



HAL
open science

Payoffs-Beliefs Duality and the Value of Information

Michel de Lara, Olivier Gossner

► **To cite this version:**

Michel de Lara, Olivier Gossner. Payoffs-Beliefs Duality and the Value of Information. 2018. hal-01941006v1

HAL Id: hal-01941006

<https://hal.science/hal-01941006v1>

Preprint submitted on 30 Nov 2018 (v1), last revised 18 Nov 2019 (v3)

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Payoffs-Beliefs Duality and the Value of Information

Michel DE LARA* Olivier GOSSNER†

November 30, 2018

Abstract

In decision problems under incomplete information, payoff vectors (indexed by states of nature) and beliefs are naturally paired by bilinear duality. We exploit this duality to analyze the value of information using convex analysis. We then derive global estimates of the value of information of any information structure from local properties of the value function and of the set of optimal actions taken at the prior belief only, and apply our results to the marginal value of information.

1 Introduction

The value of a piece of information to an economic agent depends on the information at hand, on the agent's prior on the state of nature, and on the decision problem faced. These elements are intrinsically tied, and separating the influence of one of them from that of the others is not straightforward.

The dependency of the value of information on agents' decisions and preferences is certainly a major reason why most information rankings are either uniform among agents or restricted to certain classes of agents. Blackwell (1953)'s comparison of experiments, for instance, is uniform; it states that an information structure is more informative than another if all agents, no matter their available choices and preferences, weakly prefer the former to the latter. Lehmann (1988); Persico (2000); Cabrales, Gossner, and Serrano (2013) are example of papers that build information rankings based on restricted sets of decision problems.

The flip side of this approach is that information rankings are silent as to the dependency of the value of a fixed piece information on the agent's preferences and available choices. They do not tell us what makes information more or less valuable to an arbitrary agent, and neither can they identify the agents who value a given piece of information more than others. If we

*CERMICS, École des Ponts, UPE, Champs-sur-Marne, France. E-mail: michel.delara@enpc.fr

†CREST, CNRS, École Polytechnique, and London School of Economics. Email: olivier.gossner@polytechnique.edu. Olivier Gossner acknowledges support from the French National Research Agency (ANR), "Investissements d'Avenir" (ANR-11-IDEX-0003/LabEx Ecodec/ANR-11-LABX-0047).

want to answer this type of questions, we need to examine carefully how information, priors, decisions and preferences come into play.

The effect of priors and evidence on beliefs is well understood. Given a prior belief, and after receiving some information, an agent forms a posterior belief. Posterior beliefs average out to the prior belief, and information acquisition can usefully be represented by the distribution of these posterior beliefs (see, e.g. Bohnenblust, Shapley, and Sherman, 1949; Aumann and Maschler, 1967; Aumann and Hart, 1986).

In any decision problem, to each decision and state of nature corresponds to a payoff. The decision problem can thus be represented as a set of available vector payoffs, where each payoff is indexed by a state of nature (Blackwell, 1951). Given a posterior belief, the agent makes a decision that maximizes her expected utility so that, to each (posterior) belief of the agent corresponds an expected utility at this belief. The corresponding map from beliefs to expected payoffs is called the value function. The value of a piece of information, defined as the difference in expected utilities from having or not having the information at hand, is thus the difference between the expectation of the value function at the posterior and at the prior, and is non-negative by Jensen's inequality. Thus, the value function fully captures the agent's preferences for information.

In this paper, we make use of *convex analysis* (Rockafellar, 1970) to exploit a bilinear duality structure between payoffs and beliefs, that gives expected payoff (Dentcheva and Ruszczyński, 2013). Primal variables are payoffs vectors, dual variables are beliefs (or, more generally, signed measures) and the value function appears as the (restriction to beliefs) of the support function of the set of available vector payoffs. This provides a correspondance between convex analysis concepts and tools and economic objects. The set of beliefs compatible with an optimal action is related to the *normal cone* of the set of available vector payoff at this optimal action. The *subgradient* of the value function at any belief can be represented as the set of optimal choice of vector payoffs at this belief.

We express the value of information according to the influence it has on decisions. We provide three upper and lower bounds on the value of information.

In the first upper and lower bounds, we characterize information with a positive value. We show that information has a positive value when at least one of the optimal actions at the prior becomes suboptimal for some of the posteriors. We thus define the confidence set at a prior \bar{p} as the set of posterior beliefs for which all optimal actions at \bar{p} remain optimal. We show that information has positive value if and only if posterior beliefs fall outside of the confidence set with positive probability. This result generalises insights from Hirshleifer (1971) and Mirman, Samuelson, and Urbano (1993), who already noticed that information can only be useful insofar as it influences choices. We provide corresponding lower and upper bounds to the value of information.

In the second bounds, we express the fact that the value of information is maximal when it influences actions the most, which happens when information breaks indifferences between several choices. We show that, when this is the case, the value of information can be suitably measured by an expected distance between the prior and the posterior.

Finally, our third bounds apply to cases in which the agent's optimal choice is a smooth

function of her belief around the prior. We show that, in this situation, the value function is also smooth around the prior, and the value of information is essentially a quadratic function of the expected distance between the prior and the posterior.

In a finite decision problem such as shopping behavior (McFadden, 1973) or residential location (McFadden, 1978), at any given prior, the agent either has an optimal action that is locally constant, or is indifferent between several optimal choices. The first and second upper and lower bounds are particularly useful in finite choice problems. The third bounds are most useful in decision problems with a continuum of choices, such as scoring rules (Brier, 1950) or investment decisions (Arrow, 1971). In certain decision problems, the behavior of optimal choice as a function of the belief depends on the range of parameters, and the appropriate bounds apply accordingly.

The paper is organized as follows. Section 2 presents the model and introduces the duality between actions and the value of information. The main results are presented in section 3. Section 4 is devoted to applications to the question of marginal value of information, and Section 5 concludes. The Appendix contains reviews of convex analysis and the proofs.

2 Model and payoffs-beliefs duality

We consider the classical question of an agent who faces a decision problem under imperfect information on a state of nature.

2.1 Information and action

The set of states of nature is a finite set K . We identify the set Σ of signed measures on K with \mathbb{R}^K . The agent holds a prior belief \bar{p} with full support in the set

$$\Delta = \Delta(K) \subset \Sigma \tag{1}$$

of probability distributions over K . We identify Δ with the simplex of \mathbb{R}^K .

A *decision problem* is given by an arbitrary compact convex choice set D and by a continuous payoff function $g: D \times K \rightarrow \mathbb{R}$. The convexity of D is justified by allowing the agent to randomize over decisions. Consistent with the framework of Blackwell (1953), we define the set of *actions* as the compact convex subspace of \mathbb{R}^K given by:

$$A = \{(g(d, k))_{k \in K}, d \in D\} . \tag{2}$$

The scalar product between a vector $v \in \mathbb{R}^K$ and a signed measure $s \in \mathbb{R}^K$ is

$$\langle s, v \rangle = \sum_{k \in K} s_k v_k . \tag{3}$$

This scalar product induces a duality between payoffs/actions and beliefs. Such a duality is at the core of a series of works in non-expected utility theory, such as Gilboa and Schmeidler

(1989); Maccheroni, Marinacci, and Rustichini (2006); Cerreia-Vioglio, Maccheroni, Marinacci, and Montrucchio (2011a,b).

Under belief $p \in \Delta$, the decision maker chooses a decision $d \in D$ that maximizes $\sum_k p_k g(d, k)$, or, equivalently, an action $a \in A$ that maximizes $\langle p, a \rangle$, and the corresponding *expected payoff* is $\max_{a \in A} \langle p, a \rangle \in \mathbb{R}$. We define the *value function* $v_A : \Delta \rightarrow \mathbb{R}$ by:

$$v_A(p) = \max_{a \in A} \langle p, a \rangle, \quad \forall p \in \Delta. \quad (4)$$

The value function $v_A : \Delta \rightarrow \mathbb{R}$ is convex — as the supremum of the family of linear maps $\langle \cdot, a \rangle$ for $a \in A$ — and continuous — as its domain is the whole convex set Δ (Hiriart-Urruty and Lemaréchal, 1993, p. 175).

We follow Bohnenblust, Shapley, and Sherman (1949); Blackwell (1953), and we describe information through a distribution of posterior beliefs that average to the prior belief. Hence, given the prior belief \bar{p} , we define an *information structure* as a random variable \mathbf{q} with values in Δ describing the agent's posterior beliefs, and such that

$$\mathbb{E} \mathbf{q} = \bar{p}, \quad (5)$$

where \mathbb{E} denotes the expectation operator.

Given the action set A and the information structure \mathbf{q} , the *value of information* $\mathbf{VoI}_A(\mathbf{q})$ is the difference between the expected payoff for an agent who receives information according to \mathbf{q} and one whose prior is \bar{p} . It is given by:

$$\mathbf{VoI}_A(\mathbf{q}) = \mathbb{E} v_A(\mathbf{q}) - v_A(\bar{p}). \quad (6)$$

2.2 Duality between payoffs and beliefs

Given a belief $p \in \Delta$, we let $A^*(p) \subset A$ be the *set of optimal actions at belief p* , given by

$$A^*(p) = \arg \max_{a' \in A} \langle p, a' \rangle = \{a \in A \mid \forall a' \in A, \langle p, a' \rangle \leq \langle p, a \rangle\}. \quad (7)$$

The set $A^*(p)$ is nonempty, closed, convex (as A is convex and compact), and is a subset of the *boundary* ∂A of A .

Conversely, an outside observer can make inferences on the agent's beliefs from observed actions. For an action $a \in A$, the set $\Delta_A^*(a)$ of *beliefs revealed by action a* is the set of all beliefs for which a is an optimal action, given by:

$$\Delta_A^*(a) = \{p \in \Delta \mid \forall a' \in A, \langle p, a' \rangle \leq \langle p, a \rangle\}. \quad (8)$$

Geometrically, the set $\Delta_A^*(a)$ is the intersection with Δ of the *normal cone* $N_A(a)$ (see (50) for a proper definition).

Obviously, given $a \in A$ and $p \in \Delta$, $a \in A^*(p)$ iff $p \in \Delta_A^*(a)$, as both express that action a is optimal under belief p .

In this paper, we exploit the relationship between differentiability properties of the value function v_A at \bar{p} with the value of information \mathbf{VoI}_A for an agent with prior \bar{p} .

3 On the value of information

In this section, we relate the geometry of the set A of actions both with the behavior of the agent around the prior \bar{p} , with differentiability properties of the value function v_A at the prior \bar{p} , and with the value of information \mathbf{VoI}_A . This approach allows us to derive bounds on the value of information that depend on how information influences actions.

First, in Subsection 3.1, we consider information that does not allow us to eliminate optimal actions. We introduce the *confidence set* as the set of posterior beliefs at which all optimal actions at the prior remain optimal. We show that information is valuable if and only if, with positive probability, it can lead to a posterior outside this set.

Second, in Subsection 3.2, we consider the somewhat opposite case of tie-breaking information. This corresponds to situations in which the agent is indifferent between several actions, and the information allows her to select among them. We show that the value of information can be related to an expected distance between the prior and the posterior, provided that posterior beliefs move in these tie-breaking directions.

These two first approaches are suitable in finite decision problems where the value function is piecewise linear. In the third approach, in Subsection 3.3, we look at situations in which the optimal action is locally unique around the prior and depends on information in a continuous and differentiable way. There, we show that the value of information can essentially be measured as an expected square distance from the prior to the posterior. This approach is particularly adapted to cases in which the space of actions is sufficiently rich, and where small changes of beliefs lead to corresponding changes of actions.

3.1 Valuable information

Our first task is to formalize the idea that useful information is information that affects optimal choices. Since there are potentially several optimal actions at a prior \bar{p} and at a posterior p , there are in principle many ways to formalize this idea.

We say that a belief p is in the *confidence set* $\Delta_A^c(\bar{p})$ of \bar{p} when all optimal actions at \bar{p} are also optimal at p :

$$\Delta_A^c(\bar{p}) = \bigcap_{a \in A^*(\bar{p})} \Delta_A^*(a). \quad (9)$$

Another way to look at this notion is to consider an observer who sees choices by the decision maker: $p \in \Delta_A^c(\bar{p})$ when none of the actions chosen by the agent at \bar{p} would lead the observer to refute the possibility that the agent has belief p .

The notion of a confidence set allows for the characterization of valuable information as follows.

Proposition 1 (Valuable information)

$$\mathbf{VoI}_A(\mathbf{q}) = 0 \text{ iff } \exists a^* \in A^*(\bar{p}), a^* \in A^*(\mathbf{q}) \text{ a.s.} \quad (10a)$$

$$\text{iff } \mathbf{q} \in \Delta_A^c(\bar{p}) \text{ a.s.} \quad (10b)$$

It is relatively straightforward to see that if all posteriors remain in the confidence set, information is valueless. In fact, when this is the case, the same action is optimal for all of the posteriors, which means that the agent can play this action, while ignoring the new information, and obtain the same value. The proposition shows that the converse result also holds: the value of information is positive whenever posteriors fall outside of the confidence set with some positive probability.

More can be said about estimates on the value of information. To do so, we introduce an ε -neighborhood of the confidence set $\Delta_A^c(\bar{p})$. For $\varepsilon > 0$, let

$$\Delta_{A,\varepsilon}^c(\bar{p}) = \{q \in \Delta \mid d(q, \Delta_A^c(\bar{p})) < \varepsilon\}, \quad (11)$$

where, by definition,

$$d(q, \Delta_A^c(\bar{p})) = \inf_{p \in \Delta_A^c(\bar{p})} \|p - q\|. \quad (12)$$

This leads us to a first estimate of the value of information.

Theorem 2 (Bound on the value of information based on confidence sets) *For every $\varepsilon > 0$, there exist positive constants C_A and $c_{\bar{p},A,\varepsilon}$ such that, for every information structure \mathbf{q} ,*

$$C_A \mathbb{E} d(\mathbf{q}, \Delta_A^c(\bar{p})) \geq \mathbf{VoI}_A(\mathbf{q}) \geq c_{\bar{p},A,\varepsilon} \mathbb{P}\{\mathbf{q} \notin \Delta_{A,\varepsilon}^c(\bar{p})\}. \quad (13)$$

The upper bound tells us that the value of information is bounded by (a constant times) the expected distance from the posterior to the confidence set at the prior. In particular, it is bounded by the expected distance from the posterior to the prior itself. The lower bound is a converse result, but in which we need to replace the confidence set by some ε -neighborhood. It shows us that the value of information is bounded below by (a constant times) the probability that the posterior exits the confidence set by more than ε , and, therefore, it is also larger than the expected distance from the posterior to this ε neighborhood of the confidence set. Both the lower and upper bounds depend on the confidence set $\Delta_A^c(\bar{p})$, which can be computed locally at \bar{p} . On the other hand, they apply to all information structures. The caveat is that the multiplicative constants C_A and $c_{\bar{p},A,\varepsilon}$ depend on global, and not just local, properties of the set A .

3.2 Undecided

We now consider situations in which information influences actions the most. Those are situations of indifference in which, at the prior belief \bar{p} , the agent is *undecided* between several optimal actions. A small piece of information can then be enough to break this indifference. As shown by the following proposition, the value function then exhibits a *kink* at \bar{p} .

Proposition 3 *The two conditions are equivalent:*

- the set $A^*(\bar{p})$ of optimal actions at the prior belief \bar{p} contains more than one element;
- the value function v_A is non-differentiable (in the standard sense) at the prior belief \bar{p} .

At such beliefs \bar{p} , the convexity gap of the value function v_A is maximal in the directions in which it is non-differentiable. This allows us to derive a second bound on the value of information. Cases of indifference are typical of situations with a finite number of action choices.

We call *indifference kernel* $\Sigma_A^i(\bar{p})$ at \bar{p} the vector space of signed measures

$$\Sigma_A^i(\bar{p}) = [A^*(\bar{p}) - A^*(\bar{p})]^\perp . \quad (14)$$

Beliefs in the indifference kernel $\Sigma_A^i(\bar{p})$ do not break any of the ties in $A^*(\bar{p})$, since:

$$p \in \Sigma_A^i(\bar{p}) \iff \langle p, a \rangle = \langle p, a' \rangle , \quad \forall (a, a') \in A^*(\bar{p})^2 . \quad (15)$$

We note the inclusion

$$\Delta_A^c(\bar{p}) \subset \Sigma_A^i(\bar{p}) \cap \Delta , \quad (16)$$

as every element in the confidence set is necessarily in the indifference kernel and in the simplex of probability measures.

Recall that a semi-norm on the signed measures Σ on K , identified with \mathbb{R}^K , is a mapping $\|\cdot\| : \mathbb{R}^K \rightarrow \mathbb{R}_+$ which satisfies the requirements of a norm, except that the vector subspace $\{s \in \mathbb{R}^K \mid \|s\| = 0\}$ — called the *kernel* of the semi-norm $\|\cdot\|$ — is not necessarily reduced to the null vector.

Theorem 4 (Bounds on the value of information for the undecided agent) *There exists a positive constant C_A and a semi-norm $\|\cdot\|_{\Sigma_A^i(\bar{p})}$ with kernel $\Sigma_A^i(\bar{p})$, the indifference kernel in (14), such that, for every information structure \mathbf{q} ,*

$$C_A \mathbb{E} \|\mathbf{q} - \bar{p}\| \geq \mathbf{VoI}_A(\mathbf{q}) \geq \mathbf{VoI}_{A^*(\bar{p})}(\mathbf{q}) \geq \mathbb{E} \|\mathbf{q} - \bar{p}\|_{\Sigma_A^i(\bar{p})} . \quad (17)$$

The lower bound in Theorem 4 shows that a lower bound of the value of information is the expectation of a semi-norm of the distance between the prior belief and the posterior belief. To understand the role of the kernel $\Sigma_A^i(\bar{p})$ of this semi-norm, let us first consider the set of beliefs in this set. A posterior q is in $\Sigma_A^i(\bar{p}) = [A^*(\bar{p}) - A^*(\bar{p})]^\perp$ if and only if, for any two optimal actions $a, a' \in A^*(\bar{p})$, $\langle q, a \rangle = \langle q, a' \rangle$. In words, posteriors that do not break any of the ties in $A^*(\bar{p})$ might not be valuable to the agent. On the other hand, Theorem 4 tells us that all other directions — i.e., those that allow at least one of the ties in $A^*(\bar{p})$ to be broken — are valuable to the agent, and furthermore, in these directions, the value of information behaves like an expected distance from the prior to the posterior.

The upper bound says that the value of information is bounded by an expected distance from the prior to the posterior, and the second inequality states that the value of information with decision set A is at least as large as with decision set $A^*(\bar{p})$.

Note that the bounds on Theorem 4 rely on the indifference kernel $\Sigma_A^i(\bar{p})$, which can be computed directly from the set $A^*(\bar{p})$. The multiplicative constant C_A , however, depends on more global properties of the set A .

3.3 Flexible

Finally, we consider the case in which there is a unique optimal action for each belief in the range considered, and this action depends in a certain smooth way on the belief. More precisely, we assume that around the prior, optimal actions depend on a 1-1 way on the belief in a certain differentiable way. This assumption is met when, for instance, the decision problem faced by the agent is a scoring rule (Brier, 1950), or an investment problem (Arrow, 1971; Cabrales, Gossner, and Serrano, 2013).

Our first step is to characterize a class of situations of interest, in which the agent's optimal action depends smoothly on her belief. The proposition offers three alternative characterizations of these situations, based 1) on the local behavior of the agent's optimal choices, 2) on local properties of the geometry of the boundary of the set of actions, and 3) on local second differentiability properties of the value function.

Proposition 5 *Suppose that the action set A has boundary ∂A which is a C^2 submanifold of \mathbb{R}^K . The three following conditions are equivalent:*

1. *The set-valued mapping*

$$A^*: \Delta \rightrightarrows A, \quad p \mapsto A^*(p) \tag{18}$$

is a local diffeomorphism¹ at the prior belief \bar{p} ;

2. *The set $A^*(\bar{p})$ of optimal actions at the prior belief \bar{p} is reduced to a singleton at which the curvature of the action set A is positive;*
3. *The value function v_A is twice differentiable at the prior belief \bar{p} , and the Hessian is definite positive.*

In this case, we say that the agent is flexible at \bar{p} .

Theorem 6 (Bounds on the VoI for the flexible agent) *If the agent is flexible at \bar{p} , then there exist positive constants $C_{\bar{p},A}$ and $c_{\bar{p},A}$ such that, for every information structure \mathbf{q} ,*

$$C_{\bar{p},A} \mathbb{E} \|\mathbf{q} - \bar{p}\|^2 \geq \mathbf{VoI}_A(\mathbf{q}) \geq c_{\bar{p},A} \mathbb{E} \|\mathbf{q} - \bar{p}\|^2. \tag{19}$$

Theorem 6 shows that, in the case of a flexible agent, the value of information is essentially given by the expected square distance between the prior and the posterior, up to some multiplicative constant. One of the strengths of the theorem is that its assumption that the agent is flexible is a local one, whereas its conclusion is global, as it applies to all information structures. On the other hand, the multiplicative constants themselves depend on the global behavior of the value function, and hence cannot be inferred from local properties only.

¹Meaning that the set $A^*(p)$ is a singleton for all $p \in \Delta$, in which case we identify a singleton set with its single element.

4 The marginal value of information

Radner and Stiglitz (1984) study the question of the marginal value of information. They provide joint conditions on a parametrized family of information structures together with a decision problem such that, when the agent is close to receiving no information at all, the marginal value of information is null. Their result was subsequently generalized by Chade and Shlee (2002) and De Lara and Gilotte (2007), who also characterize joint conditions on parametrized information and a decision problem leading to zero marginal value of information.

In this section, we show how our bounds on the value of information apply to the marginal value of information. In particular, we provide separate conditions on the decision problem and on the family of parametrized information structures that result in a null value of information. We then examine several parametrized families of information structures and rely on our main results to study how the marginal value of information varies depending on the decision problem faced.

Let $(\mathbf{q}^\theta)_{\theta>0}$ be a family of information structures. As in Radner and Stiglitz (1984), we are interested in the marginal value of information:

$$V^+ = \limsup_{\theta \rightarrow 0} \frac{1}{\theta} \text{VoI}_A(\mathbf{q}^\theta). \quad (20)$$

The following proposition is a straightforward consequence of Theorems 2 and 6.

Proposition 7 *Assume either that*

1. $\mathbb{E} d(\mathbf{q}^\theta, \Delta_A^c(\bar{p})) = o(\theta)$,
2. *the decision maker is flexible at \bar{p} and $\mathbb{E} \|\mathbf{q}^\theta - \bar{p}\|^2 = o(\theta)$.*

Then $V^+ = 0$.

The first condition is met automatically if $\mathbb{E} \|\mathbf{q}^\theta - \bar{p}\| = o(\theta)$. It is also met if, for instance, $\Delta_A^c(\bar{p})$ has a non-empty interior, and posteriors converge to the prior almost-surely.

We now discuss how our approach in Proposition 7 compares with the literature. In Radner and Stiglitz (1984), one finds joint conditions on the parameterized information structure $(\mathbf{q}^\theta)_{\theta>0}$ and the decision problem at hand A , leading to $V^+ = 0$. The second case in Proposition 7, when the decision maker is flexible, compares with the original Radner-Stiglitz assumptions for the smoothness part, but not for the uniqueness of optimal actions. Indeed, Assumption (A0) in Radner and Stiglitz (1984) does not require that $A^*(\mathbf{q}^\theta)$ be a singleton, for all θ .

Chade and Shlee (2002) make a step towards disentangling conditions on the parameterized information structure $(\mathbf{q}^\theta)_{\theta>0}$ from conditions on the decision problem A that lead to a null marginal value of information. However, like Radner and Stiglitz (1984), they make an assumption on how the optimal action varies with information, which makes the comparison

with Proposition 7 delicate. In addition, Chade and Shlee (2002) provide sufficient conditions for $V^+ = 0$ that bear on the conditional distribution of the signal knowing the state of nature. Our approach focuses on the posterior conditional distribution of the state of nature knowing the signal.

De Lara and Gilotte (2007) provide separate conditions on the parameterized information structure $(\mathbf{q}^\theta)_{\theta>0}$ and the decision problem A that lead to $V^+ = 0$. Their condition “IIDV=0” is that $\limsup_{\theta \rightarrow 0} \frac{1}{\theta} \mathbb{E} \|\mathbf{q}^\theta - \bar{p}\| = 0$, or, equivalently, $\mathbb{E} \|\mathbf{q}^\theta - \bar{p}\| = o(\theta)$, which implies the first item of Proposition 7. Thus, this proposition implies the main result of De Lara and Gilotte (2007).

4.1 Examples

Here, we study the marginal value of information for several typical parametrized information structures. In the first, information consists on the observation of a Brownian motion with known variance and a drift that depends on the state of nature. In the second, information consists of the observation of a Poisson process whose probability of success depends on the state of nature. In these two well studied families in the learning literature, the natural parametrization of information is the length of the interval of time during which observation takes place. In our third example, the agent observes a binary signal and the marginal value of information depends on the asymptotic informativeness of these signals close to the situation without information.

In all three following examples we assume binary states of nature: $K = \{0, 1\}$. The prior belief on the state being 1 is denoted \bar{p} . We follow the conditions under which we established bounds on the value of information, and label as “undecided” the case in which the decision problem faced by the decision maker is such that there is indifference between two actions at \bar{p} , “flexible” the case in which the optimal action is a smooth function of the belief in a neighborhood of \bar{p} , and “confident” the case in which there is a unique optimal action in an open interval of beliefs containing \bar{p} , and in this case we let (p_l, p_h) be the set of beliefs for which this action is the unique optimal one.

Our aim is to develop estimates of the marginal value of information. There are three possibilities: it can be infinite, null, or positive and finite. We denote these three cases by $V^+ = \infty$, $V^+ = 0$ and $V^+ \simeq 1$ respectively.

Example 1 (Brownian motion) *Frameworks in which agents observe a Brownian motion with known volatility and unknown drift include Bergemann and Välimäki (1997), Keller and Rady (1999), Bolton and Harris (1999), as well as reputation models like Faingold and Sannikov (2011).*

Assume the agent observes the realization of a Brownian motion with variance 1 and drift $k \in \{0, 1\}$:

$$d\mathbf{Z}_t = kdt + d\mathbf{B}_t \tag{21}$$

for a small interval of time $\theta > 0$.

If we let \mathbf{q}^t be the posterior belief at time t , it is well known² that \mathbf{q}^t follows a diffusion

²See for instance Lemma 1 in Bolton and Harris (1999) or Lemma 2 in Faingold and Sannikov (2011).

process of the form

$$d\mathbf{q}^t = \mathbf{q}^t(1 - \mathbf{q}^t)d\mathbf{w}_t,$$

where \mathbf{w} is a standard Brownian process.

Thus, for small values of θ ,

$$\mathbb{E} \|\mathbf{q}^\theta - p\| \sim \sqrt{\theta}, \quad (22a)$$

$$\mathbb{E} \|\mathbf{q}^\theta - p\|^2 \sim \theta. \quad (22b)$$

It follows from Theorems 2-6 that the marginal value of information is characterized, depending on the decision problem, as:

1. In the confident case, $V^+ = 0$,
2. In the flexible case, $V^+ \simeq 1$,
3. In the undecided case, $V^+ = \infty$.

Example 2 (Exponential learning) Exponential learning plays a central role in models of strategic experimentation such as Keller, Rady, and Cripps (2005). Assume the agent observes a Poisson process with intensity ρ_k during a small interval of time $\theta > 0$, where $\rho_1 > \rho_0$. The probability of two successes is negligible compared to the probability of one success (of order θ^2 compared to θ). A success leads to a posterior that converges, as $\theta \rightarrow 0$, from below to

$$q^+ = \frac{\bar{p}\rho_1}{\bar{p}\rho_1 + (1 - \bar{p})\rho_0} > \bar{p}, \quad (23)$$

and happens with probability of order $\sim \theta$. In the absence of success, the posterior belief converges to \bar{p} as $\theta \rightarrow 0$.

In particular:

$$\mathbb{E} \|\mathbf{q}^\theta - p\| \sim \theta, \quad (24a)$$

$$\mathbb{E} \|\mathbf{q}^\theta - p\|^2 \sim \theta. \quad (24b)$$

Hence, we obtain the following estimates on the marginal value of information:

1. In the confident case,
 - (a) $V^+ \simeq 1$ if $q^+ > p_h$,
 - (b) $V^+ \simeq 0$ if $q^+ \leq p_h$,
2. In the flexible case, $V^+ \simeq 1$,
3. In the undecided case, $V^+ \simeq 1$.

Example 3 (Equally likely signals) *In our third example, we consider binary and equally likely signals, which lead to a “split” of beliefs around the prior \bar{p} . Depending on the precision of these signals as a function of θ , the posterior beliefs are $p \pm \theta^\alpha$ for a certain parameter $\alpha > 0$. Lower values of α correspond to more spread out beliefs around the prior, hence to more accurate information.*

In this case we easily compute:

$$\mathbb{E} \|\mathbf{q}^\theta - p\| = \theta^\alpha, \quad (25a)$$

$$\mathbb{E} \|\mathbf{q}^\theta - p\|^2 = \theta^{2\alpha}. \quad (25b)$$

Here again, the marginal value of information is deduced from Theorems 2–6:

1. *In the confident case, $V^+ = 0$,*

2. *In the flexible case,*

(a) $V^+ = \infty$ if $\alpha < \frac{1}{2}$,

(b) $V^+ \simeq 1$ if $\alpha = \frac{1}{2}$,

(c) $V^+ = 0$ if $\alpha > \frac{1}{2}$,

3. *In the undecided case,*

(a) $V^+ = \infty$ if $\alpha < 1$,

(b) $V^+ \simeq 1$ if $\alpha = 1$,

(c) $V^+ = 0$ if $\alpha > 1$.

Table 1 summarizes the marginal value of information in all of our examples.

V^+	confident	flexible	undecided
Poisson	0 or 1	1	1
Brownian	0	1	∞
EL, $\alpha < \frac{1}{2}$	0	∞	∞
EL, $\alpha = \frac{1}{2}$	0	1	∞
EL, $\frac{1}{2} < \alpha < 1$	0	0	∞
EL, $\alpha = 1$	0	0	1
EL, $\alpha > 1$	0	0	0

Table 1: Marginal value of information in the different examples. EL stands for the equally likely signals case, 1 represents a positive and finite marginal value of information.

In all cases except one, the marginal value of information is completely determined by the local behavior of the value function around the prior. For the Poisson case, the marginal

value of information is 0 or positive, depending on whether the observation of a success is sufficient to lead to a decision reversal.

The marginal value of information is always weakly higher in the flexible case than in the undecided case, and weakly higher in the undecided case than in other cases. In the confident case, the marginal value of information is null, except in the Poisson case with $q^+ > p_h$. This is driven by the fact that, in all other cases, posteriors are, with high probability, too close to the prior to lead to a decision reversal. In the undecided situation, the marginal value of information is always positive or infinite, except for sufficiently uninformative binary signals ($\alpha > 1$). Finally, in the flexible case, the most representative of decision problems with a continuum of actions, the value of information is positive or infinite, except with quite uninformative binary signals ($\alpha > 1/2$).

5 Conclusion

In decision problems under incomplete information, we have formalized the natural correspondence, given by duality, between the set of available choices to a decision maker and the value function expressed as a function of her belief. This, in turn, has allowed us to derive bounds on the value of any piece of information, bounds that are based solely on local properties of the agent's behavior around her prior. Finally, we have provided applications to the question of the marginal value of information.

References

- ARROW, K. J. (1971): "The value of and demand for information," in *Decision and Organization*, ed. by C. McGuire, and R. Radner, pp. 131–139, Amsterdam. North-Holland.
- AUBIN, J.-P. (1982): *Mathematical Methods of Game and Economic Theory*. North-Holland, Amsterdam, second edn.
- AUMANN, R., AND M. MASCHLER (1967): "Repeated games with incomplete information: A survey of recent results," in *Reports to the U.S. Arms Control and Disarmament Agency, ST-116*, pp. 287–403.
- AUMANN, R. J., AND S. HART (1986): "Bi-convexity and bi-martingales," *Israel Journal of Mathematics*, 54, 159–180.
- BERGEMANN, D., AND J. VÄLIMÄKI (1997): "Market diffusion with two-sided learning," *The RAND Journal of Economics*, pp. 773–795.
- BLACKWELL, D. (1951): "Comparison of experiments," in *Proceedings of the Second Berkeley Symposium on Mathematical Statistics and Probability*, pp. 93–102. University of California Press.

- (1953): “Equivalent comparison of experiments,” *Annals of Mathematical Statistics*, 24, 265–272.
- BOHNENBLUST, H. F., L. S. SHAPLEY, AND S. SHERMAN (1949): “Reconnaissance in Game Theory,” Discussion Paper RM-208, The RAND corporation.
- BOLTON, P., AND C. HARRIS (1999): “Strategic experimentation,” *Econometrica*, 67(2), 349–374.
- BRIER, G. W. (1950): “Verification of forecasts expressed in terms of probability,” *Monthly Weather Review*, 78(1-3).
- CABRALES, A., O. GOSSNER, AND R. SERRANO (2013): “Entropy and the value of information for investors,” *American Economic Review*, 103, 360–377.
- CERREIA-VIOGLIO, S., F. MACCHERONI, M. MARINACCI, AND L. MONTRUCCHIO (2011a): “Complete monotone quasiconcave duality,” *Mathematics of Operations Research*, 36(2), 321–339.
- (2011b): “Uncertainty averse preferences,” *Journal of Economic Theory*, 146(4), 1275–1330.
- CHADE, H., AND E. SHLEE (2002): “Another Look at the Radner-Stiglitz Nonconcavity in the Value of Information,” *Journal of Economic Theory*, 107, 421–452.
- DE LARA, M., AND L. GILOTTE (2007): “A tight sufficient condition for Radner–Stiglitz nonconcavity in the value of information,” *Journal of Economic Theory*, 137(1), 696–708.
- DENTCHEVA, D., AND A. RUSZCZYŃSKI (2013): “Common Mathematical Foundations of Expected Utility and Dual Utility Theories,” *SIAM Journal on Optimization*, 23(1), 381–405.
- FAINGOLD, E., AND Y. SANNIKOV (2011): “Reputation in Continuous-Time Games,” *Econometrica*, 79, 773–876.
- GILBOA, I., AND D. SCHMEIDLER (1989): “Maxmin expected utility with non-unique prior,” *Journal of mathematical economics*, 18(2), 141–153.
- HIRIART-URURTY, J. B., AND C. LEMARÉCHAL (1993): *Convex Analysis and Minimization Algorithms I*. Springer-Verlag, Berlin.
- HIRSHLEIFER, J. (1971): “The private and social value of information and the reward to inventive activity,” *American Economic Review*, 61, 561–574.
- KELLER, G., AND S. RADY (1999): “Optimal Experimentation in a Changing Environment,” *Review of Economic Studies*, 66(3), 475–507.

- KELLER, G., S. RADY, AND M. CRIPPS (2005): “Strategic experimentation with exponential bandits,” *Econometrica*, 73(1), 39–68.
- LEHMANN, E. L. (1988): “Comparing location experiments,” *The Annals of Statistics*, 16(2), 521–533.
- MACCHERONI, F., M. MARINACCI, AND A. RUSTICHINI (2006): “Ambiguity aversion, robustness, and the variational representation of preferences,” *Econometrica*, 74(6), 1447–1498.
- McFADDEN, D. (1973): “Conditional logit analysis of qualitative choice behavior,” in *Frontiers in Econometrics*, ed. by P. Zarembka. Academic Press.
- (1978): “Modelling the choice of residential location,” in *Spatial Interaction Theory and Planning Models*, ed. by A. Karlqvist, L. Lundqvist, F. Snickars, and J. Weibull. North-Holland.
- MIRMAN, L. J., L. SAMUELSON, AND A. URBANO (1993): “Monopoly experimentation,” *International Economic Review*, pp. 549–563.
- PERSICO, N. (2000): “Information acquisition in auctions,” *Econometrica*, 68(1), 135–148.
- RADNER, R., AND J. STIGLITZ (1984): “A nonconcavity in the value of information,” in *Bayesian Models of Economic Theory*, ed. by M. Boyer, and R. Kihlstrom, pp. 33–52, Amsterdam. Elsevier.
- ROCKAFELLAR, T. R. (1970): *Convex Analysis*. Princeton University Press, Princeton, N.J.
- SCHNEIDER, R. (2014): *Convex bodies: the Brunn-Minkowski theory*. Cambridge University Press, second edn.

A Convex analysis and geometry background

The set of states of nature is a finite set K . We denote by Σ the set of *signed measures on K* , identified with \mathbb{R}^K . The set Δ of *probability distribution on K* is a convex subset of the set Σ , identified with the simplex of \mathbb{R}^K .

A.1 Some reviews of convex analysis

We rely mostly on the reference book Hiriart-Uruty and Lemaréchal (1993) for reminders on convex analysis.

Let $C \subset \mathbb{R}^K$ be a nonempty convex set. The *support function* σ_C of the set C is the convex function defined by

$$\sigma_C(s) = \sup_{x \in C} \langle s, x \rangle, \quad \forall s \in \Sigma. \quad (26)$$

For any signed measure $s \in \Sigma$, the (*exposed*) *face of C in the direction $s \in \Sigma$* is

$$F_C(s) = \arg \max_{x' \in C} \langle s, x' \rangle = \{x \in C \mid \forall x' \in C, \langle s, x' \rangle \leq \langle s, x \rangle\} \subset C. \quad (27)$$

For any x in C , the *normal cone* to the closed convex set C at $x \in C$ is

$$N_C(x) = \{s \in \Sigma \mid \forall x' \in C, \langle s, x' \rangle \leq \langle s, x \rangle\} \subset \Sigma. \quad (28)$$

Proposition 8 *Let $C \subset \mathbb{R}^K$ be a nonempty convex set.*

1. *Exposed face and normal cone are conjugate as follows:*

$$x \in F_C(y) \iff x \in C \text{ and } y \in N_C(x). \quad (29)$$

2. *Let $X \subset C$ be nonempty. Let $Y \subset \mathbb{R}^K$ be nonempty. Then*

$$X \subset \bigcap_{y \in Y} F_C(y) \iff Y \subset \bigcap_{x \in X} N_C(x) \iff \sigma_C(y) = \langle y, x \rangle, \quad \forall x \in X, \forall y \in Y. \quad (30)$$

3. *Let $y \in \mathbb{R}^K$ be such that $F_C(y) \neq \emptyset$. Then, we have*

$$\sigma_C(y') - \sigma_C(y) \geq \sigma_{F_C(y)}(y' - y) \geq \langle y' - y, x' \rangle, \quad \forall y' \in \mathbb{R}^K, \forall x' \in C. \quad (31)$$

4. *The function σ_{C-C} is a semi-norm with kernel $[C - C]^\perp$.*

Proof.

1. Exposed face and normal cone are conjugate as follows (Hiriart-Uruty and Lemaréchal, 1993, p. 220):

$$x \in C \text{ and } \langle y, x \rangle = \sigma_C(y) \iff x \in \operatorname{argmax}_{x' \in C} \langle y, x' \rangle \quad (32a)$$

$$\iff x \in F_C(y) \quad (32b)$$

$$\iff x \in C \text{ and } y \in N_C(x) . \quad (32c)$$

2. We have

$$\begin{aligned} X \subset \bigcap_{y \in Y} F_C(y) &\iff x \in F_C(y) , \forall x \in X , \forall y \in Y \\ &\iff y \in N_C(x) , \forall x \in X , \forall y \in Y \quad \text{by (29) as } X \subset C \\ &\iff Y \subset \bigcap_{x \in X} N_C(x) . \end{aligned}$$

3. The subdifferential of the support function σ_C of the (nonempty) closed convex set $C \subset \mathbb{R}^K$ at $y \in \mathbb{R}^K$ is (Aubin, 1982, p. 107), (Hiriart-Uruty and Lemaréchal, 1993, p. 258)

$$\partial\sigma_C(y) = F_C(y) = \operatorname{argmax}_{x \in C} \langle y, x \rangle , \forall y \in \mathbb{R}^K . \quad (33)$$

If $\partial\sigma_C(y) \neq \emptyset$, then (31) is a consequence of the definition of the subdifferential.

4. The support function σ_{C-C} is homogeneous and sublinear, and it is nonnegative since $0 \in C - C$. As a consequence, σ_{C-C} is a semi-norm. The kernel is easily calculated.

This ends the proof. ■

A.2 Geometric convex analysis

A nonempty, convex and compact set $A \subset \mathbb{R}^K$ is called a *convex body* of \mathbb{R}^K (Schneider, 2014, p. 8).

Regular points and smooth bodies. We say that a point $a \in A$ is *smooth* or *regular* (Schneider, 2014, p. 83) if the normal cone $N_A(a)$ is reduced to a half-line. The *set of regular points* is denoted by $\operatorname{reg}(A)$:

$$a \in \operatorname{reg}(A) \iff \exists s \in \Sigma , s \neq 0 , N_A(a) = \mathbb{R}_+ s . \quad (34)$$

Notice that a regular point a necessarily belongs to the boundary ∂A of A : $\operatorname{reg}(A) \subset \partial A$. The body A is said to be *smooth* if all boundary points of A are regular ($\operatorname{reg}(A) = \partial A$); in that case, it can be shown that its boundary ∂A is a C^1 submanifold of \mathbb{R}^K (Schneider, 2014, Theorem 2.2.4, p. 83).

Spherical image map of A . We denote by $S^{|K|-1}$ the unit sphere of the signed measures Σ on K (identified with \mathbb{R}^K with its canonical scalar product):

$$S^{|K|-1} = \{s \in \Sigma, \|s\| = 1\}. \quad (35)$$

By (34), we have that

$$a \in \text{reg}(A) \iff \exists! s \in S^{|K|-1}, N_A(a) = \mathbb{R}_+ s. \quad (36)$$

If a point $a \in A$ is regular, the unique outer normal unitary vector to A at a is denoted by $n_A(a)$, so that $N_A(a) = \mathbb{R}_+ n_A(a)$. The mapping

$$n_A : \text{reg}(A) \rightarrow S^{|K|-1}, \quad \text{where } \text{reg}(A) \subset \partial A \quad (37)$$

is called the *spherical image map of A* , or the *Gauss map*, and is continuous (Schneider, 2014, p. 88). We have

$$a \in \text{reg}(A) \Rightarrow N_A(a) = \mathbb{R}_+ n_A(a) \quad \text{where } n_A(a) \in S^{|K|-1}. \quad (38)$$

Reverse spherical image map of A . We say that a unit signed measure $s \in S^{|K|-1}$ is *regular* (Schneider, 2014, p. 87) if the (exposed) face $F_A(s)$ of A in the direction s , as defined in (47), is reduced to a singleton. The *set of regular unit signed measures* is denoted by $\text{regn}(A)$:

$$s \in \text{regn}(A) \iff s \in S^{|K|-1} \quad \text{and} \quad \exists! a \in A, F_A(s) = \{a\}. \quad (39)$$

For a regular unit signed measure $s \in S^{|K|-1}$, we denote by $f_A(s)$ the unique element of $F_A(s)$, so that $F_A(s) = \{f_A(s)\}$. The mapping

$$f_A : \text{regn}(A) \rightarrow \partial A, \quad \text{where } \text{regn}(A) \subset S^{|K|-1} \quad (40)$$

is called the *reverse spherical image map of A* , and is continuous (Schneider, 2014, p. 88). We have

$$s \in \text{regn}(A) \Rightarrow F_A(s) = \{f_A(s)\}. \quad (41)$$

Bodies with C^2 surface.

Proposition 9 (Schneider 2014, p. 113) *If the body A has boundary ∂A which is a C^2 submanifold of \mathbb{R}^K , then*

- *all points $a \in \partial A$ are regular ($\text{reg}(A) = \partial A$),*
- *the spherical image map n_A in (37) is defined over the whole boundary ∂A and is of class C^1 ,*
- *the spherical image map n_A has the reverse spherical image map f_A in (37) as right inverse, that is,*

$$n_A \circ f_A = \text{Id}_{\text{regn}(A)}. \quad (42)$$

Proof. The first two items can be found in (Schneider, 2014, p. 113). Now, we prove that $n_A \circ f_A = \text{Id}_{\text{regn}(A)}$. As $f_A : \text{regn}(A) \rightarrow \partial A$ by (40), and as $n_A : \partial A \rightarrow S^{|K|-1}$ by (37) since $\text{reg}(A) = \partial A$, the mapping $n_A \circ f_A : \text{regn}(A) \rightarrow S^{|K|-1}$ is well defined. Let $s \in \text{regn}(A)$. By (41), we have that $F_A(s) = \{f_A(s)\}$ and by (38), we have that $N_A(f_A(s)) = \mathbb{R}_+ n_A(f_A(s))$. From (52) — stating that exposed face and normal cone are conjugate — we deduce that $s \in \mathbb{R}_+ n_A(f_A(s))$. As $s \in S^{|K|-1}$, we conclude that $s = n_A(f_A(s))$ by (37). ■

Weingarten map. Let $a \in \text{reg}(A)$ be a regular point such that the spherical image map n_A in (37) is differentiable at a , with differential denoted by $T_a n_A$. The *Weingarten map* (Schneider, 2014, p. 113)

$$T_a n_A : T_a \partial A \rightarrow T_{n_A(a)} S^{|K|-1} \quad (43)$$

linearly maps the tangent space $T_a \partial A$ of the boundary ∂A at point a into the tangent space $T_{n_A(a)} S^{|K|-1}$ of the sphere $S^{|K|-1}$ at $n_A(a)$. The eigenvalues of the Weingarten map at a are called the *principal curvatures* of A at a (Schneider, 2014, p. 114); they are nonnegative (Schneider, 2014, p. 115). By definition, the body A has *positive curvature* at a if all principal curvatures at a are positive or, equivalently, if the Weingarten map is of maximal rank at a (Schneider, 2014, p. 115).

Reverse Weingarten map. Let $s \in \text{regn}(A)$ be a regular unit signed measure such that the reverse spherical image map f_A in (40) is differentiable at s , with differential denoted by $T_s f_A$. The *reverse Weingarten map*

$$T_s f_A : T_s S^{|K|-1} \rightarrow T_{f_A(s)} \partial A \quad (44)$$

maps the tangent space $T_s S^{|K|-1}$ of the sphere $S^{|K|-1}$ at s into the tangent space $T_{f_A(s)} \partial A$ of the boundary ∂A at point $f_A(s)$. The eigenvalues of the reverse Weingarten map at s are called the *principal radii of curvature* of A at s .

B Revisiting the model of Section 2

With the convex analysis tools recalled in section A.1, we revisit the model in Section 2 to prepare the proofs in Section C. We recall that $A \subset \mathbb{R}^K$ is a nonempty, convex and compact subset of \mathbb{R}^K , called the *action set*.

Support function. The *support function* σ_A of the action set A is defined by

$$\sigma_A(s) = \sup_{a \in A} \langle s, a \rangle, \quad \forall s \in \Sigma. \quad (45)$$

The value function $v_A : \Delta \rightarrow \mathbb{R}$ in (4) is the restriction of the support function σ_A to probability distributions:

$$v_A(p) = \sigma_A(p), \quad \forall p \in \Delta. \quad (46)$$

It is well known that σ_A is convex (as the supremum of the family of linear maps $\langle \cdot, a \rangle$ for $a \in A$). As the action set A is compact, $\sigma_A(s)$ takes finite values, hence its domain is Σ , hence σ_A is continuous.

(Exposed) face. For any signed measure $s \in \Sigma$, we let

$$F_A(s) = \arg \max_{a' \in A} \langle s, a' \rangle = \{a \in A \mid \forall a' \in A, \langle s, a' \rangle \leq \langle s, a \rangle\} \subset A \quad (47)$$

be the set of maximizers of $a \mapsto \langle s, a \rangle$ over A . We call $F_A(s)$ the *(exposed) face of A in the direction $s \in \Sigma$* . As the action set A is convex and compact, the face $F_A(s)$ of A in the direction s is nonempty, for any $s \in \Sigma$, and the face is a subset of the *boundary* ∂A of A :

$$F_A(s) \subset \partial A, \quad \forall s \in \Sigma. \quad (48)$$

The set $A^*(p)$ of optimal actions under belief p in (7) coincides with the (exposed) face $F_A(p)$ of A in the direction p in (47):

$$A^*(p) = F_A(p), \quad \forall p \in \Delta. \quad (49)$$

Normal cone. For any payoff vector a in A , we define

$$N_A(a) = \{s \in \Sigma \mid \forall a' \in A, \langle s, a' \rangle \leq \langle s, a \rangle\} \subset \Sigma. \quad (50)$$

We call $N_A(a)$ the *normal cone* to the closed convex set A at $a \in A$. Notice that $N_A(a)$ is made of signed measures, that are not necessarily beliefs. The set $\Delta_A^*(a)$ of beliefs compatible with optimal action a in (8) is related to the normal cone $N_A(a)$ at a in (50) by:

$$\Delta_A^*(a) = N_A(a) \cap \Delta, \quad \forall a \in A. \quad (51)$$

Conjugate subsets of actions and beliefs. Exposed face F_A and normal cone N_A are conjugate as follows:

$$s \in \Sigma \text{ and } a \in F_A(s) \iff a \in A \text{ and } s \in N_A(a). \quad (52)$$

C Proofs of the results in Section 3

Using the relations (49) and (51), we express the proofs of the results in Section 3 in terms of the sets $F_A(p)$ and $N_A(a)$ (in the set Σ of signed measures), instead of $A^*(p)$ and $\Delta_A^*(a)$ (in the set Δ of probability measures).

Value of information.

We have seen in (46) that the value function $v_A : \Delta \rightarrow \mathbb{R}$ in (4) is the restriction of the support function σ_A to beliefs in Δ . By definition (6) of the value of information, we deduce that, for any information structure \mathbf{q} , we have:

$$\mathbf{VoI}_A(\mathbf{q}) = \mathbb{E} [\sigma_A(\mathbf{q}) - \sigma_A(\bar{p})] . \quad (53)$$

Let us introduce, for all $q \in \Delta$,

$$\varphi_A^+(q) = \sigma_A(q) - \sigma_A(\bar{p}) + \sigma_{-A^*(\bar{p})}(q - \bar{p}) , \quad (54a)$$

$$\varphi_A^-(q) = \sigma_A(q) - \sigma_A(\bar{p}) - \sigma_{A^*(\bar{p})}(q - \bar{p}) . \quad (54b)$$

Proposition 10 *For any information structure \mathbf{q} , for any $a \in A$, we have that*

$$\mathbb{E} [\varphi_A^+(\mathbf{q})] = \mathbb{E} [\sigma_A(\mathbf{q}) - \sigma_A(\bar{p}) + \sigma_{-A^*(\bar{p})}(\mathbf{q} - \bar{p})] \quad (55a)$$

$$\geq \mathbf{VoI}_A(\mathbf{q}) = \mathbb{E} [\sigma_A(\mathbf{q}) - \sigma_A(\bar{p}) - \langle \mathbf{q} - \bar{p}, a \rangle] \quad (55b)$$

$$\geq \mathbb{E} [\sigma_A(\mathbf{q}) - \sigma_A(\bar{p}) - \sigma_{A^*(\bar{p})}(\mathbf{q} - \bar{p})] = \mathbb{E} [\varphi_A^-(\mathbf{q})] . \quad (55c)$$

Proof. By (54), we have, for all $q \in \Delta$,

$$\varphi_A^+(q) = \sigma_A(q) - \sigma_A(\bar{p}) + \sigma_{-A^*(\bar{p})}(q - \bar{p}) \quad (56a)$$

$$= \sup_{a \in A^*(\bar{p})} \left(\sigma_A(q) - \sigma_A(\bar{p}) - \langle q - \bar{p}, a \rangle \right) \quad (56b)$$

$$\geq \sigma_A(q) - \sigma_A(\bar{p}) - \langle q - \bar{p}, a \rangle , \quad \forall a \in A^*(\bar{p}) \quad (56c)$$

$$\geq \inf_{a \in A^*(\bar{p})} \left(\sigma_A(q) - \sigma_A(\bar{p}) - \langle q - \bar{p}, a \rangle \right) \quad (56d)$$

$$= \sigma_A(q) - \sigma_A(\bar{p}) - \sigma_{A^*(\bar{p})}(q - \bar{p}) = \varphi_A^-(q) . \quad (56e)$$

By taking the expectation, we obtain (55), using (53) and the property that $\mathbb{E} [\mathbf{q} - \bar{p}] = 0$. ■

Confidence set and indifference kernel.

We start by providing characterizations of the confidence set $\Delta_A^c(\bar{p})$ in (9) and of the indifference kernel $\Sigma_A^i(\bar{p})$ in (14), in terms of the sets $F_A(p)$ in (47) and $N_A(a)$ in (50).

Proposition 11

1. *The confidence set $\Delta_A^c(\bar{p})$ of (9) is the nonempty closed and convex set*

$$\Delta_A^c(\bar{p}) = \bigcap_{a \in A^*(\bar{p})} \Delta_A^*(a) = \bigcap_{a \in F_A(\bar{p})} N_A(a) \cap \Delta . \quad (57)$$

2. Let $p \in \Delta$. We have that

$$p \in \Delta_A^c(\bar{p}) \iff F_A(\bar{p}) \subset F_A(p) \quad (58a)$$

$$\iff \sigma_A(p) - \sigma_A(\bar{p}) - \langle p - \bar{p}, a \rangle = 0, \quad \forall a \in F_A(\bar{p}) \quad (58b)$$

$$\iff \sigma_A(p) - \sigma_A(\bar{p}) + \sigma_{-A^*(p)}(p - \bar{p}) = 0. \quad (58c)$$

3. The indifference kernel $\Sigma_A^i(\bar{p})$ of (14) is the nonempty vector subspace

$$\Sigma_A^i(\bar{p}) = [F_A(\bar{p}) - F_A(\bar{p})]^\perp = [A^*(\bar{p}) - A^*(\bar{p})]^\perp = \bigcap_{a \in F_A(\bar{p})} N_{F_A(\bar{p})}(a). \quad (59)$$

Proof.

1. Express (9) using (51).

2. We prove the three equivalences in (58).

(a) Let $p \in \Delta$. Using the property (52) that exposed face F_A and normal cone N_A are conjugate, we obtain:

$$\begin{aligned} p \in \Delta_A^c(\bar{p}) &\iff p \in \bigcap_{a \in F_A(p)} N_A(a) \text{ by (57)} \\ &\iff a \in F_A(p), \quad \forall a \in F_A(\bar{p}) \text{ by (52)} \\ &\iff F_A(\bar{p}) \subset F_A(p). \end{aligned}$$

(b) Let $p \in \Delta$. We have that

$$\begin{aligned} \sigma_A(p) - \sigma_A(\bar{p}) - \langle p - \bar{p}, a \rangle &= 0, \quad \forall a \in F_A(\bar{p}) \\ \iff \sigma_A(p) &= \langle p, a \rangle, \quad \forall a \in F_A(\bar{p}) \end{aligned}$$

because $\sigma_A(\bar{p}) = \langle \bar{p}, a \rangle$ for any $a \in F_A(\bar{p})$, since $F_A(\bar{p})$ is the set $A^*(\bar{p})$ of optimal actions under belief \bar{p} by (7) and (47)

$$\begin{aligned} \iff p &\in \bigcap_{a \in F_A(p)} N_A(a) \text{ by (30)} \\ \iff p &\in \bigcap_{a \in F_A(p)} N_A(a) \cap \Delta = \Delta_A^c(\bar{p}) \text{ by (57)}. \end{aligned}$$

(c) For any $a \in A$, we define the function

$$\varphi_a(q) = \sigma_A(q) - \sigma_A(\bar{p}) - \langle q - \bar{p}, a \rangle, \quad \forall q \in \Delta. \quad (60)$$

By (31) and (58b), we have that

$$\forall a \in F_A(\bar{p}), \forall q \in \Delta, \varphi_a(q) \geq 0, \quad (61a)$$

$$\forall a \in F_A(\bar{p}), \forall q \in \Delta_A^c(\bar{p}), \varphi_a(q) = 0. \quad (61b)$$

Let $p \in \Delta$. Using (61a), we deduce from (58b) and from the compacity of $F_A(\bar{p})$ that

$$p \in \Delta_A^c(\bar{p}) \iff \inf_{a \in F_A(\bar{p})} \left(\sigma_A(p) - \sigma_A(\bar{p}) - \langle p - \bar{p}, a \rangle \right) = 0. \quad (62)$$

We conclude with (56d)–(56e).

3. Express (14) using (49). Then, use the definition (28) of $N_{F_A(\bar{p})}(a)$.

This ends the proof. ■

We have the following inclusion between the confidence set $\Delta_A^c(\bar{p})$ in (9) and the indifference kernel $\Sigma_A^i(\bar{p})$:

$$\Delta_A^c(p) \subset \Sigma_A^i(p) \cap \Delta. \quad (63)$$

The above inclusion is strict in general. Indeed, consider a case where $F_A(p)$ is a singleton $\{a\}$. Then, on the one hand, $\Delta_A^c(p) = N_A(a) \cap \Delta$. However, on the other hand, we have by (59):

$$\Sigma_A^i(p) \cap \Delta = N_{F_A(p)}(a) \cap \Delta = N_{\{a\}}(a) \cap \Delta = \Delta.$$

As soon as $N_A(a) \cap \Delta$ only contains the belief p , we have that $\{p\} = \Delta_A^c(p)$ and $\Sigma_A^i(p) \cap \Delta = \Delta$. As an example, consider the case where the set A is the unit ball:

$$A = B(0, 1), \quad \sigma_A(s) = \|s\|, \quad N_A(a) = \mathbb{R}_+ a, \quad F_A(s) = \left\{ \frac{s}{\|s\|} \right\}, \quad N_{F_A(s)}(a) = \mathbb{R}^2, \quad (64)$$

so that $\Delta_A^c(p) = \{p\}$, $\Sigma_A^i(p) \cap \Delta = \Delta$.

C.1 Valuable information

Proof of Proposition 1.

Let $a \in F_A(\bar{p})$ and \mathbf{q} be an information structure. We have that

$$\mathbf{VoI}_A(\mathbf{q}) = 0 \iff \mathbb{E} [\sigma_A(\mathbf{q}) - \sigma_A(\bar{p})] = 0 \text{ by (53)} \quad (65a)$$

$$\iff \mathbb{E} [\sigma_A(\mathbf{q}) - \sigma_A(\bar{p}) - \langle \mathbf{q} - \bar{p}, a \rangle] = 0, \quad \text{as } \mathbb{E} [\mathbf{q} - \bar{p}] = 0 \quad (65b)$$

$$\iff \sigma_A(\mathbf{q}) - \sigma_A(\bar{p}) - \langle \mathbf{q} - \bar{p}, a \rangle = 0, \quad \mathbb{P} - \text{a.s.} \quad (65c)$$

because $\sigma_A(\mathbf{q}) - \sigma_A(\bar{p}) - \langle \mathbf{q} - \bar{p}, a \rangle \geq 0$ by (31) since $a \in F_A(\bar{p})$

$$\iff \sigma_A(\mathbf{q}) = \langle \mathbf{q}, a \rangle, \quad \mathbb{P} - \text{a.s.} \quad (65d)$$

because $\sigma_A(\bar{p}) = \langle \bar{p}, a \rangle$ since $a \in F_A(\bar{p})$

$$\iff \mathbb{P} \{a \in F_A(\mathbf{q})\} = 1 \quad (65e)$$

$$\iff \mathbb{P} \{\langle \mathbf{q}, a' - a \rangle \leq 0, \forall a' \in A\} = 1. \quad (65f)$$

Let $F \subset F_A(\bar{p})$ be a dense subset of the compact $F_A(\bar{p})$ of \mathbb{R}^K . We immediately get from the last equality that

$$\mathbf{VoI}_A(\mathbf{q}) = 0 \Rightarrow \mathbb{P} \{\langle \mathbf{q}, a' - a \rangle \leq 0, \forall a' \in A, \forall a \in F\} = 1. \quad (66a)$$

As the set $\{a \in F_A(\bar{p}) \mid \langle \mathbf{q}, a' - a \rangle \leq 0, \forall a' \in A\}$ is closed (for any outcome in the underlying sample space), we get that

$$\{\langle \mathbf{q}, a' - a \rangle \leq 0, \forall a' \in A, \forall a \in F\} \subset \{\langle \mathbf{q}, a' - a \rangle \leq 0, \forall a' \in A, \forall a \in \bar{F}\}. \quad (66b)$$

We deduce from the last equality that

$$\mathbf{VoI}_A(\mathbf{q}) = 0 \Rightarrow \mathbb{P} \{\langle \mathbf{q}, a' - a \rangle \leq 0, \forall a' \in A, \forall a \in \bar{F}\} = 1. \quad (66c)$$

Now, since $\bar{F} = F_A(\bar{p})$, we finally get that

$$\mathbf{VoI}_A(\mathbf{q}) = 0 \Rightarrow \mathbb{P} \{\langle \mathbf{q}, a' - a \rangle \leq 0, \forall a' \in A, \forall a \in F_A(\bar{p})\} = 1. \quad (66d)$$

In other words, we have obtained that, by definition (50) of the normal cone $N_A(a)$:

$$\mathbf{VoI}_A(\mathbf{q}) = 0 \Rightarrow \mathbf{q} \in \bigcap_{a \in F_A(\bar{p})} N_A(a), \quad \mathbb{P} - \text{a.s.} . \quad (66e)$$

Since $\mathbf{q} \in \Delta$, we conclude by (57) that

$$\mathbf{VoI}_A(\mathbf{q}) = 0 \Rightarrow \mathbf{q} \in \bigcap_{a \in F_A(p)} N_A(a) \cap \Delta = \bigcap_{a \in A^*(p)} \Delta_A^*(a) = \Delta_A^c(p). \quad (66f)$$

Revisiting the proof backward, or using (58b), we easily see that

$$\mathbf{q} \in \Delta_A^c(p), \quad \mathbb{P} - \text{a.s.} \Rightarrow \mathbf{VoI}_A(\mathbf{q}) = 0. \quad (67)$$

This ends the proof. ■

Proof of Theorem 2.

Let \mathbf{q} be an information structure.

First, we show the upper estimate $C_A \mathbb{E} d(\mathbf{q}, \Delta_A^c(\bar{p})) \geq \mathbf{VoI}_A(\mathbf{q})$ in (13). For this purpose, we consider $a \in A$ and we show that the function φ_a in (60) is such that

$$\varphi_a(q) \leq \sup_{a' \in A} \|a - a'\| \inf_{p \in \Delta_A^c(\bar{p})} \|p - q\|. \quad (68)$$

Indeed, we have that, for any $p \in \Delta_A^c(\bar{p})$,

$$\varphi_a(q) = \varphi_a(q) - \varphi_a(p) \text{ by (61b) since } p \in \Delta_A^c(\bar{p}) \quad (69a)$$

$$= \sigma_A(q) - \sigma_A(p) - \langle q - p, a \rangle \text{ by (60)} \quad (69b)$$

$$= \sigma_{A-a}(q) - \sigma_{A-a}(p) \text{ by (26)} \quad (69c)$$

$$\leq \sup_{a' \in A-a} \|a'\| \times \|p - q\| \text{ by (26)} \quad (69d)$$

$$= \sup_{a' \in A} \|a - a'\| \times \|p - q\|. \quad (69e)$$

By taking the infimum with respect to all $p \in \Delta_A^c(\bar{p})$, we obtain (68). Then, we deduce that

$$\mathbf{VoI}_A(\mathbf{q}) = \mathbb{E} [\varphi_a(\mathbf{q})], \quad \forall a \in A \text{ by (55b)} \quad (70a)$$

$$= \inf_{a \in A} \mathbb{E} [\varphi_a(\mathbf{q})] \quad (70b)$$

$$\leq \inf_{a \in A} \sup_{a' \in A} \|a - a'\| \times \mathbb{E} \left[\inf_{p \in \Delta_A^c(\bar{p})} \|p - q\| \right] \text{ by (68)}. \quad (70c)$$

With $C_A = \inf_{a \in A} \sup_{a' \in A} \|a - a'\|$ and (12), this gives the upper estimate $C_A \mathbb{E} d(\mathbf{q}, \Delta_A^c(\bar{p})) \geq \mathbf{VoI}_A(\mathbf{q})$ in (13).

Second, we show the lower estimate $\mathbf{VoI}_A(\mathbf{q}) \geq c_{\bar{p}, A, \varepsilon} \mathbb{P}\{\mathbf{q} \notin \Delta_{A, \varepsilon}^c(\bar{p})\}$ in (13).

We consider an open subset \mathcal{Q} of Δ that contains the confidence set $\Delta_A^c(p)$, that is, $\Delta_A^c(\bar{p}) \subset \mathcal{Q}$. By Lemma 12 right below, there exists an $a \in F_A(\bar{p})$ such that the continuous function φ_a in (60) is strictly positive on $\Delta_A^c(\bar{p})^c$. As $\mathcal{Q}^c \subset \Delta_A^c(\bar{p})^c$ and \mathcal{Q}^c is a closed subset of the compact Δ , we can define

$$c_{\bar{p}, A} = \inf_{p \notin \mathcal{Q}} \varphi_a(p) > 0. \quad (71)$$

We deduce that

$$\mathbf{VoI}_A(\mathbf{q}) = \mathbb{E} [\varphi_a(\mathbf{q})] \text{ by (55b)} \quad (72a)$$

$$= \mathbb{E} [\mathbf{1}_{\mathbf{q} \in \Delta_A^c(\bar{p})} \varphi_a(\mathbf{q}) + \mathbf{1}_{\mathbf{q} \notin \Delta_A^c(\bar{p})} \varphi_a(\mathbf{q})] \quad (72b)$$

$$= \mathbb{E} [\mathbf{1}_{\mathbf{q} \notin \Delta_A^c(\bar{p})} \varphi_a(\mathbf{q})] \text{ by (61b)} \quad (72c)$$

$$\geq \mathbb{E} [\mathbf{1}_{\mathbf{q} \notin \mathcal{Q}} \varphi_a(\mathbf{q})] \quad (72d)$$

$$\geq \mathbb{E} [\mathbf{1}_{\mathbf{q} \notin \mathcal{Q}^c} c_{\bar{p}, A}] = c_{\bar{p}, A} \mathbb{P}\{\mathbf{q} \notin \mathcal{Q}\}. \quad (72e)$$

With $\mathcal{Q} = \Delta_{A, \varepsilon}^c(\bar{p})$, we put

$$c_{\bar{p}, A, \varepsilon} = \inf_{p \notin \Delta_{A, \varepsilon}^c(\bar{p})} \varphi_a(p) > 0. \quad (73)$$

This ends the proof. ■

Lemma 12 *There exists at least one $a \in F_A(\bar{p})$ such that the function φ_a in (60) is strictly positive on the complementary set $\Delta_A^c(\bar{p})^c$.*

Proof. We consider two cases, depending whether $F_A(\bar{p})$ is a singleton or not.

Suppose that $F_A(\bar{p})$ is a singleton $\{a\}$. By (58b), we have that

$$q \notin \Delta_A^c(\bar{p}) \iff \varphi_a(q) > 0. \quad (74)$$

Suppose that $F_A(\bar{p})$ is not a singleton. Recall that the *relative interior* $\text{ri}(C)$ of a nonempty convex set $C \subset \mathbb{R}^K$ is the nonempty interior of C for the topology relative to the affine hull $\text{aff}(C)$ (Hiriart-Uruty and Lemaréchal, 1993, p. 103). We prove that any $a \in \text{ri}(F_A(q))$ answers the question. Let $a \in \text{ri}(F_A(q))$ be fixed. For any $q \notin \Delta_A^c(\bar{p})$, by (58a) we have that $F_A(\bar{p}) \not\subset F_A(q)$. Therefore, there exists $\bar{a} \in F_A(\bar{p})$ such that $\bar{a} \notin F_A(q)$, that is, such that $\sigma_A(q) > \langle q, \bar{a} \rangle$. As $a \in \text{ri}(F_A(q))$, there exists $a' \in \text{ri}(F_A(q))$ such that $a = \lambda a' + (1 - \lambda)\bar{a}$ for a certain $\lambda \in]0, 1[$. Since $\sigma_A(q) \geq \langle q, a' \rangle$ (by definition (45) of σ_A) and $\sigma_A(q) > \langle q, \bar{a} \rangle$ (as $\bar{a} \notin F_A(q)$), we deduce that

$$\sigma_A(q) = \lambda \sigma_A(q) + (1 - \lambda)\sigma_A(q) > \lambda \langle q, a' \rangle + (1 - \lambda) \langle q, \bar{a} \rangle = \langle q, a \rangle, \quad (75)$$

where we used the property that $\lambda \in]0, 1[$. Using the definition (60) of the function φ_a , we have obtained that $q \notin \Delta_A^c(\bar{p}) \Rightarrow \varphi_a(q) > 0$.

This ends the proof. ■

C.2 Undecided

Proof of Proposition 3.

We prove that the face $F_A(\bar{p})$ of A in the direction $\bar{p} \in \Delta$ is a singleton if and only if the value function v_A in (4) is differentiable at \bar{p} .

- Suppose that the face $F_A(\bar{p})$ of A in the direction $\bar{p} \in \Delta$ is a singleton.

As the face $F_A(\bar{p})$ is the subdifferential at \bar{p} of the support function σ_A (Hiriart-Uruty and Lemaréchal, 1993, p. 258), we deduce that σ_A is differentiable at \bar{p} (Hiriart-Uruty and Lemaréchal, 1993, p. 251). Therefore, the value function v_A in (4) is differentiable at \bar{p} , since $v_A : \Delta \rightarrow \mathbb{R}$ is the restriction of σ_A to probability distributions Δ , as in (46).

- Suppose the value function v_A in (4) is differentiable at \bar{p} .

We consider the *extended value function* defined by

$$\tilde{v}_A : \mathbb{R}_+^* \Delta \rightarrow \mathbb{R}, \quad s \mapsto \|s\| v_A\left(\frac{s}{\|s\|}\right). \quad (76)$$

Since the support function σ_A is positively homogeneous, we have that $\tilde{v}_A : \mathbb{R}_+^* \Delta \rightarrow \mathbb{R}$ is the restriction of σ_A to the cone $\mathbb{R}_+^* \Delta$:

$$\tilde{v}_A(s) = \sigma_A(s), \quad \forall s \in \mathbb{R}_+^* \Delta. \quad (77)$$

As the prior \bar{p} has full support, the extended value function \tilde{v}_A in (76) is well defined on a neighborhood of \bar{p} and is differentiable at \bar{p} , since v_A is also. Since $\tilde{v}_A : \mathbb{R}_+^* \Delta \rightarrow \mathbb{R}$

is the restriction of σ_A to the cone $\mathbb{R}_+^* \Delta$, we deduce that the support function σ_A is differentiable at \bar{p} .

Since, on the one hand, a convex function with domain \mathbb{R}^K is differentiable at \bar{p} if and only if the subdifferential at \bar{p} is a singleton (Hiriart-Uruty and Lemaréchal, 1993, p. 251), and, on the other hand, the face $F_A(\bar{p})$ is the subdifferential at \bar{p} of the support function σ_A (Hiriart-Uruty and Lemaréchal, 1993, p. 258), we conclude that the face $F_A(\bar{p})$ of A in the direction $\bar{p} \in \Delta$ is a singleton.

This ends the proof. ■

Proof of Theorem 4.

We prove the three inequalities in (17).

A). We prove the upper inequality $C_A \mathbb{E} \|\mathbf{q} - \bar{p}\| \geq \mathbf{VoI}_A(\mathbf{q})$ in (17).

By definition (45) of a support function, we have that $\sigma_A(\cdot) \leq \|A\| \times \|\cdot\|$, where $\|A\| = \sup\{\|a\|, a \in A\} < +\infty$. Thus $C_A = \|A\|$ in the left hand side inequality in (17).

B). We prove the middle inequality $\mathbf{VoI}_A(\mathbf{q}) \geq \mathbf{VoI}_{A^*(\bar{p})}(\mathbf{q})$ in (17).

For all $s \in \Sigma$, we have that

$$\sigma_A(s) - \sigma_A(\bar{p}) \geq \sigma_{F_A(\bar{p})}(s - \bar{p}) \text{ by (31) since } F_A(\bar{p}) \neq \emptyset \quad (78a)$$

$$= \langle s - \bar{p}, a \rangle, \quad \forall a \in F_A(\bar{p}) \text{ by definition of } \sigma_{F_A(\bar{p})} \quad (78b)$$

$$= \sigma_{F_A(\bar{p})}(s) - \sigma_{F_A(\bar{p})}(\bar{p}) \text{ by definition of } \sigma_{F_A(\bar{p})}. \quad (78c)$$

By taking the expectation \mathbb{E} , we obtain that

$$\mathbf{VoI}_A(\mathbf{q}) = \mathbb{E} [\sigma_A(\mathbf{q}) - \sigma_A(\bar{p})] \text{ by (6) and (46)} \quad (79a)$$

$$\geq \mathbb{E} [\sigma_{F_A(\bar{p})}(\mathbf{q} - \bar{p})] \text{ by (78a)} \quad (79b)$$

$$= \mathbb{E} [\sigma_{F_A(\bar{p})}(\mathbf{q}) - \sigma_{F_A(\bar{p})}(\bar{p})] \text{ by (78c)} \quad (79c)$$

$$= \mathbf{VoI}_{F_A(\bar{p})}(\mathbf{q}) \text{ by (6) and (46)}.$$

This ends the proof of the middle inequality.

C). We prove the right hand side inequality $\mathbf{VoI}_{A^*(\bar{p})}(\mathbf{q}) \geq \mathbb{E} \|\mathbf{q} - \bar{p}\|_{\Sigma_A^i(\bar{p})}$ in (17).

For this purpose, we recall that the *affine hull* $\text{aff}(S)$ of a subset S of \mathbb{R}^K is the intersection of all affine manifolds containing S . Let n be the dimension of the affine hull $\text{aff}(F_A(\bar{p}))$ of $F_A(\bar{p})$, and let a_1, \dots, a_n be n actions in $F_A(\bar{p})$ that generate $\text{aff}(F_A(\bar{p}))$. We put

$$T = \{a_1, \dots, a_n\} \subset F_A(\bar{p}) \text{ so that } \text{aff}(F_A(\bar{p})) = \text{aff}\{a_1, \dots, a_n\} = \text{aff}(T). \quad (80)$$

We will now show that

$$\|\cdot\|_{\Sigma_A^i(\bar{p})} = \frac{1}{n} \sigma_{T-T}(\cdot) \quad (81)$$

is a semi-norm with kernel $(F_A(\bar{p}) - F_A(\bar{p}))^\perp$ that satisfies the right hand side inequality in (17).

First, the support function σ_{T-T} is a semi-norm with kernel $(T - T)^\perp$ by item 4 in Proposition 8. Now, we can easily see that, for any subset $S \subset \mathbb{R}^K$, one has

$$(S - S)^\perp = (\text{aff}(S - S))^\perp = (\text{aff}(S) - \text{aff}(S))^\perp . \quad (82)$$

Using these equalities with $S = T$ and $S = F_A(\bar{p})$, we deduce that $(T - T)^\perp = (F_A(\bar{p}) - F_A(\bar{p}))^\perp$, since $\text{aff}(T) = \text{aff}(F_A(\bar{p}))$ by (80).

Second, we show that the right hand side inequality in (17) is satisfied. We have that

$$\mathbf{VoI}_A(\mathbf{q}) \geq \mathbb{E} [\sigma_{F_A(\bar{p})}(\mathbf{q} - \bar{p})] \quad \text{by (79b)} \quad (83a)$$

$$\geq \mathbb{E} [\sigma_T(\mathbf{q} - \bar{p})] \quad (83b)$$

because $T \subset F_A(\bar{p})$ and support functions (26) are monotone with respect to set inclusion,

$$= \mathbb{E} [\sigma_T(\mathbf{q} - \bar{p}) - \langle \mathbf{q} - \bar{p}, a \rangle] , \quad \forall a \in A \quad \text{because } \mathbb{E} [\langle \mathbf{q} - \bar{p}, a \rangle] = 0. \quad (83c)$$

$$= \mathbb{E} [\sigma_{T-a}(\mathbf{q} - \bar{p})] , \quad \forall a \in A \quad \text{because } \sigma_{T-a} = \sigma_{T+\{-a\}} = \sigma_T + \sigma_{\{-a\}}. \quad (83d)$$

Indeed, support functions transform a Minkowski sum of sets into a sum of support functions (Hiriart-Uruty and Lemaréchal, 1993, p. 226). Using again this property, we obtain that

$$\mathbf{VoI}_A(\mathbf{q}) \geq \frac{1}{n} \sum_{i=1}^n \mathbb{E} [\sigma_{T-a_i}(\mathbf{q} - \bar{p})] = \frac{1}{n} \mathbb{E} [\sigma_{\sum_{i=1}^n (T-a_i)}(\mathbf{q} - \bar{p})] . \quad (84)$$

Now, as $T = \{a_1, \dots, a_n\}$, it is easy to see that the sum $\sum_{i=1}^n (T - a_i)$ contains any element of the form $a_k - a_l$:

$$a_k - a_l = (a_1 - a_1) + \dots + (a_{l-1} - a_{l-1}) + (a_k - a_l) + (a_{l+1} - a_{l+1}) + \dots + (a_n - a_n) \in \sum_{i=1}^n (T - a_i) .$$

As support functions are monotone with respect to set inclusion, we deduce that

$$\sigma_{\sum_{i=1}^n (T-a_i)} \geq \sigma_{\{a_k - a_l, k, l=1, \dots, n\}} = \sigma_{T-T} , \quad (85)$$

and that

$$\mathbf{VoI}_A(\mathbf{q}) \geq \frac{1}{n} \mathbb{E} [\sigma_{\{a_k - a_l, k, l=1, \dots, n\}}(\mathbf{q} - \bar{p})] = \frac{1}{n} \mathbb{E} [\sigma_{T-T}(\mathbf{q} - \bar{p})] = \mathbb{E} \|\mathbf{q} - \bar{p}\|_{\Sigma_A^i(\bar{p})} . \quad (86)$$

This ends the proof. ■

C.3 Flexible decisions

Proof of Proposition 5.

All the reminders on geometric convex analysis in Section A.2 were done with outer normal vectors belonging to the unit sphere of signed measures. Now, as we work with beliefs — positive measures of mass 1 — we are going to adapt these concepts.

We will consider the diffeomorphism

$$\nu : S^{|K|-1} \cap \mathbb{R}_+^K \rightarrow \Delta, \quad s \mapsto \frac{s}{\langle s, 1 \rangle}, \quad (87)$$

that maps unit positive measures into probability measures, with inverse

$$\nu^{-1} : \Delta \rightarrow S^{|K|-1} \cap \mathbb{R}_+^K, \quad p \mapsto \frac{p}{\|p\|}. \quad (88)$$

Since, by assumption, the action set A has boundary ∂A which is a C^2 submanifold of \mathbb{R}^K , we know by Proposition 9 that the spherical image map $n_A : \partial A \rightarrow S^{|K|-1}$ in (37) is well defined, is of class C^1 , and has for right inverse the reverse spherical image map $f_A : \text{regn}(A) \rightarrow \partial A$ in (40), that is, $n_A \circ f_A = \text{Id}_{\text{regn}(A)}$.

The *set of relevant regular points* is the subset of the set $\text{reg}(A)$ of regular points defined by

$$a \in \text{reg}^+(A) \iff \exists p \in \Delta, \quad N_A(a) = \mathbb{R}_+ p. \quad (89)$$

For a regular action $a \in \text{reg}^+(A)$, there is only one probability $p \in \Delta$ such that $N_A(a) = \mathbb{R}_+ p$, and it is $p = \nu(n_A(a))$. We have

$$a \in \text{reg}^+(A) \Rightarrow N_A(a) = \mathbb{R}_+ \nu(n_A(a)) \quad \text{where } \nu(n_A(a)) \in \Delta. \quad (90)$$

The *set of regular probabilities* is

$$\text{regn}^+(A) = \left(\mathbb{R}_+^* \text{regn}(A) \right) \cap \Delta. \quad (91)$$

For a regular probability $p \in \text{regn}^+(A)$, there is only one action $a \in \partial A$ such that $F_A(p) = \{a\}$, and it is $a = f_A(\nu^{-1}(p))$. Indeed, by definition (47) of the (exposed) face, we have that

$$F_A(\lambda s) = F_A(s), \quad \forall \lambda \in \mathbb{R}_+^*, \quad \forall s \in \Sigma, \quad s \neq 0. \quad (92)$$

Therefore, we have that

$$p \in \text{regn}^+(A) \Rightarrow F_A(p) = \{f_A(\nu^{-1}(p))\}. \quad (93)$$

The following mappings are well defined:

$$\nu \circ n_A : \text{reg}^+(A) \rightarrow \Delta \quad \text{and} \quad f_A \circ \nu^{-1} : \text{regn}^+(A) \rightarrow \partial A, \quad (94)$$

and we have that

$$(\nu \circ n_A) \circ (f_A \circ \nu^{-1}) = \text{Id}_{\text{regn}^+(A)}. \quad (95)$$

- Item 2 \Rightarrow Item 1.

Suppose that the face $F_A(\bar{p})$ is a singleton $\{a^\sharp\}$ and the curvature of the boundary ∂A of payoffs at a^\sharp is positive.

Since, by assumption, the action set A has boundary ∂A which is a C^2 submanifold of \mathbb{R}^K , we know that the spherical image map n_A in (37) is defined over the whole boundary ∂A and is of class C^1 , and its differential is the Weingarten map.

As the curvature of the boundary ∂A of payoffs at a^\sharp is positive, the Weingarten map $T_{a^\sharp}n_A$ is of maximal rank at a^\sharp (Schneider, 2014, p. 115). Therefore, by the inverse function theorem, there exists an open neighborhood \mathcal{A} of a^\sharp in A such that $n_A(\mathcal{A})$ is an open neighborhood of $n_A(a^\sharp)$ in $S^{|K|-1}$, and such that the restriction $n_A : \mathcal{A} \rightarrow n_A(\mathcal{A})$ of the spherical image map in (37) is a diffeomorphism. By (42), we have that $n_A(a^\sharp) = \frac{\bar{p}}{\|\bar{p}\|}$ and the local inverse coincides with the restriction $f_A : n_A(\mathcal{A}) \rightarrow \mathcal{A}$ of the reverse spherical image map in (40).

As $n_A(\mathcal{A})$ is an open neighborhood of $\frac{\bar{p}}{\|\bar{p}\|}$ in $S^{|K|-1}$, and as the prior \bar{p} has full support, we deduce that $\nu(n_A(\mathcal{A}))$ is an open neighborhood of \bar{p} in Δ , where the diffeomorphism ν is defined in (87).

We easily deduce that $f_A \circ \nu^{-1} : \nu(n_A(\mathcal{A})) \rightarrow \mathcal{A}$ is a diffeomorphism. By (93), we conclude that $f_A \circ \nu^{-1}$ is the restriction of the set-valued mapping $F_A : \Delta \rightrightarrows A$, $p \mapsto F_A(p)$ in (18).

- Item 1 \Rightarrow Item 3.

Suppose that the set-valued mapping $F_A : \Delta \rightrightarrows A$, $p \mapsto F_A(p)$ in (18) is a local diffeomorphism at \bar{p} .

By definition (39) of the set of regular unit signed measures, there exists an open neighborhood Π of \bar{p} in Δ such that $\Pi \subset \text{regn}^+(A)$, where the set of relevant regular points is defined in (89). In addition, the mapping $f_A \circ \nu^{-1} : \Pi \rightarrow f_A(\nu^{-1}(\Pi))$ is a diffeomorphism.

As $F_A(p) = \{f_A(\nu^{-1}(p))\}$, for all beliefs $p \in \Pi$, we know that the support function σ_A is differentiable and that its derivative is $\nabla_p \sigma_A = f_A(\nu^{-1}(p))$ (Hiriart-Uruty and Lemaréchal, 1993, p. 251). As $f_A \circ \nu^{-1}$ is a local diffeomorphism at \bar{p} , and as the mapping ν in (87) is a diffeomorphism, we deduce that the support function σ_A is twice differentiable with Hessian having full rank. As the value function v_A is the restriction of σ_A to Δ , we conclude that v_A is twice differentiable at \bar{p} and the Hessian is definite positive.

- Item 3 \Rightarrow Item 2.

Suppose that the value function v_A is twice differentiable at \bar{p} and the Hessian is definite positive.

On the one hand, as the prior \bar{p} has full support, there exists an open neighborhood Π of \bar{p} in Δ such that v_A is differentiable on Π . On the other hand, as the support

function σ_A is positively homogeneous, and by (46), we have that

$$\sigma_A(s) = \langle s, 1 \rangle v_A \circ \nu(s), \quad \forall s \in S^{|K|-1} \cap \mathbb{R}_+^K. \quad (96)$$

Therefore, as the mapping ν in (87) is a diffeomorphism, the support function σ_A is differentiable on the open neighborhood $\nu^{-1}(\mathbf{II})$ of $\nu^{-1}(\bar{p}) = \frac{\bar{p}}{\|\bar{p}\|}$ in $S^{|K|-1} \cap \mathbb{R}_+^K$.

Since, on the one hand, a convex function with domain \mathbb{R}^K is differentiable at s if and only if the subdifferential at s is a singleton (Hiriart-Uruty and Lemaréchal, 1993, p. 251), and, on the other hand, the face $F_A(s)$ is the subdifferential at s of the support function σ_A (Hiriart-Uruty and Lemaréchal, 1993, p. 258), we conclude that the face $F_A(s)$ of A in the direction $s \in \nu^{-1}(\mathbf{II})$ is a singleton.

Therefore, by definition (39) of the set of regular unit signed measures, we have that $\nu^{-1}(\mathbf{II}) \subset \text{regn}(A)$. In addition, the restriction $f_A : \nu^{-1}(\mathbf{II}) \rightarrow f_A(\nu^{-1}(\mathbf{II}))$ of the reverse spherical image map in (40) is well defined, and we have that

$$\nabla_s \sigma_A = f_A(s), \quad \forall s \in \nu^{-1}(\mathbf{II}). \quad (97)$$

Therefore, the mapping $f_A : \nu^{-1}(\mathbf{II}) \rightarrow f_A(\nu^{-1}(\mathbf{II}))$ is differentiable at $\nu^{-1}(\bar{p}) = \frac{\bar{p}}{\|\bar{p}\|}$, and has full rank. Indeed, σ_A is twice differentiable at $\nu^{-1}(\bar{p}) = \frac{\bar{p}}{\|\bar{p}\|}$, and the Hessian is definite positive. This comes from (96), where the mapping ν in (87) is a C^∞ diffeomorphism and the value function v_A is twice differentiable at \bar{p} with definite positive Hessian.

As f_A is differentiable at $\frac{\bar{p}}{\|\bar{p}\|}$ and has full rank, the reverse Weingarten map $T_s f_A$ in (44) is well defined and has full rank. Therefore, the principal radii of curvature of A at $\frac{\bar{p}}{\|\bar{p}\|}$ are positive. Letting $a^\sharp = f_A(\frac{\bar{p}}{\|\bar{p}\|})$, we conclude that $F_A(\bar{p}) = \{a^\sharp\}$ and that the curvature of the boundary ∂A of payoffs at a^\sharp is positive.

This ends the proof. ■

Proof of Theorem 6.

We suppose that the value function v_A is twice differentiable at \bar{p} and the Hessian is definite positive. We also denote $F_A(\bar{p}) = \{a^\sharp\}$.

First, we show that the function

$$g(p) = \frac{v_A(p) - v_A(\bar{p}) - \langle p - \bar{p}, a^\sharp \rangle}{\|p - \bar{p}\|^2} \quad (98)$$

is continuous and positive on Δ . Indeed, g is continuous on $\Delta \setminus \{\bar{p}\}$, and also at \bar{p} since the value function v_A is twice differentiable at \bar{p} . In addition, $g(\bar{p}) > 0$ since the Hessian of v_A at \bar{p} is definite positive. We have $g \geq 0$ on $\Delta \setminus \{\bar{p}\}$, because $F_A(\bar{p}) = \{a^\sharp\}$ is the subdifferential at \bar{p} of the support function σ_A , and by (46). We now prove by contradiction that $g > 0$. If there existed a belief $p \neq \bar{p}$ such that $g(p) = 0$, we would have $v_A(p) - v_A(\bar{p}) - \langle p - \bar{p}, a^\sharp \rangle = 0$; this equality would then hold true over the whole segment $[p, \bar{p}]$, and we would conclude

that the second derivative of v_A at \bar{p} along the direction $p - \bar{p}$ would be zero; this would contradict the assumption that the Hessian of v_A at \bar{p} is definite positive. Therefore, we conclude that $g > 0$.

Second, letting $C_{\bar{p},A} > 0$ and $c_{\bar{p},A} > 0$ be the maximum and the minimum of the function $g > 0$ on the compact set Δ , we easily deduce (19) from (6).

This ends the proof. ■