Optimal Sequential Experimentation

Jorge D. Ramos-Mercado*

March 31, 2024

Abstract

An experimenter ($\mathcal{E}$) learns about a payoff-relevant state. He does so by flexibly managing a jump-diffusion signal with state-dependent dynamics. He controls the diffusion's precision and the arrival rate of jumps. If jumps never arrive, the signal is feasible in Moscarini and Smith (2001). $\mathcal{E}$ further faces costs rising convexly with the flow amount of information generated. My main result is that, in contrast to Zhong (2022), restricting attention to pure-diffusion signals is without loss. Intuitively, $\mathcal{E}$'s problem can be reformulated into a dynamic, constrained information acquisition problem. In said problem, $\mathcal{E}$ picks how much information to acquire from the diffusion and jumps, but some information must be acquire from the diffusion. Also, both types of information are perfect substitutes and (thus) managing a pure-diffusion signal is without loss. In a numerical example, I further illustrate how (if ever) information from jumps might be used.

I study the problem of sequential experimentation i.e., an experimenter ($\mathcal{E}$) learns about a state by explicitly managing a data generating process (DGP). In particular, $\mathcal{E}$ flexibly controls a jump-diffusion's state-dependent dynamics and faces costs convexly increasing in the flow amount of information generated. Suppose that $\mathcal{E}$ could abstract away from managing a DGP and, instead, acquires information over time. In such case, Zhong (2022) proved that it is optimal to acquire a compensated, pure-jump process signal confirming the most likely state. When $\mathcal{E}$ is required to flexibly experiment, however, I find a unique,

*Ramos-Mercado: EGADE Business School, Tecnológico de Monterrey, ramos237@tec.mx.

pure-diffusion signal. In other words, it is without loss of generality that the experimenter restricts attention to a DGP allowing him to gradually learn over time.

Wald (1947) first studied the problem of sequential experimentation. $\mathcal{E}$ acquires a sequence of iid random variables (i.e., signals) that are informative of an unknown, payoff-relevant state. Each signal models a random process through which a datum is generated: a DGP. $\mathcal{E}$ decides when to stop experimenting and make an irreversible decision. Moscarini and Smith (2001, MS) extended this setting by allowing the experimenter to flexibly control a diffusion's precision whose drift depends on the state. Intuitively, $\mathcal{E}$ flexibly controls the precision of the data generated over time. Such extension allows for a tractable characterization of the optimal experiment. MS further assume the experimentation flow costs increase convexly in the signal's precision. They find that in the optimal experiment's precision is strictly increasing in the value of experimenting.

Their model, nevertheless, forces $\mathcal{E}$ to only learn gradually i.e., acquiring Gaussian information. This assumption excludes experiments that infrequently produce very precise information i.e., Poisson Information. Such information is generated by jump processes and are commonplace in Physics, Chemistry, Earth Science, etc.. A general setting, in continuous-time, would allow $\mathcal{E}$ to generate both types of information and Zhong (2022, $Z22$) finds that such extension matters.

$Z22$ considered a reduced-form approach. $\mathcal{E}$ directly picks a Martingale, Lévy process for his beliefs i.e., he can acquire both types of information and the problem of managing a DGP is set aside. He further assumes that costs increase convexly in the flow amount of $information$ produced. These costs follow from the rational inattention literature e.g., Sims (2003), Hébert and Woodford (2021), Caplin et al (2022), Macowiak et al (2023).[1] Such assumption is further useful since it allows for the tractable comparison of costs across belief processes. $Z22$ finds no value from acquiring Gaussian information.

For an experiment to only generate Poisson information, $\mathcal{E}$ must generate noiseless information. This is a stark assumption as all experiments generate noise and issues pertaining to data cleaning are a key concern in experiment design. For example, data cleaning takes 60 percent of data scientists' (Press 2016) and 80 percent of clinical researchers'

[1]Said approach has garnered critique from Denti et al (2022) since this cost structure is inconsistent with a model of experiment specific costs.

(Rozario et al 2017) time.

For this reason, I expand the set of experiments considered in MS to allow for experiments generating data via jumps. This allows me to study how an agent optimally and flexibly learns when they are required to generate the utilized data themselves. In addition, I adopt the rational inattention cost function to tractably compare the costs of experiments. This allows me to tractably and systematically model the way in which costs are modeled in this general setting.

I then ask the following question. Is $\mathcal{E}$ strictly better off now that he can acquire both types of information? The answer is no and the intuition is straightforward. $\mathcal{E}$'s problem can be reformulated into one where he directly picks how much flow Poisson and Gaussian information to acquire. But the technological constraint forces $\mathcal{E}$ to always generate some Gaussian information. I also find Gaussian and Poisson information to be perfect substitutes. Thus, there is no added value from generating Poisson information.

This observation implies that some pure-diffusion experiment is optimal, but it does not characterize said experiment. How does the change in cost function alter the optimal experiment? I find that the flow amount of information acquired in my model is a Markov function of the current value of experimentation and precisely equals to MS's policy function for the diffusion's precision. The value of experimentation differs in both settings, however.

Next, I consider the binary state and action case to characterize when the experiment is unique and how information via jumps can be optimally used. In most cases, I find that the pure-diffusion experiment described above is the unique, optimal experiment. Nevertheless, when it may be optimal to acquire some Poisson information, I find that $\mathcal{E}$ only acquires information right before he stops experimenting.

The rest of the paper proceeds as follows. Section 1 presents the model. Sections 2,3, 4, 5, and 6 state my results. Lastly, I conclude in section 7.

1 Model

I now present the model. A Bayesian experimenter ($\mathcal{E}$) with Bernoulli preferences picks from a finite set of alternatives A, where $|A| \geq 2$. $\mathcal{E}$'s payoff depends on an unknown state $x \in X \equiv \{x_i\}_{i=1}^n$ for $n = 2, 3, \ldots$ and his initial beliefs are $p \equiv (p_i)_{i=1}^n \in \Delta^{n-1}$ where $\forall i, p_i \equiv \Pr(x = x_i) > 0$. His Payoffs are thus a function $u : A \times X \to \mathrm{R}$ and I assume that $u(\cdot, \cdot) \gg 0$ and for each pair x_i and x_j where $x_i \neq x_j$, it holds that $\operatorname{argmax}_{a \in A} u(a, x_i) \cap \operatorname{argmax}_{a \in A} u(a, x_i) = \emptyset$. This means that never making a choice is strictly suboptimal and the state matters when making an optimal choice. Next, $\mathcal{E}$ is not required to immediately decide. Instead, he can experiment and make a decision at some time $T \in [0, \infty)$. The details of how he experiments are presented below. However, if at time T he holds beliefs $p_T \in \Delta^{n-1}$, his payoff from making a decision are

$$F(p_T) \equiv \max_{a \in A} \sum_{i=1}^{n} p_{iT} u(a, x_i). \qquad \text{(Terminal Payoffs)}$$

1.1 Information Acquisition Problem

Signals I now describe the experimentation problem. $\mathcal{E}$ picks a continuous-time signal process and a signal-adapted stopping time $T < \infty$. Heuristically, the signal is the data generating processed generated by an experiment, while the stopping time is a rule when experimentation stops and a decision is made. An admissible signal $s \equiv (s_t)$ is a jump-diffusion process where $s_0 = 0$ and at each time $t \in [0, \infty)$

$$\mathrm{d}s_t = \mu(x)\,\mathrm{dt} + \frac{\mathrm{d}B_t}{\sqrt{h_t}} + \mathrm{d}N_t. \tag{1}$$

$B \equiv (B_t)$ is a Brownian motion with precision $h_t \gg 0$ and drift $\mu(x)$ where $\mu : X \to (-\infty, \infty)$ is an injective function. Meanwhile, $N \equiv (N_t)$ is an B-independent compensated jump process that jumps by $\Delta_k \in \mathrm{R}$ ($\Delta_1 < \Delta_2 < \ldots < \Delta_K$ and $K = 1, 2, \ldots$ is finite) at a rate of $\lambda_{ikt} \geq 0$ at time t iff $x = x_i$. I assume that $(h_t, ((\lambda_{ikt})_{i=1}^n)_{k=1}^K)$ satisfy the standard the standard Lipschitz condition which ensuring that (s_t) admits a weak solution[2], but no additional restrictions are made on the set of feasible parameters.[3] Lastly, note that if one were to force $\lambda_{ikt} = 0$ for each i, k, t, then the set of experiments is precisely the same as in MS.

Information and Costs I now model an experiment's flow costs. The costs of running an experiment increase in the flow amount on information generated. To do so, I first derive a measurement of the flow amount of information generated as in $Z22$.

Let $H : \Delta^{n-1} \to \mathrm{R} \in C^2$ (e.g., entropy) be a strictly concave function such that its second derivative is bounded below and away from 0 i.e., for each belief $p_t \in \Delta^{n-1}$ and $x \in \mathrm{R}^{n-1}$ such that $x \neq 0$, there exists some $\epsilon > 0$ such that $x' D^2 H(p_t) x < -\epsilon$. Further consider the belief process $(p_t \equiv p_{it}(x = x_i \mid \{s_\tau : \tau \in [0,t]\})) \subset \Delta^{n-1}$. The flow amount of information generated by (s_t) at time t is $I_t \equiv -\mathcal{L}H(p_t)$ where $\mathcal{L}(\cdot)$ is the infinitesimal generator for (p_t) i.e., for each function f, $\mathcal{L}f(p_t) \equiv \lim_{\mathrm{dt}\to 0} \frac{f(p_t) - f(p_{t-\mathrm{dt}})}{\mathrm{dt}}$ if said limit exists. Lastly, the flow cost of experimenting at time t is $c(I_t)$ for some $c(\cdot) \gg 0$ being a strictly increasing, convex, and twice differentiable function.

Payoffs I now describe payoffs. If $\mathcal{E}$ picks signal s—generating beliefs (p_t) and information (I_t)—and an s-adapted stopping time T , then at time t ($\leq T$) expected payoffs are

$$V_t(s,T) \equiv E_t\left[e^{-r(T-t)}F(p_T) - \int_t^T c(I_t)e^{-(\tau-t)}\mathrm{d}\tau \mid \{s_\tau : \tau \in [0,t]\}\right] \qquad \text{(Payoffs)}$$

Given initial beliefs p, $\mathcal{E}$'s optimization problem is

$$V(p) = \max_{s,T} V_0(s,T). \qquad \text{(Unconstrained Problem)}$$

Alternatively, $\mathcal{E}$ may be forced to only acquire pure-diffusion experiments:

[2]See for example Oksendal and Sulem (2019) among others

[3]I choose this model for its parsimony. One could allow the experimenter to observe a multidimensional Jump-diffusion and for jumps to take on a finite number of jumps—for instance. However, the results extend to such setting, but the statements become more complex at no conceptual benefit.

$$U(p) = \max_{s,T} V_0(s,T) \text{ s.t. } \forall i, t, \lambda_{it} = 0 \text{ a.s.} \qquad \text{(Constrained Problem)}$$

I conclude by noting that $U(p) \leq V(p)$, because the set of feasible experiments in the constrained problem is a strict subset of the feasible set in its unconstrained counterpart.

I now present my results. The proof has 3 parts. First, I explicitly derive beliefs. Next, I use standard techniques from stochastic calculus to derive a value function for the restricted and unrestricted problems. I then conduct a sequence of change of variables which makes my main result straightforward.

2 Belief dynamics

I first characterize how a function of Bayes beliefs, derived from (s_t), changes over time.

Lemma 2.1. *Fix $s = (s_t)$. Let (p_t) be the s-adapted Bayes consistent beliefs, then $\forall f : \Delta^{n-1} \to \mathrm{R} \in C^2$*

$$\mathcal{L}f(p_t) = \frac{h_t}{2}\sum_{ij} f_{ij}(p_t)p_{it}p_{jt}(\mu_i - \mu_t)(\mu_j - \mu_t) + \sum_{k=1}^{K} \lambda_{kt}[f(\nu_{kt}) - f(p_t) - \nabla f(p_t)'(\nu_{kt} - p_t)] \tag{2}$$

for $f_{ij}(p_t) \equiv \partial_{p_i}\partial_{p_j} f(p_t)$, $\lambda_{kt} \equiv \sum_i p_{it}\lambda_{ikt}$, $\mu_t \equiv \sum_i p_{it}\mu_i$, and for each $k = 1, 2, \ldots, K$, $\nu_{kt} \equiv (\nu_{ikt} \equiv \frac{p_{it-}\lambda_{ikt-}}{\lambda_{kt-}})$, $\tilde{\mu}_t \equiv (p_{it}(\mu_i - \mu_t))_{i=1}^n$, $p_{t-} = \lim_{dt \to 0} p_{t-dt}$.

I now sketch the proof, but delegate the derivation to appendix A.1. Since the jump and diffusion process are independent, I can approximate each process a different, conditionally independent binomial process. Fix some small dt>0, the diffusion can be approximated by a binomial that jumps up by $\sqrt{\mathrm{dt}/h_t}$ with probability $(1 - \mu_i\sqrt{h_t \mathrm{dt}})/2$ and jumps down by $-\sqrt{\mathrm{dt}/h_t}$ with probability $(1 + \mu_i\sqrt{h_t \mathrm{dt}})/2$ iff $x = x_i$. Meanwhile, the jump process can be approximated by a binomial process that for each jump of size Δ_k with probability λ_{ikt}dt and by 0 with probability $1 - \lambda_{ikt}$dt iff $x_i = x$.

I then derive Bayes posterior beliefs conditional on observing the process $(p_{t\mathrm{dt}})$ and approximate $\mathcal{L}f(p_{t\mathrm{dt}})$. I consider two cases. Case 1, the jump process jumps. Bayes posterior beliefs can disregard the diffusion process and only depend on the prior belief and the state-dependent arrival rate of jumps. The change in beliefs is then $f(p_t) - f(p_{t^-})$.

Case 2, the jump process does not jump. Bayes posterior beliefs then on both processes. I approximate $\mathcal{L}f(p_{t\mathrm{dt}}) \approx \frac{f(p_t)-f(p_{t-\mathrm{dt}})}{\mathrm{dt}}$ using a quadratic Taylor approximation centered at $p_{t-\mathrm{dt}}$. I conclude by noting that this argument is left (intentionally) as standard as possible and that for a jump-less process the generator (i.e., $\mathcal{L}(\cdot)$) simply requires taking $\lambda_{it} = 0$ for each i and t.

3 Re-formulating the Experimentation Problems

Now that the beliefs dynamics are well-defined, I define the cost function, derive the Hamilton-Jacobi-Bellman (HJB) that the experimenter's problem must solve, and then reformulate it in a more useful fashion. First, I characterize the costs function. Fix some signal $s = (s_t)$ process, then at each time t the amount of information generate is $I_t = -\mathcal{L}H(p_t) = \mathrm{I}(\phi_t, p_t)$ where

$$\mathrm{I}(\phi_t, p_t) \equiv \overbrace{\frac{h_t}{2}\sum_{ij} H_{ij}(p_t)p_{it}p_{jt}(\mu_i - \mu_t)(\mu_j - \mu_t)}^{\text{diffusion terms}} + \underbrace{\sum_{k=1}^{K} \lambda_{kt}[H(p_t) - H(\nu_{kt}) + \nabla H(p_t)'(\nu_{kt} - p_t)]}_{\text{Poisson terms}} \quad (3)$$

Now that I derived an expression for the amount of information generated, the flow costs are $c(\mathrm{I}_t)$. The formula above further illustrates several points of note. First, the signal precision enters linearly into the total amount of information and separable from the information derived from jumps. This is a feature of the continuous-time modeling choice and plays a key role in the result below.

Next, I derive an expression for the HJB describing $\mathcal{E}$'s optimal experimentation problem. $\mathcal{E}$'s problem can be written as a function of his beliefs p_t at each time t. By the principle of optimality, if at some belief $p_t \in \Delta^{n-1}, V(p_t) > F(p_t)$, then $\mathcal{E}$ picks

$\phi_t \equiv (h_t, (\lambda_{ikt}))$ to solve

$$rV(p_t) = \max_{\phi_t} \mathcal{L}V(p_t) - c[\mathrm{I}(\phi_t, p_t)] \text{ s.t. } h_t > 0,\ \forall i, \lambda_{it} \geq 0 \tag{4}$$

Oksendal and Sulem (2019) establish that the HJB equation admits a viscosity solution. This is because $\mathcal{E}$'s problem reduces to picking a locally Lipschitz collection of parameter process. They even extend the existence proof to a much broader set of problems than the one studied in this paper. This HJB equation, however, is far too general to allow for a tractable characterization. Instead, I consider a change of variables that clarifies the structure of the value function. Assume that $\mathcal{E}$ picks a Bayes posterior belief conditional that a jump arrives ${\nu_t}$[4] , the amount of information that he acquires from jumps j_t (i.e., his total amount of Poisson information), and from the diffusion β_t (his Gaussian information). More rigorously, define at each time t,

$$\forall k, j_{kt} \equiv \lambda_t[H(p_t) - H(\nu_{kt}) - \nabla H(p_t) \cdot (p_t - \nu_{kt})]$$

and

$$\beta_t \equiv \frac{h_t}{2} \sum_{ij} H_{ij}(p_t) p_{it} p_{jt} (\mu_i - \mu_t)(\mu_j - \mu_t).$$

The total amount of information is then $I_t = \beta_j + \sum_{k=1}^{K} j_{kt}$ and the generator for each function $f : \Delta^{n-1} \to \mathrm{R} \in C^2$ becomes $\mathcal{L}f(p_t) = \beta_t L(f, p_t) + j_t G(f, p_t, \nu_t)$ where

$$L(f, p_t) \equiv \frac{\frac{h_t}{2} \sum_{ij} f_{ij}(p_t) p_{it} p_{jt} (\mu_i - \mu_t)(\mu_j - \mu_t)}{\frac{h_t}{2} \sum_{ij} H_{ij}(p_t) p_{it} p_{jt} (\mu_i - \mu_t)(\mu_j - \mu_t)}$$

and

$$G(f, p_t, \nu_{kt}) \equiv \frac{f(\nu_{kt}) - f(p_t) - \nabla f(p_t)'(\nu_{kt} - p_t)}{H(p_t) - H(\nu_{kt}) + \nabla H(p_t)'(\nu_{kt} - p_t)}.$$

I can now state a more useful reformulation of the value function if $F(p_t) < V(p_t)$, then

[4]In principle he can only pick $n-1$ posterior beliefs as the beliefs regarding the n^{th} state have to be the residual enduring that beliefs are positive and add up to 1.

$$rV(p_t) = \max_{(j_{kt},\nu_{kt}),\beta_t} \beta_t L(V,p_t) + \sum_{k=1}^{K} j_{kt} G(V,p_t,\nu_{kt}) - c\left(\beta_t + \sum_{k=1}^{K} j_{kt}\right) \tag{5}$$

s.t. $(j_{kt}, \nu_{kt}, \beta_t) \in \mathrm{R}_+^{K+Kn+1}, \beta_t > 0, \forall k \sum_i \nu_{ikt} = 1$

The added value of this change of variables is made apparent above when one looks at the cost function. Notice that as the Poisson and Gaussian terms were linearly separable, Poisson and Gaussian information are perfect substitutes $and\ \mathcal{E}\ must$ acquire some Gaussian information. This last point follows from $\mathcal{E}$'s need to manage the precision of his information—a consideration that never becomes salient when one works directly in the belief space.

4 Restricted Problem

My main result is that there exist an optimal, pure-diffusion experiment. In this subsection, I $assume$ that the $\mathcal{E}$ is restricted to run a pure-diffusion experiment. I show that the optimal experiment (within this set) can be characterized by modestly adjusting the proofs in Moscarini and Smith (1998, 2002). I find that the flow amount of information acquired by the $\mathcal{E}$ is a Markov function of the value of experimenting.

First consider the simplest case of interest. Let $x, a \in \{\pm 1\}$, $\mu(x) = x$, initial belief that $x = 1$ be $p \in (0,1)$, and payoffs $\pi(a, x)$. I assume that $\Delta_1 \equiv \pi(1,1) - \pi(-1,1) > 0$, $\Delta_{-1} \equiv \pi(-1,-1) - \pi(1,-1)$, and $\Delta_0 \equiv \pi(-1,-1) - \pi(-1,1) > 0$ i.e., the optimal action is $a = x$. These assumptions imply that there exists a crucial belief $\hat{p} \in (0,1)$ such that when $p_t \equiv \Pr_t(x = 1) = \hat{p}$, the experimenter is indifferent between both choices (i.e., $\pi(-1,1) + \hat{p}\Delta_1 = \pi(-1,-1) - \hat{p}\Delta_{-1}$) and it is optimal to pick $a = 1$ iff $p_t \geq \hat{p}$. It also implies that the payoff from making a decision is

$$F(p_t) \equiv \max_{a\in\{\pm 1\}} p_t \pi(a,1) + (1-p_t)\pi(a,-1) = \begin{cases} \pi(-1,-1) - \Delta_{-1} p_t & \text{if } p_t \leq \hat{p} \\ \pi(-1,1) + \Delta_1 p_t & \text{if } p_t \geq \hat{p} \end{cases} \tag{6}$$

I now characterize the optimal experiment. First, denote the $\mathcal{E}$'s time t Bayes posterior belief that $x = 1$ given that he observes a path $\{s_\tau \mid \tau \in [0, t]\}$ as p_t. Given that the prior belief is p, and the signal process satisfies $s_0 = 0$ and at each $t \geq 0$, $\mathrm{d}s_t = x\mathrm{d}t + \mathrm{d}B_t/\sqrt{h_t}$, then Lipster and Shiryaev's (1977) theorem 9.1[5] establishes that $p_0 = p$- and at each time $t \geq 0$ beliefs evolve following

$$\mathrm{d}p_t = p_t(1 - p_t)\sqrt{h_t}\mathrm{d}\bar{B}_t \tag{7}$$

where $(\bar{B}_t)$ is a Brownian motion such that $\bar{B}_0 = 0$ and $\mathrm{d}\bar{B}_t = \mathrm{d}s_t - p_t\mathrm{d}t$. Next, notice that the flow amount of information (i.e., β_t) generated when the current belief is p_t and signal precision is h_t is $\beta_t = (h_t/2)[p_t(1 - p_t)]^2[-H''(p_t)]$. Hence, one can reformulate belief dynamics in terms of β_t as

$$\mathrm{d}p_t = \sqrt{2\beta_t/[-H''(p_t)]}\mathrm{d}\bar{B}_t. \tag{8}$$

Notice that the assumption that the second derivative of $H(\cdot)$ was bounded below a constant $-\epsilon$ for $\epsilon > 0$ implies that $\sqrt{2\beta_t/[-H''(p_t)]} < \sqrt{2\beta_t/\epsilon}$, so if the control (β_t) satisfies the standard Lipschitz conditions, then the processes (p_t) and (s_t) will have unique, weak solutions.

Next, following the appendix B proof in Moscarini and Smith (1998), the current belief (i.e., p_t) is a sufficient statistic and restricting attention to Markov controls that are a function of said beliefs is also without loss of generality. This is because there is a one-to-one correspondence between the elements of the natural filtration generated by (s_t), for each fixed control process (h_t), and belief process (p_t). Thus, one can work with the belief process (p_t) rather than the signal process (s_t) and the change of variable plays no role in this exercise. In addition, the conditions on $c(\cdot)$ and $H(\cdot)$ ensure that the rest of the argument in appendix B still follow. This implies that the value of experimentation and the optimal control can be written as a function of the current belief p_t as $U(p_t)$ and $\beta(p_t)$, respectively. Said control further solves

[5]The authors further establish the dynamics for beliefs when there are finitely many $n \geq 2$ states and the signal is a pure-diffusion. Said dynamics are associated with precisely the same generate as in t2.1 when $\lambda_{xjt} = 0$ for each x, j, t.

$$rU(p_t) = \max_{\beta_t > 0} \beta_t L(U, p_t) - c(\beta_t) \tag{9}$$

where $L(U, p_t) \equiv U''(p_t)/[-H''(p_t)]$. The first order condition (FOC) further implies that $c'(\beta_t) = L(U, p_t)$. By the principle of optimality, it consequently holds that for each current belief p_t such that $U(p_t) > F(p_t)$, then $rU(p_t) = \beta_t c'(\beta_t) - c(\beta_t)$ or that $\beta_t = \beta(u) = f^{-1}[rU(p_t)]$ for $f(x) \equiv xc'(x) - c(x)$.

Note that $\beta(\cdot)$ precisely equals the control for information precision in MS (i.e., $h(\cdot)$), but that the value of experimentation in their setting is $V(\cdot)$ and $V(\cdot) \neq U(\cdot)$, in general. Further note that proposition 7 in MS can be re-stated for the current setting be replacing $\beta(\cdot)$ for $h(\cdot)$, so the insights found from the special case studied above extends to a setting with finitely many states and actions i.e., $n, \#A \geq 2$.

I conclude this section by illustrating how $U(\cdot)$ is derived for a particular example. Suppose that $c(x) = x^2/2$ for each $x \geq 0$, $r = 0.01$, $\Delta_1 = \Delta_{-1} = \Delta_0 = 1$. This means that the experimenter wants to simply match the state and $\hat{p} = 1/2$. Plugging the FOC into the value function derived from the principle of optimality implies that

$$\text{if } U(p_t) > F(p_t), \text{then } rV(p_t) = \frac{\{[L(U, p_t)]\}^2}{2} \tag{Current Model}$$

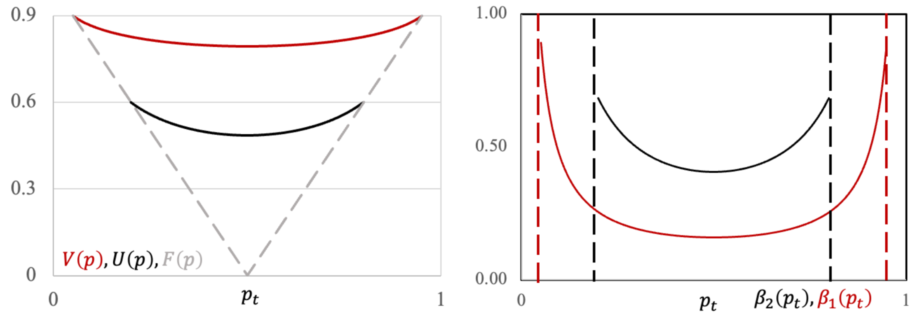
Next, it remains the case that there exist a pair of cutoff beliefs $0 < \tilde{p} < \hat{p} < \bar{p} < 1$ such that $\mathcal{E}$ experiments iff $p_t \in (\tilde{p}, \bar{p})$ and the boundary conditions are that

$$\forall p \in \{\tilde{p}, \bar{p}\}, U(p) = F(p) \tag{Value Matching}$$

and

$$\forall p \in \{\tilde{p}, \bar{p}\}, U'(p) = F'(p). \tag{Smooth pasting}$$

A similar set of standard boundary conditions are also imposed in MS. Figure 1's left panel illustrates the terminal payoff from making a decision ($F(p)$, in grey), the value of experimenting in MS when $c(h_t) = h_t/2$ ($V(p)$, in red), and $U(p)$ in black. On the other hand, the right-hand panel plots the flow amount of information acquired as a function of p_t in both settings. It shows that, relative to MS, the experimentation interval is shifted inward, but $\mathcal{E}$ always acquires more information. This is analogous to increasing the discount factor r in MS.



(a) ***Value functions*** *(b)* ***Policy functions***

Figure 1: ***Optimal policy and value function comparison between the*** MS ***model (denoted as case 1) and my own (case 2).***

5 Main Result

This section presents my main result. I will first state the result and then provide the proof.

Theorem 5.1. $\forall p_t \in \Delta^{n-1}, \;\; U(p_t) = V(p_t)$.

This theorem states that restricting attention to only pure-diffusion experiments is without loss of generality. The proof goes as follows. First, it is immediate that for each $p_t \in \Delta^{n-1}$, it must be that $U(p_t) \leq V(p_t)$. This is because the previous lemma established that there exists a control (β_t) that maximizes U. Define a new control for the general problem as $\tilde{\phi} \equiv \{\tilde{\beta}_t, (\tilde{j}_{kt}, \tilde{\nu}_{kt})\}$ such that $\tilde{\beta}_t = \beta_t$ and for each $k = 1, 2, \ldots, K, \tilde{j}_{kt} = 0$ and $\nu_{kt} = p_t$ almost surely. Control $\tilde{\phi}$ is feasible in the general problem and attains a payoff of $V(p_t, \tilde{\phi}) = U(p_t)$ for each p_t. However, said control need not be optimal, so $U(p_t) = V(p_t, \tilde{\phi}) \leq V(p_t)$ for each p_t.

The opposite inequality—i.e., $U(p_t) \geq V(p_t)$—is more subtle. First, suppose that at some belief $p_t, V(p_t) = F(p_t)$. Since at each belief p_t, $\mathcal{E}$ can always stop experimenting and make a decision, then $F(p_t) \leq U(p_t)$. As a consequence, it holds that $V(p_t) = U(p_t)$ and it suffices to show that the result holds for $p_t \in C$ where $C \equiv \{p_t \in \Delta^{n-1} : V(p_t) > F(p_t)\}$.

I now show that if an optimal experiment acquires information via jumps, it cannot yield the $\mathcal{E}$ more payoff than the experiment characterized by control $\tilde{\phi}$. Suppose that a control $\phi' = \{\beta'_t, (j'_{kt}, \nu'_{kt})\}$ attains the maximum of the HJB equation 5. At each belief $p_t \in C$, the control must satisfy two conditions. The first condition is that the control must satisfy the interior first order condition i.e.,

$$\forall t, p_t \in C, c'(I_t) = L(V, p_t)$$

for $I'_t \equiv \beta_t + \sum_{k=1}^{K} j_{kt}$. Likewise, if for some $j_{kt} > 0$ at belief p_t, then it must be the case that $c'(I'_t) = G(V, p_t, \nu_{kt})$. The second condition is that, by the principle of optimality, for each $p_t \in C$, $V(p_t)$ satisfies

$$rV(p_t) = \beta'_t L(V, p_t) + \sum_{k=1}^{K} j'_k(p_t) G(V, p_t, \nu'_{kt}) - c(I'_t) = I'_t L(V, p_t) - c(I'_t).$$

Notice that the second equality follows from the observation that the marginal benefits both types of information must be equalized whenever $\mathcal{E}$ pick a strictly positive amount of both types of information. If, on the other hand, $j_{kt} = 0$ at the optimum, then $j'_k(p_t)G(V, p_t, \nu'_{kt}) = 0$, so it does not affect the validity of the second equality. Moreover, said equality clarifies that what matters is the total amount of information acquired and the composition of information is a second-order consideration.

Define $\bar{\phi} \equiv \{\bar{\beta}_t, \bar{j}_{kt}, \bar{\nu}_{kt}\}$ such that at each belief $p_t, \bar{\beta}_t = I'_t$ and for each $k = 1, 2, \ldots, K$, $\bar{j}_{kt} = 0$ and $\bar{\nu}_{kt} = p_t$ almost surely. Then, by construction, $\bar{\phi}$ satisfies the first order conditions for the optimization problem and for each $p_t \in C$, it holds that $\bar{\beta}_t L(V, p_t) - c(\bar{\beta}_t) = I'_t L(V, p_t) - c(I'_t) = rV(p_t)$. Consequently, if $V(p_t; \bar{\phi}) = V(p_t)$. Moreover, the control $\bar{\beta} = (\bar{\beta}_t)$ and note that it is admissible in the restricted problem, so its cannot be strictly greater than $U(p_t)$ at each belief p_t. Consequently, $V(p_t) \leq U(p_t)$ for each p_t as desired.

6 $n = \#A = 2$

I now return to the binary state and action case presented in subsection 4 and show that (in such case) the pure-diffusion experiment is the only possible experiment. Intuitively, if an experiment is optimal and for some current belief p_t generates information via jumps, $then$ the posterior belief conditional on a jump arriving (denoted as ν_t) must satisfy a first order condition (FOC) and not equal to its prior. I show that both conditions are incompatible.

I first state a useful corollary.

Corllary 6.1. *The function $U(\cdot)$ is strictly convex.*

The proof for this corollary is a simple application of the FOC that β_t has to satisfy and the principle of optimality. Next, I provide a condition needed for the result.

Assumptions 6.2. *Assume that* $\lim_{p \nearrow 1} H'(p) = -\infty$, $\lim_{p \searrow 0} H'(p) = \infty$.

I now state the lemma.

Lemma 6.3. *If assumption 6.2 holds, then for every current belief $p_t \in (\tilde{p}, \bar{p})$ such that there exist an experiment for which the $\mathcal{E}$ acquires information via Poisson jumps, the belief posterior of a jump $\nu_t \in [0, \tilde{p}] \cup [\bar{p}, 1]$ satisfies*

$$\frac{F'(\nu_t) - U'(p_t)}{U''(p_t)} = \frac{H'(\nu_t) - H'(p_t)}{H''(p_t)}.$$

I now present the proof. Observe that the function $H(\cdot)$ was exogenously defined in the solution of the problem and the value of experimenting U is the same function characterized in section 4. Consequently, for each belief p_t such that $U(p_t) > V(p_t)$ (i.e., the experimenter benefits from experimenting), the posterior belief $\nu_t \in [0, 1]$ simply maximizes $\ln G(U, p_t, \nu_t)$ i.e.,

$$\max_{\nu_t \in [0,1]} \ln[U(\nu_t) - U(p_t) - U'(p)(\nu_t - p_t)] - \ln[H(p_t) - H(\nu_t) + H'(p_t)(\nu_t - p_t)]. \quad (10)$$

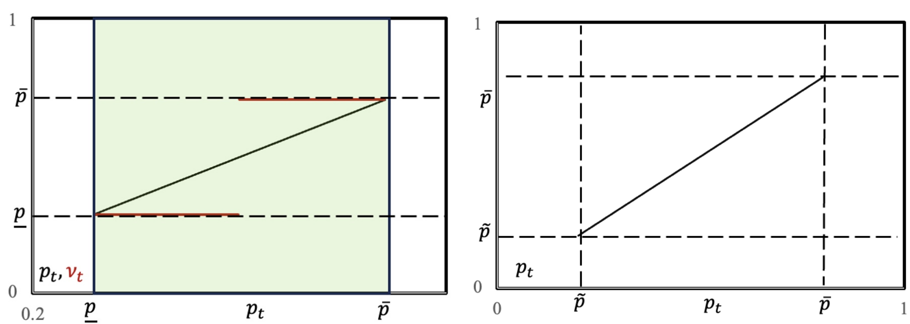
Noe that since $U(\cdot)$ is strictly convex and $H(\cdot)$ is strictly concave, then he expressions inside the logs (i.e., $U(\nu_t) - U(p_t) - U'(p_t)(\nu_t - p_t)$ and $H(p_t) - H(\nu_t) + H'(p_t)(\nu_t - p_t)$) are strictly positive. In addition, assumption 6.2 implies that the FOC is interior, so $\nu_t \in (0, 1)$ and said condition is

$$[U'(\nu_t) - U'(p_t)] = [H'(p_t) - H'(\nu_t)]G(U, p_t, \nu_t) \tag{11}$$

If information is acquired via jumps, then the FOCs found in the main result imply that $G(U, p_t, \nu_t) = U''(p_t)/[-H''(p_t)]$. Moreover, as $Z22$ points out, in any optimal experiment, the posterior belief after a jump arrive must prompt the experimenter to stop experimenting and make a decision. Combining these observations with equation 11 imply that $\nu_t \in [0, \tilde{p}] \cup [\bar{p}, 1]$ and

$$\frac{U'(\nu_t) - U'(p_t)}{U''(p_t)} = \frac{H'(\nu_t) - H'(p_t)}{H''(p_t)}. \tag{12}$$

This conclues the proof.



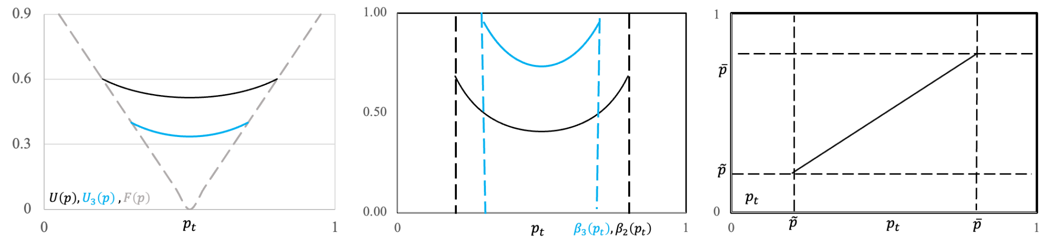
(a) ***Zhong (2022).*** *(b)* ***Current Paper***

Figure 2: ***Current belief zones where acquiring information via jumps is feasible (green), posterior beliefs conditional on a jump arriving (red), and prior belief (black).***

I now return to the example presented in section 4. Figure 2 draws the subset of experimentation beliefs for which $\mathcal{E}$ might acquire information via jumps (in green), the current belief (in black), and the posterior belief (in red). In Zhong (2022), $\mathcal{E}$ acquires information via jumps iff at a given current belief he wishes to experiment. In contrast, there does not exist any current belief for which $\mathcal{E}$ wishes to acquire information via jumps in my setting. The next subsection, however, does find a case when acquiring information via jumps might be optimal for a strict subset of the experimentation interval.

6.1 Comparative Statics

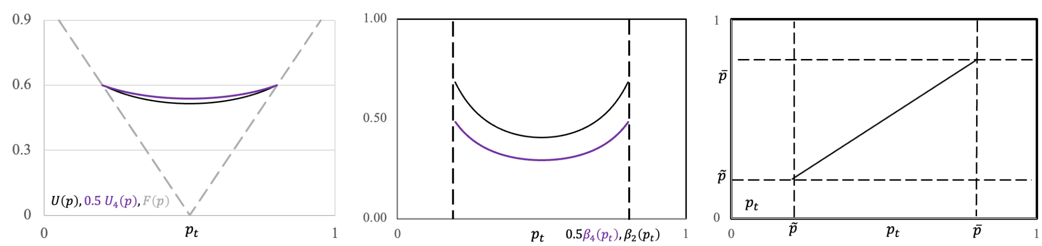
Increasing r First, I increase r from 0.001 to 0.1 and plot the outcome in figure 3. Making the experimenter more impatient lowers $U(\cdot)$, but does not affect the payoff from making a decision. As a consequence, the interval of beliefs shifts inwards, but (as in MS) the policy function remains u-shaped and shifts up. Lastly, the optimal, pure-diffusion experiment remains the unique, optimal experiment.



(a) ***Change in*** $U(\cdot)$. *(b)* ***Change in*** $\beta(\cdot)$. *(c)* ***Change in*** $\nu(\cdot)$.

Figure 3: ***Change in optimal experiments when*** r ***increases.***

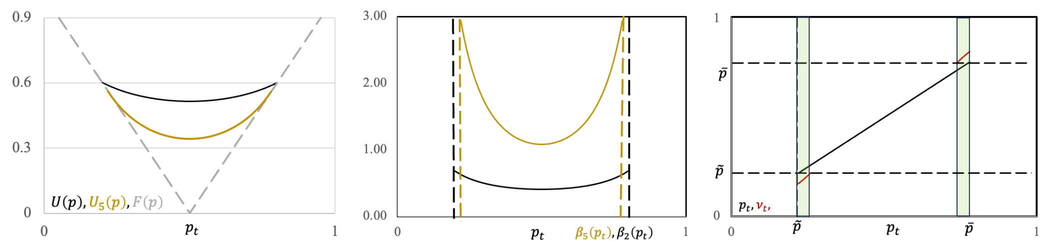
Increasing $F(\cdot)$ Next, I replace $F(\cdot)$ by $2F(\cdot)$—figure 4. The pure-diffusion experiment remains the unique optimal experiment. Meanwhile, the left-hand panel shows that the experimenter's payoff from experimenting is more than doubled relative to the benchmark in figure 2 but the experimentation interval is approximately the same. Lastly, the central sub-figure shows that the $\mathcal{E}$ acquires more flow information than in the benchmark case, but he acquires less than twice the amount of information.



(a) ***Change in*** $U(\cdot)$. *(b)* ***Change in*** $\beta(\cdot)$. *(c)* ***Change in*** $\nu(\cdot)$.

Figure 4: ***Change in optimal experiments when*** $F(\cdot)$ ***shifts up to*** $2F(\cdot)$.

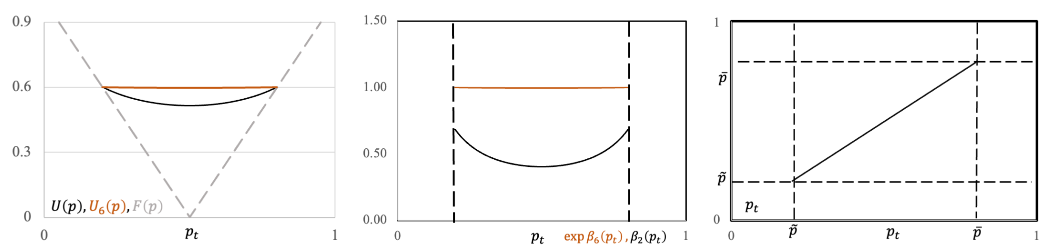
Decreasing $c(\cdot)$ Next, I change $c(I_t)$ to $c(I_t)/2$—figure 5. I first find that the payoff from experimentation shifts down uniformly and the experimentation interval shrinks inward, because (as shown in the central plot) the $\mathcal{E}$ always acquires more flow information than in the benchmark. I lastly find that the experimenter *can* benefit from acquiring Poisson information. However, he only acquires said type of information when his beliefs are near the boundaries of the experimentation region.



(a) ***Change in*** $U(\cdot)$. *(b)* ***Change in*** $\beta(\cdot)$. *(c)* ***Change in*** $\nu(\cdot)$.

Figure 5: ***Change in optimal experiments when*** $c(\cdot)$ ***shifts down uniformly to*** $c(\cdot)/2$.

Time-averaged experimentation costs The last comparative static pertains changing the cost function from $c(I_t)$ to $rc(I_t)$—figure 6. I first find that the optimal experiment characterized in section 4 is unique i.e., Poisson information is suboptimal. Next, I find that (relative to the benchmark, the value of experimenting is flat and thus the flow amount of information acquired is exponentially less than in the benchmark case.



(a) ***Change in*** $U(\cdot)$. *(b)* ***Change in*** $\beta(\cdot)$. *(c)* ***Change in*** $\nu(\cdot)$.

Figure 6: ***Change in optimal experiments when*** $c(\cdot)$ ***shifts down uniformly to*** $rc(\cdot)$.

7 Conclusion

This paper shows that the problems of sequential experimentation and information acquisition are different. This is because the dynamics behind data acquisition matter. To experiment is to carefully manage a data generation process that (by its very nature) produces noisy information. As a consequence, an experimenter not only determines what kind of information he wishes to acquire, but he also must process or clean said data. This added tasks qualitatively changes the tradeoffs faces relative to a decision-maker who is tasked to acquire pre-existing information.

This result further implies that (in many economic contexts) disregarding the process of information generation results in conclusions that could never be applied in real-world contexts. It is, therefore, important for subsequent work in dynamic experimentation, information design, Bayesian persuasion, et cetera to ensure that the information acquired in the model could be conceivably acquired in real-world applications. Otherwise, the rich set of new results in these burgeoning fields cannot be used in many real-world application or inform policy.

The paper lastly points out that the particular functional form of experimentation costs matter. At the optimum, the level of experimentation; the type of information to acquire; and when is experimenting worthwhile are all sensible to the choice of cost function. This implies that future, empirical research into experimentation should provide proper estimates of said costs.

References

[1] Bolton, Patrick, and Christopher Harris. 1999. "Strategic signal". Econometrica (ECTA) 67 (2), 349–374.

[2] Caplin, Andrew, Mark Dean, and John Leahy. 2022. "Rationally Inattentive Behavior: Characterizing and Generalizing Shannon Entropy." Journal of Political Economy 130 (6): 1676–1715.

[3] Bloedel, Alexander W., and Weijie Zhong. 2021. "The Cost of Optimally Acquired Information." Unpublished.

[4] Denti, Tommaso, Fabrizio Marinacci, and Aldo Rustichini. 2022. "Experimentation Cost of Information". American Economic Review (AER) (2022) 12(9): 3106–3123.

[5] Denti, Tommaso. Forthcoming. "Posterior Separable Cost of Information." AER.

[6] Denti, Tommaso, Massimo Marinacci, and Aldo Rustichini. 2022. "The experimental Order on Random Posteriors." Unpublished.

[7] Hébert, Benjamin, and Michael Woodford. 2021. "Rational Inattention When Decisions Take Time." Unpublished.

[8] JM Harrison. 2013. "Brownian Models of Performance and Control". Cambridge University Press.

[9] Mensch, Jeffrey. 2018. "Cardinal Representations of Information." Unpublished.

[10] Moscarini, Giuseppe, and Lones Smith. 2001. "The Optimal Level of Experimentation". Econometrica 69(6): 1629-1644.

[11] ————-. 1998. "Wald Revisited: The Optimal Level of Experimentation". Cowles Foundation Discussion Papers.

[12] Morris, Stephen, and Philipp Strack. 2019. "The Wald Problem and the Equivalence of Sequential Sampling and Static Information Costs." SSRN 2991567.

[13] Øksendal , Bernt and Agnès Sulem. 2019. ”Applied Stochastic Control of Jump Diffusions”. Springer-Verlag Switzerland.

[14] Platten, Eckhart (1982). ”An approximation method for a class of Ito processes with jump component”. Lient. Mat Rink 22(2): 124-136.

[15] Pomatto, Luciano, Philipp Strack, and Omer Tamuz. 2020. “The Cost of Information.” Unpublished.

[16] Mike Press. 2016. ”Cleaning Big Data: Most Time-Consuming, Least Enjoyable Data Science Task, Survey Says”. Forbes retrieved from: $https://www.forbes.com/sites/gilpress/2016/03/23/data-preparation-most-time-consuming-least-enjoyable-data-science-task-survey-says/amp/$

[17] Sims, Christopher A. 2003. “Implications of Rational Inattention.” Journal of Monetary Economics 50(3): 665–90.

[18] Strulovici, Bruno and Martin Szydlowski. 2015. ”On the smoothness of value functions and the existence of optimal strategies in diffusion models”. Journal of Economic Theory 159:1016–1055.

[19] Wald, A. 1947. ”Sequential analysis”. John Wiley.

[20] Zhong, Weijie. 2022 ”Optimal Dynamic Information Acquisition”. ECTA 90(4):1537-1582.

A Proofs

A.1 Deriving dynamics.

In this section, I present the characterization of beliefs. Informally, I approximate the signal process in discrete time and take limits.

Approximating signals in discrete-time Fix some small time interval dt>0, then an admissible signal $s = (s_t)$ can be approximated at times $t = 0, \text{dt}, \dots$ as $s_0 = 0$ and $\text{d}s_t \equiv s_{t+\text{dt}} - s_t$

$$\text{d}s_t = \text{d}_t^{\text{dt}} + \sum_{k=1}^{K} \text{d}J_{kt}^{\text{dt}} \tag{13}$$

for $(\text{d}_t^{\text{dt}})_{t=0}^{\infty}$ is a sequence of independent random variables such that at time t, $d_t^{\text{dt}} = \pm\sqrt{\text{dt}/h_{t-\text{dt}}}$ with probability $[1 \pm \mu_i \sqrt{h_{t-\text{dt}}\text{dt}}]/2$ iff $x = x_i$. Meanwhile, for each $k = 1, 2, \dots, K$, $(J_{kt}^{\text{dt}})_{t=0}^{\infty}$ is a sequence of independent random variables such that $\text{d}J_{kt}^{\text{dt}} = \Delta_k$ with probability $\lambda_{ikt-\text{dt}}\text{dt}$, $\text{d}J_{kt}^{\text{dt}} = 0$ otherwise, and the $\text{d}J_{kt}^{\text{dt}}$ and $\text{d}J_{k't}^{\text{dt}}$ are independent for each t, k, k' . This means that the independent diffusion and jump processes are weakly approximated. Also, the experimenter picks the parameters for realizations observed at time $t = 0\text{dt}, 2\text{dt}, \dots$ at time $t-\text{dt}$.

Approximating beliefs after a jump I first consider the case when there are jumps. Suppose that $\mathcal{E}$ held beliefs $p_{t-\text{dt}} = (p_{it-\text{dt}})$, then the Bayes posterior belief that $x = x_i$ given the jump is approximately equal to

$$\nu_{ikt} = \frac{p_{it-\text{dt}}\text{dt}\lambda_{ikt-\text{dt}}}{\sum_{j=1}^{n} p_{jt-\text{dt}}\text{dt}\lambda_{jkt-\text{dt}}} + o(\text{dt}) = \frac{p_{it-\text{dt}}\lambda_{ikt-\text{dt}}}{\sum_{j=1}^{n} p_{jkt-\text{dt}}\lambda_{jkt-\text{dt}}} + o(\text{dt})$$

where the error term $o(\text{dt})$ (such that $\lim_{\text{dt}\searrow 0} o(\text{dt})/\text{dt} = 0$) follows from the observation that distribution of d_t^{dt} approximately gives equal weight to both outcomes as dt goes to 0. Further observe that as dt goes to 0, it holds that $\nu_{ikt} = \frac{p_{it-}\lambda_{it-}}{\lambda_{kt-}}$ where $\lambda_{kt-} \equiv \sum_{j=1}^{n} p_{jt-}\lambda_{jkt-}$. It is further the case that the discontinuous change in beliefs is $\text{d}p_{ikt} \equiv \nu_{ikt} - p_{it-\text{dt}}$ and it converges to $\text{d}p_{ikt} \equiv \nu_{ikt} - p_{it-}$ as dt$\searrow 0$.

Approximating beliefs when there is no jump Next, I characterize how beliefs change when $J_t^{\mathrm{dt}} = 0$. Suppose that $\mathcal{E}$ observes $\mathrm{d}s_t = \pm\sqrt{\mathrm{dt}/h_{t-\mathrm{dt}}}$, then the probability of observing said signal realization conditional on $x = x_i$, for $i = 1, 2 \ldots, n$, is

$$\Pr_t(\mathrm{d}s_t = \pm\sqrt{\mathrm{dt}/h_{t-\mathrm{dt}}}|x = x_i) = \frac{1}{2}\left[1 \pm \mu_i\sqrt{h_{t-\mathrm{dt}}\mathrm{dt}} - \sum_{k=1}^{K}\lambda_{ikt-\mathrm{dt}}\mathrm{dt}\right] + o(\mathrm{dt})$$

Once again, if the prior belief is $p_{t-\mathrm{dt}} = (p_{it-\mathrm{dt}})$, then the Bayes posterior beliefs are

$$p_{it} = \frac{p_{it-\mathrm{dt}}\big[1 \pm \mu_i\sqrt{h_{t-\mathrm{dt}}\mathrm{dt}} - \mathrm{dt}\sum_{k=1}^{K}\lambda_{ikt-\mathrm{dt}}\big]}{\sum_{j=1}^{n} p_{jt-\mathrm{dt}}\big[1 \pm \mu_j\sqrt{h_{t-\mathrm{dt}}\mathrm{dt}} - \mathrm{dt}\sum_{k=1}^{K}\lambda_{jkt-\mathrm{dt}}\big]} + o(\mathrm{dt})$$
$$= \frac{p_{it-\mathrm{dt}}\big[1 \pm \mu_i\sqrt{h_{t-\mathrm{dt}}\mathrm{dt}} - \mathrm{dt}\sum_{k=1}^{K}\lambda_{ikt-\mathrm{dt}}\big]}{1 \pm \mu_{t-\mathrm{dt}}\sqrt{h_{t-\mathrm{dt}}\mathrm{dt}} - \mathrm{dt}\sum_{k=1}^{K}\lambda_{kt-\mathrm{dt}}} + o(\mathrm{dt})$$

where $\mu_t \equiv \sum_{i=1}^{n} p_{it}\mu_i$ and for each $k = 1, 2, \ldots, K$, $\lambda_{kt} \equiv \sum_{i=1}^{K}\lambda_{ikt}$. The change in beliefs $dp_{it} \equiv p_{it} - p_{it-\mathrm{dt}}$ can be the approximated as

$$dp_{it} = \frac{p_{it-\mathrm{dt}}\big[\pm(\mu_i - \mu_{t-\mathrm{dt}})\sqrt{h_{t-\mathrm{dt}}\mathrm{dt}} - \mathrm{dt}\sum_{k=1}^{K}(\lambda_{ikt-\mathrm{dt}} - \lambda_{kt-\mathrm{dt}})\big]}{1 \pm \mu_{t-\mathrm{dt}}\sqrt{h_{t-\mathrm{dt}}} - \mathrm{dt}\sum_{k=1}^{K}\lambda_{kt-\mathrm{dt}}} + o(\mathrm{dt})$$

and the probability that beliefs change by the amount described above occurs with a probability of approximately $(1 \pm \mu_{t-\mathrm{dt}}\sqrt{h_{t-\mathrm{dt}}\mathrm{dt}} - \mathrm{dt}\sum_{k=1}^{K}\lambda_{kt-\mathrm{dt}})/2$.

Approximating the expected change in beliefs Given the approximations for the change in beliefs given above, I now take expectations. First, conditional on there being no jump, the expected change in beliefs equals to

$$\frac{E[dp_{it}|\forall k, J_{kt}^{\text{dt}}=0]}{\text{dt}} = \frac{p_{it-\text{dt}}}{2\text{dt}}\left[(\mu_i - \mu_{t-\text{dt}})\sqrt{h_{t-\text{dt}}\text{dt}} - \text{dt}\sum_{k=1}^{K}(\lambda_{ikt-\text{dt}} - \lambda_{kt-\text{dt}})\right]$$
$$+ \frac{p_{it-\text{dt}}}{2\text{dt}}\left[-(\mu_i - \mu_{t-\text{dt}})\sqrt{h_{t-\text{dt}}\text{dt}} - \text{dt}\sum_{k=1}^{K}(\lambda_{ikt-\text{dt}} - \lambda_{kt-\text{dt}})\right] + \frac{o(\text{dt})}{\text{dt}}$$
$$= -p_{it-\text{dt}}\sum_{k=1}^{K}(\lambda_{ikt-\text{dt}} - \lambda_{kt-\text{dt}}) + \frac{o(\text{dt})}{\text{dt}}.$$

On the other hand, if jump Δ_k arrives, for some $k = 1, 2, \ldots, K$, then for each $i = 1, 2, \ldots, n$

$$E[dp_{it}|J_{kt}^{\text{dt}}=1] = \nu_{ikt} - p_{it-\text{dt}}.$$

As a consequence, the unconditional change in beliefs just equals to the expectation over the conditional expectations i.e.,

$$\frac{E[dp_{it}]}{\text{dt}} = \sum_{k=1}^{K}\frac{\text{dt}\lambda_{kt-\text{dt}}}{\text{dt}}E[dp_{it}|J_{kt}^{\text{dt}}=1] + \left(1 - \text{dt}\sum_{k=1}^{K}\lambda_{kt-\text{dt}}\right)\frac{E[dp_{it}|\forall k, J_{kt}^{\text{dt}}=0]}{\text{dt}}$$
$$= \sum_{k=1}^{K}(\nu_{ikt} - p_{it-\text{dt}})\lambda_{kt-\text{dt}} - \left(1 - \text{dt}\sum_{k=1}^{K}\lambda_{kt-\text{dt}}\right)\sum_{k=1}^{K}(\lambda_{ikt-\text{dt}}p_{it-\text{dt}} - \lambda_{kt-\text{dt}}p_{it-\text{dt}}) + \frac{o(\text{dt})}{\text{dt}}.$$
$$= \sum_{k=1}^{K}(\nu_{ikt} - p_{it-\text{dt}})\lambda_{kt-\text{dt}} - \left(1 - \text{dt}\sum_{k=1}^{K}\lambda_{kt-\text{dt}}\right)\sum_{k=1}^{K}(\nu_{ikt-\text{dt}} - p_{it-\text{dt}})\lambda_{kt-\text{dt}} + \frac{o(\text{dt})}{\text{dt}}$$
$$= -\text{dt}\sum_{k=1}^{K}\lambda_{kt-\text{dt}}\sum_{k=1}^{K}(\nu_{ikt-\text{dt}} - p_{it-\text{dt}})\lambda_{kt-\text{dt}} + \frac{o(\text{dt})}{\text{dt}}.$$

The second first line is just the law of iterated expectations and the second line replaced the conditional expectations derived above and replaces them for the expressions derived above. The third line exploits the observation that $\nu_{ikt-\text{dt}} \equiv \lambda_{ikt-\text{dt}}p_{it-\text{dt}}/\lambda_{kt-\text{dt}}$ by Bayes rule. The last line simplifies the expression and yields that as dt goes to 0, the expected change in beliefs converges to 0. Ergo, beliefs are a martingale as expected.

Approximating the co-movement of beliefs The next step is to calculate the co-movement of beliefs conditional on there being no jumps. This step is important, because the main result of this proof is to approximate the change of a twice continuously differentiable function over an instant i..e., the generator for the belief process induced from observing a given, admissible signal process. Pick some pair of state realizations x_i and x_j, then

$$\mathrm{d}p_{it}\mathrm{d}p_{jt} = \left[\frac{h_t p_{it-\mathrm{dt}} p_{jt-\mathrm{dt}}(\mu_i - \mu_t)(\mu_j - \mu_t)}{1 \pm \mu_t^2 h_t \mathrm{dt}}\right]\mathrm{dt} + o(\mathrm{dt})$$

occurs with a probability of roughly $(1 \pm \mu_t^2 h_t \mathrm{dt})/2$. The equation above holds, because the additional terms get multiplied by either a term $\epsilon_1(\mathrm{dt}) = \mathrm{dt}^{\frac{3}{2}}$ or $\epsilon_2(\mathrm{dt}) = \mathrm{dt}^2$. There terms are of the magnitude $o(\mathrm{dt})$ (i.e., $\lim_{\mathrm{dt}\searrow 0} \frac{\epsilon_j(\mathrm{dt})}{\mathrm{dt}} = 0$ for each $j = 1, 2$), so they will not add anything to the final approximation. Taking expectations of the expected co-movement then reveals that

$$\frac{E[\mathrm{d}p_{it}\mathrm{d}p_{jt} | J_t^{\mathrm{dt}} = 0]}{\mathrm{dt}} = h_t p_{it-\mathrm{dt}} p_{jt-\mathrm{dt}}(\mu_i - \mu_t)(\mu_j - \mu_t) + \frac{o(\mathrm{dt})}{\mathrm{dt}}.$$

Approximating the generator Lastly, I approximate the generator. Let $f : \Delta^{n-1} \to \mathrm{R}$ be a twice continuously differentiable function, then I need to calculate $\mathcal{L}f(p_t) \equiv E[df(p_t)]/\mathrm{dt}$ as dt goes to 0 for $df(p_t) \equiv f(p_t) - f(p_{t-\mathrm{dt}})$. I can partition the expectation by the law of iterated expectations. If there is a jump of magnitude Δ_k for some $k = 1, 2, \ldots, K$, then

$$E[df(p_t) | \mathrm{d}J_{kt}^{\mathrm{dt}} = \Delta_k] = f(\nu_{kt}) - f(p_{t-\mathrm{dt}}).$$

Alternatively, there may have been no jumps, then the change in beliefs can be approximated via a quadratic Taylor approximation as

$$\begin{aligned}
\frac{E[df(p_t)|\forall k, J_{kt}^{\text{dt}}=0]}{\text{dt}} &= \frac{1}{\text{dt}}E\left[\nabla f(p_t)'\text{d}p_t + \frac{1}{2}\text{d}p_{it}Hf(p_t)'\text{d}p_t|\forall k, J_{kt}^{\text{dt}}=0\right] + \frac{o(\text{dt})}{\text{dt}} \\
&= \frac{1}{\text{dt}}E\left[\sum_{i=1}^{n} f_i(p_t)\text{d}p_{it} + \frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n} f_{ij}(p_t)\text{d}p_{it}\text{d}p_{jt}|\forall k, J_{kt}^{\text{dt}}=0\right] + \frac{o(\text{dt})}{\text{dt}} \\
&= \sum_{i=1}^{n} f_i(p_t)\frac{E\big[\text{d}p_{it}|\forall k, J_{kt}^{\text{dt}}=0\big]}{\text{dt}} + \frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n} f_{ij}(p_t)\frac{E\big[\text{d}p_{it}\text{d}p_{jt}|\forall k, J_{kt}^{\text{dt}}=0\big]}{\text{dt}} + \frac{o(\text{dt})}{\text{dt}} \\
&= -\sum_{i=1}^{n} f_i(p_t)p_{it-\text{dt}}\sum_{k=1}^{K}(\lambda_{ikt-\text{dt}}-\lambda_{kt-\text{dt}}) + \frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n} f_{ij}(p_t)h_t p_{it-\text{dt}}p_{jt-\text{dt}}(\mu_i-\mu_{t-\text{dt}})(\mu_j-\mu_{t-\text{dt}}) + \frac{o(\text{dt})}{\text{dt}} \\
&= -\sum_{i=1}^{n}\sum_{k=1}^{K} f_i(p_t)(\lambda_{ikt-\text{dt}}-\lambda_{kt-\text{dt}})p_{it-\text{dt}} \\
&\quad + \frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n} f_{ij}(p_t)h_{t-\text{dt}}p_{it-\text{dt}}p_{jt-\text{dt}}(\mu_i-\mu_{t-\text{dt}})(\mu_j-\mu_{t-\text{dt}}) + \frac{o(\text{dt})}{\text{dt}} \\
&= -\sum_{k=1}^{K}\lambda_{kt-\text{dt}}\nabla f(p_t)'(\nu_{kt}-p_{t-\text{dt}}) + \frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n} f_{ij}(p_t)h_{t-\text{dt}}p_{it-\text{dt}}p_{jt-\text{dt}}(\mu_i-\mu_{t-\text{dt}})(\mu_j-\mu_{t-\text{dt}}) + \frac{o(\text{dt})}{\text{dt}}
\end{aligned}$$

The first line states the Taylor approximation to a second degree and the second step restates the argument in summation form. Next, the third line exploits the fact that expectations are linear operators, while the fourth line imports the approximations for these changes in value from the previous sections of the proof. The next equality regroups terms, while the last re-writes the equation above taking into account that above and replaces them for the expressions derived above. The third line exploits the observation that $\nu_{ikt-\text{dt}} \equiv \lambda_{ikt-\text{dt}}p_{it-\text{dt}}/\lambda_{kt-\text{dt}}$ by Bayes rule.

Next, I calculate $E[\text{d}f(p_t)]/\text{dt}$. Since the probability of a jump of size Δ_k is approximately $\lambda_{kt-\text{dt}}\text{dt}$, then the unconditional expectation equals to

$$
\begin{aligned}
\frac{E[\mathrm{d}f(p_t)]}{\mathrm{dt}} &= \sum_{k=1}^{K} \frac{\lambda_{kt}\mathrm{dt}}{\mathrm{dt}} E[\mathrm{d}f(p_t) \mid J_{kt}^{\mathrm{dt}} = \Delta_k] + \left[1 - \mathrm{dt} \sum_{k=1}^{K} \lambda_{kt}\right] \times \frac{E[df(p_t)|\forall k, J_{kt}^{\mathrm{dt}} = 0]}{\mathrm{dt}} \\
&= \sum_{k=1}^{K} \lambda_{kt}[f(\nu_{kt}) - f(p_{t-\mathrm{dt}})] + \left[1 - \mathrm{dt} \sum_{k=1}^{K} \lambda_{kt}\right] \times \left[\frac{1}{2} \sum_{i=1}^{n} \sum_{j=1}^{n} f_{ij}(p_t) h_{t-\mathrm{dt}} p_{it-\mathrm{dt}} p_{jt-\mathrm{dt}} (\mu_i - \mu_{t-\mathrm{dt}})(\mu_j - \mu_{t-\mathrm{dt}}) \right. \\
&\qquad \left. - \sum_{k=1}^{K} \lambda_{kt-\mathrm{dt}} \nabla f(p_{t-\mathrm{dt}})'(\nu_{kt} - p_{t-\mathrm{dt}})\right] + \frac{o(\mathrm{dt})}{\mathrm{dt}} \\
&= \sum_{k=1}^{K} \lambda_{kt}[f(\nu_{kt}) - f(p_{t-\mathrm{dt}}) - \nabla f(p_{t-\mathrm{dt}})'(\nu_{kt} - p_{t-\mathrm{dt}})] \\
&\quad + \frac{1}{2} \sum_{i=1}^{n} \sum_{j=1}^{n} f_{ij}(p_t) h_{t-\mathrm{dt}} p_{it-\mathrm{dt}} p_{jt-\mathrm{dt}} (\mu_i - \mu_{t-\mathrm{dt}})(\mu_j - \mu_{t-\mathrm{dt}}) + \frac{o(\mathrm{dt})}{\mathrm{dt}}
\end{aligned}
$$

The first equality just applies the Law of iterated expectations and the second equality replaces the conditional expectations with the approximations that were derived above. The last equality just collects terms. Lastly, take the limit as dt goes to 0 it yields that

$$
\mathcal{L}f(p_t) = \sum_{k=1}^{K} \lambda_{kt}[f(\nu_{kt}) - f(p_t) - \nabla f(p_t)'(\nu_{kt} - p_t)] + \frac{h_t}{2} \sum_{i=1}^{n} \sum_{j=1}^{n} f_{ij}(p_t) p_{it} p_{jt} (\mu_i - \mu_t)(\mu_j - \mu_t)
$$

This concludes the proof.