Rational Inattention in Continuous Time*

Benjamin Hébert †
Stanford University

Michael Woodford ‡
Columbia University

November 14, 2018

Abstract

We derive, from first principles, continuous time models of rational inattention. We begin by describing dynamic discrete time models, and stating conditions on the cost of information in such models. Using these conditions, we characterize the cost of a small amount of information, which allows us to derive tractable and parsimonious models for the continuous time limit. We then provide conditions under which the resulting belief dynamics will resemble diffusion processes, and when the beliefs will involve large jumps. There is an extensive literature in psychology, neuroscience, and behavioral economics on continuous time models of belief dynamics; our results can be viewed as a sort of micro-foundation for models in this literature.

*The authors would like to thank Mark Dean, Sebastian Di Tella, Mira Frick, Matthew Gentzkow, Emir Kamenica, Divya Kirti, Jacob Leshno, Stephen Morris, Pietro Ortoleva, Doron Ravid, José Scheinkman, Ilya Segal, Ran Shorrer, Miguel Villas-Boas, Ming Yang, and participants at the Cowles Theory conference, 16th SAET Conference, Barcelona GSE Summer Conference on Stochastic Choice, Stanford GSB research lunch, 2018 ASSA meetings, and UC Berkeley Theory Seminar for helpful discussions on this topic, and the NSF for research support. We would particularly like to thank Philipp Strack for discussing an earlier version of the paper. Portions of this paper circulated previously as the working paper "Rational Inattention with Sequential Information Sampling," which was split in two, and appeared in Benjamin Hébert's Ph.D. dissertation at Harvard University. All remaining errors are our own.

†Hébert: Stanford University. Email: bhebert@stanford.edu.
‡Woodford: Columbia University. Email: mw2230@columbia.edu.

1 Introduction

The theory of rational inattention, proposed by Christopher Sims and surveyed in Sims (2010), endogenizes the imperfect awareness that a decision maker (DM) has about the circumstances under which she must choose an action. According to the theory, the DM chooses her action on the basis of a subjective representation of the decision situation, with imperfect knowledge of the true state. The information structure is assumed to be optimal, in the sense that the DM chooses the best possible state-contingent action choice, net of a cost of information. In Sims' theory, that cost of information is proportional to the Shannon mutual information between the true state and the information observed by the DM.

It is not obvious, though, that the theorems that justify the use of mutual information in communications engineering (Cover and Thomas (2012)) provide any warrant for using it as a cost function in a theory of attention allocation, either in the case of economic decisions or that of perceptual judgments.[1] In addition, the mutual-information cost function has implications that are unappealing on their face, and that seem inconsistent with evidence on the nature of sensory processing, as discussed in, for example, Woodford (2012) and Caplin et al. (2017).

At the same time, there is a growing interest in dynamic models of information acquisition, in which the DM makes a sequence of decisions about what information to acquire before eventually choosing her action. Examples include Moscarini and Smith (2001), Woodford (2014), Fudenberg et al. (2015), and Steiner et al. (2017). We believe that it is often quite realistic to assume that information is acquired through a sequential sampling process. As discussed in Fehr and Rangel (2011) and Woodford (2014), an extensive literature in psychology and neuroscience has argued that data on both the frequency of per-

[1] As explained in Cover and Thomas (2012), these theorems rely upon the possibility of "block coding" of a large number of independent instances of a given type of message, that can be jointly transmitted before any of the messages have to be decoded by the recipient. In our situation, an action must be taken in an individual decision problem, without waiting to learn about a large number of problems of the same form.

ceptual errors and the frequency distribution of response times can be explained by models of perceptual classification based on sequential sampling. More recently, some authors have proposed that data on stochastic choice and response time in economic contexts can be similarly modeled.[2]

In this paper, we derive the continuous time limit of a dynamic rational inattention model. We place minimal restrictions on the nature of the information cost function, and yet are able to form relatively sharp conclusions about the nature of the continuous time limit. The continuous time limit we derive is reasonably tractable, especially in the limiting case in which the DM's rate of time preference is small relative to the time required to make a decision. We provide conditions under which the DM will choose beliefs that follow, in the continuous time limit, both diffusion and pure-jump processes. We apply these continuous time models in a related paper, Hébert and Woodford (2017), demonstrating their tractability and usefulness.

We begin by describing static and discrete-time dynamic rational inattention problems. The key innovations of our paper, relative to, for example, Steiner et al. (2017), are two-fold. First, we impose a cost/constraint that is convex in the amount of information acquired each period. This step builds on a similar assumption in Moscarini and Smith (2001), and is critical for generating smooth information gathering over time. Without this convexity, the DM would (for some cost functions) have no need to learn information prior to acting on that information, as in Steiner et al. (2017). Second, we consider general classes of "flow" information cost functions, characterized by certain conditions, as in Caplin and Dean (2015), De Oliveira et al. (2017), and Caplin et al. (2017). This step allow us to consider a more general family of cost functions that includes both mutual information and various proposed alternatives, including the neighborhood cost functions we advocate in

[2]In addition to the references in Fehr and Rangel (2011), recent examples include Krajbich et al. (2014) and Clithero (2016). Shadlen and Shohamy (2016) provide a neural-process interpretation of sequential-sampling models of choice.

Hébert and Woodford (2017).

After describing the discrete time dynamic rational inattention problem and the conditions we impose on our cost functions, we introduce approximations for our cost functions that become increasingly accurate as the quantity of information gathered becomes small. These approximations rely heavily on mathematical results described in Chentsov (1982), which were used in a different context in Hébert (2014). These approximations, along with the convexity described previously, are the key to our derivation of the continuous time problem. The convexity of the flow information costs guarantees that as the size of each time period grows short, less and less information is acquired, and consequently our approximations become increasingly accurate.

Our approximations can be summarized as stating that the cost of a small amount of information is entirely characterized by a divergence.[3] There are, however, two ways in which signal can contain only a small amount of information. The first way, to have frequent but uninformative signals, corresponds in the continuous time limit to a diffusion, and the cost of the diffusion is characterized by the second-order properties of the divergence. The second way is to have rare but informative signals, which corresponds to jumps in beliefs in the continuous time limit, the cost of which is also characterized by the divergence.

We then turn to our main results, which present the continuous time limit of the discrete time dynamic models. We first characterize the limit as a controlled jump-diffusion, and then show that in the absence of time discounting, the problem has a particularly simple formulation. We then describe conditions under which beliefs follow a diffusion process, with no jumps, and conditions under which beliefs involve jumps and immediate action. We also make a more general point, that pure jump processes with very frequent and very small jumps can resemble diffusions. With time discounting, the results of Zhong (2017)

[3]In the terms used by Caplin et al. (2017), every cost function is approximately posterior-separable.

show, in some special cases, that beliefs generally follow a pure jump process; we show that as the rate of time discounting converges to zero, these jumps become very frequent and small, eventually converging to a diffusion.

Our paper is not the first that seeks to derive at least some features of such models from a theory of optimal information sampling. In particular, Moscarini and Smith (2001) consider both the optimal intensity of information sampling per unit of time and the optimal stopping problem, when the only possible kind of information is given by the sample path of a Brownian motion with a drift that depends on the unknown state of the world.[4] Fudenberg et al. (2015) consider a variant of this problem with a continuum of possible states, and an exogenously fixed sampling intensity.[5] Woodford (2014) instead takes as given a stopping rule (motivated by the empirical psychology and neuroscience literatures), but allows a very flexible choice of the information sampling process, as in theories of rational inattention. Our approach differs from these earlier efforts in seeking to endogenize *both* the nature of the information that is sampled at each stage of the evidence accumulation process and the stopping rule that determines how much evidence is collected before a decision is made. Both Morris and Strack (2017) and Zhong (2017) adopt our approach and consider special cases of our convergence result, applying to mutual information (Morris and Strack (2017)) and to uniformly posterior-separable cost functions (Zhong (2017)).

We also consider decision problems with an arbitrary finite number of choice alternatives, rather than restricting attention to binary choice problems, as in both Fudenberg et al. (2015) and Woodford (2014). In the sequential information sampling problem considered here, we allow the information sampled at each stage to be chosen very flexibly, as in Woodford (2014), subject only to a "flow" information-cost function; but we also allow the

[4]Moscarini and Smith (2001) allow the instantaneous variance of the observation process to be freely chosen (subject to a cost), but this is equivalent to changing how much of the sample path of a given Brownian motion can be observed by the DM within a given amount of clock time.

[5]See also Tajima et al. (2016) for analysis of a related class of models.

decision when to stop sampling and make a decision to be made optimally, on the basis of the entire history of information sampled to that point, as in Moscarini and Smith (2001) and Fudenberg et al. (2015).

For a particular family of flow information-cost functions, the cost function for the equivalent static model is just the mutual information between the action chosen and the true state of the world; we thus provide foundations for the kind of rational inattention problem proposed by Sims (2010),[6] that do not rely on any analogy with rate-distortion theory in communications engineering. But while our dynamic model makes predictions that are equivalent to those of the rational inattention theory of Sims (2010) (and more particularly, its application to stochastic choice by Matêjka et al. (2015)) for this particular family of flow information-cost functions, we show that different predictions can be obtained under other, very plausible specifications of the flow cost function.

Section 2 by defining static rational inattention problems, and introducing the discrete time dynamic problem we study. In section 3, we describe the conditions we impose on the flow information cost function. In Section 4, we show that all flow cost functions satisfying these conditions must have a particular type of local structure when individual "experiments" are nearly uninformative. Section 5 defines our continuous time model, and demonstrates that it is the continuous time limit of the our discrete time problem. It also describes the conditions under which beliefs follow a diffusion or diffusion-like process. In section 6 we conclude.

[6] Morris and Strack (2017) provide a related foundation for the mutual-information cost function, but for the special case in which there are only two possible states.

2 Static and Dynamic Models of Rational Inattention

We begin by describing the class of static rational inattention models surveyed by Sims (2010), and then describe the discrete time dynamic model we study.

First, we introduce some notation. Given a set S, we define $\mathscr{P}(S)$ as the probability simplex associated with that set. We describe an element of the simplex $r \in \mathscr{P}(S)$ by a vector in $\mathbb{R}_+^{|S|}$ whose elements sum to one, each of which corresponds to the likelihood that a particular signal $s \in S$ is realized. Except when necessary, we will call this vector r as well, and not distinguish between the element of the simplex and its coordinate representation.

2.1 Static Models of Rational Inattention

Let $x \in X$ be the underlying state of nature, and let $s \in S$ be a signal the decision maker (DM) can receive, which might convey information about the state. We assume that X and S are finite sets. Let q denote the DM's prior belief (before receiving a signal) about the probability of state x; that is, q is element of the probability simplex $\mathscr{P}(X)$. Define $p_{s,x}$ as the probability of receiving signal s in state x, let $p_x \in \mathscr{P}(S)$ be the associated conditional probability distribution of the signals given state x, and let p be the $|S| \times |X|$ matrix whose elements are $p_{s,x}$. The matrix p, which is a set of conditional probability distributions for each state of nature, $\{p_x\}_{x \in X}$, defines an "information structure." After receiving signal s, the DM will hold a posterior, $q_s \in \mathscr{P}(X)$, which is a function of p and q, defined by Bayes' rule.

Let $a \in A$ be the action taken by the decision maker (DM). For simplicity, A is also a finite set, and we assume that the number of states is weakly larger than the number of actions, $|X| \geq |A|$. The DM's utility from taking action a in state x is $u(a,x)$. We assume that $u(a,x)$ is strictly positive and bounded above by a positive constant, $\bar{u}$.

The maximum achievable expected payoff, given an information structure p and prior

q, can be written as

$$u(p,q) \equiv \max_{\{a(s)\}} \sum_{x\in X}\sum_{s\in S} q_x p_{s,x} u(a(s),x). \tag{1}$$

The standard static rational inattention problem, given the signal alphabet S,[7] is then

$$\max_{\{p_x\in\mathscr{P}(S)\}_{x\in X}} u(p,q) - \theta C(p,q;S), \tag{2}$$

where

$$C(\cdot,\cdot;S) : \mathscr{P}(S)^{|X|} \times \mathscr{P}(X) \to \mathbb{R} \tag{3}$$

is a cost function for information structures, and $\theta > 0$ is a multiplicative factor.[8]

In the classic formulation of Sims, a problem of the form equation (2) is considered, in which the cost function $C(p,q;S)$ is given by the Shannon mutual information between the signal and the state. This can be defined using Shannon's entropy,

$$H^{Shannon}(q) \equiv -\sum_{x\in X} q_x \ln(q_x). \tag{4}$$

Shannon's entropy can in turn be used to define a measure of the degree to which the posterior q_s associated with any signal differs from the prior q, the Kullback-Leibler (KL) divergence,

$$D_{KL}(q_s||q) \equiv H^{Shannon}(q) - H^{Shannon}(q_s) + (q_s - q)^T H_q^{Shannon}(q). \tag{5}$$

[7]The full problem includes a choice over the signal alphabet S. A standard result, which will hold for all of the cost functions we study, is that $|S| = |A|$ is sufficient.

[8]The scalar θ and cost function $C(\cdot,\cdot;S)$ are redundant. Using a separate scalar parameter θ allows us to consider alternative assumptions about the tightness of the information constraint, holding fixed the measure of the informativeness of information structures represented by the function C.

Mutual information is then the expected value of the KL divergence over possible signals,

$$C^{MI}(p,q;S) \equiv \sum_{s\in S}(e_s^T pq)D_{KL}(q_s||q), \tag{6}$$

where $e_s \in \mathbb{R}^{|S|}$ is vector with one for the signal s and zero otherwise. It is a measure of the informativeness of the signal, in that it provides a measure of the degree to which the signal changes what one should believe about the state, on average.

However, Shannon's mutual information is not, however, the only possible measure of the informativeness of an information structure, or the only plausible cost function for a static rational inattention problem. We adopt a more general approach (following De Oliveira et al. (2017) and Caplin and Dean (2015)) and study the entire class of costs functions satisfying certain conditions. Before discussing these conditions, we introduce our discrete time dynamic model.

2.2 Dynamic Models of Rational Inattention

We extend the static rational inattention model just described to a discrete time dynamic setting, in which the DM has numerous opportunities to gather information before taking an action. Our extension is similar to the one described by Steiner et al. (2017), with a few key differences. First, unlike those authors, we study a setting in which the DM can take an action only once, and chooses when to stop and take an action endogenously. Second, we are interested in general cost functions of the form described by equation (3), not just mutual information. Third, we assume the DM has a motivate to smooth her information gathering over time, rather than learn all of the relevant information in a single period.

As in the static model, there is a state of the world, $x \in X$, that remains constant over time. At each time t, the DM can either stop and take an action $a \in A$, or continue and receive a signal drawn from the information structure $\{p_{t,x} \in \mathscr{P}(S)\}_{x\in X}$, for some signal alphabet S. We also assume that the signal alphabet S is finite and fixed over time, with $|S| \geq 2|X|+2$. However, the information structure $\{p_{t,x}\}_{x \in X}$ is a choice variable that can be state- and time-dependent. Fixing the signal alphabet S has no economic meaning, because the information content of receiving a particular signal $s \in S$ can change between periods. The assumption allows us to assume a finite information structure and invoke the results from section 3 below. As a technical device, we assume that S contains one signal, $\bar{s}$, that is required to be uninformative. This assumption ensures that the DM can choose to mix any arbitrary signal structure with an uninformative one, even if she has already used up her "useful" signals.

The DM's prior beliefs at time t, before receiving the signal, are denoted q_t. Each time period has a length Δ. Let τ denote the time at which the DM stops and makes a decision, with $\tau = 0$ corresponding to making a decision without acquiring any information. At this time, the DM receives utility $u(a,x) - \kappa\tau$ if she takes action a at time τ and the true state of the world is x. Let $\hat{u}(q_\tau)$ be the utility (not including the penalty for delay) associated with taking an optimal action under beliefs q_τ. We assume, as in the static model, that $u(a,x)$ is strictly positive and bounded above by the constant $\bar{u}$.

Each time period has length Δ, and the DM discounts the future exponentially, with discount factor β^Δ, for some $\beta \in (0,1]$. The parameter κ and discount factor β together govern the size of the penalty the DM faces from delaying his decision. The reason the DM does not make a decision immediately is that she is able to gather information, and make a more-informed decision.

The DM can choose an information structure that depends on the current time and past history of the signals received. As we will see, the problem has a Markov structure, and the current time's "prior," q_t, summarizes all of the relevant information that the DM needs to design the information structure. The DM is constrained to satisfy

$$E_0[\frac{\Delta}{\rho}\sum_{j=0}^{\tau\Delta^{-1}-1}\beta^{\Delta j}C(\{p_{\Delta j},q_{\Delta j};S)^{\rho}]^{\frac{1}{\rho}} \leq \Delta c E_0[\Delta\sum_{j=0}^{\tau\Delta^{-1}-1}\beta^{\Delta j}]^{\frac{1}{\rho}}, \tag{7}$$

if the DM choose to acquire any information at all ($\tau > 0$ always in this case). In words, the L^{ρ}-norm of the flow information cost function $C(\cdot)$ over time and possible histories must be less that the constant c per unit time. In the particular case in which information gathering is constant ($C(\{p_{\Delta j},q_{\Delta j};S) = \bar{C}$), this constraint simplifies to

$$\frac{\bar{C}}{\Delta} \leq \rho^{\frac{1}{\rho}}c,$$

emphasizing that the constant c can be thought of as a limit on the flow rate of information acquisition. The parameter ρ governs the substitutability of information acquisition across time. In the limit as $\rho \to \infty$, the L^{ρ} norm becomes the essential supremum, and the constraint approaches a per-period constraint on the amount of information the DM can obtain. For finite values of ρ, the DM can allocate more information gathering to states and times in which it is more advantageous to gather more information. We assume, however, that $\rho > 1$, to ensure that it is optimal for the DM to gather information gradually, rather than all at once. Our assumption of $\rho > 1$ is similar to the convex cost of the rate of experimentation assumed by Moscarini and Smith (2001). It is a critical assumption that separates our model from the model of Steiner et al. (2017), in which the DM will (under some circumstances) learn a large amount of information in a single period. We will also assume that the flow cost function $C(\cdot)$ satisfies the conditions that we introduce in the next section.

Let $V(q_0;\Delta)$ denote the value obtained in the sequence problem for a DM with prior beliefs q_0, and let q_τ denote the DM's beliefs when stopping to act. The DM's problem is

$$V(q_0;\Delta) = \max_{\{p_{\Delta j}\},\tau} E_0[\beta^\tau \hat{u}(q_\tau) - \Delta\kappa \frac{1-\beta^{\tau+\Delta}}{1-\beta^\Delta})], \tag{8}$$

subject to the information-cost constraint (7).

For an alternative interpretation of this constraint, consider the dual version of this problem, $\min_{\lambda \geq 0} W(q_0,\lambda;\Delta)$, where

$$W(q_0,\lambda;\Delta) = \max_{\{p_{\Delta j}\},\tau} E_0[\beta^\tau \hat{u}(q_\tau) - \Delta\kappa \frac{1-\beta^{\tau+\Delta}}{1-\beta^\Delta})] - \lambda E_0[\Delta^{1-\rho} \sum_{j=0}^{\tau\Delta^{-1}-1} \beta^{\Delta j} \{\frac{1}{\rho} C(p_{\Delta j}, q_{\Delta j}; S)^\rho - \Delta^\rho c^\rho\}]. \tag{9}$$

The function $W(q_0,\lambda;\Delta)$ can be thought of as the value function of a different problem, in which there is a cost of gathering information proportional to $\lambda \frac{1}{\rho} C(\cdot)^\rho$.

We will describe the continuous time limits of these functions, $W(q_0,\lambda) = \lim_{\Delta \to 0^+} W(q_0,\lambda;\Delta)$ and $V(q_0) = \lim_{\Delta \to 0^+} V(q_0;\Delta)$. Our approach is to assume standard conditions on the cost function $C(\cdot)$, which we describe in the next section, and use these conditions to characterize the cost of a small amount of information, which we discuss in section §4. Because of our $\rho > 1$ assumption, it will be optimal, as $\Delta \to 0^+$, for the DM to acquire a small amount of information each period, and consequently our results about the cost of a small amount of information will be relevant for the continuous time limit. We present these results in section §5.

3 Flow Information Costs

At each date in the discrete-time sequential rational inattention problem just described, the DM chooses an information structure. Each information structure p has a cost $C(p,q;S)$, given by a function of the form of (3), where q indicates the DM's prior in this date (that is, the posterior beliefs following from observations prior to the current date of the dynamic problem), and S is the signal alphabet.[9] Our results depend only on assuming that this flow information-cost function satisfies a set of five general conditions, stated below.

All of these conditions are satisfied by the mutual-information cost function (6) proposed by Sims, but they are also satisfied by many other cost functions (for example, the neighborhood-based cost functions we describe in Hébert and Woodford (2017)). They are closely related to conditions that other authors (De Oliveira et al. (2017) and Caplin and Dean (2015)) have also proposed as attractive general properties to assume about information-cost function, in the context of static rational inattention models.

Condition 1. Information structures that convey no information ($p_x = p_{x'}$ for all x, x' in the support of q) have zero cost. All other information structures have a strictly positive cost.

This condition ensures that the least costly strategy for the DM in the standard static rational inattention problem is to acquire no information, and make her decision based on the prior. The requirement that gathering no information has zero cost is a normalization.

The next condition is called mixture feasibility by Caplin and Dean (2015). Consider two information structures, $\{p_{1,x}\}_{x\in X}$, with signal alphabet S_1, and $\{p_{2,x}\}_{x\in X}$, with alphabet S_2. Given a parameter $\lambda \in (0,1)$, we define a mixed information structure, $\{p_{M,x}\}_{x\in X}$ over the signal alphabet $S_M = (S_1 \cup S_2) \times \{1,2\}$. For each $s = (s_1, 1)$ in the alphabet S_M, $p_{M,x}(s)$ is equal to $\lambda p_{1,x}(s)$ if $s_1 \in S_1$, and equal to 0 otherwise. Likewise, for each $s = (s_2, 2)$, $p_{M,x}(s)$ is equal to $(1-\lambda)p_{2,x}(s)$ if $s_2 \in S_2$, and equal to 0 otherwise.

That is, this information structure results, with probability λ, in a posterior associated with information structure p_1, and with probability $1-\lambda$ in a posterior associated with in-

[9] The information-cost functions that we study, like mutual information, are defined for all finite signal alphabets S. Note, however, that mutual information is also defined over alternative sets of states of nature X. We do not impose this requirement on our more general cost functions — all of our analysis takes the set of states of nature as given.

formation structure p_2. The distribution of posteriors under the mixed information structure is a convex combination of the distributions of posteriors under the two original information structures, as if the DM flipped a coin, observed the result, and then randomly chose one of the two information structures. The mixture feasibility condition requires that choosing a mixed information structure costs no more than the cost of randomizing over information structures (using a mixed strategy in the rational inattention problem).

Condition 2. Given two information structures, $\{p_{1,x}\}_{x\in X}$, with signal alphabet S_1, and $\{p_{2,x}\}_{x\in X}$, with alphabet S_2, the cost of the mixed information structure is weakly less than the weighted average of the cost of the separate information structures:

$$C(p_M, q; S_M) \leq \lambda C(p_1, q; S_1) + (1-\lambda) C(p_2, q; S_2).$$

The next condition uses Blackwell's ordering. Consider two signal structures, $\{p_x\}_{x\in X}$, with signal alphabet S, and $\{p'_x\}_{x\in X}$, with alphabet S'. The first information structure Blackwell dominates the second information structure if, for all utility functions $u(a,x)$ and all priors $q \in \mathscr{P}(X)$,

$$\bar{u}(p,q) \;\geq\; \bar{u}(p',q),$$

where $\bar{u}(p,q)$ is defined as in equation (1). In words, if one information structure Blackwell dominates another, it is weakly more useful for every decision maker, regardless of that decision maker's utility function and prior. In this sense, it conveys weakly more information. This ordering is incomplete; most information structures neither dominate nor are dominated by a given alternative information structure. However, when an information structure does Blackwell dominate another one, we assume that the dominant information structure is weakly more costly.

Condition 3. If the information structure $\{p_x\}_{x\in X}$ with signal alphabet S is more informative, in the Blackwell sense, than $\{p'_x\}_{x\in X}$, with signal alphabet S', then, for all $q \in \mathscr{P}(X)$,

$$C(\{p_x\}_{x\in X}, q; S) \geq C(\{p'_x\}_{x\in X}, q; S').$$

The first three conditions are, from a certain perspective, almost innocuous. For any joint distribution of actions and states that could have been generated by a DM solving a rational inattention type problem, with an arbitrary information cost function, there is a cost function consistent with these three conditions that also could have generated that data (Theorem 2 of Caplin and Dean (2015)). The result arises from the possibility of the DM pursuing mixed strategies over information structures, or in the mapping between signals and actions. These conditions also characterize "canonical" rational inattention cost functions, in the terminology of De Oliveira et al. (2017).

The mixture feasibility condition (Condition 2) and Blackwell monotonicity condition (Condition 3) are equivalent to requiring that the cost function be convex over information structures and Blackwell monotone.

Lemma 1. *Let p and p' be information structures with signal alphabet S. A cost function is convex in information structures if, for all $\lambda \in (0,1)$, all signal alphabets S, and all $q \in \mathscr{P}(X)$,*

$$C(\lambda p + (1-\lambda)p', q; S) \leq \lambda C(p, q; S) + (1-\lambda)C(p', q; S).$$

A cost function satisfies mixture feasibility and Blackwell monotonicity (Conditions 2 and 3) if and only if it is convex in information structures and satisfies Blackwell monotonicity.

Proof. See the appendix, section D.1. □

The fourth condition that we assume, which is not imposed by Caplin and Dean (2015), Caplin et al. (2017), or De Oliveira et al. (2017), is a differentiability condition that will allow us to characterize the local properties of our cost functions.

Condition 4. There exists an $\varepsilon > 0$ such that, for all signal alphabets S and priors q and all information structures sufficiently close to uninformative ($||p - q^T p^T \iota|| < \varepsilon$), the information cost function is continuously twice-differentiable in information structures $\{p_x\}_{x \in X}$, in all directions that do not change the support of the signal distribution, and directionally differentiable, with continuous directional derivatives, with respect to perturbations that increase the support of the signal distribution.

While this may seem a relatively innocuous regularity condition, it is not completely general; for example, it rules out the case in which the DM is constrained to use only signals in a parametric family of probability distributions, and the cost of other information structures is infinite. Thus it rules out information structures of the kind assumed in Fudenberg et al. (2015) or Morris and Strack (2017). Condition 4 also rules out other proposed alternatives, such as the channel-capacity constraint suggested by Woodford (2012). The "$\varepsilon > 0$" part of the condition indicates that that this differentiability need only hold at nearly uninformative information structures; obviously, if differentiability holds everywhere, the condition will be satisfied.

The next condition that we assume, which is also not imposed by Caplin and Dean (2015), Caplin et al. (2017), or De Oliveira et al. (2017), is a sort of local strong convexity. We will assume that the cost function exhibits strong convexity, in the neighborhood of an uninformative information structure, with respect to information structures that hold fixed the unconditional distribution of signals, uniformly over the set of possible priors.

Condition 5. There exists constants $m > 0$ and $B > 0$ such that, for all priors $q \in \mathscr{P}(X)$, and all information structures that are sufficiently close to uninformative ($C(p,q;S) < B$),

$$C(p,q;S) \geq \frac{m}{2} \sum_{s \in S} (e_s^T pq) ||q_s - q||_X^2,$$

where q_s is the posterior given by Bayes' rule and $||\cdot||_X$ is an arbitrary norm on the tangent space of $\mathscr{P}(X)$.

This condition is slightly stronger than Condition 1; it it essentially an assumption of "local strong convexity" instead of merely local strict convexity. It implies that all informative information structures have a non-trivial positive cost, and that (regardless of the DMs' current beliefs) there are no informative information structures that are "almost free." This condition allows us to assert that if the flow cost $C(p,q;S)$ is converging to zero, then either the posteriors must become close to the prior (q_s close to q) or the signals must become rare ($e_s^T pq$ close to zero), or some combination thereof.

The mutual-information cost function (6) satisfies each of these five conditions. However, it is not the only cost function to do so. For example, we can construct a family of such cost functions, using the family of "f-divergences" (see, e.g., Ali and Silvey (1966) or Amari and Nagaoka (2007)), defined as

$$D_f(q_s||q) = \sum_{x \in X} (q_x) f(\frac{q_{s,x}}{q_x}),$$

where f is any strictly convex, twice-differentiable function with $f(1) = f'(1) = 0$ and $f''(1) = 1$. The KL divergence is a member of this family, corresponding to $f(u) = u \ln u - u + 1$.

For any divergence in this family, we can define an information cost function

$$C^f(p,q;S) = \sum_{s \in S} (e_s^T pq) D_f(q_s||q) \tag{10}$$

This family of information cost functions satisfies all five of the conditions described above. Another example satisfying our conditions, as mentioned previously, are the neighborhood cost functions of described by Hébert and Woodford (2017). These cost functions fall into the "uniformly posterior-separable" family of cost functions described by Caplin et al. (2017), and (under mild regularity assumptions) all such functions satisfy our conditions. We discuss these issues in more detail in the appendix, section §A.

We are now in a position to use these conditions to characterize the cost of a small amount of information, the topic of the next section.

4 The Cost of a Small Amount of Information

In this section, we describe the cost of a small amount of information. We use Taylor's theorem to approximate the cost function and its gradient up to order Δ (a second-order approximation for the cost function, first-order for the gradient). In this section, Δ is simply "a small number squared," but when we apply our results to our dynamic rational inattention model, it will be the length of the time period.

We start by describing the local (second-order) properties of any information cost function satisfying our conditions. The condition requiring that Blackwell-dominant information structures cost weakly more (Condition 3) is of particular importance. To understand why, it is first useful to recall Blackwell's theorem.

Theorem. *(Blackwell (1953)) The information structure $\{p_x\}_{x\in X}$, with signal alphabet S, is more informative, in the Blackwell sense, than $\{p'_x\}_{x\in X}$, with signal alphabet S', if and only if there exists a Markov transition matrix $\Pi : S \to S'$ such that, for all $s' \in S'$ and $x \in X$,*

$$p'_x = \Pi p_x. \tag{11}$$

This Markov transition matrix is known as the “garbling” matrix. Another way of interpreting Condition 3 is that garbled signals are (weakly) less costly than the original signal.

There are certain kinds of garbling matrices that don’t really garble the signals. These garbling matrices have left inverses that are also Markov transition matrices. If we define an information structure $\{p_x\}_{x\in X}$, with signal alphabet S, and another information structure $\{p'_x\}_{x\in X}$, with signal alphabet S', using one of these left-invertible matrices, via equation (11), then $\{p_x\}_{x\in X}$ is more informative than $\{p'_x\}_{x\in X}$, but $\{p'_x\}_{x\in X}$ is also more informative than $\{p_x\}_{x\in X}$. These two information structures are called “Blackwell-equivalent,” and it follows that the cost of these two information structures must be equal, by Condition 3. The left-invertible Markov transition matrices associated with Blackwell-equivalent information structures are called Markov congruent embeddings by Chentsov (1982). Chentsov (1982) studied tensors and divergences that are invariant to Markov congruent embeddings (we will say “invariant” for brevity).

An invariant divergence is a divergence that is invariant to these embeddings. Let Π be a Markov congruent embedding from $\mathscr{P}(S)$ to $\mathscr{P}(S')$. The KL divergence (and the f-divergences more generally) are invariant, meaning that

$$D_f(\Pi p||\Pi r) = D_f(p||r)$$

for all p, $r \in \mathscr{P}(S)$. There are also other, non-additively-separable invariant divergences. Chentsov’s theorem (Chentsov (1982)) states that, for any invariant divergence D_I,

$$\frac{\partial^2 D_I(p||r)}{\partial p^i \partial p^j}|_{p=r} = c \cdot g_{ij}(r), \tag{12}$$

where $c > 0$ is a positive constant and $g_{ij}(r)$ is the (i, j)-element of the Fisher information matrix evaluated at r. In the coordinate system for the simplex we have adopted,

$$g_{ij}(r) = \frac{\mathbf{1}(i=j)}{r_i} - 1.$$

However, the focus of this paper is not invariant divergences, but rather invariant information cost functions. By Condition 3, all information cost functions satisfying our conditions are invariant to Markov congruent embeddings. It necessarily follows that, for any Markov congruent embedding Π, that

$$C(\{p_x\}_{x\in X}, q; S) = C(\{\Pi p_x\}_{x\in X}, q; S').$$

Using this invariance, and results from Chentsov (1982), we will describe the local structure of all information cost functions satisfying our conditions.

Chentsov establishes the following results:[10]

i) Any continuous function that is invariant over the probability simplex is equal to a constant.

ii) Any continuous, invariant 1-form tensor field over the probability simplex is equal to zero.

iii) Any continuous, invariant quadratic form tensor field over the probability simplex is proportional to the Fisher information matrix.[11]

These results allow us to characterize the local properties of rational inattention cost functions, via a Taylor expansion. Hold fixed the signal alphabet S, and consider an information

[10]See Lemma 11.1, Lemma 11.2, and Theorem 11.1 in Chentsov (1982). See also Proposition 3.19 of Ay et al. (2014), who demonstrate how to extend the Chentsov results to infinite sets X and S.

[11]A 1-form tensor field on a probability simplex $\mathscr{P}$ is a function $T : V \times \mathscr{P} \to \mathbb{R}$, where V is the tangent space of the simplex. Let $\Pi : \mathscr{P} \to \mathscr{P}'$ be a mapping from the simplex $\mathscr{P}$ to the simplex $\mathscr{P}'$, let V' be the tangent space of the simplex $\mathscr{P}'$, and let $d\Pi : V \to V'$ be the pushforward of the mapping Π. The tensor field is invariant under Π if $T(d\Pi v, \Pi p) = T(v, p)$ for all $p \in \mathscr{P}$ and v in the tangent space at p, and a similar definition holds for quadratic form tensor fields.

structure

$$p_x(\varepsilon, \nu) = r + \varepsilon \nu_x + \nu \omega_x.$$

Here, $r \in \mathscr{P}(S)$ and $\nu_x \in \mathbb{R}^{|S|}$ satisfies $\iota^T \nu_x = 0$ for all x, where ι is a vector of ones. We also assume that, for all $s \in S$, $e_s^T \nu_x \neq 0$ only if $e_s^T r > 0$. That is, ν_x is an element of the tangent space of the probability simplex at r. The same properties hold true for ω_x. As a result, for values of the perturbation parameters ε and ν sufficiently close to zero, $p_x \in \mathscr{P}(S)$ for all $x \in X$. In other words, the parameters ε and ν index a two-parameter family of perturbations of an uninformative information structure (corresponding to $\varepsilon = \nu = 0$), in which the perturbed information structures will generally be informative; the ν_x and ω_x specify two directions of perturbation. Each of the perturbed information structures has the property that p_x is absolutely continuous with respect to r.

By Condition 1, $C(\{p_x(0,0)\}_{x \in X}; q; S) = 0$. The first order term is

$$\frac{\partial}{\partial \varepsilon} C(\{p_x(\varepsilon, \nu)\}_{x \in X}, q; S)|_{\varepsilon=\nu=0} = \sum_{x \in X} C_x(\{r\}_{x \in X}, q; S) \cdot \nu_x,$$

where C_x denotes the derivative with respect to p_x. This derivative, $C_x(\{r\}; q; S)$, forms a continuous 1-form tensor field over the probability simplex $\mathscr{P}(S)$. By the invariance of $C(\cdot)$, it also follows that C_x is invariant, and therefore, by Chentsov's results, it is equal to zero.

We repeat the argument for the second derivative terms. Those terms can be written as

$$\frac{\partial}{\partial \nu} \frac{\partial}{\partial \varepsilon} C(\{p_x(\varepsilon, \nu)\}_{x \in X}, q; S)|_{\varepsilon=\nu=0} = \sum_{x' \in X} \sum_{x \in X} \omega_{x'}^T \cdot C_{xx'}(\{r\}_{x \in X}, q; S) \cdot \nu_x.$$

By the invariance of $C(\cdot)$, the quadratic form $C_{xx'}(\cdot)$ is invariant for all $x, x' \in X$, and therefore is proportional to the Fisher information matrix for all $x, x' \in X$. We can define a matrix

$k(q)$ consisting of the constants of proportionality associated with each $x, x' \in X$. That is,

$$\frac{\partial}{\partial \nu} \frac{\partial}{\partial \varepsilon} C(\{p(\cdot|\cdot;\varepsilon,\nu)\},q)|_{\varepsilon=\nu=0} = \sum_{x' \in X} \sum_{x \in X} (e_x^T k(q) e_{x'}) \omega_{x'}^T g(r) \nu_x,$$

where $g(r)$ is the Fisher information matrix evaluated at the unconditional distribution of signals $r \in \mathscr{P}(S)$. We note that the matrix-valued function $k(q)$ can depend on the prior q, but cannot depend on the unconditional distribution of signals, r; otherwise, invariance would not hold. We refer to this matrix-valued function in Hébert and Woodford (2017) as the "information cost matrix function," because it is the matrix that describes the cost of a small of amount information.

We begin by considering perturbations that preserve the support of the signal structure. As a result, this theorem should be interpreted as applying to "frequent but not very informative" signals, as opposed to "rare but informative" signals. We will discuss the latter type of signals shortly.

Theorem 1. *Suppose that an information structure $\{p_x\}_{x \in X}$, with signal alphabet S, is described by the equation*

$$p_x = r + \Delta^{\frac{1}{2}} \nu_x + o(\Delta^{\frac{1}{2}}),$$

where, for any $x \in X$ and any $\Delta \geq 0$, $e_s^T p_x \neq 0 \Rightarrow e_s^T r > 0$. Let $C(\cdot)$ be an information cost function that satisfies Conditions 1-4.

There exists a matrix valued function $k(q)$ such that, for Δ sufficiently small,

$$C(\{p_x\}_{x \in X}; q; S) = \frac{1}{2} \Delta \sum_{x' \in X} \sum_{x \in X} (e_x^T k(q) e_{x'}) \nu_{x'}^T g(r) \nu_x + o(\Delta).$$

For all q, the matrix-valued function $k(q)$ is positive semi-definite and symmetric, and satisfies $k(q)\iota = 0$.

If in addition the cost function satisfies Condition 5, then there exists a constant $m_g > 0$

such that the difference between $k(q)$ and the pseudo-inverse of the Fisher information matrix, $g^{+}(q)$, multiplied by that constant, is positive semi-definite: $k(q) \succeq m_g g^{+}(q)$.

Proof. See the appendix, section D.2. □

In the case of the mutual-information cost function, the matrix $k(q)$ is itself the pseudo-inverse of the Fisher information matrix. Written in terms of the coordinate system we employ,

$$k(q) = g^{+}(q) = Diag(q) - qq^{T}.$$

In general, however, the matrix-valued function $k(q)$ is not the pseudo-inverse of the Fisher information matrix, but rather an arbitrary matrix-valued function satisfying certain restrictions.

The theorem substantially restricts the local structure of the cost of commonly occurring, but not particularly informative, signals, relative to the most general possible alternatives (which would not satisfy our conditions). Potential information structures $\{p_x\}_{x \in X}$ can be represented as vectors of dimension $N = (|S| - 1) \times |X|$. Under the assumptions of Condition 1, convexity, and Condition 4 (but not the Blackwell ordering condition, Condition 3), the cost function must locally resemble an inner product with respect to a positive semi-definite, $N \times N$ matrix. If we impose Condition 3 as well, the results of Theorem 1 show that we can restrict this matrix to the $k(q)$ matrix, an $|X| \times |X|$ matrix. If the DM were only allowed binary signals ($|S| = 2$), this restriction would be trivial. When the DM is allowed to contemplate more general information structures, the restriction is non-trivial.

Several authors (Caplin and Dean (2015); Kamenica and Gentzkow (2011)) have observed that it is easier to study rational inattention problems by considering the space of posteriors, conditional on receiving each signal, rather than space of signals. We can redefine the cost function using the posteriors and unconditional signal probabilities, rather than the prior and the conditional probabilities of signals. The corollary below expresses the results of Theorem 1 in terms of posterior beliefs.

Corollary 1. *Under the assumptions of Theorem 1, the posterior beliefs can be written, for any $s \in S$ such that $e_s^T r > 0$, as*

$$q_{s,x} = q_x + \Delta^{\frac{1}{2}} q_x \frac{e_s^T \nu_x}{e_s^T r} + o(\Delta^{\frac{1}{2}}).$$

The cost function can be written as

$$C(\{p_x\}_{x \in X}, q; S) = \frac{1}{2} \sum_{s \in S: e_s^T r > 0} (e_s^T r)(q_s - q)^T \bar{k}(q)(q_s - q) + o(\Delta),$$

where $\bar{k}(q) = Diag(q)^+ k(q) Diag(q)^+$.

Proof. See the appendix, section D.3. □

There are, in effect, two ways for a signal to be contain a small amount of information, and different costs associated with these different types of signals. The results of Theorem 1 characterize, for any rational inattention cost function satisfying our conditions, the cost of receiving frequently, but relatively uninformative, signals. As Corollary 1 above demonstrates, the posteriors associated with these signals are close to the prior (order $\Delta^{\frac{1}{2}}$). We will discuss the cost of receiving a rare but informative signal below. Previewing the results of section 5, these two types of uninformative signals correspond, in the continuous-time limit, to the diffusion and jump components of the belief process.

Corollary 2. *Under the assumptions of Theorem 1, define the signal structure*

$$\hat{p} = \bar{p}_\Delta + \Delta \omega,$$

where p_Δ is a signal structure of the type described in Theorem 1, with $\lim_{\Delta \to 0^+} \bar{p}_\Delta = r\iota^T$, and $\sum_{s \in S} \omega e_x = 0$ for all $x \in X$, with $e_s^T \omega e_x \geq 0$ for all $s \in S$ such that $e_s^T \bar{p}_\Delta = 0$.

The cost of this information structure can be written in the form

$$C(p_\Delta;q;S) = \frac{1}{2}\sum_{s\in S:e_s^T r>0}(e_s^T r)(q_s-q)^T\bar{k}(q)(q_s-q)$$
$$+\sum_{s\in S:e_s^T r=0}(e_s^T \omega q)D^*(q_s||q)+o(\Delta), \tag{13}$$

where the divergence D^* is finite and twice-differentiable in its first argument for q' sufficiently close to q, with

$$\frac{\partial^2 D^*(r||q)}{\partial r^i \partial r^j}|_{r=q} \;=\; \bar{k}(q). \tag{14}$$

Proof. See the appendix, section D.4. □

The divergence D^* represents the cost of acquiring an infrequent, but potentially informative, signal. Naturally, if the signal is in fact not very informative, this cost must be closely related to the costs of other uninformative signals, which gives rise to the condition on the Hessian of the divergence. Note that the corollary demonstrates that the cost is additive with respect to the other signals being received (at least up to order Δ). The result follows from the directional differentiability of the cost function with respect to signals that occur with zero probability and the continuity of that directional derivative (Condition 4) and invariance.

Now that we have derived our approximation results, in particular Corollary 2, we turn to considering the continuous time limit of the dynamic rational inattention problem introduced in section §2.

5 Rational Inattention in Continuous Time

We next study the limit of the dynamic rational inattention problem described in section §2. We will derive these limits for several specific cases, depending whether the DM's has exponential discounting ($\beta < 1$ vs. $\beta = 1$) and on the nature of the flow information cost function $C(\cdot)$. In particular, we consider two cases for $C(\cdot)$: when there is a "preference for gradual learning" and when there is a "preference for discrete learning," terms we define below. We will first preview our results, then provide them formally.

To gain intuition for our results, consider equation (13) in Corollary 2. Define $\Sigma(q)$ as the variance-covariance matrix of the posterior belief, and observe that it is of order Δ (by Corollary 1). We can rewrite the expression as

$$C(p_n;q;S) = \frac{1}{2}tr[\Sigma(q)\bar{k}(q)] + \Delta \sum_{s\in S: e_s^T r=0} (e_s^T \omega q) D^*(q_s||q) + o(\Delta). \tag{15}$$

Suppose (as we will prove) that in the limit as $\Delta \to 0^+$, the process q_t converges to a jump-diffusion:

$$dq_t = -\psi_t x_t dt + x_t dJ_t + Diag(q_t)\sigma_t dB_t,$$

where dJ_t is a poisson process with intensity ψ_t and dB_t is an $|X|$-dimensional Brownian motion. For this process, it is natural to further suppose that the "frequent but not very informative" signals are associated with the diffusion, and the "rare but potentially informative" signals are associated with the jumps. Under these assumption, it is reasonable to expect that the flow information cost is

$$C(\sigma_t, \psi_t, x_t; q_t)dt = \frac{1}{2}tr[\sigma_t \sigma_t^T k(q_t)]dt + \psi_t D^*(q_t + x||q_t)dt,$$

which the continuous time analog of equation (15).

Formally, let us define the continuous time analog of the (dual) dynamic discrete time problem (equation (9)), which we will refer to as $W^+(q_t, \lambda)$.

Definition 1. The dual continuous time problem is

$$W^+(q_t,\lambda) = \sup_{\{\sigma_s,\bar{\psi}_s,x_s,\tau\}} E_t[\beta^{(\tau^*-t)}\hat{u}(q_{\tau^*}) - \frac{1}{-\ln(\beta)}(1-\beta^{(\tau^*-t)})(\kappa-\lambda c^{\rho})] -$$
$$-\frac{\lambda}{\rho}E_t[\int_t^{\tau}\beta^{(s-t)}\{\frac{1}{2}tr[\sigma_s\sigma_s^T k(q_s)] + \bar{\psi}_s D^*(q_{s^-}+x_s||q_{s^-})\}^{\rho}ds],$$

subject to the evolution of beliefs,

$$dq_t = -\bar{\psi}_t x_t dt + x_t dJ_t + Diag(q_t)\sigma_t dB_t,$$

where dJ_t is a Poisson process with intensity $\bar{\psi}_t$ and dB_t is an $|X|$-dimensional Brownian motion, and $q_{t^-}+x_t \ll q_{t^-}$, and the constraint that, for all stopping times T measurable with respect filtration generated by q_s^* ,

$$E_t[\int_t^T \beta^{(s-t)}\{\frac{1}{2}tr[\sigma_s\sigma_s^T k(q_t)] + \bar{\psi}_s D^*(q_{s^-}+x||q_{s^-})\}ds] \le (\frac{\theta}{\lambda})^{\frac{1}{\rho-1}}E_t[\frac{1-\beta^{(T-t)}}{-\ln(\beta)}],$$

where θ is a positive constant.

Our most general convergence result proves that the discrete time dual problem converges $W(q_0,\lambda;\Delta)$ converges to this value function, $W^+(q_t,\lambda)$, and, for some value of λ^*, the convergence of the original problem $V(q_0;\Delta_n)$ to $W^+(q_t,\lambda^*)$.

Theorem 2. *Let Δ_m, $m\in\mathbb{N}$, denote a sequence such that $\lim_{m\to\infty}\Delta_m = 0$, let $V(q_0;\Delta_n)$ and $W(q_0,\lambda;\Delta_n)$ denote the discrete-time value functions defined in equations (8) and (9), and suppose that the cost function satisfies conditions 1-5. Then there exists a sub-sequence $n\in\mathbb{N}$ such that, for all λ if $\beta<1$ and all $\lambda\in(0,\kappa c^{-\rho})$ if $\beta=1$,*

$$\lim_{n\to\infty} W(q_0,\lambda;\Delta_n) = W^+(q_0,\lambda).$$

There exists a λ^* *such that*

$$\lim_{n\to\infty} V(q_0;\Delta_n) = W^+(q_0,\lambda^*).$$

Moreover, it is without loss of generality to suppose that the diffusion terms (σ_s) *of the optimal policy associated with* $W^+(q_0,\lambda)$ *(definition 1) are zero.*

Proof. See the appendix, section D.10. □

Theorem 2 demonstrates the convergence of the original and dual problems, and shows (as part of the proof) that it is without loss of generality to assume there is no diffusion component. The intuition for the latter result, which might at first seem surprising, is that it is possible to synthesize a "diffusion-like" process using the jump controls. Choosing very rapid (high $\bar{\psi}_s$) and very short (small x_t) jumps can mimic a diffusion process, becoming arbitrarily similar to one as the jumps grow more frequent and shorter. Moreover, the cost of such a strategy becomes identical, in this limit, to the cost of the diffusion, by the property that the divergence D^* is, up to second order, described by the matrix k (Corollary 2). Consequently, even if the DM were restricted in the problem defined in definition 1 to have no diffusion component ($\sigma_t = 0$), the supremum of a sequence of jump controls could replicate the outcome of a diffusion, if such a policy were optimal.

In the particular case where the DM has no exponential discounting ($\beta = 1$), we demonstrate that the amount of information acquired at each moment ($C(\sigma_t,\psi_t,x_t;q_t)$ above) is constant. This property comes from the fact that the cost of delay is constant. In the case with discounting ($\beta < 1$), the cost of delay depends in part on the current level of the value function, which generates variation and leads to a non-constant flow cost function. In the $\beta = 1$ case, using the fact that the quantity of information acquired is constant, the problem can be written in a simpler form. Note that result that $\sigma_s = 0$ without loss of generality still applies.

Corollary 3. *If* $\beta = 1$*, then for all* $\lambda \in (0, \kappa c^{-\rho})$*, the problem defining* $W^+(q_0, \lambda)$ *(definition 1) can be written as*

$$W^+(q_t, \lambda) = \sup_{\{\sigma_s, \bar{\psi}_s, x_s, \tau\}} E_t[\hat{u}(q_{\tau^*}) - \tau \frac{\rho}{\rho - 1}(\kappa - \lambda c^{\rho})]$$

subject to

$$\frac{1}{2} tr[\sigma_s \sigma_s^T k(q_t)] + \bar{\psi}_s D^*(q_{s^-} + x || q_{s^-}) \leq \chi(\lambda),$$

for all times $s \in [t, \tau)$ *and some constant* $\chi(\lambda) > 0$*. For* $\lim_{n\to\infty} V(q_0; \Delta_n) = W^+(q_t, \lambda^*)$*, this simplifies to*

$$W^+(q_t, \lambda^*) = \sup_{\{\sigma_s, \bar{\psi}_s, x_s, \tau\}} E_t[\hat{u}(q_{\tau^*}) - \tau \kappa]$$

subject to

$$\frac{1}{2} tr[\sigma_s \sigma_s^T k(q_t)] + \bar{\psi}_s D^*(q_{s^-} + x || q_{s^-}) \leq \rho^{\frac{1}{\rho}} c.$$

Proof. See the appendix, section D.11. □

We next describe conditions under which the optimal policies lead gradually evolving beliefs and conditions under which the optimal policies result in beliefs that jump immediately to stopping beliefs. We begin by describing conditions for gradual learning.

5.1 Gradual Learning

We begin by defining what we call a "preference for gradual learning." This condition describes the relative costs of learning via jumps in beliefs vs. continuously diffusing beliefs, which are governed by the the matrix-valued function $\bar{k}(\cdot)$ (equivalently, $k(\cdot)$) and the divergence $D^*(\cdot||\cdot)$ (see Corollary 2).

Definition 2. The cost function $C(p,q;S)$ exhibits a "*preference for gradual learning*" if the matrix-valued function $\bar{k}(q)$ and divergence D^* associated with it satisfy, for all $q, q' \in \mathscr{P}(X)$ with $q' \ll q$,

$$D^*(q'||q) - (q'-q)^T (\int_0^1 (1-s)\bar{k}(sq' + (1-s)q)ds)(q'-q) \geq 0. \tag{16}$$

This preference is "strict" if the inequality is strict, and is "strong" if, for some $\delta > 0$ and some $m > 0$,

$$D^*(q'||q) - (q'-q)^T (\int_0^1 (1-s)\bar{k}(sq' + (1-s)q)ds)(q'-q) > m||q'-q||_2^{2+\delta}, \tag{17}$$

where $||\cdot||_2$ denotes the Euclidean norm.

In the appendix, section §A, we present examples of divergences D^* that do and do not have a preference for gradual learning.

To clarify the meaning of this preference, consider the "chain rule" (Cover and Thomas (2012)) that characterizes the Kullback-Leibler divergence (and Bregman divergences more generally),

$$D_{KL}(q'||q) + \sum_{s \in S} \pi_s D_{KL}(q_s||q') = \sum_{s \in S} \pi_s D_{KL}(q_s||q)$$

for all probability distributions π_s such that $\sum_{s \in S} \pi_s q_s = q'$. The following lemma demonstrates that, for any divergence D^* for which this chain rule is a "less-than-or-equal" inequality, there is a preference for gradual learning.

Lemma 2. *If the divergence D^* satisfies, for all $\pi_s \in \mathscr{P}(S)$ and $q, q', \{q_s\}_{s \in S} \in \mathscr{P}(X)$ such that $\sum_{s \in S} \pi_s q_s = q'$ and $q' \ll q$,*

$$D^*(q'||q) + \sum_{s \in S} \pi_s D^*(q_s||q') \leq \sum_{s \in S} \pi_s D^*(q_s||q),$$

then the cost function $C(p,q;S)$ exhibits a preference for gradual learning.

Proof. See the appendix, section D.12. □

This chain-rule inequality has a simple interpretation– it is more costly to jump directly to the beliefs $\{q_s\}$ than to have beliefs travel first to q' and then on to $\{q_s\}$. It is straightforward to see why such an assumption leads directly to gradual learning, although it is worth noting that the preference for gradual learning is a weaker condition than this chain rule inequality. This chain rule inequality could also be called "super-additivity," as the (essentially[12]) opposite assumption of the "sub-additivity" assumption discussed by Zhong (2018). In the next subsection, building on the results of Zhong (2017) and Zhong (2018), we will show in our model that that opposite inequality leads to immediate decision-making.

We begin by stating our gradual learning result for the $\beta = 1$ case. With a preference for gradual learning, we show that the value function converges to a continuous time problem with only a diffusion control. If the preference is strict, then the limiting process for beliefs must be a diffusion. If not, there may be sequences of optimal policies in the discrete time models that converge to jump-diffusions, or even pure jump processes.

Theorem 3. *Under the assumptions of Theorem 2, if $\beta = 1$ and the cost function satisfies a preference for gradual learning, there exists a sub-sequence, indexed by n, such that*

$$\lim_{n\to\infty} W(q_0,\lambda^*;\Delta_n) = \lim_{n\to\infty} V(q_0;\Delta_n) = W^+(q_0,\lambda^*) = V(q_0),$$

where

$$V(q_0) = \sup_{\{\sigma_t\},\tau} E_0[\hat{u}(q_\tau) - \kappa\tau)]$$

[12] This inequality only applies to small amounts of information, as opposed to the global assumption used by Zhong (2018).

subject to

$$dq_t = Diag(q_t)\sigma_t dB_t$$

and

$$\frac{1}{2}tr[\sigma_t \sigma_t^T k(q_t)] \leq \rho^{\frac{1}{\rho}} c.$$

If the cost function exhibits a strict preference for gradual learning, every convergent subsequence of belief processes $q^*_{t,n}$ associated with optimal policies in the discrete-time model converges in law to a diffusion.

Proof. See the appendix, section D.13. □

The sequence problem $V(q_0)$ has a straightforward interpretation– the DM controls the diffusion coefficient of her beliefs subject to a maximum rate of information acquisition defined by the information cost matrix function $k(q_t)$. We solve this problem, and prove it is equivalent in a certain sense to a static rational inattention problem, in Hébert and Woodford (2017).

We next turn to the case with discounting ($\beta < 1$). A remarkable result by Zhong (2017) (theorem 5 of that paper) demonstrates that, in a model very much like our continuous time limit, specialized to the case of only two states and with no flow costs ($\lambda c^{\rho} = \kappa$), the DM will generically choose to have no diffusion to her beliefs, only jumps. This result can be understood in two parts. First, as discussed above, it is without loss of generality to write the DM's belief process as a pure jump process, even when the law of the supremum over such processes is equivalent to the law of a diffusion.

The second part of the result of Zhong (2017) can be thought of as a (generic) lower bound on the magnitude of the jumps. That is, not only is it without loss of generality to consider a pure jump process, but the optimal policy of the DM is in fact a jump process with non-infinitesimal jumps. What is remarkable about this result, from the perspective of our results for the $\beta = 1$ case, is that it applies even when the cost function exhibits a preference for gradual learning. Zhong (2017) also shows, in the particular case of indifference to gradual learning (equality in equation (16)), which is the setting for most of his results, that the beliefs jump all the way to stopping points, a result we will replicate below.

To understand how our $\beta = 1$ results are connected to the results of Zhong (2017), for the $\beta < 1$ case, we suppose that the cost function exhibits a "strong" preference for gradual learning, as defined in definition 2. Under this assumption, we prove an upper bound on the size of the jumps, as a function of β, and show that in the limit as $\beta \to 1$, the bound converges to zero. In other words, there is no discontinuity– with discounting and a strong preference for gradual learning, the "pure jump process" for beliefs will become increasingly like a diffusion as the discount rate converges to zero.

Theorem 4. *In problem defined in definition 1, if the cost function satisfies a strong preference for gradual learning (equation (17)), then the optimal jump sizes x_t^* satisfy*

$$||x_t^*||_2 \leq (\frac{-\bar{u}\ln(\beta)}{m})^{\delta^{-1}},$$

and $\lim_{\beta \to 1^-} ||x_t^*||_2 = 0$.

Proof. See the appendix, section D.14. □

5.2 Discrete Learning

Building on Zhong (2017) and Zhong (2018), to provide contrast to our results on discrete learning, we provide conditions under which, in the $\beta = 1$ case, the DM jumps immediately to stopping beliefs. In particular, suppose that the "chain rule" inequality described previously (Lemma 2) is a "greater-than-or-equal-to" inequality,

$$D^*(q'||q) + \sum_{s\in S} \pi_s D^*(q_s||q') \geq \sum_{s\in S} \pi_s D^*(q_s||q), \tag{18}$$

for all $\pi_s, \{q_s\}, q, q'$ as in Lemma 2. In this case, it its cheaper for the DM to jump to beliefs $\{q_s\}$ rather than visit the beliefs q'. Unsurprisingly, because this holds everywhere, it leads to optimal policies that stop immediately after jumping. We conjecture, but have not proven, that similar results hold in the $\beta < 1$ case, noting also that Zhong (2017) proves specialized versions of this conjecture.

Theorem 5. *In the problem described by Corollary 3 (that is, $\beta = 1$), if equation (18) holds for all $\pi_s \in \mathscr{P}(S)$ and $q, q', \{q_s\}_{s\in S} \in \mathscr{P}(X)$ such that $\sum_{s\in S} \pi_s q_s = q'$ and $q' \ll q$, then it is without loss of generality to suppose that, for all q_{t^-} in the continuation region and any x_t^* that characterizes an optimal policy , $V(q_{t^-} + x_t^*) = \hat{u}(q_{t^-} + x_t^*)$. That is, without loss of generality, all jumps enter the stopping region.*

Proof. See the appendix, section D.15. □

The statement of Theorem 5 states that is without loss of generality to assume that the DM stops immediately after a jump in beliefs. This result is necessary, as opposed to without loss of generality, if a "strict" version of equation (18) holds for all non-degenerate distributions over the posteriors $\{q_s\}$. In contrast, in the case of indifference (equation (18) holds with equality everywhere), both Theorem 3 and Theorem 5 hold. This happens if D^* is the KL divergence, or a Bregman divergence more generally. This observation implies that the two continuous time value functions must be identical, despite one being written as controlling a diffusion process and the other (without loss of generality) a pure jump process.

We also mention that there is no reason, in general, to expect either of the chain rule inequalities to hold over the entire simplex. That is, the cases we have an analyzed are no exhaustive, and there many be cost functions that lead to large jumps in certain regions of the parameter space and small jumps or diffusions in other regions.

6 Discussion and Conclusion

We have derived a continuous-time rational-inattention model as the limit of a discrete-time sequential evidence accumulation problem. In the limit of a very large number of successive signals, each of which is only minimally informative, only the local properties of the flow cost function matter. Using these properties, we have demonstrated (without discounting) cases in which beliefs converge to either a pure diffusion process or a pure jump process. With discounting, we have described a more general limiting problem, and shown that it may result in "diffusion-like" pure jump processes. In Hébert and Woodford (2017), we apply these models and explore alternatives to the standard mutual information cost function.

We have left unresolved the question of whether it is appropriate to use the discounting or no-discounting models, and whether a preference for gradual learning is a reasonable way to characterize the evolution of beliefs. It is worth noting that in many decision-making contexts, the time it takes to make a decision (seconds or minutes) is likely to call for a cumulative discount factor (β^{τ}) very close to one. As a result, the limiting case of $\beta = 1$ may provide a useful and tractable approximation. We also note that when our key divergence D^* is a Bregman divergence, both our diffusion and jump results apply. However, this does not imply that everything is the same about the two problems– in particular, the joint distribution of decision times and actions taken may differ between the two models. This joint distribution is the subject of a great deal of existing research (e.g. Fudenberg et al. (2015)) and results in this literature might help determine whether a preference for gradual learning is a good description of decision making processes.

The continuous time limit we have derived in this paper can serve as a foundation for analytically tractable models of rational inattention. One strength of our approach is its generality– we have imposed relatively minimal assumptions on the nature of information costs, and yet derived models that are specific and tractable enough to be tested in data. Exploring the properties of these models, and in particular whether jump processes or diffusions better characterize beliefs, and what cost functions should be employed, is the next step in this research agenda.

References

Syed Mumtaz Ali and Samuel D Silvey. A general class of coefficients of divergence of one distribution from another. *Journal of the Royal Statistical Society. Series B (Methodological)*, pages 131–142, 1966.

Shun-ichi Amari and Hiroshi Nagaoka. *Methods of information geometry*, volume 191. American Mathematical Soc., 2007.

Nihat Ay, Jürgen Jost, Hông Vân Lê, and Lorenz Schwachhöfer. Information geometry and sufficient statistics. *Probability Theory and Related Fields*, pages 1–38, 2014.

David Blackwell. Equivalent comparisons of experiments. *The annals of mathematical statistics*, 24(2):265–272, 1953.

Andrew Caplin and Mark Dean. Revealed preference, rational inattention, and costly information acquisition. *American Economic Review*, 105(7):2183–2203, 2015.

Andrew Caplin, Mark Dean, and John Leahy. Rationally inattentive behavior: Characterizing and generalizing Shannon entropy. *Unpublished manuscript*, July 2017.

Nikolai Nikolaevich Chentsov. *Statistical decision rules and optimal inference*. Number 53. American Mathematical Soc., 1982.

John A. Clithero. Improving out-of-sample predictions using response times and a model of the decision process. *Unpublished manuscript*, June 2016.

Thomas M Cover and Joy A Thomas. *Elements of information theory*. John Wiley & Sons, 2012.

Henrique De Oliveira, Tommaso Denti, Maximilian Mihm, and Kemal Ozbek. Rationally inattentive preferences and hidden information costs. *Theoretical Economics*, 12:621–624, 2017.

Ernst Fehr and Antonio Rangel. Neuroeconomic foundations of economic choice — recent advances. *Journal of Economic Perspectives*, 25(4):3–30, 2011.

Drew Fudenberg, Philipp Strack, and Tomasz Strzalecki. Stochastic choice and optimal sequential sampling. *Unpublished manuscript*, 2015.

Benjamin Hébert. Moral hazard and the optimality of debt. *Unpublished manuscript*, 2014.

Benjamin Hébert and Michael Woodford. Rational inattention and sequential information sampling. Technical report, National Bureau of Economic Research, 2017.

Emir Kamenica and Matthew Gentzkow. Bayesian persuasion. *American Economic Review*, 101(6):2590–2615, 2011.

Ian Krajbich, Bastiaan Oud, and Ernst Fehr. Benefits of neuroeconomics modeling: New policy interventions and predictors of preference. *American Economic Review*, 104(5): 501–506, 2014.

Filip Matêjka, Alisdair McKay, et al. Rational inattention to discrete choices: A new foundation for the multinomial logit model. *American Economic Review*, 105(1):272–98, 2015.

Stephen Morris and Philipp Strack. The Wald problem and the equivalence of sequential sampling and static information costs. *Unpublished manuscript*, June 2017.

Giuseppe Moscarini and Lones Smith. The optimal level of experimentation. *Econometrica*, 69(6):1629–1644, 2001.

Michael Shadlen and Daphna Shohamy. Decision making and sequential sampling from memory. *Neuron*, 90(5):927–939, 2016.

Christopher A Sims. Rational inattention and monetary economics. *Handbook of Monetary Economics*, 3:155–181, 2010.

Jakub Steiner, Colin Stewart, and Filip Matějka. Rational inattention dynamics: Inertia and delay in decision-making. *Econometrica*, 85(2):521–553, 2017.

Satohiro Tajima, Jan Drugowitsch, and Alexandre Pouget. Optimal policy for value-based decision-making. *Nature communications*, 7, 2016.

Michael Woodford. Inattentive valuation and reference-dependent choice. *Unpublished manuscript*, May 2012.

Michael Woodford. An optimizing neuroeconomic model of discrete choice. Technical report, National Bureau of Economic Research, February 2014.

Weijie Zhong. Optimal dynamic information acquisition. *Unpublished manuscript*, March 2017.

Weijie Zhong. Indirect information measure and dynamic learning. *arXiv preprint arXiv:1809.00697*, 2018.

A Cost Functions Satisfying Conditions 1-5

In this appendix section, we discuss various families of cost functions that satisfy our conditions.

We first show that any "posterior-separable" cost function (in the terminology of Caplin et al. (2017)) that is sufficiently convex satisfies conditions 1-5. This includes all "uniformly posterior-separable" cost functions, such as mutual information and the neighborhood-based cost functions we describe in Hébert and Woodford (2017).

Posterior-separable cost functions are defined as

$$C(p,q;S) = \sum_{s \in S} (e_s^T pq) D(q_s || q),$$

where $D(\cdot || \cdot)$ is a divergence.

Lemma 3. *In the posterior-separable family if cost functions, if the divergence $D(\cdot||\cdot)$ is twice differentiable and strongly convex in its first argument, and differentiable in its second argument, the cost function $C(p,q;S)$ satisfies Conditions 1-5.*

Proof. See the appendix, section D.5. □

These posterior-separable cost functions a popular choice, but by no means the only possibility. As an alternative, consider "state-separable" cost functions,

$$C(p,q;S) = \sum_{x \in X} (e_x^T q) D(pe_x || pq; S),$$

where $D(pe_x||pq;S)$ is a family of divergences defined on the signal alphabets S. Mutual information is both "posterior-separable" and "state-separable," but in there are many cost functions in one family but not the other.

We will assume that the family of divergences is constant with respect to the addition of signals that never occur. That is, if $S \subset S'$, and $e_s^T p' = e_s^T p$ for $s \in S$ and zero otherwise, then

$$D(pe_x||pq;S) = D(p'e_x||p'q;S').$$

We also assume a Blackwell-type inequality for these divergences,

$$D(\Pi r||\Pi r') \leq D(r||r')$$

for all r, r' and all garbling matrices Π. Note that this implies that D is invariant in the sense described in the text. Under these assumptions, and some regularity assumptions, we prove that this family also satisfies our conditions.

Lemma 4. *In the state-separable family of cost functions, if the divergences $D(\cdot||\cdot;S)$ are convex in their first argument and twice differentiable, and satisfies, for some $m > 0$ and all $r, r' \in \mathscr{P}(S)$,*

$$D(r'||r;S) \geq m(r'-r)^T g(r)(r'-r),$$

then the cost function $C(p,q;S)$ satisfies Conditions 1-5.

Proof. See the appendix, section D.6. □

Lastly, we note that if some family of cost functions $C(p,q;S)$ satisfies our conditions, so does a convex transformation of those cost functions, or a linear combination of two such families. Consequently, the two "separable" families described above can be used to generate a huge variety of non-separable cost functions.

Lemma 5. *If $C(p,q;S)$ is a family of cost functions satisfying Conditions 1-5, then so is $C_h(p,q;S) = h(C(p,q;S))$, where $h : \mathbb{R}_+ \to \mathbb{R}_+$ is a strongly convex, strictly-increasing function with $h(0) = 0$. Likewise, if $C_1(p,q;S)$ and $C_2(p,q;S)$ are families of cost functions* *satisfying Conditions 1-5, then so is* $C_\alpha(p,q;S) = \alpha C_1(p,q;S) + (1-\alpha)C_2(p,q;S)$*, for any* $\alpha \in (0,1)$.

Proof. The proofs are almost immediate, given the the assumptions. □

B Examples of Cost Functions Generating Diffusion Learning

Here, we provide an intuitive example of families of cost functions that generate preferences for gradual learning, and hence lead to diffusion processes. We provide the examples in the context of posterior-separable cost functions,

$$C_\alpha(p,q;S) = \sum_{s\in S}(e_s^T pq) D_\alpha(q_s||q),$$

where D_α is an "alpha-divergence" (see, e.g., Amari and Nagaoka (2007)),

$$D_\alpha(p||q) = \begin{cases} \frac{4}{1-\alpha^2}\sum_{x\in X}(e_x^T q)(\frac{1+\alpha}{2}\frac{e_x^T p}{e_x^T q} + \frac{1-\alpha}{2} - (\frac{e_x^T p}{e_x^T q})^{\frac{1+\alpha}{2}}) & \alpha \notin \{-1,1\} \\ \sum_{x\in X}(e_x^T p)\ln(\frac{e_x^T p}{e_x^T q}) & \alpha = 1 \\ -\sum_{x\in X}(e_x^T q)\ln(\frac{e_x^T p}{e_x^T q}) & \alpha = -1. \end{cases}$$

First, observe that for these cost functions, $D^* = D_\alpha$ and $k(q)$ is the Fisher information matrix. Consequently, the equation defining a "preference for gradual learning" is

$$D_\alpha(q'||q) \geq D_{KL}(q'||q).$$

Because the α divergences are increasing in α (this follows from Jensen's inequality), this inequality holds (strongly, in the sense of equation (17)) for all $\alpha > 1$, and does not hold for all $\alpha < 1$.

C Additional Lemmas

Our first lemma shows that the value function $W(q_t, \lambda; \Delta)$ is well-behaved:

Lemma 6. *If $\lambda \in (0, \kappa c^{-\rho})$ and $\beta = 1$, or if $\beta < 1$, for all $\Delta \leq 1$ the value function $W(q_t, \lambda; \Delta)$ is bounded above on $q_t \in \mathscr{P}(X)$ by a constant $\bar{W}$, below by zero, and is convex in q. Moreover, for all $\Delta \leq 1$,*

$$\kappa - \lambda c^{\rho} - \ln(\beta)\bar{W} > 0.$$

Proof. See the appendix, section D.7. □

Our next lemma shows that, because of the curvature (ρ) that we impose, the DM will choose, under any optimal policy, to gather only a small amount of information in each time period, as the length of each time period shrinks.

Lemma 7. *Let $n \in \mathbb{N}$ denote a sequence such that $\Delta_n \leq 1$ and $\lim_{n\to\infty} \Delta_n = 0$. Under the assumptions of Lemma 6, any associated sequence of optimal policies $p^*_{t,n}$ satisfies, for all elements of the sequence,*

$$C(p^*_{t,n}, q_{t,n}; S) \leq (\frac{\theta}{\lambda})^{\frac{1}{\rho-1}} \Delta_n,$$

where $\theta = \lambda(\rho \frac{\kappa - \lambda c^{\rho} - \ln(\beta)\bar{W}}{\lambda(\rho-1)})^{\frac{\rho-1}{\rho}}$ and $\bar{W}$ is the upper bound of Lemma 6.

Proof. See appendix, section D.8. □

Our next lemma discuss the convergence of an arbitrary sequence of stochastic processes for beliefs (denoted $q_{t,m}$) and of stopping times (denoted τ_m) to their continuous-time limits, under the assumption that the policies generating them satisfy the bound in

Lemma 7 and the bound on expected stopping times. This lemma applies to a sequence of optimal policies, but also to sequences of sub-optimal policies. The lemma describes the convergence of the beliefs process to a martingale, which is not necessarily a diffusion (it may have jumps, or even be a semi-martingale that is not a jump-diffusion).

Lemma 8. *Let Δ_m, $m \in \mathbb{N}$, denote a sequence such that $\lim_{m\to\infty} \Delta_m = 0$. Let $p_m(q)$ denote a sequence of Markov policies satisfying the bound in Lemma 7. Let $q_{t,m}$ denote the stochastic process for the DM's beliefs at time t, under such a policy, and let τ_m be a sequence of stopping policies such that $E_0[\tau_m] \leq \bar{\tau}$.*

There exists a sub-sequence $n \in \mathbb{N}$ and a probability space such that:

i) The beliefs $q_{t,n}$ and the stopping time τ_n converge almost surely to a martingale q_t and a stopping time τ.

ii) The martingale q_t can be represented in terms of its semi-martingale characteristics,

$$B_t = -\int_0^t \left(\int_{\mathbb{R}^{|X|}\setminus\{0\}} \psi_s(x) x dx\right) dA_s$$

$$C_t = \int_0^t Diag(q_{s^-})\sigma_s \sigma_s^T Diag(q_{s^-}) dA_s$$

$$\nu_t(x) = dA_t \psi_t(x),$$

where σ_s is an $|X| \times |X|$ matrix-valued predictable stochastic process, satisfying $q_{s^-}^T \sigma_s = \vec{0}$, ψ_s is a measure on $\mathbb{R}^{|X|}\setminus\{0\}$ such that $q_{s^-} + x \in \mathscr{P}(X)$ and $q_{s^-} + x \ll q_{s^-}$ for all x in the support of ψ_s, and dA_s is the increment of a weakly increasing process.

iii) For all stopping times T,

$$E_t[\int_t^T \beta^{s-t}\{\frac{1}{2}tr[\sigma_s\sigma_s^T k(q_{s^-})] + \int_{\mathbb{R}^{|X|}\setminus\{0\}} \psi_s(x) D^*(q_{s^-}+x||q_{s^-})dx\} dA_s] \leq (\frac{\theta}{\lambda})^{\frac{1}{\rho-1}} \frac{1-\beta^{T-t}}{-\ln(\beta)}.$$

iv) The limit of the cumulative information cost is bounded below,

$$\lim_{n\to\infty} E_0[\Delta_n^{1-\rho} \sum_{j=0}^{\tau_n \Delta_n^{-1}-1} \beta^{\Delta_n j} C(p_n(q_{\Delta_n j,n}), q_{\Delta_n j,n}; S)^\rho] \geq E_t[\int_0^\tau \beta^s \{\frac{1}{2}tr[\sigma_s\sigma_s^T k(q_{s^-})] + \int_{\mathbb{R}^{|X|}\setminus\{0\}} \psi_s(x) D^*(q_{s^-}+x||q_{s^-})dx\}^\rho (\frac{dA_s}{ds})^\rho ds].$$

Proof. See the appendix, section D.9. □

D Proofs

D.1 Proof of Lemma 1

Let p and p' be information structures with signal alphabet S. First, we will show that mixture feasibility and Blackwell monotonicity imply convexity. By mixture equivalence, letting p_M denote the mixture information structure and S_M the signal alphabet,

$$C(p_M, q; S_M) \leq \lambda C(p,q;S) + (1-\lambda)C(p',q;S).$$

Consider the garbling $\Pi : S \times \{1,2\} \to S$, which maps each $(s,i) \in S_M$ to $s \in S$. By Blackwell monotonicity,

$$C(p_M, q; S_M) \geq C(\Pi p_M, q; S).$$

By construction,

$$e_s^T \Pi p_M = \lambda e_s^T p + (1-\lambda) e_s^T p',$$

and the result follows.

Now we show the other direction: that convexity and Blackwell monotonicity imply mixture feasibility. Let p_1 and p_2 be information structures with signal alphabets S_1 and S_2. Because the cost function satisfies Blackwell monotonicity, it is invariant to Markov congruent embeddings. Define $S_M = (S_1 \cup S_2) \times \{1,2\}$. There exists an embedding $\Pi_1 : S_1 \to S_M$ such that, for some $s_M = (s,i) \in S_M$,

$$e_{s_M}^T \Pi_1 p_1 = \begin{cases} 0 & i = 2 \\ 0 & s \notin S_1 \\ e_s^T p_1 & otherwise \end{cases}.$$

Define an embedding Π_2 along similar lines, and note that these embeddings are left-invertible. It follows by invariance that

$$C(\Pi_1 p_1, q; S_M) = C(p_1, q; S_1),$$

and likewise that

$$C(\Pi_2 p_2, q; S_M) = C(p_2, q; S_2).$$

By convexity,

$$C(\lambda \Pi_1 p_1 + (1-\lambda)\Pi_2 p_2; q; S_M) \leq \lambda C(\Pi_1 p_1, q; S_M) + (1-\lambda) C(\Pi_2 p_2, q; S_M).$$

Observing that

$$\lambda \Pi_1 p_1 + (1-\lambda)\Pi_2 p_2 = p_M$$

proves the result.

D.2 Proof of Theorem 1

Parts 1 and 2 of the theorem follow from a Taylor expansion of the cost function. Using the lemmas and theorem of Chentsov (1982), cited in the text, we know that for any invariant cost function with continuous second derivatives,

$$C(p,q;S) = \frac{1}{2}\Delta \sum_{x'\in X}\sum_{x\in X}(e_x^T k(q) e_{x'}) \nu_{x'}^T g(r) \nu_x + o(\Delta).$$

The second claim follows by a similar argument.

We next demonstrate the claimed properties of $k(q)$. First, $k(q)$ is symmetric and continuous in q, by the symmetry of partial derivatives and the assumption of continuous second derivatives (Condition 4). Recall the assumption that

$$p_x = r + \Delta^{\frac{1}{2}} \nu_x + o(\Delta^{\frac{1}{2}}),$$

which implies that $\sum_{s\in S} e_s^T r = 1$ and $\sum_{s\in S} e_s^T \nu_x = 0$ for all $x \in X$. Consider an information structure for which $\nu_x = \phi e_x^T v$, where $v \in \mathbb{R}^{|X|}$ and $\phi \in \mathbb{R}^{|S|}$, with $\sum_{s\in S} e_s^T \phi = 0$. Suppose that both v and ϕ are not zero. For this information structure,

$$C(p,q;S) = \frac{1}{2}\Delta \bar{g} v^T k(q) v + o(\Delta),$$

where $\phi^T g(r)\phi = \bar{g} > 0$. Suppose the information structure is uninformative for all Δ. This would be the case if v is proportional to ι, and therefore

$$\iota^T k(q)\iota = 0$$

by Condition 1. Regardless of whether the information structure is informative, by Condition 1, we must have

$$v^T k(q) v \geq 0,$$

implying that $k(q)$ is positive semi-definite. If z and $-z$ are in the tangent space of the simplex at q, there exists an x, x' $e_x^T z \neq e_{x'}^T z$ with x, x' in the support of q. Using z in the place of v above, by Condition 1, we must have

$$z^T k(q) z > 0.$$

Suppose now that the cost function satisfies Condition 5. Let v be as above, non-zero, and not proportional to ι. We have

$$C(p, q; S) = \frac{1}{2} \Delta \bar{g} v^T k(q) v + o(\Delta),$$

and therefore for the B defined in Condition 5 there exists a Δ_B such that, for all $\Delta < \Delta_B$, $C(p, q; S) < B$. Therefore, we must have

$$C(\{p_x\}_{x \in X}, q) \geq \frac{m}{2} \sum_{s \in S} (e_s^T p q) || q_s - q ||_X^2.$$

By Bayes' rule, for any signal that is received with positive probability,

$$q_s - q = \frac{(D(q) - q q^T) p^T e_s}{q^T p^T e_s}.$$

By convention, $q_s = q$ for any s such that $e_s^T p q = 0$.

The support of q_s is always a subset of the support of q, and therefore (by the equivalence of norms),

$$C(\{p_x\}_{x\in X},q) \geq \frac{m_g}{2}\sum_{s\in S}(e_s^T pq)(q_s-q)^T D^+(q)(q_s-q)$$

for some constant $m_g > 0$.

For sufficiently small Δ, $e_s^T pq > 0$ if $e_s^T r_s > 0$, and therefore

$$C(\{p_x\}_{x\in X},q) \geq \frac{m}{2}\sum_{s\in S:e_s^T r>0}\frac{(e_s^T p(D(q)-qq^T)D^+(q)(D(q)-qq^T)p^T e_s)}{(e_s^T pq)},$$

or,

$$C(\{p_x\}_{x\in X},q) \geq \frac{m}{2}\Delta\sum_{s\in S:e_s^T r>0}(e_s^T\phi)^2\frac{v^T(D(q)-qq^T)D^+(q)(D(q)-qq^T)v}{(e_s^T r)}+o(\Delta).$$

Noting that

$$\sum_{s\in S:e_s^T pq>0}\frac{(e_s^T\phi)^2}{(e_s^T pq)} = \phi^T g(r)\phi = \bar{g},$$

and that

$$(D(q)-qq^T)D^+(q)(D(q)-qq^T) = g^+(q),$$

we have

$$C(\{p_x\}_{x\in X},q) \geq \frac{m_g}{2}\Delta\bar{g}v^T g^+(q)v+o(\Delta).$$

It follows that we must have

$$\frac{1}{2}v^T k(q)v \geq \frac{m_g}{2}v^T g^+(q)v$$

for all v.

D.3 Proof of Corollary 1

Under the stated assumptions,

$$p_x = r + \Delta^{\frac{1}{2}} \nu_x + o(\Delta^{\frac{1}{2}}).$$

By Bayes' rule, for any $s \in S$ such that $e_s^T pq > 0$,

$$q_s = \frac{D(q) p^T e_s}{q^T p^T e_s}.$$

It follows immediately that

$$\lim_{\Delta \to 0^+} q_s = D(q) \frac{r^T e_s}{r_s^T} = q.$$

Next,

$$\begin{aligned}
\Delta^{-\frac{1}{2}} (q_s - q) &= \Delta^{-\frac{1}{2}} \frac{(D(q) - qq^T) p^T e_s}{q^T p^T e_s} \\
&= D(q) \frac{\nu^T e_s - \iota q^T \nu^T e_s + o(1)}{q^T p^T e_s}.
\end{aligned}$$

For any s such that $q^T p^T e_s > 0$,

$$\lim_{\Delta \to 0^+} \Delta^{-\frac{1}{2}} (q_s - q) = D(q) \frac{\nu^T e_s - \iota q^T \nu^T e_s}{r^T e_s}.$$

By Theorem 1,

$$C(p, q; S) = \frac{1}{2} \Delta \sum_{x' \in X} \sum_{x \in X} (e_x^T k(q) e_{x'}) \nu_{x'}^T g(r) \nu_x + o(\Delta).$$

By the result that $\iota^T k(q) = 0$, we have

$$C(p,q;S) = \frac{1}{2}\Delta \sum_{x'\in X}\sum_{x\in X} e_x^T k(q) e_{x'} \cdot (\nu_{x'} - q\nu)^T g(r)(\nu_x - q\nu)$$
$$+ o(\Delta).$$

By the definition of the Fisher matrix, and the observation that $\iota^T \nu_x = 0$ for all $x \in X$,

$$(\nu_{x'} - q\nu)^T g(r)(\nu_x - q\nu) = \sum_{s\in S: e_s^T r>0} (e_s^T r)\frac{(\nu_{x'} - q\nu)^T}{(e_s^T r)} e_s e_s^T \frac{(\nu_x - q\nu)}{(e_s^T r)}.$$

Substituting in the result regarding the posterior,

$$C(p,q;S) = \frac{1}{2}\sum_{s\in S: e_s^T r>0} (e_s^T r)(q_s - q)^T D^+(q) k(q) D^+(q)(q_s - q) + o(\Delta),$$

which is the result.

D.4 Proof of Corollary 2

By directional differentiability and the continuity of the directional derivatives, there exists a function

$$f(\omega, r, q;S) = \lim_{\Delta\to 0^+} \frac{C(\bar{p}_\Delta + \Delta\omega, q;S) - C(\bar{p}_\Delta, q;S)}{\Delta}.$$

Observe that, if ωe_x is in the support of r for all x in the support of q, we must have $f(\omega, \bar{p}, q;S) = 0$, by the results of Theorem 1. Relatedly, if ω and ω' differ only with respect to the frequency of signals in the support of r for all x in the support of q, we must have

$$f(\omega, r, q;S) = f(\omega', r, q;S).$$

Assuming there are signals not in the support of $\bar{p}$, we can write $\omega = \omega_1 + \omega_2 + \ldots$, where each ω_i is a perturbation that contains only one signal not the support of $\bar{p}q$. Let $N \leq |S|$ denote the number of these perturbations. We can define

$$f_i(\omega_i, r, q; S) = \lim_{\Delta \to 0^+} \frac{C(p_{i-1} + \Delta\omega_i, q; S) - C(p_{i-1}, q; S)}{\Delta},$$

where $p_{i-1} = \bar{p}_\Delta + \Delta \sum_{j=1}^{i-1} \omega_i$. By the assumption of the continuity of the directional derivatives,

$$f_i(\omega_i, r, q; S) = f(\omega_i, r, q; S).$$

It follows that

$$f(\omega, r, q; S) = \sum_{i=1}^{N} f(\omega_i, r, q; S).$$

By invariance, the function $f(\omega_i, r, q; S)$ does not depend on r or S. By the argument above, it is only a function of $e_{s_i}\omega_i$, where $s_i \in S$ is the unique signal in ω_i with $e_{s_i}^T r = 0$. By Bayes' rule,

$$e_{s_i}\omega_i = (e_{s_i}\omega_i q) Diag(q)^+ q_{s_i},$$

where q_{s_i} is the posterior associated with signal s_i. By the homogeneity of the directional derivative, we can rewrite this as

$$f(\omega_i, r, q; S) = (e_{s_i}\omega_i q) F(q_{s_i}, q).$$

By the requirement that the cost of an uninformative signal structure is zero, and everything else is costly, we must have

$$F(q, q) = 0,$$

$$F(q', q) > 0$$

for all $q' \neq q$. Therefore, F is a divergence, which we write $D^*(q'||q)$. The finiteness of

$D^*(q'||q)$ is implied by the existence of the directional derivative. The approximation of the cost function follows from this result and Corollary 1.

By invariance, there exists a Markov congruent embedding that splits each signal in S into $M > 1$ distinct signals in S'. As M becomes arbitrarily large, the probability of each signal becomes small — and in particular, can be of order Δ. It follows for all $s \in S'$ such that $||q_s - q|| = O(\Delta^{\frac{1}{2}})$ (e.g. the signals described in Corollary 1), we must have

$$D^*(q_s||q) = \frac{1}{2}\Delta(q_s^T - q)\bar{k}(q)(q_s - q) + O(\Delta).$$

Moreover, by this argument, $D^*(q'||q)$ must be twice differentiable in q' evaluated at q.

D.5 Proof of Lemma 3

We will show that Conditions 1-5 are satisfied. Recall the definition:

$$C(p,q;S) = \sum_{s \in S}(e_s^T pq)D(q_s||q).$$

D.5.1 Condition 1

Condition 1 requires that if the information structure is uninformative, the cost is zero, and if it is not, the cost is weakly positive. If the signal is uninformative, for any signal received with positive probability,

$$q_s = q,$$

and by our convention that $q_s = q$ if $(e_s^T pq) = 0$, this also holds for zero-probability signals. By the definition of a divergence, $D(q||q) = 0$ for all q, and therefore the cost of an uninformative information structure is zero.

The cost is strictly positive by the definition of a divergence (being strictly positive if

$q_s \neq q$) and the fact that probabilities are must sum to one.

D.5.2 Condition 2

Mixture feasibility requires that

$$C(p_M, q; S_M) \leq \lambda C(p_1, q; S_1) + (1-\lambda) C(p_2, q; S_2).$$

By definition,

$$\begin{aligned} C(p_M, q; S) &= \sum_{s \in S} (e_s^T p_M q) D(q_{s,M} || q) \\ &= \lambda \sum_{s \in S_1} (e_s^T p_1 q) D(q_{s,1} || q) + (1-\lambda) \sum_{s \in S_2} (e_s^T p_2 q) D(q_{s,2} || q) \\ &= \lambda C(p_1, q; S_1) + (1-\lambda) C(p_2, q; S_2). \end{aligned}$$

verifying that the condition holds.

D.5.3 Condition 3

By Blackwell's theorem, for any Markov mapping $\Pi : S \to S'$, we require that

$$C(\Pi p, q; S') \leq C(p, q; S).$$

By definition,

$$e_{s'}^T \Pi p q = \sum_{s \in S} e_{s'}^T \Pi e_s (e_s^T \Pi q)$$

and by Bayes' rule,

$$D(q) p^T \Pi^T e_{s'} = (e_{s'}^T \Pi p q) q_{s'},$$

where $q_{s'}$ is the posterior associated with $s' \in S'$. Also by Bayes' rule,

$$D(q)p^T e_s = (e_s^T pq) q_s,$$

and therefore

$$\begin{aligned} q_{s'} &= \frac{\sum_{s\in S} D(q)p^T e_s (e_s^T \Pi^T e_{s'})}{(e_{s'}^T \Pi pq)} \\ &= \frac{\sum_{s\in S} q_s (e_s^T pq)(e_s^T \Pi^T e_{s'})}{\sum_{s\in S}(e_{s'}^T \Pi e_s e_s^T pq)}. \end{aligned}$$

It follows by the convexity of D in its first argument and Jensen's inequality that

$$(e_{s'}^T \Pi pq) D(q_{s'}||q) \leq \sum_{s\in S} (e_s^T pq)(e_s^T \Pi^T e_{s'}) D(q_s||q).$$

It immediately follows that

$$\sum_{s'\in S'} (e_{s'}^T \Pi pq) D(q_{s'}||q) \leq \sum_{s\in S} (e_s^T pq) D(q_s||q).$$

D.5.4 Condition 4

We begin by showing twice-differentiability with respect to perturbations that do not change the support of the signal structure. By the definition of the cost function and the twice-differentiability of D in its first argument, it is sufficient to show that $(e_s^T pq)$ and q_s are both twice-differentiable with respect to these perturbations, in the neighborhood of an uninformative information structure.

Suppose that

$$p(\varepsilon, \nu) = r\iota^T + \varepsilon\tau + \nu\omega,$$

where $r \in \mathscr{P}(S)$ and the support of τe_x is in the support of r, and likewise for ωe_x, for all

$x \in X$.

By Bayes' rule, for all $s \in S$ such that $e_s^T r > 0$,

$$q_s(\varepsilon,\nu) = \frac{D(q)p(\varepsilon,\nu)^T e_s}{q^T p(\varepsilon,\nu)^T e_s}.$$

Simplifying,

$$\begin{aligned} q_s(\varepsilon,\nu) = {} & q\frac{r^T e_s}{r^T e_s + \varepsilon q^T \tau^T e_s + \nu q^T \omega^T e_s} + \frac{\varepsilon D(q)\tau^T e_s}{r^T e_s + \varepsilon q^T \tau^T e_s + \nu q^T \omega^T e_s} \\ & + \frac{\nu D(q)\omega^T e_s}{r^T e_s + \varepsilon q^T \tau^T e_s + \nu q^T \omega^T e_s}. \end{aligned}$$

In the neighborhood around $\varepsilon = \nu = 0$, the denominator is strictly positive, and therefore

$$\frac{\partial}{\partial \nu} q_s(\varepsilon,\nu) = -q_s(\varepsilon,\nu)\frac{q^T \omega^T e_s}{r^T e_s + \varepsilon q^T \tau^T e_s + \nu q^T \omega^T e_s} + \frac{D(q)\omega^T e_s}{r^T e_s + \varepsilon q^T \tau^T e_s + \nu q^T \omega^T e_s}$$

and

$$\begin{aligned} \frac{\partial}{\partial \varepsilon}\frac{\partial}{\partial \nu} q_s(\varepsilon,\nu) = {} & q_s(\varepsilon,\nu)\frac{q^T \omega^T e_s}{r^T e_s + \varepsilon q^T \tau^T e_s + \nu q^T \omega^T e_s}\frac{q^T \tau^T e_s}{r^T e_s + \varepsilon q^T \tau^T e_s + \nu q^T \omega^T e_s} \\ & - \frac{q^T \omega^T e_s}{r^T e_s + \varepsilon q^T \tau^T e_s + \nu q^T \omega^T e_s}\frac{D(q)\tau^T e_s}{r^T e_s + \varepsilon q^T \tau^T e_s + \nu q^T \omega^T e_s} \\ & - q_s(\varepsilon,\nu)\frac{q^T \omega^T e_s}{r^T e_s + \varepsilon q^T \tau^T e_s + \nu q^T \omega^T e_s}\frac{q^T \tau^T e_s}{r^T e_s + \varepsilon q^T \tau^T e_s + \nu q^T \omega^T e_s} \\ & - \frac{D(q)\omega^T e_s}{r^T e_s + \varepsilon q^T \tau^T e_s + \nu q^T \omega^T e_s}\frac{q^T \tau^T e_s}{r^T e_s + \varepsilon q^T \tau^T e_s + \nu q^T \omega^T e_s}. \end{aligned}$$

For $s \in S$ such that $e_s^T r = 0$, $q_s(\varepsilon,\nu) = q$, and therefore $\frac{\partial}{\partial \varepsilon}\frac{\partial}{\partial \nu} q_s(\varepsilon,\nu) = 0$. Therefore, $\frac{\partial}{\partial \nu} q_s(\varepsilon,\nu)$ can be written as a quadratic form in $vec(\tau)$ and $vec(\omega)$. It follows that $q_s(\varepsilon,\nu)$, in the neighborhood of an uninformative information structure, is twice-differentiable in the directions that do not change the support of the distribution of signals. By construction,

$(e_s^T pq)$ is twice-differentiable.

Now consider a perturbation that changes the support of the signals,

$$p(\varepsilon) = r\iota^T + \varepsilon\tau + \varepsilon\omega,$$

where $e_s^T \omega = 0$ for all s such that $e_s^T r > 0$, and greater than or equal to zero otherwise, and the support of τe_x is in the support of r for all $x \in X$. We have

$$q_s(\varepsilon) = q\frac{r^T e_s}{r^T e_s + \varepsilon q^T \tau^T e_s + \varepsilon q^T \omega^T e_s} + \frac{\varepsilon D(q)\tau^T e_s}{r^T e_s + \varepsilon q^T \tau^T e_s + \varepsilon q^T \omega^T e_s}$$
$$+ \frac{\varepsilon D(q)\omega^T e_s}{r^T e_s + \varepsilon q^T \tau^T e_s + \varepsilon q^T \omega^T e_s}.$$

For s such that $e_s^T \omega > 0$,

$$q_s(\varepsilon) = \frac{D(q)\omega^T e_s}{q^T \omega^T e_s},$$

and hence does not depend on ε. We also have $(e_s^T pq) = \varepsilon q^T \omega^T e_s$ for such s. Directional differentiability, continuous in (ω, τ), follows immediately.

D.5.5 Condition 5

This condition requires that, for some $m > 0$ and $B > 0$, for all $C(p,q;S) < B$,

$$C(p,q;S) \geq \frac{m}{2}\sum_{s \in S}(e_s^T pq)||q_s - q||_X^2,$$

where $||\cdot||_X$ is an arbitrary norm on the tangent space of $\mathscr{P}(X)$. It follows immediately by the strong convexity of the divergence.

D.6 Proof of Lemma 4

We will show that Conditions 1-5 are satisfied. Recall the definition:

$$C(p,q;S) = \sum_{x\in X}(e_x^T q)D(pe_x||pq;S).$$

D.6.1 Condition 1

Condition 1 requires that if the information structure is uninformative, the cost is zero, and if it is not, the cost is strictly positive. If the signal is uninformative, $pe_x = pq$ for all $x \in X$, and the result holds by the definition of a divergence. The cost for informative signals is strictly positive by the definition of a divergence.

D.6.2 Condition 2

Mixture feasibility requires that

$$C(p_M,q;S_M) \leq \lambda C(p_1,q;S_1) + (1-\lambda)C(p_2,q;S_2).$$

This follows by the convexity of the divergence, the Blackwell condition, and Lemma 1.

D.6.3 Condition 3

The result follows immediately by the Blackwell assumption on the divergence.

D.6.4 Condition 4

Twice differentiability follows by assumption. Directional differentiability, with continuous directional derivatives, follows from convexity (for the existence of directional derivatives) and twice-differentiability in the interior (which ensures continuity), and the assumption of continuity in the limit (as the signal probability reaches zero, and the signal alphabet changes).

D.6.5 Condition 5

This condition requires that, for some $m > 0$ and $B > 0$, for all $C(p,q;S) < B$,

$$C(p,q;S) \geq \frac{m}{2} \sum_{s \in S} (e_s^T pq) ||q_s - q||_X^2,$$

where $|| \cdot ||_X$ is an arbitrary norm on the tangent space of $\mathscr{P}(X)$.

By assumption,

$$D(r'||r) \geq m(r' - r)^T g(r)(r' - r),$$

where $g(r)$ is the Fisher information matrix.

Consequently,

$$\sum_{x \in X} (e_x^T q) D(pe_x || pq; S) \geq m \sum_{x \in X} \sum_{s \in S} \frac{(e_x^T q)}{e_s^T pq} (e_x^T - q^T) p^T e_s e_s^T p(q - e_x),$$

which by Bayes' rule is

$$\sum_{x \in X} (e_x^T q) D(pe_x || pq; S) \geq m \sum_{s \in S} (e_s^T pq)(q_s^T - q^T) g(q)(q_s - q).$$

Therefore, by $g(q) \succeq I$,

$$\sum_{x \in X} (e_x^T q) D(pe_x || pq; S) \geq m \sum_{s \in S} (e_s^T pq) ||q_s - q||_2^2.$$

D.7 Proof of Lemma 6

Write the value function in sequence-problem form, for the $\beta < 1$ case:

$$W(q_0,\lambda;\Delta) = \max_{\{p_{\Delta j}\},\tau} E_0[\beta^{\tau}\hat{u}(q_{\tau}) - \kappa\Delta\frac{1-\beta^{\tau}}{1-\beta^{\Delta}})] -$$

$$\lambda E_0[\Delta^{1-\rho}\sum_{j=0}^{\tau\Delta^{-1}}\beta^{\Delta j}\{\frac{1}{\rho}C(\{p_{\Delta j,x}\}_{x\in X}, q_{\Delta j}(\cdot))^{\rho} - \Delta^{\rho}c^{\rho}\}],$$

Define

$$\bar{u} = \max_{a\in A, x\in X} u(a,x).$$

By the weak positivity of the cost function $C(\cdot)$, it follows that

$$W(q_0,\lambda;\Delta) \leq \bar{u} + \Delta E_0[\frac{1-\beta^{\tau}}{1-\beta^{\Delta}}](\lambda c^{\rho} - \kappa).$$

If $\lambda \in [0, \kappa c^{-\rho}]$, the value function is bounded above by $\bar{u}$. If $\lambda > \kappa c^{-\rho}$,

$$W(q_0,\lambda;\Delta) \leq \bar{u} + \frac{\Delta}{1-\beta^{\Delta}}(\lambda c^{\rho} - \kappa),$$

and

$$1-\beta^{\Delta} > \frac{-\Delta\ln(\beta)}{1-\Delta\ln(\beta)},$$

implying

$$\frac{\Delta}{1-\beta^{\Delta}} < \frac{1-\ln(\beta)}{-\ln(\beta)}$$

for all $\Delta \leq 1$. Therefore,

$$\bar{W} = \bar{u} + \frac{1-\ln(\beta)}{-\ln(\beta)}\max\{\lambda c^{\rho} - \kappa, 0\}.$$

It follows immediately that

$$\kappa - \lambda c^{\rho} - \ln(\beta)\bar{W} = \begin{cases} -\ln(\beta)\bar{u} + \kappa - \lambda c^{\rho} & \kappa \geq \lambda c^{\rho} \\ -\ln(\beta)\bar{u} & \kappa < \lambda c^{\rho}, \end{cases}$$

and therefore

$$\kappa - \lambda c^{\rho} - \ln(\beta)\bar{W} > 0.$$

For the $\beta = 1$ case, by the assumption that $\lambda c^{\rho} \leq \kappa$, $W(q_0, \lambda; \Delta) \leq \bar{u} = \bar{W}$, and the result holds immediately.

There is a smallest possible decision utility which is strictly positive, and because stopping now and deciding is always feasible,

$$W(q_0, \lambda; \Delta) \geq 0.$$

We can define the "state-specific" value function, $W(q_t, \lambda; \Delta, x)$, which is the value function conditional on the true state being x. The state-specific value function has a recursive representation, in the region in which the DM continues to gather information:

$$\begin{aligned} W(q_t, \lambda; \Delta, x) &= -\kappa\Delta + \lambda\Delta^{1-\rho}(\Delta^{\rho} c^{\rho} - \frac{1}{\rho} C(\cdot)^{\rho}) + \\ &\quad \beta^{\Delta} \sum_{s \in S:\, e_s^T p_t e_x > 0} (e_s^T p_t^* e_x) W(q^*_{t+\Delta,s}, \lambda; \Delta, x). \end{aligned}$$

In this equation, we take the optimal information structure as given. Note that, by construction, wherever the DM does not choose to stop, the expected value of the state-specific value functions is equal to the value function.

$$\sum_{x\in X} q_{t,x} W(q_t,\lambda;\Delta,x) = W(q_t,\lambda;\Delta).$$

By the optimality of the policies, we have

$$W(q_t,\lambda;\Delta) \geq \sum_{x\in X} q_{t,x} W(q',\lambda;\Delta,x),$$

for any q' in $\mathscr{P}(X)$. Suppose not; then the DM could simply adopt the information structure associated with beliefs q' and achieve higher utility, contradicting the optimality of the policy.

The convexity of the value function follows from the observation that

$$\begin{aligned} W(\alpha q+(1-\alpha)q',\lambda;\Delta) &= \alpha \sum_{x\in X} q_x W(\alpha q+(1-\alpha)q',\lambda;\Delta,x)+ \\ &\quad (1-\alpha)\sum_{x\in X} q'_x W(\alpha q+(1-\alpha)q',\lambda;\Delta,x), \end{aligned}$$

and using the inequality above,

$$W(\alpha q+(1-\alpha)q',\lambda;\Delta) \leq \alpha W(q,\lambda;\Delta)+(1-\alpha)W(q',\lambda;\Delta).$$

D.8 Proof of Lemma 7

Consider an alternative policy that mixes (in the sense of Condition 2) the optimal signal structure and an uninformative signal, with probabilities $1-a$ and a, respectively. We must have

$$-\beta^{\Delta_n}\sum_{s\in S}(e_s^T r_{t,n}^*)(W(q_{t,n,s}^*,\lambda;\Delta_n)-W(q_{t,n},\lambda;\Delta_n))-\lambda\Delta_n^{1-\rho}C(p_{t,n}^*,q_{t,n})^{\rho-1}\frac{\partial C(p_{t,n}(a),q_{t,n})}{\partial a}|_{a=0^+}\leq 0,$$

which is the first-order condition at the optimal policy in the direction of adding a little bit of the uninformative signal (decreasing a). By the convexity of $C(\cdot)$ and Condition 1,

$$C(p^*_{t,n}, q_{t,n}) + \frac{\partial C(p_{t,n}(a), q_{t,n})}{\partial a}|_{a=0^+} \leq 0,$$

and therefore we must have

$$\beta^{\Delta_n} \sum_{s \in S} (e_s^T r^*_{t,n})(W(q^*_{t,n,s}, \lambda; \Delta_n) - W(q_{t,n}, \lambda; \Delta_n)) \geq \lambda \Delta_n^{1-\rho} C(p^*_{t,n}, q_{t,n})^\rho.$$

Applying the Bellman equation in the continuation region,

$$(1-\beta^{\Delta_n})W(q_{t,n}, \lambda; \Delta_n) + (\kappa - \lambda c^\rho)\Delta_n + \frac{\lambda}{\rho}\Delta_n^{1-\rho} C(p^*_{t,n}, q_{t,n})^\rho \geq \lambda \Delta_n^{1-\rho} C(p^*_{t,n}, q_{t,n})^\rho.$$

Therefore,

$$\lambda(1 - \frac{1}{\rho})\Delta_n^{-\rho} C(p^*_{t,n}, q_{t,n})^\rho \leq (\kappa - \lambda c^\rho) + \frac{(1-\beta^{\Delta_n})}{\Delta_n} W(q_{t,n}, \lambda; \Delta_n).$$

If $\beta = 1$, then

$$C(p^*_{t,n}, q_{t,n}) \leq \Delta_n (\frac{\theta}{\lambda})^{\frac{1}{\rho-1}},$$

for the constant $\theta = \lambda(\rho \frac{\kappa - \lambda c^\rho}{\lambda(\rho-1)})^{\frac{\rho-1}{\rho}} > 0$.

If $\beta < 1$, note that

$$\frac{(1-\beta^{\Delta_n})}{\Delta_n} < -\ln(\beta).$$

Let $\bar{W}$ denote the upper bound on $W(q_{t,n}, \lambda; \Delta_n)$, which exists by Lemma 6. We have

$$C(p^*_{t,n}, q_{t,n}) \leq \Delta_n (\frac{\theta}{\lambda})^{\frac{1}{\rho-1}},$$

where

$$\theta = \lambda(\rho\frac{\kappa - \lambda c^{\rho} - \ln(\beta)\bar{W}}{\lambda(\rho-1)})^{\frac{\rho-1}{\rho}}.$$

The constant θ is positive by (6). Note that this generalizes the formula of the $\beta = 1$ case.

D.9 Proof of Lemma 8

We begin by discussing the convergence of stopping times. Let $\bar{W}$ denote the upper bound on $W(q_{t,n}, \lambda; \Delta_n)$, which exists by Lemma 6. Suppose that under an optimal policy,

$$\lim_{T\to\infty} Pr\{\tau_n < T\} = 1 - \alpha < 1.$$

The value function at time T must be bounded above by

$$W(q_T, \lambda; \Delta) \le (1-\alpha)\bar{W},$$

as the payoff conditional on never stopping is negative. Now consider an alternative policy that follows the optimal policy until time T, and then stops. The difference in the initial value functions is bounded above by the possibility of making the best possible decision under the optimal policy vs. the worst possible decision under the alternative policy, with utility $\underline{u} > 0$:

$$(1-\alpha - Pr\{\tau_n < T\})\beta^T \bar{W} \ge (1 - Pr\{\tau_n < T\})\beta^T \underline{u}.$$

This inequality cannot hold in the limit as $T \to \infty$. Therefore, by the positivity of τ_n, the laws of τ_n are tight, and therefore there exists a sub-sequence that converges in measure. Pass to this sub-sequence (which we will also index by n), and let τ denote the limit of this sub-sequence.

The beliefs $q_{t,n}$ are a family of $\mathbb{R}^{|X|}$-valued stochastic processes, with $q_{t,n} \in \mathscr{P}(X)$ for all $t \in [0,\infty)$ and $n \in \mathbb{N}$. Construct them as RCLL processes by assuming that $q_{\Delta_n j+\varepsilon,n} = q_{\Delta_n j,n}$ for all m, $\varepsilon \in [0,\Delta_n)$, and $j \in \mathbb{N}$. We next establish that the laws of $q_{t,n}$ are tight. By Condition 5 and Lemma 7,

$$\frac{m}{2}\sum_{s\in S}(e_s^T p_n(q_{t,n})q_{t,n})||q_{s,n}(q_{t,n}) - q_{t,n}||_2^2 \leq C(p_n(q_{t,n}), q_{t,n}; S) \leq \Delta_n(\frac{\theta}{\lambda})^{\frac{1}{\rho-1}},$$

where $q_{s,n}(q)$ is defined by $p_n(q)$ and Bayes' rule. It follows that, for any $\varepsilon > 0$, there exists an N_ε such that, for all $n > N_\varepsilon$,

$$P(||q_{t+\Delta_n,n} - q_{t,n}|| > \varepsilon) \leq K_\varepsilon \Delta_n,$$

for the constant $K_\varepsilon = 2m^{-1}\varepsilon^{-2}\theta^{\frac{1}{\rho-1}}$. By Theorem 3.21 in chapter 6 of Jacod and Shiryaev (2013), and the boundedness of $q_{t,n}$, it follows that the laws of $q_{t,n}$ are tight. By Prokhorov's theorem (Theorem 3.9 in chapter 6 of Jacod and Shiryaev (2013)), it follows that there exists a convergent sub-sequence. Pass to this sub-sequence, and let q_t denote the limiting stochastic process. By Proposition 1.1 in chapter 9 of Jacod and Shiryaev (2013), q_t is a martingale with respect to the filtration it generates. By Skorohod's representation theorem, there exists a probability space and random variables (which we will also denote with $q_{t,n}$ and q_t) such the convergence is almost sure. We refer to this probability space and these random variables in what follows.

Note that, by Bayes' rule, if $e_x^T q_{t,n} = 0$ for some $x \in X$ and time t, then $e_x^T q_{s,n} = 0$ for all $s > t$. By Proposition 2.9 and Corollary 2.38 in chapter 2 of Jacod and Shiryaev (2013), we can write the "good" version of the martingale with characteristics

$$B = -\int_0^t (\int_{\mathbb{R}^{|X|}\setminus\{0\}} \psi_s(x)xdx)dA_s$$

$$C = \int_0^t \Sigma_s dA_s$$

$$\nu = dA_s \psi_s(x).$$

Because beliefs remain in the simplex, $\psi_s(x)$ has support only on x such that $q_s + x \in \mathscr{P}(X)$. Relatedly, $\iota^T \Sigma_s = 0$. By the property mentioned above, $q_s + x \ll q_s$, and Σ_s can be decomposed as $\Sigma_s = D(q_{s^-})\sigma_s \sigma_s^T D(q_{s^-})$.

By the convexity of the cost function and Corollary 2,

$$C(p_n(q_{t,n}), q_{t,n}; S) \geq \sum_{s \in S} (e_s^T p_n(q_{t,n}) q_{t,n}) D^*(q_{s,n}(q_{t,n}) || q_{t,n}).$$

Defining the process, for arbitrary stopping time T,

$$D_{s,n} = \lim_{\varepsilon \to 0^+} D^*(q_{s^- + \varepsilon, n} || q_{s^-, n})$$

and

$$D_{t,T,n} = E_t[\int_t^T \beta^{\Delta_n \lfloor \Delta_n^{-1}(s-t) \rfloor} D_{s,n} ds] \leq (\frac{\theta}{\lambda})^{\frac{1}{\rho-1}} \Delta_n E_t[\sum_{j=0}^{\lfloor \Delta_n^{-1}(s-t) \rfloor} \beta^{j\Delta_n}],$$

we have by Ito's lemma, almost sure convergence, and the dominated convergence theorem,

$$D_{t,T} = \lim_{n \to \infty} D_{t,T,n} = E_t[\int_t^T \beta^{s-t} \{\frac{1}{2} tr[\sigma_s \sigma_s^T k(q_{s^-})] + \int_{\mathbb{R}^{|X|} \setminus \{0\}} \psi_s(x) D^*(q_{s^-} + x || q_{s^-}) dx\} dA_s].$$

Hence, for all such stopping times T,

$$E_t[\int_t^T \beta^{s-t} \{\frac{1}{2} tr[\sigma_s \sigma_s^T k(q_{s^-})] + \int_{\mathbb{R}^{|X|} \setminus \{0\}} \psi_s(x) D^*(q_{s^-} + x || q_{s^-}) dx\} dA_s] \leq (\frac{\theta}{\lambda})^{\frac{1}{\rho-1}} \frac{1 - \beta^{T-t}}{-\ln(\beta)}.$$

Note also by this argument that

$$\lim_{n\to\infty} E_0[\Delta_n^{1-\rho} \sum_{j=0}^{\tau_n\Delta_n^{-1}-1} \beta^{\Delta_n j} C(p_n(q_{\Delta_n j,n}), q_{\Delta_n j,n}; S)^\rho]$$

$$= \lim_{n\to\infty} E_0[\int_0^{\tau_n} \beta^{\Delta_n \lfloor \Delta_n^{-1} t\rfloor} \Delta_n^{-\rho} C(p_n(q_{t,n}), q_{t,n}; S)^\rho dt]$$

$$\geq E_t[\int_0^{\tau} \beta^s \{\frac{1}{2} tr[\sigma_s \sigma_s^T k(q_{s^-})] + \int_{\mathbb{R}^{|X|}\setminus\{0\}} \psi_s(x) D^*(q_{s^-}+x||q_{s^-})dx\}^\rho (\frac{dA_s}{ds})^\rho ds]$$

D.10 Proof of Theorem 2

Assume that $\lambda \in (0, \kappa c^{-\rho})$ if $\beta = 1$, $\lambda > 0$ if $\beta < 1$. Under this assumption, lemmas 6, 7, and 8 apply.

Let m index a sequence of Markov optimal policies, $p_m^*(q)$, and of stopping times τ_m^*. Let $q_{t,n}^*$ denote the associated process for beliefs. By the uniform boundedness and convexity of the family of value functions $W(q, \lambda; \Delta_m)$, a uniformly convergent sub-sequence exists. Rockafellar (1970) Theorem 10.9 demonstrates that a uniformly convergent sub-sequence exists on the relative interior of the simplex, and Rockafellar (1970) Theorem 10.3 demonstrates that there is a unique extension to a convex and continuous function on the boundary of the simplex.

Pass to this sub-sequence, which (for simplicity) we also index by m, and let $W(q, \lambda)$ denote its limit. By Lemmas 6 and 7, the sequence of optimal policies and stopping time satisfies the conditions of Lemma 8. It follows by that lemma that

$$\lim_{n\to\infty} E_0[\int_0^{\tau_n^*} \beta^{\Delta_n \lfloor \Delta_n^{-1} t\rfloor} \Delta_n^{1-\rho} C(p_n^*(q_{t,n}^*), q_{t,n}^*; S)^\rho dt]$$

$$\geq E_t[\int_0^{\tau} \beta^s \{\frac{1}{2} tr[\sigma_s^* \sigma_s^{*T} k(q_{s^-}^*)] + \int_{\mathbb{R}^{|X|}\setminus\{0\}} \psi_s^*(x) D^*(q_{s^-}+x||q_{s^-})dx\}^\rho\}^\rho (\frac{dA_s^*}{ds})^\rho ds],$$

where q_s^* is the limiting stochastic process and $\sigma_s^*, \psi_s^*, dA_s^*$ are associated with the characteristics of the martingale q_s^*.

We also have, by weak convergence,

$$\lim_{n\to\infty} E_0[\beta^{\tau_n}\hat{u}(q_{\tau_n^*,n}) - \Delta_n \frac{1-\beta^{\tau_n^*}}{1-\beta^{\Delta_n}}(\kappa - \lambda c^\rho))] = E_0[\beta^{\tau^*}\hat{u}(q_{\tau^*}) - \frac{1}{-\ln(\beta)}(1-\beta^{\tau^*})(\kappa-\lambda c^\rho))].$$

Recall also the bound, for any stopping time T measurable with respect filtration generated by q_s^* ,

$$E_t[\int_t^T \beta^s\{\frac{1}{2}tr[\sigma_s^*\sigma_s^{*T}k(q_{s^-})] + \int_{\mathbb{R}^{|X|}\setminus\{0\}} \psi_s^*(x)D^*(q_{s^-}^* + x||q_{s^-}^*)dx\}dA_s^*] \leq (\frac{\theta}{\lambda})^{\frac{1}{\rho-1}} E_t[\frac{1-\beta^{(T-t)}}{-\ln(\beta)}].$$

It follows that

$$W(q,\lambda) \leq W^+(q,\lambda)$$

for all $q \in \mathscr{P}(X)$, where

$$\begin{aligned} W^+(q_t,\lambda) = \sup_{\{\sigma_s,\psi_s,dA_s,\tau\}} & E_t[\beta^{(\tau^*-t)}\hat{u}(q_{\tau^*}) - \frac{1}{-\ln(\beta)}(1-\beta^{(\tau^*-t)})(\kappa-\lambda c^\rho)] - \\ & -\frac{\lambda}{\rho}E_t[\int_t^\tau \beta^{(s-t)}\{\frac{1}{2}tr[\sigma_s\sigma_s^T k(q_s)] + \int_{\mathbb{R}^{|X|}\setminus\{0\}} \psi_s(x)D^*(q_{s^-}+x||q_{s^-})dx\}^\rho (\frac{dA_s}{ds})^\rho ds], \end{aligned}$$

subject to the constraints, for all stopping times T measurable with respect filtration generated by q_s^* ,

$$E_t[\int_t^T \beta^{(s-t)}\{\frac{1}{2}tr[\sigma_s\sigma_s^T k(q_t)] + \int_{\mathbb{R}^{|X|}\setminus\{0\}} \psi_s(x)D^*(q_{s^-}+x||q_{s^-})dx\}dA_s] \leq (\frac{\theta}{\lambda})^{\frac{1}{\rho-1}} E_t[\frac{1-\beta^{(T-t)}}{-\ln(\beta)}]$$

and the evolution of beliefs as implied by the characteristics derived from σ_s, ψ_s, dA_s. Observe, by the arguments in the proof of Lemma 6, that $W^+(q,\lambda)$ is bounded, and convex in q, and satisfies $\kappa - \lambda c^\rho - \ln(\beta)W^+(q_t,\lambda) > 0$.

Also note that, for W^+, it is without loss of generality to set $dA_s = ds$. Scaling dA_s up and scaling $\sigma_s \sigma_s^T$ and ψ_s down, or vice versa, does not change the constraint, and setting $dA_s = 0$ is clearly sub-optimal by the result that $\kappa - \lambda c^\rho - \ln(\beta) W^+(q_t, \lambda) > 0$. Note also that there is a version of the optimal policies which satisfy the constraint everywhere:

$$\frac{1}{2} tr[\sigma_s \sigma_s^T k(q_{s^-})] + \int_{\mathbb{R}^{|X|} \setminus \{0\}} \psi_s(x) D^*(q_{s^-} + x || q_{s^-}) dx \leq (\frac{\theta}{\lambda})^{\frac{1}{\rho - 1}}.$$

Next, observe that for any x such that $\psi_s^+(x) > 0$, the first-order condition requires that

$$\lambda \{ \frac{1}{2} tr[\sigma_s^+ \sigma_s^{+T} k(q_{s^-})] + \int_{\mathbb{R}^{|X|} \setminus \{0\}} \psi_s^+(x'') D^*(q_{s^-} + x'' || q_{s^-}) dx'' \}^{\rho - 1} D^*(q_{s^-} + x || q_{s^-}) =$$
$$W^+(q_{s^-} + x, \lambda) - W^+(q_{s^-}, \lambda) - x^T \cdot W_q^+(q_{s^-}, \lambda), \tag{19}$$

where the differentiability of W^+ in the continuation region follows from the envelope theorem. Similarly, increase $\sigma_s \sigma_s^T$ by a quantity $\varepsilon x x^T$ results in a first order condition, anywhere W^+ is twice-differentiable, of

$$\lambda \{ \frac{1}{2} tr[\sigma_s^+ \sigma_s^{+T} k(q_{s^-})] + \int_{\mathbb{R}^{|X|} \setminus \{0\}} \psi_s^+(x'') D^*(q_{s^-} + x'' || q_{s^-}) dx'' \}^{\rho - 1} \frac{1}{2} tr[x x^T k(q_{s^-})] \geq \frac{1}{2} x^T W_{qq}^+(q_{s^-}) x,$$

with equality if the diffusion terms are non-zero in that direction. Note that the bound that the optimal policies satisfies imply that $W_{qq}^+(q_{s^-})$, interpreted in a distributional sense, is finite.

Combining these two first order conditions, consider a perturbation that decreases $\sigma_s \sigma_s^T$ by $\varepsilon x x^T$ and increases $\psi_s(\nu x)$ and $\psi_s(-\nu x)$ by $\frac{1}{2} \frac{\varepsilon}{\nu^2}$. The first-order conditions for this perturbation is

$$\lambda\{\frac{1}{2}tr[\sigma_s^{+}\sigma_s^{+T}k(q_{s^-})]+\int_{\mathbb{R}^{|X|}\setminus\{0\}}\psi_s^{+}(x'')D^*(q_{s^-}+x''||q_{s^-})dx''\}^{\rho-1}\times$$
$$\{\frac{1}{2\nu^2}D^*(q_{s^-}+\nu x||q_{s^-})+\frac{1}{2\nu^2}D^*(q_{s^-}-\nu x||q_{s^-})-\frac{1}{2}tr[xx^Tk(q_{s^-})]\}=$$
$$\frac{1}{2\nu^2}(W^{+}(q_{s^-}+\nu x,\lambda)+W^{+}(q_{s^-}-\nu x,\lambda)-2W^{+}(q_{s^-},\lambda))-\frac{1}{2}x^TW_{qq}^{+}(q_{s^-})x.$$

In the limit as $\nu\to 0^+$, this equation is always satisfied, and therefore it is without loss of generality to suppose that the diffusion term is zero.

Lastly, if there exists an x,x' with $\psi_s^{+}(x)>0$ and $\psi_s^{+}(x')>0$, and alternative policy that sets $\tilde{\psi}_s^{+}(x)=\psi_s^{+}(x)+\frac{D^*(q_{s^-}+x'||q_{s^-})}{D^*(q_{s^-}+x||q_{s^-})}\psi_s^{+}(x')$ and $\tilde{\psi}_s^{+}(x')=0$ generates the same cost, and changes utility by

$$\frac{D^*(q_{s^-}+x'||q_{s^-})}{D^*(q_{s^-}+x||q_{s^-})}(W^{+}(q_{s^-}+x,\lambda)-W^{+}(q_{s^-},\lambda)-x^T\cdot W_q^{+}(q_{s^-},\lambda))\psi_s^{+}(x')-(W^{+}(q_{s^-}+x',\lambda)-W^{+}(q_{s^-},\lambda)-$$

It follows that it is without loss of generality to assume that $\psi_s^{+}(x)>0$ for at most one value of x. Recalling that the optimal policies are Markov, let $\sigma^{+}(q_s)$ denote the optimal policy for the diffusion, let $\bar{\psi}^{+}(q)$ denote the optimal jump intensity, and let $x^{+}(q)$ denote the Markov optimal jump direction. Any semi-martingale with these characteristics generates a law that is identical to the jump-diffusion process described in the lemma.

Noting that $W^{+}(q,\lambda)\geq W(q,\lambda)$, it follows that if there exists a sequence of policies that converge to the stochastic process q_t^{+}, characterized by σ^{+}, $\bar{\psi}^{+}$, x^{+}, and whose cumulative information costs $\Delta_n^{-1}C(\cdot)$ converge to the total information costs in definition 1, then such a sequence of policies achieves, in the limit, at least as much utility as any other sequence of policies. It would then be the case that there must be sequence of optimal policies that converges a.s. (as in Lemma 8) to some optimal policy of W^{+} (not necessarily the policies that generate q_t^{+}). Note also by the result above that it is without loss of generality to suppose $\sigma^+ = 0$.

We can rewrite our controls in terms of the jump destination, $q^+(q_s) = q_s + x^+(q_s)$. To construct such a sequence of convergent policies, consider the "constant control" described in chapter 13.2 of Kushner and Dupuis (2013) ("constant controls", in this context, being a constant $q^+, \bar{\psi}_t^+$ pair over the interval $[t, t+\Delta_n)$, switching to $\psi_t = 0$ after the first jump). By theorem 2.3 of that chapter, there exists a sequence of constant controls that converge (weakly) to the optimal policies of W^+. Moreover, these controls result, of the intervals $[t, t+\Delta_n)$, in a two-point distribution, with support on $q^+(q_t)$ and $q_t - \bar{\psi}_t^+(q^+(q_t) - q_t)\Delta_n$ for the left limit of the process at time $t + \Delta_n$.

Define the constant

$$\theta^+ = \frac{E_t[\int_t^\tau \beta^{(s-t)} \bar{\psi}_s^+(q_{s^-}) D^*(q_{s^-} + x^+(q_{s^-}) || q_{s^-}) ds]}{E_t[\frac{1-\beta^{(\tau-t)}}{-\ln(\beta)}]}.$$

Now consider a modification of these constant control policies, which scale the intensity $\bar{\psi}_t^+$ by the quantity $\alpha_n(q_t)$, so that, for the modified policy,

$$\Delta_n C(\cdot) = \theta^+.$$

By the first-order condition with respect to $\bar{\psi}_t$, and the Bellman equation,

$$\kappa - \lambda c^\rho - \ln(\beta) W^+(q_t, \lambda) = \lambda(1 - \frac{1}{\rho})(\bar{\psi}_t^+(q_{t^-}) D^*(q_{t^-} + x^+(q_{t^-}) || q_{t^-}))^\rho.$$

By the convexity of $C(\cdot)$,

$$C(\cdot) \geq \alpha_n(q_t) \bar{\psi}_t^+(q_{t^-}) D^*(q_{t^-} + x^+(q_{t^-}) || q_{t^-}).$$

Observe that the lower bound on the value function that

$$\kappa - \lambda c^{\rho} - \ln(\beta) W^{+}(q_t, \lambda) > 0$$

for all q_t. It follows that

$$\alpha_n(q_t) \in [0, \frac{\theta^{+}}{\kappa - \lambda c^{\rho} - \ln(\beta)\underline{\mathrm{W}}}],$$

where $\underline{\mathrm{W}} = \min_{q \in \mathscr{P}(X)} W^{+}(q, \lambda)$, and that

$$\lim_{n \to \infty} \alpha_n(q_t) = 1.$$

Therefore, by this uniform bound, the modified policies converge weakly to the same limit as the constant control policies, and hence to an optimal policy of W^{+}. Moreover, by construction, the costs converge, and hence the dual value function W^{+} is achievable.

We next demonstrate equality of the primal and dual. The associated Bellman equation for the dual value function W^{+}, in the continuation region, is

$$\begin{aligned} -\ln(\beta) W^{+}(q_s, \lambda) = \max_{\sigma_s, \psi_s} & E[dW^{+}(q_s, \lambda)] - (\kappa - \lambda c^{\rho}) ds \\ & - \frac{\lambda}{\rho} \{ \frac{1}{2} tr[\sigma_s \sigma_s^T k(q_s)] + \int_{\mathbb{R}^{|X|} \setminus \{0\}} \psi_s(x) D^{*}(q_{s^-} + x || q_{s^-}) dx \}^{\rho}. \end{aligned}$$

Consider a perturbation which scales $\sigma_s^{+} \sigma_s^{+T}$ and ψ_s^{+} be some constant $(1+\varepsilon)$. Note that such a perturbation would also scale $E[dW^{+}]$ by $(1+\varepsilon)$, and that at least one of σ_s^{+} and ψ_s^{+} must be non-zero by the assumption that $-\ln(\beta) W^{+}(q_s, \lambda) + \kappa - \lambda c^{\rho} > 0$. The first order condition for this perturbation is

$$-\ln(\beta)W^{+}(q_s,\lambda)+(\kappa-\lambda c^{\rho})+\frac{\lambda}{\rho}\{\frac{1}{2}tr[\sigma_s^{+}\sigma_s^{+T}k(q_{s^-})]+\int_{\mathbb{R}^{|X|}\setminus\{0\}}\psi_s^{+}(x)D^{*}(q_{s^-}+x||q_{s^-})dx\}^{\rho}=$$
$$\lambda\{\frac{1}{2}tr[\sigma_s^{+}\sigma_s^{+T}k(q_{s^-})]+\int_{\mathbb{R}^{|X|}\setminus\{0\}}\psi_s^{+}(x)D^{*}(q_{s^-}+x||q_{s^-})dx\}^{\rho},$$

which must hold at the optimal policies for this problem. We can rewrite the Bellman equation as

$$-\ln(\beta)W^{+}(q_s,\lambda)ds+(\kappa-\lambda c^{\rho})ds=$$
$$E[dW^{+}(q_s,\lambda)]-\frac{\lambda}{\rho}(\rho\frac{rW^{+}(q_s,\lambda)+(\kappa-\lambda c^{\rho})}{\lambda(\rho-1)})ds,$$

or

$$(-\ln(\beta)W^{+}(q_s,\lambda)+(\kappa-\lambda c^{\rho}))\frac{\rho}{\rho-1}=E[dW^{+}(q_s,\lambda)].$$

Solving this equation,

$$W^{+}(q_s,\lambda)=E_s[\beta^{\frac{\rho}{\rho-1}(\tau^{*}-s)}\hat{u}(q_{\tau^{*}})-\frac{\rho}{\rho-1}(\kappa-\lambda c^{\rho})\int_s^{\tau^{*}}\beta^{\frac{\rho}{\rho-1}(l-s)}dl].$$

Define λ^{*} by

$$E_s[\beta^{\frac{\rho}{\rho-1}(\tau^{*}-s)}\hat{u}(q_{\tau^{*}})-\frac{\rho}{\rho-1}(\kappa-\lambda^{*}c^{\rho})\int_s^{\tau^{*}}\beta^{\frac{\rho}{\rho-1}(l-s)}dl]=$$
$$E_0[\beta^{(\tau^{*}-s)}\hat{u}(q_{\tau^{*}})-\kappa\int_s^{\tau^{*}}\beta^{(l-s)}dl].$$

We can rewrite this as

$$(\frac{1}{\rho-1}\kappa - \frac{\rho}{\rho-1}\lambda^* c^{\rho})E_0[\int_s^{\tau^*} \beta^{\frac{\rho}{\rho-1}(l-s)} dl] =$$
$$E_0[\beta^{\frac{\rho}{\rho-1}(\tau^*-s)}\hat{u}(q_{\tau^*}) - \kappa \int_s^{\tau^*} \beta^{\frac{\rho}{\rho-1}(l-s)} dl] -$$
$$E_0[\beta^{(\tau^*-s)}\hat{u}(q_{\tau^*}) - \kappa \int_0^{\tau^*} \beta^{(l-s)} dl].$$

The right-hand side is weakly negative, and zero if $\beta = 1$. Consequently, $\lambda^* > 0$, and $\lambda^* = \frac{\kappa}{\rho c^{\rho}} < \kappa c^{-\rho}$ if $\beta = 1$.

Consider a convergent sub-sequence of $V(q_0;\Delta_n)$ (which exists by the uniform boundedness and convexity of the problem), and denote its limit $V(q_0)$ (again, we will index this sequence by n). By the standard duality inequalities, for all λ,

$$V(q_0;\Delta_n) \leq W(q_0,\lambda;\Delta_n),$$

for all n, and therefore

$$V(q_0) \leq W^+(q_0,\lambda^*).$$

Consider the value function $\tilde{V}(q_0)$, which is the value function under the feasible optimal policies for $W^+(q_0,\lambda^*)$. It follows that $\tilde{V}(q_0) = W(q_0,\lambda^*)$, and $\tilde{V}(q_0) \leq V(q_0)$, and therefore $V(q_0) = W(q_0,\lambda^*)$.

Note that every convergent sub-sequence of $V(q_0;\Delta_n)$ converges to the same function. It follows that

$$V(q_0) = \lim_{\Delta \to 0^+} V(q_0;\Delta).$$
$$= E_0[\beta^{\tau^*}\hat{u}(q_{\tau^*}) - \kappa \int_0^{\tau^*} \beta^l dl].$$

By the definition of λ^* and the Bellman equation,

$$E_0[\int_0^{\tau^*} \beta^s \frac{1}{\rho}\{\frac{1}{2}tr[\sigma_s^* \sigma_s^{*T} k(q_{s^-})] + \int_{\mathbb{R}^{|X|}\setminus\{0\}} \psi_s^*(x) D^*(q_{s^-}+x||q_{s^-})dx\}^\rho ds] \leq c^\rho E_0[\int_0^{\tau^*} \beta^l dl],$$

as required. It follows that the value function is the maximized over all policies satisfying the above constraint (which is the limiting constraint, by the dominated convergence theorem), concluding the proof.

D.11 Proof of Corollary 3

With $\beta = 1$, the associated Bellman equation for the candidate value function W^+, in the continuation region, is

$$0 = \max_{\sigma_s,\psi_s} E[dW^+(q_s,\lambda)] - (\kappa - \lambda c^\rho)ds - \frac{\lambda}{\rho}\{\frac{1}{2}tr[\sigma_s \sigma_s^T k(q_s)] + \int_{\mathbb{R}^{|X|}\setminus\{0\}} \psi_s(x) D^*(q_{s^-}+x||q_{s^-})dx\}^\rho.$$

Let σ_s^+ and ψ_s^+ denote optimal policies for this problem. Consider a perturbation which scales $\sigma_s^+ \sigma_s^{+T}$ and ψ_s^+ be some constant $(1+\varepsilon)$. Note that such a perturbation would also scale $E[dW^+]$ by $(1+\varepsilon)$, and that at least one of σ_s^+ and ψ_s^+ must be non-zero by the assumption that $\kappa - \lambda c^\rho > 0$. The first order condition for this perturbation is

$$(\kappa - \lambda c^\rho) + \frac{\lambda}{\rho}\{\frac{1}{2}tr[\sigma_s^+ \sigma_s^{+T} k(q_{s^-})] + \int_{\mathbb{R}^{|X|}\setminus\{0\}} \psi_s^+(x) D^*(q_{s^-}+x||q_{s^-})dx\}^\rho = \lambda\{\frac{1}{2}tr[\sigma_s^+ \sigma_s^{+T} k(q_{s^-})] + \int_{\mathbb{R}^{|X|}\setminus\{0\}} \psi_s^+(x) D^*(q_{s^-}+x||q_{s^-})dx\}^\rho,$$

which must hold at the optimal policies for this problem. It follows by the definition of θ in the $\beta = 1$ case (see the proof of Lemma 7),

$$\theta = \lambda(\rho\frac{\kappa - \lambda c^{\rho}}{\lambda(\rho - 1)})^{\frac{\rho-1}{\rho}},$$

that the constraint

$$\frac{1}{2}tr[\sigma_s\sigma_s^T k(q_{s^-})] + \bar{\psi}_s D^*(q_{s^-} + x_s||q_{s^-}) \leq (\frac{\theta}{\lambda})^{\frac{1}{\rho-1}}$$

binds with equality everywhere. Consequently, the Bellman equation can be rewritten as

$$\max_{\sigma_s,\psi_s} E[dW^+(q_s,\lambda)] - (\kappa - \lambda c^{\rho})ds - \frac{\lambda}{\rho}(\frac{\theta}{\lambda})^{\frac{\rho}{\rho-1}}ds$$

subject to

$$\frac{1}{2}tr[\sigma_s\sigma_s^T k(q_{s^-})] + \bar{\psi}_s D^*(q_{s^-} + x_s||q_{s^-}) \leq (\frac{\theta}{\lambda})^{\frac{1}{\rho-1}}.$$

Defining

$$\chi(\lambda) = (\rho\frac{\kappa - \lambda c^{\rho}}{\lambda(\rho - 1)})^{\frac{1}{\rho}}$$

and observing that

$$\begin{aligned}\kappa - \lambda c^{\rho} + \frac{\lambda}{\rho}(\frac{\theta}{\lambda})^{\frac{\rho}{\rho-1}} &= \kappa - \lambda c^{\rho} + \frac{\kappa - \lambda c^{\rho}}{\rho - 1}\\ &= (\kappa - \lambda c^{\rho})\frac{\rho}{\rho - 1},\end{aligned}$$

the result follows, noting from the proof of Theorem 2 that $\lambda^* = \frac{1}{\rho}\kappa c^{-\rho}$ when $\beta = 1$, and therefore

$$\chi(\lambda^*) = c\rho^{\frac{1}{\rho}}.$$

D.12 Proof of Lemma 2

Consider a two-signal alphabet, $s \in \{s_1, s_2\}$, with $\pi_{s_1} = \pi_{s_2}$, and $q_{s_1} = (1+\varepsilon)q' - \varepsilon q$ and $q_{s_2} = (1-\varepsilon)q' + \varepsilon q$. Applying the "chain rule" inequality,

$$
\begin{aligned}
D^*(q'||q) + \frac{1}{2}D^*(q' + \varepsilon(q'-q)||q') + \frac{1}{2}D^*(q' - \varepsilon(q'-q)||q') \\
\leq \frac{1}{2}D^*(q' + \varepsilon(q'-q)||q) + \frac{1}{2}D^*(q' - \varepsilon(q'-q)||q).
\end{aligned}
$$

Dividing by ε^2 and taking the limit as $\varepsilon \to 0^+$,

$$
(q'-q)^T \cdot \bar{k}(q') \cdot (q'-q) \leq \frac{d^2}{d\varepsilon^2} D^*(q' + \varepsilon(q'-q)||q)|_{\varepsilon=0}.
$$

Since this must hold for all $q' \ll q$, it holds for $q' = q + t(q''-q)$, with some arbitrary $q'' \ll q$ and $t \in [0,1]$. Therefore,

$$
\frac{d^2}{dt^2} D^*(q + t(q''-q)||q)|_{\varepsilon=0} \geq (q''-q)^T \cdot \bar{k}(q + t(q''-q)) \cdot (q''-q).
$$

Integrating,

$$
D^*(q''||q) \geq \int_0^1 \int_0^s (q''-q)^T \cdot \bar{k}(q + t(q''-q)) \cdot (q''-q) dt ds,
$$

which is

$$
D^*(q''||q) \geq \int_0^1 (1-t)(q''-q)^T \cdot \bar{k}(q + t(q''-q)) \cdot (q''-q) dt.
$$

D.13 Proof of Theorem 3

Conjecture that $\lambda \in (0, \kappa c^{-\rho})$. Under this conjecture, lemmas 6, 7, 8, and 2 apply.

Consider a possibly sub-optimal policy which sets $\psi_s(x) = 0$ and satisfies the constraint. The above FOC applies, and therefore we must have

$$tr[\tilde{\sigma}_s \tilde{\sigma}_s^T (D(q_{s^-}) W_{qq}^+(q_{s^-}, \lambda) D(q_{s^-}) - \theta k(q_{s^-}))] \leq 0,$$

where W_{qq}^+ is understood in a distributional sense. It follows that, for all feasible x,

$$W^+(q_{s^-} + x, \lambda) - W^+(q_{s^-}, \lambda) - x^T W_q^+(q_{s^-}, \lambda; -x) \leq \int_0^1 \int_0^s x^T \bar{k}(q_{s^-} + lx) x dl ds.$$

By our assumption of gradual learning (definition 2), this implies that

$$W^+(q_{s^-} + x, \lambda) - W^+(q_{s^-}, \lambda) - x^T W_q^+(q_{s^-}, \lambda; -x) \leq \theta D^*(q_{s^-} + x || q_{s^-}).$$

Hence, it is without loss of generality to assume that $\psi_s^+(x) = 0$ for all x. Note that, if there is a strict preference for gradual learning, the above inequality is strict for all non-zero x. As a result, in this case $\psi_s^+(x) = 0$ for all x. Note also that our control problem involves direct control of the diffusion coefficients, and hence satisfies the standard requirements for the existence and uniqueness of a strong solution to the resulting SDE (Pham (2009) sections 1.3 and 3.2).

D.14 Proof of Theorem 4

The associated Bellman equation, in the continuation region, is

$$\begin{aligned} 0 = \max_{\sigma_s, \psi_s} & E[dW^+(q_s, \lambda)] + \ln(\beta) W^+(q_s, \lambda) ds - (\kappa - \lambda c^\rho) ds \\ & - \frac{\lambda}{\rho} \{ \frac{1}{2} tr[\sigma_s \sigma_s^T k(q_s)] + \int_{\mathbb{R}^{|X|} \setminus \{0\}} \psi_s(x) D^*(q_{s^-} + x || q_{s^-}) dx \}^\rho ds. \end{aligned} \tag{20}$$

Let σ_s^+ and ψ_s^+ denote optimal policies for this problem. Suppose that the constraint does not bind, and consider a perturbation which scales $\sigma_s^+\sigma_s^{+T}$ and ψ_s^+ be some constant $(1+\varepsilon)$. Note that such a perturbation would also scale $E[dW^+]$ by $(1+\varepsilon)$, and that at least one of σ_s^+ and ψ_s^+ must be non-zero by the assumption that $-\ln(\beta)W^+(q_s,\lambda)+\kappa-\lambda c^\rho >$ 0. The first order condition for this perturbation is

$$-\ln(\beta)W^+(q_s,\lambda)+(\kappa-\lambda c^\rho)+\frac{\lambda}{\rho}\{\frac{1}{2}tr[\sigma_s^+\sigma_s^{+T}k(q_{s^-})]+\int_{\mathbb{R}^{|X|}\setminus\{0\}}\psi_s^+(x)D^*(q_{s^-}+x||q_{s^-})dx\}^\rho =$$
$$\lambda\{\frac{1}{2}tr[\sigma_s^+\sigma_s^{+T}k(q_{s^-})]+\int_{\mathbb{R}^{|X|}\setminus\{0\}}\psi_s^+(x)D^*(q_{s^-}+x||q_{s^-})dx\}^\rho,$$

which must hold at the optimal policies for this problem.

Consider a sub-optimal policy which sets $\psi_s(x)=0$ and satisfies

$$\frac{1}{2}tr[\tilde{\sigma}_s\tilde{\sigma}_s^T k(q_{s^-})]=(\frac{\tilde{\theta}}{\lambda})^{\frac{1}{\rho-1}},$$

where

$$(\frac{\tilde{\theta}}{\lambda})^{\frac{\rho}{\rho-1}}\frac{\lambda}{\rho}+(\kappa-\lambda c^\rho)=\tilde{\theta}.$$

For such a policy, the Bellman equation must be an inequality,

$$tr[\tilde{\sigma}_s\tilde{\sigma}_s^T(D(q_{s^-})W_{qq}^+(q_{s^-},\lambda)D(q_{s^-})]\leq -\ln(\beta)W^+(q_s,\lambda)ds+\tilde{\theta}\frac{1}{2}tr[\tilde{\sigma}_s\tilde{\sigma}_s^T k(q_{s^-})].$$

where W_{qq}^+ is understood in a distributional sense. It follows that, for all feasible x,

$$W^+(q_{s^-}+x,\lambda)-W^+(q_{s^-},\lambda)-x^T W_q^+(q_{s^-},\lambda;-x)\leq \int_0^1\int_0^s x^T\bar{k}(q_{s^-}+lx)xdlds$$
$$-\ln(\beta)\int_0^1\int_0^t W^+(q_{s^-}+lx)dldt.$$

By the strong preference for gradual learning and the upper bound on utility,

$$W^{+}(q_{s^{-}}+x,\lambda)-W^{+}(q_{s^{-}},\lambda)-x^{T}W_{q}^{+}(q_{s^{-}},\lambda;-x)-\tilde{\theta}D^{*}(q_{s^{-}}+x||q_{s^{-}})\leq-\ln(\beta)\bar{u}\frac{1}{2}||x||_{2}^{2}-m||x||_{2}^{2+\delta}. \tag{21}$$

If a jump is optimal, we must have (by the first-order condition)

$$W^{+}(q_{s^{-}}+x,\lambda)-W^{+}(q_{s^{-}},\lambda)-x^{T}W_{q}^{+}(q_{s^{-}},\lambda;-x)=\lambda\{\psi_{s}^{+}(x)D^{*}(q_{s^{-}}+x||q_{s^{-}})\}^{\rho-1}D^{*}(q_{s^{-}}+x||q_{s^{-}}),$$

and

$$\lambda(1-\rho^{-1})\{\psi_{s}^{+}(x)D^{*}(q_{s^{-}}+x||q_{s^{-}})\}^{\rho}=-\ln(\beta)W^{+}(q_{s},\lambda)+(\kappa-\lambda c^{\rho}).$$

Therefore, by the positivity of the value function,

$$\lambda\{\psi_{s}^{+}(x)D^{*}(q_{s^{-}}+x||q_{s^{-}})\}^{\rho-1}\geq\lambda(\frac{(\kappa-\lambda c^{\rho})}{\lambda(1-\rho^{-1})})^{\frac{\rho-1}{\rho}},$$

which implies

$$W^{+}(q_{s^{-}}+x,\lambda)-W^{+}(q_{s^{-}},\lambda)-x^{T}W_{q}^{+}(q_{s^{-}},\lambda;-x)\geq\tilde{\theta}D^{*}(q_{s^{-}}+x||q_{s^{-}}).$$

Using equation (21) above,

$$m||x||_{2}^{\delta}\leq-\ln(\beta)\bar{u},$$

which is

$$||x||_{2}\leq(-\frac{\bar{u}\ln(\beta)}{m})^{\delta^{-1}}.$$

D.15 Proof of Theorem 5

The problem described in Corollary 3, using the fact that it is without loss of generality to assume a pure jump process, is

$$W^{+}(q_t,\lambda) = \sup_{\{\bar{\psi}_s,x_s,\tau\}} E_t[\hat{u}(q_{\tau^*}) - \tau\frac{\rho}{\rho-1}(\kappa-\lambda c^{\rho})]$$

subject to

$$\bar{\psi}_s D^*(q_{s^-}+x||q_{s^-}) \leq \chi(\lambda).$$

Suppose that the theorem is false– that for some q_{t^-} and x_t^*, $q_{t^-}+x_t^* = q'$ is in the continuation region. The first-order condition (see equation (19)) can be written as

$$W^{+}(q_{t^-}+x_t^*,\lambda) - W^{+}(q_{t^-},\lambda) - x_t^{*T}\cdot W_q^{+}(q_{t^-},\lambda) = \theta D^*(q'||q).$$

If q' is in the continuation region, there must be some $q'' \ll q'$ such that

$$W^{+}(q'',\lambda) - W^{+}(q',\lambda) - (q''-q')\cdot W_q^{+}(q',\lambda) = \theta D^*(q''||q').$$

Adding these two equations together and re-arranging,

$$W^{+}(q'',\lambda) - W^{+}(q_{t^-},\lambda) - (q''-q_{t^-})W_q^{+}(q_{t^-},\lambda) =$$
$$\theta D^*(q'||q) + \theta D^*(q''||q') + (q''-q')\cdot(W_q^{+}(q',\lambda) - W_q^{+}(q_{t^-},\lambda)).$$

By the fact that x_t^* is optimal, it must be the case that

$$(q''-q')^T\cdot W_q^{+}(q',\lambda) - (q''-q')^T\cdot W_q^{+}(q_{t^-},\lambda) = \theta(q''-q')^T\cdot D_q^*(q'||q).$$

Now consider the chain rule inequality, supposing that $s \in \{s_1, s_2\}$ with $\pi_{s_1} = \varepsilon$, $\pi_{s_2} = 1-\varepsilon$, $q_{s_1} = q''$, and $q_{s_2} = q' - \frac{\varepsilon}{1-\varepsilon}(q''-q')$,

$$D^*(q'||q) + \varepsilon D^*(q''||q') + (1-\varepsilon)D^*(q' - \frac{\varepsilon}{1-\varepsilon}(q''-q')||q') \geq \varepsilon D^*(q''||q) + (1-\varepsilon)D^*(q' - \frac{\varepsilon}{1-\varepsilon}(q''-q')||q).$$

Dividing by ε and taking limits,

$$D_q^*(q'||q)(q''-q') + D^*(q''||q') \geq D^*(q''||q) - D^*(q'||q).$$

Consequently,

$$W^+(q'',\lambda) - W^+(q',\lambda) - (q''-q) \cdot W_q^+(q,\lambda) \geq \theta D^*(q''||q),$$

meaning that it is without loss of generality to suppose that beliefs jump directly to q'' instead of to q'. Therefore, it is without loss of generality to suppose beliefs jump directly to the stopping region.

E Appendix References

Shun-ichi Amari and Hiroshi Nagaoka. *Methods of information geometry*, volume 191. American Mathematical Soc., 2007.

Andrew Caplin, Mark Dean, and John Leahy. Rationally inattentive behavior: Characterizing and generalizing Shannon entropy. *Unpublished manuscript*, July 2017.

Nikolai Nikolaevich Chentsov. *Statistical decision rules and optimal inference*. Number 53. American Mathematical Soc., 1982.

Benjamin Hébert and Michael Woodford. Rational inattention and sequential information sampling. Technical report, National Bureau of Economic Research, 2017.

Jean Jacod and Albert Shiryaev. *Limit theorems for stochastic processes*, volume 288. Springer Science & Business Media, 2013.

Harold Kushner and Paul G Dupuis. *Numerical methods for stochastic control problems in continuous time*, volume 24. Springer Science & Business Media, 2013.

Huyên Pham. *Continuous-time stochastic control and optimization with financial applications*, volume 61. Springer Science & Business Media, 2009.

R Tyrrell Rockafellar. Convex analysis. 1970.