Optimal and Myopic Information Acquisition*

Annie Liang† Xiaosheng Mu‡ Vasilis Syrgkanis§

First version: March 11, 2017
This version: April 15, 2019

Abstract

A decision-maker (DM) faces an intertemporal decision problem, where his payoff depends on actions taken across time as well as on an unknown Gaussian state. The DM can learn about the state from different (correlated) information sources, and allocates a budget of samples across these sources each period. A simple information acquisition strategy for the DM is to neglect dynamic considerations and allocate samples myopically. How inefficient is this strategy relative to the optimal information acquisition strategy? We show that if the budget of samples is sufficiently large then there is no inefficiency: myopic information acquisition is exactly optimal.

JEL codes: C44, D81, D83
Keywords: Information Acquisition, Correlation, Endogenous Attention, Myopic Choice, Robustness, Value of Information

*This paper was previously circulated under the title of "Dynamic Information Acquisition from Multiple Sources." We thank Yash Deshpande, Johannes Hörner, Michihiro Kandori, Elliot Lipnowski, George Mailath, Juuso Toikka, Anton Tsoy, and Muhamet Yildiz for helpful comments. We are especially grateful to Drew Fudenberg, Eric Maskin and Tomasz Strzalecki for their constant guidance and support.

†Department of Economics, University of Pennsylvania
‡Cowles Foundation and Columbia University
§Microsoft Research

1 Introduction

In a classic problem of sequential information acquisition, a Bayesian decision-maker (DM) repeatedly acquires information and takes actions. His payoff depends on the sequence of actions taken, as well as on an unknown payoff-relevant state. We consider a setting in which the DM acquires information from a limited number of *flexibly correlated* information sources, and allocates a fixed number of observations across these sources each period.

Neglecting dynamic considerations, a simple strategy for information acquisition is to acquire at each period the set of signals that maximally reduces uncertainty about the payoff-relevant state. We refer to this as the *myopic* (or greedy) rule, as it is the optimal rule if the DM mistakenly believes each period to be the last possible period of information acquisition.

This myopic rule turns out to possess strong optimality properties when the available signals are jointly normal. First, if signal observations are acquired in sufficiently large blocks each period, then myopic information acquisition is exactly optimal (Theorem 1). We provide a sufficient condition on the required size of the block; this condition depends on primitives of the informational environment but not on the payoff function.

Second, under a joint condition on the prior and signal structure, myopic information acquisition is exactly optimal for *all* block sizes (Theorem 2). And finally, for generic signal structures, and for any block size, the optimal strategy proceeds by myopic acquisition after finitely many periods (Theorem 3). These results hold across all payoff functions (and in particular, independently of discounting).

Why does the myopic rule perform so well? The main inefficiency of myopic planning is that it neglects potential complementarities across signals. A signal that is individually uninformative can be very informative when paired with other signals; thus, repeated (greedy) acquisition of the best single signal need not result in the best sequence of signals.[1] A key observation is that whether the DM perceives two signals as providing complementary information depends on his current belief over the state space.[2] This means that complementarities across signals are not intrinsic to the underlying signal correlation structure: As the DM's beliefs about the states evolve, so too do his perceptions of the correlations

[1]See Section 4 for a concrete example and further discussion.

[2]As a simple example, suppose the payoff-relevant state is θ_1, and the available signals are about $\theta_1 + \theta_2$ and θ_2. These signals are "complementary" when the agent's prior belief is that θ_1 and θ_2 are independent: observations of the first signal improve the value of observing the second signal, and vice versa. But if the DM's prior is such that $\theta_2 = \theta_1$, then these signals become substitutes.

across signals. We show that as information accumlates, the DM's beliefs evolve in such a way that the signals *endogenously de-correlate* from his perspective, and are eventually perceived as providing approximately independent information.[3] At the limit in which all signals are independent, the value of any given signal can be evaluated separately of the others. The dynamic problem is thus "separable," and can be replaced with a sequence of static problems.

The mechanism we identify is different from the one underlying a classic result from the experimentation literature. In "learning by experimentation" settings, myopic behavior is eventually near-optimal: In the long run, the DM's beliefs converge, so the value of exploration (i.e. learning) becomes second-order relative to the value of exploitation of the perceived best arm.[4] In our paper, signal acquisition decisions are driven by learning concerns exclusively, as there is by design no exploitation incentive. To see this, recall that in the classic multi-armed bandit problem (Gittins, 1979; Bergemann and Välimäki, 2008), actions play the dual role of influencing the evolution of beliefs and also determining flow payoffs. In our setting (which does not fall into the multi-armed bandit framework), there is a separation between *signal choices*, which influence the evolution of beliefs, and *actions*, which determine (unobserved) payoffs. Myopic signal choices become optimal in our framework because *they maximize the speed of learning*, and not because they optimize a tradeoff between learning and payoffs. Additionally, a myopic strategy is immediately optimal in multi-armed bandit problems only under very restrictive assumptions (Berry and Fristedt, 1988; Banks and Sundaram, 1992).

Our results simplify the analysis of optimal dynamic information acquisition in an informational environment that is commonly used in economics: normal signals. However, the core of our analysis—the "endogenous de-correlation" of signals described above—does not rely on the assumption of normality. As we discuss further in Section 4, this de-correlation derives from a Bayesian version of the Central Limit Theorem, which holds for arbitrary signal distributions. This suggests that our eventual optimality result (Theorem 3) generalizes.[5]

[3] It is clear that the DM's beliefs eventually become very precise about each of the unknown states. However, this does not by itself lead to optimality of myopic information acquisition. See Section 5.1 for more detail.

[4] Easley and Kiefer (1988) and Aghion et al. (1991) show that if there is a unique myopically optimal policy at the limiting beliefs, then the optimal policies in each period must converge to this policy. In our setting, every policy (signal choice) is trivially myopic at the limiting beliefs (a point mass at the true parameter), so we do not have uniqueness and cannot use this argument to characterize long-run behavior.

[5] See Section 5.2 for discussion.

1.1 Related Literature

Information acquisition from Gaussian signals has been studied in a large literature (Angeletos and Pavan, 2007; Hellwig and Veldkamp, 2009; Myatt and Wallace, 2012; Bergemann et al., 2015; Lambert et al., 2018), primarily focusing on strategic settings where information is given exogenously or acquired at one time. Here we consider repeated information acquisitions by a single decision-maker. The closest works in this regard are Meyer and Zwiebel (2007), Sethi and Yildiz (2016), Sethi and Yildiz (2017), and Liang and Mu (2018), all of which focus on long-run or stationary outcomes. In contrast, here we consider the entire path of information acquisitions and show that in many cases it is identical to the myopic strategy.

A related recent literature (Bubeck et al., 2009; Russo, 2016) studies "best-arm identification" in a multi-armed bandit setting: A DM samples for a number of periods before selecting an arm and receiving its payoff. Our results for the case of two states exactly apply to the problem of identifying the better of two correlated normal arms. However, due to our assumption of an one-dimensional payoff-relevant state, we are not able to handle more than two arms.[6] We note that correlation is the key feature of our setting, and are not aware of many papers that study correlated bandits, either under the classical framework or under best-arm identification. See Rothschild (1974), Keener (1985) and Mersereau et al. (2009) for a few stylized cases.

Our results on the comparison of sequential normal experiments (see the discussion in Section 4, and results in Appendix B) generalize the main result in Greenshtein (1996). Greenshtein (1996) compares two *deterministic (i.e. history-independent)* sequences of signals, where each signal is the payoff-relevant state plus *independent* normal noise. His Theorem 3.1 implies that one sequence is Blackwell-dominant if and only if its cumulative precision is higher at every time. Note that this statement does not refer to the prior beliefs, but if we impose a normal prior, then higher cumulative precision is equivalent to lower posterior variance. Thus, the result of Greenshtein (1996) coincides with ours when there is only one persistent (and payoff-relevant) state, and all signals are independent conditional on it. Our setting features additional correlation across different signals. Consequently, dynamic

[6] With two arms, the DM only cares about the difference in their expected payoffs. Choosing among more than two arms would involve *multi-dimensional payoff uncertainty* and *a decision problem that is not prediction*. The lack of a complete Blackwell ordering limits the generalization of our argument. Incidentally, in related sequential search settings, Sanjurjo (2017), Ke and Villas-Boas (2017) and Chick and Frazier (2012) also highlight the challenge of characterizing the optimal strategy once there are at least three alternatives.

Blackwell comparison in our model depends on prior beliefs.[7] This feature, together with the endogenous choice of signals (which may be history-dependent), complicates our analysis relative to Greenshtein (1996).

Finally, our work is closely related to *optimal design*, a field initiated by the the early work of Robbins (1952) (see Chernoff (1972) for a survey). Specifically, the problem of one-shot allocation of t signals (our t-optimal criterion in Section 5) is equivalent to a Bayesian optimal design problem with respect to the "c-optimality criterion", which seeks to minimize the variance of an unknown parameter. Our analysis is however focused on dynamics, and we demonstrate here the optimality of "greedy design" for a broad class of *intertemporal* objectives.

2 Preliminaries

2.1 Model

Time is discrete. At each time $t = 1, 2, \dots$, the DM first allocates a budget of *samples* or *observations* across K information sources, and then chooses an action a_t from a set A_t.[8]

The DM's payoff $U(a_1, a_2, \dots; \omega)$ is an arbitrary function that depends on the sequence of action choices and a payoff-relevant state $\omega \in \mathbb{R}$. We assume that payoffs are realized only at an (exogenously or endogenously determined) end date; thus, the information sources described below are the only channel through which the DM learns. This assumption distinguishes our model from multi-armed bandit problems.

To illustrate a familiar subclass of such decision problems, imagine that the DM takes an action just once at a final period T, which may be determined by an exogenous distribution (e.g. geometric distribution) or endogenously chosen by the DM. Payoff in this case is given by $U(a_1, a_2, \dots; \omega) = u_T(a_T, \omega)$, where T is the (random) exogenous or endogenous final time period, and a_T is the action taken. The time-dependent payoff function $u_T(a_T, \omega)$ may incorporate discounting and/or a per-period cost to signal acquisition.[9]

Apart from the decision problem, there are K information sources, which depend on

[7] This is already the case for static comparisons, since as the prior beliefs vary, it is not always the same signal that leads to the lowest posterior variance about the payoff-relevant state.

[8] Thus, the action a_t can be based on the information received in period t. The timing of these choices is not important for our results.

[9] In the latter case, we think of costs as fixed across sources in a given period, but can vary across periods. See See e.g. Fudenberg et al. (2018) and Che and Mierendorff (2019) for recent models with constant waiting cost per period.

the unknown and persistent state vector $\theta = (\theta_1, \ldots, \theta_K)' \sim \mathcal{N}(\mu^0, V^0)$.[10] We assume that the prior covariance matrix V^0 has full rank. The payoff-relevant state ω is some linear combination of these unknown states, which we write as $\omega = \alpha' \cdot \theta$ for some $\alpha \in \mathbb{R}^K$.

In each period, the DM chooses B sources (allowing for repetition), where $B \in \mathbb{N}^+$ is interpreted as a fixed time/attention constraint (see Section 5.3 for extension to endogenous choice of B). Choice of source $k = 1, \ldots, K$ produces an observation of

$$X_k = c_k' \cdot \theta + \epsilon_k, \quad \epsilon_k \sim \mathcal{N}(0, \sigma_k^2)$$

where the coefficient vectors $c_k = (c_{k1}, \ldots, c_{kK})'$ and signal variances σ_k^2 are fixed (and known), but the Gaussian error terms are independent across realizations and sources. Throughout, we use C to denote the matrix of coefficients whose k-th row is c_k'. We impose the following assumption on the informational environment:

Assumption 1 (Non-Redundancy)**.** *The matrix C has full rank, and no proper subset of row vectors of C spans α (the vector determining the payoff-relevant state).*

This implies that the DM *can and must* observe each source infinitely often to recover the value of the payoff-relevant state ω.[11]

2.2 Interpretations

We provide below several interpretations of this framework.

News Sources with Correlated Biases. On election day T, a DM will choose which of two candidates 1 and 2 to vote for, where his payoff depends on $\omega = \theta_1 - \theta_2$, the difference between the candidates' qualities θ_1 and θ_2. In each period up to time T, the DM can acquire information from different news sources. All sources provide biased information, and moreover the biases are correlated across the sources. As the DM acquires information, he learns not only about the payoff-relevant state ω, but also how to de-bias (and aggregate) information from the various news sources.

Attribute Discovery. A product has K unknown attributes $\theta_1, \ldots, \theta_K$. Its value ω is some arbitrary linear combination of these attribute values. For example, the DM may want to learn the value of a conglomerate composed of several companies, where each company i is

[10] Here and later, we use the apostrophe to denote vector or matrix transpose.

[11] When the DM does not need to observe all sources in order to learn ω, a new question emerges regarding which subset of sources the DM will choose from. Characterization of that subset is the focus of Liang and Mu (2018).

valued at θ_i and the value of the conglomerate is $\omega := \alpha_1\theta_1 + \cdots + \alpha_K\theta_K$. The DM has access to (noisy) observations of different linear combinations of the attributes; for example, he might have access to evaluations of each θ_i individually. At some endogenously chosen end time, the DM decides whether or not to invest in the conglomerate.

Sequential Polling. A polling organization seeks to predict the average opinion in the population towards an issue. There are K demographic groups in the population, and opinions in demographic group i are normally distributed with unknown mean θ_i and known variance σ_i^2. The fraction of the population in each demographic group i is p_i, so the average opinion is $\omega := \sum_i p_i\theta_i$. It is not feasible to directly sample individuals according to the true distribution p_i, but the organization can sample individuals according to other non-representative distributions $\hat{p}_i \neq p_i$. Each period, the polling organization allocates a fixed budget of opinion samples across the available distributions (polling technologies), and posts a prediction for ω. Its payoff is the average prediction error across some fixed number of periods.

Intertemporal Investment. Each action a_t is a decision of how to allocate capital between consumption, and two investment possibilities: a liquid asset (bond), and an illiquid asset (pension fund). The return to the liquid asset is known: 1 dollar saved today is worth e^r dollars tomorrow. The return to the illiquid asset is unknown: every dollar invested today in the pension fund deterministically yields e^ω dollar(s) tomorrow, where ω is the payoff-relevant state. The worker works for T periods, and in each of these periods he learns about ω (from some information sources) and then allocates his wealth across consumption, saving and investment. After period T, the worker retires and receives all the returns from his investments into the illiquid asset. His objective is to maximize the sum of his discounted consumption utilities and his payoff after retirement.[12]

[12]An important assumption of this example is that the return to investment is deterministic and only observed at the end. However, our model and results extend to a situation where there are "free" signals arriving each period that do not count toward the capacity constraint B. By considering the realized log return as a particular free signal, the extension of our model covers the case where investment returns are stochastic and the DM observes past return realizations.

3 (Eventual) Optimality of Myopic Rule

3.1 Myopic Information Acquisition

A strategy consists of an information acquisition strategy and a decision strategy. An *information acquisition strategy* is a measurable map from possible histories of signal realizations to multi-sets of B signals, and a *decision strategy* is a map from histories to actions.

We will say that an information acquisition strategy is myopic if it proceeds by choosing signals that maximally reduce (next period) uncertainty about the payoff-relevant state.

Definition 1. *An information acquisition strategy is* myopic*, if at every period, it prescribes choosing the B signals that (combined with the history of observations) lead to the lowest posterior variance about ω.*

Note that the B signals which minimize the posterior variance also Blackwell dominate any other multi-set of B signals (see e.g. Hansen and Torgersen (1974)). Thus, myopic acquisition is optimal if the current period is the last chance for information acquisition, and this is true no matter what the payoff function is.

Our results below reveal a close relationship between the optimal information acquisition strategy and the myopic rule, even when the DM takes into account the possibility of future information acquisition. We do not pursue a characterization of the optimal *decision* strategy, which in general depends on the payoff function, although we point out one application of our main results towards simplification of this characterization.

3.2 Main Results

We present three results regarding optimality of the myopic information acquisition rule. Theorem 1 says that myopic information acquisition is optimal from period 1 if B (the number of observations acquired each period) is sufficiently large. Our next two results hold for arbitrary budgets B: Theorem 2 provides a sufficient condition on the prior and the coefficient matrix C under which myopic information acquisition is optimal from period 1, and Theorem 3 states that the optimal rule is *eventually* myopic in generic environments.

Theorem 1 (Immediate Optimality under Many Observations)**.** *Fix any prior and signal structure, and suppose B is sufficiently large. Then the DM has an optimal strategy that acquires information myopically.*[13]

[13] Without further assumptions on the payoff function U, we cannot assert *strict* optimality of the myopic

The theorem tells us that, optimally, the DM chooses the most informative B signals in the first period based on his prior, then chooses the most informative B signals in the second period based on his updated posterior, and so on. Note that since posterior variances are independent of signal realizations (and since we assume no feedback from past actions), the myopic strategy is *history-independent*, and can be represented as a deterministic signal path. We mention that Theorem 1 can be strengthened to optimality of the myopic rule *following every history*, including those in which the DM has previously deviated from the myopic rule.

A precise bound for how large B must be for Theorem 1 to hold appears in Section 5.1. Our next two results hold for arbitrary block sizes B. First, the myopic rule is again optimal from period 1 in a class of "separable" environments. Let $f(q_1, \ldots, q_K)$ denote the DM's posterior variance about ω, updating from q_i observations of each signal X_i. An informational environment is separable if its posterior variance function can be decomposed in the following way:

Definition 2. *The informational environment* $(V^0, C, \{\sigma_i^2\})$ *is* separable *if there exist convex functions* $g_1, \ldots, g_K$ *and a strictly increasing function* F *such that*

$$f(q_1, \ldots, q_K) = F(g_1(q_1) + \cdots + g_K(q_K)).$$

Intuitively, separability ensures that observing signal i does not change the relative value of other signals, but strictly decreases the marginal value of signal i relative to every other signal.

Note that separability is not defined directly on the primitives of the informational environment (V^0, C, and $\{\sigma_i^2\}_{i=1}^K$), as it is based instead on the posterior variance function f. Nevertheless, f can be directly computed from these primitives. For example, it is simple to verify that the following two informational environments are separable:

Example (Orthogonal Signals). The DM's prior is standard Gaussian ($V^0 = \mathbf{I}_K$), and the row vectors of C form an orthogonal basis.

Example (Multiple Biases). There is a single payoff-relevant state $\omega = \theta_1 \sim \mathcal{N}(0, v_1)$. The DM has access to observations of $X_1 = \theta_1 + \theta_2 + \cdots + \theta_K + \epsilon_1$, where each θ_i ($i > 1$) is a persistent "bias" independently drawn from $\mathcal{N}(0, v_i)$, and $\epsilon_1 \sim \mathcal{N}(0, \sigma_1^2)$ is a noise term i.i.d.

rule. For instance, this would not be true if there exists a "dominant" action sequence that maximizes $U(a_1, a_2, \ldots; \omega)$ for every value of ω. But the myopic rule would be uniquely optimal in any decision problem where more precise beliefs lead to strictly higher expected payoffs.

over time. Additionally, the DM has access to signals about each bias $X_i = \theta_i + \epsilon_i$ $(i > 1)$, where $\epsilon_i \sim \mathcal{N}(0, \sigma_i^2)$.[14]

The result below says that myopic information acquisition is optimal in all separable informational environments.

Theorem 2 (Immediate Optimality in Separable Environments)**.** *Suppose the informational environment is separable. Then for every $B \in \mathbb{N}^+$, the DM has an optimal strategy that acquires information myopically.*

In all remaining cases, optimal signal choices are eventually myopic.

Theorem 3 (Generic Eventual Myopia)**.** *Fix any prior covariance matrix V^0 and signal variances $\{\sigma_i^2\}_{i=1}^K$. For generic coefficient matrices C, there exists a time $T^* \in \mathbb{N}$ depending on the information environment such that in any decision problem, the DM has an optimal strategy that acquires information myopically after T^* periods.*

Thus at all late periods, the optimal signal acquisitions are those that maximally reduce posterior variance in the given period.

This result tells us that the optimal rule eventually proceeds by myopic signal acquisition; this is different from the statement that acquisition of signals myopically *from period 1* eventually leads to the optimal signal path. We show in Appendix H that the latter statement is also true.

4 Intuition for Theorems 1-3

When Myopic is Optimal. We begin by considering acquisition of $B = 1$ observations each period. Even with this small block size, myopic information acquisition is sometimes optimal.

Example 1. Suppose the DM wants to learn $\omega = \theta_1 + \theta_2 + \theta_3$, where $\theta_1, \theta_2, \theta_3 \sim_{i.i.d.} \mathcal{N}(0, 1)$

[14] To check separability, note that the DM's posterior variance about ω is given by

$$f(q_1, \ldots, q_K) = v_0 - \frac{v_0^2}{v_0 + \frac{\sigma_1^2}{q_1} + \sum_{i=2}^K \left(v_i - \frac{v_i^2}{v_i + \sigma_i^2/q_i}\right)}.$$

and the available signals are:

$$X_1 = \theta_1 + \epsilon_1$$
$$X_2 = \theta_2 + \epsilon_2$$
$$X_3 = \theta_3 + \epsilon_3$$

It is simple to see that myopic information acquisition proceeds by "balancing" acquisition of information across the sources—e.g. $X_1X_2X_3X_1X_2X_3\dots$—and moreover, this strategy minimizes posterior variance about ω at every period. We generalize a result from Greenshtein (1996) to show that any strategy which pointwise minimizes posterior variance must be Blackwell-dominant among the class of all information acquisition strategies, implying its optimality across all payoff criteria.

An important feature of this example is that the best signal to acquire in period 1 is contained in the best pair of signals to acquire, which is contained in the best triple, and so forth. Thus, no tradeoffs are necessary across periods, and the myopic rule minimizes posterior variance period by period. This property is formalized in our notion of "separability" in Theorem 2, and is a sufficient condition for optimality of myopic acquisition.

The Role of Complementarities. In general however, myopic information acquisition fails to be optimal when there are strong complementarities across signals, as in the following example:

Example 2. There are three states $\theta_1, \theta_2, \theta_3 \sim_{i.i.d.} \mathcal{N}(0, 1)$ and the available signals are:

$$X_1 = \theta_1 - \theta_2 + \epsilon_1$$
$$X_2 = \theta_2 - \theta_3 + \epsilon_2$$
$$X_3 = \theta_3 + \epsilon_3$$

The agent wants to learn $\omega = \theta_1$.

In this example, the optimal allocation of four observations across X_1, X_2, and X_3 is $(3, 1, 0)$, and the best subsequent *single* signal to acquire is X_1. But because signals X_2 and X_3 are strong complements (observation of either signal increases the value of observation of the other), the best subsequent *pair of signals* is $\{X_2, X_3\}$. Thus the optimal allocation of five observations is $(4, 1, 0)$ while the optimal allocation of six observations is $(3, 2, 1)$, which cannot be produced from the history $(4, 1, 0)$. Myopic information acquisition fails to be optimal in this environment.[15]

[15] Observe that with five observations to allocate, $f(4, 1, 0) = \frac{11}{23} < \frac{14}{29} = f(3, 1, 1) = f(3, 2, 0)$. Whereas with six observations, $f(3, 2, 1) = \frac{5}{11} < \frac{17}{37} = f(4, 1, 1) = f(4, 2, 0)$.

"De-correlation" of Posterior Beliefs. A key part of our argument is that complementarities "wash out" as information is acquired, so that environments which start off like the second example become "like" the first example (as observations of each signal accumulate). To facilitate this comparison, rewrite Example 2 in the following way: Define a new set of states $\tilde{\theta}_1 = \theta_1 - \theta_2$, $\tilde{\theta}_2 = \theta_2 - \theta_3$, and $\tilde{\theta}_3 = \theta_3$. The original prior over $(\theta_1, \theta_2, \theta_3)$ defines a new prior over $(\tilde{\theta}_1, \tilde{\theta}_2, \tilde{\theta}_3)$, and the original payoff-relevant state can be re-expressed in the new states as $\omega = \tilde{\theta}_1 + \tilde{\theta}_2 + \tilde{\theta}_3$.

Example 2'. There are three states $\tilde{\theta}_1, \tilde{\theta}_2, \tilde{\theta}_3 \sim_{i.i.d.} \mathcal{N}(0, \Sigma)$ and the available signals are:

$$X_1 = \tilde{\theta}_1 + \epsilon_1$$
$$X_2 = \tilde{\theta}_2 + \epsilon_2$$
$$X_3 = \tilde{\theta}_3 + \epsilon_3$$

The agent wants to learn $\omega = \tilde{\theta}_1 + \tilde{\theta}_2 + \tilde{\theta}_3$.

The signal structure and the payoff-relevant state in Example 2' are the same as in Example 1, with the crucial difference that the prior over $(\tilde{\theta}_1, \tilde{\theta}_2, \tilde{\theta}_2)$ is correlated in Example 2', while the prior over $(\theta_1, \theta_2, \theta_3)$ was independent in Example 1. Given sufficiently many observations of each signal in Example 2', however, the DM's posterior beliefs over $(\tilde{\theta}_1, \dots, \tilde{\theta}_K)$ become *almost independent*. That is, not only does learning about each $\tilde{\theta}_k$ occur, but the states $\tilde{\theta}_k$ "de-correlate." Thus, eventually the signals can be viewed as approximately independent.

The above heuristic statements about de-correlation roughly follow from a Bayesian version of the Central Limit Theorem. We establish a technical lemma (Lemma 2 in Appendix A) that strengthens this with a comparison of the value of different signals. Specifically, we characterize the externality that observation of a given signal X_i has on the marginal value of future observations. We show that once posterior beliefs are sufficiently de-correlated, the effect of observing X_i on the value of future observations of X_i far outweighs its effect on future observations of any X_j, $j \neq i$. Thus, the only effect that observation of X_i can have on the ranking of signals is by reducing the value of signal i relative to other signals. This means that we eventually have a situation much like Example 1, where the different signals are *almost separable*.

Exact Optimality. In near-separable environments, the difference between the myopic strategy and the optimal strategy is small, but the two strategies may not be identical. We show that the "gaps" between myopic and optimal acquisition occur for a purely technical reason related to integer approximations. Using results from Diophantine approximation theory, we demonstrate that the environments in which these inefficiencies emerge are non-generic: this gives us our result in Theorem 3. A different way to resolve the integer approximation problem is to allow for acquisition of a larger number of signals each period. We show that given a sufficiently large block size B, the optimal allocations of kB samples are nested for different integers k, re-producing exact separability. This gives us Theorem 1.

5 Discussion and Extensions

5.1 Precision vs. Correlation

With sufficiently many observations, the DM's beliefs simultaneously become more precise and less correlated, and these two effects are confounded in our main results. It is tempting to think that Theorem 1 (or Theorem 3) follows from the eventual precision of beliefs. However, as our discussion above suggests, the key feature is not how precise beliefs are, but how correlated they are. Specifically, the block size B needed in Theorem 1 depends on how many observations are required for the transformed states $\tilde{\theta}_1, \dots, \tilde{\theta}_K$ (see below) to "de-correlate," at which point complementarities across signals are weak.

Below we make this formal with a bound on B. To state our result, we define transformed states $\tilde{\theta}_k = \frac{1}{\sigma_k}\langle c_k, \theta \rangle$ (dividing through by σ_k normalizes all error variances to 1), and let $\tilde{V}$ denote the prior covariance matrix over these transformed states. The payoff-relevant state $\omega = \alpha' \cdot \theta$ can be rewritten as another linear combination of the transformed states: $\omega = \tilde{\alpha}' \cdot \tilde{\theta}$ for some fixed payoff weight vector $\tilde{\alpha} \in \mathbb{R}^K$. In the following result we assume for simplicity that $\tilde{\alpha} = \mathbf{1}$ is the vector of ones, although our analysis can be easily adapted to any vector $\tilde{\alpha}$.

Proposition 1. *Let R denote the operator norm of the matrix $(\tilde{V})^{-1}$.*[16] *Suppose $\omega = \sum_{i=1}^{K} \tilde{\theta}_i$, then $B \geq 8(R+1)K^{1.5}$ is sufficient for Theorem 1 to hold.*

Notice that this bound is increasing in the norm of $\tilde{V}^{-1}$. To interpret this, suppose first that we adjust the precision of the DM's prior beliefs over $\tilde{\theta}_1, \dots, \tilde{\theta}_K$ but fix the degree of correlation, for example by scaling $\tilde{V}$ by a factor less than 1. Then the norm of $\tilde{V}^{-1}$ increases, and a larger block size B is needed. This is because a more precise prior can be understood

[16]The operator norm of a matrix M is defined as $\|M\|_{op} = \sup\left\{\frac{\|Mx\|}{\|x\|} : x \in \mathbb{R}^K \text{ with } x \neq \mathbf{0}\right\}$.

as "re-scaling" the state space by shrinking all states towards zero. Since signal noise is not correspondingly re-scaled, each signal now reveals less about the states, and de-correlation takes longer.

In contrast, suppose we hold prior precision fixed and increase the degree of prior correlation. This would correspond to fixing the diagonal entries of $\tilde{V}$ and increasing the off-diagonal entries, so that the variances about individual states are unchanged but their covariances become larger in magnitude. Then the entire matrix $\tilde{V}$ becomes closer to being singular, the norm of its inverse increases and a larger B is required. That is, a more correlated prior requires more observations to de-correlate.

To summarize, optimality of myopic information acquisition obtains quickly when prior beliefs are *imprecise* and *weakly correlated*.

5.2 How Important is Normality?

The key part of our argument is that signals eventually de-correlate. This de-correlation derives from a Bayesian version of the Central Limit Theorem, and does not rely on special properties of normality. Consider a more general setting with signals $X_i = \tilde{\theta}_i + \epsilon_i$, where the noise term ϵ_i has an arbitrary distribution with zero mean and finite variance. Then, we have that the suitably normalized posterior distribution over $(\tilde{\theta}_1, \dots, \tilde{\theta}_K)$ converges towards a standard normal distribution (so that signals are approximately independent). We thus expect that given sufficiently many observations, our previous comparisons on the value of information extend beyond normal signals.

But if we drop normality, our results weaken in the following ways:

Immediate Optimality of the Myopic Rule. For normal signals, we can establish a T such that given T observations of each signal, the posterior covariance matrix (over the transformed states) is almost independent. Notably, this bound T holds uniformly across all histories of signal realizations, thanks to the fact that posterior variances do not depend on signal realizations under normality. As mentioned above, we can use a Bayesian Central Limit Theorem to argue a similar property for other signal distributions. The difference is that the CLT gives us approximate independence *almost surely*, so that at every period t, there is still positive probability (albeit vanishing as t increases) that the posterior covariance matrix is far from being diagonal. This precludes us from demonstrating a block size B given which the optimal rule would be myopic *from period 1* (Theorem 1). For general signal distributions, we thus conjecture that *almost surely* the optimal rule is eventually myopic, but we do not know what conditions would guarantee this equivalence from the beginning.

General Intertemporal Payoffs. The place where we rely most heavily on normality is the statement that our results hold for *all* payoff criteria that depends on ω (and actions). Indeed, when payoff-relevant uncertainty is one-dimensional (as it is here), all normal signals can be Blackwell-ordered based on their posterior variances, and this is not the case for general signals. We note moreover that normality is important for comparison of signal sequences. As discussed in Section 4, the ranking of sequences of normal signals is the same whether we consider the class of static decision problems or the broader class of intertemporal decisions. This equivalence does not hold in general; see Greenshtein (1996) for a counterexample involving Bernoulli signals.

5.3 Endogenous Block Sizes

So far we have imposed an exogenous capacity constraint of B signals per period. Suppose now that in each period t, the DM can choose to observe any number $N_t \in \mathbb{Z}_+$ of signal realizations (which are then optimally allocated across signals). The DM incurs a flow cost of information acquisition, modeled as $\kappa(N_t)$ for some increasing cost function $\kappa(\cdot)$ with $\kappa(0) = 0$. This framework embeds our main model if we define $\kappa(N) = 0$ for $N \leq B$ and $\kappa(N) = \infty$ for $N > B$.

Let the DM's payoff be $U(a_1, a_2, \ldots; \omega) - \sum_t \delta^{t-1} \cdot \kappa(N_t)$ for some discount factor δ.[17] For the special case of endogenous stopping, the payoff function simplifies to

$$\delta^\tau \cdot u(a_\tau; \omega) - \sum_{t=1}^{\tau} \delta^{t-1} \cdot \kappa(N_t)$$

whenever the DM stops after τ periods. This is a discrete-time generalization of the framework proposed in Moscarini and Smith (2001), although our focus is on allocation of the signals instead of choice of "learning intensities" N_t.[18]

Theorems 1 and 2 generalize to this setting:

Corollary 1. *Suppose B is sufficiently large or the informational environment is separable. Then, even with endogenous block sizes, the DM has an optimal strategy that chooses signals myopically.*

[17] Our analysis can accommodate more general payoff functions of the form $U(N_1, a_1, N_2, a_2, \ldots; \omega)$.

[18] Moscarini and Smith (2001) has a single state and a single signal ($K = 1$), so the DM only chooses N_t. Unlike Moscarini and Smith (2001), we do not characterize the optimal sequence of $(N_t)_{t \geq 1}$, but instead show how this problem can be separated from allocation of those N_t observations across different signals.

In the above corollary, "myopic acquisition" means the following: In any period t, given (endogenously chosen) block size N_t, optimal acquisitions are the N_t signals that minimize posterior variance about ω. We emphasize that while myopic signal choices are optimal, myopic intensity choices are likely not. However, knowing that the signal choices must follow the myopic path provides a simplifying first step towards the characterization of optimal block sizes. Generic eventual myopia (Theorem 3) also extends, but we omit the details.

Appendix

A Preliminary Results

Throughout this appendix, we assume that the payoff-relevant state ω in fact coincides with the first unknown state θ_1, which will simplify the exposition. This is without loss of generality, since we can always define "transformed states" to include ω. Moreover, note that Assumption 1 translates into the requirement that the matrix of signal coefficients C *admits an inverse matrix, whose first row contains non-zero entries.* That is, we work under the assumption that $[C^{-1}]_{1i} \neq 0, \forall 1 \leq i \leq K$.[19]

A.1 Posterior Variance Function

We begin by presenting basic results that will be used repeatedly. The following lemma characterizes the posterior variance function f mentioned in the main text, which maps signal counts to the DM's posterior variance about the payoff-relevant state θ_1.

Lemma 1. *Given prior covariance matrix V^0 and q_i observations of each signal i, the DM's posterior variance about θ_1 is given by*

$$f(q_1, \ldots, q_K) = [V^0 - V^0 C' \Sigma^{-1} C V^0]_{11} \tag{1}$$

where $\Sigma = CV^0C' + D^{-1}$ and $D = \operatorname{diag}\left(\frac{q_1}{\sigma_1^2}, \ldots, \frac{q_K}{\sigma_K^2}\right)$. The function f is decreasing and convex in each q_i whenever these arguments take non-negative extended real values: $q_i \in \overline{\mathbb{R}_+} = \mathbb{R}_+ \cup \{+\infty\}$.[20]

Proof. The expression (1) comes directly from the conditional variance formula for multivariate Gaussian distributions. To prove $\frac{\partial f}{\partial q_i} \leq 0$, consider the partial order $\succeq$ on positive semi-definite matrices so that $A \succeq B$ if and only if $A - B$ is positive semi-definite. As q_i increases, the matrices D^{-1} and Σ decrease in this order. Thus Σ^{-1} increases in this order, which implies that $V^0 - V^0C'\Sigma^{-1}CV^0$ decreases in this order. In particular, the diagonal entries of $V^0 - V^0C'\Sigma^{-1}CV^0$ are uniformly smaller, so that f becomes smaller. Intuitively, more information always improves the decision-maker's estimates.

[19] When M is matrix, we let M_{ij} denote its (i, j)-th entry.

[20] We allow the function f to take $+\infty$ as arguments. This extension does not affect the properties of f, and it is convenient for some of our analysis.

To prove f is convex, it suffices to prove f is *midpoint-convex* since the function is clearly continuous. Take $q_1, \dots, q_K,\ r_1, \dots, r_K \in \overline{\mathbb{R}_+}$ and let $s_i = \frac{q_i + r_i}{2}$. Define the corresponding diagonal matrices to be $D_q,\ D_r,\ D_s$. We need to show $f(q_1, \dots, q_K) + f(r_1, \dots, r_K) \geq 2f(s_1, \dots, s_K)$. For this, we first use the Woodbury inversion formula to write

$$\Sigma^{-1} = (CV^0C' + D^{-1})^{-1} = J - J(J+D)^{-1}J,$$

with $J = (CV^0C')^{-1}$. Plugging this back into (1), we see that it suffices to show the following matrix order:

$$\frac{(J+D_q)^{-1} + (J+D_r)^{-1}}{2} \succeq (J+D_s)^{-1}.$$

Inverting both sides, we need to show $2\left((J+D_q)^{-1} + (J+D_r)^{-1}\right)^{-1} \preceq J + D_s$. From definition, $D_q + D_r = \operatorname{diag}(\frac{q_1+r_1}{\sigma_1^2}, \dots, \frac{q_K+r_K}{\sigma_K^2}) = 2D_s$. Thus the above follows from the AM-HM inequality for positive definite matrices, see for instance Ando (1983). □

A.2 The Matrix Q_i

Let us define for each $1 \leq i \leq K$,

$$Q_i = C^{-1}\Delta_{ii}C'^{-1} \tag{2}$$

where Δ_{ii} is the matrix with '1' in the (i,i)-th entry, and zeros elsewhere. We note that $[Q_i]_{11} = ([C^{-1}]_{1i})^2$, which is strictly positive under Assumption 1. These matrices Q_i play an important role in our proofs.

A.3 Order Difference Lemma

Here we establish the asymptotic order for the second derivatives of f.

Lemma 2. *As $q_1, \dots, q_K \to \infty$, $\frac{\partial^2 f}{\partial q_i^2}$ is positive with order $\frac{1}{q_i^3}$, whereas $\frac{\partial^2 f}{\partial q_i \partial q_j}$ has order at most $\frac{1}{q_i^2 q_j^2}$ for any $j \neq i$. Formally, there is a positive constant L depending on the informational environment, such that $\frac{\partial^2 f}{\partial q_i^2} \geq \frac{1}{Lq_i^3}$ and $|\frac{\partial^2 f}{\partial q_i \partial q_j}| \leq \frac{L}{q_i^2 q_j^2}$.*

To interpret, the second derivative $\partial^2 f/\partial q_i^2$ is the effect of observing signal i on the marginal value of the next observation of signal i. Our lemma says that this second derivative is always eventually positive, so that each observation of signal i makes the next observation of signal i less valuable. The cross-partial $\partial^2 f/\partial q_i \partial q_j$ is the effect of observing signal i on the marginal value of the next observation of a different signal j, and its sign is ambiguous.

The key content of the lemma is that regardless of the sign of the cross partial, it is always of lower order compared to the second derivative. In words, the effect of observing a signal on the marginal value of other signals (as quantified by the cross-partial) is eventually second-order to its effect on the marginal value of further realizations of the same signal (as quantified by the second derivative). This is true for any signal path in which the signal counts $q_1, \ldots, q_K$ go to infinity proportionally, which we will justify later.

Proof. Recall from Lemma 1 that $f(q_1, \ldots, q_K) = [V^0 - V^0 C' \Sigma^{-1} C V^0]_{11}$ and therefore

$$\frac{\partial^2 f}{\partial q_i \partial q_j} = [\partial_{ij}(V^0 - V^0 C' \Sigma^{-1} C V^0)]_{11} \qquad \frac{\partial^2 f}{\partial q_i^2} = [\partial_{ii}(V^0 - V^0 C' \Sigma^{-1} C V^0)]_{11} \tag{3}$$

Using properties of matrix derivatives,

$$\partial_{ii}(\Sigma^{-1}) = \Sigma^{-1}(\partial_i \Sigma)\Sigma^{-1}(\partial_i \Sigma)\Sigma^{-1} - \Sigma^{-1}(\partial_{ii}\Sigma)\Sigma^{-1} + \Sigma^{-1}(\partial_i \Sigma)\Sigma^{-1}(\partial_i \Sigma)\Sigma^{-1}. \tag{4}$$

The relevant derivatives of the covariance matrix Σ are

$$\partial_i \Sigma = -\frac{\sigma_i^2}{q_i^2}\Delta_{ii} \qquad \partial_{ii}\Sigma = \frac{2\sigma_i^2}{q_i^3}\Delta_{ii}$$

Plugging these into (4), we obtain $\partial_{ii}(\Sigma^{-1}) = -\frac{2\sigma_i^2}{q_i^3}(\Sigma^{-1}\Delta_{ii}\Sigma^{-1}) + O\left(\frac{1}{q_i^4}\right)$. Thus by (3),

$$\frac{\partial^2 f}{\partial q_i^2} = \left[-V^0 C' \cdot \frac{\partial^2(\Sigma^{-1})}{\partial q_i^2} \cdot C V^0\right]_{11} = \frac{2\sigma_i^2}{q_i^3} \cdot \left[V^0 C' \Sigma^{-1} \Delta_{ii} \Sigma^{-1} C V^0\right]_{11} + O\left(\frac{1}{q_i^4}\right). \tag{5}$$

As $q_1, \ldots, q_k \to \infty$, $\Sigma \to C V^0 C'$ which is symmetric and non-singular. Thus the matrix $V^0 C' \Sigma^{-1} \Delta_{ii} \Sigma^{-1} C V^0$ converges to the matrix Q_i defined earlier in (2). From (5) and $[Q_i]_{11} > 0$, we conclude that $\frac{\partial^2 f}{\partial q_i^2}$ is positive with order $\frac{1}{q_i^3}$. Similarly, for $i \neq j$, we have

$$\partial_{ij}(\Sigma^{-1}) = \Sigma^{-1}(\partial_j \Sigma)\Sigma^{-1}(\partial_i \Sigma)\Sigma^{-1} - \Sigma^{-1}(\partial_{ij}\Sigma)\Sigma^{-1} + \Sigma^{-1}(\partial_i \Sigma)\Sigma^{-1}(\partial_j \Sigma)\Sigma^{-1}.$$

The relevant derivatives of the covariance matrix Σ are

$$\partial_i \Sigma = -\frac{\sigma_i^2}{q_i^2}\Delta_{ii} \qquad \partial_j \Sigma = -\frac{\sigma_j^2}{q_j^2}\Delta_{jj} \qquad \partial_{ij}\Sigma = \mathbf{0}$$

From this it follows that $\partial_{ij}(\Sigma^{-1}) = O\left(\frac{1}{q_i^2 q_j^2}\right)$. The same holds for $\frac{\partial^2 f}{\partial q_i \partial q_j}$ because of (3), completing the proof of the lemma. □

A.4 t-optimality

We introduce a central definition that will go into our proofs:

Definition 3. *Say a vector of signal counts $(n_1, \ldots, n_K)'$ is t-optimal if it minimizes the posterior variance about ω among all vectors that allocate t observations. That is,*

$$(n_1, \ldots, n_K) \in \underset{(q_1,\ldots,q_K): q_i \in \mathbb{Z}_+, \sum_i q_i = t}{\operatorname{argmin}} f(q_1, \ldots, q_K).$$

Such vectors need not be unique, but we will often write $n(t)$ as a generic t-optimal vector. We will also use the phrase "division vector" interchangeably with "signal count vector."

B Dynamic Blackwell Comparison

B.1 Statement of the Lemma

This subsection establishes a dynamic version of Blackwell dominance for sequences of normal signals. As an overview, we first generalize Greenshtein (1996) and show that a *deterministic* (i.e. history-independent) signal sequence yields higher expected payoff than another in every intertemporal decision problem if (and only if) the former sequence induces lower posterior variances about θ_1 at every period. This will be a corollary of the lemma below, which also covers strategies that may condition on signal realizations.

We introduce some notation: Since θ_1 is the only payoff-relevant state, the DM in our model only needs to remember the expected value of θ_1 and the covariance matrix over all of the states (that is, expected values of the other states do not matter). Thus, we can summarize any history of beliefs by $h^T = (\mu_1^0, V^0; \ldots, \mu_1^T, V^T)$, with μ_1^t representing the posterior expected value of θ_1 after t periods and V^t the posterior covariance matrix. Since the posterior covariance matrix is a function of signal counts, we can also keep track of the evolution of posterior covariance matrices by a sequence of division vectors. That is, we will write the history as $h^T = (\mu_1^0, d(0); \ldots; \mu_1^T, d(T))$, where each $d(t) = (d_1(t), \ldots, d_K(t))$ counts the number of each signal acquired by time t. We can then view any information acquisition strategy S as a mapping from such sequences of expected values and division vectors to signal choices.

Consider a mapping $\tilde{G}$ from possible sequences of divisions to these sequences themselves: For each $(d(0), \ldots, d(T))$, $\tilde{G}$ maps to another sequence $(\tilde{d}(0), \ldots, \tilde{d}(T))$, subject to the following "consistency" requirements. First, $\sum_i \tilde{d}_i(t) = t$, meaning that each $\tilde{d}(t)$ must be a possible division at time t. Second, $\tilde{d}_i(t) \geq \tilde{d}_i(t-1)$, meaning that the sequence $\tilde{d}$ can be attained via a sequential sampling rule. Lastly, we require

$$(\tilde{d}(0), \ldots, \tilde{d}(T-1)) = \tilde{G}(d(0), \ldots, d(T-1))$$

so that nesting sequences are mapped to nesting sequences.

The following lemma says that if $d(\cdot)$ represents the division vectors under an information acquisition strategy S, and if $\tilde{G}$ is a consistent mapping that uniformly reduces the posterior variance, then we can find another information acquisition strategy $\tilde{S}$ whose division vectors are given by $\tilde{d}(\cdot)$. Moreover, our construction ensures that $\tilde{S}$ leads to more dispersed posterior beliefs than S at every period, so that in any decision problem, acquiring signals according to $\tilde{S}$ is weakly better than S (when actions are taken optimally).

Lemma 3. *Fix any information acquisition strategy S and any consistent mapping $\tilde{G}$ defined above. Suppose that for every sequence of divisions $(d(0), \ldots, d(T))$ realized under S, it holds that*

$$f(\tilde{d}(T)) \leq f(d(T)).$$

Then there exists deviation strategy $\tilde{S}$ such that, at every period T, any history $h^T = (\mu_1^0, d(0); \ldots; \mu_1^T, d(T))$ under S can be "associated with" a distribution of histories $\tilde{h}^T = (\nu_1^0, \tilde{d}(0); \ldots; \nu_1^T, \tilde{d}(T))$ with the following properties:

1. *the probability of h^T occurring under S is the same as the probability of its associated $\tilde{h}^T$ (integrated with respect to the probability of "association") occurring under $\tilde{S}$;*

2. *the total probability that any $\tilde{h}^T$ is associated to (integrated with respect to different possible h^T) is 1;*

3. *under the association, the distribution of ν_1^t is normal with mean μ_1^t and variance $f(d(t)) - f(\tilde{d}(t))$ for each t.*

Consequently, for any decision strategy A, there exists another decision strategy $\tilde{A}$ such that the expected payoff under $(\tilde{S}, \tilde{A})$ is no less than the expected payoff under (S, A).

To interpret, the first two properties require that the association relation is a *Markov kernel* between histories under S and histories under $\tilde{S}$; this enables us to compare payoffs under $\tilde{S}$ to those under S. The third property guarantees that the alternative strategy $\tilde{S}$ is more informative than S.

We note the following corollary, which is obtained from the previous lemma by considering a constant mapping $\tilde{G}$. Recall the t-optimal division vectors $n(t)$ from Definition 3.

Corollary 2. *Suppose each coordinate of $n(Bt)$ increases in t. Then it is possible and optimal for the DM to achieve cumulated signal count $n(Bt)$ at every period t. In these cases the myopic rule is optimal.*

B.2 Proof of Lemma 3

We construct $\tilde{S}$ iteratively as follows. In the first period, consider the signal choice under S (given the null history). This signal leads to the division $d(1)$. Let $\tilde{S}$ observe the unique signal that would achieve the division $\tilde{d}(1)$.

After the first observation, the DM's *distribution of posterior beliefs* about θ_1 under S is $\theta_1 \sim \mathcal{N}(\mu_1^1, f(d(1)))$ with μ_1^1 a normal random variable with mean μ_1^0 and variance $f(\mathbf{0}) - f(d(1))$. By comparison, the distribution of posterior beliefs under $\tilde{S}$ is $\theta_1 \sim \mathcal{N}(\nu_1^1, f(\tilde{d}(1)))$ with ν_1^1 drawn from $\mathcal{N}(\mu_1^0, f(\mathbf{0}) - f(\tilde{d}(1)))$. Since $f(\tilde{d}(1)) \leq f(d(1))$, the latter distribution of beliefs (under $\tilde{S}$) is more informative a la Blackwell. Thus, we can associate each belief $\theta_1 \sim \mathcal{N}(\mu_1^1, f(d(1)))$ under S with a more informative distribution of beliefs $\mathcal{N}(\nu_1^1, f(\tilde{d}(1)))$ under $\tilde{S}$. To be more specific, for fixed μ_1^1, the associated ν_1^1 is distributed normally with mean μ_1^1 and variance $f(d(1)) - f(\tilde{d}(1)))$. Thus by construction, all three properties are satisfied at period 1. To facilitate the discussion below, we say this distribution of beliefs under $\tilde{S}$ "imitates" the belief $(\mu_1^1, f(d(1)))$ under S.

In the second period, the deviation strategy $\tilde{S}$ takes the current belief $(\nu_1^1, f(\tilde{d}(1)))$ and randomly selects some μ_1^1 (with conditional probabilities under the Markov kernel) to "imitate." That is, given any selection of μ_1^1, find the signal that S would observe in period 2 given belief $(\mu_1^1, f(d(1)))$. This signal choice under S leads to the division sequence $(d(0), d(1), d(2))$, which is mapped to $(\tilde{d}(0), \tilde{d}(1), \tilde{d}(2))$. Naturally, we let $\tilde{S}$ observe the signal that would lead to the division $\tilde{d}(2)$. Such a signal is well-defined due to our consistency requirements on $\tilde{G}$.

To proceed with the analysis, let us fix μ_1^1 and study the distribution of posterior beliefs about θ_1 after two observations. Under S, the distribution of posterior beliefs is $\theta_1 \sim \mathcal{N}(\mu_1^2, f(d(2)))$ with μ_1^2 normally distributed with mean μ_1^1 and variance $f(d(1)) - f(d(2))$. While under $\tilde{S}$, the distribution of posterior beliefs is $\theta_1 \sim (\nu_1^2, f(\tilde{d}(2)))$ with ν_1^2 drawn from $\mathcal{N}(\mu_1^1, f(d(1)) - f(\tilde{d}(2)))$.[21]

[21] Here we use the following technical result: suppose the DM is endowed with *a distribution of prior beliefs* $\theta \sim \mathcal{N}(\mu, V)$, with μ_1 normally distributed with mean y and variance σ^2, then upon observing signal i and performing Bayesian updating, his distribution of posterior beliefs is $\theta \sim \mathcal{N}(\hat{\mu}, \hat{V})$, with $\hat{\mu}_1$ normally distributed with mean y and variance $\sigma^2 + [V - \hat{V}]_{11}$. This is proved by noting that the DM's distribution of beliefs about θ_1 must integrate to the same ex-ante distribution of θ_1.

Since $f(\tilde{d}(2)) \leq f(d(2))$, the distribution of beliefs under $\tilde{S}$ Blackwell-dominates the distribution under S, for each μ_1^1. We can thus associate each history $(\mu_1^1, d(1); \mu_1^2, d(2))$ under S with a distribution of histories $(\nu_1^1, \tilde{d}(1); \nu_1^2, \tilde{d}(2))$ under $\tilde{S}$, such that the corresponding beliefs under $\tilde{S}$ are more informative at both periods. Repeating this procedure completes the construction of $\tilde{S}$, which satisfies all three properties stated in the lemma.

Finally, suppose A is any decision strategy that maps histories to actions. We need to find $\tilde{A}$ such that the pair $(\tilde{S}, \tilde{A})$ does no worse than (S, A). This is straightforward given what we have done: at any history $\tilde{h}^T$ under $\tilde{S}$, let $\tilde{h}^T$ randomly select h^T to imitate, and define $\tilde{A}(\tilde{h}^T) = A(h^T)$. Then we see that a DM who follows the decision strategy A obtains the same payoff along any belief history h as another DM who uses the decision strategy $\tilde{A}$ and faces the distribution of belief histories $\tilde{h}$. Integrating over h, we have shown that $(\tilde{S}, \tilde{A})$ achieves the same payoff as (S, A). The lemma is proved.

C Proof of Theorem 1 (Large Block Sizes)

By Corollary 2, it suffices to show that for sufficiently large B, each coordinate $n(Bt)$ is increasing in t. To do this, we first argue that the t-optimal signal counts for different signals grow to infinity (roughly) proportionally. In more detail, define

$$\lambda_i = \frac{|[C^{-1}]_{1i}| \cdot \sigma_i}{\sum_{j=1}^{K} |[C^{-1}]_{1j}| \cdot \sigma_j}. \tag{6}$$

Then we will show that for each signal i, $n_i(t) - \lambda_i \cdot t$ remains bounded even as $t \to \infty$.

Indeed, we must at least have $n_i(t) \to \infty$; otherwise the posterior variance $f(n(t))$ would be bounded away from zero, which would contradict the optimality of $n(t)$ since $f(t/K, \dots, t/K) \to 0$. Additionally, we compute from (1) that

$$\partial_i f(n(t)) = -\frac{\sigma_i^2}{n_i^2} \cdot [V^0 C' \Sigma^{-1} \Delta_{ii} \Sigma^{-1} C V^0]_{11}. \tag{7}$$

As each $n_i \to \infty$, the matrix $\Sigma = CV^0C' + D^{-1}$ (see Lemma 1) converges to CV^0C'. So $V^0 C' \Sigma^{-1} \Delta_{ii} \Sigma^{-1} C V^0$ converges to the matrix Q_i defined in (2). It follows from (7) that $\partial_i f \sim \frac{-\sigma_i^2}{n_i^2} \cdot [Q_i]_{11}$, where $\sim$ means their ratio converges to 1. Since a t-optimal division must satisfy $\partial_i f \sim \partial_j f$ (because we are doing discrete optimization, $\partial_i f$ and $\partial_j f$ need only be approximately equal), we deduce that n_i and n_j must grow proportionally. Using $[Q_i]_{11} = ([C^{-1}]_{1i})^2$, we have $n_i(t) \sim \lambda_i t$.

Next, note that because $n_i(t) \sim \lambda_i t$, $\Sigma = CV^0C' + D^{-1} = CV^0C' + O(\frac{1}{t})$, where we follow the standard "Big O" notation for the limit $t \to \infty$. Thus in fact $V^0 C' \Sigma^{-1} \Delta_{ii} \Sigma^{-1} C V^0$

converges to Q_i at the rate of $\frac{1}{t}$. From (7), we obtain $\partial_i f = \frac{-\sigma_i^2 \cdot [Q_i]_{11} + O(\frac{1}{t})}{n_i^2}$. Optimality of the division vector $n(t)$ gives us the first-order condition $\partial_i f = \partial_j f + O(\frac{1}{t^3})$.[22] So

$$\frac{\lambda_i^2 + O(\frac{1}{t})}{n_i^2} = \frac{\lambda_j^2 + O(\frac{1}{t})}{n_j^2}.$$

This is equivalent to $\lambda_i^2 n_j^2 - \lambda_j^2 n_i^2 = O(t)$, which yields $\lambda_i n_j - \lambda_j n_i = O(1)$ after factorization. Hence $n_i(t) = \lambda_i \cdot t + O(1)$ as we claimed.

Having completed this asymptotic characterization of the t-optimal division vectors, we will now show that $n(t+K-1) \geq n(t)$ (in each coordinate) whenever t is sufficiently large. Theorem 1 will follow once this is proved.[23]

Suppose for the sake of contradiction that $n_1(t+K-1) \leq n_1(t) - 1$. Note we have $\sum_{i=1}^{K}(n_i(t+K-1) - n_i(t)) = K-1$. So $\sum_{i=2}^{K}(n_i(t+K-1) - n_i(t)) \geq K$, and we can without loss of generality assume $n_2(t+K-1) \geq n_2(t)+2$. To summarize, when transitioning from t-optimality to $t+K-1$-optimality, signal 1 is acquired at least once less and signal 2 at least twice more. Below we will obtain a contradiction by arguing that at period $t+K-1$, the posterior variance could be further reduced by observing signal 1 once more and signal 2 once less.

Indeed, let us write $n_i = n_i(t)$ and $\tilde{n}_i = n_i(t+K-1)$. Then t-optimality of $n(t)$ gives us

$$f(n_1 - 1, n_2 + 1, \ldots, n_K) \geq f(n_1, n_2, \ldots, n_K).$$

With a slight abuse of notation, we let $\partial_i f$ to denote the *discrete* partial derivative of f: $\partial_i f(q) = f(q_i + 1, q_{-i}) - f(q)$. Then the above display is equivalent to

$$\partial_2 f(n_1 - 1, n_2, \ldots, n_K) \geq \partial_1 f(n_1 - 1, n_2, \ldots, n_K). \tag{8}$$

We claim this implies the following:

$$\partial_2 f(\tilde{n}_1, \tilde{n}_2 - 1, \ldots, \tilde{n}_K) > \partial_1 f(\tilde{n}_1, \tilde{n}_2 - 1, \ldots, \tilde{n}_K). \tag{9}$$

This would lead to

$$f(\tilde{n}_1, \tilde{n}_2, \ldots, \tilde{n}_K) > f(\tilde{n}_1 + 1, \tilde{n}_2 - 1, \ldots, \tilde{n}_K),$$

[22] Since $n(t)$ minimizes the posterior variance function f only among integer vectors, we need not have the exact equality $\partial_i f = \partial_j f$. However, error terms that arise due to discreteness are bounded by $O(\frac{1}{t^3})$.

[23] To be fully rigorous, this only proves Theorem 1 when B is sufficiently large and is a multiple of $K-1$. However, we can similarly show $n(t+K) \geq n(t)$ for sufficiently large t. The two inequalities $n(t+K-1) \geq n(t)$ and $n(t+K) \geq n(t)$ together are sufficient to deduce Theorem 1 for all large B.

which would be our desired contradiction.

It remains to show (8) $\implies$ (9). By assumption, we have $\tilde{n}_1 \leq n_1 - 1$, $\tilde{n}_2 \geq n_2 + 2$ and the difference between any $\tilde{n}_j$ and n_j is bounded uniformly over t. Thus the LHS of (9) exceeds the LHS of (8) by (at least) a second derivative ∂_{22} minus a finite number of cross partial derivatives ∂_{2j}. By Lemma 2, this difference on the LHS is positive with order $\frac{1}{t^3}$. The difference between the RHS of (9) and the RHS of (8) can be positive or negative, but either way it has order $O(\frac{1}{t^4})$. This shows (9) is a consequence of (8), and the theorem follows.

D Proof of Theorem 2 (Separable Environments)

Suppose the informational environment is separable. We will show $n(t)$ increases in t, which implies the theorem via Corollary 2.

Note that in a separable environment, the definition of t-optimality reduces to:

$$n(t) = (n_1(t), \ldots, n_K(t)) \in \operatorname*{argmin}_{(q_1,\ldots,q_K):q_i \in \mathbb{Z}_+, \sum_{i=1}^K q_i = t} \sum_{i=1}^K g_i(q_i)$$

where $g_1, \ldots, g_K$ are convex functions.

In this setting, the myopic rule sequentially chooses the signal i that minimizes the difference $g_i(q_i + 1) - g_i(q_i)$, given the current division vector q. But since the g-functions are convex, the outcome under this greedy procedure actually achieves t-optimality at every period t.[24] Hence $n(t)$ increases in t and is achieved by the myopic rule.

E Preparation for the Proof of Theorem 3

E.1 Switch Deviations

We now introduce several results that will be used to show that the optimal rule eventually proceeds myopically in generic environments. Relative to the proofs of Theorems 1 and 2, the new difficulty that arises is that in general, the optimal information acquisition strategy conditions on signal realizations. As a result, the induced division vectors $d(\cdot)$ are stochastic, and we will need the full power of our dynamic Blackwell lemma.

In what follows, we will apply Lemma 3 using a particular class of mappings $\tilde{G}$.

[24]This can be easily proved by induction.

Definition 4. *Fix a particular sequence of divisions* $(d^*(0), d^*(1), \ldots, d^*(t_0))$. *Let* i *be the signal observed in period* t_0 *and* j *be any other signal. An* (i,j)-switch *mapping* $\tilde{G}$ *specifies the following:*

1. *Suppose* $T < t_0$ *or* $d(t) \neq d^*(t)$ *for some* $t \leq t_0$, *then let* $\tilde{G}(d(0), \ldots, d(T))$ *be itself.*

2. *Otherwise* $T \geq t_0$ *and* $d(t) = d^*(t), \forall t \leq t_0$. *If* $d_j(T) = d_j(t_0)$, *then let* $\tilde{d}(T) = (d_i(T) - 1, d_j(T) + 1, d_{-ij}(T))$. *If* $d_j(T) > d_j(t_0)$, *then let* $\tilde{d}(T) = d(T)$.

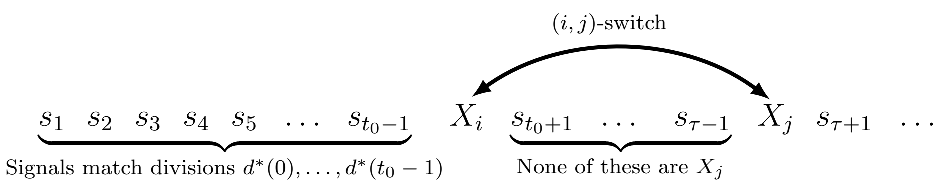

Figure 1: *Pictorial representation of an* (i,j)*-switch based on a sequence of divisions* $d^*(0), \ldots, d^*(t_0)$.

Let us interpret this definition by relating to the resulting deviation strategy $\tilde{S}$ constructed in Lemma 3. The first case above says that $\tilde{S}$ only deviates when the history of divisions is $d^*(0), \ldots, d^*(t_0 - 1)$ and S is about to observe signal i in period t_0. The second case says that $\tilde{S}$ dictates observing signal j instead at that history; subsequently, $\tilde{S}$ observes the same signal as S (at the imitated belief) until the first period at which S is about to observe signal j. If that period exists, the deviation strategy $\tilde{S}$ *switches back* to observing signal i and coincides with S afterwards.

The benefit of these "switch deviations" is that their posterior variances can be easily compared to the original strategy. Specifically, $\tilde{d}(t) = d(t)$ except at those histories that begin with $d^*(0), d^*(1), \ldots, d^*(t_0 - 1)$ (and before signal j is observed again under S). At such histories, the posterior variance is strictly lower under $\tilde{S}$ if and only if

$$f(d_i(t) - 1, d_j(t) + 1, d_{-ij}(t)) < f(d(t)).$$

Using (absolute values of) the discrete partial derivatives, we can rewrite this condition as

$$|\partial_i f(d_i(t) - 1, d_j(t), d_{-ij}(t))| < |\partial_j f(d_i(t) - 1, d_j(t), d_{-ij}(t))|. \qquad (*)$$

We can thus obtain the following corollary:

Corollary 3. *Suppose we can find a history of divisions* $d(0), \ldots, d(t_0)$ *realized under* S *such that* $d_i(t_0) = d_i(t_0 - 1) + 1$ *and moreover (*) holds for all divisions* $d(t)$ *with* $d_j(t) = d_j(t_0)$ *and* $d_k(t) \geq d_k(t_0), \forall k$. *Then the switch deviation* $\tilde{S}$ *constructed above improves upon* S.

Note that the condition $d_j(t) = d_j(t_0)$ captures the fact that $\tilde{d}(t)$ differs from $d(t)$ only until signal j is chosen again by S. Meanwhile, $d_k(t) \geq d_k(t_0), \forall k$ holds because we only compare posterior variances after t_0 periods.

E.2 Asymptotic Characterization of Optimal Strategy

Below we will use the contrapositive of Corollary 3 to argue that if S is the optimal information acquisition strategy, then we cannot find any history of realized divisions such that (*) always holds. Technically speaking, we might worry that although $\tilde{S}$ strictly improves upon S in terms of posterior variances, it might achieve the same expected payoff as S (for instance, when the DM faces a constant payoff function). Nonetheless, by Zorn's lemma we can choose S to be an optimal strategy that is additionally "un-dominated" in terms of posterior variances. With that choice, the deviation $\tilde{S}$ cannot exist, and our arguments remain valid.

To illustrate, we now derive the asymptotic signal proportions for the optimal information acquisition strategy S.

Lemma 4. *Suppose S is the optimal information acquisition strategy, and $d(\cdot)$ is its induced divisions. Let λ_k be defined as in (6). In generic informational environments, the difference $d_k(T) - \lambda_k \cdot T$ remains bounded as $T \to \infty$, for any realized division $d(T)$ and each signal k.*

Proof. For this proof, we only need the informational environment to be such that each signal has *strictly positive marginal value*. That is, for any signal k and any possible division q, we require

$$f(q_k + 1, q_{-k}) < f(q).$$

This is "generically" satisfied because any equality $f(q_k + 1, q_{-k}) = f(q)$ would impose a non-trivial polynomial equation over the signal linear coefficients, and the number of such constraints is at most countable.

Under this genericity assumption, let us first show $d_k(T) \to \infty$ holds for each signal k, and the speed of divergence depends only on the informational environment. For contradiction, suppose this is not true. Then we can find a sequence of histories $\{h^{T_m}\}$ such that $T_m \to \infty$ but $d_1(T_m)$ remains bounded (these histories need not nest one another). By passing to a subsequence, we may assume $q_k = \lim_{m\to\infty} d_k(T_m)$ exists for every signal k, where this limit may be infinity. Define I to be the non-empty subset of signals (not including signal 1) with $q_k = \infty$. Furthermore, we assume that the signal observed in the last period of each of these histories h^{T_m} is the same signal i. We also assume $i \in I$; otherwise just truncate the histories by finitely many periods.

Take any signal $j \notin I$ (for instance, $j = 1$ works). Choose T_m sufficiently large and consider the (i, j)-switch deviation $\tilde{S}$ that deviates from h^{T_m} by observing signal j instead of i in period T_m. We will verify (*) for all possible divisions $d(t)$ with $d_j(t) = d_j(T_m)$ and $d_i(t) \geq d_i(T_m)$, which will contradict the optimality of S via Corollary 3. Indeed, note that as $T_m \to \infty$, $d_i(T_m) \to \infty$ because $i \in I$. Since $d_i(t) \geq d_i(T_m)$, the LHS of (*) approaches zero as T_m increases. By comparison, the RHS of (*) is bounded away from zero because $d_j(t) = d_j(T_m)$ is bounded, and we assume each signal has strictly positive marginal value. Hence (*) holds and we have shown that $d(T) \to \infty$ in each coordinate.

Next, from (7), we have the following approximations for the partial derivatives:

$$|\partial_i f(d_i(t) - 1, d_j(t), d_{-ij}(t))| \sim \frac{\sigma_i^2 \cdot [Q_i]_{11}}{d_i(t)^2} \qquad |\partial_j f(d_i(t) - 1, d_j(t), d_{-ij}(t))| \sim \frac{\sigma_j^2 \cdot [Q_j]_{11}}{d_j(t)^2}.$$

If $\limsup_{t_0 \to \infty} \frac{d_i(t_0)}{d_j(t_0)} > \frac{\lambda_i}{\lambda_j}$ (recall that λ_i is proportional to $\sigma_i \cdot \sqrt{[Q_i]_{11}}$), then the above estimates would imply (*) whenever $d_i(t) \geq d_i(t_0)$ (because $t \geq t_0$) and $d_j(t) = d_j(t_0)$. That would contradict the optimality of S. Hence, $\limsup_{t_0 \to \infty} \frac{d_i(t_0)}{d_j(t_0)} \leq \frac{\lambda_i}{\lambda_j}$ for every pair of signals i and j. It follows that $d_k(t_0) \sim \lambda_k \cdot t_0, \forall k$.

Once these asymptotic proportions are proved, we know that the matrix $\Sigma = CV^0C' + D^{-1}$ converges to CV^0C' at the rate of $\frac{1}{t}$. By (7), we can deduce more precise approximations:

$$|\partial_i f(d_i(t) - 1, \cdots)| = \frac{\sigma_i^2 \cdot [Q_i]_{11} + O(\frac{1}{t})}{d_i(t)^2} \qquad |\partial_j f(d_i(t) - 1, \cdots)| = \frac{\sigma_j^2 \cdot [Q_j]_{11} + O(\frac{1}{t})}{d_j(t)^2}.$$

If $\frac{d_i(t_0)}{d_j(t_0)} > \frac{\lambda_i}{\lambda_j} + O(\frac{1}{t_0})$, then these refined estimates would again imply (*) whenever $d_i(t) \geq d_i(t_0)$ and $d_j(t) = d_j(t_0)$. To avoid the resulting contradiction, we must have $\frac{d_i(t_0)}{d_j(t_0)} \leq \frac{\lambda_i}{\lambda_j} + O(\frac{1}{t_0})$ for every signal pair. This enables us to conclude $d_k(t_0) = \lambda_k \cdot t_0 + O(1)$ as desired. □

F Proof of Theorem 3 (Generic Eventual Myopia)

F.1 Outline

To guide the reader through this appendix, we begin by outlining the proof of the theorem, which is broken down into several steps. Throughout, we focus on the case of $B = 1$ (one signal each period), but our proof easily extends to arbitrary B. We will first show a simpler (and weaker) result that, in generic environments, the number of periods in which the optimal strategy differs from the t-optimal division has natural density 1. Our proof of this result is based on the observation that if equivalence does not hold at some time t, there must be *two different divisions* over signals for which the resulting posterior variances about θ_1 are within $O(\frac{1}{t^4})$ from each other. This leads to a Diophantine approximation inequality, which we can show only occurs at a vanishing fraction of periods t.

To improve the result and demonstrate equivalence at *all late periods*, we show that the number of "exceptional periods" t is generically finite if there are *three different divisions* over signals whose posterior variances are within $O(\frac{1}{t^4})$ from each other. This allows us to conclude that in generic environments, the t-optimal divisions eventually monotonically increase in t.

In such environments, t-optimality can be achieved at every late period. Thus, whenever t-optimality obtains in *some* late period, it will be sustained in all future periods. Since we have already established that the optimal strategy achieves t-optimality infinitely often, we conclude equivalence at all large t.

We mention that in this appendix, we use a slightly different notion of "generic" where we fix the signal coefficient matrix C and instead (randomly) vary the signal variances $\{\sigma_i^2\}$. This concept implies (and is stronger than) the previous genericity concept defined on C.

F.2 Equivalence at Almost All Times

We begin by proving a weaker result, that the optimal strategy induces the t-optimal division $n(t)$ at almost all periods t.

Proposition 2. *Suppose the informational environment $(V^0, C, \{\sigma_i^2\})$ has the property that for any $i \neq j$, the ratio $\frac{\lambda_i}{\lambda_j}$ is an irrational number. Then, at a set of times with natural density 1,*[25] *$d(t) = n(t)$ (which is unique) holds for every decision problem. In particular, the optimal strategy induces a deterministic division vector at such times.*

Proof of Proposition 2. Suppose that $d_1(T) \geq n_1(T) + 1$ and $d_2(T) \leq n_2(T) - 1$. Consider the last period $t_0 \leq T$ in which the optimal strategy observed signal 1. Then

$$d_1(t_0) = d_1(T) \geq n_1(T) + 1; \quad d_2(t_0) \leq d_2(T) \leq n_2(T) - 1.$$

Using the contrapositive of Corollary 3 with the $(1,2)$-switch, we know that (*) cannot always hold. Thus there exists a division $d(t)$ such that the inequality (*) is reversed. That

[25] Formally, for any set of positive integers A, let $A(N)$ count the number of integers in A no greater than N. Then we define the natural density of A to be $\lim_{N\to\infty} \frac{A(N)}{N}$, when this limit exists.

is, we can find a division $d(t)$ with $d_1(t) \geq d_1(t_0)$ and $d_2(t) = d_2(t_0)$ such that (getting rid of the absolute values)

$$\partial_1 f(d_1(t) - 1, d_2(t), d_{-12}(t)) \leq \partial_2 f(d_1(t) - 1, d_2(t), d_{-12}(t)). \tag{10}$$

On the other hand, t-optimality of $n(t)$ gives us

$$\partial_1 f(n_1(T), n_2(T) - 1, n_{-12}(T)) \geq \partial_2 f(n_1(T), n_2(T) - 1, n_{-12}(T)). \tag{11}$$

Note that $d_2(t) = d_2(t_0)$ implies $t - t_0$ is bounded (due to Lemma 4). On the other hand, we have $d_1(t) = d_1(T_0)$ by construction (t_0 is the last period signal 1 was observed). Hence $t_0 - T$ is also bounded. Combining both, we deduce $t - T$ must be bounded. Applying Lemma 4 again, we know that any difference $d_i(t) - n_i(T)$ is bounded.

Now because $d_1(t) - 1 \geq d_1(t_0) - 1 \geq n_1(T)$, the LHS of (10) has size at least the LHS of (11) minus a finite number of cross partial derivatives ∂_{1j}. Similarly, the RHS of (10) is at most bigger than the RHS of (10) by a number of cross partials. Together with the order difference lemma, these imply that the only way (10) and (11) can both hold is if the two sides of (11) differ by at most $O\left(\frac{1}{T^4}\right)$.

To summarize: A necessary condition for $d_1(T) \geq n_1(T) + 1$ and $d_2(T) \leq n_2(T) + 1$ to occur is that

$$|f(n_1(T) + 1, n_2(T) - 1, \ldots, n_K(T)) - f(n(T))| = O\left(\frac{1}{T^4}\right). \tag{12}$$

Hence, to prove the Proposition we only need to show that (12) holds at a set of times with natural density 0. The following lemma proves exactly this property. □

Lemma 5. *Suppose $\frac{\lambda_1}{\lambda_2}$ is an irrational number. For positive constants c_0, c_1, define $\mathcal{A}(c_0, c_1)$ to be the following set of positive integers:*

$$\{t : \exists\ q_1, q_2, \ldots, q_K \in \mathbb{Z}_+, s.t.\ |q_i - \lambda_i \cdot t| \leq c_0, \forall i$$
$$\wedge\ \ |f(q_1, q_2 + 1, \ldots, q_K) - f(q_1 + 1, q_2, \ldots, q_K)| \leq c_1/t^4\}.$$

Then $\mathcal{A}(c_0, c_1)$ has natural density zero.

Proof of Lemma 5. The proof relies on the following technical result, which gives a precise approximation of the *discrete partial derivatives* of f:

Lemma 6. *Fix the informational environment. There exists a constant a_j such that*

$$f(q_j, q_{-j}) - f(q_j + 1, q_{-j}) = \frac{\sigma_j^2 \cdot [Q_j]_{11}}{(q_j - a_j)^2} + O\left(\frac{1}{t^4}; c_0\right) \tag{13}$$

holds for all $q_1, \ldots, q_K$ with $|q_i - \lambda_i t| \leq c_0, \forall i$. The notation $O\left(\frac{1}{t^4}; c_0\right)$ means an upper bound of $\frac{L}{t^4}$, where the constant L may depend on the informational environment as well as on c_0.[26]

Assuming (13), we see that the condition

$$|f(q_1, q_2 + 1, \ldots, q_K) - f(q_1 + 1, q_2, \ldots, q_K)| \leq \frac{c_1}{t^4}$$

implies $\left|\frac{\sigma_1^2 \cdot [Q_1]_{11}}{(q_1 - a_1)^2} - \frac{\sigma_2^2 \cdot [Q_2]_{11}}{(q_2 - a_2)^2}\right| \leq \frac{c_2}{t^4}$ and thus $\left|(\frac{\lambda_1}{q_1 - a_1})^2 - (\frac{\lambda_2}{q_2 - a_2})^2\right| \leq \frac{c_3}{t^4}$ for some larger positive constants c_2, c_3. This further implies $\left|\frac{\lambda_1}{q_1 - a_1} - \frac{\lambda_2}{q_2 - a_2}\right| \leq \frac{c_4}{t^3}$, which reduces to

$$\left|q_2 - a_2 - \frac{\lambda_2}{\lambda_1}(q_1 - a_1)\right| \leq \frac{c_5}{t}. \tag{14}$$

This inequality says that the fractional part of $\frac{\lambda_2}{\lambda_1} q_1$ is very close to the fractional part of $\frac{\lambda_2}{\lambda_1} a_1 - a_2$. But since $\frac{\lambda_2}{\lambda_1}$ is an irrational number, the fractional part of $\frac{\lambda_2}{\lambda_1} q_1$ is "equi-distributed" in (0,1) as q_1 ranges in the positive integers.[27] Thus the Diophantine approximation (14) only has solution at a set of times t with natural density 0, proving Lemma 5. Below we supply the technically involved proof of (13). □

Proof of Lemma 6. Fix $q_1, \ldots, q_K$ and the signal j. Recall the diagonal matrix $D = \text{diag}(\frac{q_1}{\sigma_1^2}, \ldots, \frac{q_K}{\sigma_K^2})$. Consider any $\hat{q}_j \in [q_j, q_j + 1]$ and let $\hat{D}$ be the analogue of D for the division $(\hat{q}_j, q_{-j})$. That is, $\hat{D} = D$ except that $[\hat{D}]_{jj} = \frac{\hat{q}_j}{\sigma_j^2}$. Let $\hat{\Sigma} = CV^0C' + \hat{D}^{-1}$. From (7), we have

$$\partial_j f(\hat{q}_j, q_{-j}) = -\frac{\sigma_j^2}{\hat{q}_j^2} \cdot \left[V^0 C' \hat{\Sigma}^{-1} \Delta_{jj} \hat{\Sigma}^{-1} C V^0\right]_{11}. \tag{15}$$

Here and later in this proof, $\partial_j f$ represents the usual continuous derivative rather than the discrete derivative.

Let $D_0 = \text{diag}\left(\frac{\lambda_1 t}{\sigma_1^2}, \ldots, \frac{\lambda_K t}{\sigma_K^2}\right)$ and $\Sigma_0 = CV^0C' + D_0^{-1}$. For $|q_i - \lambda_i t| \leq c_0, \forall i$ we have $\hat{D} - D_0 = O(c_0)$, where the Big O notation applies entry-wise. It follows that

$$\hat{\Sigma} = CV^0C' + \hat{D}^{-1} = CV^0C' + D^{-1} + O(\frac{1}{t^2}; c_0) = \Sigma_0 + O(\frac{1}{t^2}; c_0).$$

[26] In applying Lemma 5 to prove Proposition 2, c_0 is taken to be the bound on $n_i - \lambda_i \cdot t$.

[27] The Equi-distribution Theorem states that for any irrational number α and any sub-interval $(a, b) \subset (0, 1)$, the set of positive integers n such that the fractional part of αn belongs to (a, b) has natural density $b - a$. It is a special case of the Ergodic Theorem.

Observe that the matrix inverse is a differentiable mapping at Σ_0 (which is $CV^0C' + D_0^{-1} \succeq CV^0C'$ and thus positive definite). Thus we have

$$\hat{\Sigma}^{-1} = \Sigma_0^{-1} + O\left(\frac{1}{t^2}; c_0\right).$$

Plugging this into (15) and using $\hat{q}_j \sim \lambda_j t$, we obtain that

$$\partial_j f(\hat{q}_j, q_{-j}) = -\frac{\sigma_j^2}{\hat{q}_j^2} \cdot \left[V^0C'\Sigma_0^{-1}\Delta_{jj}\Sigma_0^{-1}CV^0\right]_{11} + O\left(\frac{1}{t^4}; c_0\right). \tag{16}$$

Since $\Sigma_0 = CV^0C' + \frac{1}{t} \cdot \operatorname{diag}\left(\frac{\sigma_1^2}{\lambda_1}, \ldots, \frac{\sigma_K^2}{\lambda_K}\right)$, we can apply Taylor expansion (to the matrix inverse map) and write

$$\Sigma_0^{-1} = (CV^0C')^{-1} - \frac{1}{t}(CV^0C')^{-1} \cdot \operatorname{diag}\left(\frac{\sigma_1^2}{\lambda_1}, \ldots, \frac{\sigma_K^2}{\lambda_K}\right) \cdot (CV^0C')^{-1} + O\left(\frac{1}{t^2}\right). \tag{17}$$

This implies

$$\begin{aligned} V^0C'\Sigma_0^{-1}\Delta_{jj}\Sigma_0^{-1}CV^0 &= V^0C'(CV^0C')^{-1}\Delta_{jj}(CV^0C')^{-1}CV^0 - \frac{M_j}{t} + O\left(\frac{1}{t^2}\right) \\ &= Q_j - \frac{M_j}{t} + O\left(\frac{1}{t^2}\right), \end{aligned} \tag{18}$$

where M_j is a fixed $K \times K$ matrix depending only on the informational environment. For future use, we note that

$$\begin{aligned} M_j = {} & V^0C'(CV^0C')^{-1} \operatorname{diag}\left(\frac{\sigma_1^2}{\lambda_1}, \ldots, \frac{\sigma_K^2}{\lambda_K}\right)(CV^0C')^{-1}\Delta_{jj}(CV^0C')^{-1}CV^0 \\ & + V^0C'(CV^0C')^{-1}\Delta_{jj}(CV^0C')^{-1} \operatorname{diag}\left(\frac{\sigma_1^2}{\lambda_1}, \ldots, \frac{\sigma_K^2}{\lambda_K}\right)(CV^0C')^{-1}CV^0 \\ = {} & C^{-1} \operatorname{diag}\left(\frac{\sigma_1^2}{\lambda_1}, \ldots, \frac{\sigma_K^2}{\lambda_K}\right)(CV^0C')^{-1}\Delta_{jj}C'^{-1} \\ & + C^{-1}\Delta_{jj}(CV^0C')^{-1} \operatorname{diag}\left(\frac{\sigma_1^2}{\lambda_1}, \ldots, \frac{\sigma_K^2}{\lambda_K}\right)C'^{-1}. \end{aligned} \tag{19}$$

Using (18), we can simplify (16) to

$$\partial_j f(\hat{q}_j, q_{-j}) = -\frac{\sigma_j^2}{\hat{q}_j^2} \cdot \left[Q_j - \frac{M_j}{t}\right]_{11} + O\left(\frac{1}{t^4}; c_0\right). \tag{20}$$

Integrating this over $\hat{q}_j \in [q_j, q_j + 1]$, we conclude that

$$f(q_j, q_{-j}) - f(q_j + 1, q_{-j}) = \frac{\sigma_j^2}{q_j(q_j + 1)} \cdot \left[Q_j - \frac{M_j}{t}\right]_{11} + O\left(\frac{1}{t^4}; c_0\right). \tag{21}$$

We set $a_j = -\left(\frac{\lambda_j \cdot [M_j]_{11}}{2[Q_j]_{11}} + \frac{1}{2}\right)$. Then

$$\frac{\sigma_j^2}{q_j(q_j+1)} \cdot \left[Q_j - \frac{M_j}{t}\right]_{11} = (\sigma_j^2 \cdot [Q_j]_{11}) \cdot \frac{1 + \frac{2a_j+1}{\lambda_j t}}{q_j(q_j+1)} = \frac{\sigma_j^2 \cdot [Q_j]_{11}}{(q_j - a_j)^2} + O\left(\frac{1}{t^4}; c_0\right),$$

implying the desired approximation (13). The last equality above uses $\frac{1+\frac{2a_j+1}{\lambda_j t}}{q_j(q_j+1)} = \frac{1}{(q_j-a_j)^2} + O\left(\frac{1}{t^4}; c_0\right)$, which is because

$$\frac{q_j(q_j+1)}{(q_j-a_j)^2} = 1 + \frac{2(a_j+1)}{q_j - a_j} + O\left(\frac{1}{(q_j-a_j)^2}\right) = 1 + \frac{2a_j+1}{\lambda_j t} + O\left(\frac{1}{t^2}; c_0\right)$$

dividing through by $q_j(q_j+1)$. $\square$

F.3 Simultaneous Diophantine Approximation

The above Lemma 5 tells us that at most times t, there do not exist a *pair* of divisions (differing minimally on two signal counts) that lead to posterior variances close to each other (with a difference of $\frac{c_1}{t^4}$). We obtain a stronger result if a *triple* of such divisions were to exist.

Lemma 7. *Fix V^0 and C, and let signal variances vary. For positive constants c_0, c_1, define $\mathcal{A}^*(c_0, c_1)$ to be the following set of positive integers:*

$$\begin{aligned}\{t : \exists\ q_1, q_2, q_3, \ldots, q_K \in \mathbb{Z}_+, s.t.\ |q_i - \lambda_i t| \le c_0, \forall i \\ \wedge\ |f(q_1, q_2+1, q_3, \ldots, q_K) - f(q_1+1, q_2, q_3, \ldots, q_K)| \le c_1/t^4 \\ \wedge\ |f(q_1, q_2, q_3+1, \ldots, q_K) - f(q_1+1, q_2, q_3, \ldots, q_K)| \le c_1/t^4\}\end{aligned}$$

Then, for generic signal variances, $\mathcal{A}^*(c_0, c_1)$ has finite cardinality.

Proof. So far we have been dealing with fixed informational environments. However, a number of parameters defined above depend on the signal variances $\sigma = \{\sigma_i^2\}_{i=1}^K$. Specifically, while the matrix $Q_i = C^{-1}\Delta_{ii}C'^{-1}$ is independent of σ, the asymptotic proportions λ_i (which is proportional to $\sigma_i \cdot [Q_i]_{11}$) do vary with σ. In this proof, we write $\lambda_i(\sigma)$ to highlight this dependence.

Next, we recall the matrix M_j introduced earlier in (19). We note that for fixed matrices V^0 and C, each entry of $M_j(\sigma)$ is a fixed linear combination of $\frac{\sigma_1^2}{\lambda_1(\sigma)}, \ldots, \frac{\sigma_K^2}{\lambda_K(\sigma)}$.

Then, the parameter $a_j(\sigma)$ in (13) is given by (see the previous proof)

$$a_j(\sigma) = -\frac{1}{2} - \frac{\lambda_j(\sigma) \cdot [M_j(\sigma)]_{11}}{2[Q_j]_{11}} = -\frac{1}{2} + \lambda_j(\sigma) \sum_{i=1}^K \tilde{b}_{ji} \frac{\sigma_i^2}{\lambda_i(\sigma)} = -\frac{1}{2} + \sum_{i=1}^K b_{ji}\sigma_i\sigma_j \qquad (22)$$

for some constants $\tilde{b}_{ji}, b_{ji}$ independent of σ. In the last equality above, we used the fact that $\frac{\lambda_j(\sigma)}{\lambda_i(\sigma)}$ equals a constant times $\frac{\sigma_j}{\sigma_i}$.

Thus Lemma 6 gives

$$f(q_j, q_{-j}) - f(q_j + 1, q_{-j}) = \frac{\sigma_j^2 \cdot [Q_j]_{11}}{(q_j - a_j(\sigma))^2} + O\left(\frac{1}{t^4}; c_0\right)$$

whenever $|q_i - \lambda_i(\sigma) \cdot t| \leq c_0, \forall i$. We comment that the Big O constant here may depend on σ. However, a single constant suffices if we restrict each σ_i to be bounded above and bounded away from zero. Since measure-zero sets are closed under countable unions, this restriction does not affect the result we want to prove.

By the above approximation, a necessary condition for $t \in \mathcal{A}^*(c_0, c_1)$ is that q_1, q_2, q_3 satisfy

$$\left|(q_2 - a_2(\sigma)) - \frac{\eta \cdot \sigma_2}{\sigma_1}(q_1 - a_1(\sigma))\right| \leq \frac{c_6}{q_1} \tag{23}$$

as well as

$$\left|(q_3 - a_3(\sigma)) - \frac{\kappa \cdot \sigma_3}{\sigma_1}(q_1 - a_1(\sigma))\right| \leq \frac{c_6}{q_1} \tag{24}$$

for some constant c_6 independent of σ (c_6 may depend on c_0, c_1 stated in the lemma). The constant η is given by $\eta = \sqrt{[Q_2]_{11}/[Q_1]_{11}}$, and similarly for κ.

It remains to show that for generic σ, there are only finitely many positive integer triples (q_1, q_2, q_3) satisfying the *simultaneous Diophantine approximation* (23) and (24). To prove this, we assume that each σ_i is i.i.d. drawn from the uniform distribution on $[\frac{1}{L}, L]$, where L is a large constant. Denote by $F(q_1, q_2, q_3)$ the event that (23) and (24) hold simultaneously. We claim that there exists a constant c_7 such that $\mathbb{P}(F(q_1, q_2, q_3)) \leq \frac{c_7}{q_1^4}$ holds for all q_1, q_2, q_3. Since $F(q_1, q_2, q_3)$ cannot occur for $q_2, q_3 > c_8 q_1$, this claim will imply

$$\sum_{q_1, q_2, q_3} \mathbb{P}(F(q_1, q_2, q_3)) < \sum_{q_1} \sum_{q_2, q_3 \leq c_8 q_1} \frac{c_7}{q_1^4} < \sum_{q_1} \frac{c_7 c_8^2}{q_1^2} < \infty. \tag{25}$$

Generic finiteness of tuples (q_1, q_2, q_3) will then follow from the Borel-Cantelli Lemma.[28]

To prove this claim, it suffices to show that if $\sigma = (\sigma_1, \sigma_2, \sigma_3, \sigma_4, \ldots, \sigma_K)$ and $\sigma' = (\sigma_1, \sigma_2', \sigma_3', \sigma_4, \ldots, \sigma_K)$ both satisfy (23) and (24), then $|\sigma_2 - \sigma_2'|, |\sigma_3 - \sigma_3'| \leq \frac{c}{q_1^2}$ for some

[28] Because of the use of Borel-Cantelli Lemma, this proof (unlike Lemma 5 above) does not allow us to effectively determine for given σ whether (23) and (24) only have finitely many integer solutions. Nonetheless, a modification of this proof does imply the following finite-time probabilistic statement: when $\sigma_1, \ldots, \sigma_K$ are independently drawn, the probability that the optimality strategy coincides with t-optimality at every period $t \geq T$ is at least $1 - O(\frac{1}{T})$, where the constant involved only depends on the distribution of σ.

constant c.[29] Without loss, we assume $|\sigma_2 - \sigma_2'| \geq |\sigma_3 - \sigma_3'|$. Using (22), we can rewrite the condition (23) as

$$\Bigg| \underbrace{\left(q_2 + \frac{1}{2}\right) - \frac{\eta \cdot \sigma_2}{\sigma_1}\left(q_1 + \frac{1}{2}\right)}_{A} + \underbrace{\sum_i \beta_i \sigma_2 \sigma_i}_{B} \Bigg| \leq \frac{c_6}{q_1}$$

for some constants β_i independent of σ. A similar inequality holds at σ':

$$\Bigg| \underbrace{\left(q_2 + \frac{1}{2}\right) - \frac{\eta \cdot \sigma_2'}{\sigma_1}\left(q_1 + \frac{1}{2}\right)}_{A'} + \underbrace{\sum_i \beta_i \sigma_2' \sigma_i'}_{B'} \Bigg| \leq \frac{c_6}{q_1}.$$

It follows from the above two inequalities that $|A + B - A' - B'| \leq \frac{2c_6}{q_1}$. Furthermore, since $|A - A'| \leq |A + B - A' - B'| + |B - B'|$ (by triangle inequality), we deduce

$$\left| \frac{\eta \cdot (\sigma_2' - \sigma_2)}{\sigma_1} \cdot \left(q_1 + \frac{1}{2}\right) \right| \leq \frac{2c_6}{q_1} + \left| \sum_i \beta_i (\sigma_2' \sigma_i' - \sigma_2 \sigma_i) \right|. \tag{26}$$

Because $\sigma_i' = \sigma_i$ for $i \neq 2, 3$, we have

$$\begin{aligned} \left| \sum_i \beta_i (\sigma_2' \sigma_i' - \sigma_2 \sigma_i) \right| &= \left| \sum_i \beta_i (\sigma_2' - \sigma_2) \sigma_i + \sum_i \beta_i \sigma_2' (\sigma_i' - \sigma_i) \right| \\ &= \left| \left(\sum_i \beta_i (\sigma_2' - \sigma_2) \sigma_i \right) + \beta_2 \sigma_2' (\sigma_2' - \sigma_2) + \beta_3 \sigma_2' (\sigma_3' - \sigma_3) \right| \\ &\leq (K+2) L \cdot \max_i |\beta_i| \cdot |\sigma_2' - \sigma_2| \,. \end{aligned}$$

Plugging this estimate into (26), we obtain the desired result $|\sigma_2 - \sigma_2'| \leq \frac{c}{q_1^2}$. This completes the proof of the lemma. □

F.4 Monotonicity of t-Optimal Divisions

We apply Lemma 7 to prove the eventual monotonicity of t-optimal divisions in generic informational environments.

Proposition 3. *Fix V^0 and C. For generic signal variances $\{\sigma_i^2\}_{i=1}^K$, there exists T_0 such that for $t \geq T_0$, the t-optimal division $n(t)$ is unique, and it satisfies $n_i(t+1) \geq n_i(t), \forall i$.*

[29] This implies that the probability of the event $F(q_1, q_2, q_3)$ conditional on any value of $\sigma_1, \sigma_4, \ldots, \sigma_K$ is bounded by $\frac{c_7}{q_1^4}$, which is stronger than the claim.

Proof. Uniqueness follows from the stronger fact that in generic informational environments, $f(q_1, \ldots, q_K)$ differs from $f(q'_1, \ldots, q'_K)$ whenever $q \neq q'$. Below we focus on monotonicity.

Using the order difference lemma, we can already deduce the difference $|n_i(t+1) - n_i(t)|$ is no more than 1 at sufficiently late periods t. Suppose that $n_1(t+1) = n_1(t) - 1$. Then because $\sum_i (n_i(t+1) - n_i(t)) = 1$, we can without loss assume $n_2(t+1) = n_2(t) + 1$ and $n_3(t+1) = n_3(t) + 1$.

For notational ease, write $n_i = n_i(t), n'_i = n_i(t+1)$. By t-optimality, we have

$$f(n_1, n_2, n_3, \ldots, n_K) \leq f(n_1 - 1, n_2 + 1, n_3, \ldots, n_K)$$

$$f(n'_1, n'_2, n'_3, \ldots, n'_K) \leq f(n'_1 + 1, n'_2 - 1, n'_3, \ldots, n'_K)$$

These inequalities are equivalent to

$$\partial_2 f(n_1 - 1, n_2, n_3, \ldots, n_K) \geq \partial_1 f(n_1 - 1, n_2, n_3, \ldots, n_K) \tag{27}$$

$$\partial_2 f(n'_1, n'_2 - 1, n'_3, \ldots, n'_K) \leq \partial_1 f(n'_1, n'_2 - 1, n'_3, \ldots, n'_K) \tag{28}$$

with $\partial_i f$ representing the *discrete partial derivative.*

Since $n'_2 - 1 = n_2$, the LHS of (28) is at least the LHS of (27) minus a number of cross partials. Similarly, the RHS of (28) is at most bigger than the RHS of (27) by a number of cross partials. Thus the only way (27) and (28) can both hold is if the two sides of (27) differ by no more than $O(\frac{1}{t^4})$. That is, for some absolute constant c_1,[30] we have

$$|f(n_1 - 1, n_2 + 1, n_3, \ldots, n_K) - f(n_1, n_2, n_3, \ldots, n_K)| \leq \frac{c_1}{t^4}. \tag{29}$$

An analogous argument yields

$$|f(n_1 - 1, n_2, n_3 + 1, \ldots, n_K) - f(n_1, n_2, n_3, \ldots, n_K)| \leq \frac{c_1}{t^4}. \tag{30}$$

But now we can apply Lemma 7 to show that in generic environments, there are only finitely many integer tuples $(n_1, \ldots, n_K)$ that satisfy both (29) and (30). This proves the result. □

F.5 Completing the Proof of Theorem 3

By Proposition 3, generically there exists T_0 such that $n(t)$ is monotonic in t after T_0 periods. Thus, using our dynamic Blackwell lemma, if the DM achieves t-optimality at some period $t \geq T_0$, he will continue to do so in the future. By Proposition 2, such a time t does exist. This proves Theorem 3.

[30] As discussed in the proof of Lemma 7, we can find a single constant c_1 that works for all σ bounded above and bounded away from zero.

G Proof of Proposition 1 (Bound on B)

G.1 Some Estimates

Throughout, we work with the transformed model, where each signal X_i is simply $\tilde{\theta}_i$ plus standard Gaussian noise, and the DM's prior covariance matrix over the transformed states is $\tilde{V}$. Let $\gamma = \gamma(q_1, \ldots, q_K)$ represent the following $K \times 1$ vector:

$$\gamma = (\tilde{V} + E)^{-1} \cdot \tilde{V} \cdot \tilde{\alpha} \tag{31}$$

with $E = \mathrm{diag}(\frac{1}{q_1}, \ldots, \frac{1}{q_K})$ and $\omega = \tilde{\alpha}' \cdot \tilde{\theta}$. For $1 \leq i \leq K$, γ_i is the i-th coordinate of γ.

Here we re-derive the posterior variance function f, its derivatives and second derivatives. Our formulae below take as primitives $\tilde{V}$ and $\tilde{\alpha}$, but they are equivalent to those presented in Appendix A (for the original model).

Fact 1 (Posterior Variance)**.** $f(q_1, \ldots, q_K) = \tilde{\alpha}'(\tilde{V} - \tilde{V}(\tilde{V} + E)^{-1}\tilde{V})\tilde{\alpha}$.

Fact 2 (Partial Derivatives of Posterior Variance)**.** $\partial_i f(q_1, \ldots, q_K) = -\frac{1}{q_i^2} \cdot \tilde{\alpha}'\tilde{V}(\tilde{V}+E)^{-1}\Delta_{ii}(\tilde{V}+E)^{-1}\tilde{V}\tilde{\alpha} = -\frac{\gamma_i^2}{q_i^2}$.

Fact 3 (Second-Order Partial Derivatives of Posterior Variance)**.**

$$\begin{aligned}
&\partial_{ii} f(q_1, \ldots, q_K) \\
&\quad= \frac{2 \cdot \tilde{\alpha}'\tilde{V}(\tilde{V} + E)^{-1}\Delta_{ii}(\tilde{V} + E)^{-1}\tilde{V}\tilde{\alpha}}{q_i^3} - \frac{2 \cdot \tilde{\alpha}'\tilde{V}(\tilde{V} + E)^{-1}\Delta_{ii}(\tilde{V} + E)^{-1}\Delta_{ii}(\tilde{V} + E)^{-1}\tilde{V}\tilde{\alpha}}{q_i^4} \\
&\quad= \frac{2\gamma_i^2}{q_i^3} \cdot \left(1 - \frac{[(\tilde{V} + E)^{-1}]_{ii}}{q_i}\right)
\end{aligned}$$

Fact 4 (Cross-Partial Derivatives of Posterior Variance)**.**

$$\begin{aligned}
\partial_{ij} f(q_1, \ldots, q_K) &= \frac{-2}{q_i^2 q_j^2} \cdot \tilde{\alpha}'\tilde{V}(\tilde{V} + E)^{-1}\Delta_{ii}(\tilde{V} + E)^{-1}\Delta_{jj}(\tilde{V} + E)^{-1}\tilde{V}\tilde{\alpha} \\
&= \frac{-2\gamma_i\gamma_j}{q_i^2 q_j^2} \cdot [(\tilde{V} + E)^{-1}]_{ij}.
\end{aligned}$$

All of the above facts can be proved by simple linear algebra, so we omit the details.

G.2 Refined Asymptotic Characterization of $n(t)$

We now specialize to $\tilde{\alpha} = \mathbf{1}$ and establish the next lemma, which refines our asymptotic characterization of $n(t)$ in Appendix C.[31] Proposition 1 will immediately follow.

[31] It is easy to see that in the transformed model, λ_i is proportional to $|\tilde{\alpha}_i|$. So $\lambda_i = \frac{1}{K}$ here.

Lemma 8. *For $t \geq 8(R+1)K\sqrt{K}$, it holds that $|n_i(t) - \frac{t}{K}| \leq 4(R+1)\sqrt{K}$.*

Proof. Note from (31) that $(\tilde{V}+E)\gamma = \tilde{V}\mathbf{1}$. So $\tilde{V}(\mathbf{1}-\gamma) = E\gamma = (\frac{\gamma_1}{q_1}, \dots, \frac{\gamma_K}{q_K})'$, and

$$\mathbf{1} - \gamma = (\tilde{V})^{-1} \cdot \left(\frac{\gamma_1}{q_1}, \dots, \frac{\gamma_K}{q_K}\right)'.$$

From the definition of the operator norm, we deduce

$$\sum_{i=1}^{K}(1-\gamma_i)^2 = \|w-\gamma\|^2 \leq R^2 \cdot \left(\sum_{j=1}^{K} \frac{\gamma_j^2}{q_j^2}\right). \tag{32}$$

This holds for any division vector q and the corresponding γ (which is a function of q).

Now suppose without loss of generality that $n_1(t) \geq \frac{t}{K}$. Let $q = (n_1(t)-1, n_2(t), \dots, n_K(t))$ and consider the corresponding γ. Then from t-optimality we have

$$|f(q_1+1, q_{-1}) - f(q)| \geq |f(q_j+1, q_{-j}) - f(q)|, \quad \forall j.$$

Note that the discrete partial derivatives above are related to the usual continuous partials by the following inequalities:[32]

$$\frac{\gamma_j^2}{q_j(q_j+1)} \leq |f(q_j+1, q_{-j}) - f(q)| \leq \frac{\gamma_j^2}{q_j^2}.$$

We therefore deduce

$$\frac{\gamma_1^2}{q_1^2} \geq \frac{\gamma_j^2}{q_j(q_j+1)}, \quad \forall j. \tag{33}$$

Combining (32) and (33) and using $\frac{1}{q_j^2} \leq \frac{2}{q_j(q_j+1)}$, we see that

$$\sum_{i=1}^{K}(1-\gamma_i)^2 \leq 2R^2K \cdot \frac{\gamma_1^2}{q_1^2}. \tag{34}$$

In particular, we know that $\gamma_1 - 1 \leq R\sqrt{2K} \cdot \frac{\gamma_1}{q_1}$. It is easy to see this implies

$$\gamma_1 \leq 1 + \frac{2R\sqrt{K}}{q_1} \leq \sqrt{2} \tag{35}$$

whenever $q_1 = n_1(t) - 1 \geq \frac{t}{K} - 1 \geq (2\sqrt{2}+2)R\sqrt{K}$. Plugging this back into the RHS of (34), we then obtain

$$\gamma_j \geq 1 - \frac{2R\sqrt{K}}{q_1} \geq 2 - \sqrt{2}. \tag{36}$$

[32] The RHS follows from the convexity of f. The LHS can be proved by using Fact 2, Fact 3 and noting that γ_j^2 is an increasing function in q_j, because $\frac{\partial \gamma_j(q)}{\partial q_j} = \frac{\gamma_j}{q_j^2} \cdot [(V+E)^{-1}]_{jj}$ has the same sign of γ_j.

Now use (33), (35) and (36) to deduce that

$$q_j + 1 \geq \frac{\gamma_j}{\gamma_1} \cdot q_1 \geq \frac{1 - \frac{2R\sqrt{K}}{q_1}}{1 + \frac{2R\sqrt{K}}{q_1}} \cdot q_1 \geq \left(1 - \frac{4R\sqrt{K}}{q_1}\right) \cdot q_1 = q_1 - 4R\sqrt{K}.$$

Recall $q_j = n_j(t)$ for $j > 1$ and $q_1 = n_1(t) - 1$. We thus have

$$n_j(t) \geq n_1(t) - 4R\sqrt{K} - 2. \tag{37}$$

Since $n_1(t) \geq \frac{t}{K}$, the above implies $n_j(t) \geq \frac{t}{K} - 4(R+1)\sqrt{K}$ for each signal j. This proves half of the lemma.

For the other half, note that $n_j(t) \leq \frac{t}{K}$ must hold for *some* signal j. Thus (37) yields $n_1(t) \leq \frac{t}{K} + 4(R+1)\sqrt{K}$. This is not just true for signal 1, but in fact for any signal i with $n_i(t) \geq \frac{t}{K}$. So we conclude $n_i(t) \leq \frac{t}{K} + 4(R+1)\sqrt{K}$ for each signal i. The proof of the lemma is complete. □

H Eventual Optimality of the Myopic Rule

Below, write $m(t)$ for the division vector at time t achieved under the (history-independent) myopic rule.[33] We have discussed that when Theorems 1 or 2 apply, the myopic division vector $m(t)$ is t-optimal at every period t. In this appendix, we argue that generically, division vectors $m(t)$ at late periods are t-optimal. This result complements our Theorem 3 (which characterizes the eventual path of the *optimal* strategy), and suggests that a DM who naively follows the myopic rule all the way from the beginning will not do poorly.

To avoid repetition, here we only sketch the core argument. The main new step is to show that the division vectors $m(t)$ under the myopic rule grow to infinity in each coordinate. That is, a myopic DM *would not get stuck* observing a subset of signals forever. Once this is shown, we can repeat the (rest of the) proof of Lemma 4 and deduce that $m_i(t) - \lambda_i \cdot t$ remains bounded. And with these asymptotic characterizations, we can reproduce the proof of Theorem 3 (now for the myopic strategy instead of the optimal strategy) without muh modification.[34]

To see myopic signal choices never get stuck, we establish the following lemma.

[33]That is, $m(t) = (m_1(t), \ldots, m_K(t))$ where $m_i(t)$ is the number of times signal i has been observed under myopic information acquisition prior to and including period t.

[34]These latter steps are actually simpler to carry out for the myopic strategy. This is because in constructing a profitable deviation from the myopic strategy, we only need to achieve lower posterior variance at a single period. The switch deviations we used before are no longer needed.

Lemma 9. *Fix an arbitrary division vector $q \in \mathbb{R}_+^K$ (need not be integers). The partial derivatives of f at q are all zero if and only if $q_1 = \cdots = q_K = \infty$.*

This holds because for normal-linear signals, the posterior variance is *globally convex*. So if each signal has zero marginal value given a division vector q, then q must be a global minimizer of posterior variance, which only occurs when each $q_i = \infty$.

We mention that a similar result (i.e. myopic information acquisition does not get stuck) would not in general be true for other signal structures. The following is a counterexample with normal but non-linear signals.

Example 3. Consider three states $\theta_1, \theta_2, \theta_3$ drawn independently. The DM has access to the following three signals:

$$X_1 = \theta_1 + \text{sign}(\theta_2) + \epsilon_1$$
$$X_2 = \text{sign}(\theta_2\theta_3) + \epsilon_2$$
$$X_3 = \theta_3 + \epsilon_3$$

where $\epsilon_1, \epsilon_2, \epsilon_3$ are Gaussian noise terms. We focus on the prediction problem, in which (at a random time) the DM makes a prediction about θ_1 and receives negative of the squared prediction error.

Note that prior to the first observation of X_2, signal X_3 is completely uninformative about the payoff-relevant state θ_1 (even when combined with previous observations of X_1). Similarly, signal X_2 is individually uninformative about θ_2,[35] and thus about θ_1. These imply that the DM's uncertainty about θ_1 is not reduced upon the first observation of either X_2 or X_3. Hence, the myopic rule in this example is to always observe X_1, contrary to Lemma 9.

Thus, if the DM follows the myopic rule in this example, he will never completely learn the value of θ_1. By contrast, if he is sufficiently patient, then his optimal strategy will observe each signal infinitely often and identify the value of θ_1 in the long run. This distinction suggests that *eventual optimality of the myopic rule* may not hold in general informational environments beyond those considered in this paper. Nonetheless, we conjecture that the *optimal rule eventually proceeds myopically* (that is, Theorem 3 generalizes).

[35] This is because the sign of $\theta_2\theta_3$ does not contain any new information about θ_2 when θ_3 is equally likely to be positive or negative.

References

Aghion, P., P. Bolton, C. Harris, and B. Jullien (1991): "Optimal Learning by Experimentation," *Review of Economic Studies*, 58, 621–654.

Ando, T. (1983): "On the Arithmetic-Geometric-Harmonic-Mean Inequalities for Positive Definite Matrices," *Linear Algebra and its Applications*, 52, 31–37.

Angeletos, G.-M. and A. Pavan (2007): "Efficient Use of Information and the Social Value of Information," *Econometrica*, 75, 1103–1142.

Banks, J. S. and R. K. Sundaram (1992): "A Class of Bandit Problems Yielding Myopically Optimal Strategies," *Journal of Applied Probability*, 29, 625–632.

Bergemann, D., T. Heumann, and S. Morris (2015): "Information and Volatility," *Journal of Economic Theory*.

Bergemann, D. and J. Välimäki (2008): "Bandit Problems," in *The New Palgrave Dictionary of Economics*, ed. by S. N. Durlauf and L. E. Blume, Basingstoke: Palgrave Macmillan.

Berry, D. and B. Fristedt (1988): "Optimality of Myopic Stopping Times for Geometric Discounting," *Journal of Applied Probability*, 25, 437–443.

Bubeck, S., R. Munos, and G. Stoltz (2009): "Pure Exploration in Multi-armed Bandits Problems," in *Algorithmic Learning Theory. ALT 2009. Lecture Notes in Computer Science, vol 5809*, Springer, Berlin, Heidelberg.

Che, Y.-K. and K. Mierendorff (2019): "Optimal Dynamic Allocation of Attention," *American Economic Review*, forthcoming.

Chernoff, H. (1972): *Sequential Analysis and Optimal Design*, Society for Industrial and Applied Mathematics.

Chick, S. E. and P. Frazier (2012): "Sequential Sampling with Economics of Selection Procedures," *Management Science*, 58, 550–569.

Easley, D. and N. M. Kiefer (1988): "Controlling a Stochastic Process with Unknown Parameters," *Econometrica*, 56, 1045–1064.

Fudenberg, D., P. Strack, and T. Strzalecki (2018): "Stochastic Choice and Optimal Sequential Sampling," *American Economic Review*.

Gittins, J. C. (1979): "Bandit processes and dynamic allocation indices," *Journal of the Royal Statistical Society, Series B*, 148–177.

Greenshtein, E. (1996): "Comparison of Sequential Experiments," *The Annals of Statistics*, 24, 436–448.

Hansen, O. H. and E. N. Torgersen (1974): "Comparison of Linear Normal Experiments," *The Annals of Statistics*, 2, 367–373.

Hellwig, C. and L. Veldkamp (2009): "Knowing What Others Know: Coordination Motives in Information Acquisition," *The Review of Economic Studies*, 76, 223–251.

Ke, T. T. and J. M. Villas-Boas (2017): "Optimal Learning Before Choice," Working Paper.

Keener, R. (1985): "Further Contributions to the Two-armed Bandit Problem," *Annals of Statistics*, 13, 418–422.

Lambert, N., M. Ostrovsky, and M. Panov (2018): "Strategic Trading in Informationally Complex Environments," *Econometrica*.

Liang, A. and X. Mu (2018): "Complementary Information and Learning Traps," Working Paper.

Mersereau, A. J., P. Rusmevichientong, and J. N. Tsitsiklis (2009): "A Structured Multiarmed Bandit Problem and the Greedy Policy," *IEEE Transactions On Automatic Control*, 54, 2787–2802.

Meyer, M. and J. Zwiebel (2007): "Learning and Self-Reinforcing Behavior," Working Paper.

Moscarini, G. and L. Smith (2001): "The Optimal Level of Experimentation," *Econometrica*, 69, 1629–1644.

Myatt, D. P. and C. Wallace (2012): "Endogenous Information Acquisition in Coordination Games," *The Review of Economic Studies*, 79, 340–374.

Robbins, H. (1952): "Some aspects of the sequential design of experiments," *Bull. Amer. Math. Soc.*, 58, 527–535.

Rothschild, M. (1974): "A Two-Armed Bandit Theory of Market Pricing," *Journal of Economic Theory*, 9, 185–202.

Russo, D. (2016): "Simple Bayesian Algorithms for Best-Arm Identification," Working Paper.

Sanjurjo, A. (2017): "Search with Multiple Attributes: Theory and Empirics," *Games and Economic Behavior*, 535–562.

Sethi, R. and M. Yildiz (2016): "Communication with Unknown Perspectives," *Econometrica*, 84, 2029–2069.

——— (2017): "Culture and Communication," Working Paper.