PAYOFFS-BELIEFS DUALITY AND THE VALUE OF INFORMATION

MICHEL DE LARA* AND OLIVIER GOSSNER†

Abstract. In decision problems under incomplete information, actions (identified to 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 concepts and tools from convex analysis. We define the value function as the support function of the set of available actions: the subdifferential at a belief is the set of optimal actions at this belief; the set of beliefs at which an action is optimal is the normal cone of the set of available actions at this point. Our main results are 1) a necessary and sufficient condition for positive value of information 2) 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. We apply our results to the marginal value of information at the null, that is, when the agent is close to receiving no information at all, and we provide conditions under which the marginal value of information is infinite, null, or positive and finite.

Keywords: value of information, convex analysis, payoffs-beliefs duality.
AMS classification: 46N10, 91B06.

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.

Most information rankings are either uniform among agents or restricted to certain classes of agents. Blackwell's comparison of experiments [8], 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. Papers [26, 31, 12] are examples 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 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. [9, 3]).

In any decision problem, to each decision and state of nature corresponds 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 [7]. 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 nonnegative. Thus, the value function fully captures the agent's preferences for information.

In this paper, we make use of *convex analysis* [33] to exploit a bilinear duality structure between payoffs and beliefs, that gives expected payoff [17]. 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 correspondence between convex analysis concepts and tools, on the one hand, and economic objects, on the other hand. The set of beliefs compatible with an optimal action is related to the *normal cone* of the set of available vector payoffs at this optimal action. The *subdifferential* 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 if and only if at least one of the optimal actions at the prior becomes

*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@cnrs.fr

suboptimal for some of the posteriors. We thus define the confidence set at a prior belief $\bar{p}$ as the set of posterior beliefs for which all optimal actions at $\bar{p}$ remain optimal. Our result says that information has positive value if and only if posterior beliefs fall outside of the confidence set with positive probability. This result generalizes insights from [23] and [30], who had 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. There are several optimal actions at the prior, and information that allows to break indifferences has highest value.

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 this intermediate case, information impacts actions in a continuous way. The optimal actions at the prior belief and at a posterior close to it are themselves close; so choosing one instead of the other has a mild, albeit positive, impact on the expected payoff.

In a finite decision problem — such as shopping behavior [28] or residential location [29] — 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 [11] or investment decisions [1].

The paper is organized as follows. Sect. 2 presents the model and introduces the duality between actions/payoffs and beliefs. The main results are presented in Sect. 3. Sect. 4 is devoted to an illustration of our results in an insurance example and Sect. 5 to applications to the question of marginal value of information. Sect. 6 concludes by discussing related literature. The Appendix contains background on convex analysis and the proofs.

2. Model, payoffs-beliefs duality and information. We consider the classical question of an agent who faces a decision problem under imperfect information on a state of nature. 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 $\Delta = \Delta(K) \subset \Sigma = \mathbb{R}^K$ of probability distributions over K. We identify Δ with the simplex of $\mathbb{R}^K$.

A *decision problem* is given by an arbitrary compact choice set D and by a continuous payoff function $g \colon D \times K \to \mathbb{R}$. Consistent with the framework of [8], we define the set of *actions* as the compact convex subspace of $\mathbb{R}^K$ given by the *closed convex hull*:

$$A = \overline{\text{co}}\{\big(g(d,k)\big)_{k\in K}, d \in D\} \subset \mathbb{R}^K \ . \tag{2.1}$$

The convexity of A is justified by allowing the agent to randomize over actions.

Duality between actions/payoffs and beliefs. The scalar product between a vector $v \in \mathbb{R}^K$ and a signed measure $s \in \mathbb{R}^K$ is $\langle s\,, v\rangle = \sum_{k\in K} s_k v_k$. This scalar product induces a duality between payoffs/actions and beliefs. Such a duality is at the core of a series of works in nonexpected utility theory, such as [21, 27, 14].

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* $v_A : \Delta \to \mathbb{R}$ by:

$$v_A(p) = \max_{a\in A} \langle p\,, a\rangle \ , \ \ \forall p \in \Delta \ . \tag{2.2}$$

The value function $v_A \colon \Delta \to \mathbb{R}$ is convex — as the supremum of the family of affine functions $\langle \cdot\,, a\rangle$ for $a \in A$ — and continuous — as its effective domain is the whole convex set Δ [22, p. 175].

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

$$A^\star(p) = \arg\max_{a'\in A} \langle p\,, a'\rangle = \{a \in A \mid \forall a' \in A\,,\ \langle p\,, a'\rangle \le \langle p\,, a\rangle\} \ . \tag{2.3}$$

Geometrically, the set $A^\star(p)$ is the *(exposed) face of A in the direction $p \in \Delta$* (see (A.3) in Appendix for a proper definition). The set $A^\star(p)$ is nonempty, closed and convex (as A is convex and compact).

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

$$\Delta_A^\star(a) = \{p \in \Delta \mid \forall a' \in A\ ,\ \ \langle p\,, a'\rangle \le \langle p\,, a\rangle\}\ . \tag{2.4}$$

Geometrically, the set $\Delta_A^\star(a)$ is the intersection with Δ of the *normal cone* $N_A(a)$ (see (A.6) for a proper definition).

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

Information structure. We follow [9, 8], 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 $\mathbf{q}$, defined over a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ and with values in Δ, describing the agent's posterior beliefs, and such that (where $\mathbb{E}$ denotes the expectation operator with respect to $\mathbb{P}$)

$$\mathbf{q} : (\Omega, \mathcal{F}, \mathbb{P}) \to \Delta\ ,\ \ \mathbb{E}\big[\mathbf{q}\big] = \bar{p}\ . \tag{2.5}$$

Given the action set A in (2.1) and the information structure $\mathbf{q}$ in (2.5), 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 belief is $\bar{p}$. It is given by:

$$\mathbf{VoI}_A(\mathbf{q}) = \mathbb{E}\big[v_A(\mathbf{q})\big] - v_A(\bar{p})\ . \tag{2.6}$$

The following example illustrates relations between the set A of actions and the value function v_A.

Example 1. *Consider two states of nature, $K = \{1, 2\}$, decisions $D = \{d_1, d_2, d_3, d_4\}$, and payoffs given by Table 1. In this case, A is the convex hull of the four points $(3, 0)$, $(2, 2)$, $(0, 5/2)$ and $(0, 0)$. The value*

	$k = 1$	$k = 2$
d_1	3	0
d_2	2	2
d_3	0	5/2
d_4	0	0

Table 1
Table of payoffs

function v_A, expressed as a function of the probability p of state 2, is the maximum of the following three affine functions: $3(1-p)$, 2, and $5p/2$. Action $(3, 0)$ is optimal for $p \le 1/3$, $(2, 2)$ is optimal for $p \in [1/3, 4/5]$, and $(0, 5/2)$ is optimal for $p \ge 4/5$. Both the set A and the function v_A are represented in Figure 1.

At $p = 4/5$, the optimal actions are $(2, 2)$, $(0, 5/2)$, and their convex combinations. At this point, the mapping v_A is not differentiable. However, its subdifferential — which can be visualized as the set of straight lines that are below v_A and tangent to it at $p = 4/5$ — is still well defined and corresponds precisely to the optimal actions $A^\star(4/5)$, i.e. the convex hull of $\{(2, 2), (0, 5/2)\}$.

The set $\Delta_A^\star(3, 0)$ of beliefs revealed by action $(3, 0)$ consists of the range $p \in [0, 1/3]$, and it can be seen on the right side of Figure 1 that, for this range of probabilities, the action $(3, 0)$ is optimal and that v_A is linear and equal to $3(1 - p)$.

3. On the value of information. In this section, we relate the geometry of the set A of actions in (2.1) with the behavior of the agent around the prior belief $\bar{p}$, with differentiability properties of the value function v_A in (2.2) at the prior belief $\bar{p}$, and with the value of information $\mathbf{VoI}_A$ in (2.6). This approach allows us to derive bounds on the value of information that depend on how information influences actions.

First, in Subsect. 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

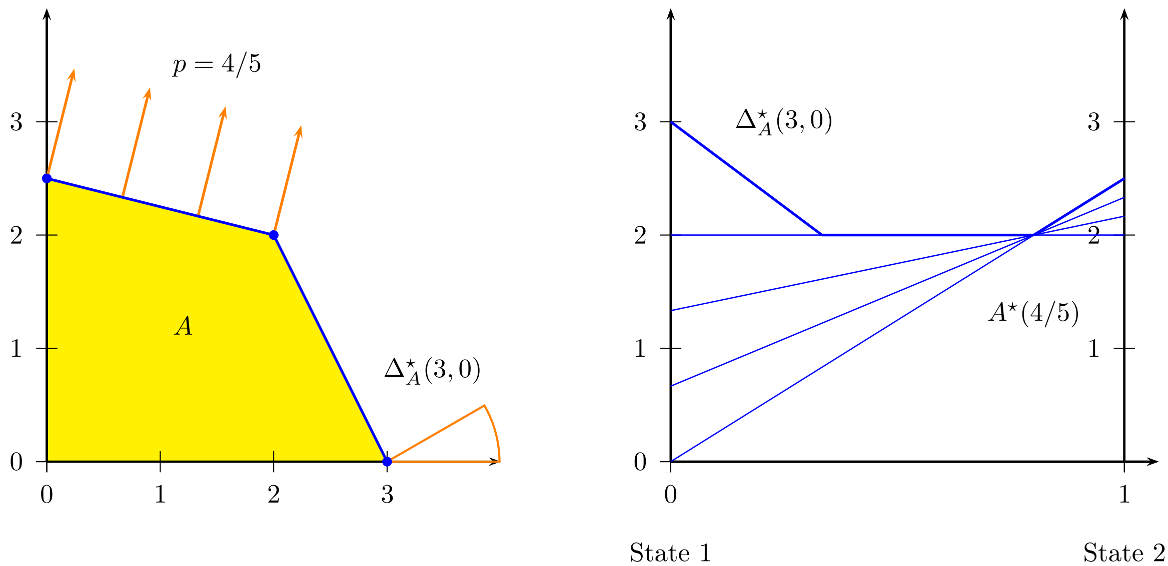

FIG. 1. *The set A of actions on the left, and the value function v_A on the right. Each of the four arrows on the left represents an action a such that $p = 4/5$ belongs to the set $\Delta^\star_A(a)$ of beliefs revealed by action a. On the right side, these four actions (each attached to an arrow) can be seen as four elements of the subdifferential of the value function v_A at $p = 4/5$. The set $\Delta^\star_A(3, 0) = [0, 1/3]$ can be visualized both as the normal cone at $(3, 0)$ on the left side, and as the range of values of probabilities p for which $(3, 0)$ is optimal on the right.*

optimal. We show that information is valuable if and only if, with positive probability, it can lead to a posterior outside this set. Therefore, information is valuable whenever it allows to eliminate some actions from the set of optimal ones.

Second, in Subsect. 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 Subsect. 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 smooth 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 small changes of actions.

3.1. Valuable information. Our first task is to formalize the idea that useful information is information that affects optimal choices (quoting [23], “Information is of value only if it can affect action”). Since there are potentially several optimal actions at a prior belief $\bar{p}$ 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* $\Delta^{\mathrm{c}}_A(\bar{p})$ of prior belief $\bar{p}$ iff all optimal actions at $\bar{p}$ (those in $A^\star(\bar{p})$) are also optimal at p. In other words, we define the *confidence set of prior belief* $\bar{p}$ by:

$$\Delta^{\mathrm{c}}_A(\bar{p}) = \bigcap_{a \in A^\star(\bar{p})} \Delta^\star_A(a) \ . \tag{3.1}$$

Another way to look at this notion is to consider an observer who sees choices by the decision maker:

$p \in \Delta^{c}_{A}(\bar{p})$ when none of the actions chosen by the agent at prior belief $\bar{p}$ 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 3.1 (Valuable information). *For every information structure* $\mathbf{q}$ *as in* (2.5)*, we have:*

$$\mathbf{VoI}_A(\mathbf{q}) = 0 \iff \exists a^\star \in A^\star(\bar{p}) \ , \ \ a^\star \in A^\star(\mathbf{q}) \ , \ \ \mathbb{P} - a.s. \tag{3.2a}$$

$$\iff \mathbf{q} \in \Delta^{c}_{A}(\bar{p}) \ , \ \ \mathbb{P} - a.s. \tag{3.2b}$$

In Example 1, the confidence set at $\bar{p} = 1/2$ is the closed interval $[1/3, 4/5]$ (the flat portion of the function to the right of Figure 1). Information is valuable whenever, with some positive probability, the posterior does not belong to this set. When the posterior falls in this set with probability one, the value function averaged at the prior precisely equals the value at prior belief $\bar{p}$, hence information has no value.

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 $\Delta^{c}_{A}(\bar{p})$. For $\varepsilon > 0$, let

$$\Delta^{c}_{A,\varepsilon}(\bar{p}) = \{q \in \Delta \mid d\big(q, \Delta^{c}_{A}(\bar{p})\big) < \varepsilon\} \ \text{ where } \ d\big(q, \Delta^{c}_{A}(\bar{p})\big) = \inf_{p \in \Delta^{c}_{A}(\bar{p})} \|p - q\|. \tag{3.3}$$

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

Theorem 3.2 (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* $\mathbf{q}$ *as in* (2.5)*:*

$$C_A \mathbb{E}\left[d\big(\mathbf{q}, \Delta^{c}_{A}(\bar{p})\big)\right] \geq \mathbf{VoI}_A(\mathbf{q}) \geq c_{\bar{p},A,\varepsilon} \mathbb{P}\{\mathbf{q} \notin \Delta^{c}_{A,\varepsilon}(\bar{p})\} \ . \tag{3.4}$$

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 is at least distance ε from the confidence set, 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 $\Delta^{c}_{A}(\bar{p})$ in (3.1), which can be computed locally at prior belief $\bar{p}$. On the other hand, they apply to all information structures. The caveat is that the multiplicative constants C_A and $c_{\bar{p},A,\varepsilon}$ in (3.4) depend on global, and not just local, properties of the action 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 (whose proof we do not give, as it is well-known in convex analysis [22, p. 251]), the value function then exhibits a *kink* at prior belief $\bar{p}$.

Proposition 3.3. *The two following conditions are equivalent:*
- *the set* $A^\star(\bar{p})$ *of optimal actions at the prior belief* $\bar{p}$ *in* (2.3) *contains more than one element;*
- *the value function* v_A *in* (2.2) *is nondifferentiable (in the standard sense) at the prior belief* $\bar{p}$*.*

Cases of indifference are typical of situations with a finite number of action choices. Coming back to Example 1, the agent is undecided for $\bar{p} = 1/2$ and $\bar{p} = 3/4$: at these priors, the agent has several optimal choices, and the value function is nondifferentiable. At all other priors, the optimal choice is unique, and the value function is differentiable.

At prior beliefs $\bar{p}$ satisfying the conditions of Proposition 3.3, the convexity gap of the value function v_A is maximal in the directions in which it is nondifferentiable. This allows us to derive a second bound on the value of information. For this purpose, we call *indifference kernel* $\Sigma^{\mathrm{i}}_A(\bar{p})$ at prior belief $\bar{p}$ the vector space of signed measures that are orthogonal to all differences of optimal actions $A^\star(\bar{p})$ at $\bar{p}$, that is,

$$\Sigma^{\mathrm{i}}_A(\bar{p}) = [A^\star(\bar{p}) - A^\star(\bar{p})]^\perp \; . \tag{3.5}$$

Beliefs in the indifference kernel $\Sigma^{\mathrm{i}}_A(\bar{p})$ do not break any of the ties in $A^\star(\bar{p})$, since $p \in \Sigma^{\mathrm{i}}_A(\bar{p}) \iff \langle p \, , a \rangle = \langle p \, , a' \rangle \; , \; \forall (a, a') \in A^\star(\bar{p})^2$. We note the inclusion $\Delta^{\mathrm{c}}_A(\bar{p}) \subset \Sigma^{\mathrm{i}}_A(\bar{p}) \cap \Delta$ as every element in the confidence set is necessarily in the indifference kernel and in the simplex of probability measures.

Recall that a *seminorm* on the signed measures Σ on K, identified with $\mathbb{R}^K$, is a mapping $\|\cdot\| : \mathbb{R}^K \to \mathbb{R}_+$ 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 seminorm $\|\cdot\|$ — is not necessarily reduced to the null vector.

Theorem 3.4 (Bounds on the value of information for the undecided agent). *There exists a positive constant C_A and a seminorm $\|\cdot\|_{\Sigma^i_A(\bar{p})}$ with kernel $\Sigma^i_A(\bar{p})$, the indifference kernel in (3.5), such that, for every information structure $\mathbf{q}$ as in (2.5):*

$$C_A \mathbb{E} \, \|\mathbf{q} - \bar{p}\| \geq \mathbf{VoI}_A(\mathbf{q}) \geq \mathbf{VoI}_{A^\star(\bar{p})}(\mathbf{q}) \geq \mathbb{E} \, \|\mathbf{q} - \bar{p}\|_{\Sigma^i_A(\bar{p})} \; . \tag{3.6}$$

For $\bar{p} = 1/2$ or $\bar{p} = 3/4$ in Example 1, Theorem 3.4 shows that the value of information for these priors is bounded above and below by a constant times the norm-1 between the prior and the posterior. Since any small amount of information allows to break the indifference between the optimal actions at these priors, information is very valuable.

The lower bound in Theorem 3.4 shows that a lower bound of the value of information is the expectation of a seminorm of the distance between the prior belief and the posterior belief. To understand the role of the kernel $\Sigma^{\mathrm{i}}_A(\bar{p})$ of this seminorm, let us first consider the set of beliefs in this set. A posterior q is in $\Sigma^{\mathrm{i}}_A(\bar{p}) = [A^\star(\bar{p}) - A^\star(\bar{p})]^\perp$ if and only if, for any two optimal actions $a, a' \in A^\star(\bar{p})$, $\langle q \, , a \rangle = \langle q \, , a' \rangle$. In words, posteriors that do not break any of the ties in $A^\star(\bar{p})$ might not be valuable to the agent. On the other hand, Theorem 3.4 tells us that all other directions — i.e., those that allow at least one of the ties in $A^\star(\bar{p})$ 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 inner inequality states that the value of information with decision set A is at least as large as with action set $A^\star(\bar{p})$.

Note that the bounds on Theorem 3.4 rely on the indifference kernel $\Sigma^{\mathrm{i}}_A(\bar{p})$ in (3.5), which can be computed directly from the set $A^\star(\bar{p})$ by (3.5). The multiplicative constant C_A in (3.6), however, depends on more global properties of the action 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 smoothly on the belief. More precisely, we assume that, around the prior, optimal actions smoothly depend in a 1-1 way on the belief. This assumption is met when, for instance, the decision problem faced by the agent is a scoring rule [11], or an investment problem [1, 12].

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 following proposition offers three alternative characterizations of these situations, based 1) on the local behavior of the agent's optimal 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. For background on geometric convex analysis, the reader can consult §A.2 in the Appendix.

Proposition 3.5. *Suppose that the action set A in (2.1) has boundary ∂A which is a C^2 submanifold of $\mathbb{R}^K$ of dimension $|K| - 1$. The three following conditions are equivalent:*

1. *The set-valued mapping of optimal actions at the prior belief $\bar{p}$ in (2.3)*

$$A^\star : \Delta \rightrightarrows \partial A \; , \;\; p \mapsto A^\star(p) \tag{3.7}$$

is a local diffeomorphism[1] *at the prior belief* $\bar{p}$*;*

2. *The set* $A^\star(\bar{p})$ *of optimal actions at the prior belief* $\bar{p}$ *in* (2.3) *is reduced to a singleton at which the curvature of the action set* A *is positive;*
3. *The value function* v_A *in* (2.2) *is twice differentiable at the prior belief* $\bar{p}$*, with positive definite Hessian at* $\bar{p}$*.*

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

THEOREM 3.6 (Bounds on the VoI for the flexible agent). *If the agent is flexible at prior belief* $\bar{p}$*, then there exist positive constants* $C_{\bar{p},A}$ *and* $c_{\bar{p},A}$ *such that, for every information structure* $\mathbf{q}$ *as in* (2.5)*:*

$$C_{\bar{p},A}\mathbb{E}\,||\mathbf{q}-\bar{p}||^2 \geq \mathbf{VoI}_A(\mathbf{q}) \geq c_{\bar{p},A}\mathbb{E}\,||\mathbf{q}-\bar{p}||^2 \,. \tag{3.8}$$

Theorem 3.6 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 $C_{\bar{p},A}$ and $c_{\bar{p},A}$ in (3.8) themselves depend on the global behavior of the value function, and hence cannot be inferred from local properties only.

4. An insurance example. In this example, we study an insurance problem and illustrate how the results of Sect. 3 apply. The insuree chooses whether to insure, or not, and at which indemnity level to insure if she does. The uncertainty is about the level of risk she incurs, and she may receive some partial information about it.

EXAMPLE 2. *The model is drawn from the classical insurance framework (see [6, 18]).*

An insuree faces the decision of partially or fully insuring a good of value ϖ *against the possibility of its total loss. Pricing is assumed to be linear, so that, for an indemnity* I*, the insurance company charges*

$$P(I) = \alpha I + f \text{ where } \alpha \in]0,1[\,, \quad f > 0\,. \tag{4.1}$$

In exchange for the premium $P(I)$*, the insuree gets compensation of an amount* I *from the insurance company in case of a loss. For the range of wealth* w *considered, the insuree's utility function* u *is considered to have constant absolute risk aversion* R*, that is,*

$$u(w) = 1 - e^{-Rw}\,. \tag{4.2}$$

By (2.1)*, the set of* actions *is the closed convex hull*

$$A = \overline{co}\Big\{\Big(u(\varpi), u(0)\Big)\,,\ \Big(u\big(-P(I)+\varpi\big), u\big(-P(I)+I\big)\Big)\Big\} \tag{4.3}$$

where, by convention, the first coordinate corresponds to no loss and the second corresponds to the loss.

The insuree's subjective perception that a loss may arise is $p \in]0,1[$*, probability of loss. The insuree chooses either not to insure, and obtains expected utility*

$$U_0(p) = (1-p)u(\varpi) + pu(0) = (1-p)\big(1-e^{-R\varpi}\big)\,, \tag{4.4a}$$

or to insure for an indemnity $I > 0$ *that maximizes the expected utility*

$$U(p,I) = (1-p)u\big(-P(I)+\varpi\big) + pu\big(-P(I)+I\big) = 1 - pe^{-R\left(-P(I)+I\right)} - (1-p)e^{-R\left(-P(I)+\varpi\right)}\,. \tag{4.4b}$$

The question now becomes whether no insurance or a positive level of indemnity is chosen.

[1] In particular, the set $A^\star(p)$ is a singleton for all $p \in \Delta$, in which case we identify a singleton set with its single element.

PROPOSITION 4.1. *There exists a threshold belief* $p^* \in]0,1[$ *and a smooth function* $\hat{I} : [p^*, 1] \to]0, +\infty[$ *such that*

1. *for* $p < p^*$, *it is optimal not to insure,*
2. *for* $p = p^*$, *the insuree is indifferent between no insurance and insurance at the positive indemnity level* $\hat{I}(p^*)$,
3. *for* $p > p^*$, *it is optimal to insure at the positive indemnity level* $\hat{I}(p)$.

Proof. It is easy to see that the function $I \in \mathbb{R} \mapsto U(p, I)$ in (4.4b) is strictly concave with a unique maximum, characterized by $\partial U/\partial I = 0$, and achieved at

$$\hat{I}(p) = \varpi - \frac{1}{R} \ln\left(\frac{1-p}{p} \frac{\alpha}{1-\alpha}\right) , \ \forall p \in]0, 1[\, . \tag{4.5}$$

We denote by $\hat{p}$ the unique $p \in]0, 1[$ such that $\hat{I}(p) > 0 \iff p > \hat{p}$. To determine whether no insurance or a nonnegative level of indemnity is chosen, we introduce the difference of expected utilities

$$\delta(p) = \max_{I \geq 0} U(p, I) - U_0(p) = \begin{cases} U(p, 0) - U_0(p) & \text{if } p \leq \hat{p} \, , \\ U\big(p, \hat{I}(p)\big) - U_0(p) & \text{if } p \geq \hat{p} \, . \end{cases} \tag{4.6}$$

We study the behavior of the function δ when p is small and when p is close to one. After computation, we find that, for all $p \in [0, 1]$, $U(p, 0) - U_0(p) = -\big(e^{Rf} - 1\big)\big(p + (1 - p)e^{-R\varpi}\big) < 0$. Therefore, $\delta(p) < 0$ for all $p \leq \hat{p}$. On the other hand, when p goes to 1, $\delta(p)$ goes to 1 because $U_0(p) \to 0$ and $U\big(p, \hat{I}(p)\big) = (1-p)\Big(1 - e^{-R\big(-P(\hat{I}(p))+\varpi\big)}\Big) + p\Big(1 - e^{-R\big(-P(\hat{I}(p))+\hat{I}(p)\big)}\Big) = 1 - (1-p)\big(\frac{1-p}{p}\frac{\alpha}{1-\alpha}\big)^{\alpha} e^{R(1-\alpha)\varpi} - p\big(\frac{1-p}{p}\frac{\alpha}{1-\alpha}\big)^{1-\alpha} e^{-R(1-\alpha)\varpi} \to 1$ (as $\alpha \in]0, 1[$). As a consequence, we can define $p^* = \inf\{p \in [0, 1] \mid \delta(p) > 0\}$, which belongs to $[\hat{p}, 1[$. Indeed, since $\delta(p) < 0$ for $p \leq \hat{p}$, we deduce that $p^* \geq \hat{p}$; and $p^* < 1$ because $\delta(p) \to 1$ when $p \to 1$. We now check that p^* and $\hat{I}$ in (4.5) satisfy the three assertions of the Proposition.

By definition of p^* and of the function δ, for $p < p^*$, it is optimal not to insure.

As the function δ is continuous, we have $\delta(p^*) = 0$ and the insuree is indifferent between no insurance and insurance at the positive indemnity level $\hat{I}(p^*)$.

To finish, we will now show that $\delta(p) > 0$ when $p > p^*$, leading to the conclusion that it is optimal to insure at the positive indemnity level $\hat{I}(p)$. Indeed, for $p > p^*$, we have

$$\begin{aligned}
\delta(p) &= \delta(p) - \delta(p^*) \quad \text{as } \delta(p^*) = 0 \\
&= U\big(p, \hat{I}(p)\big) - U\big(p, \hat{I}(p^*)\big) + U\big(p, \hat{I}(p^*)\big) - U_0(p) - \big[U\big(p^*, \hat{I}(p^*)\big) - U_0(p^*)\big] \text{ by (4.6)} \\
&> U\big(p, \hat{I}(p^*)\big) - U_0(p) - U\big(p^*, \hat{I}(p^*)\big) + U_0(p^*) \qquad \text{as } U\big(p, \hat{I}(p)\big) - U\big(p, \hat{I}(p^*)\big) > 0 \\
&\quad \text{by definition of the maximizer } \hat{I}(p) \text{ and since } \hat{I}(p) > \hat{I}(p^*) \geq 0 \text{ as } p > p^* \geq \hat{p} \\
&= (1-p)\big[u\big(-P(\hat{I}(p^*)) + \varpi\big) - u(\varpi)\big] + p\big[u\big(-P(\hat{I}(p^*)) + \hat{I}(p^*)\big) - u(0)\big] \\
&\quad - (1-p^*)\big[u\big(-P(\hat{I}(p^*)) + \varpi\big) - u(\varpi)\big] - p^*\big[u\big(-P(\hat{I}(p^*)) + \hat{I}(p^*)\big) - u(0)\big] \text{ by (4.4)} \\
&= (p - p^*)\Big[\big[u\big(-P(\hat{I}(p^*)) + \hat{I}(p^*)\big) - u(0)\big] + \big[u(\varpi) - u\big(-P(\hat{I}(p^*)) + \varpi\big)\big]\Big] \geq 0
\end{aligned}$$

since both terms between inner brackets are increments of the increasing function u, where $-P(\hat{I}(p^*)) + \hat{I}(p^*) \geq 0$ (to be seen below) and $P(\hat{I}(p^*)) \geq 0$ (because $\hat{I}(p^*) \geq 0$). If we had $-P(\hat{I}(p^*)) + \hat{I}(p^*) < 0$, we would arrive at the contradiction that $0 = \delta(p^*) = (1-p^*)\big[u\big(-P(\hat{I}(p^*)) + \varpi\big) - u(\varpi)\big] + p^*\big[u\big(-P(\hat{I}(p^*)) + \hat{I}(p^*)\big) - u(0)\big] < 0$ since both terms between brackets are (negative) increments of the increasing function u. ∎

Now, we assume that the insuree has access to a small piece of information concerning her probability of loss. Once informed, she discovers that the probability q of a loss is either $p - \varepsilon$ or $p + \varepsilon$, where both possibilities are equally likely and $\varepsilon > 0$ is a small positive number. Let $v(q)$ be the utility of the insuree with beliefs q, once the optimal policy is chosen:

$$v(q) = \max\left\{U_0(q), \max_{I \geq 0} U(q, I)\right\} \, . \tag{4.7}$$

As v is the value function in (2.2), the value of information in the decision problem is defined as the expected utility with the information minus the expected utility absent the information, as in (2.6):

$$\mathbf{VoI}(\varepsilon) = \frac{1}{2}v(p+\varepsilon) + \frac{1}{2}v(p-\varepsilon) - v(p) \ . \tag{4.8}$$

Note that $\mathbf{VoI}(\varepsilon)$ measures the value of information in terms of utility; the equivalent measure in monetary terms would be $-\frac{1}{R}\ln(1-\mathbf{VoI}(\varepsilon))$. The following proposition characterizes the value of a small amount of information, in terms of the agent's optimal insurance behavior.

Proposition 4.2. *Depending on the probability of loss p, the value of information for small ε behaves as follows:*

1. *In the confident case, for $p < p^*$, $\mathbf{VoI}(\varepsilon) = 0$ for small ε,*
2. *In the undecided case, for $p = p^*$, $\mathbf{VoI}(\varepsilon) \sim C^*\varepsilon$ for a constant $C^* > 0$,*
3. *In the flexible case, for $p > p^*$, $\mathbf{VoI}(\varepsilon) \sim C(p)\varepsilon^2$ for a constant $C(p) > 0$.*

Proof. The confident and undecided cases are immediate consequences of Theorems 3.2 and 3.4, together with Proposition 4.1. In the flexible case, the optimal indemnity level is given by $\hat{I}(p) > 0$, and the function $\hat{I} :]p^*, 1] \to]0, +\infty[$ in (4.5) is differentiable with $\frac{d\hat{I}(p)}{dp} \neq 0$. The set of optimal actions $A^\star(p)$ in (2.3) is reduced to the single point $A^\star(p) = \Big(1 - e^{-R\big(-P\big(\hat{I}(p)\big)+\varpi\big)}, 1 - e^{-R\big(-P\big(\hat{I}(p)\big)+\hat{I}(p)\big)}\Big)$. As the curve $p \in]p^*, 1] \mapsto A^\star(p)$ has a derivative that never vanishes, we deduce that it is a local diffeomorphism (onto its image in ∂A) at p, and Theorem 3.6 applies. ∎

The results of Proposition 4.1 are intuitive. First, a small piece of information is valueless if the agent is not buying insurance. For such agents, a small bit of information does not affect behavior, as even bad news is not enough to trigger insurance purchase. For an undecided agent who is indifferent between no insurance and insurance at a positive indemnity level $I(p^*)$, a small piece of information is enough to break the indifference and significantly influences her behavior; this is the situation in which information is the most valuable. Finally, for an agent who takes a positive level of indemnity, information may affect the level of indemnity chosen. But, because the change of indemnity level is itself of order ε, and the indemnity level $I(p^*)$ is ε-optimal at the posterior, the value of information is a second order in ε.

Figure 2 represents the set A of actions (4.3) to the left, and the corresponding value function $v = v_A$ in (4.7) to the right. In the representation of A, the horizontal axis corresponds to the payoff without loss, and the vertical axis to the payoff in case of a loss. The circled dot to the right corresponds to the choice of no insurance; it maximizes payoff in case of no loss. The thick curve represents the set of payoffs that are achieved by different coverage levels. Finally, A is the convex hull of this set of points; it appears under the dashed contour. As seen on the value function graph, for low values of the probability p of loss, the value function is linear as the insuree chooses not to purchase insurance. At p^* (which is approximately 0.334), the value function exhibits a kink, and the agent is indifferent between no insurance and a positive indemnity level. Finally, for larger values of p, the value function v is twice continuously differentiable with a positive second derivative, and the optimal insurance level is a smooth and positive function of the insuree's belief.

5. The marginal value of information. The question of the marginal value of information is studied in [32]. They provide joint conditions on a parameterized 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 in [15] and [16], where are provided joint conditions on parameterized 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, obtained in Sect. 3, apply to the marginal value of information. In Subsect. 5.1, we provide separate conditions on the decision problem and on the family of parameterized information structures that result in a null value of information. We then examine, in Subsect. 5.2, several parameterized 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.

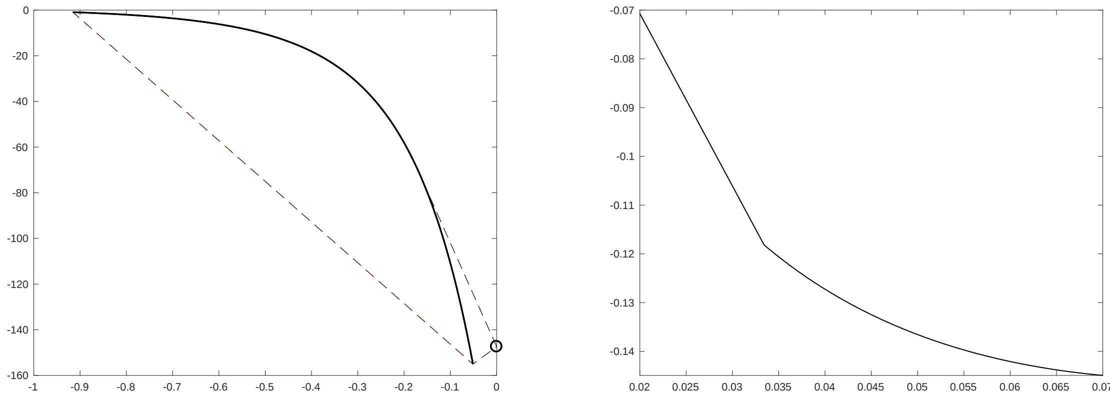

FIG. 2. *The action set A on the left and the corresponding value function $v = v_A$ in (4.7) for the insurance example on the right. Parameter values are $\alpha = 0.08$, $f = 10$, $\varpi = 1000$, $R = 10$.*

5.1. Model and first result. Let $(\mathbf{q}^\theta)_{\theta>0}$ be a family of information structures as in (2.5). As in [32], we are interested in the so-called *marginal value of information*:

$$V^+ = \limsup_{\theta\to 0} \frac{1}{\theta} \mathbf{VoI}_A(\mathbf{q}^\theta) \ . \tag{5.1}$$

The following proposition is a straightforward consequence of Theorems 3.2 and 3.6.

PROPOSITION 5.1. *Assume that*
- *either* $\mathbb{E}\left[d\big(\mathbf{q}^\theta, \Delta^c_A(\bar{p})\big)\right] = o(\theta)$,
- *or the decision maker is* flexible *at prior belief* $\bar{p}$ *and* $\mathbb{E}\,\|\mathbf{q}^\theta - \bar{p}\|^2 = o(\theta)$.

Then the marginal value of information $V^+ = 0$.

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

We now discuss how our approach in Proposition 5.1 compares with the literature. In [32], one finds joint conditions on the parameterized information structure $(\mathbf{q}^\theta)_{\theta>0}$ and the decision problem at hand, leading to $V^+ = 0$. The second case in Proposition 5.1, 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 [32] does not require that $A^\star(\mathbf{q}^\theta)$ be a singleton, for all θ.

The authors of [15] make a step towards disentangling conditions on the parameterized information structure $(\mathbf{q}^\theta)_{\theta>0}$ from conditions on the decision problem that lead to a null marginal value of information. However, like [32], they make an assumption on how the optimal action varies with information, which makes the comparison with Proposition 5.1 delicate. In addition, [15] 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.

The authors of [16] provide separate conditions on the parameterized information structure $(\mathbf{q}^\theta)_{\theta>0}$ and the decision problem (represented by the action set A) that lead to $V^+ = 0$. Their condition "IIDV=0" is that $\limsup_{\theta\to 0} \frac{1}{\theta}\mathbb{E}\,\|\mathbf{q}^\theta - \bar{p}\| = 0$, or, equivalently, $\mathbb{E}\,\|\mathbf{q}^\theta - \bar{p}\| = o(\theta)$, which implies the first item of Proposition 5.1. Thus, this latter proposition implies the main result of [16].

5.2. Examples. Here, we study the marginal value of information for several typical parameterized information structures. In the first example, 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 example, 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 parameterization of information is the length of the interval of time during which observation takes place. In the 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, $K = \{0,1\}$, and (by a slight abuse of notation) the prior belief on the state being 1 is denoted $\bar{p} \in]0,1[$. We follow the conditions in Sect. 3 under which we established bounds on the value of information, and label as: "confident" the case in which $\bar{p}$ lies in the interior of the confidence set $\Delta_A^c(\bar{p})$ (in this case, $\Delta_A^c(\bar{p})$ is a closed nonempty interval $\left[p_l, p_h\right]$ by Proposition A.3, and the value function is linear on this range); "undecided" the case in which the decision problem faced by the decision maker is such that there is indifference between two actions at prior belief $\bar{p}$; "flexible" the case in which the optimal action is a smooth function of the belief in a neighborhood of prior belief $\bar{p}$.

Our aim is to develop estimates of the marginal value of information V^+ in (5.1). 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 3 (Brownian motion). *Frameworks in which agents observe a Brownian motion with known volatility and unknown drift include [5, 24, 10], as well as reputation models like [19].*

Assume the agent observes the realization of a Brownian motion with variance 1 and drift $k \in \{0,1\}$, namely $d\mathbf{Z}_t = kdt + d\mathbf{B}_t$, for a small interval of time $\theta > 0$. If we let $\mathbf{q}^t$ be the posterior belief at time t, it is well-known[2] that $\mathbf{q}^t$ follows a diffusion 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 Browian process. Thus, for small values of θ, we have the estimates

$$\mathbb{E}\,\|\mathbf{q}^\theta - p\| \sim \sqrt{\theta}\ ,\ \ \mathbb{E}\,\|\mathbf{q}^\theta - p\|^2 \sim \theta\ .$$

It follows from Theorems 3.2-3.6 that the marginal value of information is characterized, depending on the decision problem, as:

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

Example 4 (Poisson learning). *An important class of models of strategic experimentation (see [25]) are those in which the agent's observations are driven by a Poisson process of unknown intensity. Assume the agent observes, during a small interval of time $\theta > 0$, a Poisson process with intensity ρ_k, $k \in \{0,1\}$, where $\rho_1 > \rho_0 > 0$. 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 from below, as $\theta \to 0$, to*

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

and happens with probability of order $\sim \theta$. In the absence of success, the posterior belief converges to the prior belief $\bar{p}$ as $\theta \to 0$. As we have seen that the confidence set $\Delta_A^c(\bar{p})$ is a closed interval $\left[p_l, p_h\right]$, we note that $\mathbb{E}\left[d\big(\mathbf{q}^\theta, \Delta_A^c(\bar{p})\big)\right] \sim \theta$ if $q^+ > p_h$, and $\mathbb{E}\left[d\big(\mathbf{q}^\theta, \Delta_A^c(\bar{p})\big)\right] = o(\theta)$ otherwise. This implies:

1. *In the confident case,*
   (a) $V^+ \simeq 1$ *if* $q^+ > p_h$,
   (b) $V^+ \simeq 0$ *if* $q^+ \leq p_h$.

We also have the estimates

$$\mathbb{E}\,\|\mathbf{q}^\theta - p\| \sim \theta\ ,\ \ \mathbb{E}\,\|\mathbf{q}^\theta - p\|^2 \sim \theta\ ,$$

which imply the following estimates on the marginal value of information:

2. *In the undecided case, $V^+ \simeq 1$,*
3. *In the flexible case, $V^+ \simeq 1$.*

Example 5 (Equally likely signals). *Here, we consider binary and equally likely signals, which lead to a "split" of beliefs around the prior belief $\bar{p}$. Depending on the precision of these signals as a function of θ,*

[2] See for instance Lemma 1 in [10] or Lemma 2 in [19].

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

$$\mathbb{E}\,\|\mathbf{q}^\theta - p\| = \theta^\alpha \ , \ \ \mathbb{E}\,\|\mathbf{q}^\theta - p\|^2 = \theta^{2\alpha} \ ,$$

and we observe that $\mathbb{E}\left[d(\mathbf{q}^\theta, \Delta_A^c(\bar{p})\right] = 0$ for θ small enough. Here again, the marginal value of information is deduced from Theorems 3.2–3.6:

1. *In the confident case, $V^+ = 0$,*
2. *In the undecided case,*
   (a) *$V^+ = \infty$ if $\alpha < 1$,*
   (b) *$V^+ \simeq 1$ if $\alpha = 1$,*
   (c) *$V^+ = 0$ if $\alpha > 1$,*
3. *In the flexible case,*
   (a) *$V^+ = \infty$ if $\alpha < 1/2$,*
   (b) *$V^+ \simeq 1$ if $\alpha = 1/2$,*
   (c) *$V^+ = 0$ if $\alpha > 1/2$.*

Table 2 summarizes the marginal value of information in all of our examples.

Marginal value of information V^+	confident	undecided	flexible
Brownian	0	∞	1
Poisson learning	0 or 1	1	1
Equally likely signals, $\alpha < 1/2$	0	∞	∞
Equally likely signals, $\alpha = 1/2$	0	∞	1
Equally likely signals, $1/2 < \alpha < 1$	0	∞	0
Equally likely signals, $\alpha = 1$	0	1	0
Equally likely signals, $\alpha > 1$	0	0	0

TABLE 2

Marginal value of information in the different examples. The value 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 lower 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 ($\alpha > 1/2$).

6. Related literature. The value of information in decision problems is a well-studied question in economics and in statistics. The central work in this area is [8], which defines a source of information α as *more informative* than another, β, whenever all agents, independently of their preferences and decision problems faced, weakly prefer α to β. Blackwell [8] characterizes precisely this relationship in the following terms: α is more informative than β if and only if information from β can be obtained as a garbling of the information from α.

The requirement that all agents agree on their preferences between two statistical experiments is a strong one. It implies that this ranking is incomplete, as many such pairs of experiments cannot be ranked according to this ordering. Some authors have considered subclasses of decision problems in order to obtain rankings that are more complete than Blackwell's. For instance, [26], [31] and [2] restrict attention to families of decision problems that generate monotone decision rules. Focusing on investment decision problems, [12] obtains and characterizes a complete ranking of information sources based on a uniform criterion; [13] uses a duality approach to characterize the value of an information purchase that consists of an information structure with a price attached to it.

The present work departs from this literature in the sense that we focus on the value of information for a given agent, instead of trying to measure the value of information independently of the agent. Papers [20] and [4] characterize the possible preferences for information that any agent can have, letting the decision problem vary and the agent's preferences vary.

The question of marginal value of information is studied in [32, 15, 16]. They consider parameterized information structures, and derive general conditions on the couple consisting of the information structures and the decision problem under which the marginal value of information close to no information is zero. Our work contributes to this question by allowing us to derive estimates on the value of information based on separate conditions on the decision problem and on the information structure. This is the approach we have taken in Sect. 5. Our contribution considerably opens the spectrum of possibilities for the marginal value of information, by giving conditions under which it can be infinite, null, or positive and finite.

Acknowledgements. Olivier Gossner acknowledges support from the French National Research Agency (ANR), "Investissements d'Avenir" (ANR-11-IDEX-0003/LabEx Ecodec/ANR-11-LABX-0047), and thanks Rafael Veiel for his excellent research assistantship. The Authors thank the Editor and two Referees for their insightful comments, that have helped improve the manuscript.

REFERENCES

[1] K. J. Arrow, *The value of and demand for information*, in Decision and Organization, C. McGuire and R. Radner, eds., Amsterdam, 1971, North-Holland, pp. 131–139.
[2] S. Athey and J. Levin, *The value of information in monotone decision problems*, Research in Economics, (2017).
[3] R. Aumann and M. Maschler, *Repeated games with incomplete information: A survey of recent results*, in Reports to the U.S. Arms Control and Disarmament Agency, ST-116, 1967, pp. 287–403.
[4] Y. Azrieli and E. Lehrer, *The value of a stochastic information structure*, Games and Economic Behavior, 63 (2008), pp. 679–693.
[5] D. Bergemann and J. Välimäki, *Market diffusion with two-sided learning*, The RAND Journal of Economics, (1997), pp. 773–795.
[6] D. Bernoulli, *Specimen theoriae novae de mensura sortis*, in Commentarii Academiae Scientiarum Imperialis Petropolitanae, vol. 5, 1738, pp. 175–192.
[7] D. Blackwell, *Comparison of experiments*, in Proceedings of the Second Berkeley Symposium on Mathematical Statistics and Probability, University of California Press, 1951, pp. 93–102.
[8] D. Blackwell, *Equivalent comparison of experiments*, Annals of Mathematical Statistics, 24 (1953), pp. 265–272.
[9] H. F. Bohnenblust, L. S. Shapley, and S. Sherman, *Reconnaissance in game theory*, Tech. Report RM-208, The RAND corporation, 1949.
[10] P. Bolton and C. Harris, *Strategic experimentation*, Econometrica, 67 (1999), pp. 349–374.
[11] G. W. Brier, *Verification of forecasts expressed in terms of probability*, Monthly Weather Review, 78 (1950).
[12] A. Cabrales, O. Gossner, and R. Serrano, *Entropy and the value of information for investors*, American Economic Review, 103 (2013), pp. 360–377.
[13] A. Cabrales, O. Gossner, and R. Serrano, *A normalized value for information purchases*, Journal of Economic Theory, 170 (2017), pp. 266–288.
[14] S. Cerreia-Vioglio, F. Maccheroni, M. Marinacci, and L. Montrucchio, *Complete monotone quasiconcave duality*, Mathematics of Operations Research, 36 (2011), pp. 321–339.
[15] H. Chade and E. Shlee, *Another look at the Radner-Stiglitz nonconcavity in the value of information*, Journal of Economic Theory, 107 (2002), pp. 421–452.
[16] M. De Lara and L. Gilotte, *A tight sufficient condition for Radner–Stiglitz nonconcavity in the value of information*, Journal of Economic Theory, 137 (2007), pp. 696–708.
[17] D. Dentcheva and A. Ruszczyński, *Common mathematical foundations of expected utility and dual utility theories*, SIAM Journal on Optimization, 23 (2013), pp. 381–405.
[18] L. Eeckhoudt, C. Gollier, and H. Schlesinger, *Economic and Financial Decisions under Risk*, Princeton University Press, 2005.
[19] E. Faingold and Y. Sannikov, *Reputation in continuous-time games*, Econometrica, 79 (2011), pp. 773–876.
[20] I. Gilboa and E. Lehrer, *The value of information - an axiomatic approach*, Journal of Mathematical Economics, 20 (1991), pp. 443–459.
[21] I. Gilboa and D. Schmeidler, *Maxmin expected utility with non-unique prior*, Journal of Mathematical Economics, 18 (1989), pp. 141–153.
[22] J. B. Hiriart-Ururty and C. Lemaréchal, *Convex Analysis and Minimization Algorithms I*, Springer-Verlag, Berlin, 1993.

[23] J. Hirshleifer, *The private and social value of information and the reward to inventive activity*, American Economic Review, 61 (1971), pp. 561–574.
[24] G. Keller and S. Rady, *Optimal experimentation in a changing environment*, Review of Economic Studies, 66 (1999), pp. 475–507.
[25] G. Keller, S. Rady, and M. Cripps, *Strategic experimentation with exponential bandits*, Econometrica, 73 (2005), pp. 39–68.
[26] E. L. Lehmann, *Comparing location experiments*, The Annals of Statistics, 16 (1988), pp. 521–533.
[27] F. Maccheroni, M. Marinacci, and A. Rustichini, *Ambiguity aversion, robustness, and the variational representation of preferences*, Econometrica, 74 (2006), pp. 1447–1498.
[28] D. McFadden, *Conditional logit analysis of qualitative choice behavior*, in Frontiers in Econometrics, P. Zarembka, ed., Academic Press, 1973.
[29] D. McFadden, *Modelling the choice of residential location*, in Spatial Interaction Theory and Planning Models, A. Karlqvist, L. Lundqvist, F. Snickars, and J. Weibull, eds., North-Holland, 1978.
[30] L. J. Mirman, L. Samuelson, and A. Urbano, *Monopoly experimentation*, International Economic Review, (1993), pp. 549–563.
[31] N. Persico, *Information acquisition in auctions*, Econometrica, 68 (2000), pp. 135–148.
[32] R. Radner and J. Stiglitz, *A nonconcavity in the value of information*, in Bayesian Models of Economic Theory, M. Boyer and R. Kihlstrom, eds., Amsterdam, 1984, Elsevier, pp. 33–52.
[33] T. R. Rockafellar, *Convex Analysis*, Princeton University Press, Princeton, N.J., 1970.
[34] R. Schneider, *Convex bodies: the Brunn-Minkowski theory*, Cambridge University Press, second ed., 2014.

Appendix A. Appendix.

A.1. Revisiting the model of Sect. 2. We revisit the model in Sect. 2 with convex analysis tools to prepare the proofs in Sect. A.3. We recall that $A \subset \mathbb{R}^K$ in (2.1) is a nonempty, convex and compact subset of $\mathbb{R}^K$, called the *action set*, and that we identify the set Σ of signed measures on K with $\mathbb{R}^K$.

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

$$\sigma_A(s) = \sup_{a \in A} \langle s \,, a \rangle \;\;, \;\; \forall s \in \Sigma \;. \tag{A.1}$$

The value function $v_A : \Delta \to \mathbb{R}$ in (2.2) is the restriction of σ_A to probability distributions $\Delta = \Delta(K) \subset \Sigma$:

$$v_A(p) = \sigma_A(p) \;, \;\; \forall p \in \Delta \;. \tag{A.2}$$

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 effective domain is Σ, hence σ_A is continuous.

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

$$F_A(s) = \arg\max_{a' \in A} \langle s \,, a' \rangle = \{ a \in A \mid \forall a' \in A \;, \;\; \langle s \,, a' \rangle \leq \langle s \,, a \rangle \} \subset A \tag{A.3}$$

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: $F_A(s) \subset \partial A \;, \;\; \forall s \in \Sigma$. We will use the following property: for any nonempty convex set $C \subset \mathbb{R}^K$ and $y \in \mathbb{R}^K$ such that $F_C(y) \neq \emptyset$, we have

$$\sigma_C(y') - \sigma_C(y) \geq \sigma_{F_C(y)}(y' - y) \geq \langle y' - y \,, x' \rangle \;\;, \;\; \forall y' \in \mathbb{R}^K \;, \;\; \forall x' \in C \;. \tag{A.4}$$

The set $A^\star(p)$ of optimal actions under belief p in (2.3) coincides with the (exposed) face $F_A(p)$ of A in the direction p in (A.3):

$$A^\star(p) = F_A(p) \;, \;\; \forall p \in \Delta \;. \tag{A.5}$$

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

$$N_A(a) = \{ s \in \Sigma \mid \forall a' \in A \;, \;\; \langle s \,, a' \rangle \leq \langle s \,, a \rangle \} \subset \Sigma \;. \tag{A.6}$$

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 in Σ, that are not necessarily beliefs. The set $\Delta^\star_A(a)$ of beliefs compatible with optimal action a in (2.4) is related to the normal cone $N_A(a)$ at a in (A.6) by:

$$\Delta^\star_A(a) = N_A(a) \cap \Delta \;, \;\; \forall a \in A \;. \tag{A.7}$$

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

$$s \in \Sigma \text{ and } a \in F_A(s) \iff a \in A \text{ and } s \in N_A(a) . \tag{A.8}$$

A.2. Background on geometric convex analysis. A nonempty, convex and compact set $A \subset \mathbb{R}^K$ is called a *convex body* of $\mathbb{R}^K$ [34, p. 8].

Regular points and smooth bodies. We say that a point $a \in A$ is *smooth* or *regular* [34, p. 83] if the normal cone $N_A(a)$ in (A.3) is reduced to a half-line. The *set of regular points* is denoted by $\text{reg}(A)$:

$$a \in \text{reg}(A) \iff \exists s \in \Sigma\,,\ s \neq 0\,,\ N_A(a) = \mathbb{R}_+ s . \tag{A.9}$$

Notice that a regular point a necessarily belongs to the boundary ∂A of A: $\text{reg}(A) \subset \partial A$. The body A is said to be *smooth* if all boundary points of A are regular ($\text{reg}(A) = \partial A$); in that case, it can be shown that its boundary ∂A is a C^1 submanifold of $\mathbb{R}^K$ [34, Theorem 2.2.4, p. 83].

Spherical image map of A. We denote by $S^{|K|-1} = \{s \in \Sigma\,,\ \|s\| = 1\}$ the unit sphere of the signed measures Σ on K (identified with $\mathbb{R}^K$ with its canonical scalar product). By (A.9), we have that $a \in \text{reg}(A) \iff \exists! s \in S^{|K|-1}\,,\ N_A(a) = \mathbb{R}_+ s$. 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

$$n_A : \text{reg}(A) \to S^{|K|-1}\,, \ \text{ where } \text{reg}(A) \subset \partial A\,, \tag{A.10}$$

is called the *spherical image map of A*, or the *Gauss map*, and is continuous [34, p. 88]. We have

$$a \in \text{reg}(A) \Rightarrow N_A(a) = \mathbb{R}_+ n_A(a) \ \text{ where } \ n_A(a) \in S^{|K|-1} . \tag{A.11}$$

Reverse spherical image map of A. We say that a unit signed measure $s \in S^{|K|-1}$ is *regular* [34, p. 87] if the (exposed) face $F_A(s)$ of A in the direction s, as defined in (A.3), is reduced to a singleton. The *set of regular unit signed measures* is denoted by $\text{regn}(A)$:

$$s \in \text{regn}(A) \iff s \in S^{|K|-1} \text{ and } \exists! a \in A\,,\ F_A(s) = \{a\} . \tag{A.12}$$

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

$$f_A : \text{regn}(A) \to \partial A\,, \ \text{ where } \text{regn}(A) \subset S^{|K|-1}\,, \tag{A.13}$$

is called the *reverse spherical image map of A*, and is continuous [34, p. 88]. We have

$$s \in \text{regn}(A) \Rightarrow F_A(s) = \{f_A(s)\} . \tag{A.14}$$

Bodies with C^2 surface.

PROPOSITION A.1 (Schneider 2014, p. 113). *If the body A has boundary ∂A which is a C^2 submanifold of $\mathbb{R}^K$, then i) all points $a \in \partial A$ are regular ($\text{reg}(A) = \partial A$), ii) the spherical image map n_A in (A.10) is defined over the whole boundary ∂A and is of class C^1, iii) the spherical image map n_A has the reverse spherical image map f_A in (A.10) as right inverse, that is, $n_A \circ f_A = \text{Id}_{\text{regn}(A)}$.*

Proof. The first two items can be found in [34, p. 113]. Now, we prove that $n_A \circ f_A = \text{Id}_{\text{regn}(A)}$. As $f_A : \text{regn}(A) \to \partial A$ by (A.13), and as $n_A : \partial A \to S^{|K|-1}$ by (A.10) since $\text{reg}(A) = \partial A$, the mapping $n_A \circ f_A : \text{regn}(A) \to S^{|K|-1}$ is well defined. Let $s \in \text{regn}(A)$. By (A.14), we have that $F_A(s) = \{f_A(s)\}$ and by (A.11), we have that $N_A\big(f_A(s)\big) = \mathbb{R}_+ n_A\big(f_A(s)\big)$. From (A.8) — 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 $s = n_A\big(f_A(s)\big)$ by (A.10). ∎

Weingarten map. Let $a \in \text{reg}(A)$ be a regular point, as in (A.9), such that the spherical image map n_A in (A.10) is differentiable at a, with differential denoted by $T_a n_A$. The *Weingarten map* [34, p. 113] $T_a n_A : T_a \partial A \to T_{n_A(a)} S^{|K|-1}$ 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 [34, p. 114]; they are nonnegative [34, 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 [34, p. 115].

Reverse Weingarten map. Let $s \in \operatorname{regn}(A)$ be a regular unit signed measure such that the reverse spherical image map f_A in (A.13) is differentiable at s, with differential denoted by $T_s f_A$. The *reverse Weingarten map*

$$T_s f_A : T_s S^{|K|-1} \to T_{f_A(s)} \partial A \tag{A.15}$$

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.

A.3. Proofs of the results in Sect. 3. Using the relations (A.5) and (A.7), we express the proofs of the results in Sect. 3 in terms of the sets $F_A(p)$ in (2.1) and $N_A(a)$ in (A.6) (in the set Σ of signed measures), instead of $A^\star(p)$ in (2.3) and $\Delta^\star_A(a)$ in (2.4) (in the set Δ of probability measures).

Value of information. We have seen in (A.2) that the value function $v_A : \Delta \to \mathbb{R}$ in (2.2) is the restriction of the support function σ_A to beliefs in Δ. By definition (2.6) of the value of information, we deduce that, for any information structure $\mathbf{q}$ as in (2.5), we have:

$$\mathbf{VoI}_A(\mathbf{q}) = \mathbb{E}\left[\sigma_A(\mathbf{q}) - \sigma_A(\bar{p})\right] . \tag{A.16}$$

LEMMA A.2. *Let us introduce, for all $q \in \Delta$,*

$$\varphi^+_A(q) = \sigma_A(q) - \sigma_A(\bar{p}) + \sigma_{-A^\star(\bar{p})}(q - \bar{p}) \; , \tag{A.17a}$$

$$\varphi^-_A(q) = \sigma_A(q) - \sigma_A(\bar{p}) - \sigma_{A^\star(\bar{p})}(q - \bar{p}) \; . \tag{A.17b}$$

Then, for any information structure $\mathbf{q}$ and for any $a \in A$, we have that

$$\mathbb{E}\left[\varphi^+_A(\mathbf{q})\right] = \mathbb{E}\left[\sigma_A(\mathbf{q}) - \sigma_A(\bar{p}) + \sigma_{-A^\star(\bar{p})}(\mathbf{q} - \bar{p})\right] \tag{A.18a}$$

$$\geq \mathbf{VoI}_A(\mathbf{q}) = \mathbb{E}\left[\sigma_A(\mathbf{q}) - \sigma_A(\bar{p}) - \langle \mathbf{q} - \bar{p} \, , a \rangle\right] \tag{A.18b}$$

$$\geq \mathbb{E}\left[\sigma_A(\mathbf{q}) - \sigma_A(\bar{p}) - \sigma_{A^\star(\bar{p})}(\mathbf{q} - \bar{p})\right] = \mathbb{E}\left[\varphi^-_A(\mathbf{q})\right] . \tag{A.18c}$$

Proof. By (A.17), we have, for all $q \in \Delta$,

$$\varphi^+_A(q) = \sigma_A(q) - \sigma_A(\bar{p}) + \sigma_{-A^\star(\bar{p})}(q - \bar{p}) \tag{A.19a}$$

$$= \sup_{a \in A^\star(\bar{p})} \Big(\sigma_A(q) - \sigma_A(\bar{p}) - \langle q - \bar{p} \, , a \rangle \Big) \tag{A.19b}$$

$$\geq \sigma_A(q) - \sigma_A(\bar{p}) - \langle q - \bar{p} \, , a \rangle \; , \;\; \forall a \in A^\star(\bar{p}) \tag{A.19c}$$

$$\geq \inf_{a \in A^\star(\bar{p})} \Big(\sigma_A(q) - \sigma_A(\bar{p}) - \langle q - \bar{p} \, , a \rangle \Big) \tag{A.19d}$$

$$= \sigma_A(q) - \sigma_A(\bar{p}) - \sigma_{A^\star(\bar{p})}(q - \bar{p}) = \varphi^-_A(q) \; . \tag{A.19e}$$

By taking the expectation, we obtain (A.18), using (A.16) and the property that $\mathbb{E}\left[\mathbf{q} - \bar{p}\right] = 0$ in (2.5). ☐

Confidence set and indifference kernel. We start by providing characterizations of the confidence set $\Delta^{\mathrm{c}}_A(\bar{p})$ in (3.1) and of the indifference kernel $\Sigma^{\mathrm{i}}_A(\bar{p})$ in (3.5), in terms of $F_A(p)$ in (A.3) and $N_A(a)$ in (A.6).

PROPOSITION A.3.

1. *The confidence set $\Delta^c_A(\bar{p})$ of (3.1) is the nonempty closed and convex set*

$$\Delta^c_A(\bar{p}) = \bigcap_{a \in A^\star(\bar{p})} \Delta^\star_A(a) = \bigcap_{a \in F_A(\bar{p})} N_A(a) \cap \Delta \; . \tag{A.20}$$

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

$$p \in \Delta^c_A(\bar{p}) \iff F_A(\bar{p}) \subset F_A(p) \tag{A.21a}$$

$$\text{(A.21b)} \qquad \iff \sigma_A(p) - \sigma_A(\bar{p}) - \langle p - \bar{p}\,, a\rangle = 0 \;, \;\; \forall a \in F_A(\bar{p})$$

$$\text{(A.21c)} \qquad \iff \sigma_A(p) - \sigma_A(\bar{p}) + \sigma_{-A^\star(p)}(p - \bar{p}) = 0 \;.$$

3. *The indifference kernel* $\Sigma^i_A(\bar{p})$ *of* (3.5) *is the vector subspace*

$$\Sigma^i_A(\bar{p}) = \left[F_A(\bar{p}) - F_A(\bar{p})\right]^\perp = \left[A^\star(\bar{p}) - A^\star(\bar{p})\right]^\perp = \bigcap_{a \in F_A(\bar{p})} N_{F_A(\bar{p})}(a) \;.$$

Proof.
1. Express (3.1) using (A.7).
2. We prove the three equivalences in (A.21).
   (a) Let $p \in \Delta$. Using the property (A.8) that exposed face F_A and normal cone N_A are conjugate,

$$\text{we obtain: } p \in \Delta^{\mathrm{c}}_A(\bar{p}) \iff p \in \bigcap_{a \in F_A(p)} N_A(a) \;\text{ by (A.20)}$$
$$\iff a \in F_A(p) \;, \;\; \forall a \in F_A(\bar{p}) \;\text{ by (A.8)} \quad \iff \quad F_A(\bar{p}) \subset F_A(p) \;.$$

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

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

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 $A^\star(p)$ of optimal actions under prior belief $\bar{p}$ by (2.3) and (A.3)

$$\text{(by definition (A.6) of } N_A(a)\text{)} \qquad \iff p \in \bigcap_{a \in F_A(\bar{p})} N_A(a)$$
$$\iff p \in \bigcap_{a \in F_A(\bar{p})} N_A(a) \cap \Delta = \Delta^{\mathrm{c}}_A(\bar{p}) \;\text{ by (A.20).}$$

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

$$\text{(A.22)} \qquad \varphi_a(q) = \sigma_A(q) - \sigma_A(\bar{p}) - \langle q - \bar{p}\,, a\rangle \;\;, \;\; \forall q \in \Delta \;.$$

   By (A.4) and (A.21b), we have that

$$\text{(A.23a)} \qquad \forall a \in F_A(\bar{p}) \;, \;\; \forall q \in \Delta \;, \;\; \varphi_a(q) \geq 0 \;,$$
$$\text{(A.23b)} \qquad \forall a \in F_A(\bar{p}) \;, \;\; \forall q \in \Delta^{\mathrm{c}}_A(\bar{p}) \;, \;\; \varphi_a(q) = 0 \;.$$

   Let $p \in \Delta$. Using (A.23a), we deduce from (A.21b) and from the compacity of $F_A(\bar{p})$ that $p \in \Delta^{\mathrm{c}}_A(\bar{p}) \iff \inf_{a \in F_A(\bar{p})} \Big(\sigma_A(p) - \sigma_A(\bar{p}) - \langle p - \bar{p}\,, a\rangle\Big) = 0$. We conclude with (A.19d)–(A.19e).
3. Express (3.5) using (A.5). Then, use the definition of $N_{F_A(\bar{p})}(a)$ in (A.6).

This ends the proof. ∎

A.3.1. Valuable information.

Proof of Proposition 3.1. Let $a \in F_A(\bar{p})$ and $\mathbf{q}$ be an information structure as in (2.5). We have that

$$\mathbf{VoI}_A(\mathbf{q}) = 0 \iff \mathbb{E}\left[\sigma_A(\mathbf{q}) - \sigma_A(\bar{p})\right] = 0 \;\text{ by (A.16)}$$
$$\iff \mathbb{E}\left[\sigma_A(\mathbf{q}) - \sigma_A(\bar{p}) - \langle \mathbf{q} - \bar{p}\,, a\rangle\right] = 0 \;, \quad \text{as } \mathbb{E}\left[\mathbf{q} - \bar{p}\right] = 0$$

(because $\sigma_A(\mathbf{q}) - \sigma_A(\bar{p}) - \langle \mathbf{q} - \bar{p}\,, a\rangle \geq 0$ by (A.4) since $a \in F_A(\bar{p})$)

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

$$
\begin{aligned}
\text{(because } \sigma_A(\bar p) = \langle \bar p \,, a\rangle \text{ since } a \in F_A(\bar p)) \quad &\iff \sigma_A(\mathbf{q}) = \langle \mathbf{q} \,, a\rangle \ , \ \mathbb{P}-\text{a.s.} \\
&\iff \mathbb{P}\left\{ a \in F_A(\mathbf{q}) \right\} = 1 \\
&\iff \mathbb{P}\left\{ \langle \mathbf{q} \,, a'-a\rangle \leq 0 \ , \ \forall a' \in A \right\} = 1 \ .
\end{aligned}
$$

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 $\mathbf{VoI}_A(\mathbf{q}) = 0 \Rightarrow \mathbb{P}\left\{ \langle \mathbf{q} \,, a'-a\rangle \leq 0 \ , \ \forall a' \in A \ , \ \forall a \in F \right\} = 1$. 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 Ω), we get that $\left\{ \langle \mathbf{q} \,, a'-a\rangle \leq 0 \ , \ \forall a' \in A \ , \ \forall a \in F \right\} \subset \left\{ \langle \mathbf{q} \,, a'-a\rangle \leq 0 \ , \ \forall a' \in A \ , \ \forall a \in \overline{F} \right\}$. We deduce from the last equality that $\mathbf{VoI}_A(\mathbf{q}) = 0 \Rightarrow \mathbb{P}\left\{ \langle \mathbf{q} \,, a'-a\rangle \leq 0 \ , \quad \forall a' \in A \ , \ \forall a \in \overline{F} \right\} = 1$. Now, since $\overline{F} = F_A(\bar p)$, we finally get that $\mathbf{VoI}_A(\mathbf{q}) = 0 \Rightarrow \mathbb{P}\left\{ \langle \mathbf{q} \,, a'-a\rangle \leq 0 \ , \quad \forall a' \in A \ , \ \forall a \in F_A(\bar p) \right\} = 1$. In other words, we have obtained that, by definition (A.6) of the normal cone $N_A(a)$: $\mathbf{VoI}_A(\mathbf{q}) = 0 \Rightarrow \mathbf{q} \in \bigcap_{a \in F_A(\bar p)} N_A(a) \ , \ \mathbb{P}-\text{a.s.}$. Since $\mathbf{q} \in \Delta$, we conclude by (A.20) that

$$
\mathbf{VoI}_A(\mathbf{q}) = 0 \Rightarrow \mathbf{q} \in \bigcap_{a \in F_A(p)} N_A(a) \cap \Delta = \bigcap_{a \in A^\star(p)} \Delta^\star_A(a) = \Delta^{\mathrm{c}}_A(p) \ .
$$

Revisiting the proof backward, or using (A.21b), we easily see that $\mathbf{q} \in \Delta^{\mathrm{c}}_A(p) \ , \ \mathbb{P}-\text{a.s.} \Rightarrow \mathbf{VoI}_A(\mathbf{q}) = 0$. This ends the proof. ∎

Proof of Theorem 3.2. Let $\mathbf{q}$ be an information structure as in (2.5).

First, we show the upper estimate $C_A \mathbb{E}\, d\big(\mathbf{q}, \Delta^{\mathrm{c}}_A(\bar p)\big) \geq \mathbf{VoI}_A(\mathbf{q})$ in (3.4). For this purpose, we consider $a \in A$ and we show that the function φ_a in (A.22) is such that

$$
\text{(A.26)} \qquad \varphi_a(q) \leq \sup_{a' \in A} \| a - a' \| \inf_{p \in \Delta^{\mathrm{c}}_A(\bar p)} \| p - q \| \ .
$$

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

$$
\begin{aligned}
\varphi_a(q) &= \varphi_a(q) - \varphi_a(p) \ \text{ by (A.23b) since } p \in \Delta^{\mathrm{c}}_A(\bar p) \\
&= \sigma_A(q) - \sigma_A(p) - \langle q - p \,, a\rangle \ \text{ by (A.22)} \\
&= \sigma_{A-a}(q) - \sigma_{A-a}(p) \ \text{ by (A.1)} \\
&\leq \sup_{a' \in A-a} \| a' \| \times \| p - q \| \ \text{ by (A.1)} \ = \sup_{a' \in A} \| a - a' \| \times \| p - q \| \ .
\end{aligned}
$$

By taking the infimum with respect to all $p \in \Delta^{\mathrm{c}}_A(\bar p)$, we obtain (A.26). Then, we deduce that

$$
\begin{aligned}
\mathbf{VoI}_A(\mathbf{q}) &= \mathbb{E}\left[\varphi_a(\mathbf{q}) \right] \ , \ \forall a \in A \ \text{ by (A.18b)} \\
&= \inf_{a \in A} \mathbb{E}\left[\varphi_a(\mathbf{q}) \right] \leq \inf_{a \in A} \sup_{a' \in A} \| a - a' \| \times \mathbb{E}\left[\inf_{p \in \Delta^{\mathrm{c}}_A(\bar p)} \| p - q \| \right] \ \text{ by (A.26)}.
\end{aligned}
$$

With $C_A = \inf_{a \in A} \sup_{a' \in A} \| a - a' \|$ and (3.3), this gives the upper estimate $C_A \mathbb{E}\, d\big(\mathbf{q}, \Delta^{\mathrm{c}}_A(\bar p)\big) \geq \mathbf{VoI}_A(\mathbf{q})$ in (3.4).

Second, we show the lower estimate $\mathbf{VoI}_A(\mathbf{q}) \geq c_{\bar p, A, \varepsilon} \mathbb{P}\{ \mathbf{q} \not\in \Delta^{\mathrm{c}}_{A,\varepsilon}(\bar p) \}$ in (3.4). We consider an open subset $\mathcal{Q}$ of Δ that contains the confidence set $\Delta^{\mathrm{c}}_A(p)$, that is, $\Delta^{\mathrm{c}}_A(\bar p) \subset \mathcal{Q}$. By Lemma A.4 right below, there exists an $a \in F_A(\bar p)$ such that the continuous function φ_a in (A.22) is strictly positive on $\Delta^{\mathrm{c}}_A(\bar p)^c$. As $\mathcal{Q}^c \subset \Delta^{\mathrm{c}}_A(\bar p)^c$ and $\mathcal{Q}^c$ is a closed subset of the compact Δ, we can define $c_{\bar p, A} = \inf_{p \not\in \mathcal{Q}} \varphi_a(p) > 0$. We deduce that

$$
\begin{aligned}
\mathbf{VoI}_A(\mathbf{q}) &= \mathbb{E}\left[\varphi_a(\mathbf{q}) \right] \ \text{ by (A.18b)} \\
&= \mathbb{E}\left[\mathbf{1}_{\mathbf{q} \in \Delta^{\mathrm{c}}_A(\bar p)} \varphi_a(\mathbf{q}) + \mathbf{1}_{\mathbf{q} \not\in \Delta^{\mathrm{c}}_A(\bar p)} \varphi_a(\mathbf{q}) \right] \\
&= \mathbb{E}\left[\mathbf{1}_{\mathbf{q} \not\in \Delta^{\mathrm{c}}_A(\bar p)} \varphi_a(\mathbf{q}) \right] \ \text{ by (A.23b)} \\
&\geq \mathbb{E}\left[\mathbf{1}_{\mathbf{q} \not\in \mathcal{Q}} \varphi_a(\mathbf{q}) \right] \geq \mathbb{E}\left[\mathbf{1}_{\mathbf{q} \not\in \mathcal{Q}} c_{\bar p, A} \right] = c_{\bar p, A} \mathbb{P}\{ \mathbf{q} \not\in \mathcal{Q} \} \ .
\end{aligned}
$$

With $\mathcal{Q} = \Delta^{\mathrm{c}}_{A,\varepsilon}(\bar p)$, we put $c_{\bar p, A, \varepsilon} = \inf_{p \not\in \Delta^{\mathrm{c}}_{A,\varepsilon}(\bar p)} \varphi_a(p) > 0$.

This ends the proof. ∎

Lemma A.4. *There exists at least one $a \in F_A(\bar{p})$ such that the function φ_a in (A.22) 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 (A.21b), we have that $q \notin \Delta_A^{c}(\bar{p}) \iff \varphi_a(q) > 0$.

Suppose that $F_A(\bar{p})$ is a not singleton. Recall that the *affine hull* $\mathrm{aff}(S)$ of a subset S of $\mathbb{R}^K$ is the intersection of all affine manifolds containing S, and that the *relative interior* $\mathrm{ri}(C)$ of a nonempty convex set $C \subset \mathbb{R}^K$ is the nonempty interior of C for the topology relative to the affine hull $\mathrm{aff}(C)$ [22, p. 103]. We prove that any $a \in \mathrm{ri}\big(F_A(q)\big)$ answers the question. Let $a \in \mathrm{ri}\big(F_A(q)\big)$ be fixed. For any $q \notin \Delta_A^{c}(\bar{p})$, by (A.21a) 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 $a \in \mathrm{ri}\big(F_A(q)\big)$, there exists $a' \in \mathrm{ri}\big(F_A(q)\big)$ 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 (A.1) of σ_A) and $\sigma_A(q) > \langle q\,, \bar{a}\rangle$ (as $\bar{a} \notin F_A(q)$), we deduce that $\sigma_A(q) = \lambda\sigma_A(q) + (1-\lambda)\sigma_A(q) > \lambda\langle q\,, a'\rangle + (1-\lambda)\langle q\,, \bar{a}\rangle = \langle q\,, a\rangle$, where we used the property that $\lambda \in]0,1[$. Using the definition (A.22) 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. ∎

A.3.2. Undecided.

Proof of Theorem 3.4. We prove the three inequalities in (3.6).

I). We prove the upper inequality $C_A\mathbb{E}\|\mathbf{q}-\bar{p}\| \geq \mathbf{VoI}_A(\mathbf{q})$ in (3.6).

By definition (A.1) 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 (3.6).

II). We prove the middle inequality $\mathbf{VoI}_A(\mathbf{q}) \geq \mathbf{VoI}_{A^\star(\bar{p})}(\mathbf{q})$ in (3.6).

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

$$\sigma_A(s) - \sigma_A(\bar{p}) \geq \sigma_{F_A(\bar{p})}(s-\bar{p}) \text{ by (A.4) since } F_A(\bar{p}) \neq \emptyset \tag{A.30a}$$
$$= \langle s-\bar{p}\,, a\rangle\ ,\ \forall a \in F_A(\bar{p}) \text{ by definition of } \sigma_{F_A(\bar{p})} \tag{A.30b}$$
$$= \sigma_{F_A(\bar{p})}(s) - \sigma_{F_A(\bar{p})}(\bar{p}) \text{ by definition of } \sigma_{F_A(\bar{p})}. \tag{A.30c}$$

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

$$\mathbf{VoI}_A(\mathbf{q}) = \mathbb{E}\left[\sigma_A(\mathbf{q}) - \sigma_A(\bar{p})\right] \text{ by (2.6) and (A.2)} \tag{A.31a}$$
$$\geq \mathbb{E}\left[\sigma_{F_A(\bar{p})}(\mathbf{q}-\bar{p})\right] \text{ by (A.30a)} \tag{A.31b}$$
$$= \mathbb{E}\left[\sigma_{F_A(\bar{p})}(\mathbf{q}) - \sigma_{F_A(\bar{p})}(\bar{p})\right] \text{ by (A.30c)} \tag{A.31c}$$
$$= \mathbf{VoI}_{F_A(\bar{p})}(\mathbf{q}) \text{ by (2.6) and (A.2).}$$

This ends the proof of the middle inequality.

III). We prove the right hand side inequality $\mathbf{VoI}_{A^\star(\bar{p})}(\mathbf{q}) \geq \mathbb{E}\|\mathbf{q}-\bar{p}\|_{\Sigma_A^{i}(\bar{p})}$ in (3.6).

Let n be the dimension of the affine hull $\mathrm{aff}\big(F_A(\bar{p})\big)$ of $F_A(\bar{p})$, and let $a_1, \ldots, a_n$ be n actions in $F_A(\bar{p})$ that generate $\mathrm{aff}\big(F_A(\bar{p})\big)$. We put

$$T = \{a_1, \ldots, a_n\} \subset F_A(\bar{p}) \text{ so that } \mathrm{aff}\big(F_A(\bar{p})\big) = \mathrm{aff}\{a_1, \ldots, a_n\} = \mathrm{aff}(T)\,. \tag{A.32}$$

We will now show that $\|\cdot\|_{\Sigma_A^{i}(\bar{p})} = \frac{1}{n}\sigma_{T-T}(\cdot)$ is a seminorm with kernel $(F_A(\bar{p}) - F_A(\bar{p}))^\perp$ that satisfies the right hand side inequality in (3.6).

First, the support function σ_{T-T} is a seminorm with kernel $(T-T)^\perp$, as easily seen. Now, we also easily see that, for any subset $S \subset \mathbb{R}^K$, one has $(S-S)^\perp = \big(\mathrm{aff}(S-S)\big)^\perp = \big(\mathrm{aff}(S) - \mathrm{aff}(S)\big)^\perp$. 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 $\mathrm{aff}(T) = \mathrm{aff}\big(F_A(\bar{p})\big)$ by (A.32). Second, we show that the right hand side inequality in (3.6) is satisfied. We have

$$\mathbf{VoI}_A(\mathbf{q}) \geq \mathbb{E}\left[\sigma_{F_A(\bar{p})}(\mathbf{q}-\bar{p})\right] \text{ by (A.31b)}$$

(because $T \subset F_A(\bar{p})$ and support functions (A.1) are monotone with respect to set inclusion)

$$\geq \mathbb{E}\left[\sigma_T(\mathbf{q}-\bar{p})\right]$$

$$= \mathbb{E}\left[\sigma_T(\mathbf{q}-\bar{p}) - \langle \mathbf{q}-\bar{p}\,, a\rangle\right]\ ,\ \ \forall a \in A \ \text{ because } \mathbb{E}\left[\langle \mathbf{q}-\bar{p}\,, a\rangle\right] = 0.$$
$$= \mathbb{E}\left[\sigma_{T-a}(\mathbf{q}-\bar{p})\right]\ ,\ \ \forall a \in A \ \text{ because } \sigma_{T-a} = \sigma_{T+\{-a\}} = \sigma_T + \sigma_{\{-a\}}.$$

Indeed, support functions transform a Minkowski sum of sets into a sum of support functions [22, p. 226]. Using again this property, we obtain that $\mathbf{VoI}_A(\mathbf{q}) \geq \frac{1}{n}\sum_{i=1}^{n}\mathbb{E}\left[\sigma_{T-a_i}(\mathbf{q}-\bar{p})\right] = \frac{1}{n}\mathbb{E}\left[\sigma_{\sum_{i=1}^{n}(T-a_i)}(\mathbf{q}-\bar{p})\right]$. Now, as $T=\{a_1,\ldots,a_n\}$, it is easy to see that the sum $\sum_{i=1}^{n}(T-a_i)$ contains any element of the form $a_k - a_l = (a_1-a_1)+\cdots+(a_{l-1}-a_{l-1})+(a_k-a_l)+(a_{l+1}-a_{l+1})+\cdots+(a_n-a_n) \in \sum_{i=1}^{n}(T-a_i)$. As support functions are monotone with respect to set inclusion, we deduce that $\sigma_{\sum_{i=1}^{n}(T-a_i)} \geq \sigma_{\{a_k-a_l,k,l=1,\ldots,n\}} = \sigma_{T-T}$ and that $\mathbf{VoI}_A(\mathbf{q}) \geq \frac{1}{n}\mathbb{E}\left[\sigma_{\{a_k-a_l,k,l=1,\ldots,n\}}(\mathbf{q}-\bar{p})\right] = \frac{1}{n}\mathbb{E}\left[\sigma_{T-T}(\mathbf{q}-\bar{p})\right] = \mathbb{E}\left\|\mathbf{q}-\bar{p}\right\|_{\Sigma_A^{\mathrm{i}}(\bar{p})}$.

This ends the proof. ∎

A.3.3. Flexible.

Proof of Proposition 3.5. All the reminders on geometric convex analysis in Sect. 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 consider the diffeomorphism

$$\nu : S^{|K|-1} \cap \mathbb{R}_+^K \to \Delta\ ,\ \ s \mapsto \frac{s}{\langle s\,, 1\rangle}\ , \tag{A.34}$$

that maps unit positive measures into probability measures, with inverse $\nu^{-1} : \Delta \to S^{|K|-1} \cap \mathbb{R}_+^K\ ,\ \ p \mapsto \frac{p}{\|p\|}$.

Since, by assumption, the action set A has boundary ∂A which is a C^2 submanifold of $\mathbb{R}^K$, we know by Proposition A.1 that the spherical image map $n_A : \partial A \to S^{|K|-1}$ in (A.10) is well defined, is of class C^1, and has for right inverse the reverse spherical image map $f_A : \mathrm{regn}(A) \to \partial A$ in (A.13), that is, $n_A \circ f_A = \mathrm{Id}_{\mathrm{regn}(A)}$.

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

$$a \in \mathrm{reg}^+(A) \iff \exists p \in \Delta\ ,\ \ N_A(a) = \mathbb{R}_+ p\ . \tag{A.35}$$

For a regular action $a \in \mathrm{reg}^+(A)$, there is only one probability $p \in \Delta$ such that $N_A(a) = \mathbb{R}_+ p$, and it is $p = \nu\big(n_A(a)\big)$. We have $a \in \mathrm{reg}^+(A) \Rightarrow N_A(a) = \mathbb{R}_+\nu\big(n_A(a)\big)$ where $\nu\big(n_A(a)\big) \in \Delta$. The *set of regular probabilities* is $\mathrm{regn}^+(A) = \Big(\mathbb{R}_+^*\mathrm{regn}(A)\Big) \cap \Delta$. For a regular probability $p \in \mathrm{regn}^+(A)$, there is only one action $a \in \partial A$ such that $F_A(p) = \{a\}$, and it is $a = f_A\big(\nu^{-1}(p)\big)$. Indeed, by definition (A.3) of the (exposed) face, we have that $F_A(\lambda s) = F_A(s)\ ,\ \forall \lambda \in \mathbb{R}_+^*\ ,\ \forall s \in \Sigma\ ,\ \ s \neq 0$. Therefore, we have that

$$p \in \mathrm{regn}^+(A) \Rightarrow F_A(p) = \{f_A\big(\nu^{-1}(p)\big)\}\ . \tag{A.36}$$

The following mappings are well defined: $\nu \circ n_A : \mathrm{reg}^+(A) \to \Delta$ and $f_A \circ \nu^{-1} : \mathrm{regn}^+(A) \to \partial A$, and we have that $(\nu \circ n_A) \circ (f_A \circ \nu^{-1}) = \mathrm{Id}_{\mathrm{regn}^+(A)}$.

- Item 2 ⇒ 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 (A.10) 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 $a^\sharp$ [34, 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 $n_A\big(a^\sharp\big)$ in $S^{|K|-1}$, and such that the restriction $n_A : \mathcal{A} \to n_A(\mathcal{A})$ of the spherical image map in (A.10) is a diffeomorphism. By item iii) in Proposition A.1, we have that $n_A\big(a^\sharp\big) = \frac{\bar{p}}{\|\bar{p}\|}$ and the local inverse coincides with the restriction $f_A : n_A(\mathcal{A}) \to \mathcal{A}$ of the reverse spherical image map in (A.13). As $n_A(\mathcal{A})$ is an open neighborhood of $\frac{\bar{p}}{\|\bar{p}\|}$ in $S^{|K|-1}$, and as the prior belief $\bar{p}$ has full support, we deduce that $\nu\big(n_A(\mathcal{A})\big)$ is an open neighborhood of $\bar{p}$ in Δ, where the diffeomorphism ν is defined in (A.34). We easily deduce that $f_A \circ \nu^{-1} : \nu\big(n_A(\mathcal{A})\big) \to \mathcal{A}$ is a diffeomorphism. By (A.36), 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 (3.7).

- Item 1 ⇒ Item 3.
Suppose that the set-valued mapping $F_A : \Delta \rightrightarrows A$, $p \mapsto F_A(p)$ in (3.7) is a local diffeomorphism at $\bar{p}$. By definition (A.12) of the set of regular unit signed measures, there exists an open neighborhood $\amalg$ of $\bar{p}$ in Δ such that $\amalg \subset \mathrm{regn}^+(A)$, where the set of relevant regular points is defined in (A.35). In addition, the mapping $f_A \circ \nu^{-1} : \amalg \to f_A\big(\nu^{-1}(\amalg)\big)$ is a diffeomorphism.
As $F_A(p) = \{f_A\big(\nu^{-1}(p)\big)\}$, for all beliefs $p \in \amalg$, we know that the support function σ_A is differentiable and that its gradient is $\nabla_p \sigma_A = f_A\big(\nu^{-1}(p)\big)$ [22, p. 251]. As $f_A \circ \nu^{-1}$ is a local diffeomorphism at $\bar{p}$, and as the mapping ν in (A.34) 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 positive definite.
- Item 3 ⇒ Item 2.
Suppose that the value function v_A is twice differentiable at $\bar{p}$ and the Hessian is positive definite. On the one hand, as the prior $\bar{p}$ has full support, there exists an open neighborhood $\amalg$ of $\bar{p}$ in Δ such that v_A is differentiable on $\amalg$. On the other hand, as the support function σ_A is positively homogeneous, and by (A.2), we have that

$$\sigma_A(s) = \langle s\,, 1\rangle \times \big(v_A \circ \nu\big)(s)\ , \ \ \forall s \in S^{|K|-1} \cap \mathbb{R}^K_+\ . \tag{A.37}$$

Therefore, as the mapping ν in (A.34) is a diffeomorphism, the support function σ_A is differentiable on the open neighborhood $\nu^{-1}(\amalg)$ 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 effective domain $\mathbb{R}^K$ is differentiable at s if and only if the subdifferential at s is a singleton [22, p. 251], and, on the other hand, the face $F_A(s)$ is the subdifferential at s of the support function σ_A [22, p. 258], we conclude that the face $F_A(s)$ of A in the direction $s \in \nu^{-1}(\amalg)$ is a singleton.
Therefore, by definition (A.12) of the set of regular unit signed measures, we have that $\nu^{-1}(\amalg) \subset \mathrm{regn}(A)$. In addition, the restriction $f_A : \nu^{-1}(\amalg) \to f_A\big(\nu^{-1}(\amalg)\big)$ of the reverse spherical image map in (A.13) is well defined, and we have that $\nabla_s \sigma_A = f_A(s)\ ,\ \ \forall s \in \nu^{-1}(\amalg)$. Therefore, the mapping $f_A : \nu^{-1}(\amalg) \to f_A\big(\nu^{-1}(\amalg)\big)$ 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 positive definite. This comes from (A.37), where the mapping ν in (A.34) is a C^∞ diffeomorphism and the value function v_A is twice differentiable at $\bar{p}$ with positive definite Hessian.
As f_A is is differentiable at $\frac{\bar{p}}{\|\bar{p}\|}$ and has full rank, the reverse Weingarten map $T_s f_A$ in (A.15) is well defined and has full rank. Therefore, the principal radii of curvature of A at $\frac{\bar{p}}{\|\bar{p}\|}$ are positive. Letting $a^\sharp = f_A\big(\frac{\bar{p}}{\|\bar{p}\|}\big)$, 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 3.6. We suppose that the value function v_A in (2.2) is twice differentiable at $\bar{p}$, with positive definite Hessian. We denote $F_A(\bar{p}) = \{a^\sharp\}$.

First, we show that the function $g(p) = \frac{v_A(p) - v_A(\bar{p}) - \langle p - \bar{p}\,, a^\sharp\rangle}{\|p-\bar{p}\|^2}$ is continuous and positive on Δ. Indeed, g is continuous on $\Delta \backslash \{\bar{p}\}$, 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 positive definite. We have $g \geq 0$ on $\Delta \backslash \{\bar{p}\}$, because $F_A(\bar{p}) = \{a^\sharp\}$ is the subdifferential at $\bar{p}$ of the support function σ_A, and by (A.2). 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 (nonzero) direction $p - \bar{p}$ would be zero; this would contradict the assumption that the Hessian of v_A at $\bar{p}$ is positive definite. 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 (3.8) from (2.6).

This ends the proof. ∎