

# 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](mailto:michel.delara@enpc.fr)

†CREST, CNRS, École Polytechnique, and London School of Economics. Email: [olivier.gossner@cirs.fr](mailto:olivier.gossner@cirs.fr)

50 suboptimal for some of the posteriors. We thus define the confidence set at a prior belief  $\bar{p}$  as the set of  
 51 posterior beliefs for which all optimal actions at  $\bar{p}$  remain optimal. Our result says that information has  
 52 positive value if and only if posterior beliefs fall outside of the confidence set with positive probability.  
 53 This result generalizes insights from [23] and [30], who had already noticed that information can only be  
 54 useful insofar as it influences choices. We provide corresponding lower and upper bounds to the value of  
 55 information.

56 In the second bounds, we express the fact that the value of information is maximal when it influences  
 57 actions the most, which happens when information breaks indifferences between several choices. We show  
 58 that, when this is the case, the value of information can be suitably measured by an expected distance  
 59 between the prior and the posterior. There are several optimal actions at the prior, and information that  
 60 allows to break indifferences has highest value.

61 Finally, our third bounds apply to cases in which the agent’s optimal choice is a smooth function of her  
 62 belief around the prior. We show that, in this situation, the value function is also smooth around the prior,  
 63 and the value of information is essentially a quadratic function of the expected distance between the prior  
 64 and the posterior. In this intermediate case, information impacts actions in a continuous way. The optimal  
 65 actions at the prior belief and at a posterior close to it are themselves close; so choosing one instead of the  
 66 other has a mild, albeit positive, impact on the expected payoff.

67 In a finite decision problem — such as shopping behavior [28] or residential location [29] — at any given  
 68 prior the agent either has an optimal action that is locally constant, or is indifferent between several optimal  
 69 choices. The first and second upper and lower bounds are particularly useful in finite choice problems. The  
 70 third bounds are most useful in decision problems with a continuum of choices, such as scoring rules [11] or  
 71 investment decisions [1].

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

77 **2. Model, payoffs-beliefs duality and information.** We consider the classical question of an agent  
 78 who faces a decision problem under imperfect information on a state of nature. The set of states of nature  
 79 is a finite set  $K$ . We identify the set  $\Sigma$  of signed measures on  $K$  with  $\mathbb{R}^K$ . The agent holds a prior belief  $\bar{p}$   
 80 with full support in the set  $\Delta = \Delta(K) \subset \Sigma = \mathbb{R}^K$  of probability distributions over  $K$ . We identify  $\Delta$  with  
 81 the simplex of  $\mathbb{R}^K$ .

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

$$85 \quad (2.1) \quad A = \overline{\text{co}}\{(g(d, k))_{k \in K}, d \in D\} \subset \mathbb{R}^K .$$

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

87 **Duality between actions/payoffs and beliefs.** The scalar product between a vector  $v \in \mathbb{R}^K$  and a  
 88 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  
 89 and beliefs. Such a duality is at the core of a series of works in nonexpected utility theory, such as [21, 27, 14].

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

$$93 \quad (2.2) \quad v_A(p) = \max_{a \in A} \langle p, a \rangle , \quad \forall p \in \Delta .$$

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

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

$$97 \quad (2.3) \quad A^*(p) = \arg \max_{a' \in A} \langle p, a' \rangle = \{a \in A \mid \forall a' \in A, \langle p, a' \rangle \leq \langle p, a \rangle\} .$$

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

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

$$103 \quad (2.4) \quad \Delta_A^*(a) = \{p \in \Delta \mid \forall a' \in A, \langle p, a' \rangle \leq \langle p, a \rangle\}.$$

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

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

108 **Information structure.** We follow [9, 8], and we describe information through a distribution of pos-  
 109 terior beliefs that average to the prior belief. Hence, given the prior belief  $\bar{p}$ , we define an *information*  
 110 *structure* as a random variable  $\mathbf{q}$ , defined over a probability space  $(\Omega, \mathcal{F}, \mathbb{P})$  and with values in  $\Delta$ , describing  
 111 the agent’s posterior beliefs, and such that (where  $\mathbb{E}$  denotes the expectation operator with respect to  $\mathbb{P}$ )

$$112 \quad (2.5) \quad \mathbf{q} : (\Omega, \mathcal{F}, \mathbb{P}) \rightarrow \Delta, \quad \mathbb{E}[\mathbf{q}] = \bar{p}.$$

113 Given the action set  $A$  in (2.1) and the information structure  $\mathbf{q}$  in (2.5), the *value of information*  $\mathbf{VoI}_A(\mathbf{q})$   
 114 is the difference between the expected payoff for an agent who receives information according to  $\mathbf{q}$  and one  
 115 whose prior belief is  $\bar{p}$ . It is given by:

$$116 \quad (2.6) \quad \mathbf{VoI}_A(\mathbf{q}) = \mathbb{E}[v_A(\mathbf{q})] - v_A(\bar{p}).$$

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

118 **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

119 function  $v_A$ , expressed as a function of the probability  $p$  of state 2, is the maximum of the following three  
 120 affine functions:  $3(1-p)$ ,  $2$ , and  $5p/2$ . Action  $(3, 0)$  is optimal for  $p \leq 1/3$ ,  $(2, 2)$  is optimal for  $p \in [1/3, 4/5]$ ,  
 121 and  $(0, 5/2)$  is optimal for  $p \geq 4/5$ . Both the set  $A$  and the function  $v_A$  are represented in Figure 1.  
 122

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

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

130 **3. On the value of information.** In this section, we relate the geometry of the set  $A$  of actions  
 131 in (2.1) with the behavior of the agent around the prior belief  $\bar{p}$ , with differentiability properties of the value  
 132 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  
 133 allows us to derive bounds on the value of information that depend on how information influences actions.

134 First, in Subsect. 3.1, we consider information that does not allow us to eliminate optimal actions. We  
 135 introduce the *confidence set* as the set of posterior beliefs at which all optimal actions at the prior remain

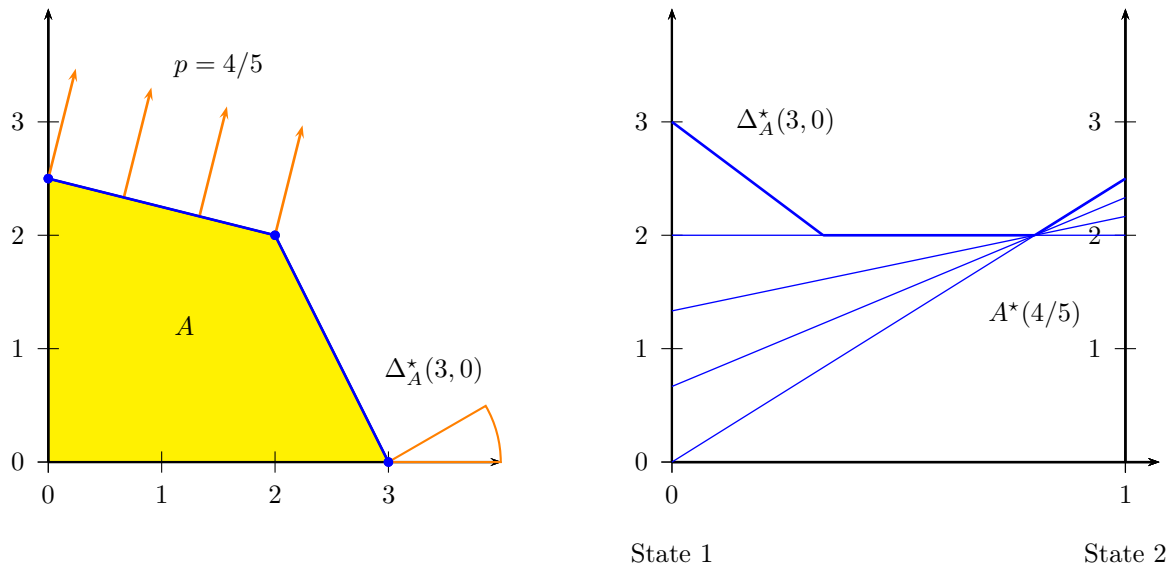


FIG. 1. The set  $A$  of actions on the left, and the value function  $v_A$  on the right. Each of the four arrows on the left represents an action  $a$  such that  $p = 4/5$  belongs to the set  $\Delta_A^*(a)$  of beliefs revealed by action  $a$ . On the right side, these four actions (each attached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

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

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

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

150 **3.1. Valuable information.** Our first task is to formalize the idea that useful information is informa-  
 151 tion that affects optimal choices (quoting [23], “Information is of value only if it can affect action”). Since  
 152 there are potentially several optimal actions at a prior belief  $\bar{p}$  and at a posterior  $p$ , there are in principle  
 153 many ways to formalize this idea.

154 We say that a belief  $p$  is in the *confidence set*  $\Delta_A^c(\bar{p})$  of prior belief  $\bar{p}$  iff all optimal actions at  $\bar{p}$  (those  
 155 in  $A^*(\bar{p})$ ) are also optimal at  $p$ . In other words, we define the *confidence set of prior belief*  $\bar{p}$  by:

$$156 \quad (3.1) \quad \Delta_A^c(\bar{p}) = \bigcap_{a \in A^*(\bar{p})} \Delta_A^*(a).$$

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

158  $p \in \Delta_A^c(\bar{p})$  when none of the actions chosen by the agent at prior belief  $\bar{p}$  would lead the observer to refute  
 159 the possibility that the agent has belief  $p$ .

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

161 PROPOSITION 3.1 (Valuable information). *For every information structure  $\mathbf{q}$  as in (2.5), we have:*

162 (3.2a) 
$$\mathbf{VoI}_A(\mathbf{q}) = 0 \iff \exists a^* \in A^*(\bar{p}), a^* \in A^*(\mathbf{q}), \mathbb{P} - a.s.$$

163 (3.2b) 
$$\iff \mathbf{q} \in \Delta_A^c(\bar{p}), \mathbb{P} - a.s.$$

165

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

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

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

177 (3.3) 
$$\Delta_{A,\varepsilon}^c(\bar{p}) = \{q \in \Delta \mid d(q, \Delta_A^c(\bar{p})) < \varepsilon\} \text{ where } d(q, \Delta_A^c(\bar{p})) = \inf_{p \in \Delta_A^c(\bar{p})} \|p - q\|.$$

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

179 THEOREM 3.2 (Bound on the value of information based on confidence sets). *For every  $\varepsilon > 0$ , there*  
 180 *exist positive constants  $C_A$  and  $c_{\bar{p},A,\varepsilon}$  such that, for every information structure  $\mathbf{q}$  as in (2.5):*

181 (3.4) 
$$C_A \mathbb{E} [d(\mathbf{q}, \Delta_A^c(\bar{p}))] \geq \mathbf{VoI}_A(\mathbf{q}) \geq c_{\bar{p},A,\varepsilon} \mathbb{P}\{\mathbf{q} \notin \Delta_{A,\varepsilon}^c(\bar{p})\}.$$

182

183 The upper bound tells us that the value of information is bounded by (a constant times) the expected  
 184 distance from the posterior to the confidence set at the prior. In particular, it is bounded by the expected  
 185 distance from the posterior to the prior itself. The lower bound is a converse result, but in which we need  
 186 to replace the confidence set by some  $\varepsilon$ -neighborhood. It shows us that the value of information is bounded  
 187 below by (a constant times) the probability that the posterior is at least distance  $\varepsilon$  from the confidence set,  
 188 and, therefore, it is also larger than the expected distance from the posterior to this  $\varepsilon$ -neighborhood of the  
 189 confidence set. Both the lower and upper bounds depend on the confidence set  $\Delta_A^c(\bar{p})$  in (3.1), which can be  
 190 computed locally at prior belief  $\bar{p}$ . On the other hand, they apply to all information structures. The caveat  
 191 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  
 192 the action set  $A$ .

193 **3.2. Undecided.** We now consider situations in which information influences actions the most. Those  
 194 are situations of indifference in which, at the prior belief  $\bar{p}$ , the agent is *undecided* between several optimal  
 195 actions. A small piece of information can then be enough to break this indifference. As shown by the  
 196 following proposition (whose proof we do not give, as it is well-known in convex analysis [22, p. 251]), the  
 197 value function then exhibits a *kink* at prior belief  $\bar{p}$ .

198 PROPOSITION 3.3. *The two following conditions are equivalent:*

- 199
  - the set  $A^*(\bar{p})$  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  $\bar{p}$ .

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

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

$$209 \quad (3.5) \quad \Sigma_A^i(\bar{p}) = [A^*(\bar{p}) - A^*(\bar{p})]^\perp .$$

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

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

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

$$219 \quad (3.6) \quad C_A \mathbb{E} \|\mathbf{q} - \bar{p}\| \geq \mathbf{VoI}_A(\mathbf{q}) \geq \mathbf{VoI}_{A^*(\bar{p})}(\mathbf{q}) \geq \mathbb{E} \|\mathbf{q} - \bar{p}\|_{\Sigma_A^i(\bar{p})} .$$

220

221 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  
 222 is bounded above and below by a constant times the norm-1 between the prior and the posterior. Since any  
 223 small amount of information allows to break the indifference between the optimal actions at these priors,  
 224 information is very valuable.

225 The lower bound in Theorem 3.4 shows that a lower bound of the value of information is the expectation  
 226 of a seminorm of the distance between the prior belief and the posterior belief. To understand the role  
 227 of the kernel  $\Sigma_A^i(\bar{p})$  of this seminorm, let us first consider the set of beliefs in this set. A posterior  $q$  is  
 228 in  $\Sigma_A^i(\bar{p}) = [A^*(\bar{p}) - A^*(\bar{p})]^\perp$  if and only if, for any two optimal actions  $a, a' \in A^*(\bar{p})$ ,  $\langle q, a \rangle = \langle q, a' \rangle$ . In  
 229 words, posteriors that do not break any of the ties in  $A^*(\bar{p})$  might not be valuable to the agent. On the other  
 230 hand, Theorem 3.4 tells us that all other directions — i.e., those that allow at least one of the ties in  $A^*(\bar{p})$   
 231 to be broken — are valuable to the agent, and furthermore, in these directions, the value of information  
 232 behaves like an expected distance from the prior to the posterior.

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

236 Note that the bounds on Theorem 3.4 rely on the indifference kernel  $\Sigma_A^i(\bar{p})$  in (3.5), which can be  
 237 computed directly from the set  $A^*(\bar{p})$  by (3.5). The multiplicative constant  $C_A$  in (3.6), however, depends  
 238 on more global properties of the action set  $A$ .

239 **3.3. Flexible.** Finally, we consider the case in which there is a unique optimal action for each belief  
 240 in the range considered, and this action depends smoothly on the belief. More precisely, we assume that,  
 241 around the prior, optimal actions smoothly depend in a 1-1 way on the belief. This assumption is met when,  
 242 for instance, the decision problem faced by the agent is a scoring rule [11], or an investment problem [1, 12].

243 Our first step is to characterize a class of situations of interest, in which the agent's optimal action  
 244 depends smoothly on her belief. The following proposition offers three alternative characterizations of these  
 245 situations, based 1) on the local behavior of the agent's optimal choices, 2) on local properties of  
 246 the geometry of the boundary of the set of actions, and 3) on local second differentiability properties of the  
 247 value function. For background on geometric convex analysis, the reader can consult §A.2 in the Appendix.

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

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

$$251 \quad (3.7) \quad A^* : \Delta \rightrightarrows \partial A, \quad p \mapsto A^*(p)$$



252 is a local diffeomorphism<sup>1</sup> at the prior belief  $\bar{p}$ ;  
 253 2. The set  $A^*(\bar{p})$  of optimal actions at the prior belief  $\bar{p}$  in (2.3) is reduced to a singleton at which the  
 254 curvature of the action set  $A$  is positive;  
 255 3. The value function  $v_A$  in (2.2) is twice differentiable at the prior belief  $\bar{p}$ , with positive definite  
 256 Hessian at  $\bar{p}$ .  
 257 In this case, we say that the agent is flexible at  $\bar{p}$ .

258 **THEOREM 3.6** (Bounds on the VoI for the flexible agent). *If the agent is flexible at prior belief  $\bar{p}$ , then*  
 259 *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):*

$$260 \quad (3.8) \quad C_{\bar{p},A} \mathbb{E} \|\mathbf{q} - \bar{p}\|^2 \geq \mathbf{VoI}_A(\mathbf{q}) \geq c_{\bar{p},A} \mathbb{E} \|\mathbf{q} - \bar{p}\|^2 .$$

261  
 262 Theorem 3.6 shows that, in the case of a flexible agent, the value of information is essentially given by the  
 263 expected square distance between the prior and the posterior, up to some multiplicative constant. One of the  
 264 strengths of the theorem is that its assumption that the agent is flexible is a local one, whereas its conclusion  
 265 is global, as it applies to all information structures. On the other hand, the multiplicative constants  $C_{\bar{p},A}$  and  
 266  $c_{\bar{p},A}$  in (3.8) themselves depend on the global behavior of the value function, and hence cannot be inferred  
 267 from local properties only.

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

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

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

$$275 \quad (4.1) \quad P(I) = \alpha I + f \text{ where } \alpha \in ]0, 1[, \quad f > 0 .$$

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

$$279 \quad (4.2) \quad u(w) = 1 - e^{-Rw} .$$

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

$$281 \quad (4.3) \quad A = \overline{\text{co}} \left\{ \left( u(\varpi), u(0) \right), \left( u(-P(I) + \varpi), u(-P(I) + I) \right) \right\}$$

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

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

$$285 \quad (4.4a) \quad U_0(p) = (1 - p)u(\varpi) + pu(0) = (1 - p)(1 - e^{-R\varpi}) ,$$

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

$$287 \quad (4.4b) \quad U(p, I) = (1 - p)u(-P(I) + \varpi) + pu(-P(I) + I) = 1 - pe^{-R(-P(I)+I)} - (1 - p)e^{-R(-P(I)+\varpi)} .$$

288

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

---

<sup>1</sup>In particular, the set  $A^*(p)$  is a singleton for all  $p \in \Delta$ , in which case we identify a singleton set with its single element.

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

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

296 *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  
 297 maximum, characterized by  $\partial U / \partial I = 0$ , and achieved at

$$298 \quad (4.5) \quad \hat{I}(p) = \varpi - \frac{1}{R} \ln\left(\frac{1-p}{p} \frac{\alpha}{1-\alpha}\right), \quad \forall p \in ]0, 1[.$$

299 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  
 300 a nonnegative level of indemnity is chosen, we introduce the difference of expected utilities

$$301 \quad (4.6) \quad \delta(p) = \max_{I \geq 0} U(p, I) - U_0(p) = \begin{cases} U(p, 0) - U_0(p) & \text{if } p \leq \hat{p}, \\ U(p, \hat{I}(p)) - U_0(p) & \text{if } p \geq \hat{p}. \end{cases}$$

302 We study the behavior of the function  $\delta$  when  $p$  is small and when  $p$  is close to one. After computa-  
 303 tion, we find that, for all  $p \in [0, 1]$ ,  $U(p, 0) - U_0(p) = -(e^{Rf} - 1)(p + (1-p)e^{-R\varpi}) < 0$ . Therefore,  
 304  $\delta(p) < 0$  for all  $p \leq \hat{p}$ . On the other hand, when  $p$  goes to 1,  $\delta(p)$  goes to 1 because  $U_0(p) \rightarrow 0$  and  
 305  $U(p, \hat{I}(p)) = (1-p)\left(1 - e^{-R(-P(\hat{I}(p)) + \varpi)}\right) + p\left(1 - e^{-R(-P(\hat{I}(p)) + \hat{I}(p))}\right) = 1 - (1-p)\left(\frac{1-p}{p} \frac{\alpha}{1-\alpha}\right)^\alpha e^{R(1-\alpha)\varpi} -$   
 306  $p\left(\frac{1-p}{p} \frac{\alpha}{1-\alpha}\right)^{1-\alpha} e^{-R(1-\alpha)\varpi} \rightarrow 1$  (as  $\alpha \in ]0, 1[$ ). As a consequence, we can define  $p^* = \inf \{p \in [0, 1] \mid \delta(p) > 0\}$ ,  
 307 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  $\delta(p) \rightarrow 1$   
 308 when  $p \rightarrow 1$ . We now check that  $p^*$  and  $\hat{I}$  in (4.5) satisfy the three assertions of the Proposition.

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

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

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

$$314 \quad \delta(p) = \delta(p) - \delta(p^*) \quad \text{as } \delta(p^*) = 0$$

$$315 \quad = U(p, \hat{I}(p)) - U(p, \hat{I}(p^*)) + U(p, \hat{I}(p^*)) - U_0(p) - [U(p^*, \hat{I}(p^*)) - U_0(p^*)] \text{ by (4.6)}$$

$$316 \quad > U(p, \hat{I}(p^*)) - U_0(p) - U(p^*, \hat{I}(p^*)) + U_0(p^*) \quad \text{as } U(p, \hat{I}(p)) - U(p, \hat{I}(p^*)) > 0$$

$$317 \quad \text{by definition of the maximizer } \hat{I}(p) \text{ and since } \hat{I}(p) > \hat{I}(p^*) \geq 0 \text{ as } p > p^* \geq \hat{p}$$

$$318 \quad = (1-p)[u(-P(\hat{I}(p^*)) + \varpi) - u(\varpi)] + p[u(-P(\hat{I}(p^*)) + \hat{I}(p^*)) - u(0)]$$

$$319 \quad - (1-p^*)[u(-P(\hat{I}(p^*)) + \varpi) - u(\varpi)] - p^*[u(-P(\hat{I}(p^*)) + \hat{I}(p^*)) - u(0)] \text{ by (4.4)}$$

$$320 \quad = (p-p^*)\left[[u(-P(\hat{I}(p^*)) + \hat{I}(p^*)) - u(0)] + [u(\varpi) - u(-P(\hat{I}(p^*)) + \varpi)]\right] \geq 0$$

322 since both terms between inner brackets are increments of the increasing function  $u$ , where  $-P(\hat{I}(p^*)) +$   
 323  $\hat{I}(p^*) \geq 0$  (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  
 324 would arrive at the contradiction that  $0 = \delta(p^*) = (1-p^*)[u(-P(\hat{I}(p^*)) + \varpi) - u(\varpi)] + p^*[u(-P(\hat{I}(p^*)) +$   
 325  $\hat{I}(p^*)) - u(0)] < 0$  since both terms between brackets are (negative) increments of the increasing function  $u$ .  $\square$

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

$$330 \quad (4.7) \quad v(q) = \max \left\{ U_0(q), \max_{I \geq 0} U(q, I) \right\}.$$



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

$$333 \quad (4.8) \quad \mathbf{VoI}(\varepsilon) = \frac{1}{2}v(p + \varepsilon) + \frac{1}{2}v(p - \varepsilon) - v(p).$$

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

337 **PROPOSITION 4.2.** *Depending on the probability of loss  $p$ , the value of information for small  $\varepsilon$  behaves*  
 338 *as follows:*

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

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

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

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

366 **5. The marginal value of information.** The question of the marginal value of information is studied  
 367 in [32]. They provide joint conditions on a parameterized family of information structures together with  
 368 a decision problem such that, when the agent is close to receiving no information at all, the marginal  
 369 value of information is null. Their result was subsequently generalized in [15] and [16], where are provided  
 370 joint conditions on parameterized information and a decision problem leading to zero marginal value of  
 371 information.

372 In this Section, we show how our bounds on the value of information, obtained in Sect. 3, apply to  
 373 the marginal value of information. In Subsect. 5.1, we provide separate conditions on the decision problem  
 374 and on the family of parameterized information structures that result in a null value of information. We  
 375 then examine, in Subsect. 5.2, several parameterized families of information structures and rely on our main  
 376 results to study how the marginal value of information varies depending on the decision problem faced.

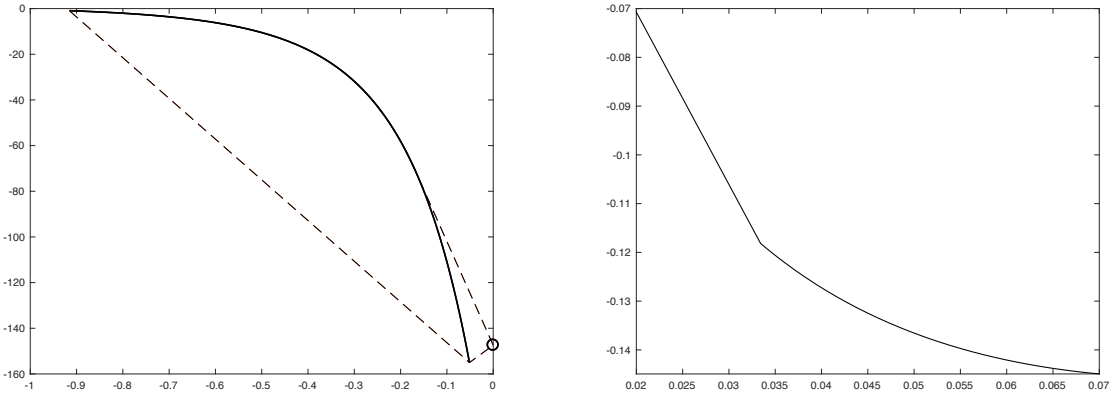


FIG. 2. The action set  $A$  on the left and the corresponding value function  $v = v_A$  in (4.7) for the insurance example on the right. Parameter values are  $\alpha = 0.08$ ,  $f = 10$ ,  $\varpi = 1000$ ,  $R = 10$ .

377 **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],  
 378 we are interested in the so-called *marginal value of information*:

379 (5.1) 
$$V^+ = \limsup_{\theta \rightarrow 0} \frac{1}{\theta} \text{VoI}_A(\mathbf{q}^\theta).$$

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

381 PROPOSITION 5.1. *Assume that*

- 382 • either  $\mathbb{E}[d(\mathbf{q}^\theta, \Delta_A^c(\bar{p}))] = o(\theta)$ ,  
 383 • or the decision maker is flexible at prior belief  $\bar{p}$  and  $\mathbb{E}\|\mathbf{q}^\theta - \bar{p}\|^2 = o(\theta)$ .

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

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

387 We now discuss how our approach in Proposition 5.1 compares with the literature. In [32], one finds joint  
 388 conditions on the parameterized information structure  $(\mathbf{q}^\theta)_{\theta>0}$  and the decision problem at hand, leading to  
 389  $V^+ = 0$ . The second case in Proposition 5.1, when the decision maker is flexible, compares with the original  
 390 Radner-Stiglitz assumptions for the smoothness part, but not for the uniqueness of optimal actions. Indeed,  
 391 Assumption (A0) in [32] does not require that  $A^*(\mathbf{q}^\theta)$  be a singleton, for all  $\theta$ .

392 The authors of [15] make a step towards disentangling conditions on the parameterized information  
 393 structure  $(\mathbf{q}^\theta)_{\theta>0}$  from conditions on the decision problem that lead to a null marginal value of information.  
 394 However, like [32], they make an assumption on how the optimal action varies with information, which makes  
 395 the comparison with Proposition 5.1 delicate. In addition, [15] provide sufficient conditions for  $V^+ = 0$  that  
 396 bear on the conditional distribution of the signal knowing the state of nature. Our approach focuses on the  
 397 posterior conditional distribution of the state of nature knowing the signal.

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

402 **5.2. Examples.** Here, we study the marginal value of information for several typical parameterized  
 403 information structures. In the first example, information consists on the observation of a Brownian motion  
 404 with known variance and a drift that depends on the state of nature. In the second example, information  
 405 consists of the observation of a Poisson process whose probability of success depends on the state of nature.  
 406 In these two well studied families in the learning literature, the natural parameterization of information is  
 407 the length of the interval of time during which observation takes place. In the third example, the agent

408 observes a binary signal and the marginal value of information depends on the asymptotic informativeness  
 409 of these signals close to the situation without information.

410 In all three following examples we assume binary states of nature,  $K = \{0, 1\}$ , and (by a slight abuse  
 411 of notation) the prior belief on the state being 1 is denoted  $\bar{p} \in ]0, 1[$ . We follow the conditions in Sect. 3  
 412 under which we established bounds on the value of information, and label as: “confident” the case in which  $\bar{p}$   
 413 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  $[p_l, p_h]$  by  
 414 Proposition A.3, and the value function is linear on this range); “undecided” the case in which the decision  
 415 problem faced by the decision maker is such that there is indifference between two actions at prior belief  $\bar{p}$ ;  
 416 “flexible” the case in which the optimal action is a smooth function of the belief in a neighborhood of prior  
 417 belief  $\bar{p}$ .

418 Our aim is to develop estimates of the marginal value of information  $V^+$  in (5.1). There are three  
 419 possibilities: it can be infinite, null, or positive and finite. We denote these three cases by  $V^+ = \infty$ ,  $V^+ = 0$   
 420 and  $V^+ \simeq 1$  respectively.

421 **EXAMPLE 3** (Brownian motion). *Frameworks in which agents observe a Brownian motion with known*  
 422 *volatility and unknown drift include [5, 24, 10], as well as reputation models like [19].*

423 *Assume the agent observes the realization of a Brownian motion with variance 1 and drift  $k \in \{0, 1\}$ ,*  
 424 *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*  
 425 *is well-known<sup>2</sup> 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*  
 426 *Browian process. Thus, for small values of  $\theta$ , we have the estimates*

427 
$$\mathbb{E} \|\mathbf{q}^\theta - p\| \sim \sqrt{\theta}, \quad \mathbb{E} \|\mathbf{q}^\theta - p\|^2 \sim \theta.$$

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

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

433 **EXAMPLE 4** (Poisson learning). *An important class of models of strategic experimentation (see [25])*  
 434 *are those in which the agent’s observations are driven by a Poisson process of unknown intensity. Assume*  
 435 *the agent observes, during a small interval of time  $\theta > 0$ , a Poisson process with intensity  $\rho_k$ ,  $k \in \{0, 1\}$ ,*  
 436 *where  $\rho_1 > \rho_0 > 0$ . The probability of two successes is negligible compared to the probability of one success*  
 437 *(of order  $\theta^2$  compared to  $\theta$ ). A success leads to a posterior that converges from below, as  $\theta \rightarrow 0$ , to*

438 
$$q^+ = \frac{\bar{p}\rho_1}{\bar{p}\rho_1 + (1 - \bar{p})\rho_0} > \bar{p},$$

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

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

445 *We also have the estimates*

446 
$$\mathbb{E} \|\mathbf{q}^\theta - p\| \sim \theta, \quad \mathbb{E} \|\mathbf{q}^\theta - p\|^2 \sim \theta,$$

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

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

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

<sup>2</sup>See for instance Lemma 1 in [10] or Lemma 2 in [19].

452 the posterior beliefs are  $p \pm \theta^\alpha$  for a certain parameter  $\alpha > 0$  (lower values of  $\alpha$  correspond to more spread  
453 out beliefs around the prior, hence to more accurate information). In this case we easily compute

454 
$$\mathbb{E} \|\mathbf{q}^\theta - p\| = \theta^\alpha, \quad \mathbb{E} \|\mathbf{q}^\theta - p\|^2 = \theta^{2\alpha},$$

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

- 457 1. In the confident case,  $V^+ = 0$ ,  
458 2. In the undecided case,  
459 (a)  $V^+ = \infty$  if  $\alpha < 1$ ,  
460 (b)  $V^+ \simeq 1$  if  $\alpha = 1$ ,  
461 (c)  $V^+ = 0$  if  $\alpha > 1$ ,  
462 3. In the flexible case,  
463 (a)  $V^+ = \infty$  if  $\alpha < 1/2$ ,  
464 (b)  $V^+ \simeq 1$  if  $\alpha = 1/2$ ,  
465 (c)  $V^+ = 0$  if  $\alpha > 1/2$ .

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

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

467 In all cases except one, the marginal value of information is completely determined by the local behavior  
468 of the value function around the prior. For the Poisson case, the marginal value of information is 0 or  
469 positive, depending on whether the observation of a success is sufficient to lead to a decision reversal.

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

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

484 The requirement that all agents agree on their preferences between two statistical experiments is a strong  
485 one. It implies that this ranking is incomplete, as many such pairs of experiments cannot be ranked according  
486 to this ordering. Some authors have considered subclasses of decision problems in order to obtain rankings  
487 that are more complete than Blackwell’s. For instance, [26], [31] and [2] restrict attention to families of  
488 decision problems that generate monotone decision rules. Focusing on investment decision problems, [12]

489 obtains and characterizes a complete ranking of information sources based on a uniform criterion; [13] uses  
490 a duality approach to characterize the value of an information purchase that consists of an information  
491 structure with a price attached to it.

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

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

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

507

#### REFERENCES

- 508 [1] K. J. ARROW, *The value of and demand for information*, in Decision and Organization, C. McGuire and R. Radner, eds.,  
509 Amsterdam, 1971, North-Holland, pp. 131–139.
- 510 [2] S. ATHEY AND J. LEVIN, *The value of information in monotone decision problems*, Research in Economics, (2017).
- 511 [3] R. AUMANN AND M. MASCHLER, *Repeated games with incomplete information: A survey of recent results*, in Reports to  
512 the U.S. Arms Control and Disarmament Agency, ST-116, 1967, pp. 287–403.
- 513 [4] Y. AZRIELI AND E. LEHRER, *The value of a stochastic information structure*, Games and Economic Behavior, 63 (2008),  
514 pp. 679–693.
- 515 [5] D. BERGEMANN AND J. VÄLIMÄKI, *Market diffusion with two-sided learning*, The RAND Journal of Economics, (1997),  
516 pp. 773–795.
- 517 [6] D. BERNOULLI, *Specimen theoriae novae de mensura sortis*, in Commentarii Academiae Scientiarum Imperialis Petropoli-  
518 tanae, vol. 5, 1738, pp. 175–192.
- 519 [7] D. BLACKWELL, *Comparison of experiments*, in Proceedings of the Second Berkeley Symposium on Mathematical Statistics  
520 and Probability, University of California Press, 1951, pp. 93–102.
- 521 [8] D. BLACKWELL, *Equivalent comparison of experiments*, Annals of Mathematical Statistics, 24 (1953), pp. 265–272.
- 522 [9] H. F. BOHNEBLUST, L. S. SHAPLEY, AND S. SHERMAN, *Reconnaissance in game theory*, Tech. Report RM-208, The  
523 RAND corporation, 1949.
- 524 [10] P. BOLTON AND C. HARRIS, *Strategic experimentation*, Econometrica, 67 (1999), pp. 349–374.
- 525 [11] G. W. BRIER, *Verification of forecasts expressed in terms of probability*, Monthly Weather Review, 78 (1950).
- 526 [12] A. CABRALES, O. GOSSNER, AND R. SERRANO, *Entropy and the value of information for investors*, American Economic  
527 Review, 103 (2013), pp. 360–377.
- 528 [13] A. CABRALES, O. GOSSNER, AND R. SERRANO, *A normalized value for information purchases*, Journal of Economic  
529 Theory, 170 (2017), pp. 266–288.
- 530 [14] S. CERREIA-VIOLIO, F. MACCHERONI, M. MARINACCI, AND L. MONTRUCCHIO, *Complete monotone quasiconcave*  
531 *duality*, Mathematics of Operations Research, 36 (2011), pp. 321–339.
- 532 [15] H. CHADE AND E. SHLEE, *Another look at the Radner-Stiglitz nonconcavity in the value of information*, Journal of  
533 Economic Theory, 107 (2002), pp. 421–452.
- 534 [16] M. DE LARA AND L. GILOTTE, *A tight sufficient condition for Radner–Stiglitz nonconcavity in the value of information*,  
535 Journal of Economic Theory, 137 (2007), pp. 696–708.
- 536 [17] D. DENTCHEVA AND A. RUSZCZYŃSKI, *Common mathematical foundations of expected utility and dual utility theories*,  
537 SIAM Journal on Optimization, 23 (2013), pp. 381–405.
- 538 [18] L. ECKHOUDT, C. GOLLIER, AND H. SCHLESINGER, *Economic and Financial Decisions under Risk*, Princeton University  
539 Press, 2005.
- 540 [19] E. FAINGOLD AND Y. SANNIKOV, *Reputation in continuous-time games*, Econometrica, 79 (2011), pp. 773–876.
- 541 [20] I. GILBOA AND E. LEHRER, *The value of information - an axiomatic approach*, Journal of Mathematical Economics, 20  
542 (1991), pp. 443–459.
- 543 [21] I. GILBOA AND D. SCHMEIDLER, *Maxmin expected utility with non-unique prior*, Journal of Mathematical Economics, 18  
544 (1989), pp. 141–153.
- 545 [22] J. B. HIRIART-URURTY AND C. LEMARÉCHAL, *Convex Analysis and Minimization Algorithms I*, Springer-Verlag, Berlin,  
546 1993.



- 547 [23] J. HIRSHLEIFER, *The private and social value of information and the reward to inventive activity*, American Economic  
548 Review, 61 (1971), pp. 561–574.
- 549 [24] G. KELLER AND S. RADY, *Optimal experimentation in a changing environment*, Review of Economic Studies, 66 (1999),  
550 pp. 475–507.
- 551 [25] G. KELLER, S. RADY, AND M. CRIPPS, *Strategic experimentation with exponential bandits*, Econometrica, 73 (2005),  
552 pp. 39–68.
- 553 [26] E. L. LEHMANN, *Comparing location experiments*, The Annals of Statistics, 16 (1988), pp. 521–533.
- 554 [27] F. MACCHERONI, M. MARINACCI, AND A. RUSTICHINI, *Ambiguity aversion, robustness, and the variational representation  
555 of preferences*, Econometrica, 74 (2006), pp. 1447–1498.
- 556 [28] D. MCFADDEN, *Conditional logit analysis of qualitative choice behavior*, in Frontiers in Econometrics, P. Zarembka, ed.,  
557 Academic Press, 1973.
- 558 [29] D. MCFADDEN, *Modelling the choice of residential location*, in Spatial Interaction Theory and Planning Models, A. Kar-  
559 lqvist, L. Lundqvist, F. Snickars, and J. Weibull, eds., North-Holland, 1978.
- 560 [30] L. J. MIRMAN, L. SAMUELSON, AND A. URBANO, *Monopoly experimentation*, International Economic Review, (1993),  
561 pp. 549–563.
- 562 [31] N. PERSICO, *Information acquisition in auctions*, Econometrica, 68 (2000), pp. 135–148.
- 563 [32] R. RADNER AND J. STIGLITZ, *A nonconvexity in the value of information*, in Bayesian Models of Economic Theory,  
564 M. Boyer and R. Kihlstrom, eds., Amsterdam, 1984, Elsevier, pp. 33–52.
- 565 [33] T. R. ROCKAFELLAR, *Convex Analysis*, Princeton University Press, Princeton, N.J., 1970.
- 566 [34] R. SCHNEIDER, *Convex bodies: the Brunn-Minkowski theory*, Cambridge University Press, second ed., 2014.

## 567 Appendix A. Appendix.

568 **A.1. Revisiting the model of Sect. 2.** We revisit the model in Sect. 2 with convex analysis tools to  
569 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  
570 of  $\mathbb{R}^K$ , called the *action set*, and that we identify the set  $\Sigma$  of signed measures on  $K$  with  $\mathbb{R}^K$ .

571 *Support function.* The *support function*  $\sigma_A$  of the action set  $A$  is defined by

$$572 \text{ (A.1)} \quad \sigma_A(s) = \sup_{a \in A} \langle s, a \rangle, \quad \forall s \in \Sigma.$$

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

$$574 \text{ (A.2)} \quad v_A(p) = \sigma_A(p), \quad \forall p \in \Delta.$$

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

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

$$578 \text{ (A.3)} \quad F_A(s) = \arg \max_{a' \in A} \langle s, a' \rangle = \{a \in A \mid \forall a' \in A, \langle s, a' \rangle \leq \langle s, a \rangle\} \subset A$$

579 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  $s \in \Sigma$* .  
580 As the action set  $A$  is convex and compact, the face  $F_A(s)$  of  $A$  in the direction  $s$  is nonempty, for any  $s \in \Sigma$ ,  
581 and the face is a subset of the *boundary*  $\partial A$  of  $A$ :  $F_A(s) \subset \partial A$ ,  $\forall s \in \Sigma$ . We will use the following property:  
582 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

$$583 \text{ (A.4)} \quad \sigma_C(y') - \sigma_C(y) \geq \sigma_{F_C(y)}(y' - y) \geq \langle y' - y, x' \rangle, \quad \forall y' \in \mathbb{R}^K, \forall x' \in C.$$

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

$$586 \text{ (A.5)} \quad A^*(p) = F_A(p), \quad \forall p \in \Delta.$$

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

$$588 \text{ (A.6)} \quad N_A(a) = \{s \in \Sigma \mid \forall a' \in A, \langle s, a' \rangle \leq \langle s, a \rangle\} \subset \Sigma.$$

589 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  
590 measures in  $\Sigma$ , that are not necessarily beliefs. The set  $\Delta_A^*(a)$  of beliefs compatible with optimal action  $a$   
591 in (2.4) is related to the normal cone  $N_A(a)$  at  $a$  in (A.6) by:

$$592 \text{ (A.7)} \quad \Delta_A^*(a) = N_A(a) \cap \Delta, \quad \forall a \in A.$$



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

$$594 \quad (\text{A.8}) \quad s \in \Sigma \text{ and } a \in F_A(s) \iff a \in A \text{ and } s \in N_A(a).$$

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

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

$$599 \quad (\text{A.9}) \quad a \in \text{reg}(A) \iff \exists s \in \Sigma, \quad s \neq 0, \quad N_A(a) = \mathbb{R}_+ s.$$

600 Notice that a regular point  $a$  necessarily belongs to the boundary  $\partial A$  of  $A$ :  $\text{reg}(A) \subset \partial A$ . The body  $A$  is  
 601 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  
 602 its boundary  $\partial A$  is a  $C^1$  submanifold of  $\mathbb{R}^K$  [34, Theorem 2.2.4, p. 83].

603 *Spherical image map of  $A$ .* We denote by  $S^{|K|-1} = \{s \in \Sigma, \quad \|s\| = 1\}$  the unit sphere of the signed  
 604 measures  $\Sigma$  on  $K$  (identified with  $\mathbb{R}^K$  with its canonical scalar product). By (A.9), we have that  $a \in$   
 605  $\text{reg}(A) \iff \exists! s \in S^{|K|-1}, \quad N_A(a) = \mathbb{R}_+ s$ . If a point  $a \in A$  is regular, the unique outer normal unitary  
 606 vector to  $A$  at  $a$  is denoted by  $n_A(a)$ , so that  $N_A(a) = \mathbb{R}_+ n_A(a)$ . The mapping

$$607 \quad (\text{A.10}) \quad n_A : \text{reg}(A) \rightarrow S^{|K|-1}, \quad \text{where } \text{reg}(A) \subset \partial A,$$

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

$$609 \quad (\text{A.11}) \quad a \in \text{reg}(A) \Rightarrow N_A(a) = \mathbb{R}_+ n_A(a) \quad \text{where } n_A(a) \in S^{|K|-1}.$$

610 *Reverse spherical image map of  $A$ .* We say that a unit signed measure  $s \in S^{|K|-1}$  is *regular* [34, p. 87]  
 611 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*  
 612 *regular unit signed measures* is denoted by  $\text{regn}(A)$ :

$$613 \quad (\text{A.12}) \quad s \in \text{regn}(A) \iff s \in S^{|K|-1} \text{ and } \exists! a \in A, \quad F_A(s) = \{a\}.$$

614 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  
 615  $F_A(s) = \{f_A(s)\}$ . The mapping

$$616 \quad (\text{A.13}) \quad f_A : \text{regn}(A) \rightarrow \partial A, \quad \text{where } \text{regn}(A) \subset S^{|K|-1},$$

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

$$618 \quad (\text{A.14}) \quad s \in \text{regn}(A) \Rightarrow F_A(s) = \{f_A(s)\}.$$

619 *Bodies with  $C^2$  surface.*

620 **PROPOSITION A.1** (Schneider 2014, p. 113). *If the body  $A$  has boundary  $\partial A$  which is a  $C^2$  submanifold*  
 621 *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*  
 622 *defined over the whole boundary  $\partial A$  and is of class  $C^1$ , iii) the spherical image map  $n_A$  has the reverse*  
 623 *spherical image map  $f_A$  in (A.10) as right inverse, that is,  $n_A \circ f_A = \text{Id}_{\text{regn}(A)}$ .*

624 *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  
 625  $f_A : \text{regn}(A) \rightarrow \partial A$  by (A.13), and as  $n_A : \partial A \rightarrow S^{|K|-1}$  by (A.10) since  $\text{reg}(A) = \partial A$ , the mapping  
 626  $n_A \circ f_A : \text{regn}(A) \rightarrow S^{|K|-1}$  is well defined. Let  $s \in \text{regn}(A)$ . By (A.14), we have that  $F_A(s) = \{f_A(s)\}$  and  
 627 by (A.11), we have that  $N_A(f_A(s)) = \mathbb{R}_+ n_A(f_A(s))$ . From (A.8) — stating that exposed face and normal  
 628 cone are conjugate — we deduce that  $s \in \mathbb{R}_+ n_A(f_A(s))$ . As  $s \in S^{|K|-1}$ , we conclude that  $s = n_A(f_A(s))$   
 629 by (A.10).  $\square$

630 *Weingarten map.* Let  $a \in \text{reg}(A)$  be a regular point, as in (A.9), such that the spherical image map  $n_A$   
 631 in (A.10) is differentiable at  $a$ , with differential denoted by  $T_a n_A$ . The *Weingarten map* [34, p. 113]  $T_a n_A :$   
 632  $T_a \partial A \rightarrow T_{n_A(a)} S^{|K|-1}$  linearly maps the tangent space  $T_a \partial A$  of the boundary  $\partial A$  at point  $a$  into the tangent  
 633 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  
 634 the *principal curvatures* of  $A$  at  $a$  [34, p. 114]; they are nonnegative [34, p. 115]. By definition, the body  $A$   
 635 has *positive curvature* at  $a$  if all principal curvatures at  $a$  are positive or, equivalently, if the Weingarten  
 636 map is of maximal rank at  $a$  [34, p. 115].

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

$$640 \quad (\text{A.15}) \quad T_s f_A : T_s S^{|K|-1} \rightarrow T_{f_A(s)} \partial A$$

641 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 bound-  
642 ary  $\partial A$  at point  $f_A(s)$ . The eigenvalues of the reverse Weingarten map at  $s$  are called the *principal radii of*  
643 *curvature* of  $A$  at  $s$ .

644 **A.3. Proofs of the results in Sect. 3.** Using the relations (A.5) and (A.7), we express the proofs of  
645 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  $\Sigma$  of signed measures),  
646 instead of  $A^*(p)$  in (2.3) and  $\Delta_A^*(a)$  in (2.4) (in the set  $\Delta$  of probability measures).

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

$$650 \quad (\text{A.16}) \quad \mathbf{VoI}_A(\mathbf{q}) = \mathbb{E} [\sigma_A(\mathbf{q}) - \sigma_A(\bar{p})] .$$

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

$$652 \quad (\text{A.17a}) \quad \varphi_A^+(q) = \sigma_A(q) - \sigma_A(\bar{p}) + \sigma_{-A^*(\bar{p})}(q - \bar{p}) ,$$

$$653 \quad (\text{A.17b}) \quad \varphi_A^-(q) = \sigma_A(q) - \sigma_A(\bar{p}) - \sigma_{A^*(\bar{p})}(q - \bar{p}) .$$

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

$$655 \quad (\text{A.18a}) \quad \mathbb{E} [\varphi_A^+(\mathbf{q})] = \mathbb{E} [\sigma_A(\mathbf{q}) - \sigma_A(\bar{p}) + \sigma_{-A^*(\bar{p})}(\mathbf{q} - \bar{p})]$$

$$656 \quad (\text{A.18b}) \quad \geq \mathbf{VoI}_A(\mathbf{q}) = \mathbb{E} [\sigma_A(\mathbf{q}) - \sigma_A(\bar{p}) - \langle \mathbf{q} - \bar{p}, a \rangle]$$

$$657 \quad (\text{A.18c}) \quad \geq \mathbb{E} [\sigma_A(\mathbf{q}) - \sigma_A(\bar{p}) - \sigma_{A^*(\bar{p})}(\mathbf{q} - \bar{p})] = \mathbb{E} [\varphi_A^-(\mathbf{q})] .$$

658  
659

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

$$661 \quad (\text{A.19a}) \quad \varphi_A^+(q) = \sigma_A(q) - \sigma_A(\bar{p}) + \sigma_{-A^*(\bar{p})}(q - \bar{p})$$

$$662 \quad (\text{A.19b}) \quad = \sup_{a \in A^*(\bar{p})} \left( \sigma_A(q) - \sigma_A(\bar{p}) - \langle q - \bar{p}, a \rangle \right)$$

$$663 \quad (\text{A.19c}) \quad \geq \sigma_A(q) - \sigma_A(\bar{p}) - \langle q - \bar{p}, a \rangle , \quad \forall a \in A^*(\bar{p})$$

$$664 \quad (\text{A.19d}) \quad \geq \inf_{a \in A^*(\bar{p})} \left( \sigma_A(q) - \sigma_A(\bar{p}) - \langle q - \bar{p}, a \rangle \right)$$

$$665 \quad (\text{A.19e}) \quad = \sigma_A(q) - \sigma_A(\bar{p}) - \sigma_{A^*(\bar{p})}(q - \bar{p}) = \varphi_A^-(q) .$$

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

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

669 **PROPOSITION A.3.**

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

$$671 \quad (\text{A.20}) \quad \Delta_A^c(\bar{p}) = \bigcap_{a \in A^*(\bar{p})} \Delta_A^*(a) = \bigcap_{a \in F_A(\bar{p})} N_A(a) \cap \Delta .$$

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

$$673 \quad (\text{A.21a}) \quad p \in \Delta_A^c(\bar{p}) \iff F_A(\bar{p}) \subset F_A(p)$$

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

$$(A.21c) \quad \iff \sigma_A(p) - \sigma_A(\bar{p}) + \sigma_{-A^*(p)}(p - \bar{p}) = 0.$$

679

680 3. The indifference kernel  $\Sigma_A^i(\bar{p})$  of (3.5) is the vector subspace

$$681 \quad \Sigma_A^i(\bar{p}) = [F_A(\bar{p}) - F_A(\bar{p})]^\perp = [A^*(\bar{p}) - A^*(\bar{p})]^\perp = \bigcap_{a \in F_A(\bar{p})} N_{F_A(\bar{p})}(a).$$

682 *Proof.*

683 1. Express (3.1) using (A.7).

684 2. We prove the three equivalences in (A.21).

685 (a) Let  $p \in \Delta$ . Using the property (A.8) that exposed face  $F_A$  and normal cone  $N_A$  are conjugate,

$$686 \quad \text{we obtain: } p \in \Delta_A^c(\bar{p}) \iff p \in \bigcap_{a \in F_A(p)} N_A(a) \text{ by (A.20)}$$

$$687 \quad \iff a \in F_A(p), \quad \forall a \in F_A(\bar{p}) \text{ by (A.8)} \iff F_A(\bar{p}) \subset F_A(p).$$

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

$$690 \quad \sigma_A(p) - \sigma_A(\bar{p}) - \langle p - \bar{p}, a \rangle = 0, \quad \forall a \in F_A(\bar{p})$$

$$691 \quad \iff \sigma_A(p) = \langle p, a \rangle, \quad \forall a \in F_A(\bar{p})$$

692 because  $\sigma_A(\bar{p}) = \langle \bar{p}, a \rangle$  for any  $a \in F_A(\bar{p})$ , since  $F_A(\bar{p})$  is the set  $A^*(\bar{p})$  of optimal actions under prior belief  $\bar{p}$  by (2.3) and (A.3)

$$694 \quad (\text{by definition (A.6) of } N_A(a)) \quad \iff p \in \bigcap_{a \in F_A(\bar{p})} N_A(a)$$

$$695 \quad \iff p \in \bigcap_{a \in F_A(\bar{p})} N_A(a) \cap \Delta = \Delta_A^c(\bar{p}) \text{ by (A.20).}$$

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

$$698 \quad (A.22) \quad \varphi_a(q) = \sigma_A(q) - \sigma_A(\bar{p}) - \langle q - \bar{p}, a \rangle, \quad \forall q \in \Delta.$$

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

$$700 \quad (A.23a) \quad \forall a \in F_A(\bar{p}), \quad \forall q \in \Delta, \quad \varphi_a(q) \geq 0,$$

$$701 \quad (A.23b) \quad \forall a \in F_A(\bar{p}), \quad \forall q \in \Delta_A^c(\bar{p}), \quad \varphi_a(q) = 0.$$

702 Let  $p \in \Delta$ . Using (A.23a), we deduce from (A.21b) and from the compacity of  $F_A(\bar{p})$  that  
 704  $p \in \Delta_A^c(\bar{p}) \iff \inf_{a \in F_A(\bar{p})} (\sigma_A(p) - \sigma_A(\bar{p}) - \langle p - \bar{p}, a \rangle) = 0$ . We conclude with (A.19d)–  
 705 (A.19e).

706 3. Express (3.5) using (A.5). Then, use the definition of  $N_{F_A(\bar{p})}(a)$  in (A.6).

707 This ends the proof.  $\square$

### 708 A.3.1. Valuable information.

709 *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

$$710 \quad \mathbf{VoI}_A(\mathbf{q}) = 0 \iff \mathbb{E} [\sigma_A(\mathbf{q}) - \sigma_A(\bar{p})] = 0 \text{ by (A.16)}$$

$$711 \quad \iff \mathbb{E} [\sigma_A(\mathbf{q}) - \sigma_A(\bar{p}) - \langle \mathbf{q} - \bar{p}, a \rangle] = 0, \quad \text{as } \mathbb{E} [\mathbf{q} - \bar{p}] = 0$$

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

$$712 \quad \iff \sigma_A(\mathbf{q}) - \sigma_A(\bar{p}) - \langle \mathbf{q} - \bar{p}, a \rangle = 0, \quad \mathbb{P} - \text{a.s.}$$

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

717 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  
718 equality that  $\mathbf{VoI}_A(\mathbf{q}) = 0 \Rightarrow \mathbb{P}\{\langle \mathbf{q}, a' - a \rangle \leq 0, \forall a' \in A, \forall a \in F\} = 1$ . As the set  $\{a \in F_A(\bar{p}) \mid$   
719  $\langle \mathbf{q}, a' - a \rangle \leq 0, \forall a' \in A\}$  is closed (for any outcome in the underlying sample space  $\Omega$ ), we get that  
720  $\{\langle \mathbf{q}, a' - a \rangle \leq 0, \forall a' \in A, \forall a \in F\} \subset \{\langle \mathbf{q}, a' - a \rangle \leq 0, \forall a' \in A, \forall a \in \bar{F}\}$ . We deduce from the last  
721 equality that  $\mathbf{VoI}_A(\mathbf{q}) = 0 \Rightarrow \mathbb{P}\{\langle \mathbf{q}, a' - a \rangle \leq 0, \forall a' \in A, \forall a \in \bar{F}\} = 1$ . Now, since  $\bar{F} = F_A(\bar{p})$ , we  
722 finally get that  $\mathbf{VoI}_A(\mathbf{q}) = 0 \Rightarrow \mathbb{P}\{\langle \mathbf{q}, a' - a \rangle \leq 0, \forall a' \in A, \forall a \in F_A(\bar{p})\} = 1$ . In other words, we have  
723 obtained that, by definition (A.6) of the normal cone  $N_A(a)$ :  $\mathbf{VoI}_A(\mathbf{q}) = 0 \Rightarrow \mathbf{q} \in \bigcap_{a \in F_A(\bar{p})} N_A(a)$ ,  $\mathbb{P} - \text{a.s.}$   
724 Since  $\mathbf{q} \in \Delta$ , we conclude by (A.20) that

$$725 \quad \mathbf{VoI}_A(\mathbf{q}) = 0 \Rightarrow \mathbf{q} \in \bigcap_{a \in F_A(\bar{p})} N_A(a) \cap \Delta = \bigcap_{a \in A^*(\bar{p})} \Delta_A^*(a) = \Delta_A^c(\bar{p}).$$

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

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

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

$$731 \quad (\text{A.26}) \quad \varphi_a(q) \leq \sup_{a' \in A} \|a - a'\| \inf_{p \in \Delta_A^c(\bar{p})} \|p - q\|.$$

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

$$\begin{aligned}
733 \quad & \varphi_a(q) = \varphi_a(q) - \varphi_a(p) \text{ by (A.23b) since } p \in \Delta_A^c(\bar{p}) \\
734 \quad & = \sigma_A(q) - \sigma_A(p) - \langle q - p, a \rangle \text{ by (A.22)} \\
735 \quad & = \sigma_{A-a}(q) - \sigma_{A-a}(p) \text{ by (A.1)} \\
736 \quad & \leq \sup_{a' \in A-a} \|a'\| \times \|p - q\| \text{ by (A.1)} = \sup_{a' \in A} \|a - a'\| \times \|p - q\|. \\
737 \quad &
\end{aligned}$$

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

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

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

744 Second, we show the lower estimate  $\mathbf{VoI}_A(\mathbf{q}) \geq c_{\bar{p}, A, \varepsilon} \mathbb{P}\{\mathbf{q} \notin \Delta_{A, \varepsilon}^c(\bar{p})\}$  in (3.4). We consider an open  
745 subset  $\mathcal{Q}$  of  $\Delta$  that contains the confidence set  $\Delta_A^c(\bar{p})$ , that is,  $\Delta_A^c(\bar{p}) \subset \mathcal{Q}$ . By Lemma A.4 right below,  
746 there exists an  $a \in F_A(\bar{p})$  such that the continuous function  $\varphi_a$  in (A.22) is strictly positive on  $\Delta_A^c(\bar{p})^c$ . As  
747  $\mathcal{Q}^c \subset \Delta_A^c(\bar{p})^c$  and  $\mathcal{Q}^c$  is a closed subset of the compact  $\Delta$ , we can define  $c_{\bar{p}, A} = \inf_{p \notin \mathcal{Q}} \varphi_a(p) > 0$ . We deduce  
748 that

$$\begin{aligned}
749 \quad & \mathbf{VoI}_A(\mathbf{q}) = \mathbb{E} [\varphi_a(\mathbf{q})] \text{ by (A.18b)} \\
750 \quad & = \mathbb{E} \left[ \mathbf{1}_{\mathbf{q} \in \Delta_A^c(\bar{p})} \varphi_a(\mathbf{q}) + \mathbf{1}_{\mathbf{q} \notin \Delta_A^c(\bar{p})} \varphi_a(\mathbf{q}) \right] \\
751 \quad & = \mathbb{E} \left[ \mathbf{1}_{\mathbf{q} \notin \Delta_A^c(\bar{p})} \varphi_a(\mathbf{q}) \right] \text{ by (A.23b)} \\
752 \quad & \geq \mathbb{E} [\mathbf{1}_{\mathbf{q} \notin \mathcal{Q}} \varphi_a(\mathbf{q})] \geq \mathbb{E} [\mathbf{1}_{\mathbf{q} \notin \mathcal{Q}} c_{\bar{p}, A}] = c_{\bar{p}, A} \mathbb{P}\{\mathbf{q} \notin \mathcal{Q}\}. \\
753 \quad &
\end{aligned}$$

754 With  $\mathcal{Q} = \Delta_{A, \varepsilon}^c(\bar{p})$ , we put  $c_{\bar{p}, A, \varepsilon} = \inf_{p \notin \Delta_{A, \varepsilon}^c(\bar{p})} \varphi_a(p) > 0$ .

755 This ends the proof.  $\square$

756 LEMMA A.4. *There exists at least one  $a \in F_A(\bar{p})$  such that the function  $\varphi_a$  in (A.22) is strictly positive*  
 757 *on the complementary set  $\Delta_A^c(\bar{p})^c$ .*

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

759 Suppose that  $F_A(\bar{p})$  is a singleton  $\{a\}$ . By (A.21b), we have that  $q \notin \Delta_A^c(\bar{p}) \iff \varphi_a(q) > 0$ .

760 Suppose that  $F_A(\bar{p})$  is not a singleton. Recall that the *affine hull*  $\text{aff}(S)$  of a subset  $S$  of  $\mathbb{R}^K$  is the  
 761 intersection of all affine manifolds containing  $S$ , and that the *relative interior*  $\text{ri}(C)$  of a nonempty convex  
 762 set  $C \subset \mathbb{R}^K$  is the nonempty interior of  $C$  for the topology relative to the affine hull  $\text{aff}(C)$  [22, p. 103].  
 763 We prove that any  $a \in \text{ri}(F_A(q))$  answers the question. Let  $a \in \text{ri}(F_A(q))$  be fixed. For any  $q \notin \Delta_A^c(\bar{p})$ ,  
 764 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  
 765 that  $\sigma_A(q) > \langle q, \bar{a} \rangle$ . As  $a \in \text{ri}(F_A(q))$ , there exists  $a' \in \text{ri}(F_A(q))$  such that  $a = \lambda a' + (1 - \lambda)\bar{a}$  for a certain  
 766  $\lambda \in ]0, 1[$ . Since  $\sigma_A(q) \geq \langle q, a' \rangle$  (by definition (A.1) of  $\sigma_A$ ) and  $\sigma_A(q) > \langle q, \bar{a} \rangle$  (as  $\bar{a} \notin F_A(q)$ ), we deduce  
 767 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  
 768  $\lambda \in ]0, 1[$ . Using the definition (A.22) of the function  $\varphi_a$ , we have obtained that  $q \notin \Delta_A^c(\bar{p}) \Rightarrow \varphi_a(q) > 0$ .

769 This ends the proof.  $\square$

### 770 A.3.2. Undecided.

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

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

773 By definition (A.1) of a support function, we have that  $\sigma_A(\cdot) \leq \|A\| \times \|\cdot\|$ , where  $\|A\| = \sup\{\|a\|, a \in$   
 774  $A\} < +\infty$ . Thus  $C_A = \|A\|$  in the left hand side inequality in (3.6).

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

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

$$\begin{aligned} 777 \text{(A.30a)} \quad \sigma_A(s) - \sigma_A(\bar{p}) &\geq \sigma_{F_A(\bar{p})}(s - \bar{p}) \text{ by (A.4) since } F_A(\bar{p}) \neq \emptyset \\ 778 \text{(A.30b)} \quad &= \langle s - \bar{p}, a \rangle, \quad \forall a \in F_A(\bar{p}) \text{ by definition of } \sigma_{F_A(\bar{p})} \\ 779 \text{(A.30c)} \quad &= \sigma_{F_A(\bar{p})}(s) - \sigma_{F_A(\bar{p})}(\bar{p}) \text{ by definition of } \sigma_{F_A(\bar{p})}. \end{aligned}$$

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

$$\begin{aligned} 782 \text{(A.31a)} \quad \mathbf{VoI}_A(\mathbf{q}) &= \mathbb{E} [\sigma_A(\mathbf{q}) - \sigma_A(\bar{p})] \text{ by (2.6) and (A.2)} \\ 783 \text{(A.31b)} \quad &\geq \mathbb{E} [\sigma_{F_A(\bar{p})}(\mathbf{q} - \bar{p})] \text{ by (A.30a)} \\ 784 \text{(A.31c)} \quad &= \mathbb{E} [\sigma_{F_A(\bar{p})}(\mathbf{q}) - \sigma_{F_A(\bar{p})}(\bar{p})] \text{ by (A.30c)} \\ 785 &= \mathbf{VoI}_{F_A(\bar{p})}(\mathbf{q}) \text{ by (2.6) and (A.2)}. \end{aligned}$$

787 This ends the proof of the middle inequality.

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

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

$$791 \text{(A.32)} \quad T = \{a_1, \dots, a_n\} \subset F_A(\bar{p}) \text{ so that } \text{aff}(F_A(\bar{p})) = \text{aff}\{a_1, \dots, a_n\} = \text{aff}(T).$$

792 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  
 793 right hand side inequality in (3.6).

794 First, the support function  $\sigma_{T-T}$  is a seminorm with kernel  $(T - T)^\perp$ , as easily seen. Now, we also  
 795 easily see that, for any subset  $S \subset \mathbb{R}^K$ , one has  $(S - S)^\perp = (\text{aff}(S - S))^\perp = (\text{aff}(S) - \text{aff}(S))^\perp$ . Using  
 796 these equalities with  $S = T$  and  $S = F_A(\bar{p})$ , we deduce that  $(T - T)^\perp = (F_A(\bar{p}) - F_A(\bar{p}))^\perp$ , since  $\text{aff}(T) =$   
 797  $\text{aff}(F_A(\bar{p}))$  by (A.32). Second, we show that the right hand side inequality in (3.6) is satisfied. We have

$$\begin{aligned} 798 \quad \mathbf{VoI}_A(\mathbf{q}) &\geq \mathbb{E} [\sigma_{F_A(\bar{p})}(\mathbf{q} - \bar{p})] \text{ by (A.31b)} \\ &\text{(because } T \subset F_A(\bar{p}) \text{ and support functions (A.1) are monotone with respect to set inclusion)} \\ 799 \quad &\geq \mathbb{E} [\sigma_T(\mathbf{q} - \bar{p})] \end{aligned}$$

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

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  $\mathbf{Vol}_A(\mathbf{q}) \geq \frac{1}{n} \sum_{i=1}^n \mathbb{E} [\sigma_{T-a_i}(\mathbf{q} - \bar{p})] = \frac{1}{n} \mathbb{E} [\sigma_{\sum_{i=1}^n (T-a_i)}(\mathbf{q} - \bar{p})]$ . Now, as  $T = \{a_1, \dots, a_n\}$ , it is easy to see that the sum  $\sum_{i=1}^n (T - a_i)$  contains any element of the form  $a_k - a_l = (a_1 - a_1) + \dots + (a_{l-1} - a_{l-1}) + (a_k - a_l) + (a_{l+1} - a_{l+1}) + \dots + (a_n - a_n) \in \sum_{i=1}^n (T - a_i)$ . As support functions are monotone with respect to set inclusion, we deduce that  $\sigma_{\sum_{i=1}^n (T-a_i)} \geq \sigma_{\{a_k - a_l, k, l=1, \dots, n\}} = \sigma_{T-T}$  and that  $\mathbf{Vol}_A(\mathbf{q}) \geq \frac{1}{n} \mathbb{E} [\sigma_{\{a_k - a_l, k, l=1, \dots, n\}}(\mathbf{q} - \bar{p})] = \frac{1}{n} \mathbb{E} [\sigma_{T-T}(\mathbf{q} - \bar{p})] = \mathbb{E} \|\mathbf{q} - \bar{p}\|_{\Sigma_A^i(\bar{p})}$ . This ends the proof.  $\square$

### 810 A.3.3. Flexible.

811 *Proof of Proposition 3.5.* All the reminders on geometric convex analysis in Sect. A.2 were done with  
812 outer normal vectors belonging to the unit sphere of signed measures. Now, as we work with beliefs —  
813 positive measures of mass 1 — we are going to adapt these concepts. We consider the diffeomorphism

$$814 \quad (\text{A.34}) \quad \nu : S^{|K|-1} \cap \mathbb{R}_+^K \rightarrow \Delta, \quad s \mapsto \frac{s}{\langle s, \mathbf{1} \rangle},$$

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

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

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

$$820 \quad (\text{A.35}) \quad a \in \text{reg}^+(A) \iff \exists p \in \Delta, \quad N_A(a) = \mathbb{R}_+ p.$$

821 For a regular action  $a \in \text{reg}^+(A)$ , there is only one probability  $p \in \Delta$  such that  $N_A(a) = \mathbb{R}_+ p$ , and it is  
822  $p = \nu(n_A(a))$ . We have  $a \in \text{reg}^+(A) \Rightarrow N_A(a) = \mathbb{R}_+ \nu(n_A(a))$  where  $\nu(n_A(a)) \in \Delta$ . The *set of regular*  
823 *probabilities* is  $\text{regn}^+(A) = (\mathbb{R}_+^* \text{regn}(A)) \cap \Delta$ . For a regular probability  $p \in \text{regn}^+(A)$ , there is only one  
824 action  $a \in \partial A$  such that  $F_A(p) = \{a\}$ , and it is  $a = f_A(\nu^{-1}(p))$ . Indeed, by definition (A.3) of the (exposed)  
825 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

$$826 \quad (\text{A.36}) \quad p \in \text{regn}^+(A) \Rightarrow F_A(p) = \{f_A(\nu^{-1}(p))\}.$$

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

829 • Item 2  $\Rightarrow$  Item 1.

830 Suppose that the face  $F_A(\bar{p})$  is a singleton  $\{a^\sharp\}$  and the curvature of the boundary  $\partial A$  of payoffs  
831 at  $a^\sharp$  is positive. Since, by assumption, the action set  $A$  has boundary  $\partial A$  which is a  $C^2$  submanifold  
832 of  $\mathbb{R}^K$ , we know that the spherical image map  $n_A$  in (A.10) is defined over the whole boundary  $\partial A$   
833 and is of class  $C^1$ , and its differential is the Weingarten map. As the curvature of the boundary  $\partial A$  of  
834 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,  
835 by the inverse function theorem, there exists an open neighborhood  $\mathcal{A}$  of  $a^\sharp$  in  $A$  such that  $n_A(\mathcal{A})$   
836 is an open neighborhood of  $n_A(a^\sharp)$  in  $S^{|K|-1}$ , and such that the restriction  $n_A : \mathcal{A} \rightarrow n_A(\mathcal{A})$  of the  
837 spherical image map in (A.10) is a diffeomorphism. By item iii) in Proposition A.1, we have that  
838  $n_A(a^\sharp) = \frac{\bar{p}}{\|\bar{p}\|}$  and the local inverse coincides with the restriction  $f_A : n_A(\mathcal{A}) \rightarrow \mathcal{A}$  of the reverse  
839 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  
840 prior belief  $\bar{p}$  has full support, we deduce that  $\nu(n_A(\mathcal{A}))$  is an open neighborhood of  $\bar{p}$  in  $\Delta$ , where  
841 the diffeomorphism  $\nu$  is defined in (A.34). We easily deduce that  $f_A \circ \nu^{-1} : \nu(n_A(\mathcal{A})) \rightarrow \mathcal{A}$  is a  
842 diffeomorphism. By (A.36), we conclude that  $f_A \circ \nu^{-1}$  is the restriction of the set-valued mapping  
843  $F_A : \Delta \rightrightarrows A$ ,  $p \mapsto F_A(p)$  in (3.7).



844

- Item 1  $\Rightarrow$  Item 3.

845

846

847

848

849

850

851

852

853

854

- Item 3  $\Rightarrow$  Item 2.

855

856

857

858

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  $\Pi$  of  $\bar{p}$  in  $\Delta$  such that  $\Pi \subset \text{regn}^+(A)$ , where the set of relevant regular points is defined in (A.35). In addition, the mapping  $f_A \circ \nu^{-1} : \Pi \rightarrow f_A(\nu^{-1}(\Pi))$  is a diffeomorphism.

As  $F_A(p) = \{f_A(\nu^{-1}(p))\}$ , for all beliefs  $p \in \Pi$ , we know that the support function  $\sigma_A$  is differentiable and that its gradient is  $\nabla_p \sigma_A = f_A(\nu^{-1}(p))$  [22, p. 251]. As  $f_A \circ \nu^{-1}$  is a local diffeomorphism at  $\bar{p}$ , and as the mapping  $\nu$  in (A.34) is a diffeomorphism, we deduce that the support function  $\sigma_A$  is twice differentiable with Hessian having full rank. As the value function  $v_A$  is the restriction of  $\sigma_A$  to  $\Delta$ , we conclude that  $v_A$  is twice differentiable at  $\bar{p}$  and the Hessian is positive definite.

854

855

856

857

858

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  $\Pi$  of  $\bar{p}$  in  $\Delta$  such that  $v_A$  is differentiable on  $\Pi$ . On the other hand, as the support function  $\sigma_A$  is positively homogeneous, and by (A.2), we have that

859

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

860

861

862

863

864

865

866

867

868

869

870

871

872

873

874

875

876

877

Therefore, as the mapping  $\nu$  in (A.34) is a diffeomorphism, the support function  $\sigma_A$  is differentiable on the open neighborhood  $\nu^{-1}(\Pi)$  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  $\sigma_A$  [22, p. 258], we conclude that the face  $F_A(s)$  of  $A$  in the direction  $s \in \nu^{-1}(\Pi)$  is a singleton.

Therefore, by definition (A.12) of the set of regular unit signed measures, we have that  $\nu^{-1}(\Pi) \subset \text{regn}(A)$ . In addition, the restriction  $f_A : \nu^{-1}(\Pi) \rightarrow f_A(\nu^{-1}(\Pi))$  of the reverse spherical image map in (A.13) is well defined, and we have that  $\nabla_s \sigma_A = f_A(s)$ ,  $\forall s \in \nu^{-1}(\Pi)$ . Therefore, the mapping  $f_A : \nu^{-1}(\Pi) \rightarrow f_A(\nu^{-1}(\Pi))$  is differentiable at  $\nu^{-1}(\bar{p}) = \frac{\bar{p}}{\|\bar{p}\|}$ , and has full rank. Indeed,  $\sigma_A$  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  $\nu$  in (A.34) is a  $C^\infty$  diffeomorphism and the value function  $v_A$  is twice differentiable at  $\bar{p}$  with positive definite Hessian.

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

This ends the proof.  $\square$

878

879

*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  $F_A(\bar{p}) = \{a^\sharp\}$ .

880

881

882

883

884

885

886

887

888

889

890

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  $\Delta$ . Indeed,  $g$  is continuous on  $\Delta \setminus \{\bar{p}\}$ , and also at  $\bar{p}$  since the value function  $v_A$  is twice differentiable at  $\bar{p}$ . In addition,  $g(\bar{p}) > 0$  since the Hessian of  $v_A$  at  $\bar{p}$  is positive definite. We have  $g \geq 0$  on  $\Delta \setminus \{\bar{p}\}$ , because  $F_A(\bar{p}) = \{a^\sharp\}$  is the subdifferential at  $\bar{p}$  of the support function  $\sigma_A$ , 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  $\Delta$ , we easily deduce (3.8) from (2.6).

This ends the proof.  $\square$