f(\mu_{t'}) - f(\mu_t)}{t' - t} \bigg| \mathcal{F}_t \right]
\]

By definition, \( \mathcal{L}_t f \) captures the expected speed at which \( f(\mu_t) \) is increasing. If \( \langle \mu_t \rangle \) is a well-behaved markov process and \( f \) is \( C^2 \), then \( \mathcal{L}_t f \) is the standard infinitesimal generator (subscript \( t \) omitted).  

Cost of information: I assume that the flow cost of information depends on how fast it reduces uncertainty. \( H : \Delta(X) \to \mathbb{R} \) is a concave and continuous function. I call \( H \) an uncertainty measure, because \( E[H(\mu)] \) is larger when \( \mu \) is less certain (i.e., \( \mu \) is Blackwell less informative about the state). Flow cost of information is assumed to be \( h(I_t) \), where \( I_t \) is defined as:

\[
I_t = -\mathcal{L}_t H(\mu_t)
\]

\( I_t \) is by definition the speed at which uncertainty reduces due to belief updating. I call \( I_t \) the (flow) informativeness measure. One example for \( H \) function is the Entropy function: \( H(\mu) = -\sum \mu_x \log(\mu_x) \). It is well known that receiving information reduces Entropy, and \( I_t \) is exactly the speed of Entropy reduction. The form of informativeness measure \( I_t \) is the main technical assumption for my analysis. I show that its discrete-time foundation is "posterior separability" in Section 4. Moreover, in Section 6.1.1, I show an axiom characterizing posterior separability. In Section 6.1.2, I generalize this assumption on \( I_t \).

Stochastic control problem: The DM solves the following stochastic control problem:

\[
V(\mu) = \sup_{\langle \mu_t \rangle \in \mathcal{M}, \tau} E \left[ e^{-\rho \tau} F(\mu_\tau) - \int_0^\tau e^{-\rho t} h(I_t) dt \right]
\]

(1)

where \( \mathcal{M} \) is the set of admissible processes, defined as:

1. \( \langle \mu_t \rangle : \mathbb{R}^+ \to \Delta(X) \) is a martingale with cadlag path;
2. \( \mu_0 = \mu; \)

---

\( ^2 \)For all Feller processes, \( \mathcal{L} \) is well-defined. For example \( \mu_t = a t + \sigma W_t \) (W is standard brownian motion) is Fellerian, and \( \mathcal{L} f(\mu) = a f(\mu) + \frac{1}{2} \sigma^2 f''(\mu) \). For a general martingale and general \( f \), \( \mathcal{L}_t f \) is not necessarily well-defined. I will impose minimal assumptions on \( \langle \mu_t \rangle \) to ensure that \( \mathcal{L} \) is well-defined later when defining admissibility.

\( ^3 \)\( H \) is concave if and only if \( \forall \mu \) Blackwell more informative than \( \nu \), \( E[H(\mu)] \leq E[H(\nu)] \). Concave function is continuous on interior. So the only loss of generality is continuity of \( H \) on \( \partial \Delta(X) \).
3. The objective function in Equation (1) is integrable. The objective function in Equation (1) is fairly standard in a canonical information acquisition problem. The DM acquires information to affect \( \langle \mu_t \rangle \) and chooses stopping time \( \tau \) to maximize the expected stopping payoff \( E[e^{-\tau}F(\mu_\tau)] \) less the total information cost \( E[\sum_1^\tau e^{-\tau}h(\mu_t)dt] \). The novel feature is that the DM is allowed to fully control \( \langle \mu_t \rangle \), as opposed to canonical models where the DM controls only a few parameters determining \( \langle \mu_t \rangle \). The non-parametric control of belief process exactly models the flexible design of information by the DM.

I make the following assumption on cost function to rule out a trivial case that learning everything immediately at \( t = 0 \) is optimal.

**Assumption 1.** \( h : \mathbb{R}^+ \rightarrow \mathbb{R}^+ \) is weakly increasing, convex and continuous. \( \lim_{I \rightarrow \infty} h'(I) = \infty \).

Assuming an increasing and continuous cost is standard. The interpretation of convexity and the Inada condition \( \lim h'(I) = \infty \) is that the DM has strict incentive to smooth the acquisition of information. Given Assumption 1, if the DM acquires all information immediately then uncertainty reduces at infinite speed and the marginal cost \( h'(I) \) is infinite, hence suboptimal. I solve a special case violating Assumption 1 in Section 6.2, where I assume \( h \) to be linear. In this case the optimal strategy is to learn everything at \( t = 0 \) (a static strategy).

Now in Section 3.1, I show a few examples of canonical Wald-type learning models, each being a variant of Equation (1) with extra constraint on the set of admissible belief processes. The examples also illustrate that the non-parametric model is in fact necessary if we want to study the benchmark of flexible information acquisition.

### 3.1. Examples

State is binary \( X = \{0, 1\} \). The prior belief of state \( x = 1 \) is \( \mu \in (0, 1) \). \( A = \{a, b\} \). The DM wants to match the state: \( u(a, 0) = u(b, 1) = 1; u(a, 1) = u(b, 0) = -1 \). Discount rate \( \rho = 1 \). \( H \) is standard Entropy function: \( H(\mu) = -\mu \log(\mu) - (1 - \mu) \log(1 - \mu) \). Information cost \( h(I) = \frac{1}{2}I^2 \). Now consider three simple learning strategies:

1. **Gaussian learning:** the signal follows a diffusion process whose drift is the true state, and whose variance is controlled by the DM. It is well known that the posterior belief follows a diffusion process: \( d\mu_t = \sigma_t dW_t \), where \( W \) is a Wiener process (Bolton and Harris (1999)). By Ito’s lemma, \( -LH(\mu_t) = -\frac{1}{2}\sigma_t^2H''(\mu_t) = \frac{\sigma_t^2}{2\mu_t(1-\mu_t)} \). This control problem is studied in Moscarini and Smith (2001), where optimal information acquisition is characterized by HJB:

\[
\rho V_1(\mu) = \sup_{\sigma > 0} \frac{1}{2} \sigma^2 V_1''(\mu) - \frac{1}{2} \left( \frac{\sigma^2}{2\mu(1-\mu)} \right)^2
\]

---

4 Integrability means: if we replace \( L \) with finite difference and replace integral with Riemann sum based on any discretization of time, the limit exists when time period shrinks to 0. For the formal definition, see Section 4.2.

5 Convexity and Inada condition are not the weakest condition guaranteeing information smoothing. A weaker sufficient condition is: let \( \bar{X} = \lim_{\rho \rightarrow 0} \frac{\rho}{h(\bar{X})}, \) then \( \rho \bar{X} - h(\bar{X}) > \rho \sup F \). This condition explicitly states that when \( f \) is sufficiently large, \( f \) is just sufficiently convex that utility gain from smoothing information dominates loss from longer waiting. All following theorems in this paper are proved under this weaker condition.

6 With “belief elasticity” defined as \( \phi(\mu) = \mu(1-\mu) \) in my model.
with smooth pasting when $V_1$ touches $F$. The solution to the HJB is plotted as the blue curve in Figure 1-a. The blue shaded region $(v_1, v_2)$ is the experimentation region and the non-shaded region is the stopping region.

2. **Perfect revealing evidence:** the DM observes discrete breakthroughs that reveals the true state. She controls the arrival rate $\lambda_t$ of evidence. Therefore, belief follows a Poisson process: $d\mu_t = (1 - \mu_t)dJ_1^\gamma(\lambda_t \mu_t) + (0 - \mu_t)dJ_0^\gamma(\lambda_t(1 - \mu_t))$, where $J^\gamma(\cdot)$ are independent Poisson counting processes with Poisson rate $\cdot$. The Entropy reduction speed is $\lambda_t H(\mu)$. The value function is characterized by:

$$\rho V_2(\mu) = \sup_{\lambda > 0} \lambda(\mu F(1) + (1 - \mu)F(0) - V_2(\mu)) - \frac{1}{2}(\lambda H(\mu))^2$$

$V_2$ is plotted as the red curve in Figure 1-(b). To compare the two learning strategies, I also plot $V_1$ as the blue curve in Figure 1-(b). In the blue shaded region $V_1$ is higher and in the red shaded region $V_2$ is higher. The two arrows show the belief jumps induced by breakthroughs at $\mu$.

3. **Confirmatory evidence:** the DM seeks only confirmatory news. With fixed Poisson rate $\gamma = 2$, a good news arrives and reveals the most likely state (state 1 when $\mu_t > 0.5$ and 0 when $\mu_t < 0.5$). If the news doesn’t arrive, belief drifts towards opposite direction. This is a standard good news model (in each half space). Belief follows a compensated Poisson process:

$$d\mu_t = \begin{cases} (1 - \mu_t)(dJ_1^\gamma(\gamma \mu_t) - \gamma \mu_t dt) & \text{when } \mu > 0.5 \\ (0 - \mu_t)(dJ_1^\gamma(\gamma(1 - \mu_t)) - \gamma(1 - \mu_t) dt) & \text{when } \mu < 0.5 \end{cases}$$

The HJB of value function:

$$\rho V_3(\mu) = \begin{cases} \gamma \mu(F(1) - V_3(\mu) - V_1^\gamma(\mu)(1 - \mu)) - \frac{1}{2}(\gamma \mu(H(\mu) + H'(\mu)(1 - \mu))^2 & \text{when } \mu > 0.5 \\ \gamma(1 - \mu)(F(0) - V_3(\mu) + V_1^\gamma(\mu)\mu(1 - \mu) - \frac{1}{2}(\gamma(1 - \mu)(H(\mu) - H'(\mu)) & \text{when } \mu < 0.5 \end{cases}$$

$V_3$ is plotted as the black curve in Figure 1-(c). To compare the three strategies, I also plot $\max\{V_1, V_2\}$ (blue when $V_1$ is higher and red when $V_2$ is higher). In the grey shaded region, $V_3$ is higher than $\max\{V_1, V_2\}$. The arrow shows the belief jump.

In the examples, Equation (1) resembles three commonly studied information acquisition models when we impose different parametrization constraints on $\langle \mu_t \rangle$. Figure 1-c shows that the three strategies are suboptimal individually — there is non-degenerate region for each of them being optimal. It is a reasonable conjecture that
if we make more stochastic processes admissible, the DM becomes better off by using a more complex learning strategy. Therefore, removing all extra constraints on admissible processes $M$ is necessary to prevent the loss of generality from exogenous restrictions on the information process.

3.2. Dynamic programming and HJB equation

Solving Equation (1) is not an easy task due to the abstract strategy space. To the best of my knowledge, there is no existing general theory applicable to this stochastic control problem. The closest problems are studied in a set of remarkable papers on martingale method in stochastic control (Davis (1979), Boel and Kohlmann (1980), Striebel (1984)). These papers introduce abstract formulations of stochastic control problems with general (semi)martingale control process. These papers study finite horizon and specific forms of objective functions, hence they do not cover Equation (1).

Nevertheless, it is useful to introduce the general dynamic programming principle and HJB characterization from this literature. Based on the intuition of dynamic programming, it is a reasonable conjecture that $V(\mu_t)$ satisfies the following HJB:

$$
\max \left\{ F(\mu_t) - V(\mu_t), \quad -\rho V(\mu_t) + \sup_{\mu_t} \left\{ \mathcal{L}_t V(\mu_t) - h(-\mathcal{L}_t H(\mu_t)) \right\} \right\} = 0
$$

Equation (2) is conceptually the same as the standard HJB equation. Recall the definition for infinitesimal generator, $\mathcal{L}_t V(\mu_t)$ is the flow utility gain from continuing. The exact form of $\mathcal{L}_t V$ and $\mathcal{L}_t H$ depends on the underlying probability space, the filtration and the control process in a neighbourhood of $t$ (which are summarized by the symbol $d\mu_t$). So Equation (2) essentially states the dynamic programming principle: at any instance when the control is chosen optimally, either stopping is optimal (the first term is 0), or continuing is optimal and the net continuation gain equals loss from discounting (the second term is 0).

For a simple example, let $M$ be a family of Markov jump-diffusion belief processes:

$$
d\mu_t = (v(\mu_t) - \mu_t)(dJ_t(p_t) - p_t dt) + \sigma(\mu_t)dW_t
$$

where $(p, v, \sigma) : \mu_t \rightarrow \mathbb{R}^+ \otimes \Delta(\text{Supp}(\mu)) \otimes \mathbb{R}^{\text{Supp}(\mu)}$ are control parameters. $J_t(\cdot)$ is Poisson counting process with Poisson rate $(\cdot)$, $W_t$ is a standard Wiener process. Itô’s lemma implies an explicit form for infinitesimal generator:

$$
\mathcal{L} V(\mu) = pV(v) - V(\mu) - \nabla V(\mu)(v - \mu) + \frac{1}{2} \sigma^T H(\mu) \sigma
$$

where $H$ is Hessian matrix operator. The corresponding HJB becomes:

$$
\rho V(\mu) = \max \left\{ \rho F(\mu), \sup_{p,v,\sigma} \left\{ p(V(v) - V(\mu) - \nabla V(\mu)(v - \mu)) + \frac{1}{2} \sigma^T H(\mu) \sigma \right\} \right\}
$$

$$
- h \left( p(H(\mu) - H(v) + \nabla H(\mu)(v - \mu)) - \frac{1}{2} \sigma^T HH(\mu) \sigma \right)
$$
Now consider the general martingale \( \langle \mu_t \rangle \). Of course the first gap to be filled is to formally establish Equation (2), which the general martingale method is not applicable to. Even if Equation (2) is true, the martingale method only states the existence of such \( \mathcal{L}_t V \) (for example theorem 4.3.1 of Boel and Kohlmann (1980)), instead of an explicit representation. Not knowing the explicit form of the map from \( d \mu_t \) to \( \mathcal{L}_t V \), Equation (2) is practically intractable. This is considered as the main drawback of the martingale method (see discussions in Davis (1979)).

In this paper, I bypass the difficulties faced by existing stochastic control theories by using an indirect method and 1) formally establish the HJB Equation (2), 2) simplify the operator \( \mathcal{L}_t \) to a practically tractable form:

**Theorem 1.** Assume \( H \) is strictly concave and \( C^{(2)} \) on interior beliefs, and Assumption 1 is satisfied. Let \( V(\mu) \in C^{(1)} \) be a solution\(^7\) to Equation (4) then \( V(\mu) \) solves Equation (1).

Theorem 1 first states that \( V(\mu) \) is characterized by a HJB equation, as one could expect. More surprisingly, Theorem 1 also states that the HJB is exactly Equation (4). It is known in jump-diffusion control theory that the solution of Equation (4) is the solution of Equation (1) with admissible controls defined by SDE (3) (theorem 6.3 of Hanson (2007)). Therefore, Equation (1) can be solved with a Markov jump-diffusion process, where the control variables are arrival rate \( (p) \) and size \( (v - \mu) \) of a Poisson jump, and the flow variance \( (\sigma) \) of Gaussian diffusion.

The compensated Poisson jump part and Gaussian diffusion part in SDE (3) each represents a simple learning strategy. The Poisson jump in belief process can be induced by observing a non-conclusive news whose arrival follows a Poisson process. The compensating belief drift is induced by observing no news arriving. I say that the DM is using Poisson learning or is acquiring Poisson signal when there is a compensated Poisson part in belief process. The Gaussian diffusion in belief process can be induced by observing the realization of a Gaussian process, with state \( x \) being its unobservable drift. I say that the DM is using Gaussian learning or is acquiring Gaussian signal when there is a diffusion part in belief process.

The control variables in Equation (4) represent four kinds of trade-offs in a one-step decision problem. The first is the standard continuing v.s. stopping trade-off in dynamic programming problems, captured by the outer layer maximization. The second is information cost v.s. utility gain trade-off. More informative signal leads to higher utility gain, it also incurs higher cost due to increasing \( h \). The third is Poisson v.s. Gaussian trade-off. Total amount of informativeness is allocated to the Poisson signal \( (p, v) \) and the Gaussian signal \( \sigma \). The forth is precision v.s. frequency trade-off. Fixing the informativeness of a Poisson signal, when DM increases the precision of a signal (chooses \( v \) further away from prior \( \mu \)), she has to sacrifice the arriving frequency of the signal (lower arrival rate \( p \)). These trade-offs, especially the precision-frequency trade-off, will be discussed in details to characterize the solution to Equation (4) in Section 5.

\(^7\) Solution to second order ODE is not well defined. To be precise, \( V \) is viscosity solution (see Crandall et al. (1992)) to Equation (4). In the viscosity solution, \( \sigma^T H V(\mu, \sigma) \sigma \) is replaced by \( D^2 V(\mu, \sigma) ||\sigma||^2 \), where \( D^2 V(\mu, \sigma) = \lim_{s \to 0} \frac{1}{s^2} \mathbb{E}_{\text{law}(\sigma)} \{V(\mu + \sigma) - V(\mu) - DV(\mu) \sigma\} \).
The indirect method I use to prove Theorem 1 is to characterize Equation (1) as a limit of discrete-time information acquisition problems when time period is going to zero. On the one hand, the discrete-time problem serves as foundation for Equation (1). Modeling assumptions like using martingale $\langle \mu_t \rangle$ as control, and using $-\mathcal{L}_t H$ as measure of informativeness are justified by more fundamental assumptions. On the other hand, Theorem 1 is proved by showing that Equation (4) could be obtained as the limit of discrete time value functions. The discrete-time analyses are presented in Section 4. Readers more interested in the solution to HJB Equation (4) can jump to Section 5.

4. Discrete-time foundation

In this section, I first propose a discrete-time problem in Section 4.1 that explicitly models flexible design of signal process. Then I show in Section 4.2 that it equivalently characterizes the discretization of the continuous-time stochastic control problem Equation (1). In Section 4.3 I introduce a key lemma that links all the discrete-time analysis and proves Theorem 1.

4.1. General discrete-time problem

Decision problem: Time is discrete $t \in \mathbb{N}$. Period length is $dt > 0$. The other primitives $(A, X, u, \mu, \rho)$ are the same as in Section 3. The utility of action-state pair $(a, x)$ in period $t$ is $e^{-\rho dt} u(a, x)$.

Strategy: a strategy is a triplet $(S^t, \tau, A^t)$. $S^t$ is a random process correlated with state, called information structure. Realization of $S^t$ is called a signal history. Signal history up to period $t$ is denoted by $S^t$. Each $S^t$ specifies the signal structure acquired in period $t$ conditional on all histories up to period $t$. $\tau$ is a random variable whose realization is in $\mathbb{N}$. $\tau$ specifies a random decision time. Action choice $A^t$ is a random process whose realization is in $A$. Each $A^t$ specifies the joint distribution of action choice and state conditional on making decision in period $t$. Let the marginal distribution of the state be denoted by random variable $X$.

Cost of information: The cost of information is assumed to be a discretization of that in Equation (1). Define $h_{dt}(I) = h(\frac{1}{dt}) dt$. Define the measure of signal informativeness:

**Assumption A.** $I(S; X|\mu) = E_s[H(\mu) - H(v(|s)))]$, where $v$ is the posterior belief about $x$ induced by signal $s$ according to Bayes rule.

The per-period cost of information conditional on history $(S^{t-1}, 1_{t \leq t})$ is $h_{dt}(I(S^t; X|S^{t-1}, 1_{t \leq t}))$. It is not difficult to see that $I(S^t; X|S^{t-1}, 1_{t \leq t})$ is the difference formulation of $-\mathcal{L}_t H(\mu_t) dt$.

**Assumption A** is called (uniform) posterior separability in the literature. If $H$ is the standard Entropy function, then $I$ is the mutual information between signal $S$ and unknown state $X$.

---

8 $S^{-1}$ is defined as a degenerate random variable that induces belief same as prior belief $\mu$ of the DM for notation simplicity.
Dynamic Optimization: The dynamic optimization problem of the DM is:

\[
V_{dt}(\mu) = \sup_{S^t, \tau, A^t} E \left[ e^{-\rho dt \cdot \tau} u(S^t, A^t) - \sum_{t=0}^{n} e^{-\rho dt \cdot t} h_{dt}(I(S^t; X^t | S^{t-1}, 1_{\tau \leq t})) \right]
\]  

(5)

The two constraints in Equation (5) are called information processing constraints. The first constraint states that signal history prior to action time is sufficient for action time. The second constraint states that signal history prior to period t is sufficient for action at time t. They are extensions to the standard measurability requirement, allowing randomness unrelated to unknown state to be added.

4.2. Discretization of continuous-time problem

For every admissible in Equation (1), the objective function \( E[e^{-\rho t} F(\mu_t)] - \int_0^t e^{-\rho t} h(I_t) dt \) is defined as a limit of Riemann sum:

\[
\lim_{dt \to 0} \sum_{i=0}^{\infty} \text{Prob}(x \in [i dt, (i+1) dt]) \left( e^{-(i+1) \rho dt} - \sum_{j=0}^{i} e^{-j \rho dt} h \left( E \left[ \frac{H(\mu_{jdt}) - H(\mu_{(j+1)dt})}{dt} F_{jdt} \right] \right) dt \right)
\]

Noticing that the limit exists and is unique for any discretization, the uniform discretization I am using. Since \( \langle \mu_t \rangle \) is a martingale and \( \tau \) is measurable to \( F_t \), the discretized \( \langle \mu_{jdt} \rangle \) is a discrete-time martingale, and discretized \( \tau \) is still a stopping time measurable to the natural filtration. So any feasible strategy to Equation (1) naturally defines a feasible strategy in the following optimization problem:

\[
V_{dt}(\mu) = \sup_{\mu_t, \tau} E \left[ e^{-\rho dt \cdot \tau} F(\mu_{\tau}) - \sum_{t=0}^{\tau} e^{-\rho dt \cdot t} h_{dt}(E[H(\mu_t) - H(\mu_{t+1}) | F_t]) \right]
\]  

(6)

where \( \mu_t \) is a discrete-time martingale and \( \tau \in \mathbb{N} \) is measurable to \( \mu_t \)’s natural filtration \( F_t \). In fact, when Assumption A is satisfied, Equation (6) is equivalent to Equation (5). Equation (5) can be modified by removing redundant information and rewriting based on beliefs to become Equation (6). So for simplicity I use same notation \( V_{dt} \) to represent the value function for both Equations (5) and (6). (for the formal proof see Lemma A.3).

Now consider the relation between \( V \) and \( V_{dt} \). \( V \) is defined by first taking the limit \( dt \to 0 \), then take the supremum w.r.t. continuous-time strategies. On the contrary, \( V_{dt} \) is defined with the order switched, first taking supremum for each discrete-time problems then take \( dt \to 0 \). The following lemma shows that the limit and supremum operators are interchangeable:

Lemma 1. \( \lim_{dt \to 0} V_{dt}(\mu) = V(\mu) \).

\(^9\)Noticing that in every period, the information in current period has not been acquired yet. So decision can only be taken based on the information already acquired in the past. As a result in information processing constraints information is advanced by one period. This within period timing issue does not make a difference when going to continuous-time limit. It matters when going to the linear cost limit and I will highlight this in Theorem 7.
Equation (6) is a discrete-time optimization problem, with exponential discounting and bounded utility functions. Therefore, standard dynamic programming principle guarantees the Bellman characterization of $V_{dt}$:

**Lemma 2** (Discrete-time Bellman). Assume Assumption A is satisfied, then $V_{dt}$ solves the following functional equation:

$$V_{dt}(\mu) = \max \left\{ F(\mu), \sup_{p_i, v_i} e^{-\rho dt} \sum_{i=1}^{N} p_i V_{dt}(v_i) - h_{dt}(\mu) - \sum p_i H(v_i) \right\}$$  \hspace{1cm} (7)

where $N = 2|X|$, $p \in \Delta(N)$, $v_i \in \Delta X$.

Equation (7) is a standard Bellman equation, except that it covers a restricted space of strategies. The choice of signal structure is restricted to have support size no larger than $2|X|$, while the original space contains signal structures with arbitrary number of realizations. This simplification is based on a generalized concavification methodology developed in Lemma 1 of Zhong (2017a). The original concavification methodology is an application of Carathéodory theorem on graph of objective function on belief space.\(^{10}\) Equation (7) involves an additional term $h_{dt}(H(\mu) - \sum p_i H(v_i))$, which is not linear in expectation but fits in the generalized concavification method. Finally, Proposition 2 of Zhong (2017a) applies the standard contract mapping to guarantee a unique solution to Equation (7).

**Lemma 3.** With Assumption A, $\forall dt > 0$, there exists unique $V_{dt} \in C(\Delta X)$ solving Equation (7) and the supremum is attained.

4.3. Convergence and verification theorem

The following figure illustrates the roadmap for proving Theorem 1: the red arrow on the left. The discrete-time problem has value function $V_{dt}$ and corresponding Bellman equation Equation (7) (the double arrow on the right, proved in Lemma 2). I show that $V_{dt}$ converges to the continuous-time optimal control value $V$ (the arrow on the top, proved in Lemma 1). In the next lemma, I show that solution to continuous-time Bellman Equation (4) is the limit of solution to Equation (7) (the arrow on the bottom, to be proved in Lemma 4). Therefore, the value function solving Equation (4) is the solution to the continuous-time stochastic control problem Equation (1).

\(^{10}\) see Aumann et al. (1995) and Kamenica and Gentzkow (2009))
Lemma 4. Assume $H$ is strictly concave and $C^{(2)}$ on interior beliefs, and Assumption A and Assumption 1 are satisfied. Suppose $V(\mu) \in C^{(1)}$ is a solution to Equation (4). Then $V_{dt} \xrightarrow{dt \to 0} V$.

Lemma 4 proves that whenever Equation (4) has a solution, this solution is unique and coincides with the limit of solution to discrete-time problem Equation (7). Verification theorem Theorem 1 is a direct corollary of Lemmas 1, 2 and 4.

5. Optimal information acquisition

In this section I prove existence and fully characterize the value and policy functions of the continuous time HJB Equation (4) with binary states and two forms of flow cost function: a hard cap or a smooth convex function. In both cases, the optimal strategy is confirmatory Poisson learning, namely signal arrival induces Poisson jumps of belief towards more likely state. The optimal stopping time is immediately after signal arrival, and absence of signal is followed by searching for more precise signals. Optimal information cost decreases overtime when cost is flexible. Then in Section 5.2 I discuss the key trade-offs in the optimization problem and provide intuition for the optimal strategy. In Section 5.3 I show a sketched constructive proof for existence and the characterization of optimal strategy.

First of all, I introduce the assumptions which make the problem mathematically tractable:

Assumption 2.

1. (Binary states): $|X| = 2$.
2. (Positive payoff): $\forall \mu \in [0,1], \ F(\mu) > 0$.
3. (Uncertainty measure): $H''(\mu) < 0$ and locally Lipschitz on $(0,1), \lim_{\mu \to 0,1} |H'(\mu)| = \infty$.

Assumption 2 contains three parts. First, I restrict the state space to be binary. Therefore belief space is one dimensional and I can apply tools from differential equations to construct a candidate solution. Although existence of solution technically relies on the binary state assumption, most of the characterization results generalize to general state spaces as is discussed in Section 6.5. Second, I assume that utility from decision making is strictly positive so that “delay forever” is strictly suboptimal. In fact property 2 is without loss of generality in the sense that we can always add a dummy “outside action” that gives utility close to zero. Third, I assume that $H$ is sufficiently smooth, strictly convex (which rules out free information) and satisfies an Inada condition (which guarantees non-degenerate stopping region).

5.1. Main characterization theorem

Theorem 1 states that to characterize $V(\mu)$, it is sufficient to find a smooth solution to HJB Equation (4). I prove the existence of such a solution and provide characterization under either Assumption 1-a or Assumption 1-b, two slightly stronger variants of Assumption 1.
Assumption 1-a (Capacity constraint). There exists $c$ s.t. $h(I) = \begin{cases} 0 & \text{when } I \leq c \\ +\infty & \text{when } I > c \end{cases}$

Assumption 1-a restricts the cost function $h$ to be a hard cap: information is essentially free when its measure is below flow capacity $c$ and infinitely costly when it exceeds this capacity. This condition forces the DM to smooth his information acquisition process over time.

Theorem 2. Given Assumptions 1-a and 2, there exists quasi-convex value function $V \in C^{(1)}(0, 1)$ solving Equation (4). Let $E = \{\mu \in [0, 1] | V(\mu) > F(\mu)\}$ be the experimentation region, there exists policy function $v : E \rightarrow [0, 1]$ s.t.:

$$\rho V(\mu) = -c \frac{F(v(\mu)) - V(\mu) - V'(\mu)(v(\mu) - \mu)}{H(v(\mu)) - H(\mu) - H'(\mu)(v(\mu) - \mu)}$$

where $v(\mu)$ is unique a.e. and satisfies following properties:

1. $\exists \mu^* \in \arg\min V$ s.t. $\mu > \mu^* \implies v(\mu) > \mu$ and $\mu < \mu^* \implies v(\mu) < \mu$.
2. $v(\mu) \in E^C$. (successful experiment lands in stopping region)
3. $v(\mu)$ is strictly decreasing on each interval of $E \cap [0, \mu^*)$ and $E \cap [\mu^*, 1]$.
4. $\rho V(\mu) > -c \frac{V''(\mu)}{H'(\mu)} \forall \mu \in E \setminus \mu^*$.

Theorem 2 proves existence of solution to Equation (4) and characterizes the optimal policy function. The theorem first states that the optimal value function is implemented by a Poisson signal, i.e. seeking a breakthrough that drives belief to jump to $v(\mu)$. Property 1 says that the optimal signal is confirmatory: when $\mu > \mu^*$, the DM holds high prior belief for state 1 and she acquires information that induces even higher posterior belief. Vice versa for $\mu < \mu^*$. Conditional on receiving no signal, the DM’s belief drifts towards $\mu^*$. Property 2 says that the image of $v$ is always in the stopping region. In other words, the optimal stopping time is exactly the signal arrival time. Property 3 says that when the prior belief is drifting towards $\mu^*$, the optimal posterior belief induced by signal is moving against $\mu^*$, i.e. of increasing precision. Since Assumption 1-a means that the total informativeness of signal is bounded, a signal of increasing precision is achieved at the expense of decreasing frequency of observing a signal. Finally, the optimal policy is essentially unique: Gaussian signal is almost never optimal (Property 4) and optimal Poisson signal is almost everywhere unique.

Assumption 1-b (Convex flow cost). $h : \mathbb{R}^+ \rightarrow \mathbb{R}^+$ is $C^{(2)}$ smooth, $h(0) = 0$, $h''(c) > 0$, $\lim_{c \rightarrow \infty} h'(c) = \infty$.

Assumption 1-b restricts the cost function $h$ to be $C^{(2)}$ smooth and strictly convex: acquiring additional unit of information is of strictly increasing marginal cost. The Inada condition in Assumption 1 is still retained. If we replace Assumption 1-a with Assumption 1-b, we have the following characterization theorem:

Theorem 3. Given Assumptions 1-b and 2, there exists quasi-convex value function $V \in C^{(1)}(0, 1)$ solving Equation (4). Let $E = \{\mu \in [0, 1] | V(\mu) > F(\mu)\}$ be the experimentation
region, there \( \exists \) policy function \( \nu : E \rightarrow [0,1], I \in C^{(1)}(E) \) s.t.
\[
\rho V(\mu) = -I(\mu) \cdot \frac{F(\nu(\mu)) - V(\mu) - V'(\mu)(\nu(\mu) - \mu)}{H(\nu(\mu)) - H(\mu) - H'(\mu)(\nu(\mu) - \mu)} - h(I(\mu))
\]
where \( \nu \) and \( I \) are unique a.e. and satisfy the following properties:

1. \( \exists \mu^* \in \arg \min V \text{ s.t. } \mu > \mu^* \implies \nu(\mu) > \mu \text{ and } \mu < \mu^* \implies \nu(\mu) < \mu. \)
2. \( \nu(\mu) \in E^C. \)
3. \( \nu(\mu) \) is strictly decreasing on each interval of \( E \cap [0,\mu^*) \) and \( E \cap (\mu^*,1]. \)
4. \( \rho V(\mu) > \max_{\sigma} \frac{1}{2}\sigma^2 V''(\mu) - h(-\frac{1}{2}\sigma^2 H''(\mu)) \forall \mu \in E \setminus \mu^*. \)
5. \( I(\mu) \) is isomorphic to \( V(\mu) \).

Other than property 5, the optimal strategy shares the same set of properties as in Theorem 2. The optimal value function can be achieved through Poisson signals. Optimal stopping time is arrival time of signals. The unique optimal signal takes a form of confirmatory evidence that arrives at increasing precision and decreasing frequency conditional on continuation. Property 5 states that the optimal intensity of experimentation (parametrized by informativeness measure \( I \)) is higher when continuation value is higher. Since the belief process is drifts downward value function conditional on continuation, this means that the DM invests less in information acquisition when time goes.

The intuition for property 5 is already discussed in Moscarini and Smith (2001). The marginal gain from experimentation is proportional to the continuation value while marginal cost is increasing in \( I \). Therefore, the optimal cost is isomorphic to the value function. This property is called “value-level monotonicity” in Moscarini and Smith (2001), where level (flow variance of diffusion process) is a parameter for both intensity and precision of experimentation. My analysis identifies this intuition separately from another important trade-off between signal precision and frequency. I refer to property 5 as “value-intensity monotonicity” in this paper.

Examples

In this section, I first provide a minimal working example that illustrates Theorem 3 in Example 5.1. Then I provide some supplementary examples showing a rich set of results predicted in my framework, including multiple phases experimentation in Example 5.2, and learning from one-sided search in Example 5.3.

**Example 5.1.** Consider the problem studied in Section 3.1. \( F(\mu) = \max\{2\mu - 1, 1 - 2\mu\}, \)
\[
H(\mu) = -\mu \log(\mu) - (1 - \mu) \log(1 - \mu), \rho = 1, h(I) = \frac{1}{2}I^2.
\]
Now no parametric assumption is placed on belief process.

The solution is presented in Figures 2 and 3. In Figure 2-(a), dashed lines depicts \( F(\mu) \), blue curve depicts \( V(\mu) \). The blue shaded region is experimentation region \( E \). Figure 2-(b) shows the optimal posterior \( \nu(\mu) \) as a function of prior. As stated in Theorem 2, the policy function is piecewise smooth and decreasing. The three arrows in

---

11 Noticing that given \( \nu \), choosing \( I \) and \( p \) are equivalent. They uniquely pin down each other according to equation \( I(\mu) = p(\mu)(-H(\nu(\mu)) + H(\mu) + H'(\mu)(\nu(\mu) - \mu)). \)
**Figure 2**-(a) depict optimal strategies prescribed at three different priors. The arrows start at priors and point to optimal posteriors. The blue curve in **Figure 2**-(c) shows optimal intensity \( I(\mu) \) as a function of prior. It is easy to see that \( I(\mu) \) is isomorphic to \( V(\mu) \) in the experimentation region.

**Figure 3** illustrate the dynamics of optimal policy. **Figure 3**-(a) shows the optimal belief process. Conditional on no signal arrival, posterior belief drifts towards critical belief level \( \mu^* \). It is clear that in this example, there are two phases of experimentation (represented by colors of shaded region in **Figure 3**). In the first phase (blue region), DM seeks a Poisson signal to confirm one action. As time goes, the signal precision is increasing while signal frequency and learning intensity are decreasing (as in **Figure 3**-(b)&(c)). Eventually, the DM switches to the second phase (grey region). In the second phase, she seeks two signals confirming each state in a balanced way such that before any signal arrives her posterior belief is stationary.

Now it is clear why the three simple learning strategies studied in **Section 3.1** are optimal in the corresponding regions. It is an approximation to the full solution in Example 5.1. The confirmatory signal in **Section 3.1** is non-flexible, so it is optimal when the preference is not extreme. When belief is close to the boundary, optimal Poisson signal should be short jump with high frequency — approximated by Gaussian learning. When belief is close to \( \mu^* \), optimal Poisson signal should be low cost, and jumping to each side does not differ very much — approximated by a perfect revealing signal.

**Example 5.2** (Multiple phases). **Figure 4** depicts an example with four actions, whose expected payoffs are represented by the four dashed lines in **Figure 4**-(a). The two blue dashed lines are called riskier actions and the two red dashed lines are called safer actions. The upper envelope of the four lines is \( F(\mu) \). The experimentation region
now contains three intervals. Looking the middle interval, in red regions, the DM has more extreme belief and searches for a safer action (red arrow). In blue region, the DM has more ambiguous belief and searches for a riskier action (blue arrow). Figure 4-(c) depicts optimal belief process with prior belief in the red region. Experimentation now has three phases, the DM searches for a safer action in phase 1, searches for a riskier action in phase 2 and searches in a balanced way in phase 3.

**Example 5.3** (One-sided search). Figure 5 depicts an example with only one-sided search. In this example, state 1 dominates state 0 for both action. By property 1, \( \mu^* \) can only be 0. Figure 5-(b) shows that optimal strategy is to search for a Poisson signal inducing a posterior belief higher than the prior. Figure 5-(c) shows that in this example, there is only one phase. If no signal arrives before belief drops to the critical belief, it is optimal for the DM to stop learning and choose the safe action.

This example illustrates the precise definition for confirmatory evidence: optimal belief jump is always in the direction value function is increasing — the myopically optimal direction.

**5.2. Proof methodology and key intuitions**

In Section 3, I introduce four kinds of trade-offs. Now I discuss them in details and illustrate how these trade-offs leads to the properties of optimal strategy in Theorems 2 and 3. I first derive a geometric characterization for optimal policy in Section 5.2.1. Then I discuss how key tradeoffs are represented by the geometric characterization and provide intuition for the optimal policy. In Section 5.2.2 I show a sketched proof for Theorem 2.
5.2.1. Key trade-offs and geometric representation

In order to gain intuition it is useful to conduct a thought experiment. Fix a value function \( V \) and consider a simplified optimization problem:

\[
\sup_{p \geq 0, v} p \left( V(v) - V(\mu) - V'(\mu)(v - \mu) \right) - h(p(H(\mu) - H(v) + H'(\mu)(v - \mu)))
\]  

Equation (8) is more restrictive than Equation (4). I assume that the DM acquires only a Poisson signal. Let us ignore Gaussian signal for a moment. Define:

\[
\begin{align*}
U(\mu, v) &= V(v) - V(\mu) - V'(\mu)(v - \mu) \\
J(\mu, v) &= H(\mu) - H(v) + H'(\mu)(v - \mu)
\end{align*}
\]

The interpretation of \( U(\mu, v) \) is the flow value per unit arrival rate from a Poisson signal with posterior \( v \). Similarly, \( J(\mu, v) \) is the flow uncertainty reduction per unit arrival rate from the Poisson signal. Then Equation (8) can be rewritten as:

\[
\sup_{p \geq 0, v} p \cdot U(\mu, v) - h(p \cdot J(\mu, v))
\]

The problem is separable in choosing \( I \) and \( v \). The solution \((v^*, I^*)\) is characterized by:

\[
\begin{align*}
v^* &\in \arg \max_v \frac{U(\mu, v)}{J(\mu, v)} \\
h'(I^*) &= \max_v \frac{U(\mu, v)}{J(\mu, v)}
\end{align*}
\]

The optimal posterior \( v^* \) maximizes \( \frac{U(\mu, v)}{J(\mu, v)} \) — the value to uncertainty reduction ratio. Let \( \lambda = h'(I^*) = \max_v \frac{U(\mu, v)}{J(\mu, v)} \), then \( U(\mu, v) \leq \lambda J(\mu, v) \) and equality holds at \( v^* \).\(^{12}\)

Define \( G(\mu) = V(\mu) + \lambda H(\mu) \). I call \( G(\mu) \) the gross value function. Then definition of \( U \) and \( V \) implies: \( U(\mu, v) - \lambda J(\mu, v) = G(v) - G(\mu) - G'(\mu)(v - \mu) \). Hence, gross value function has the following property:

\[
\begin{align*}
G(v) &\leq G(\mu) + G'(\mu)(v - \mu) \quad \forall v \in [0, 1] \\
G(v^*) &= G(\mu) + G'(\mu)(v^* - \mu)
\end{align*}
\]

Equation (9) states that \( G(v) \) is everywhere (weakly) below the tangent line of \( G \) at \( \mu \), except that \( G(\mu) \) and \( G(v^*) \) touch the tangent line. The tangent line is linear (hence concave), so it weakly dominates \( G \)’s upper concave hull \( \text{co}(G) \). Therefore, \( G(\mu) = \text{co}(G)(\mu) \) and \( G(v^*) = \text{co}(G)(\mu^*) \). See Figure 6 for the graphical illustration. Figure 6-(a) and Figure 6-(b) depict value function \( V \) and uncertainty measure \( H \) respectively. Figure 6-(c) depicts the gross value function \( G = V + \lambda H \) where \( \lambda \) is calculated for prior \( \mu \). As I have discussed, \( G \) touches its upper concave hull at both \( \mu \) and \( v^* \). When

\(^{12}\)With Assumption 1-a, \( I^* = c \) and \( \lambda = \max \frac{U(\mu, c)}{J(\mu, c)} \) is the Lagrangian multiplier for constraint \( I \leq c \).

19
\( \nu^* \) is unique, \( \mu \) and \( \nu^* \) are the two boundary points of the concavified region (the set on which \( G < \text{co}(G) \)).

Equation (9) is called a concavification characterization as it is an analog to the concavification method in Bayesian persuasion problems. The difference is that in a Bayesian persuasion problem, the boundary points of a concavified region are optimal posteriors, while in current problem the prior is also on the boundary of a concavified region. This property has clear economic meaning. \( G \) is called the gross value function because it integrates value function \( V \) and uncertainty measure \( H \) using marginal cost level \( l \). \( l \) is the multiplier capturing the marginal effect of reducing uncertainty on flow cost. So solving:

\[
\sup_{p \geq 0, \nu} p(G(\nu) - G(\mu) - G'(\nu)(\nu - \mu))
\]

is equivalent to solving Equation (8). Whether Equation (10) yields a positive payoff depends on whether \( G(\mu) < \text{co}(G)(\mu) \). Suppose \( G(\mu) < \text{co}(G)(\mu) \) then there is a strictly positive gain from information and Equation (10) is strictly positive. However, Equation (10) is linear in the signal arrival rate \( p \). As a result the DM has incentive to increase \( p \). Increasing \( p \) drives up marginal cost \( \frac{d}{d\nu} \). So when the optimum is reached, \( \frac{d}{d\nu} \) (or \( \lambda \)) must be such that solving Equation (10) yields exactly zero utility: \( G(\mu) = \text{co}(G)(\mu) \). This characterization illustrates that in the continuous time limit, information is smoothed such that only infinitesimal amount of uncertainty is reduced at every instant of time.

Now suppose that the HJB is satisfied, i.e. Equation (8) equals the flow discounting loss \( \rho V(\mu) \). Then applying \( I^* = p^* \cdot f(\mu, \nu^*) \) and \( h'(I^*) = \frac{U(\mu, \nu^*)}{f(\mu, \nu^*)} \) to the HJB implies:

\[
\rho V(\mu) = p^* \cdot U(\mu, \nu^*) - h(p^* \cdot f(\mu, \nu^*))
\implies \rho V(\mu) = I^* h'(I^*) - h(I^*)
\]

Combining Equation (9) and Equation (11) pins down the value function \( V \) and corresponding strategies \( p, \nu \).\(^{13}\) Now I can analyze key trade-offs in dynamic information acquisition problem by studying Equations (9) and (11).

\(^{13}\) With Assumption 1-a, \( h(I^*) = 0 \) and \( I^* = c \). So \( \rho V(\mu) = \lambda c \).
1. Utility gain v.s. information cost

Equation (11) illustrates the utility gain v.s. information cost trade-off. Since $h$ is a convex function, $Ih'(I) - h(I)$ is increasing in $I^{14}$. That is to say, the optimal flow informativeness measure $I$ is isomorphic in continuation value $V(\mu)$. This property is exactly the "value-intensity monotonicity" I introduced in Section 5.1.

The intuition for this property is quite simple. The marginal cost of increasing informativeness of the signal proportionately is $Ih'(I)$. The marginal gain is from increasing arrival rates proportionately (keeping the signal precision fixed as in envelope theorem). Increasing arrival rate by unit proportion reduces waiting time by same proportion. So marginal gain from increasing $I$ by unit proportion is discount $\rho V$ plus cost $h(I)$. At the optimum, marginal cost equates marginal gain, therefore we get Equation (11) and flow informativeness is monotonic in value function.

If we consider the case with Assumption 1-a, then $\lambda$ in Equation (9) is replaced by shadow cost of increasing informativeness (see Footnotes 12 and 13). And Equation (11) can be written as $\rho V(\mu) = c\lambda$. Although in this case intensity is fixed, there is a monotonicity between shadow cost and value function.

To sum up, by studying the utility gain v.s. information cost trade-off, I established a monotonicity between shadow/marginal cost $\lambda$ and continuation value $V(\mu)$. (I refer to both of them as "value-intensity monotonicity" for notational simplicity). Now that I characterized $\lambda$, we can move on and focus on Equation (9).

2. Precision v.s. frequency

A novel trade-off characterized by Equation (9) is precision v.s. frequency trade-off. The value-intensity monotonicity determines $I$ from value function. Now the DM allocates total informativeness $I$ into precision (parametrized by size of belief jumps) and frequency (parametrized by arrival rate of jumps). Equation (9) suggests that optimal signal precision can be solved by concavifying gross value function $G(\mu)$. In this section, I illustrate how this trade-off changes at different prior and explain the intuition behind it.

Figure 7: Precision-frequency trade-off

Figure 7 shows how varying $\lambda$ affects the optimal size of jump. In Figure 7-(a) the blue curve is $G(\mu)$ and the dashed curve is $co(G)$. I call the blue region where $G(\mu) < co(G)(\mu)$ the concavified region and the white region where $G(\mu) = co(G)(\mu)$

---

\[^{14}\frac{d}{dI}(Ih'(I) - h(I)) = Ih''(I) \geq 0\]
the *globally concave region*. The prior $\mu$ and optimal posterior $\nu$ are on the boundary of a concavified region. Now consider $G_1 = V + \lambda_1 H$, where $\lambda_1 > \lambda$. Figure 7-(b) depicts both $G$ (the dashed curve) and $G_1$ (the blue curve). Since $G_1$ is $G$ plus a strictly concave function, any belief in the globally concave region of $G$ is still in the globally concave region of $G_1$. As a result with a larger $\lambda$ the white region is expanding and the blue region is shrinking (see Figure 7-(c)). So prior and optimal posterior are closer to each other. Recall that $\lambda$ is monotonic $V$. This means with higher continuation value, the DM is more willing to choose signal inducing shorter belief jump.

The intuition for this property is as follows. When the DM is more sure about the state, continuation value is higher, hence utility loss from discounting is higher. The DM wants to receive a signal more frequently to enjoy the high value sooner. Therefore, the value of frequency outweighs the value of precision and the optimal strategy is to give up some precision to get higher arrival frequency. In this analysis, continuation value is isomorphic to $\lambda$, which controls the shape of $G$. Relative value from signal precision (to value from frequency) is determined by global concavity of gross value function. So the analysis presented earlier by Figure 7 exactly illustrates the intuition.

**Confirming v.s. contradicting:** The analysis above pins down the absolute size of optimal belief jump. To pin down the optimal posterior, it remains to be seen which direction of jump is optimal. Now I show that the precision-frequency trade-off also implies the optimality of confirmatory learning.

Let us hypothetically consider a belief $\mu$ at which jumping towards right is optimal (weakly). In both panels of Figure 8, $\mu$ is the prior and $\nu_L,\nu_R$ are optimal posteriors on each side of $\mu$ respectively. So jumping to $\nu_R$ (the black arrow) is better than jumping to $\nu_L$ (the dashed black arrow). Let $V$ be increasing around $\mu$. Now consider the DM’s incentive at $\mu_1$ slightly larger than $\mu$ (on Figure 8-(a)). By an envelop theorem argument, $\nu_L$ and $\nu_R$ are optimal posteriors also for $\mu_1$ on each side. To pin down the optimal posterior for $\mu_1$, we just need to compare $\nu_L$ and $\nu_R$. Since $\mu_1 > \mu$, $\nu_R$ is closer to prior, while $\nu_L$ is further away from prior. Meanwhile, $V(\mu_1) > V(\mu)$ implies that the DM prefers frequency to precision more with belief $\mu_1$. Therefore, $\nu_R$ is strictly preferred at $\mu_1$. Consider $\mu_2$ slightly smaller than $\mu$ (on Figure 8-(b)). Similar analysis shows that now size of jumping to $\nu_R$ is larger, and DM prefers precision more with belief $\mu_2$. So $\nu_R$ is also strictly optimal for $\mu_2$. In this analysis, jumping towards

![Figure 8: Confirmatory v.s. contradictory](image)

the direction that value function is increasing means the signal is confirmatory. When value function is quasi-convex, it is equivalent to property 1 of Theorems 2 and 3. Therefore, the precision-frequency trade-off implies that the incentive for confirmatory learning is self-enforcing.

22
3. Poisson vs. Gaussian

So far I ignored the possibility of Gaussian signals. In fact, they are implicitly modeled in Equation (9). Consider the optimization w.r.t. Gaussian signals:

\[
\begin{align*}
\sup_{\sigma} \sigma^2 V''(\mu) - h(-\sigma^2 H''(\mu)) \\
\Rightarrow \text{FOC} : V''(\mu) + \lambda H''(\mu) = 0 \\
\iff G''(\mu) = 0
\end{align*}
\]

where \( \lambda = h'(-\sigma^2 H''(\mu)) \) with Assumption 1-a or \( \lambda = \frac{\beta}{\tau} V(\mu) \) with Assumption 1-b. Comparing Equations (9) and (12), it is not difficult to notice that Equation (12) is exactly a limit of Equation (9) when optimal posterior \( \nu \) is converging to prior \( \mu \). This is intuitive since Gaussian signal can be approximated by a Poisson signal with very low precision and high arrival rate.

The comparison between Gaussian signal and Poisson signal is effectively the comparison between a special imprecise Poisson signal and other Poisson signals. So this trade-off is a special case of the precision v.s. frequency trade-off. Gaussian signal is optimal when the DM wants to trade off almost all precision for frequency — a slightly less patient DM is willing to avoid any waiting and stop immediately, while a slightly more patient DM is willing to wait for a Poisson signal with positive jump. Therefore, Gaussian signal is optimal only on boundaries of experimentation regions. Given this intuition, one could imagine that Gaussian signal is generically suboptimal except for special cases where there is no precision-frequency trade-off at all. Since the preference between precision and frequency depends on the loss from delaying, the trade-off diminishes only when the DM does not discount future. This intuition is confirmed in a no-discounting special case in Section 6.3 as well as in the main model of Hébert and Woodford (2016).

4. Continuing vs. stopping

Now consider the optimal stopping rule. Theorems 2 and 3 states that optimal posteriors are in the stopping region, i.e., it is suboptimal to do repeated jumps. Now I show that repeated jumps can be improved by a direct jump. Let \( \nu \) be the optimal posterior for prior \( \mu \) (see Figure 9). Then Equation (9) and Equation (11) jointly implies that \( \frac{U_0}{J_0} = \frac{U'_0}{J'_0} = \lambda(\mu) \).
Hypothetically imaging that at \( \nu \) it is optimal to continue and optimal posterior is \( \nu' \). Then \( \frac{U_1}{\lambda_1} = \lambda(\nu) \), and \( \lambda(\nu) > \lambda(\mu) \) by the confirmatory evidence property & value-intensity monotonicity. I want to show that this implies \( \frac{U(\mu, \nu')}{U(\mu, \nu')} = \frac{U_1 + U'_1}{\lambda_1 + \lambda'_1} > \lambda(\mu) \), i.e. jumping to posterior \( \nu' \) directly is strictly better than the two-step jump. By elementary geometry there exists \( \alpha \) s.t \( U'_1 = \alpha U_0 \) and \( J'_1 = \alpha J_0 \). Therefore, the value to uncertain reduction ratio \( \frac{U(\mu, \nu')}{J(\mu, \nu')} = \frac{U_1 + \alpha U_0}{\lambda_1 + \alpha \lambda_0} \) is a weighted average of \( \frac{U_0}{\lambda_0} \) and \( \frac{U_1}{\lambda_1} \), which is larger than \( \lambda(\mu) \).

Now the intuition for the stopping rule is clear. If we combine a two-step jump into a direct jump, then the flow utility gain is a weighted sum of that of the two jumps. The flow uncertainty reduction is exactly the same weighted sum of that of the two jumps. Therefore, the net value from a direct jump is a weighted average of the net values from each jumps. Sequentially jumping to higher values is dominated by directly jumping to the highest value.

**Remark.**

The intuition behind the value-intensity monotonicity is purely driven by convexity of cost function \( h \) and is obviously independent to the formulation of the information measure. The intuition behind the optimality of Poisson signal over Gaussian signal is to use the precision-frequency trade-off to compare generic Poisson signal with extremely imprecise Poisson signal. It does not depend on the exact form of \( I \) (or posterior separability). I generalize the optimality of Poisson with generic cost function in Theorem 6, Section 6.1. I also discuss confirmatory evidence and immediate stopping properties with generic cost function in Section 6.1.

The precision-frequency trade-off does not depend on the size of state space either. I confirm this with a general characterization result with more states Theorem 10 in Section 6.5. However, the existence of solution (which has not been shown) relies on the binary state assumption. A constructive proof for the binary state case based on ODE theory is introduced in Section 5.2.2.

Our discussion so far does not rely on the exact form of \( \lambda \). The qualitative properties of all these trade-offs depend only on monotonicity of \( \lambda \) in continuation value, which is true with both Assumptions 1-a and 1-b. So when I introduce the sketched proof, I discuss only Theorem 2 and it extends to Theorem 3.

5.2.2. **Sketched proof**

I prove Theorem 2 by construction and verification. I conjecture that the optimal policy to Equation (4) takes the form in Theorem 2: a single confirmatory signal associated with immediate action. I first construct \( V(\mu) \) and \( V(\nu) \) with three steps:

- **Step 1.** Determine \( \mu^* \). \( \mu^* \) can be calculated as essentially the unique belief at which \( V'(\mu^*) = 0 \) and searching for posteriors at either side of \( \mu^* \) are equally good. Notic-

---

\[ \text{[15] See Figure 9.} \quad \frac{U_1}{\lambda_1} = \frac{U_0}{\lambda_0} = \lambda(\mu) \implies \frac{U_1}{\lambda_1} = \lambda(\mu). \quad \text{Hence,} \quad \frac{U_1}{\lambda_1} = \frac{U_0}{\lambda_0}. \quad \text{I assume the ratio to be} \ a. \]
ing that if $V'(\mu^*) = 0$, HJB implies:

$$V(\mu^*) = \frac{F(v)}{1 + \frac{\partial}{\partial \mu} J(\mu, v)}$$

$\mu^*$ is solved as the unique belief at which:

$$\sup_{v \leq \mu^*} \frac{F(v)}{1 + \frac{\partial}{\partial \mu} J(\mu, v)} = \sup_{v \geq \mu^*} \frac{F(v)}{1 + \frac{\partial}{\partial \mu} J(\mu, v)}$$

$V(\mu^*), V'(\mu^*)$ and $v(\mu^*)$ are all pinned down correspondingly.

- **Step 2.** Search for optimal posterior for a fixed action. Let $a$ be the optimal action for optimal posterior $v \geq \mu^*$ solved from step 1. Let $F_a(\mu) = E[\mu(a, x)]$. Now solve for optimal posterior $v$ on $F_a$:

$$\rho V(\mu) = \max_{v \geq \mu} -c \frac{F_a(v) - V(\mu) - V'(\mu)(v - \mu)}{H(v) - H(\mu) - H'(\mu)(v - \mu)}$$

The primitives in objective function are all sufficiently smooth in $v$. Then, first order condition w.r.t. $v$ yields a well behaved first order ordinary differential equation characterizing $v(\mu)$ with initial condition $v(\mu^*)$. Therefore we can solve for optimal policy $v$ and calculate value $V(\mu)$ accordingly for $\mu \geq \mu^*$. $V(\mu)$ and $v(\mu)$ for all $\mu \leq \mu^*$ is solved by a symmetric process.

- **Step 3.** Update the value function w.r.t. all alternative actions and smoothly paste the solved value function piece by piece. This step starts from solving the ODE defined in step 2 at $\mu^*$. Then I extend the value function towards $\mu = \{0, 1\}$. Whenever I reach a belief at which two actions yield same payoff, I setup a new ODE with the new action. This process continues until the calculated value function $V(\mu)$ smoothly pastes to $F(\mu)$. This procedure generates a quasi-convex value function (minimized at $\mu^*$).

Solving the ODE characterizing $v(\mu)$ directly implies monotonicity of $v(\mu)$ in each connected continuation region. Now I need to verify the optimality of the constructed strategy. The verification takes three steps, each ruling out repeated jumps, contradictory evidence and Gaussian signals respectively. The intuitions for suboptimality of these three alternative strategies are explained in Section 5.2 already. The formal proof is relegated to Section 7.5.

5.3. **Convergence of policy**

Although I obtain an essentially unique optimal policy for Equation (4), Theorem 1 does not rule out other possible optimal policies for Equation (1). To get behavior predictions from my model, I show the convergence of discrete time optimal policy towards the solutions defined in Theorems 2 and 3. I define a modified version of Lévy distance to characterize difference between policy functions:
Definition 1 (Lévy metric). Let $F, G : [0, 1] \to [0, 1]$ be two correspondence. Define the Lévy metric between them to be:

$$d_L(F, G) := \inf \left\{ \epsilon > 0 \left| \inf_{|y-x| \leq \epsilon} d_H(F(x), G(y)) \leq \epsilon \forall x \in [0, 1] \right. \right\}$$

where $d_H$ is standard Hausdorff metric on $\mathbb{R}^1$.

$d_L(F, G) = \epsilon$ means the $\forall \mu \in [0, 1], y \in F(\mu)$, there must be some $\mu'$ in $\epsilon$-neighbourhood of $\mu$ such that $y$ is in $\epsilon$-neighbourhood of $G(\mu')$. When $G$ is continuous at $\mu$, and $\epsilon$ is sufficiently small, it simply states that image of $F$ and $G$ at $\mu$ are close to each other (measured by $d_H$). If $d_L(F, G) = 0$ then $F$ and $G$ are identical.

Theorem 4 (Convergence of policy). With Assumptions 1-a and 2 or Assumptions 1-b and 2 satisfied. Let $v(\mu)$ be policy correspondence solving Equation (4). Let $N(\mu) = \{\mu\} \cup v(\mu)$. Let $N_{dt}(\mu)$ be the support of optimal posteriors solving Equation (7). Then:

$$\lim_{dt \to 0} d_L(N, N_{dt}) = 0$$

Theorem 4 states that the graph of policy function of discrete-time problem Equation (5) converges to the graph of the continuous solution defined in Theorems 2 and 3. The convergence is illustrated in Figure 10. I calculate the discrete-time policy function using parameters in Example 5.1. The red, blue and green lines represents optimal posteriors when $V_{dt} > F$ with $dt = 10^{-5}, 10^{-3}$ and $10^{-2}$. As is shown in the figure, when $dt \to 0$, one of optimal posterior is converging to prior, and the other optimal posterior is converging to the continuous time solution. The posterior converging to prior captures a drift term and the other posterior captures a Poisson jump in the limit.

![Figure 10: Convergence of policy function](image-url)
6. Discussions

In this section, I discuss in details the assumptions I make in the baseline model: Assumptions 1 and 2, and Assumption A for the discrete-time analysis. All assumptions I made can be categorized into three classes:

1. **Economic assumptions:** Posterior separability of information measure (Assumption A). Convexity of cost function (Assumption 1). Impatienences (positive ρ).

2. **Restrictive assumptions:** Finite actions and binary states (Assumption 2).

3. **Technical assumptions:** Smoothness and positiveness assumptions (Assumption 2).

The economic assumptions are crucial for my methodology and deserve in-depth discussion. In Section 6.1, I first show an axiomatic characterization for Assumption A to illustrate its economics meaning. Then I extend the continuous time problem to impose a general information measure on jump-diffusion processes and show that Poisson signal strictly dominates Gaussian signal except for non-generic cases. Finally I explain that immediate action and confirmatory learning properties are tightly tied to posterior separability. To illustrate the role of Assumption 1, I discuss the case when cost function is linear in Section 6.2 and show that without convexity, the optimal strategy is static. To illustrate the role of discounting, I discuss the case with no discounting but flow waiting cost in Section 6.3, and show that without discounting, the trade-off between precision and frequency diminishes and the dynamics of information become irrelevant.

Restrictive assumptions do restrict generality of my model. However, relaxing them does not fundamentally alter the key intuition and the key methodology generalizes. In Section 6.4, I relax the finite action assumption. I showed that problem with continuum of actions can be approximated well by adding actions, in the sense of both value function and policy function. In Section 6.5, I relax the binary state assumption. Although the constructive proof for existence no longer works with general state space, I showed that the properties in Theorem 2 extend to general state space. Technical assumptions do not restrict my model in a meaningful way so I will not discuss them.

6.1. Posterior separable measure

In this section, I first provide an axiom for Assumption A and extend my baseline model to generic flow information measures.

6.1.1. Axiom for Assumption A

**Theorem 5.** \( I(S; X|μ) \) is a non-negative informativeness measure. \( I \) satisfies Assumption A if and only if the following axiom holds:

**Axiom:** \( \forall μ, \forall \text{ information structure } S_1 \text{ and information structure } S_2|S_1 \text{ whose distribution is conditional on realization of } S_1: \)

\[
I((S_1, S_2); X|μ) = I(S_1; X|μ) + E[I(S_2; X|S_1, μ)]
\]

**Theorem 5** states that the *chain rule* (the name for a key property of mutual information in Cover and Thomas (2012)) is not only a necessary condition but also a
sufficient condition for posterior separability. Given any experiment, we can divide it into multiple stages of “smaller” experiments. This axiom requires that the total informativeness of this sequence of small experiments is "path-independent": it always equals to the informativeness of the compound experiment.

Given Theorem 5, Assumption A is essentially a consistency requirement on cost of compound experiments. It helps me throughout the whole analysis. First, separability of information measure establishes Lemma A.1. It helps me eliminate redundant information and inter-temporal complexity to establish equivalence between continuous time model and limit of discrete time model. Second, the methodology of concavifying “gross value function” is only possible when expected utility gain and information measure takes a consistent form. Finally, as I will show later the immediate action and confirmatory evidence results are closely tied to this assumption. However, this axiom seems to contradict another form of consistency: take any signal structure as an object, its cost should be prior independent. To accommodate prior-independence, and many other possible information measures, I study a problem with more general information measure (but simplified in other dimensions) in next section.

6.1.2. General informativeness measure

I setup a continuous time HJB equation with a generic information cost structure which imposes no specific link between prior and posterior. I want to show that one key feature—optimality of Poisson learning I identified in the baseline model is generic. Let \( J(\mu, v) \) be a sufficiently smooth function. Consider the following functional equation:

\[
\rho V(\mu) = \max \left\{ \rho F(\mu), \sup_{p,v,c^2} p(V(v) - V(\mu) - V'(\mu)(v - \mu)) + \frac{1}{2} \sigma^2 V''(\mu) \right\} 
\]

s.t. \( pJ(\mu,v) + \kappa(\mu,\sigma) \leq c \)  

(13)

The objective function of Equation (13) is exactly the same as that of Equation (4) with Assumption 1-a. I assume that DM chooses a combination of Poisson signal and Gaussian signal. The gain from information is same as before. I assume \( J(\mu, v) \) to be an arbitrary function which is both prior and posterior dependent. Cost of diffusion signal is \( \kappa(\mu,\sigma) \). We impose following assumptions on \( J(\mu, v) \) and \( \kappa(\mu,\sigma) \):

**Assumption 3.**
1. \( J \in C^4(0,1)^2 \).
2. \( \forall \mu \in (0,1), J(\mu,\mu) = J_v'(\mu,\mu) = 0, \) and \( J_{vv}(\mu,\mu) > 0 \).
3. \( \kappa(\mu,\sigma) = \frac{1}{2} \sigma^2 J_{vv}(\mu,\mu) \).

First \( J \) is assumed to be sufficiently smooth to eliminate technical difficulties. \( J(\mu,\mu) = 0 \) is an implication of “an uninformative Poisson signal is free”.\(^{16}\) \( J_v'(\mu,\mu) = 0 \) and \( J_{vv}(\mu,\mu) > 0 \) are implications of “any informative Poisson signal is costly”. Within

\(^{16}\) In this setup, \( J(\mu,\mu) = 0 \) is WLOG. If uninformative signal has strictly positive, we can always shift capacity constraint \( c \) to make cost 0.
this continuous time framework, the assumptions imposed on $J$ is without loss of generality. The crucial assumption is the third condition: $\kappa(\mu, \sigma) = \frac{1}{2} \sigma^2 J^n_{vv} (\mu, \mu)$. This assumption is essentially saying the cost functional is 'continuous' in the space of signal structures. Consider a Poisson signal $(p, v)$. When $v \to \mu$, the utility gain from learning this signal is:

$$p(V(v) - V(\mu) - V'(\mu)(v - \mu)) = p\left(\frac{1}{2} V''(\mu)(v - \mu)^2 + O|v - \mu|^3\right)$$

So $(p, v)$ is approximating a Gaussian signal with flow variance $p(v - \mu)^2$. Meanwhile, the cost of this signal is:

$$pJ(\mu, v) = p\left(J(\mu, \mu) + J'_v(\mu, \mu)(v - \mu) + \frac{1}{2} J''_v(\mu, \mu)(v - \mu)^2 + O(|v - \mu|^3)\right)$$

$$= \frac{1}{2} p(v - \mu)^2 J''_v(\mu, \mu) + pO(|v - \mu|^3)$$

Hence, if cost of Gaussian signal is consistent with cost of imprecise Poisson signal in the limit, $\kappa(\mu, \sigma) = \frac{1}{2} \sigma^2 J^n_{vv} (\mu, \mu)$.

**Theorem 6.** Given Assumption 3, suppose $V \in C^{(3)}(0,1)$ solves Equation (13) and let $L(\mu)$ be defined by:

$$L(\mu) = \frac{\rho}{c} J^n_{vv} (\mu, \mu)^2 - \frac{2J^{(3)}_{vv} (\mu, \mu)^2 + J^{(3)}_{vvv} (\mu, \mu) J^{(3)}_{vv} (\mu, \mu)}{J^n_{vv} (\mu, \mu)} + J^{(4)}_{vvv} (\mu, \mu) + J^{(4)}_{vvvv} (\mu, \mu)$$

Then in the open region: $D = \{ \mu | V(\mu) > F(\mu) \text{ and } L(\mu) \neq 0 \}$, the set of $\mu$ s.t.:

$$\rho V(\mu) = c\frac{V''(\mu)}{J^n_{vv} (\mu, \mu)}$$

is of zero measure.

The interpretation of Theorem 6 is that Poisson signal is almost always strictly superior than diffusion signal. In the experimentation region where $L(\mu) \neq 0$, $V(\mu)$ can be achieved by a diffusion signal only at a zero measure of points. $L(\mu) = 0$ is a partial differential equation on $J(\mu, v)$ in the diagonal of space. Therefore, the set of points that $L(\mu) = 0$ could contain interval only when $J(\mu, v)$ is locally solution to the PDE. Solution to a specific PDE is a non-generic set in the set of all functions satisfying Assumption 3. In this sense, for an arbitrary information measure $J(\mu, v)$, the optimal policy function contains diffusion signal almost nowhere.

A trivial sufficient condition for $L(\mu) \neq 0$ is invariance of $J^{(2)}_{vv} (\mu, v)$. In this case $L(\mu) = \frac{\rho}{c} J^n_{vv} (\mu, \mu)^2 > 0$ for sure. The first corollary of Theorem 6 states that when $J^n_{vv} (\mu, v)$ is approximately invariant, $D$ is almost the whole experimentation region. $\forall f \in C^{(1)}(0,1)^2$ define a norm: $\| f(\cdot) \|_\delta = \sup_{x \in [\delta,1-\delta]} \{ |f(x, x)|, \| \nabla f(x, x) \|_{L^2} \}$. 

29
Corollary 6.1. Given Assumption 3, suppose \( V \in C^3(0,1) \) solves Equation (13), then for any \( \delta > 0 \), there exists \( \epsilon \) s.t. if \( \left\| \frac{V^{(3)}}{V'} \right\|_{\delta} \leq \epsilon \):

\[
\left\{ \mu \in [\delta, 1 - \delta] \left\| \rho V(\mu) = c \frac{V''(\mu)}{\int_{\nu} V'(\mu, \mu)} \right\} \right
\]

is of zero measure.

Posterior separability implies \( \int_{VV} V(\mu, \nu) = 0 \). The condition in Corollary 6.1 states that \( \int_{\nu} V'(\mu, \nu) \) is approximately a constant over \( \mu \) for \( \nu \) close to \( \mu \). This result illustrates my analysis in Section 5.2.1. The comparison of Poisson and Gaussian signal relies only on local properties of \( J \). Another simple sufficient condition for \( L(\mu) \neq 0 \) is high impatience or low learning capacity:

Corollary 6.2. Given Assumption 3, suppose \( V \in C^3(0,1) \) solves Equation (13), then for any \( \delta > 0 \), there exists \( \Delta \) s.t. if \( \frac{\rho}{\epsilon} \geq \Delta \):

\[
\left\{ \mu \in [\delta, 1 - \delta] \left\| \rho V(\mu) = c \frac{V''(\mu)}{\int_{\nu} V'(\mu, \mu)} \right\} \right
\]

is of zero measure.

In Section 5.2.1, I show that the precision-frequency trade-off determines generates gap between Gaussian signals and Poisson signals. When \( \frac{\rho}{\epsilon} \) is high, the dependence of discount loss on continuation value is magnified. Hence, in this case Poisson is more likely to be optimal even though the information cost could be distorted in an arbitrary way that is in favor of Gaussian signals.

Although Poisson learning is generic optimal, immediate action and confirmatory evidence are not. They are implications of linear additivity of information measure on compound signal structures. Imagine a case in which signals of high precision are relatively cheap (say an extreme case where \( J(\mu, \nu) \) is truncated both below and above). Then, when prior is close to boundary of experimentation region, seeking for confirmatory evidence (with low precision and high frequency) results in very high cost, while seeking for a precise contradictory signal is quite cheap. Therefore, confirmatory evidence seeking will be dominated. In fact, this example shares same intuition with the findings in Che and Mierendorff (2016). In their setup, there is no cost on signal but total attention is bounded. When allocating more attention to signals revealing states of higher prior probability, DM is effectively increasing arrival rate of signal. Of course signals revealing higher prior states induces shorter jumps, thus less precise. So the precision-frequency trade-off will drive confirmatory evidence seeking in intermediate region. On boundary, since total arrival rate is bounded and signals are costless, it is as if contradictory signal is free but high frequency confirmatory signal is infinitely expensive. As a result contradictory evidence becomes optimal.

On the other hand, consider the immediate action property. Imagine a case in which signals of low precision is relatively cheap. Then, break a long jump of posterior into multiple short jumps might become profitable. Immediate action property is named single experiment property (SEP) in Che and Mierendorff (2016). In their paper, it is also documented that SEP is not a robust property in a generic Poisson learning model.
6.2. Linear flow cost

In this subsection, I study the case where the flow cost \( h(I) \) is a linear function.

**Assumption 1’ (Linear flow cost).** Function \( h \) is defined by \( h(I) = \lambda I, \lambda > 0 \).

As discussed in Assumption 1, the convexity of \( h(I) \) gives the DM incentive to smooth the acquisition of information. If \( h(I) \) is a linear function, then it is optimal to acquire all information and make decision immediately:

**Theorem 7.** Given Assumption 1’, suppose \( V(\mu) \) solves Equation (1), then:

\[
V(\mu) = \sup_{P \in \Delta^2(X)} E_P[F(v)] - \lambda E_P[H(\mu) - H(v)]
\]

The intuition for this result is simple. At any instant of time, suppose it is optimal to continue learning for positive amount of time. The value is discounted future value at next instant of time \((t + dt)\) less flow cost of information. Now consider moving the learning strategy at \( t + dt \) to current period. Then both future value at \( t + dt \) and cost are discounted \( dt \) less. If the net utility gain from learning at \( t + dt \) is non-negative, then this operation increases current utility by reducing waiting time.\(^{17}\) If the net utility gain from learning at \( t + dt \) is negative, then stopping learning immediately increases current utility. This operation can always be applied repeated and strictly improves the strategy until all information is acquired at period 0.\(^{18}\)

6.3. Linear discounting

As discussed in Section 5.2, discounting is one key factor driving all the dynamics. With exponential discounting, DM’s trade-off between arrival frequency and precision of signals changes according to continuation value. One can easily imagine that if we replace exponential discounting with linear discounting, i.e. DM pays a fixed per period waiting cost, time distribution of utility gain and loss does not matter for the DM anymore. DM’s problem is separable in two sub-problems: she first solves for the optimal information smoothing strategy, fixing total expected waiting time and aggregate information structure. Then she can solve for the optimal aggregate information. Consider the following problem:

\[
V(\mu) = \sup_{\langle \mu, \eta \rangle \in \mathcal{M}, \tau} E \left[ F(\mu_\tau) - m\tau - \int_0^\tau h(I_t)dt \right]
\]

**Theorem 8.** Given Assumption 1, suppose \( V(\mu) \) solves Equation (15), then:

\[
V(\mu) = \sup_{P \in \Delta^2(X), \lambda > 0} E_P[F(v)] - \frac{m + h(\lambda)}{\lambda} E_P[H(\mu) - H(v)]
\]

\(^{17}\) This step utilizes Assumption 1’, which implies that cost of a combined signal structure is sum of cost of each of them.

\(^{18}\) Strictly speaking, an immediate learning strategy is not admissible because its belief path is not cadlag. However, there always exists a way to implement a signal structure in arbitrarily short period of time, and the payoff approximates the immediate learning payoff.
Theorem 8 illustrates the two step method. The second step problem is exactly a static information acquisition problem, with $\frac{m + f(\lambda)}{\lambda}$ being marginal cost on information measure (same as in Caplin and Dean (2013) and Matejka and McKay (2014)). Parameter $\frac{m + f(\lambda)}{\lambda}$ measures per period cost of information from solving a dynamic information allocation problem.

A further generalization of time preference is in Zhong (2017b), which relates optimality of Poisson to risk loving on time dimension. It is shown in Zhong (2017b) that within the set of all decision time distributions induced by dynamic learning strategies implementing a same target information structure, all distributions have same expected decision time and Poisson learning generates the most dispersed (in MPS sense) decision time distribution. This result gives us a more complete picture: convex discounting generically generates the precision-frequency trade-off. And optimality of Poisson learning is a result of the trade-off. Only in a special case when there is not discounting, there is no such trade-off and all learning dynamics are essentially equivalent.

6.4. Continuum of actions

In this section, I extend my model to accommodate infinite actions (or even continuum of actions) in the underlying decision problem, i.e. $|A| = \infty$. Mathematically, the difference is that value from immediate action $F(\mu) = \sup_{a \in A} E[u(a, x)]$ is no-longer a piecewise linear function. There will be several technical problems arising from a continuum of actions. For example whether the supremum is achievable and whether $F$ has bounded subdifferentials. I impose the following assumption:

**Assumption 4.** $F(\mu) = \max_{a \in A} E[u(a, x)]$ has bounded subdifferentials.

Assumption 4 rules out two cases. First is that the supremum is not achievable. Second case is some optimal action being infinitely risky: the optimal action with belief approaching $x = 0$ has utility approaching $-\infty$ at state 1 (and similar case with states swapped). A sufficient condition for Assumption 4 will be:

**Assumption 4’.** $A$ is a compact set. $\forall x \in X, u(a, x) \in C(A) \cap TB(A)$.

The proof of Lemma 4 does not rely on the fact that $F(\mu)$ is piecewise linear. Actually the only necessary properties of $F(\mu)$ is boundedness and continuity in Lemma 3 that proves existence of solution to discrete time functional equation Equation (17). Therefore Assumption 4 guarantees that Lemma 3 and Lemma A.8 still hold when there is a continuum of actions. With Assumption 4, problem with continuum of actions can be approximated well by a sequence of problems with discrete actions. I first define the following notation: $\forall F$ satisfying Assumption 4, $\mathcal{V}_{dt}(F)$ is the unique solution of Equation (7) and $\mathcal{V}(F) = \lim_{dt \to 0} \mathcal{V}_{dt}(F)$.

**Lemma 5.** Given Assumptions 1 and 4, $\mathcal{V}$ is a Lipschitz continuous functional under $L_{\infty}$ norm.
Lemma 5 states that a problem with continuum of actions can be approximated well by a sequence of problems with discrete actions in the sense of value function convergence. Next, I push the convergence criteria further to convergence of policy function.

Theorem 9. Given Assumptions 1-a, 2 and 4, let \( \{F_n\} \) be a set of piecewise linear functions on \([0, 1]\) satisfying:

1. \( \|F_n - F\|_\infty \to 0; \)
2. \( \forall \mu \in [0, 1], \lim F_n'(\mu) = F'(\mu). \)

Then \( |V(F) - V(F_n)| \to 0 \) and:

1. \( V(F) \) solves Equation (4).
2. \( \forall \mu \text{ s.t. } V(\mu) > F(\mu), \) if each \( \nu_n \) is maximizer of \( V(F_n) \) and \( \nu = \lim_{n \to \infty} \nu_n \) exists, then \( \nu \) is the optimal posterior in Equation (4) at \( \mu \).

Theorem 9 states that to solve for a continuous time problem with a continuum of actions, one can simply use both value function and policy function from problem with finite actions to approximate. As long as the immediate action values \( F_n \) converge both uniformly in value and pointwise in first derivative, the optimal value functions have a uniform limit. The limit will solve Equation (4) and the optimal policy function will be the pointwise limit of policy functions for finite action problems.

![Figure 11: Approximation of a continuum of actions](image)

Figure 11 illustrates this approximation process. On both panels, only \( \mu \in [0.5, 1] \) is plotted. On the right panel, the thin black curve shows a smooth \( F(\mu) \) associated with continuum of actions. Since optimal policy only utilizes a subset of actions, I approximate the smooth function only locally as the upper envelope of dashed lines (each represents one action). The optimal value function with continuous actions is blue curve and the discrete action approximation is red curve. Left panel shows the approximation of policy function. The blue smooth curve is optimal policy of continuous action problem and the red curve with breaks is optimal policy of discrete action problem.
To approximate a smooth $F(\mu)$, one can simply add more and more actions to the finite action problem and use $F$’s supporting hyper planes to approximate it. Then the optimal policy function will have more and more breaks as optimal policy will involve more frequent jumping among actions. In the limit, as number of breaks grows to infinity, the size of breaks shrinks to zero and approaches a continuous policy function.

6.5. General state space

In this section, I extend state space to higher dimensional. The constructive proof for Theorems 2 and 3 relies on ODE theory to guarantee existence of solution. With a larger state space, construction of value function relies on existence of PDE. There is no general theory ensuring existence. Nevertheless, the verification part still works. In fact, the discussion in Section 5.2 seems to extend to higher dimensional spaces in a natural way. I formalize a partial characterization result in the section.

Let $n = |X|$. Consider value function $V(\mu)$ on $\Delta(X)$. Let $V(\mu) \in C\Delta(X)$ and $C^{(2)}$ smooth when $V(\mu) > F(\mu)$. Consider a functional equation on $V(\mu)$:

$$pV(\mu) = \max \left\{ \rho F(\mu), \max_{v,p,\sigma} p(V(v) - V(\mu) - \nabla V(\mu) \cdot (v-\mu)) + \sigma^T HV(\mu)\sigma \right\}$$

$$\text{s.t. } -p(H(v) - H(\mu) - \nabla H(\mu) \cdot (v-\mu)) - \sigma^T HH(\mu)\sigma \leq c$$

(16)

where $v \in \Delta(\text{supp}(\mu))$, $p \in \Delta I$ and $\sigma \in \mathbb{R}^{[\text{supp}(\mu)]}$. Equation (16) comes from applying Assumption 1-a and smoothness condition to Equation (4). I only discuss Assumption 1-a because the intuition is the same and similar proof methodology can be applied to Assumption 1-b to show an analog result.

**Theorem 10.** Let $E = \{\mu \in \Delta(X) | V(\mu) > F(\mu)\}$ be experimentation region. Suppose there exists $C^{(2)}$ smooth $V(\mu)$ on $E$ solving Equation (16), then $\exists$ policy function $v : E \rightarrow \Delta(X)$ s.t.

$$pV(\mu) = -c \frac{F(v(\mu)) - V(\mu) - \nabla V(\mu) \cdot (v(\mu) - \mu)}{H(v(\mu)) - H(\mu) - \nabla H(\mu) \cdot (v(\mu) - \mu)}$$

and $v$ satisfies the following properties:

1. $D_{\mu}V(\mu) \geq 0$.
2. $v(\mu) \in E^{C}$.
3. $D_{\mu}V(\mu) \cdot HH(v(\mu)) \leq 0$.
4. $pV(\mu) \geq \sup_{\sigma} -\sigma^T HV(\mu)\sigma$.

There exists a nowhere dense set $D$ s.t. strict inequality holds on $E \setminus D$ in property 1,3 and 4.

**Theorem 10** states that if a solution $V(\mu)$ to Equation (16) exists, then $V(\mu)$ can be solved with only Poisson signals. The four properties are extensions to the four properties in Theorem 2 respectively. Property 2 and 4 are exactly the immediate action property and the suboptimality of Gaussian signal. Property 1 and 3 are weaker

---

19 The maximization problem can be translated into a PDE system. What is problematic is boundary condition. In fact, to solve for $V(\mu)$ searching over one action, I need to use the value function at regions where DM is indifferent between two actions as boundary condition. That is unknown, in contrast to the one dimensional analog $V(\mu^*)$ which can be easily calculated.

20 $HH(\mu)$ is defined on boundary where $V(\mu) = F(\mu)$ as continuous extension of interior Hessian’s by Kirszbraun theorem.
than in Theorem 2. Property 1 is the extension to confirmatory signal property. It states that optimal direction of jump is in the mayopic direction that value function increases. Property 3 is the extension to the increasing precision property. $D_{\mu \rightarrow \nu} V(\mu)$ is the direction $\nu$ is moving when $\mu$ is moving against $\nu$. $HH(\nu)(\nu - \mu)$ is the direction $(\nu - \mu)$ distorted by a negative definite matrix $HH(\nu)$. In a special case when $H(\mu) = \|\mu - \mu_0\|_2^2$, $HH(\nu)(\nu - \mu)$ is in the same direction as $(\mu - \nu)$, which means the distance between $\mu$ and $\nu$ is increasing when $\mu$ is drifting against $\nu$. In a generic case, this property does not directly predict how $\|\nu - \mu\|$ changes.

Figures 12 and 13 illustrate Theorem 10 in a numerical example. There are three states and three actions. Belief space is a two-dimensional simplex. $F(\mu)$ is assumed to be a centrally symmetric function on belief space (Figure 12-(a)). Value function $V(\mu)$ is the meshed manifold in Figure 12-(c). Each blue curve in Figure 12-(b) shows a drifting path of posterior beliefs. Take a prior in lower right region. Then optimal policy is to search for one posterior (red points in lower right corner of Figure 13-(c)), and posterior belief conditional on receiving no signal drifts along the curve in arrowed direction as in Figure 13-(c). Once belief reaches the boundary, optimal policy becomes searching for two posteriors in a balanced way and posterior drifts towards center of belief space (see Figure 13-(b), arrowed blue curve is belief trajectory and dashed arrows points to optimal posterior). Finally, if belief reaches center, optimal policy is to search for three posteriors in a balanced way (Figure 13-(a)).
7. Conclusion

This paper provides a dynamic information acquisition framework which allows fully general design of signal process, and characterizes the optimal information acquisition strategy. My first contribution is an optimization foundation for a family of simple information generating process: for an information acquisition problem with flexible design of information, optimal information structure induces beliefs following a jump-diffusion process. Second, I characterize the optimal policy: it is optimal to seek for a Poisson signal whose arrival confirms prior belief. Arrival of the signal leads to immediate action. Absence of the signal is followed by continuing learning at increasing precision and decreasing intensity. This paper also provides a dynamic foundation for rational inattention: static rational inattention is both the full patience limit and linear flow cost limit of the dynamic information acquisition problem.
Appendix: Omitted proofs

7.1. Roadmap for proofs

Figure 14: Roadmap for proofs

Figure 14 illustrates the roadmap for proofs in this paper. Each node in the figure displays theorem/lemma’s name and its page number. Proof of each node depends (indirectly) on all nodes linked (indirectly) to it on...
the right. From top to bottom, the nodes are ordered by order of proofs: each node only depends on nodes on the right of it or above it. So it is clear that there is no circular argument. Dependent nodes that have been proved earlier are boxed by dashed lines. From left to right, the nodes are ordered by importance. Lemmas in the first layer are directly supporting the proof for theorems. Lemmas in the second layer or above are more technical lemmas.

7.2. proof of Lemma 1
Proof. First of all, by Lemma A.8 the limit $\lim_{dt \to 0} V_{dt}(\mu)$ exists. As discussed in the main text, take any strategy $(\mu_t, \tau)$ feasible in the continuous-time problem. Discretize the objective function on a uniform grid with size $dt$. Then the utility is weakly less than $V_{dt}(\mu)$, since $V_{dt}(\mu)$ is optimized among all discrete-time martingales and stopping times. Therefore, $V(\mu) \leq \lim_{dt \to 0} V_{dt}(\mu)$.

Now prove $V(\mu) \geq \lim_{dt} V_{dt}(\mu)$. Given time period $dt$, by Lemma 2 there exists optimal solution $\mu^*_t$ and $\tau^*$. Where $\mu^*_{t+1} | F_t$ has support size $N$. Now define a continuous-time belief process $\mu_t$. $\forall i \in \mathbb{N}, \mu_{idt} = \mu^*_{idt}$. $\forall i$ and conditional on $F_t$, apply Lemma A.6 to the distribution of $\mu^*_{t+1}$ to smooth it on $[idt, (i+1)dt]$. Lemma A.6 states that there exists martingale $\langle \tilde{\mu} \rangle$ satisfying: $\forall s, t \in [0,1], s > t$: $E[H(\mu_t) - H(\mu_s) | F_t] = (s - t)E[H(\mu^*_t) - H(\mu^*_s) | F_t]$. For $t \in [idt, (i+1)dt]$, define $\mu_t | F_{idt} = \tilde{\mu}_{i+1} dt | F_t$. Therefore, $\forall t \in [idt, (i+1)dt]$: $\frac{1}{s-t}E[H(\mu^*_{t+1}) - \sum_j p_j^i H(\mu^*_{t+1})] = \int_{s-t}^{idt} E[H(\mu^*_{t+1}) - \sum_j p_j^i H(\mu^*_{t+1})]dt$.

Let $\tau = \tau^* dt$. It is easy to see that by construction $\tau$ is measurable to the natural filtration of $\mu_t$. Therefore:

$$V(\mu) \geq \int_{0}^{idt} e^{-\rho t} E(h(I_t)) dt$$

$$= \int_{0}^{idt} e^{-\rho dt} \cdot \sum_{i=0}^{\tau} e^{-\rho t} \cdot H(\mu^*_t) - \sum_{j=0}^{\tau} p_j^i H(\mu^*_t) dt$$

$$= \int_{0}^{idt} e^{-\rho dt} \cdot \sum_{i=0}^{\tau} p_j^i H(\mu^*_t) dt$$

$$= \int_{0}^{idt} e^{-\rho dt} \cdot \sum_{i=0}^{\tau} p_j^i H(\mu^*_t) dt$$

Second inequality is from $1 - e^{-x} \leq x$. Therefore, $V(\mu) \geq \lim_{dt} V_{dt}(\mu)$. Q.E.D.

7.3. Proof of Lemmas 2 and 3: generalized concavification

Lemma 2 is a corollary of Lemmas A.2, A.3 and A.4 and Theorem 11. Lemmas A.2, A.3 and A.4 proves the Bellman characterization. Theorem 11 shrinks the number of optimal posteriors using a generalized concavification method. Lemma 3 is a direct corollary of Theorem 11.

Theorem 11 (Concavification). Let $X$ be a finite state space, $V \in C(\Delta X), \mu \in \Delta X$. $H \in C(\Delta X)$ is non-negative. $f : \mathbb{R}^+ \mapsto \mathbb{R}^+$ continuous, increasing and convex. Then there exists $\tau$ s.t. $|\text{supp}(\tau)| \leq 2|X|$ solving:

$$\sup_{\tau \in \Delta X} E_{\tau} [V(v)] - f(H(\mu) - E_{\tau} [H(v)])$$

s.t. $E_{\tau} [v] = \mu$

Let $C^* = H(\mu) - E_{\tau} [H(v)]$, there exists $\lambda \in \text{dom} (C)$ such that:

$$\text{co}(V + \lambda H)(\mu) = E_{\tau} [(V + \lambda H)(\mu)]$$
Proof. Theorem 11 is a corollary of Lemma 1 and Theorem 4 of Zhong (2017a).

Support size: since objective function is monototic in \((E_{\tau}[V], E_{\tau}[H])\), optimal solution must be on the boundary of set \(\{(E_{\tau}[V], E_{\tau}[H])|E_{\tau}[v] = \mu\}\). Lemma 1 of Zhong (2017a) implies that there exists \(\tau\) solving Equation (17) and \(|\text{supp}(\tau)| \leq 2|X|\).

Concavification: Suppose \(f(\mathcal{C}) = \infty \iff \mathcal{C} \supseteq \tilde{\mathcal{C}}\). Since \(v - f(H(\mu) - h)\) is a concave function in \((v, h)\), and \(H(\mu) - h \leq \tilde{\mathcal{C}}\) is a linear constraint, we can apply Theorem 4 of Zhong (2017a): let \(V^*\) be maximum of Equation (17), there exists \(\lambda\) s.t.

\[
\begin{cases}
\tau \in \arg \max_{\tau \in \Delta(X)} E_{\tau}[\lambda_1 V - \lambda_2 H] \\
\text{subject to } E_{\tau}[v] = \mu
\end{cases}
\]

\((E_{\tau}[V], \mathcal{C}^*) \in \arg \min_{\mathcal{C} \subseteq \mathcal{C}} \min_{v \in \mathcal{C}} \lambda_1 v - \lambda_2 C \quad \text{subject to } v - f(C) > V^*
\]

Then by Kuhn-Tucker condition (generalized to subgradients), there exists \(\eta, \gamma > 0\) such that:

\[
\begin{cases}
\lambda_1 = \eta \\
\lambda_2 = -\eta \partial f(C^*) - \gamma
\end{cases}
\]

If \(\eta = 0\), then \(\gamma > 0\) and \(\tau\) maximizes \(E_{\tau}[\gamma H]\), then optimal \(\tau\) is uninformative and \(C^* = 0\), contradiction. So \(\eta > 0\). If \(\gamma = 0\), then we can normalize \((\lambda_1, -\lambda_2)\) to \((1, \lambda)\) and \(\lambda \in \partial f(C^*)\). If \(\gamma > 0\), the complementary slackness condition implies that \(C^* = \tilde{\mathcal{C}}\) and \(\lambda^2/\eta \in \partial f(C^*)\). So we can also normalize \((\lambda_1, -\lambda_2)\) to \((1, \lambda)\) and \(\lambda \in \partial f(C^*)\). Q.E.D.

7.4. proof of Lemma 4

Equation (4) only applies to smooth functions. However \(\nabla\) is not necessarily smooth. First define a space of functions which includes \(\nabla(\mu)\) on \(\Delta(X)\):

\[
\mathcal{L} = \left\{ V: \Delta(X) \to \mathbb{R}^+ \middle| \forall \mu \in \Delta X, \mu' \in \Delta(\text{supp}(\mu)), \lim_{\mu' \to \mu} \frac{|V(\mu') - V(\mu)|}{\|\mu' - \mu\|} \in \mathbb{R} \right\}
\]

where \(\|\cdot\|\) is Euclidean norm on \(\Delta X\). \(\mathcal{L}\) is the set of pointwise Lipschitz functions on \(\Delta(X)\). \(\nabla \in \mathcal{L}\) is guaranteed by Lemma A.9. Consider the following Bellman equation defined on \(\mathcal{L}\):

\[
\rho V(\mu) = \max \left\{ \rho F(\mu), \sup_{\substack{\nu \in \Delta(\text{supp}(\mu)) \cap \mathbb{R}^+, \\
\n \partial \in \mathbb{R}^{\text{supp}(\mu)} \right\} \right.
\]

\[
\begin{aligned}
&= \rho f(\mu) + \sum_{i \in \text{supp}(\mu)} p_i (V(\nu_i) - V(\mu)) \\
&\quad - \nabla V(\mu) \cdot \nabla f(\mu) \\
&\quad + \frac{1}{2} \|\partial\|^2 D^2 V(\mu, \partial) \\
&\quad - f(\mu) - \sum_{i \in \text{supp}(\nu_i)} p_i (H(\nu_i) - H(\mu) - \nabla H(\mu) \cdot (\nu_i - \mu)) \quad \text{(18)}
\end{aligned}
\]

\(\nabla\) and \(H\) represents gradient and Hessian operator (well-defined on all interior points). Since \(\nabla\) is not necessarily differentiable, we use operator \(D\) and \(D^2\) to replace Jacobian and Hessian. \(D\) and \(D^2\) are defined as following \(\forall y \in B^{|\text{supp}(x)| - 1}\) (Unit ball in \(|\text{supp}(x)| - 1\) dimensional space):

Definition 2 (General differential). \(\forall f \in \mathcal{L}\):

\[
\begin{cases}
D f(x, y) = \lim_{\delta \to 0} \frac{f(x + \delta y) - f(x) - \delta f(x, y)}{\delta |y|} \\
D^2 f(x, y) = \lim_{\delta \to 0} 2 \frac{f(x + \delta y) - f(x) - \delta f(x, y)}{\delta |y|^2}
\end{cases}
\]
Noticing that if \( f \in C^{(1)}(\Delta X) \), then \( Df(x,y) = \nabla f(x,y) \). If \( f \in C^{(2)}(\Delta X) \) then \( D^2f(x,y) = \frac{\nabla^T H(x,y)}{||y||} \). It is not hard to verify that for \( C^{(1)} \) smooth value function \( V(\mu) \), Equation (18) is equivalent to Equation (4).

**Proof.**

Consider Lemma 4 by replacing Equation (4) with Equation (18). If the statement is proved with Equation (18), then since \( \overline{V} = V \) is \( C^{(1)} \) smooth, \( \overline{V} \) is smooth and Equation (4) automatically holds. I prove by induction on dimensionality of \( \mu \). First of all, Lemma 4 is trivially true when \( \mu = \delta_x \) since \( V(\mu) = \overline{V}(\mu) = F(\mu) \) when state is deterministic. Now it is sufficient to prove \( \overline{V} = V \) on interior of any \( \Delta X \) conditional on \( \overline{V} = V \) being true on \( \partial \Delta X \).

**Unimprovability:** Suppose \( \overline{V} \) is improvable in Equation (18) at interior \( \mu \), there exists \( p_i, v_i, \hat{\sigma}, c \) such that:

\[
\rho \overline{V}(\mu) < \sum p_i (\overline{V}(v_i) - \overline{V}(\mu)) - D \overline{V}(\mu, \mu - \sum p_i v_i) \| \sum p_i v_i - \mu \| + \sum D^2 \overline{V}(\mu, \hat{\sigma}) \| \hat{\sigma} \|^2 - h(c)
\]

where \( c = - \sum p_i (H(v_i) - H(\mu) - \nabla H(\mu) \cdot (v_i - \mu)) - \sum \hat{\sigma} H H(\mu) \hat{\sigma} \)

Then if we compare the following two ratios:

\[
\frac{\sum p_i (\overline{V}(v_i) - \overline{V}(\mu)) - D \overline{V}(\mu, \mu - \sum p_i v_i) \| \sum p_i v_i - \mu \|}{- \sum p_i (H(v_i) - H(\mu) - \nabla H(\mu) \cdot (v_i - \mu))} \frac{D^2 \overline{V}(\mu, \hat{\sigma}) \| \hat{\sigma} \|^2}{- \hat{\sigma} H H(\mu) \hat{\sigma}}
\]

At least one of them must be larger than \( \frac{\rho \overline{V}(\mu) + h(c)}{c} \).

- **Case 1:**

\[
\frac{\sum p_i (\overline{V}(v_i) - \overline{V}(\mu)) - D \overline{V}(\mu, \mu - \sum p_i v_i - \mu) \| \sum p_i v_i - \mu \|}{- \sum p_i (H(v_i) - H(\mu) - \nabla H(\mu) \cdot (v_i - \mu))} \geq \frac{\rho}{c} \overline{V}(\mu) + \frac{h(c)}{c} + \varepsilon
\]

By **Definition 2**, there exists \( \delta, \varepsilon > 0 \) s.t.:

\[
\frac{\sum p_i (\overline{V}(v_i) - \overline{V}(\mu)) - D \overline{V}(\mu, \mu - \sum p_i v_i - \mu)}{\sum p_i (H(v_i) - H(\mu) + H(v_i) - H(\mu) - \nabla H(\mu) \cdot (v_i - \mu))} \geq \frac{\rho}{c} \overline{V}(\mu) + \frac{h(c)}{c} + \varepsilon
\]

where \( \delta \) is sufficiently small that \( \mu_0 = \mu - \delta (\sum p_i v_i - \mu) \in \Delta X^0 \). Then by construction, if we assume:

\[
\begin{aligned}
\begin{cases}
p_i' = \frac{1}{1+\delta} \\
p_i'' = \frac{\delta}{1+\delta} p_i
\end{cases}
\end{aligned}
\]

Then \((p_i', v_i')\) is Bayesian plausible:

\[
\begin{aligned}
\sum p_i' &= 1 \\
\sum p_i' v_i' &= \mu
\end{aligned}
\]

where 0 is also included in indices \( i \)'s. Replacing terms in Equation (19) and let \( I(v_i | \mu) = H(\mu) - \sum p_i' H(v_i) \):

\[
\frac{\sum p_i' (\overline{V}(v_i) - \overline{V}(\mu))}{- \sum p_i' H(v_i) + H(\mu)} \geq \frac{\rho}{c} \overline{V}(\mu) + \frac{h(c)}{c} + \varepsilon
\]

\[
\Longrightarrow \sum p_i' (\overline{V}(v_i) - \frac{I(v_i | \mu)}{c} h(c)) \geq \left( 1 + \rho \frac{I(v_i | \mu)}{c} \right) \overline{V}(\mu) + \varepsilon I(v_i | \mu)
\]

It is easy to verify that \( I(v_i | \mu) \) is continuous in \( \delta \) and it is zero when \( \delta = 0 \). So \( \delta \) can be chosen sufficiently small that

\[
e^{\frac{I(v_i | \mu)}{c}} - \left( 1 + \rho \frac{I(v_i | \mu)}{c} \right) \leq \sum_{k=1}^{\infty} \frac{1}{(k+1)!} \left( \frac{\rho}{c} \right)^{k+1} I(v_i | \mu)^k, I(v_i | \mu) \leq \frac{\varepsilon I(v_i | \mu)}{4 \sup F}
\]

(21)
The equality is from Taylor expansion of exponential function. Plug Equation (21) into Equation (20):

\[
\sum p'_i \nabla(v_i) - \frac{I(v_i)}{c} h(c) \geq \epsilon e^{\frac{I(v_i)}{c}} \nabla(\mu) + \frac{\epsilon}{4} I(v_i) \mu
\]

\[
\implies e^{-\frac{I(v_i)}{c}} \left( \sum p'_i \nabla(v_i) \right) - \frac{I(v_i)}{c} h(c) \geq \nabla(\mu) + e^{-\frac{\epsilon}{4} I(v_i)} \frac{\epsilon}{4} \frac{I(v_i)}{c} h(c)
\]

(22)

Noticing that \(1 - e^{-\frac{\epsilon}{4} I(v_i)}\) is a second order small term. Then we can pick \(\delta\) such that Equation (22) implies:

\[
e^{-\frac{\epsilon}{4} I(v_i)} \left( \sum p'_i \nabla(v_i) \right) - \frac{I(v_i)}{c} h(c) \geq \nabla(\mu) + \frac{\epsilon}{8} I(v_i) \mu
\]

From now on, we fix \(\epsilon\) and \(\delta\). Pick \(dt = \frac{I(v_i)}{c}, dt_m = \frac{d}{m}\). By uniform convergence, there exists \(N\) s.t. \(\forall m \geq N:\)

\[
e^{-\rho dt_m} \left( \sum p'_i \nabla(v_i) \right) - dt \cdot \left( \frac{I(v_i)}{m} \right) > V_{dt_m}(\mu)
\]

\[
\implies e^{-\rho dt_m} \left( \sum p'_i \nabla(v_i) \right) - \sum_{\tau=0}^{m-1} e^{-\rho \tau dt_m} \left( \frac{I(v_i)}{m} \right) > V_{dt_m}(\mu)
\]

That is to say we find a feasible experiment, whose cost can be spread into \(m\) periods (by applying Lemma A.6 recursively). This experiment strictly dominates the optimal experiment at \(\mu\) for discrete time problem with \(dt_m\). Contradiction. Thus \(\nabla\) must be unimprovable at \(\mu\).

* Case 2:

\[
\frac{D^2 \nabla(\mu, \hat{\sigma})}{-\sigma^T HH(\mu) \sigma} \geq \frac{\rho}{c} \nabla(\mu) + \frac{h(c)}{c} + 2\epsilon
\]

The by Definition 2, there exists \(\hat{\sigma}, \delta, \epsilon > 0\) s.t.:

\[
\frac{\nabla(\mu + \delta \hat{\sigma}) - \nabla(\mu) - \delta D \nabla(\mu, \hat{\sigma})}{-H(\mu + \delta \hat{\sigma}) + H(\mu) + \delta \nabla H(\mu) \hat{\sigma}} \geq \frac{\rho}{c} \nabla(\mu) + \frac{h(c)}{c} + 2\epsilon
\]

Then by Definition 2, there exists \(\delta'\) s.t.:

\[
\frac{\nabla(\mu + \delta \hat{\sigma}) - \nabla(\mu) - \delta \nabla(\mu) - \nabla(\mu - \delta \hat{\sigma})}{-H(\mu + \delta \hat{\sigma}) + H(\mu) + \delta \nabla H(\mu - \delta \hat{\sigma})} \geq \frac{\rho}{c} \nabla(\mu) + \frac{h(c)}{c} + \epsilon
\]

If we let \(\mu_1 = \mu - \delta \hat{\sigma}\) and \(\mu_2 = \mu + \delta \hat{\sigma}\), \(p_1 = \frac{\hat{\sigma}}{\delta + \hat{\sigma}}, p_2 = \frac{\hat{\sigma}}{\delta + \hat{\sigma}}\), then:

\[
\sum p_i \nabla(v_i) \geq \left(1 + \rho \frac{I(v_i)}{c}\right) \nabla(\mu) + \frac{I(v_i)}{c} h(c) + \epsilon I(v_i) \mu
\]

Noticing that the expression we are studying is exactly the same as in Case 1. Then using same argument, This case is also ruled out.

**Equality:** \(\forall\) smooth function \(V\) solving Equation (18), \(\nabla = V\). Noticing that this automatically proves uniqueness of solution to Equation (18). I prove inequality from both directions for \(\mu \in \Delta(X)^o:\)

* \(\nabla(\mu) \geq V(\mu)\): Suppose not, then consider \(\nabla(\mu) = V(\mu) - V(\mu)\). Since both \(V\) and \(\nabla\) are continuous, \(\nabla\) is continuous. Therefore argmin \(\nabla U(\mu)\) is non empty and \(\min \nabla U < 0\) according to our assumption. Choose \(\mu \in \argmin U (\mu \in \Delta X^o \text{ since } V = \nabla \text{ on boundary})\). Since \(\nabla(\mu) \geq F(\mu), V(\mu) > F(\mu)\). Let \((p_i, v_i, \hat{\sigma})\) be a strategy solving \(V(\mu)\):

\[
\rho V(\mu) = \sum_{i} p_i (V(v_i) - V(\mu)) - DV(\mu, \sum_{i} p_i v_i - \mu) \left\| \sum p_i (v_i - \mu) \right\| + \frac{1}{2} D^2 V(\mu, \hat{\sigma}) \| \hat{\sigma} \|^2
\]

(23)
Now compare $D\nabla$ and $DV$:

$$\frac{\nabla(\mu) - \nabla(\mu')}{{\mu-\mu'}} = \frac{V(\mu) - V(\mu') + U(\mu) - U(\mu')}{\|\mu - \mu'\|} \leq \frac{V(\mu) - V(\mu')}{\|\mu - \mu'\|}$$

$$\Rightarrow \text{liminf} \frac{\nabla(\mu) - \nabla(\mu')}{\|\mu - \mu'\|} \leq \text{lim} \frac{V(\mu) - V(\mu')}{\|\mu - \mu'\|}$$

$$\Rightarrow D^2 \nabla(\mu, \mu' - \mu) \|\mu' - \mu\| \leq \nabla V(\mu) \cdot (\mu' - \mu)$$

Compare $D^2 \nabla$ and $D^2 V$:

$$\frac{\nabla(\mu') - \nabla(\mu) - D\nabla(\mu) \cdot (\mu' - \mu) \|\mu' - \mu\|}{\|\mu' - \mu\|^2} \leq \frac{V(\mu') - V(\mu) - \nabla V(\mu) \cdot (\mu' - \mu) + U(\mu') - U(\mu)}{\|\mu - \mu'\|^2}$$

$$\Rightarrow D^2 \nabla(\mu, \nu) \geq D^2 V(\mu, \nu)$$

Therefore Equation (23) implies:

$$\rho V(\mu) \leq \sum p_i(\nabla (v_i) - \nabla (v_i) - (U(v_i) - U(\mu)))$$

$$- D\nabla(\mu) \sum v_i - \mu \left\| \sum v_i - \mu \right\| + \frac{1}{2} D^2 \nabla(\mu, \nu) \|\nu\|^2$$

$$- h \left( - \sum p_i(H(v_i) - H(\mu) + \nabla H(\mu)(v_i - \mu)) - \frac{1}{2} \sigma^T H(\mu) \hat{\sigma} \right)$$

$$\leq \rho \nabla V(\mu)$$

The first inequality comes from replacing $DV$ and $D^2 V$ with $D\nabla$ and $D^2 \nabla$. The second inequality comes from $U(v_i) - U(\mu) \geq 0$ and unimprovability of $V$. Contradiction.

- $V(\mu) \geq \nabla(\mu)$: We prove by showing that $\forall dt > 0, V \geq V_{dt}$. Suppose not, then there exists $\mu', dt$ s.t. $V_{dt}(\mu') > V(\mu')$. Let $dt_n = \frac{dt}{n}$. Since $V_{dt_n}$ is increasing in $n$, there exists $\epsilon > 0$ s.t. $V_{dt_n}(\mu') - V(\mu') \geq \epsilon \forall n \in \mathbb{N}$. Now consider $U_n = V - V_{dt_n}$. $U_n$ is continuous by Lemma 3 and $U_n(\mu') \leq \epsilon$. Pick $\mu^n \in \text{argmin} U_n$. Since $\Delta(X)$ is compact, there exists a converging sequence $\lim mu_n = \mu$. By assumption, $U_n(\mu^n) \leq \epsilon$, therefore since $U(\mu) = \lim U_n(\mu^n) \leq \epsilon$, $\mu$ must be in interior of $\Delta(X)$. So without loss, $\mu^n$ can be picked that $\mu^n \in \Delta(X)^0$. Now consider the optimal strategy of discrete time problem:

$$\left\{ \begin{array}{l}
V_{dt_n}(\mu^n) = e^{-\rho dt_n} \sum p_i^n V_{dt_n}(v_i^n) - dt_n h(c_c) \\
\sum p_i^n (H(\mu^n) - H(v_i^n)) = c_n dt_n \\
\sum p_i^n v_i^n = \mu^n; \sum p_i^n = 1
\end{array} \right.$$

By definition of $U_n(\mu)$:

$$\sum p_i^n (V(v_i^n) - V(\mu^n)) = \sum p_i^n (V_{dt_n}(v_i^n) - V_{dt_n}(\mu^n) - U_n(\mu^n) + U(v_i^n))$$

$$\geq \sum p_i^n (V_{dt_n}(v_i^n) - V_{dt_n}(\mu^n))$$

$$= (e^{\rho dt_n} - 1) V_{dt_n}(\mu^n) + e^{\rho dt_n} dt_n h(c_n)$$

$$\geq \rho dt_n V_{dt_n}(\mu^n) + e^{\rho dt_n} dt_n h(c_n)$$

$$\Rightarrow \rho dt_n \mu + e^{\rho dt_n} dt_n h(c_n)$$

$$\Rightarrow \rho \mu + e^{\rho dt_n} dt_n h(c_n)$$

$$\Rightarrow \rho \mu \leq -\rho \epsilon + \sum p_i^n dt_n (V(v_i^n) - V(\mu^n)) - e^{\rho dt_n} h(c_n)$$

$$\Rightarrow \rho \mu \leq -\rho \epsilon + \sum p_i^n dt_n (V(v_i^n) - V(\mu^n)) - h(c_n)$$

(24)
The first equality is definition of $U_\mu$. The first inequality is from $\mu^n \in \text{argmin} U_\mu$. The second inequality is from $e^{x-1} \geq x$. Now since number of posteriors $v_i^n$ is no more than $2|X|$, we can take a subsequence of $n$ such that all $\lim v_i^n = v_i$. We partition $v_i^n$ into two kinds: $\lim v_i^n = v_i \neq \mu$, $\lim v_i^n = \mu$. Since $V$ is unimprovable, $\forall \epsilon, \sigma$ we have $D^2 V(\mu, \sigma) \| \sigma \|^2 \leq - \sigma^T HH(\mu) \sigma \left( \frac{\rho}{c} V(\mu) + \frac{h(c)}{c} \right)$. Since $V \in C^{(1)}$, $H \in C^{(2)}$, $\forall \eta$, there exists $\delta$ s.t. $\forall |\mu' - \mu| \leq \delta$:

$$
\begin{align*}
& \left\{ \| HH(\mu) - HH(\mu') \| \leq \eta \\
& \| V(\mu) - V(\mu') \| \leq \eta \\
\rightarrow & \ D^2 V(\mu', \sigma) \leq \left( \frac{\rho}{c} V(\mu') + \frac{h(c)}{c} \right) \left( - \frac{\sigma^T HH(\mu') \sigma}{\| \sigma \|^2} \right) \\
& \leq \left( \frac{\rho}{c} V(\mu) + \frac{h(c)}{c} \right) \left( - \frac{\sigma^T HH(\mu) \sigma}{\| \sigma \|^2} \right) + \left( \frac{\rho}{c} \sup F + \frac{h(c)}{c} \right) \eta \Rightarrow \frac{\rho}{c} \eta \| HH(\mu) \| \\
\end{align*}
$$

If we pick $\eta$ and $\delta$ properly:

$$
D^2 V(\mu', \sigma) \leq \left( \frac{\rho}{c} V(\mu) + \frac{h(c)}{c} \right) \left( - \frac{\sigma^T HH(\mu) \sigma}{\| \sigma \|^2} \right) + 1 + h(c) \frac{\eta}{c}
$$

Then there exists $N$ s.t. $\forall n \geq N$, $|v_i^n - \mu| < \delta$, $|\mu^n - \mu| < \delta$. Applying Lemma A.10 to $g(\alpha) = V(\alpha v_i^n + (1 - \alpha) \mu^n)$:

$$
\begin{align*}
V(v_i^n) - V(\mu^n) - \nabla V(\mu^n)(v_i^n - \mu^n) &= g(1) - g(0) - g'(0) \\
\leq & \frac{1}{2} \sup_{\alpha \in (0,1)} D^2 g(\alpha, 1) = \sup_{\alpha \in (0,1)} \limsup_{d \to 0} \frac{g(\alpha + d) - g(\alpha) - \alpha g'(0)}{d^2} \\
= & \sup_{\xi \in (\mu^n, v_i^n)} \limsup_{d \to 0} \frac{V(\xi + d(v_i^n - \mu^n) - V(\xi) - dJ V(\xi)(v_i^n - \mu^n)}{d^2} \\
\leq & \frac{1}{2} \sup_{|\xi - \mu| \leq \delta} \left\| v_i^n - \mu^n \right\|^2 \\
\leq & -\frac{1}{2} \left( \frac{\rho}{c} V(\mu) + \frac{h(c)}{c} \right) \left( v_i^n - \mu^n \right)^T HH(\mu)(v_i^n - \mu^n) + \frac{1 + h(c)}{2c} \eta \left\| v_i^n - \mu^n \right\|^2 \\
\end{align*}
$$

Therefore, by applying Equation (25):

$$
\begin{align*}
& \sum p_i^n \left( V(v_i^n) - V(\mu^n) \right) \\
= & \sum p_i^n \left( V(v_i^n) - V(\mu^n) - \nabla V(\mu^n)(v_i^n - \mu^n) \right) + \sum p_i^n \left( V(v_i^n) - V(\mu^n) - \nabla V(\mu^n)(v_i^n - \mu^n) \right) \\
\leq & \sum p_i^n \left( V(v_i^n) - V(\mu^n) - \nabla V(\mu^n)(v_i^n - \mu^n) \right) \\
& - \frac{1}{2} \left( \frac{\rho}{c} V(\mu) + \frac{h(c)}{c} \right) \sum p_i^n \left( v_i^n - \mu^n \right)^T HH(\mu)(v_i^n - \mu^n) + \frac{1 + h(c)}{2c} \eta \sum p_i^n \left\| v_i^n - \mu^n \right\|^2 \\
\end{align*}
$$

Noticing that Equations (25) and (26) are true uniform to $c$, so we can replace $c$ with $c_n$ and Equation (26) is still true. Now let $\tilde{p}_i^n = \frac{p_i^n}{d\mu^n/\mu^n - \tilde{\sigma}_n^T HH(\mu^n) \tilde{\sigma}_n dt_n} = \sum p_i^n \left( H(\mu^n) - H(v_i^n) + \nabla H(\mu)(v_i^n - \mu^n) \right)$, we will have:

$$
\sum p_i^n \left( H(\mu^n) - H(v_i^n) + H'(\mu^n)(v_i^n - \mu^n) \right) - \tilde{\sigma}_n^T HH(\mu^n) \tilde{\sigma}_n = c_n
$$

($\tilde{p}_i^n, \tilde{\sigma}_n$) is a feasible experiment for Equation (18). Therefore, by optimality of $V$ at $\mu^n$, we have

$$
\begin{align*}
& \left\{ \sum p_i^n \left( V(v_i^n) - V(\mu^n) - \nabla V(\mu^n)(v_i^n - \mu^n) \right) \leq \left( c_n + \tilde{\sigma}_n^T HH(\mu^n) \tilde{\sigma}_n \right) \left( \frac{\rho}{c_n} V(\mu^n) + \frac{h(c_n)}{c_n} \right) \right. \\
& \left. D^2 V(\mu^n, \tilde{\sigma}_n) \leq - \frac{\tilde{\sigma}_n^T HH(\mu^n) \tilde{\sigma}_n}{\| \tilde{\sigma}_n \|^2} \left( \frac{\rho}{c_n} \nabla V(\mu^n) + \frac{h(c_n)}{c_n} \right) \right.
\end{align*}
$$

43
Then we study term $\sum p^n_j (v^n_j - \mu^n)^2$. Apply Lemma A.10 to $g(\alpha) = H(\alpha v^n_j + (1-\alpha) \mu^n)$:

$$\sum p^n_j \left( H(\mu^n) - H(v^n_j) + \nabla H(\mu^n)(v^n_j - \mu^n) \right) \geq \frac{1}{2} \inf_{\tilde{\mu} \in [\mu^n, \mu^n]} \sum p^n_j \left( (v^n_j - \mu^n)^T H H(\tilde{\mu})(v^n_j - \mu^n) \right)$$

$$\geq -\frac{1}{2} \sum p^n_j ((v^n_j - \mu^n)^T H H(\mu^n)(v^n_j - \mu^n)) - \frac{1}{2} \eta \sum p^n_j \left\| v^n_j - \mu^n \right\|^2$$

(29)

Therefore, to sum up:

$$\sum \frac{p^n_{i,j}}{d_t} (V(v^n_{i,j}) - V(\mu^n)) \leq \sum p^n_i (V(v^n_i) - V(\mu^n) - \nabla V(\mu^n)(v^n_i - \mu^n))$$

$$+ \frac{1}{2} \sum \frac{p^n_{i,j}}{d_t} \left( - (v^n_j - \mu^n)^T H H(\mu)(v^n_j - \mu^n) \left( \frac{\rho}{c_n} V(\mu) + \frac{h(c_n)}{c_n} \right) \right)$$

$$+ \sum \frac{p^n_{i,j}}{d_t} \left( 1 + h(c_n) \frac{\| v^n_j - \mu^n \|^2}{2c_n} \eta \right)$$

$$\leq \left( c_n + \delta^n H H(\mu^n) \right) \left( \frac{\rho}{c_n} V(\mu) + \frac{h(c_n)}{c_n} \right)$$

$$+ \left( \frac{1}{2} \sum \frac{p^n_{i,j}}{d_t} (H(\mu^n) - H(v^n_j) + \nabla H(\mu^n)(v^n_j - \mu^n)) \right)$$

$$+ \frac{1}{d_t} \sum \frac{p^n_{i,j}}{2} \left( \| v^n_j - \mu^n \|^2 \right) \left( \frac{\rho}{c_n} V(\mu) + \frac{h(c_n)}{c_n} \right)$$

$$+ \frac{1}{d_t} \sum \frac{p^n_{i,j}}{2} \left( \| v^n_j - \mu^n \|^2 \right) \left( \frac{1 + h(c_n)}{2c_n} \eta \right)$$

$$\leq \rho V(\mu^n) + h(c_n) + \frac{1}{d_t} \sum \frac{p^n_{i,j}}{2} \left( \| v^n_j - \mu^n \|^2 \right) \left( \frac{1 + \rho V(\mu) + 2h(c_n)}{2c_n} \right) \eta + \rho \eta$$

The first inequality is Equation (26). The second inequality comes from Equation (28) and Equation (29). The next equality comes from definition of $\delta^n$. The last inequality comes from canceling out terms and $-\delta^n H H(\mu^n) \leq c_n$ (Noticing the difference between $V(\mu)$ and $V(\mu^n)$). Then by plug into Equation (24):

$$\rho V(\mu^n) \leq -\rho \epsilon + \rho V(\mu^n) + \frac{1}{d_t} \sum \frac{p^n_{i,j}}{2} \left( \| v^n_j - \mu^n \|^2 \right) \left( \frac{1 + \rho V(\mu) + 2h(c_n)}{2c_n} \right) \eta + \rho \eta$$

What’s more:

$$\sum p^n_i \left\| v^n_i - \mu^n \right\|^2 \inf_{\sigma} \frac{\left\| \sigma^T H H(\mu) \sigma \right\|}{\left\| \sigma \right\|^2}$$

$$\leq \sum p^n_i (v^n_i - \mu^n) H H(\mu)(\mu^n - \mu^n) \leq c_n d_t n + \eta \sum p^n_i \left\| v^n_i - \mu^n \right\|^2$$

$$\Rightarrow \sum p^n_i \left\| v^n_i - \mu^n \right\|^2 \leq \frac{c_n d_t n}{\inf_{\sigma} \frac{\left\| \sigma^T H H(\mu) \sigma \right\|}{\left\| \sigma \right\|^2} - \eta}$$

$$\Rightarrow \rho \epsilon \leq \frac{1}{2} \left( 1 + \rho V(\mu) + 2h(c_n) \right) \frac{\eta}{\inf_{\sigma} \frac{\left\| \sigma^T H H(\mu) \sigma \right\|}{\left\| \sigma \right\|^2} - \eta} + \rho \eta$$

44
By Lemma A.7, \( h(c_n) \) is uniformly bounded above. Since \( H \) is strictly concave \( \inf_{\sigma} \frac{e^T H(\mu) e}{|\sigma|^2} \) is positive. The inequality holds when \( \eta \) is chosen smaller than \( \inf_{\sigma} \frac{e^T H(\mu) e}{|\sigma|^2} \). By taking \( \eta \to 0 \), LHS will eventually be larger than RHS. Contradiction. Therefore, we proved that:

\[
V(\mu) = \limsup_{\Delta t \to 0} V_{dt}(\mu) = \bar{V}(\mu)
\]

Q.E.D.

7.5. Proof of Theorem 2

**Proof.** I prove Theorem 2 by guess and verification. First step is to construct the point \( \mu^* \) and function \( v(\mu) \) which is a feasible mechanism in Equation (4). Second step is to prove unimprovability of \( V(\mu) \). To simplify notation, I define a flow version of information measure:

\[
J(\mu, v) = H(\mu) - H(v) + H'(\mu)(v - \mu)
\]

Then total flow information cost will be \( p \cdot J(\mu, v) \). Let \( F_m = E_p[\mu(a_m, x)] \) and reorder \( a_m \) s.t. \( F_m \) is increasing in \( m \). Let \( \mu_k \) be each kink points of \( F: F(\mu) = F_k(\mu) \iff \mu \in [\mu_{k-1}, \mu_k] \). \( m \) is the smallest index s.t. \( F_m' \geq 0 \).

**Algorithm:**

In this part, I introduce the algorithm constructing \( V(\mu) \) and \( v(\mu) \). I only discuss the case \( \mu \geq \mu^* \) and the case \( \mu \leq \mu^* \) will follow by a symmetric method. The main steps are illustrated in Figure 15. First step is to find critical belief \( \mu^* \) at which two sided stationary Poisson signal is optimal (\( \mu^* = 0.5 \) in Figure 15). Then valued function is solved by searching over posterior beliefs associated with taking immediate actions. Then I maximize over all those value functions and claim it solves Equation (4).

![Figure 15: Construction of optimal value function.](image)

- **Step 1:** Define:

\[
\nabla^+(\mu) = \max_{v \neq \mu} \frac{F_m(v)}{1 + \frac{\epsilon}{\epsilon} J(\mu, v)}
\]
\[ \nabla^-(\mu) = \max_{v \leq \mu} \frac{F_m(v)}{1 + \frac{c}{\xi} J(\mu, v)} \]

\( \nabla^+ \) is increasing and \( \nabla^- \) is decreasing. There exists \( \mu^* \in [0,1] \) s.t. \( \nabla^+(\mu) > \nabla^-(\mu) \) when \( \mu > \mu^* \) and \( \nabla^-(\mu) < \nabla^-(\mu) \) when \( \mu < \mu^* \) (See Lemma 10). Define \( \nabla(\mu) = \max \{ \nabla^+(\mu), \nabla^-(\mu) \} \).

- **Step 2:** I construct the first piece of \( V(\mu) \) to the right of \( \mu^* \). There are three possible cases of \( \mu^* \) to be discussed (I omitted \( \mu^* = 1 \) by symmetry).

**Case 1:** Suppose \( \mu^* \in (0,1) \) and \( \nabla(\mu^*) > F(\mu^*) \). Then, there exists \( m \) and \( v(\mu^*) \in (\mu^*, 1) \) s.t.

\[ \nabla(\mu^*) = \frac{F_m(v(\mu^*))}{1 + \frac{c}{\xi} J(\mu^*, v(\mu^*))} \]

Initial condition \( (\mu_0 = \mu^*, V_0 = \nabla(\mu^*), V_0' = 0) \) satisfies Lemma 11, which implies that there exists \( V_m(\mu) \) solving:

\[ V_m(\mu) = \max_{v > \mu} \frac{c F_m(v) - V(\mu) - V'(\mu)(v-\mu)}{J(\mu, v)} \]

This refers to Figure 15-1. Define

\[ V_{\mu^*}(\mu) = \begin{cases} F(\mu) & \text{if } \mu \leq \mu^* \\ V_m(\mu) & \text{if } \mu > \mu^* \end{cases} \]

Be Lemma 11, when \( V_{\mu^*}(\mu) > F(\mu), V_{\mu^*} \) is smoothly increasing and optimal \( v(\mu) \) is smoothly decreasing.

Now Update \( V_{\mu^*}(\mu) \) with respect to more actions. First consider \( F_{m-1} \) and let \( \tilde{\mu}_m \) be the smallest \( \mu > \mu^* \):

\[ V_{\mu^*}(\tilde{\mu}_m) = \max_{v > \mu} \frac{c F_{m-1}(v) - V_{\mu^*}(\tilde{\mu}_m) - V'_{\mu^*}(\tilde{\mu}_m)(v-\tilde{\mu}_m)}{J(\mu, v)} \]

At \( \tilde{\mu}_m \), searching posterior on \( F_{m-1} \) first dominates searching posterior on \( F_m \). This step refers to Figure 15-2. \( \tilde{\mu}_m \) is the smallest intersection point of blue curve \( (V_{\mu^*}(\mu)) \) and thin red curve (right hand side). If \( V_m(\tilde{\mu}_m) > F_{m-1}(\tilde{\mu}_m) \) then solve for \( V_{m-1} \) with initial condition \( \mu_0 = \tilde{\mu}_m, V_0 = V_m(\tilde{\mu}_m), V_0' = V_{m}^*(\tilde{\mu}_m) \) according to Lemma 11 and redefine \( V_{\mu^*} = V_{m-1}(\mu) \) when \( \mu > \tilde{\mu}_m \). Otherwise skip to looking for \( \tilde{\mu}_{m-1} \). If \( m - 1 > m \), continue this procedure by looking for \( \tilde{\mu}_{m-1} \) and update \( V_{\mu^*} \mid \mu > \tilde{\mu}_{m-1} \) with corresponding \( V_{m-2} \ldots \) until \( m = m \). This refers to Figure 15-3. Now suppose \( V_m \) first hits \( F(\mu) \) at some point \( \mu^{**} \) (\( \mu^{**} > \mu^* \) since \( V_m(\mu^*) > F(\mu^*) \)). \( V_{\mu^*} \) is a smooth function on \( [\mu^*, \mu^{**}] \) such that:

\[ V_{\mu^*}(\mu) = \begin{cases} F(\mu) & \text{if } \mu \leq \mu^* \text{ or } \mu > \mu^{**} \\ V_k(\mu) & \text{if } \mu \in [\tilde{\mu}_k, \tilde{\mu}_{k-1}] \end{cases} \]

By construction, optimal posterior \( v_{\mu^*}(\mu) \) is smoothly decreasing on each \( (\tilde{\mu}_{k+1}, \tilde{\mu}_k) \) and jumps down at each \( \tilde{\mu}_k \). By Lemma B.4 and our construction, \( \forall \mu \in [\mu^*, \mu^{**}] \):

\[ V_{\mu^*}(\mu) = \max_{v > \mu} \frac{c F_k(v) - V_{\mu^*}(\mu) - V'_{\mu^*}(\mu)(v-\mu)}{J(\mu, v)} \quad (30) \]

**Case 2:** Suppose \( \mu^* \in (0,1) \) but \( \nabla(\mu^*) = F(\mu^*) \), let \( \mu^{**} = \inf \{ \mu > \mu^* \mid \nabla(\mu) > F(\mu) \} \).

---

21 Existence is guaranteed by smoothness of \( V_{\mu^*} \) and \( J \). Noticing that \( V_m(\tilde{\mu}_m) > F_{m-1}(\tilde{\mu}_m) \). Otherwise, there will be a \( \tilde{\mu}_m < \mu^*_m \) s.t. \( V_m(\tilde{\mu}_m) = F_{m-1}(\tilde{\mu}_m) \) and it is easy to verify that \( V_m \) is weakly larger than the maximum. So there is an even smaller \( \mu^*_m \), contradiction.

22 Define \( \tilde{\mu}_{m+1} = \mu^* \) and \( \tilde{\mu}_m = \mu^{**} \) for consistency.

23 Since \( F_{k+1} \) always crosses \( F_k \) above, when indifference between choosing \( F_{k-1} \) and \( F_k \), the posterior corresponding to \( F_{k-1} \) must be smaller.
Case 3: Suppose $\mu^*=0$, then $F'(0) \geq 0$ (by Lemma 10). Consider

$$\tilde{V}(\mu) = \max_{v \geq \mu, k} \frac{c \cdot F_k(v) - F_1(\mu) - F_1'(v - \mu)}{f(\mu, v)}$$

Define $\mu^{**} = \inf \{ \mu | \tilde{V}(\mu) > F_1(\mu) \}$. By Assumption 2, $\lim_{\mu \to 0} |H'(\mu)| = \infty$, then there exists $\delta$ s.t. $\forall \mu < \delta$, $\forall v \geq m \sup_{f(\mu, v)} \leq \inf F$. Therefore $\mu^{**} \geq \delta > 0$. This step refers to Figure 15-4.

- Step 3: Solve for $V$ to the right of $\mu^{**}$. For all $\mu^{\circ} \geq \mu^{**}$ such that:

$$F(\mu^{\circ}) = \max_{v \geq \mu^{\circ}, k} \frac{c \cdot F_k(v) - F(\mu^{\circ}) - F'(\mu^{\circ})(v - \mu^{\circ})}{f(\mu^{\circ}, v)}$$

(31)

Let $m$ be the index of optimal action. Solve for $V_m$ with initial condition $\mu_0 = \mu^{\circ}, V_0 = F(\mu^{\circ}), V_0' = F'(\mu^{\circ})$. Then take same steps in Step 3 and solve for $\hat{\mu}_k$ and $V_{k-1}$ sequentially until $V_{m_0}$ first hits $F$. This step refers to Figure 15-4,5. Now suppose $V_{m_0}$ first hits $F(\mu)$ at some point $\mu^\infty$ (can potentially be $\mu$), define:

$$V_{\mu^\infty}(\mu) = \begin{cases} V(\mu) & \text{if } \mu < \mu^\circ \text{ or } \mu \geq \mu^\infty \\ V_k(\mu) & \text{if } \mu \in [\hat{\mu}_k+1, \hat{\mu}_k] \end{cases}$$

By Lemma 11, $V_\mu$ is piecewise smooth and pasted smoothly. So $V_\mu$ is a smooth function on $[\mu, \mu^{**}]$. Optimal posterior $v_{\mu^\infty}(\mu)$ is smoothly decreasing on each $(\hat{\mu}_{k+1}, \hat{\mu}_k)$ and jumps down at each $\hat{\mu}_k$. By Lemma B.4 and our construction, $\forall \mu \in [\mu^{\circ}, \mu^\infty]$:

$$V_\mu(\mu) = \max_{v \geq \mu^{\circ}, k} \frac{c \cdot F_k(v) - V_{\mu^\circ}(\mu) - V_{\mu^\circ}(\mu)(v - \mu)}{f(\mu, v)}$$

(30)

Let $\Omega$ be the set of all such $\mu^{\circ}$'s.

- Step 4: Define:

$$V(\mu) = \begin{cases} V_{\mu^\circ}(\mu) & \text{if } \mu \in [\mu^*, \mu^{**}] \\ \sup_{\mu^\circ \in \Omega} \{ V_{\mu^\circ}(\mu) \} & \text{if } \mu \geq \mu^{**} \end{cases}$$

In the algorithm, I only discussed the case $\mu^* < 1$ and constructed the value function on the right of $\mu^*$. On the left of $\mu^*$, $V$ can be defined using a totally symmetric argument by referring to Lemma 11' and Lemma B.4'.

Smoothness:
I want to show that $V(\mu)$ is piecewisely defined as $V_{\mu_0}$'s. This is true when $\mu < \mu^{**}$ by definition of $V_{\mu^*}$. So I prove this for $\mu > \mu^{**}$. First prove some useful lemmas:

**Lemma 6.** $\forall k$, there exists $\mu_0 \in \Omega$ s.t. $V_{\mu_0}(\mu^*_k) > F(\mu^*_k)$.

**Proof.** Suppose $F=F_{k-1}$ at $\mu^{**}$. Equation (30) implies $\mu^*_k > \mu^{**} > \mu^*_{k-1}$. Consider:

$$U_k(\mu) = \max_{v \geq \mu} \frac{c \cdot F(v) - F(\mu) - F'(\mu)(v - \mu)}{f(\mu, v)}$$

$U_k$ is continuous by maximum theorem on $[\mu^{**}, \mu^*_k]$. Since $U_k(\mu^{**}) = F(\mu^{**})$, $\lim_{\mu \to \mu^*_k} U_k(\mu) = +\infty$, there exists $\mu_0$ s.t. $U_k(\mu_0) = F(\mu_0)$ and $U_k(\mu) > F(\mu) \forall \mu \in (\mu_0, \mu^*_k)$. Now consider $V_{\mu_0}(\mu)$. I claim that $V_{\mu_0}(\mu) > F(\mu) \forall \mu \in$

---

24 By definition of $\mu^{**}$, $\mu_0$ is bounded away from $0,1$ and Equation (31) implies conditions in Lemma 11 are satisfied.

25 Define $\mu_{m+1} = \mu^\circ$ and $\mu_{m_0} = \mu^{**}$ for consistency.
Assume not, then by intermediate value theorem, there exists $\mu'$ s.t $V_\mu'(\mu') \leq F(\mu)$ and $V'_{\mu_0}(\mu') \leq F(\mu)$. However, this implies

$$V_{\mu_0}(\mu') = \frac{c}{J(\mu', \mu)} \frac{F(\mu') - V_\mu(\mu') - V_{\mu_0}(\mu')}{\mu' - \mu} \geq U_k(\mu') > F(\mu')$$

Contradiction. Now assume $V_{\mu_0}$ hits $F$ at $\mu_0$. Then $U_{k+1}(\mu_0) \leq 0$ and $\lim_{\mu \to \mu_{k+1}(\mu)} = +\infty$, so we can find $V_{\mu_1}(\mu_{k+1}) > F(\mu_{k+1})$. By induction on $k$, Lemma 6 is true. Q.E.D.

**Lemma 7.** $\forall \mu_0 < \mu_1 \in \Omega$, let $I_1 = \{ \mu | V_\mu(\mu) > F(\mu) \}$. Then either $I_0 \cap I_1 = \emptyset$, or $I_1 \subset I_0$ and $V_{\mu_0} \geq V_{\mu_1}$.

**Proof.** The only possible contradiction of Lemma 7 is that $\exists \mu' \in \mu_0 \cap I_1$ s.t. $V_{\mu_1}(\mu') > V_{\mu_0}(\mu')$. Since at $\mu_1$, $V_{\mu_0}(\mu_1) > V_{\mu_1}(\mu_1) = F(\mu_1)$, by intermediate value theorem, there exists $\xi \in (\mu_1, \mu')$ s.t. $V_{\mu_1}(\xi) > V_{\mu_0}(\xi)$ and $V'_{\mu_1}(\xi) > V'(\xi)$. Since $\xi \in I_1$, there exists $\nu, m$ solving Equation (30) for $V_{\mu_1}(\xi)$:

$$V_{\mu_0}(\xi) = \frac{c}{\rho} \frac{F_n(\nu) - V_{\mu_0}(\nu) - V_{\mu_0}(\xi)}{J(\xi, \nu)} > \frac{c}{\rho} \frac{F_n(\nu) - V_{\mu_1}(\nu) - V_{\mu_1}(\xi)}{J(\xi, \nu)} = V_{\mu_1}(\xi) > V_{\mu_0}(\xi)$$

Contradiction. So Lemma 7 is true. Q.E.D.

**Lemma 8.** $\forall \nu \in \Omega \cup \{ \mu \} \cup \{ \text{inf } F, \sup F \}$ is totally bounded and equi-continuous on $[\mu^{**}, 1]$.

**Proof.** $V_{\nu}^\nu$ is bounded above by $\sup F(\mu)$ and below by $\inf F(\mu)$. Consider $V_{\nu}^\nu$. When $V_{\nu}^\nu(\mu) = F(\mu)$, obviously derivative is bounded by $\max |F'|$. When $V_{\nu}^\nu(\mu) > F(\mu)$, suppose $V_{\nu}^\nu(\mu) > \max |F'|$, then $F(\nu) - V_{\nu}^\nu(\mu) - V_{\nu}^\nu(\mu)(\nu - \mu) < F(\nu) - F(\mu) - F'(\nu)(\nu - \mu) \leq 0$, contradiction. By Lemma 11, $V_{\nu}^\nu > 0$. So $V_{\nu}^\nu$ are uniformly bounded in $[0, \max |F'|]$. Now consider $\forall \nu, \forall \nu \in \Omega, V_{\nu} \in \{ \text{inf } F, \sup F \} \Rightarrow \max \{ V_{\nu} \} \in \{ \text{inf } F, \sup F \}$. By Lemma 7, max$\{ V_{\nu} \}$ is piecewisely defined as $V_{\nu}$ on finite disjoint intervals. So its derivative is piecewisely defined as $V_{\nu}^\nu$, therefore bounded in $[0, \max |F'|]$. Therefore $\forall \nu$ is totally bounded and equi-continuous on $[\mu^{**}, 1]$. Q.E.D.

**Lemma 9.** There exists $\Delta$ s.t. $\forall \nu \in \Omega$, on $\{ \mu | V_\nu(\mu) > F(\mu) \}$, $V'(\mu)$ has Lipschitz parameter $\Delta$.

**Proof.** $\forall \mu \in (\mu_{k+1}, \mu_k), \nu$ is smooth in $\mu$ and $V_\nu^\nu > 0$, by envelope theorem:

$$V_\nu^\nu(\mu) = -\frac{c}{\rho} \frac{v - \mu}{J(\mu, \nu)} \left( V_\nu^\nu(\mu) + \frac{\rho}{c} V(\mu) H''(\mu) \right) > 0$$

$$V_\nu^\nu(\mu) + \frac{\rho}{c} V_\nu(\mu) H''(\mu) < 0$$

$V_\nu(\mu)$ is bounded by $\sup F$. It is easy to see that $\sup \Omega < \mu_n$ (where $n$ is the largest index). By Lemma 6, there is $\mu_0 \in \Omega$ s.t. $V_{\mu_0}(\mu_0) > F(\mu_0)$, By Lemma 7, $\sup \{ \mu | V_{\mu_0}(\mu) > F(\mu) \} < \nu(\mu_0) < 1$. Therefore, $\mu$ is bounded away from 1. Then by Assumption 2, $-H''(\mu)$ is bounded above. Therefore, $\Delta$ exists for all such $\mu$.

Then consider $\mu = \mu_n$, since $V_{\mu_n}^\nu$ is bounded on both side by $\Delta$, $V_{\mu_n}^\nu(\mu) \leq \Delta$. Therefore at $\mu V_\nu$ has Lipschitz parameter $\Delta$ by Kirszbraun theorem. Q.E.D.

1. Step 1: prove $V \in C[\mu^{**}, 1]$.
2. Sort all rational numbers in $[\mu^{**}, 1]$ as $\{ r_n \}$. $\forall N$, there exists $\mu_{n,M} \in \Omega$ s.t. $V(r_n) - V_{\mu_{n,M}}(r_n) < \frac{1}{N}$. Let $V_N = \max \{ V_{\mu_{n,N}} \}$, then $\{ V_N \} \subset V$ and $V_N$ converges to $V$ pointwisely on $\{ r_n \}$. Let $\hat{V} = \lim V_{\mu_{n}}$ by Lemma 8, $\hat{V} \in C[\mu^{**}, 1]$. By definition $\hat{V} \leq V$. Suppose $\hat{V}(\mu) < F(\mu)$, then there exists $V_{\mu_0}(\mu) > \hat{V}(\mu)$. Since both $V_{\mu_0}$ and $\hat{V}$ are continuous, $V_{\mu_0} > \hat{V}$ on an open interval, containing some $r_n$. Contradiction. So $\hat{V} = V \in C[\mu^{**}, 1]$. Let $\{ \mu \geq \mu^{**} | V(\mu) > F(\mu) \} = \bigcup I_m$ where $I_m$ are disjoint open intervals.
• Step 2: prove \( \forall I_m, \exists \mu_n \in \Omega \text{ s.t. } V(\mu) = \lim V_{\mu_n}(\mu) \) and \( V'(\mu) = \lim V_{\mu_n}'(\mu) \) on \( I_m \).

Pick any \( \mu \in I_m \). Let \( \Theta(\mu) = \{ \mu \in \Omega | V_{\mu}(\mu) > F(\mu) \} \). Then by definition of \( V(\mu) \), \( \Theta(\mu) \) is non-empty. Let \( \tilde{V} = \sup_{\mu \in \Theta(\mu)} V_{\mu} \). \forall N, there exists \( \mu_n, M \in \Theta(\mu) \text{ s.t. } \tilde{V}(\mu_n) - V_{\mu_n}(\mu_n) \leq \frac{1}{N} \). Since \( V_{\mu_n}(\mu) > F(\mu) \), by Lemma 7, there exists \( V_{\mu_n} = \max\{ V_{\mu_n} \} \). Therefore, \( \lim V_{\mu_n} = \tilde{V} \) on \( \{ r_n \} \). By Lemma 8 \( \tilde{V} = \lim \sup V_{\mu_n} \in C[\mu^{**},1] \). Now suppose \( V(\mu) > \tilde{V}(\mu) \), then there exists \( V_{\mu_n}(\mu) > F(\mu) \). \( \mu \in \Theta(\mu) \) by Lemma 7, contradiction. Therefore, \( \lim V_{\mu_n} = V \) on \( I_m \).

Let \( I_m = \{ a_m, b_m \} \). Now consider \( \{ V_{\mu_n} \} \). \( V_{\mu_n}(a_m) = F'(a_m) \). Lemma 9 implies that \( V_{\mu_n} \) are totally bounded and equi-continuous on \( I_m \). Therefore, there exists subsequence \( V_{\mu_k} \) being Cauchy w.r.t. sup norm on \([a_m,b_m]\). So \( V \) as limit of \( V_{\mu_k} \) is differentiable on \([a_m,b_m]\) and \( V' = \lim V_{\mu_n}' \).\(^{26}\)

• Step 3: prove \( \forall I_m, \exists \mu^m \in \Omega \text{ s.t. } V(\mu) = V_{\mu^m} \) on \( I_m \).

Let \( \mu^m = \inf I_m \). By step 2, it is easy to verify that \( \mu_n \rightarrow \mu^m \). Then since Equation (31) is continuous in \( \mu \), it is satisfied at \( \mu^m \) and \( \mu^m \in \Omega \). Since both \( V_{\mu_n} \) and \( V_{\mu_k} \) converges on \( I_m \), Equation (30) is satisfied for \( V \) on \( I_m \).

Let \( F(\mu^m) = F_k(\mu^m) \).

As an intermediate step, I first prove that Equation (30) is solved for \( k' > k \) in a non-degenerate neighbour of \( \mu^m \). Take any \( \mu' > \mu \) s.t. \( V(\mu') > F(\mu') \), since \( V(\mu^m) = F_k(\mu^m) \), there exists \( \mu^* \in (\mu^m, \mu') \) and \( \epsilon > 0 \) s.t. \( \forall \mu \in (\mu^m, \mu^*) \) \( V(\mu) - F_k(\mu^m) < V(\mu') - F_k(\mu') - \epsilon \). I claim that Equation (30) is solved at all \( \mu \in (\mu^m, \mu^*) \) with \( k' > k \). Suppose not, then for \( n \) sufficiently large:

\[
V_{\mu_n}(\mu) = \frac{c}{\rho} F_k(v) - V_{\mu_n}(\mu) - V_{\mu_n}'(\mu)(v - \mu) \leq \frac{c}{\rho} F_k(v) - F_k(\mu^m) - V_{\mu_n}'(\mu)(v - \mu) = (F_k - V_{\mu_n}'(\mu))(v - \mu)
\]

Therefore \( F_k \geq V_{\mu_n}'(\mu) \). By construction of \( V_{\mu_n} \) at any \( \mu^m \geq \mu \) Equation (30) is solved with \( k \), therefore \( F_k \geq V_{\mu_n}(\mu) \) holds for all \( \mu^m \geq \mu \). This implies \( \forall \mu^m \geq \mu, V_{\mu_n}(\mu') - F_k(\mu^m) \leq V_{\mu_n}(\mu) - V(\mu') - F_k(\mu') - \epsilon \). Take \( n \rightarrow \infty \) and \( \mu^m = \mu^* \), contradiction. Therefore, Equation (30) is solved at all \( \mu \in (\mu^m, \mu^*) \) for \( V(\mu) \) with \( k' > k \).

Now consider \( V_{\mu^n}(\mu) \). By my construction, suppose \( V_{\mu^n} \) is updated up to action \( k + 1 \). I claim that \( V_{\mu^n} = V \) when \( \mu \in [\mu^*, \mu^*] \). Suppose not true, then there exists \( \mu \) at which \( V_{\mu^n}(\mu) < V(\mu), V_{\mu^m}(\mu) < V'(\mu) \). It is easy to verify that Equation (30) is violated at \( V_{\mu_n}(\mu) \). Therefore, if \( V_{\mu^n} \neq V \), it must happen in \( (\mu^*, b_m) \). Again we can find \( \mu \in (\mu^*, b_m) \) s.t. \( V_{\mu^m}(\mu) < V(\mu), V_{\mu^n}(\mu) < V'(\mu) \), which is not possible. So \( V(\mu) = V_{\mu^m}(\mu) \) on \( I_m \).

To sum up, \( V \) can be represented as:

\[
V(\mu) = \begin{cases} V_{\mu^*}(\mu) & \text{if } \mu \in [\mu^*, \mu^{**}] \\ V_{\mu^n}(\mu) & \text{if } \mu \in \mu^n \\ F(\mu) & \text{otherwise} \end{cases}
\]

Now I prove smoothness of \( V(\mu) \) on \( [\mu^*, 1] \). By Lemma 9, \( \forall \mu \in I_m \):

\[
\begin{align*}
F'(a_m) - \Delta |\mu - a_m| &\leq V'(\mu) \leq F'(a_m) + \Delta |\mu - a_m| \\
F'(b_m) + \Delta |\mu - b_m| &\geq V'(\mu) \geq F'(b_m) - \Delta |\mu - b_m|
\end{align*}
\]

Therefore \( |V'(\mu) - F'(\mu)| \) is bounded by \( \Delta |I_m| \). Define:

\[
V_n(\mu) = \begin{cases} V_{\mu^n}(\mu) & \text{when } \mu \in I_m, m \leq n \\ F(\mu) & \text{otherwise} \end{cases}
\]

Then \( V_n(\mu) \rightarrow V(\mu) \). By Lemma 6, we can without loss assume first \( n \) \( V_{\mu^n} \) have \( I_m \) covering \( \mu_n \). Fix \( n, \forall \mu, \forall m \geq n, \text{if } \mu \in I_m \text{ and } m \leq n \text{ or } \mu \notin \bigcup I_m, \text{then } V_n'(\mu) = V_m'(\mu), \text{else if } \mu \in I_m \text{, } m > n \text{, then } |V_n'(\mu) - F'(\mu)| \text{ and } |V_m'(\mu) - F'(\mu)|

\(^{26}\) This result is ex. 14.2.7 from Tao (2016).
are all bounded by $\Delta |I_m|$. Therefore, $V'_n(\mu)$ is a Cauchy sequence. Then $V'_n(\mu) \to V'(\mu)$ pointwise. Since each $V'_n$ is continuous, $V$ is a smooth function on $[0,1]$ and $V'=F'$ when $V=F$.

**Unimprovability:**

Finally, I prove unimprovability of $V(\mu)$.

- **Step 1:** I first show that $V(\mu)$ solves the following problem:

\[
V(\mu) = \max_{\nu, m} \left\{ F(\mu) \max_{\nu, m} \frac{c}{\rho} \frac{F_m(v) - V(\mu) - V'(\mu)(v-\mu)}{J(\mu, v)} \right\}
\]

Equation (P-C) is the maximization problem over all confirmatory evidence seeking with immediate decision making upon arrival of signals. Equation (P-C) is implied by Equation (30) for $\mu \in E$. So it is sufficient to prove Equation (P-C) for $\mu \in E^c$. Suppose there exists $\mu \geq \mu^*$ s.t. Equation (P-C) is violated. Let $F(\mu) = F_k(\mu)$. Then without loss we can assume that:

\[
U(\mu) = \max_{\nu, k > k} \frac{c}{\rho} \frac{F_k(v) - F_k(\mu) - F_k(v-\mu)}{J(\mu, v)} > F_k(\mu)
\]

By Lemma 6, there exists $l_k$ s.t. $\mu_k \in I_k$. At $b_k = \sup I_k$, $U(b_k) \leq F_k(b_k)$. Therefore, since $U(\mu)$ is continuous there exists largest $\mu' < \mu$ s.t. $U(\mu') = F_k(\mu')$. Then Equation (31) is satisfied at $\mu'$ so consider $V_{\mu'}$. Since $V_{\mu'}(\mu) \leq V(\mu) = F_k(\mu)$, there exists $\mu''(\mu, \mu)$ s.t. $V_{\mu''}(\mu) \leq F_k(\mu)$ and $V_{\mu''}(\mu) < F_k(\mu)$. Therefore $U(\mu'') > F_k(\mu'')$ implies $V_{\mu''}(\mu'') > F_k(\mu'')$, contradiction. Apply a symmetric argument to $\mu < \mu^*$, I proved Equation (P-C).

- **Step 2:** Then I will show that $V(\mu)$ solves the following problem:

\[
V(\mu) = \max_{\nu, m} \left\{ F(\mu) \max_{\nu, m} \frac{c}{\rho} \frac{V(v) - V(\mu) - V'(\mu)(v-\mu)}{J(\mu, v)} \right\}
\]

Equation (P-D) is the maximization problem over all confirmatory evidence seeking strategies. It has less constraint than Equation (P-C): when a signal arrives and posterior belief $v$ is realized, the DM is allowed to continue experimentation instead of being forced to take an action.

We only show that case $\mu \geq \mu^*$ and a totally symmetric argument applies to $\mu < \mu^*$. Suppose Equation (P-C) is violated at $\mu$, then there exists $\nu'$ such that:

\[
V(\mu) = \max_{\nu > \mu, m} \frac{c}{\rho} \frac{F_m(v) - V(\mu) - V'(\mu)(v-\mu)}{J(\mu, v)} < \frac{c}{\rho} \frac{V(\nu') - V(\mu) - V'(\mu)(\nu' - \mu)}{J(\mu, \nu')}
\]

Let $\tilde{V} = V(\mu)$. Suppose the maximizer is $\nu, m$. Optimality implies first order conditions Equation (37) and Equation (36):

\[
\begin{aligned}
& F_m(v) + \frac{c}{\rho} \tilde{V} H'(v) = V'(\mu) + \frac{c}{\rho} \tilde{V} H'(\mu) \\
& \left( F_m(v) + \frac{c}{\rho} \tilde{V} H'(v) \right) - \left( V(\mu) + \frac{c}{\rho} \tilde{V} H(\mu) \right) = \left( V'(\mu) + \frac{c}{\rho} \tilde{V} H'(\mu)(v-\mu) \right)
\end{aligned}
\]

We define $L(V, \lambda, \mu)(v)$ and $G(V, \lambda, \mu)(v)$ as:

\[
\begin{aligned}
& L(V, \lambda, \mu)(v) = (V(\mu) + \lambda H(\mu)) + (V'(\mu) + \lambda H'(\mu))(v-\mu) \\
& G(V, \lambda, \mu)(v) = V(v) + \lambda H(v)
\end{aligned}
\]
Consider:

First, since diffusion experiments. 2) evidence seeking of all possible posteriors instead of just confirmatory evidence.

Therefore is convex and weakly larger than zero. However by Equation (37)

Then since $L$ is a linear function of $v$ and $G(F_m, \frac{p}{c} \hat{V})(v)$ is a strictly concave smooth function of $v$. Consider:

$$L\left(V, \frac{p}{c} \hat{V}, \mu \right)(v) - G\left(F_m, \frac{p}{c} \hat{V} \right)(v)$$

Equation (37) implies that it attains minimum 0 at $v$. For any $m'$ other than $m$,

$$L\left(V, \frac{p}{c} \hat{V}, \mu \right)(v) - G\left(F_m', \frac{p}{c} \hat{V} \right)(v)$$

is convex and weakly larger than zero. However by Equation (32):

$$L\left(V, \frac{p}{c} \hat{V}, \mu \right)(v') - G\left(V, \frac{p}{c} \hat{V} \right)(v') = -\left( V'(v') - V(\mu) - V'(\mu)(v' - \mu) + \frac{p}{c} \hat{V} J(\mu, v') \right) < 0$$

Therefore $L\left(V, \frac{p}{c} \hat{V}, \mu \right)(v) - G\left(V, \frac{p}{c} \hat{V} \right)(v)$ has strictly negative minimum. Suppose it’s minimized at $\tilde{\mu}$ ($\tilde{\mu} > \mu$ since $L(V, \lambda, \mu)(\mu) = G(V, \lambda)(\mu)$). Then FOC is a necessary condition:

$$V'(\mu) + \frac{p}{c} \hat{V} H'(\mu) = V'(\tilde{\mu}) + \frac{p}{c} \hat{V} H'(\tilde{\mu})$$

Consider:

$$L\left(V, \frac{p}{c} \hat{V}, \mu \right)(v(\tilde{\mu})) - G\left(F_m, \frac{p}{c} \hat{V} \right)(\tilde{\mu}) = L\left(V, \frac{p}{c} \hat{V}, \mu \right)(v(\tilde{\mu})) - G\left(F_m, \frac{p}{c} \hat{V} \right)(v(\tilde{\mu}))$$

$$+ V(\tilde{\mu}) - V(\mu) + \frac{p}{c} \hat{V} (H(\tilde{\mu}) - H(\mu)) - \left( V'(\mu) + \frac{p}{c} \hat{V} H'(\mu) \right)(\tilde{\mu} - \mu)$$

$$\geq V(\tilde{\mu}) - V(\mu) + \frac{p}{c} \hat{V} (H(\tilde{\mu}) - H(\mu)) - \left( V'(\mu) + \frac{p}{c} \hat{V} H'(\mu) \right)(\tilde{\mu} - \mu)$$

$$= G\left(V, \frac{p}{c} \hat{V}, \mu \right)(\tilde{\mu}) - L\left(V, \frac{p}{c} \hat{V}, \mu \right)(\tilde{\mu}) > 0$$

In the first equality I used Equation (37) at $\tilde{\mu}$. In first inequality I used suboptimality of $\tilde{\mu}$ at $\mu$. However for $m'$ and $v(\tilde{\mu})$ being optimizer at $\tilde{\mu}$:

$$0 = L\left(V, \frac{p}{c} \hat{V}, \mu \right)(v(\tilde{\mu})) - G\left(F_m, \frac{p}{c} \hat{V}(\tilde{\mu}) \right)(v(\tilde{\mu}))$$

$$= L\left(V, \frac{p}{c} \hat{V}, \mu \right)(v(\tilde{\mu})) - G\left(F_m, \frac{p}{c} \hat{V} \right)(v(\tilde{\mu}))$$

$$+ \frac{p}{c} \left( V(\tilde{\mu}) - V(\mu) \right) (H(\tilde{\mu}) - H(v(\tilde{\mu}))) + H'(\tilde{\mu})(v(\tilde{\mu}) - \tilde{\mu})$$

$$> \frac{p}{c} \left( V(\tilde{\mu}) - V(\mu) \right) (H(\tilde{\mu}) - H(v(\tilde{\mu}))) + H'(\tilde{\mu})(v(\tilde{\mu}) - \tilde{\mu})$$

Contradiction. Therefore, I proved Equation (P-D).

- **Step 3:** We show that $V$ satisfies Equation (18), which is less restrictive than Equation (P-D) by allowing 1) diffusion experiments. 2) evidence seeking of all possible posteriors instead of just confirmatory evidence.

First, since $V$ is smoothly increasing and has a differentiable optimizer $v$, envelope theorem implies:

$$V'(\mu) = \frac{c - V''(\mu)(v - \mu)}{H'(\mu)} + V(\mu) \frac{-H''(\mu)(v - \mu)}{J(\mu, v)}$$

$$= -\frac{c}{\rho} \frac{v - \mu}{J(\mu, v)} \left( V''(\mu) + \frac{p}{c} V(\mu) H'(\mu) \right) > 0$$

$$\implies V''(\mu) + \frac{p}{c} V(\mu) H'(\mu) < 0$$

Therefore, allocating to diffusion experiment will always be suboptimal. What’s more consider:

$$V^+(\mu) = \max_{v < \mu} \frac{c}{\rho} \frac{V(v) - V(\mu) - V'(\mu)(v - \mu)}{J(\mu, v)}$$
\[
\rightarrow V^{-}(\mu) = -\frac{c}{\rho} V''(\mu) \left( V''(\mu) + \frac{\rho}{c} V''(\mu) \right)
\]

\[V^{-}(\mu^*) = V(\mu^*)\] and whenever \(V(\mu) = V^{-}(\mu)\), so \(V^{-}(\mu) < 0\). Therefore, \(V^{-}(\mu)\) can never cross \(V(\mu)\) from below. That is to say:

\[
\rho V(\mu) = \max \left\{ \rho F(\mu), \max_{v,p} (V(v) - V(\mu) - V'(\mu)(v-\mu)) + \frac{1}{2} V''(\mu) \sigma^2 \right\}
\]

s.t. \( p \frac{\partial}{\partial \mu} V(\mu, v) + \frac{1}{2} H''(\mu) \sigma^2 \leq c \)

To sum up, I constructed a policy function \(v(\mu)\) and value function \(V(\mu)\) solving Equation (18). Now consider the four properties in Theorem 2. First, by my construction algorithm, in the case \(\mu^* \in [0,1]\), I can replace \(\mu^*\) with \(\mu^* \in (0,1)\). Therefore WLOG \(\mu^* \in (0,1)\). Second, \(E = \{ \mu \in [0,1] \mid V(\mu) > F(\mu) \}\) is a union of disjoint open intervals. By my construction, \(V(\mu) = V_{\mu^*}(\mu)\). On each \(I_m\), \(v_{\mu^*}(\mu)\) is strictly decreasing and jumps down at finite \(\tilde{\mu}_k\)'s. Finally, uniqueness argument in Lemma 11 implies that \(v\) is uniquely determined by FOC. Therefore, except for those discontinuous points of \(v\), \(v\) is uniquely defined. Number of such discontinuous points is countable, thus of zero measure.

Q.E.D.

**Lemma 10.** Define \(\nabla^+\) and \(\nabla^-\):

\[
\nabla^+(\mu) = \max_{v \geq \mu, m} \frac{F_m(v)}{1 + \frac{c}{\rho} J(\mu, v)}
\]

\[
\nabla^-(\mu) = \max_{v \leq \mu, m} \frac{F_m(v)}{1 + \frac{c}{\rho} J(\mu, v)}
\]

There exists \(\mu^* \in [0,1]\) s.t. \(\nabla^+(\mu) > \nabla^-(\mu) \forall \mu \geq \mu^*; \nabla^+(\mu) \leq \nabla^-(\mu) \forall \mu \leq \mu^*\).

**Proof.** We define function \(U_m^+\) and \(U_m^-\) as following:

\[
U_m^+(\mu) = \max_{v > \mu} \frac{F_m(v)}{1 + \frac{c}{\rho} J(\mu, v)}
\]

\[
U_m^- (\mu) = \max_{v < \mu} \frac{F_m(v)}{1 + \frac{c}{\rho} J(\mu, v)}
\]

First of all, I solve \(U_m^+, U_m^-\) on interior \(\mu \in (0,1)\). Since \(F_m(\mu)\) is a linear function, \(J(\mu, v) \equiv 0\) is smooth, the objective function will be a continuous function on compact domain. Therefore both maximization operators are well defined. Existence is already guaranteed, therefore I can refer to first order condition to characterize the maximizer:

FOC: \(F_m' \left( 1 + \frac{\rho}{c} J(\mu, v) \right) + F_m(v) \frac{\rho}{c} (H'(v) - H'(\mu)) = 0 \) \hspace{1cm} (33)

SOC: \(\frac{\rho}{c} F_m' \left( H'(v) - H'(\mu) \right) \) \hspace{1cm} (34)

First discuss solving for \(v \geq \mu\). Since \(1 + \frac{c}{\rho} J > 0\), \(H'' < 0\), \(H'(v) - H'(\mu) < 0\) and inequality is strict when \(v > \mu\). Therefore, if \(F_m' < 0\), FOC being held will imply SOC being strictly positive at \(v > \mu\). So \(\forall F_m' < 0\), optimal \(v\) will be boundary. What’s more:

\[
\frac{F_m(\mu)}{1 + \frac{c}{\rho} J(\mu, \mu)} = F_m(\mu) > F_m(1) > \frac{F_m'(1)}{1 + \frac{c}{\rho} J(\mu, 1)}
\]

So \(U_m^+(\mu) = F_m(\mu)\). If \(F_m' = 0\), then \(\forall v > \mu\):

\[
\frac{F_m(\mu)}{1 + \frac{c}{\rho} J(\mu, \mu)} = F_m(\mu) = F_m(\mu) > \frac{F_m'(\mu)}{1 + \frac{c}{\rho} J(\mu, \mu)}
\]
Therefore $\forall F'_m \leq 0$, $U'_m(\mu)=F_m(\mu)$. Then consider the case $F'_m>0$. SOC is strictly negative when FOC holds and $\nu>\mu$. Therefore solution of FOC characterizes maximizer. Consider:

$$\lim_{\nu \to \mu} F'_m(1+\frac{\rho}{c}J(\mu,\nu)) + F_m(\nu)\frac{\rho}{c}(H'(\nu)-H'(\mu)) = F'_m > 0$$

Moreover, Equation (33) is strictly positive when $\nu=\mu$. This implies $U'_m(\mu) > F_m(\mu)$ when $F'_m > 0$.

New consider limit of $U'_m$ when $\mu \to 0,1$. When $\mu \to 1$, $U'_m(\mu) \leq \max_{\nu \geq \nu} F_m(\nu) = F(1)$. When $\mu \to 0$, consider FOC Equation (33):

$$\lim_{\mu \to 0} F'_m(1+\frac{\rho}{c}J(\mu,\nu)) + F_m(\nu)\frac{\rho}{c}(H'(\nu)-H'(\mu)) = \lim_{\mu \to 0} F'_m(1+\frac{\rho}{c}J(\nu,\mu)) + F_m(\mu)\frac{\rho}{c}(H'(\nu)-H'(\mu)) = -\infty$$

Therefore, when $\mu \to 0$, optimal $\nu \to 0$. Therefore $F_m(\nu) \leq F_m(\nu) \to F_m(0)$. To conclude, $U'_m(\mu) = F_m(\mu)$ when $\mu = 0,1$.

Let $\bar{m}$ be the first $F'_m > 0$ (not necessarily exists). Let:

$$U^+(\mu) = \max_{m \geq \bar{m}} U'_m(\mu)$$

Then $U^+(\mu)$ will be a strictly increasing function when $\bar{m}$ exists. Symmetrically I can define $\underline{m}$ to be last $F'_m < 0$ and:

$$U^-(\mu) = \max_{m \leq \underline{m}} U'_m(\mu)$$

There are three cases:

- **Case 1:** when $F$ is not monotonic, then both $U^+$ and $U^-$ exists. Moreover, $F(0) > F_m(0)$ and $F(1) > F_m(1)$. Therefore, $U^+(0) < U^-(0)$ and $U^+(1) > U^-(1)$. There must exists unique $\mu^* \in (0,1)$ s.t. $U^+(\mu^*) = U^-(\mu^*)$.

- **Case 2:** when $F' \geq 0$, then define $\mu^* = 0$.

- **Case 3:** when $F' \leq 0$, then define $\mu^* = 1$.

Finally, define $\bar{V}$:

$$\bar{V}^+(\mu) = \max\{F(\mu), U^+(\mu)\}$$

$$\bar{V}^-(\mu) = \max\{F(\mu), U^-(\mu)\}$$

$$\bar{V}(\mu) = \max\{\bar{V}^+(\mu), \bar{V}^-(\mu)\}$$

Given our construction, $\mu^*$ always exists and satisfies the conditions in Lemma 10. Q.E.D.
Lemma 11. Assume $\mu_0 \geq \mu^*$, $F_m' \geq 0$, $V_0, V_0' \geq 0$ satisfies:

$$\begin{cases}
V(\mu_0) \geq V_0 \geq F_m(\mu_0) \\
V_0 = \max_{v \geq \mu_0} \frac{c F_m(v) - V_0 - V_0'(v - \mu_0)}{J(\mu_0, v)}
\end{cases}$$

Then there exists a $C^{(1)}$ smooth and strictly increasing $V(\mu)$ defined on $[\mu_0, 1]$ satisfying

$$V(\mu) = \max_{v \geq \mu} \frac{c F_m(v) - V(\mu) - V'(\mu)(v - \mu)}{J(\mu, v)}$$  \hspace{1cm} (35)

and initial condition $V(\mu_0) = V_0, V'(\mu_0) = V_0'$.

**Proof.** We start from deriving the FOC and SOC for Equation (35):

**FOC:**

$$F_m' - V'(\mu) + \frac{F_m(v) - V(\mu) - V'(\mu)(v - \mu)}{J(\mu, v)}(H'(v) - H'(\mu)) = 0$$  \hspace{1cm} (36)

**SOC:**

$$H'(v) - H'(\mu) \left( \frac{F_m'(v) - V'(\mu)}{J(\mu, v)} + \frac{F_m(v) - V(\mu) - V'(\mu)(v - \mu)}{J(\mu, v)^2}(H'(v) - H'(\mu)) \right) + \frac{H''(v)}{J(\mu, v)}(F_m(v) - V(\mu) - V'(\mu)(v - \mu)) \leq 0$$  \hspace{1cm} (37)

If we impose feasibility:

$$V(\mu) = \frac{c F_m(v) - V(\mu) - V'(\mu)(v - \mu)}{J(\mu, v)}$$  \hspace{1cm} (38)

FOC and SOC reduces to:

**FOC:**

$$F_m' - V'(\mu) + \frac{F_m(v) - V(\mu) - V'(\mu)(v - \mu)}{J(\mu, v)}(H'(v) - H'(\mu)) = 0$$  \hspace{1cm} (39)

**SOC:**

$$H'(v) - H'(\mu) \left( \frac{F_m'(v) - V'(\mu)}{J(\mu, v)} + \frac{F_m(v) - V(\mu) - V'(\mu)(v - \mu)}{J(\mu, v)^2}(H'(v) - H'(\mu)) \right) + \frac{H''(v)}{J(\mu, v)}(F_m(v) - V(\mu) - V'(\mu)(v - \mu)) \leq 0$$  \hspace{1cm} (40)

Let us proceed as following, I use FOC and feasibility to derive an ODE system with initial value defined by $V_0, V_0'$. Then I prove that the solution $V$ must be strictly positive. Therefore, SOC is strict at the point where FOC is satisfied, the solution must be locally maximizer. Moreover, since $H'(v) - H'(\mu) < 0$, when FOC is positive, SOC must be negative, then FOC will have a unique solution. Therefore the solution I get from the ODE system will be solution to problem Equation (35).

$$\begin{cases}
\text{Equation (36)} \implies V(\mu) = \frac{F_m(v) - V'(\mu)(v - \mu)}{1 + \frac{c}{\xi} J(v, \mu)} \\
\text{Equation (37)} \implies V'(\mu) = F_m' + \frac{\xi}{c} V'(\mu)(H'(v) - H'(\mu)) \\
V(\mu) = \frac{F_m(\mu)}{1 - \frac{\xi}{c} J(\nu, \mu)} \\
V'(\mu) = F_m' + \frac{\xi}{c} F_m(\mu)(H'(v) - H'(\mu)) - \frac{\xi}{c} J(\nu, \mu) \cdot \frac{\xi}{c} F_m(\mu)(H'(v) - H'(\mu)) $$
\end{cases}$$  \hspace{1cm} (39)

Consistency of Equation (39) implies that $v = v(\mu)$ is characterized by the following ODE:

$$\frac{\partial}{\partial \mu} F_m(\mu) + \frac{\partial}{\partial \nu} F_m(\mu) \cdot \gamma = F_m' + \frac{\xi}{c} F_m(\mu)(H'(v) - H'(\mu)) \cdot \frac{\xi}{c} F_m(\mu)(H'(v) - H'(\mu))$$  \hspace{1cm} (40)

Simplifying Equation (40):

$$\frac{F_m'}{1 - \frac{\xi}{c} J(\nu, \mu)} + \frac{\xi}{c} F_m(\mu)(H'(v) - H'(\mu)) \cdot \frac{\xi}{c} F_m(\mu) \cdot \frac{\xi}{c} F_m(\mu)(H'(v) - H'(\mu)) \cdot \frac{\xi}{c} F_m(\mu)(H'(v) - H'(\mu))$$

54
\[ F_m^\prime + \frac{\xi}{\rho} (-F_m J(v,\mu) + F_m(\mu)(H'(v) - H'(\mu))) \]
\[
\Rightarrow F_m(\mu)(H'(v) - H'(\mu)) + F_m(\mu)H''(v)(\mu - v) \dot{v} = (-F_m J(v,\mu) + F_m(\mu)(H'(v) - H'(\mu))) (1 - \frac{\rho}{\xi} J(v,\mu))
\]
\[
\Rightarrow F_m(\mu)H''(v)(\mu - v) \dot{v} = -F_m J(v,\mu)(1 - \frac{\rho}{\xi} J(v,\mu)) F_m(\mu)(H'(v) - H'(\mu))
\]
\[
\Rightarrow \dot{v} = \frac{F_m(1 - \frac{\rho}{\xi} J(v,\mu)) + \frac{1}{\xi} F_m(\mu)(H'(v) - H'(\mu))}{F_m(\mu)H''(v)(\mu - v)}
\]

Since I want to solve for \( V_0 \) on \([\mu_0, 1] \), I solve for \( \nu \) at \( \mu_0 \) as the initial condition of ODE for \( \nu \). To utilize \textbf{Lemma B.2}, I need to first verify the inequality condition in \textbf{Lemma B.2}.

The FOC characeterizing \( \nu \) is \textbf{Equation (39)}:

\[
(F_m^\prime - V_0')(1 - \frac{\rho}{\xi} J(\nu_0,\mu_0)) + \frac{\rho}{\xi} F_m(\mu_0)(H'(\nu_0) - H'(\mu_0)) = 0
\]

\[
\Leftrightarrow F_m^\prime (1 + \frac{\rho}{\xi} J(\nu_0,\mu_0)) + \frac{\rho}{\xi} F_m(\nu_0)(H'(\nu_0) - H'(\mu)) = V_0'(1 - \frac{\rho}{\xi} J(\nu_0,\mu_0))
\]

\[
\Leftrightarrow F_m(\mu_0)(F_m^\prime (1 + \frac{\rho}{\xi} J(\nu_0,\mu_0)) + \frac{\rho}{\xi} F_m(\nu_0)(H'(\nu_0) - H'(\mu))) = V_0'F_m(\mu_0)(1 - \frac{\rho}{\xi} J(\nu_0,\mu_0))
\]

Since \( V_0 = \frac{F_m(\mu_0)}{1 - \frac{\rho}{\xi} J(\nu_0,\mu_0)} \geq 0 \), LHS is weakly positive. This satisfies the condition in \textbf{Lemma B.2}. Then \textbf{Lemma B.2} guarantees existence of \( v(\mu) \), which is continuously decreasing from \( \mu_0 \) until it hits \( v(\mu) = \mu \). Suppose \( \nu \) is minimized at \( \bar{\mu}_m < 1 \), define \( V(\mu) \) as following:

\[
V(\mu) = \begin{cases} 
\frac{F_m(\mu)}{1 - \frac{\rho}{\xi} J(v(\mu),\mu)} & \text{if } \mu \in [\mu_0, \bar{\mu}_m) \\
F_m(\mu) & \text{if } \mu \in [\bar{\mu}_m, 1]
\end{cases}
\]

Then I prove the properies of \( V \):

1. When \( \mu \to \bar{\mu}_m, v(\mu) \to \mu \). Therefore \( J(v,\mu) \to 0 \). This implies \( V(\mu) \to F_m(\mu) \). So \( V \) is continuous.

2. By \textbf{Equation (39)}, when \( \mu \in [\mu_0, \bar{\mu}_m) \):

\[
V'(\mu) = F_m^\prime + \frac{F_m(\mu)(H'(v(\mu)) - H'(\mu))}{1 - \frac{\rho}{\xi} J(v(\mu),\mu)}
\]

When \( \mu \to \bar{\mu}_m, H'(v(\mu)) - H'(\mu) \to 0, J(v(\mu),\mu) \to 0 \). Thus \( V'(\mu) \to F_m^\prime \). So \( V' \) will be continuous everywhere on \([\mu_0, 1] \). \( V \in C^1([\mu_0, 1]) \).

3. Rewrite \textbf{Equation (39)} on \([\mu_0, 1] \):

\[
V'(\mu) = \frac{F_m^\prime (1 + \frac{\rho}{\xi} J(\mu,\nu)) + F_m(\nu)(H'(v) - H'(\mu))}{1 - \frac{\rho}{\xi} J(v,\mu)}
\]

According to proof of \textbf{Lemma B.2}, \( V'(\mu) > 0 \forall \mu \in (\mu_0, 1] \). Moreover since \( V^0 \geq 0, V(\mu) > 0 \forall \mu \in (\mu_0, 1] \).

\[ Q.E.D. \]
Reference


Che, Y.-K. and Mierendorff, K. (2016). Optimal sequential decision with limited attention. 3, 4, 30

Clark, A. (2016). Contracts for information acquisition. unpublished manuscript. 4


Hébert, B. and Woodford, M. (2016). Rational inattention with sequential information sampling. 4, 23

Hörner, J. and Skrzypacz, A. (2016). Learning, experimentation and information design. 4


Ke, T. T. and Villas-Boas, J. M. (2016). Optimal learning before choice. 4


Matyskova, L. (2018). Bayesian persuasion with costly information acquisition. unpublished manuscript.


