HAL open science

Evaluating information in zero-sum games with incomplete information on both sides

Bernard de Meyer, Ehud Lehrer, Dinah Rosenberg

To cite this version:

Bernard de Meyer, Ehud Lehrer, Dinah Rosenberg. Evaluating information in zero-sum games with incomplete information on both sides. 2009. halshs-00390625

HAL Id: halshs-00390625

https://shs.hal.science/halshs-00390625v1

Submitted on 2 Jun 2009

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

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

Distributed under a Creative Commons Attribution 4.0 International License

Documents de Travail du
Centre d'Economie de la Sorbonne

CES Working Papers

Evaluating information in zero-sum games with incomplete information on both sides

Bernard DE MEYER, Ehud LEHRER, Dinah ROSENBERG

2009.35

Maison des Sciences Économiques, 106-112 boulevard de L'Hôpital, 75647 Paris Cedex 13
http://ces.univ-paris1.fr/cesdp/CES-docs.htm
ISSN : 1955-611X

Evaluating information in zero-sum games with incomplete information on both sides

Bernard De Meyer[*], Ehud Lehrer[†] and Dinah Rosenberg[‡§]

May 14, 2009

Abstract

In a Bayesian game some players might receive a noisy signal regarding the specific game actually being played before it starts. We study zero-sum games where each player receives a partial information about his own type and no information about that of the other player, and analyze the impact the signals have on the payoffs. It turns out that the functions that evaluate the value of information share two property. The first is Blackwell monotonicity, which means that each player gains from knowing more. The second is concavity on the space of conditional probabilities.

[*]CES-équipe cermsem, Université Paris 1-Panthéon Sorbonne, 112 Boulevard de l'hôpital, 73013 Paris, France; e-mail:demeyer@univ-paris1.fr

[†]School of Mathematical Sciences, Tel Aviv University, Tel Aviv 69978, Israel; e-mail: lehrer@post.tau.ac.il; home page: www.math.tau.ac.il/∼lehrer

[‡]LAGA Institut Galilée, Université Paris 13, avenue Jean Baptiste Clément, 93430 Villetaneuse, France; e-mail: dinah@math.univ-paris13.fr

[§]This paper extends a paper by the last two authors that bears the same title.

[0]*Journal of Economic Literature* classification numbers: C72, C73, D80, D82, D83.

1 Introduction

In strategic interactions some relevant aspects of the environment might be imperfectly known to players. However, in order to maximize their payoffs, players might look for additional information. The impact of different information on the outcome of a strategic interaction is the subject of this paper.

The set-up we chose to analyze this question is of games with incomplete information. Before the game starts each player obtains a noisy signal about the state of nature. The players' signals are determined by an information structure that specifies how they stochastically depend on the state of nature. Upon receiving the signals the players take actions and receive payoffs that depend on the actions taken and on the state of nature.

In general, evaluating the impact of the information structure on the outcome of the game is not an easy task for a few reasons. First, the interpretation of 'outcome' depends on the solution concept applied. Second, for most solution concepts there are typically multiple outcomes (equilibria). Finally, players might get correlated information, which typically has a significant effect on the outcome. These reasons do not exist when dealing with zero-sum games. In these games, there is one natural solution concept, the value, which induces a unique payoff.

There are two main approaches to analyze the connection between information and payoffs. The first is to compare between two information structures and find which is payoff-wise better than the other. This direction has been widely studied in the literature. Blackwell (1953) initiated this direction and compared between information structures in one-player decision problems. He proved that an information structure always yields a higher value of the problem than another one if and only if it is more informative in a certain sense. Blackwell's result was extended to zero-sum games by Gossner and Mertens (2001). It is well known that this property does not extend to non-zero-sum games and various attempts have been made to understand when it might be extended (see for instance, Hirshleifer (1971), Bassan et al. (2003), Kamien and al. (1990), Neyman (1991), Lehrer and al. (2006)).

The second approach is to study the impact of an information structure on the outcome of the interaction. A typical question in this line is whether the outcome

depends in any particular and discernable way on the information. Since any specific game might have its own idiosyncrasies, an insight into the subject can be obtained only by looking at all possible interactions.

For a given game we define the a value-of-information-function which associates any information structure with the value of the corresponding Bayesian game. We study the properties that are common to *all* such functions and thereby, the connection between information and outcomes that is not specific to a particular game.

Gilboa and Lehrer (1991) treated deterministic information structures and characterized those functions that are value-of-information-functions of *one-player* decision problems. Blackwell (1953)'s result implies that monotonicity is a necessary condition but it turns out not to be sufficient. Their result has been extended to random information structures by Azrieli and Lehrer (2004).

Lehrer and Rosenberg (2006) studied the nature of value-of-information-functions in two cases: (a) one-sided deterministic information: the game depends on a state of nature which is partially known only to player 1 (i.e., player 2 gets no information); and (b) symmetric deterministic information: the game depends on a state of nature which is equally known to both players. In case (a) the more refined the partition the higher the value. Lehrer and Rosenberg (2006) showed that *any* function defined over partitions which is increasing with respect to refinements is a value-of-information-function of some game with incomplete information on one side.

This paper extends the discussions to the case where both players get some partial information on the state of nature. We focus on zero-sum games with lack of information on both sides and independent information. The type k of player 1 and the type l of player 2 are drawn independently of each other. Each player obtains no information about the other player's type and just a partial information about his own. This special case of independent types has been extensively studied in the context of repeated games with incomplete information (see Aumann and Maschler (1996), Mertens and Zamir (1971)).

Formally, the information structure is characterized by a pair of probability transitions (μ, ν) where μ is a probability transition between K and the set of messages of player 1 and ν is a probability transition between L and the set of messages of player 2. Each player is informed only of his message which therefore endows him with some partial information on k for player 1 and l for player 2. Then, a zero-sum game whose payoffs depend on the players' types is played. The value of this game depends on

the information structure, (μ, ν). The function that associates to every information structure the corresponding value is called the value-of-information-function of this game. The goal of this paper is to characterize those functions that are a value-of-information-function of some game.

The result concerning games with one-sided information is certainly relevant to the current two-sided information case. In the one-sided and deterministic information case it is proved that a function is a value of information function if and only it is increasing when the information of the informed agent is refined. When the lack of information is on both sides, we show that any value of information function has to be increasing when player 1's information gets refined (and decreasing when the information of player 2 is refined). The notion of refining information has two meanings: monotonicity related to Blackwell's partial order over information structures and concavity over the space of conditional probabilities. That these conditions are necessary is a consequence of known results. Our contribution is to prove that they are sufficient.

The implication of this result is that essentially no further condition beyond monotonicity with respect to information is required to characterize the value-of-information functions. This means that the model of Bayesian game with varying information structures can be refuted only by observations that contradict monotonicity.

The center of the proof is the duality between payoffs and information. This means that giving more information to a player amounts to giving him more payoffs in some sense (see also Gossner (2006)). The proof method uses duality technique inspired by Fenchel conjugate of convex functions (see, Rockafellar 1970). The duality technique has been introduced to games with incomplete information by De Meyer and widely investigated since then (see De Meyer (1996), De Meyer and Rosenberg (1999), De Meyer and Marino (2005), De Meyer and Moussa-Saley (2003), Laraki (2002), Sorin (2002)). For any game with incomplete information on one side, he defined a dual game for which the value is the Fenchel conjugate of the value of the initial game.

The paper is organized as follows. We first present the model with the notion of game and of information structure. We define the value of information function. We introduce the notion of standard information structure so that we can state properly the main result that characterizes value of information functions. We define the main notions of concavity, convexity and Blackwell monotonicity, and state the result. We

then sketch the proof and state the main structure theorem before proceeding to the proof itself.

2 The model

2.1 Game and information structures

Two players play a zero-sum game with action spaces A and B and whose payoff function g depends on a pair of parameters (k, l), which is randomly chosen from $K \times L$ according to an independent distribution $\mathbf{p} \otimes \mathbf{q}$. The sets K and L are assumed to be finite and the distributions $\mathbf{p}$ and $\mathbf{q}$ are assumed to have a full support.

Before the game starts the players obtain a noisy signal that depends on the realized parameter. When (k, l) is realized, player i obtains a signal m_i from measurable space $(M_i, \mathcal{M}_i)$ that is randomly chosen according to the distribution $\nu_i = \nu_i(h)$, where $h = k, l$, depending on i. ν_1 (resp. ν_2) is thus a transition probability from K (resp. L) to $(M_1, \mathcal{M}_1)$ (resp. $(M_2, \mathcal{M}_2)$). $I_i := (M_i, \mathcal{M}_i, \nu_i)$ is called the information structure of player i and $\mathcal{I}_i$ denotes the set of all player i's information structures.
We will see that the relevant aspect of an information structure is that it induces a probability distribution over conditional probabilities over the state space. Upon receiving the message m_1 player 1 computes the conditional probability that the realized state is k. However, the message m_1 is random and hence, the conditional probability on K given m_1 is random so that I_1 actually induces (ex ante) a distribution over $\Delta(K)$.

2.2 Strategies and payoffs

Player 1 has action set A and player 2 action set B. The action spaces can contain infinitely many actions and we therefore have to assume that they are endowed with σ-algebras $\mathcal{A}$ and $\mathcal{B}$. [1]. Also, for all k, l, the payoff map $(a, b) \to g(a, b, k, l) \in \mathbb{R}$ is

[1] We further will make the technical hypothesis that $(A, \mathcal{A})$ and $(B, \mathcal{B})$ are standard Borel. A probability space $(A, \mathcal{A})$ is standard Borel if there is a one to one bi-measurable map between $(A, \mathcal{A})$ and a Borelian subset of $\mathbb{R}$. This hypothesis will ensure the existence of regular conditional probabilities, when required. We could have dispensed with this hypothesis at the price of restricting the definition of a strategy in the next section: a strategy σ would be defined in this context as a measurable map from $([0, 1] \times M_1, \mathcal{B}_{[0,1]} \otimes \mathcal{M}_1)$ to $(A, \mathcal{A})$, where $\mathcal{B}_{[0,1]}$ is the Borel σ-algebra on $[0, 1]$. Upon receiving the message m_1, player 1 selects a uniformly distributed random number $u \in [0, 1]$ and plays the action $\sigma(u, m_1)$.

assumed to be $\mathcal{A}\otimes\mathcal{B}$-measurable. A game Γ is thus represented by $\Gamma = \langle A, \mathcal{A}, B, \mathcal{B}, g\rangle$.

Upon receiving the messages player 1 takes an action $a \in A$ and player 2 takes an action $b \in B$. Then player 2 pays player 1 $g(a, b, k, l)$. Note that g depends on both actions and on the parameters k and l. Players are allowed to make mixed moves and a strategy of player 1 is thus a transition probability $\sigma : M_1 \to A$ while a strategy of player 2 is a transition probability $\tau : M_2 \to B$. In other words, for all m_1, $\sigma(m_1)$ is a probability on $(A, \mathcal{A})$ according to which player 1 will choose his action if he receives the message m_1. Furthermore, for all $A' \in \mathcal{A}$, the map $m_1 \to \sigma(m_1)[A']$ is measurable. A similar condition applies to τ.

A strategy σ (resp. τ) induces, together with $\mathbf{p}, \nu_1$ (resp. $\mathbf{q}, \nu_2$), a probability distribution over $M_1 \times A \times K$ (resp. $M_2 \times B \times L$), which we denote by Π_σ (resp. Π_τ). We also denote π_σ and π_τ the marginals of Π_σ and Π_τ on $A \times K$ and $B \times L$ respectively. Let also $\pi_{\sigma,\tau}$ denote $\pi_\sigma \otimes \pi_\tau$.

The payoff is then $g(\sigma, \tau) := E_{\pi_{\sigma,\tau}}[g(a, b, k, l)]$. However, since we deal with general functions g, this expectation could fail to exist for all σ, τ. A strategy σ of player 1 is called *admissible* if it induces payoffs bounded from below for any strategy of player 2. Formally, σ is *admissible* if there exists $m \in \mathbb{R}$ such that, $\forall l, b,\ \ E_{\pi_\sigma}[|g(k, a, l, b))|] < \infty$, and $E_{\pi_\sigma}[g(k, a, l, b)] \geq m$. Admissible strategies of player 2 are defined in a similar fashion[2]. The sets of admissible strategies are denoted Σ_a and T_a. [3]

2.3 Value of information function

The game with the information structures and I_1 and I_2 is denoted by $\Gamma(I_1, I_2)$. It has a value, denoted $V_\Gamma(I_1, I_2)$, if

$$V_\Gamma(I_1, I_2) = \sup_{\sigma\in\Sigma_a} \inf_{\tau} g(\sigma, \tau) = \inf_{\tau\in T_a} \sup_{\sigma} g(\sigma, \tau).$$

Our goal is to study the relationship between the value of the game and the information structures.

Definition 2.1 *A function $V : \mathcal{I}_1 \times \mathcal{I}_2 \to \mathbb{R}$ is called a value-of-information function if there is a game Γ such that for every I_1, I_2, the value of the game $\Gamma(I_1, I_2)$ is $V(I_1, I_2)$.*

[2]In fact it is enough that $E_{\pi_\sigma}[\max(-g(k, a, l, b), 0)] < \infty$ and $g(k, a, l, b)$ is integrable, with an integral that is a real number or infinity.

[3]Note that $g(\sigma, \tau)$ is well defined if either $\sigma \in \Sigma_a$ or $\tau \in T_a$.

The purpose of this paper is to characterize the value-of-information functions.

2.4 Standard information structures

The relevant aspect of an information structure is that it induces a probability distribution over conditional probabilities over the state space. Upon receiving the message m_1 player 1 computes the conditional probability that the realized state is k, denoted by $[I_1](m_1)(k)$. Since the message m_1 is random, I_1 actually induces (*ex* ante) a distribution over[4] $\Delta(K)$. We refer to this distribution as the law of I_1 and denote it by $[I_1]$. Denote by $\Delta_{\mathbf{p}}(\Delta(K))$ the set of all probability measures ν over $\Delta(K)$ whose average value is $\mathbf{p}$ (i.e., $E_\nu(p) = \mathbf{p}$). Note that $[I_1]$ belongs to this set.

Different values of m_1 could lead to the same conditional probability $[I_1](m_1)$. There is thus apparently more information embedded in m_1 than in $[I_1](m_1)$ alone. However, this additional information is irrelevant for predicting k since k and m_1 are independent given $[I_1](m_1)$. Therefore, after computing $[I_1](m_1)$, player 1 may forget the value of m_1 and play in a game Γ a strategy that just depends on $[I_1](m_1)$.[5]

The standard information structure I_1' that is associated to I_1 is defined as follows. The message set is $M_1' := \Delta(K)$ and the message is $m_1' := [I_1](m_1)$ when the message under I_1 was m_1. Clearly, the posterior probability on K given m_1' is in this case equal to m_1'. It is therefore more convenient to use the letter P for the message sent by I_1', since it is precisely the posterior. The standard information structure I_2' for player 2 corresponding to I_2 is defined in the same fashion and the message sent by I_2' will be denoted Q.

In a game Γ, player i will play identically with I_i an with I_i' and have the same payoff. Therefore, if V_Γ is a value-of-information function, then $V_\Gamma(I_1, I_2) = V_\Gamma(I_1', I_2')$. This implies that the value of information function depends only on the standard information structure, or equivalently on the law induced by the information structure on $\Delta(K)$. This fact is captured in the following proposition.

Proposition 2.2 *If V is a value of information function, then there exists a map*

[4]For a measurable set M we denote by $\Delta(M)$ the set of all probability measures over M.

[5]Indeed, consider an arbitrary strategy σ and the probability Π_σ it induces jointly with $\mathbf{p}, \nu_1$ on $M_1 \times A \times K$. Denote $\overline{\sigma}(P)$ de conditional probability on a given $[I_1](m_1) = P$. Since $(A, \mathcal{A})$ is Borel standard, such a transition probability from $\Delta(K)$ to A exists. Denote next $\overline{\sigma}$ the strategy that consists in playing $\overline{\sigma}([I_1](m_1))$ upon receiving the message m_1. It appears that the marginals π_σ and $\pi_{\overline{\sigma}}$ of Π_σ and $\Pi_{\overline{\sigma}}$ on $K \times A$ coincide. Since the payoff in Γ just depends on π_σ, we infer that for every strategy τ of player 2, $g(\sigma, \tau) = g(\overline{\sigma}, \tau)$.

$v : \Delta_{\mathbf{p}}(\Delta(K)) \times \Delta_{\mathbf{q}}(\Delta(L)) \to \mathbb{R}$, *such that for all* I_1, I_2, $V(I_1, I_2) = v([I_1], [I_2])$.

The rest of the paper is devoted to characterizing value of information functions by focusing on functions defined over distribution over posteriors.

3 The main theorem

Before we state the main theorem, we need a few definitions concerning the notions of concave-convex functions and of Blackwell monotonicity. These turn out to be central in the characterization of value-of-information functions.

3.1 Concavity-convexity and Blackwell monotonicity

Definition 3.1 *Let* μ_1, μ_2 *be two probability measures in* $\Delta_{\mathbf{p}}(\Delta(K))$. *We will say that* μ_1 *is* more informative *than* μ_2, *denoted* $\mu_1 \succeq \mu_2$, *if there exist two random variables* X_1 *and* X_2 *such that the distribution of* X_i *is* μ_i, $i = 1, 2$ *and* $E(X_1|X_2) = X_2$. *A similar definition holds for distributions in* $\Delta_{\mathbf{q}}(\Delta(L))$.

Note that 'being more informative than' is a partial order.

Blackwell (Blackwell, 1953) proved that μ_1 is more informative than μ_2 iff in any one-player optimization problem (with the state space being K with probability $\mathbf{p}$) the optimal payoff corresponding to the law of μ_1 is no lower than that corresponding to the law of μ_2.

Definition 3.2 *A function* $v : \Delta_{\mathbf{p}}(\Delta(K)) \times \Delta_{\mathbf{q}}(\Delta(L)) \to \mathbb{R}$ *is* concave-convex, semi-continuous and Blackwell monotonic *if*

1. $\forall \nu$, *the map* $\mu \to v(\mu, \nu)$ *is concave, upper semi-continuous (usc) with respect to the weak topology on* $\Delta_{\mathbf{p}}(\Delta(K))$ *and increasing with respect to* $\succeq$.

2. $\forall \mu$, *the map* $\nu \to v(\mu, \nu)$ *is convex, lower semi-continuous (lsc) with respect to the weak topology on* $\Delta_{\mathbf{q}}(\Delta(L))$ *and decreasing with respect to* $\succeq$.

Note that the definition involves two distinct properties: concavity-convexity and Blackwell monotonicity. Both are known to be related to the fact that information is valuable in zero-sum games.

The concavity-convexity property of the value appears in the context of repeated games with incomplete information (see, Aumann and Maschler, 1995). Concavity of a function v states that when μ and μ' are distributions over $\Delta(K)$, the value the function takes at $\lambda\mu + (1-\lambda)\mu'$, $v(\lambda\mu + (1-\lambda)\mu')$ is higher than the weighted average of the values it takes at μ and μ', $\lambda v(\mu) + (1-\lambda)v(\mu')$. We interpret this property assuming that v is the value of a game. Let I_1 be an information structure. Its law, $[I_1]$, represents what player 2 knows about player 1's information. Suppose that $[I_1] = \lambda\mu + (1-\lambda)\mu'$. The fact that player 2 knows the information structure means that he knows that the information of player 1 is determined by a lottery that determines whether the law will be μ (with probability λ) or μ' (with probability $1-\lambda$). But player 2 does not know the outcome of this lottery. When player 2 is better informed (knowing the outcome of the lottery), the value of the game, $\lambda v(\mu) + (1-\lambda)v(\mu')$, is smaller (recall, player 2 is the minimizer) than when he is less informed, $v([I_1])$. This is precisely the concavity condition.

Blackwell monotonicity, on the other hand, relates to the information owned by the player about the state space he is directly informed of. Blackwell monotonicity compares the value the function takes at distributions μ_1 and μ_2 that relate to each other in the following way. There exists a random vector (X_1, X_2) such that $E[X_1|X_2] = X_2$ and X_i is μ_i distributed, $i = 1, 2$. This means that the conditional probability over K given by μ_1 is more precise than the one given by μ_2 (it is a martingalization), and the value of the game will therefore be higher. [6]

None of of concavity-convexity and Blackwell monotonicity properties implies the other. To illustrate the difference between them consider the following example. Let f be a function defined on $\Delta_{\mathbf{p}}(\Delta(K))$ as $f(\mu) = -E_\mu[||p - \mathbf{p}||]$. This function is linear in μ and, in particular, concave. Suppose now that μ_2 is a Dirac mass at $\mathbf{p}$ and $\mu_1 \in \Delta_{\mathbf{p}}(\Delta(K))$ puts no weight at $\mathbf{p}$. Clearly, μ_1 is more informative than μ_2. However, $f(\mu_1) < 0 = f(\mu_2)$. It implies that f is not Blackwell monotonic.

[6] More formally, consider then the following information structures: Nature first picks at random the vector (X_1, X_2) with the corresponding joint law and then selects k with the lottery $X_1 \in \Delta(K)$. In information structure number 1, the message sent by nature to player 1 is the pair (X_1, X_2). The posterior of player 1 in this case is thus X_1 and the law of this information structure is thus μ_1. In information structure number 2, the message sent by nature to player 1 is X_2. If he had observed X_1 in addition to X_2, his posterior would have been X_1. Since he has not observed X_1, his posterior is thus $E[X_1|X_2] = X_2$. Therefore the law of information structure number 1 is μ_2. Going from information structure number 1 to number 2 consists for player 1 in a loss of information: it reduces strategies the set of available strategies. Therefore $v(\mu_1, \nu) \geq v(\mu_2, \nu)$.

3.2 Main theorem

The main result of this paper which will be proved in the next sections is:

Theorem 3.3 *The map $V : \mathcal{I}_1 \times \mathcal{I}_2 \to \mathbb{R}$ defined by $V(I_1, I_2)$ is a value-of-information function iff $V(I_1, I_2) = v([I_1], [I_2])$ for a concave-convex, semi-continuous and Blackwell-monotonic function v.*

Blackwell proved that monotonicity is a necessary condition for a function to be the value of a one-player optimization problem. Since the value of a game is also the optimal value of an optimization problem of player i, assuming implicitly that player $-i$ plays a best reply, Blackwell monotonicity is a necessary condition here also. Lehrer and Rosenberg (2006) proved that this is also the case in games with incomplete information on one side with deterministic information.

As in Lehrer and Rosenberg (2006), Theorem 3.3 proves that only the concavity-convexity property and Blackwell monotonicity are common to all value-of-information functions. It implies that when one observes the results of interactions under various information functions, the only properties of the observed results that one may expect to obtain are concavity-convexity and Blackwell monotonicity, which are necessary. Thus, the main implication of Theorem 3.3 is that only violations of concavity-convexity property or Blackwell monotonicity can refute the hypothesis that the interaction can be modeled as a game played by Bayesian players. For an elaboration of this point, see Lehrer and Rosenberg (2006).

3.3 The deterministic case

In the special case of deterministic information, more can be said. It turn out that in this case, the concave-convex property and Blackwell-monotonicity coincide.

When μ_1 and ν_2 are restricted to be deterministic functions, it is equivalent for player 1 to know m_1 or to know the subset of K for which the message would be m_1. A deterministic information structure can be therefore modeled as a pair of partitions $\mathcal{P}, \mathcal{Q}$ respectively on K and L with the interpretation that if the true state is k, l player 1 is informed of the cell of $\mathcal{P}$ that contains k and player 2 is informed of the cell of $\mathcal{Q}$ that contains l. We denote the set of partitions on K and L respectively by $\mathcal{K}$ and $\mathcal{L}$. Note that there are only finitely many such partitions. A value of information function is now a real function defined over the set of partitions. In this

case Blackwell monotonicity is translated to the following. Recall that a partition $\mathcal{P}$ is said to refine a partition $\mathcal{P}'$ if any atom of $\mathcal{P}'$ is a union of atoms of $\mathcal{P}$.

A function V from $\mathcal{K} \times \mathcal{L}$ to $\mathbb{R}$ is increasing (resp. decreasing) in $\mathcal{P}$ (resp. $\mathcal{Q}$) if for any $\mathcal{P}$, $\mathcal{P}'$ in $\mathcal{K}$, such that $\mathcal{P}$ refines $\mathcal{P}'$, and any $\mathcal{Q}$ in $\mathcal{L}$, $V(\mathcal{P}, \mathcal{Q}) \geq V(\mathcal{P}', \mathcal{Q})$ (resp. for any $\mathcal{P} \in \mathcal{K}$ and $\mathcal{Q}, \mathcal{Q}' \in \mathcal{L}$ such that $\mathcal{Q}$ refines $\mathcal{Q}'$, $V(\mathcal{P}, \mathcal{Q}) \leq V(\mathcal{P}, \mathcal{Q}')$).

Theorem 3.4 *Let v be a function on pairs of partitions respectively on K and L. It is a value of information function iff it is Blackwell monotonic. Moreover it can then be obtained as the value of a finite game (i.e., A and B are finite).*

This theorem differs from Theorem 3.3 in two respects. First, the games are finite. This stronger result can be derived from Theorem 3.3, taking into account that there are only finitely many partitions. Second, the condition of concavity-convexity of the function is dropped here. This is so because any function of partitions that is Blackwell monotonic can be extended to a function over laws of posterior probabilities that is concave-convex, semi continuous and Blackwell monotonic (this result is proved in section 10). Thus, Theorem 3.4 derives from Theorem 3.3 which extends Lehrer and Rosenberg (2006) that applies to games with incomplete information with deterministic information.

4 A structure theorem

The first main step of the proof is to prove a structure theorem that characterizes concave-convex and Blackwell monotonic functions.

This theorem extends the well-known characterization that states that a convex function is the maximum of the linear functions that are below it. In our setup a linear function on μ (a distribution over $\Delta(K)$) is the integral of some function on $\Delta(K)$ with respect to μ. Therefore, a concave function f satisfies that for any law μ_0 on $\Delta(K)$, $f(\mu_0)$ is the infimum of $\int \psi(p) d\mu_0(p)$ over continuous functions ψ (defined on $\Delta(K)$)that satisfy $\Delta(K)$, $E_\mu[\psi(P)] \geq f(\mu)$ for every law μ on $\Delta(K)$.

The second part of the theorem characterizes functions that have the concave-convex property and are also Blacwell monotonic. It gives a clearer sense of the difference between both notions. In order for a function to be in addition Blackwell monotonic it needs to be the infimum of $\int \psi(p) d\mu(p)$ over continuous functions ψ such that (i) the integral is greater than $f(\mu)$, and (ii) ψ is convex.

The intuition behind this result is that $\psi(p)$ represents the payoff that can be achieved by some strategy of player 2 when the conditional probability on K is p and player 1 plays optimally. The integral $\int \psi(p)d\mu(p)$ represents the expected payoff and the minimum is taken with respect to available strategies of player 2. Blackwell monotonicity implies that a martingalization of the conditional probabilities gives more information to player 1 and enables him to increase the payoff so that ψ will be convex.

In order to state formally the result, we need a few notations.

Ψ^0 and Ψ^1 denote respectively the set of continuous and continuously differentiable functions from $\Delta(K)$ to $\mathbb{R}$. The set of convex functions in Ψ^i will be denoted $\Psi^{i,vex}$, $i = 0, 1$.

For a given function $\psi \in \Psi^0$, and a law μ in $\Delta_{\mathbf{p}}(\Delta(K))$, we denote by $\tilde{\psi}(\mu)$ the expectation of ψ with respect to μ (i.e., $\int \psi(p)d\mu(p)$).
For $f : \Delta_{\mathbf{p}}(\Delta(K)) \to \mathbb{R}$, we define Ψ^i_f (resp. $\Psi^{i,vex}_f$) to be the set of functions $\psi \in \Psi^i$ (resp. $\psi \in \Psi^{i,vex}$) such that $\forall \mu \in \Delta_{\mathbf{p}}(\Delta(K))$, $f(\mu) \leq \tilde{\psi}(\mu)$.

Theorem 4.1

1. *If a function $f : \Delta_{\mathbf{p}}(\Delta(K)) \to \mathbb{R}$ is concave and usc with respect to the weak topology, then $\forall \mu$, $f(\mu) = \inf_{\psi \in \Psi^0_f} \tilde{\psi}(\mu)$.*

2. *If f is also Blackwell increasing, then $\forall \mu$, $f(\mu) = \inf_{\psi \in \Psi^{0,vex}_f} \tilde{\psi}(\mu)$;*

3. *furthermore, $\forall \mu : f(\mu) = \inf_{\psi \in \Psi^{1,vex}_f} \tilde{\psi}(\mu)$*

This theorem will be proved in section 7.

5 The conditions of Theorem 3.3 are necessary

In this section we show that any value-of-information function is concave-convex and Blackwell monotonic.

Theorem 5.1 *If the map $V(I_1, I_2) = v([I_1], [I_2])$ is a value of information function, then v must be concave-convex semi-continuous and Blackwell monotonic.*

Indeed, one can compute the expected payoff given a pair of strategies (σ, τ), denoted by $g(\sigma, \tau)$. One expression of this payoff is

$$g(\sigma, \tau) = E_{[I_1]}\left[\sum_{k \in K} P(k) E_{\sigma(P), \pi_\tau}[g(a, b, k, l)]\right],$$

where $\sigma(P)$ denotes the strategy followed given the conditional probability is P. Note that since σ is the family of all possible $\sigma(P)$ for all P, one has

$$\sup_\sigma g(\sigma, \tau) = E_{[I_1]}\left[\sup_{\sigma(P)} \sum_{k \in K} P(k) E_{\sigma(P), \pi_\tau}[g(a, b, k, l)]\right] = E_{[I_1]}\left[\sup_{a \in A} \sum_{k \in K} P(k) E_{\pi_\tau}[g(a, b, k, l)]\right]$$

This function is the expectation of a convex function in P so that by Jensen's inequality it is Blackwell increasing in $[I_1]$ and so is $v([I_1][I_2]) = \inf_\tau \sup_\sigma g(\sigma, \tau)$. Moreover the above expression proves that $\sup_\sigma g(\sigma, \tau)$ is linear and continuous in $[I_1]$ so that $v([I_1][I_2])$ is concave and usc.

6 Sketch of the proof of Theorem 3.3

Our goal is now to prove the reciprocal of the above theorem. For an arbitrary concave-convex, Blackwell monotonic and semi-continuous function v, we construct a game Γ such that $V_\Gamma(I_1, I_2) = v([I_1], [I_2]), \forall I_1, I_2$.

6.1 Games with incomplete information on one side

6.1.1 A first construction: one-sided information game

We first sketch the proof of the main result in the special case where L is a singleton, which is the case of games with incomplete information on one side. In this case the value functions depend only on their first argument. Given a concave, semi-continuous and Blackwell monotonic function v, we construct a two stage game whose value is $v([I_1])$.

The first intuition is to construct a two stage game. At the first stage player 2 picks a function $\psi \in \Psi_v^{1,vex}$, and his choice is observed by player 1. Then from stage 2 on, the game will proceed in a way that is continuation value will be $\tilde{\psi}(\mu)$. The structure theorem then implies that the value of such a game is indeed given by v.

In order for the value of the continuation game to be $\tilde{\psi}(\mu)$, we use the convexity of ψ and allow player 1 to choose a payoff a_k for each state k, with the constraint that the overall expected payoff is dominated by ψ. This amount to choosing a linear function of the conditional probability P on K that is dominated by ψ.

At the first stage Player 2 selects $\psi \in \Psi_v^{1,vex}$. Player 1 is then informed of ψ. The message given by player 1's standard information structure (which is defined in section 2.4) is the conditional probability P on k given his message. He then selects an affine functional $\langle a(P,\psi), p\rangle + \alpha(P,\psi)$ dominated by ψ. The payoff of the game is $\langle a(P,\psi), P\rangle + \alpha(P,\psi)$[7].

The normal form of this game has value v, because of the constraint on the choices of α and a and of the definition of $\Psi_v^{1,vex}$.

6.1.2 Another parametrization of the same game

Observe that when informed of ψ, P, player 1 may select an affine functional tangent to ψ at some point $\tilde{p}(P,\psi)$, for otherwise his strategy would be dominated. Since ψ is differentiable, the corresponding affine functional of p is $\psi(\tilde{p}(P,\psi)) + \langle \nabla\psi(\tilde{p}(P,\psi)); p - \tilde{p}(P,\psi)\rangle$. One might thus think of the previous game as the following two stage game. Player 2 selects $\psi \in \Psi_v^{1,vex}$. Player 1 is informed of ψ and P, and then he selects an action $\tilde{p} \in \Delta(K)$. The payoff is given by

$$\psi(\tilde{p}(P,\psi)) + \langle \nabla\psi(\tilde{p}(P,\psi)); P - \tilde{p}(P,\psi)\rangle, \tag{1}$$

Put it differently, if the chosen action are $\psi, \tilde{p}$ and the state is k, the payoff is

$$g(k,\tilde{p},\psi) := \psi(\tilde{p}) + \frac{\partial}{\partial p_k}\psi(\tilde{p}) - \langle \nabla\psi(\tilde{p}); \tilde{p}\rangle. \tag{2}$$

This game formulation has the advantage that player 1's action space is independent of player 2's move.

6.1.3 A simultaneous-move game

In order to extend the previous construction to games with incomplete information on both sides, player 2 should be given a way to condition his move on his information.

[7] This is a game with incomplete information: if the moves of the players are ψ and (a,α) and the state is k, the payoff is $a_k + \alpha$. This payoff depends on ψ because the possible choices of a are constrained by ψ.

However, letting player 2 play first will enable player 1 to extract information from the chosen act of player 2. For this reason we need to construct a simultaneous move game with the same value.

The simultaneous move game is described as follows. Once informed of P, Player 1 selects a point $\tilde{p}(P) \in \Delta(K)$. Simultaneously, Player 2 picks $\psi \in \Psi_v^{1,vex}$ and the payoff $g(k, \tilde{p}, \psi)$ is given by (2). Note that this game has the same value like the previous one, because in the previous game an optimal strategy for player 1 was to ignore ψ and to choose $\tilde{p}(\psi, P) = P$.

6.2 The extension to games with incomplete information on both sides: a first intuition

We turn now to games with incomplete information on both sides. We will find a game Γ whose value is $v([I_1], [I_2])$ for every $[I_1], [I_2]$. From the section on games with incomplete information above, we know that for a fixed $[I_2]$, one can construct a game $\Gamma_{[I_2]}$ with value v. However, the strategy space of player 2 in $\Gamma_{[I_2]}$ varies with $[I_2]$. That is, $\Psi^{1,vex}_{v(\cdot,[I_2])}$ might differ from $\Psi^{1,vex}_{v(\cdot,[I_2'])}$ when $[I_2] \neq [I_2']$,

One way to unify the strategy spaces in $\Gamma_{[I_2]}$ is the following. If $\psi \in \Psi^{1,vex}$, then the function $\psi + \alpha$ belongs to $\Psi^{1,vex}_{v(\cdot,\nu)}$ whenever the constant α is large enough. More precisely, we require that $\forall \mu \in \Delta_{\mathbf{p}}(\Delta(K)) : \tilde{\psi}(\mu) + \alpha \geq v(\mu, \nu)$. This turns out to be equivalent to

$$\alpha \geq w(\psi, \nu) := \sup_{\mu \in \Delta_{\mathbf{p}}(\Delta(K))} v(\mu, \nu) - \tilde{\psi}(\mu). \tag{3}$$

The idea is to let player 2 choose any function $\psi \in \Psi^{1,vex}$ and to use $\psi + w$ in the previous game, leading thus to a payoff of:

$$\psi(\tilde{p}(P, \psi)) + \langle \nabla \psi(\tilde{p}(P, \psi)); P - \tilde{p}(P, \psi) \rangle + w(\psi, [I_2]).$$

Note that the function w is the Fenchel transform of the value v. This Fenchel transform can be viewed as the value of a dual game. This might help interpreting it as the value of information function of a game.

6.3 Reminder on dual games

The notion of the dual game was first introduced by De Meyer. Suppose that v is a value of information function, and G is the game whose value is V. Furthermore,

suppose that ψ is a function on $\Delta(K)$ that we will call a cost. In the dual game there are two stages. First, player 1 can buy an information structure μ at the cost of $\tilde{\psi}(\mu)$ and this choice is publicly observed. Then both players take part in the game G and get the corresponding payoff. Therefore the overall payoff in the dual game is the sum of the payoff in the original game that has been played and the cost of the information structure μ that has been chosen. The idea is that on the one hand it is always better for player 1 to get more information but on the other hand since this information is costly, depending on the cost, he might choose to buy a less informative structure. It turns out that the value of such a game is the Fenchel conjugate of v defined by $w(\psi, [I_2]) := \sup_{\mu \in \Delta_{\mathbf{p}}(\Delta(K))} v(\mu, [I_2]) - \tilde{\psi}(\mu)$. Since the function v is concave there is a bi-dual equality:

$$v([I_1], [I_2]) = \inf_{\psi \in \Psi^{1,vex}} w(\psi, [I_2]) + \tilde{\psi}([I_1])$$

Moreover w is convex and decreasing for the Blackwell order in $[I_2]$.

A dual version of the structure theorem can be obtained as follows[8]. We know that

$$v(\mu, \nu) = \sup_{\phi \in \Phi^{1,cav}_{v(\mu,.)}} \tilde{\phi}(\nu).$$

Therefore

$$w(\psi, [I_2]) = \sup_{(\mu,\phi) \in \Delta_{\mathbf{p}}(\Delta(K)) \times \Phi^{1,cav}_{v(\mu,.)}} \tilde{\phi}([I_2]) - \tilde{\psi}(\mu)$$

In the sequel we denote by $\mathcal{F}$ the set $\Delta_{\mathbf{p}}(\Delta(K)) \times \Phi^{1,cav}_{v(\mu,.)}$.

All these results will be proved formally in the sequel.

6.4 The construction of the game

We now turn to sketching the final proof in the case of games with lack of information on both sides.

The case of games with incomplete information on one side might compel one to introduce a game $\Gamma'(I_1, I_2)$ where player 1 chooses a P-specific $\tilde{p}$ in $\Delta(K)$, player 2

[8]For $f : \Delta_{\mathbf{q}}(\Delta(L)) \to \mathbb{R}$, we define Φ^0_f (resp. Φ^1_f) as the set of continuous (resp. continuously differentiable) functions $\phi : \Delta(L) \to \mathbb{R}$ such that $\forall \nu \in \Delta_{\mathbf{q}}(\Delta(L)), f(\nu) \geq \tilde{\phi}(\nu)$.

For $i = 1, 0$, $\Phi^{i,cav}_f$ will denote the set of concave functions $\phi \in \Phi^i_f$.

chooses $\psi \in \Psi^{1,vex}$, and the payoff is

$$\psi(\tilde{p}(P)) + \langle \nabla\psi(\tilde{p}(P)), P - \tilde{p}(P)\rangle + w(\psi, [I_2]).$$

This game clearly has the value $v([I_1], [I_2])$, like $\Gamma_{[I_2]}(I_1)$. However, this is not a game with incomplete information on both sides, because the payoff function should depend linearly on the posterior Q over L induced by I_2.

The dual structure theorem states that

$$w(\psi, [I_2]) = \sup_{(\phi,\mu)\in\mathcal{F}} -\tilde{\psi}(\mu) + \tilde{\phi}([I_2]).$$

This property enables us to separate ψ and $[I_2]$ and therefore to avoid the definition of a game that would depend on ψ and yield a two stage game.

From the above we know that $\tilde{\phi}([I_2])$ can be interpreted as the value of a game with incomplete information on one side, where player 2 is the informed player.

This leads us to the following definition of Γ. Once informed of P, player 1 selects $\tilde{p} \in \Delta(K)$, and $(\mu, \phi) = f \in \mathcal{F}$. Player 2, after being informed of Q, selects $\psi \in \Psi^{1,vex}$ and $\tilde{q} \in \Delta(L)$.

In words, each player selects two items. The first is a fictitious conditional probability to which he plays a best response as if this probability was the true one. The second is a payoff function for his opponent so that playing a best response with respect to this function leads to the desired value-of-function function by the structure theorem.

The payoff is now,

$$\psi(\tilde{p}) + \langle \nabla\psi(\tilde{p}), P - \tilde{p}\rangle - \tilde{\psi}(\mu) + \phi(\tilde{q}) + \langle \nabla\phi(\tilde{q}), Q - \tilde{q}\rangle$$

In other words, if $a = (\tilde{p}, (\mu, \phi))$, and $b = (\psi, \tilde{q})$, the payoff is

$$g(k, l, a, b) := \psi(\tilde{p}) + \frac{\partial}{\partial p_k}\psi(\tilde{p}) - \langle \nabla\psi(\tilde{p}), \tilde{p}\rangle - \tilde{\psi}(\mu) + \phi(\tilde{q}) + \frac{\partial}{\partial q_l}\phi(\tilde{q}) - \langle \nabla\phi(\tilde{q}), \tilde{q}\rangle$$

We will prove that for all information structures (I_1, I_2), the values of $\Gamma(I_1, I_2)$ is $v([I_1], [I_2])$ and Theorem 3.3 will then follow.

6.5 Intuition of a construction through dual games

Another approach is to try to construct the game using known results on games with incomplete information on one side. Indeed, the result on games with incomplete

information on one side proves that for each ψ, $w(\psi, .)$ is a value of information function of a game with incomplete information on one side where player 2 is the informed player (call G_ψ the corresponding game) and that $\tilde{\psi}$ is a value of information function of a game with incomplete information on one side where player 1 is the informed player (call G' the corresponding game). The idea is to define a game in which first player 2 chooses ψ and this choice is observed, then games G and G' are played independently for this choice of ψ and the total payoff is the sum of payoffs. The problem with such a construction is that in the true game player 2 can use his information to choose ψ and in order for the value not to be affected by this possibility one has to prove that it is optimal for player 2 not to use his information in the first stage. This can be done at least in the deterministic case. As we did in the case of games with incomplete information on one side we may also avoid this problem by designing a simultaneous move game.

7 Proof of the structure Theorem 4.1

Recall that for $f : \Delta_{\mathbf{p}}(\Delta(K)) \to \mathbb{R}$, we define Ψ_f^0 (resp. Ψ_f^1) as the set of continuous (resp. continuously differentiable) functions $\psi : \Delta(K) \to \mathbb{R}$ such that $\forall \mu \in \Delta_{\mathbf{p}}(\Delta(K)), f(\mu) \leq \tilde{\psi}(\mu)$.

For $i = 1, 0$, $\Psi_f^{i,vex}$ will denote the set of convex functions $\psi \in \Psi_f^i$. Recall Theorem 4.1.

Theorem 7.1

1. *If a function $f : \Delta_{\mathbf{p}}(\Delta(K)) \to \mathbb{R}$ is concave usc with respect to the weak topology, then $\forall \mu, f(\mu) = \inf_{\psi \in \Psi_f^0} \tilde{\psi}(\mu)$.*

2. *If f is further Blackwell increasing, then $\forall \mu : f(\mu) = \inf_{\psi \in \Psi_f^{0,vex}} \tilde{\psi}(\mu)$.*

3. *In this case, we also have $\forall \mu : f(\mu) = \inf_{\psi \in \Psi_f^{1,vex}} \tilde{\psi}(\mu)$.*

Proof: We prove the first claim. Let f be a concave weakly-usc function on $\Delta_{\mathbf{p}}(\Delta(K))$. Since $\Delta_{\mathbf{p}}(\Delta(K))$ is weakly compact, f is bounded from above so that $\Psi_f^0 \neq \emptyset$. The definition of Ψ_f^0 implies that $\forall \mu, f(\mu) \leq \inf_{\psi \in \Psi_f^0} \tilde{\psi}(\mu)$

As for the converse inequality, for a given $\overline{\mu} \in \Delta_{\mathbf{p}}(\Delta(K))$, let t be such that $t > f(\overline{\mu})$. The set $A := \{\mu | f(\mu) < t\}$ is then an open set in the weak topology, and

it contains $\overline{\mu}$. Therefore, there exists a positive ϵ and a finite family $\{\psi_1, \ldots, \psi_n\}$ of continuous functions such that $B := \{\mu | \forall i : \tilde{\psi}_i(\mu) < \tilde{\psi}_i(\overline{\mu}) + \epsilon\}$ is included in A. In other words, if we define $L(\mu, s) \in \mathbb{R}^{n+1}$ as

$$L(\mu, s) := (\tilde{\psi}_1(\overline{\mu}) + \epsilon - \tilde{\psi}_1(\mu), \ldots, \tilde{\psi}_n(\overline{\mu}) + \epsilon - \tilde{\psi}_n(\mu), f(\mu) - t - s),$$

then $\forall \mu, \forall s \geq 0 : L(\mu, s) \notin]0, \infty[^{n+1}$. Since f is concave and $\tilde{\psi}_i$ is linear, the set $H := \{L(\mu, s) | \mu \in \Delta_{\mathbf{p}}(\Delta(K)), s \geq 0\}$ is convex and can thus be separated from the open convex cone $]0, \infty[^{n+1}$. There exists therefore $(\alpha_1, \ldots, \alpha_n, \alpha) \neq 0$ in $\mathbb{R}_+^{n+1}$ such that $\forall \mu, \forall s \geq 0$,

$$\sum_{i=1}^{n} \alpha_i (\tilde{\psi}_i(\overline{\mu}) + \epsilon - \tilde{\psi}_i(\mu)) + \alpha(f(\mu) - t - s) \leq 0.$$

Note that $\alpha > 0$. Otherwise, one of the α_i should be strictly positive ($(\alpha_1, \ldots, \alpha_n, \alpha) \neq 0$) and the left hand side of the last inequality evaluated at $\mu = \overline{\mu}$ would then be strictly positive. This is a contradiction.

Therefore, with $\lambda(x) := t + \sum_{i=1}^{n} \frac{\alpha_i}{\alpha}(\psi_i(x) - \epsilon - \tilde{\psi}_i(\overline{\mu}))$, we infer that $\lambda \in \Psi_f$. We have thus $t \geq \tilde{\lambda}(\overline{\mu}) \geq \inf_{\psi \in \Psi_f^0} \tilde{\psi}(\overline{\mu})$. Since this inequality holds for all $t > f(\overline{\mu})$, we conclude that $\inf_{\psi \in \Psi_f^0} \tilde{\psi}(\overline{\mu}) \leq f(\overline{\mu})$, as needed.

We now prove claim 2. Let f be in addition Blackwell increasing. Let ψ be in Ψ_f^0 and let $\phi := vex(\psi)$ be the convexifiaction of ψ. That is, ϕ is the largest lsc convex function dominated by ψ. Since for all x in $\Delta(K)$ there is a probability distribution μ_x on $\Delta(K)$ such that $E_{\mu_x}[p] = x$ and that $\phi(x) = E_{\mu_x}[\psi(p)]$, for any random variable X with probability distribution μ in $\Delta_{\mathbf{p}}(\Delta(K))$, we may consider a random vector Y whose conditional distribution given X is μ_X. (X, Y) is thus a martingale.

So, if $Y \sim \mu'$ we get $\mu' \succeq \mu$. Since $\psi \in \Psi_f$ and f is Blackwell increasing, we infer that $\tilde{\phi}(\mu) = E[\phi(X)] = E[E[\psi(Y)|X]] = E[\psi(Y)] = \tilde{\psi}(\mu') \geq f(\mu') \geq f(\mu)$. This is true for all μ. Thus, we have actually proved that if $\psi \in \Psi_f^0$, then $vex(\psi) = \phi \in \Psi_f^0$. Therefore $f(\mu) \leq \inf_{\psi \in \Psi_f^{0,vex}} \tilde{\psi}(\mu) \leq \inf_{\psi \in \Psi_f^0} \tilde{\psi}(\mu) = f(\mu)$.

Claim 3 follows from the fact that a convex continuous function ψ on $\Delta(K)$ can be uniformly approximated by a function $\psi_2 \in \Psi^{1,vex}$, up to an arbitrary $\epsilon > 0$. Indeed, ψ is the supremum of the set $\mathcal{G}$ of affine functionals g dominated by ψ on $\Delta(K)$. For $\epsilon > 0$, the family $O_g := \{p \in \Delta(K) : g(p) > \psi(p) - \epsilon/2\}$ indexed by $g \in \mathcal{G}$ is thus an open covering of the compact set $\Delta(K)$. It contains therefore a finite sub-covering

$\mathcal{G}'$, and we conclude that $\psi_1(p) := \max\{g(p) + \epsilon/2 : g \in \mathcal{G}'\}$ is a convex function that satisfies $\psi \leq \psi_1 \leq \psi + \epsilon/2$. The function ψ_1, as a maximum of finitely many affine functional is the restriction to $\Delta(K)$ of a Lipschitz concave function defined on $\mathbb{R}^K$ (also denoted ψ_1). Let us next define, for an arbitrary $\rho > 0$, $\psi_2(x) := E[\psi_1(x+\rho Z)]$, where Z is a random vector whose components are independent standard normal random variables. Then ψ_2 is continuously differentiable due the smoothing property of the convolution operator. ψ_2 is also convex, since if $\lambda \in [0,1], x, x' \in \mathbb{R}^K$:

$$\begin{aligned}
\lambda\psi_2(x) + (1-\lambda)\psi_2(x') &= E[\lambda\psi_1(x+\rho Z) + (1-\lambda)\psi_1(x'+\rho Z)] \\
&\geq E[\psi_1(\lambda x + (1-\lambda)x' + \rho Z)] \\
&= \psi_2(\lambda x + (1-\lambda)x')
\end{aligned}$$

From Jensen's inequality, we get $\psi_1(x) \leq \psi_2(x) \leq \psi_1(x) + \rho\kappa E[\|Z\|]$, where κ is the Lispchitz constant of ψ_1. For ρ small enough, we get thus $\psi \leq \psi_2 \leq \psi + \epsilon$. ■

8 The duality between v and w

We here prove the results stated in Section 6.3: the duality relation between v and w as well as the dual structure theorem.

The dual function w of v is defined by,

$$w(\psi, \nu) := \sup_{\mu \in \Delta_{\mathbf{p}}(\Delta(K))} v(\mu, \nu) - \tilde{\psi}(\mu).$$

w is thus a function on $U := \Psi^0 \times \Delta_{\mathbf{q}}(\Delta(L))$. It is convenient to introduce the following sets. Φ^0 and Φ^1 refer to the sets of functions ϕ on $\Delta(L)$ that are respectively continuous and continuously differentiable. For $i = 1, 0$, $\Phi^{i,cav}$ will denote the set of concave functions in Φ^i.

Consider next the set $F := \Delta_{\mathbf{p}}(\Delta(K)) \times \Phi^0$. This set is in a sense a dual of the set U: to each $f = (\mu, \phi) \in F$ corresponds a linear functional $\hat{f}$ on U defined by,

$$\hat{f} : (\psi, \nu) \in U \to \hat{f}(\psi, \nu) := -\tilde{\psi}(\mu) + \tilde{\phi}(\nu).$$

We define the F-topology on U as the minimal topology that makes continuous the map $(\psi, \nu) \to \hat{f}(\psi, \nu), \forall f \in F$. The F topology is thus the product topology of the $\Delta_{\mathbf{p}}(\Delta(K))$-weak topology on Ψ^0 (i.e., the minimal topology that makes $\psi \to E_\mu[\psi(p)]$ continuous, $\forall \mu \in \Delta_{\mathbf{p}}(\Delta(K))$) and of the weak topology on $\Delta_{\mathbf{q}}(\Delta(L))$.

Theorem 8.1

1. *Let v be a concave-convex semi-continuous function on $\Delta_{\mathbf{p}}(\Delta(K)) \times \Delta_{\mathbf{q}}(\Delta(L))$. Then the dual function w is convex on U, l.s.c. with respect to the F-topology.*

2. *For all ψ, the map $\nu \to w(\psi, \nu)$ is further decreasing with respect to Blackwell's order on $\Delta_{\mathbf{q}}(\Delta(L))$.*

3. *Furthermore $\forall \mu, \nu : v(\mu, \nu) = \inf_{\psi \in \Psi^{1,vex}} w(\psi, \nu) + \tilde{\psi}(\mu)$. (Fenchel duality equation).*

Proof: At fixed μ, the map $h_\mu : (\psi, \nu) \to v(\mu, \nu) - E_\mu[\psi(p)]$ is convex F-lsc in (ψ, ν), as $\nu \to v(\mu, \nu)$ is convex lsc in ν with respect to the weak topology on $\Delta_{\mathbf{q}}(\Delta(L))$ and $\psi \to -E_\mu[\psi(p)]$ is linear continuous with respect to the $\Delta_{\mathbf{p}}(\Delta(K))$-weak topology on $\Psi^{0,vex}$.

As a supremum of convex F-lsc maps, $w(\psi, \nu) := \sup_{\mu \in \Delta_{\mathbf{p}}(\Delta(K))} h_\mu(\psi, \nu)$ is thus convex F-lsc as desired.

If $\nu \preceq \nu'$, then for all $\mu : v(\mu, \nu) \geq v(\mu, \nu')$. Therefore, $\forall \mu, \psi : v(\mu, \nu) - \tilde{\psi}(\mu) \geq v(\mu, \nu') - \tilde{\psi}(\mu)$, and thus $w(\psi, \nu) \geq w(\psi, \nu')$ as stated in claim 2).

We next prove claim 3). From the definition of w in Eq. (3), we have clearly that $\forall \psi, \nu, \forall \mu : w(\psi, \nu) \geq v(\mu, \nu) - \tilde{\psi}(\mu)$.
Therefore, $\forall \mu, \nu, \forall \psi : w(\psi, \nu) + \tilde{\psi}(\mu) \geq v(\mu, \nu)$. It follows that $\forall \mu, \nu, \inf_{\psi \in \Psi^{1,vex}} w(\psi, \nu) + \tilde{\psi}(\mu) \geq v(\mu, \nu)$.

On the other hand, observe that if $\psi \in \Psi^{1,vex}_{v(.,\nu)}$ then $\forall \mu : \tilde{\psi}(\mu) \geq v(\mu, \nu)$. Therefore $\forall \mu : 0 \geq v(\mu, \nu) - \tilde{\psi}(\mu)$, and thus $w(\psi, \nu) \leq 0$. From claim 3 in theorem 4.1, we conclude then that

$$\begin{array}{rcl} v(\mu, \nu) & = & \inf_{\psi \in \Psi^{1,vex}_{v(.,\nu)}} \tilde{\psi}(\mu) \\ & \geq & \inf_{\psi \in \Psi^{1,vex}_{v(.,\nu)}} w(\psi, \nu) + \tilde{\psi}(\mu) \\ & \geq & \inf_{\psi \in \Psi^{1,vex}} w(\psi, \nu) + \tilde{\psi}(\mu) \end{array}$$

Claim 3) is thus proved. ■

For $i = 0$ or 1, let us define $F^{i,cav}$ as $F^{cav} := \Delta_{\mathbf{p}}(\Delta(K)) \times \Phi^{i,cav}$. For $\mu \in \Delta_{\mathbf{p}}(\Delta(K))$, we define $\Phi^{i,cav}_{v(\mu,.)}$ as the set of $\phi \in \Phi^{i,cav}$ such that for all $\nu : \tilde{\phi}(\nu) \leq v(\mu, \nu)$.

Theorem 8.2 *Let $\mathcal{F}$ be $\{f \in F^{1,cav} : \forall (\psi, \nu) \in U : \hat{f}(\psi, \nu) \leq w(\psi, \nu)\}$.*

1. *If* $\mu \in \Delta_{\mathbf{p}}(\Delta(K))$ *and* $\phi \in \Phi^{1,cav}_{v(\mu,.)}$, *then* $(\mu, \phi) \in \mathcal{F}$.

2. $\forall(\psi, \nu) \in U : w(\psi, \nu) = \sup_{f \in \mathcal{F}} \hat{f}(\psi, \nu)$.

Proof: Consider a pair $f = (\mu, \phi)$ such that $\phi \in \Phi^{1,cav}_{v(\mu,.)}$. From the definitions of $\Phi^{1,cav}_{v(\mu,.)}$ and $w(\psi, \nu)$, we get

$$\forall(\psi, \nu) \in U : \hat{f}(\psi, \nu) = \tilde{\phi}(\nu) - \tilde{\psi}(\mu) \leq v(\mu, \nu) - \tilde{\psi}(\mu) \leq w(\psi, \nu),$$

which is equivalent to $f \in \mathcal{F}$.

We next prove claim 2. It results from the definition of $\mathcal{F}$ that $\forall(\psi, \nu) \in U$: $w(\psi, \nu) \geq \sup_{f \in \mathcal{F}} \hat{f}(\psi, \nu)$. On the other hand, the dual result of Theorem 4.1 indicates that

$$\forall \mu \in \Delta_{\mathbf{p}}(\Delta(K)), \forall \nu \in \Delta_{\mathbf{q}}(\Delta(L)), v(\mu, \nu) = \sup_{\phi \in \Phi^{1,cav}_{v(\mu,.)}} \tilde{\phi}(\nu).$$

Therefore, it follows from claim 1 that

$$\sup_{f \in \mathcal{F}} \hat{f}(\psi, \nu) \geq \sup_{\mu} \sup_{\phi \in \Phi^{1,cav}_{v(\mu,.)}} \tilde{\phi}(\nu) - \tilde{\psi}(\mu) = \sup_{\mu} v(\mu, \nu) - \tilde{\psi}(\mu) = w(\psi, \nu),$$

and the theorem is proved. ■

9 Proof of the main result: Theorem 3.3

9.1 Definition of Γ.

The game Γ is defined as follows: The action space A for player 1 is $\Delta(K) \times \mathcal{F}$ and the action space B for player 2 is $\Psi^{1,vex} \times \Delta(L)$. In state (k, l), if $a = (\tilde{p}, (\mu, \phi)) \in A$ and $b = (\psi, \tilde{q})$, the payoff $g(k, l, a, b)$ is:

$$g(k, l, a, b) := \psi(\tilde{p}) + \frac{\partial}{\partial p_k}\psi(\tilde{p}) - \langle \nabla\psi(\tilde{p}), \tilde{p}\rangle - \tilde{\psi}(\mu) + \phi(\tilde{q}) + \frac{\partial}{\partial q_l}\phi(\tilde{q}) - \langle \nabla\phi(\tilde{q}), \tilde{q}\rangle$$

We only will need to consider pure strategies in this game. Once informed of P, player 1 selects deterministically an action a_P, i.e., $\tilde{p}_P \in \Delta(K)$, and $(\mu_P, \phi_P) \in \mathcal{F}$. Once informed of Q, player 2 selects b_Q: $\psi_Q \in \Psi^{1,vex}$ and $\tilde{q}_Q \in \Delta(L)$.

The expected payoff conditionally to (P, Q) is then $\sum_{k,l} P_k Q_l g(k, l, a_P, b_Q) =$

$$\psi_Q(\tilde{p}_P) + \langle \nabla\psi_Q(\tilde{p}_P), P - \tilde{p}_P\rangle - \tilde{\psi}_Q(\mu_P) + \phi_P(\tilde{q}_Q) + \langle \nabla\phi_P(\tilde{q}_Q), Q - \tilde{q}_Q\rangle$$

9.2 Player 1 can guarantee $v([I_1],[I_2])$ in $\Gamma([I_1],[I_2])$.

Consider the following strategy of player 1 in $\Gamma([I_1],[I_2])$: $\tilde{p}_P := P$, $\mu_P := [I_1]$ and $\phi_P := \phi$ for an arbitrary $\phi \in \Phi^{1,cav}_{v([I_1],.)}$. Notice that it results from the first claim in theorem 8.2 that $([I_1],\phi) \in \mathcal{F}$, as required for a strategy of player 1.

Then the conditional expected payoff given (P,Q) is

$$\psi_Q(P) - \tilde{\psi_Q}([I_1]) + \phi(\tilde{q}_Q) + \langle \nabla\phi(\tilde{q}_Q), Q - \tilde{q}_Q\rangle.$$

Since P and Q are independent, and since the law of P is $[I_1]$, the conditional expectation of the first term given Q is $E_P[\psi_Q(P)] = \tilde{\psi_Q}([I_1])$: When taking expectations, the first two terms cancel each other in the above payoff. Since, furthermore, ϕ is concave, we get $\phi(\tilde{q}_Q) + \langle \nabla\phi(\tilde{q}_Q), Q - \tilde{q}_Q\rangle \geq \phi(Q)$, and the payoff of player 1 is at least: $E_Q[\phi(Q)] = \tilde{\phi}([I_2])$.

Since $\phi \in \Phi^{1,cav}_{v([I_1],.)}$ is arbitrary, we conclude that player 1 can guarantee $\sup_{\phi\in\Phi^{1,cav}_{v([I_1],.)}} \tilde{\phi}([I_2])$ which is equal to $v([I_1],[I_2])$ as it follows from a dual version of Theorem 4.1

9.3 Player 2 can guarantee $v([I_1],[I_2])$ in $\Gamma([I_1],[I_2])$.

The strategy $\tilde{q}_Q := Q$ and $\psi_Q := \psi$ for an arbitrary $\psi \in \Psi^{1,vex}$ leads to the following expression for the expected payoff conditionally to (P,Q):

$$\psi(\tilde{p}_P) + \langle \nabla\psi(\tilde{p}_P), P - \tilde{p}_P\rangle - \tilde{\psi}(\mu_P) + \phi_P(Q).$$

Since ψ is convex: $\psi(\tilde{p}_P) + \langle \nabla\psi(\tilde{p}_P), P - \tilde{p}_P\rangle \leq \psi(P)$. Conditionally to P, the payoff of player 1 is thus at most:

$$\psi(P) - \tilde{\psi}(\mu_P) + \tilde{\phi_P}([I_2]).$$

The last two terms in this expression are equal to $\hat{f}_P(\psi,[I_2])$, where $f_P := (\mu_P,\phi_P)$. From claim 2 in theorem 8.2, $\hat{f}_P(\psi,[I_2]) \leq w(\psi,[I_2])$. Player 1's payoff, conditionally to P is less than $\psi(P) + w(\psi,[I_2])$, and thus the above strategy of player 2 guarantees him $\tilde{\psi}([I_1]) + w(\psi,[I_2])$. Since $\psi \in \Psi^{1,vex}$ is arbitrary, we conclude that player 2 can guarantee $\inf_{\psi\in\Psi^{1,vex}} \tilde{\psi}([I_1]) + w(\psi,[I_2]) = v([I_1],[I_2])$, as it follows from claim 3 in Theorem 8.1.

10 Finite partitions

In this section we prove theorem 3.4. Indeed, the main effort of the section will be put in proving that that for any function $V(\mathcal{P},\mathcal{Q})$ of partitions which is monotonic for the refinement ordering is equal to $V(\mathcal{P},\mathcal{Q}) = v([\mathcal{P}],[\mathcal{Q}])$, where v is a concave-convex Blackwell monotonic semicontinuous function on $(\Delta_{\mathbf{p}}(\Delta(K)) \times \Delta_{\mathbf{q}}(\Delta(L)))$. Starting from a given V, defined on finitely many points of $(\Delta_{\mathbf{p}}(\Delta(K))\times\Delta_{\mathbf{q}}(\Delta(L)))$, we have to extend it to the whole set.This will be enough in order to prove the theorem. Indeed, having the function v, one can apply the previous result and find a (possibly infinite) game Γ with strategy sets A, B which has v as a value of information function. Then we restrict Γ to a new game Γ' with strategy sets $A' \subset A$ and $B' \subset B$ where A' is a finite subset of A and B' is a finite subset of B: A' (resp. B') is a finite collection of strategies that takes one optimal strategy in A (resp. B) for each pair of partitions $\mathcal{P},\mathcal{Q}$. It is clear that Γ' is a finite game which has V as a value of deterministic information function.

Let $\mathcal{C}$ denote the family of non empty subsets of K. For $C \in \mathcal{C}$, denote P_C the conditional probability on K given C. In other words $P_{C,k} := (\mathbf{p}_k/\mathbf{p}(C))\mathbb{1}_{k\in C}$. Note that if $C \neq C'$ then $P_C \neq P_{C'}$. Let S denote $S := \{P_C : C \in \mathcal{C}\}$. Any partition $\mathcal{P}$ will induce a law $[\mathcal{P}]$ of the posterior P. $[\mathcal{P}]$ is thus an element of $\Delta_{\mathbf{p}}(S)$ which denotes the set of probability distribution on S with expectation $\mathbf{p}$. For a set $C \in \mathcal{C}$ and a partition $\mathcal{P}$, the probability that the posterior P takes the value P_C is denoted $[\mathcal{P}](P_C)$ and is equal to $\mathbf{p}(C)$ if C belongs to $\mathcal{P}$, 0 otherwise. We get thus

$$[\mathcal{P}](P_C) := \mathbf{p}(C)\mathbb{1}_{C\in\mathcal{P}} \tag{4}$$

We first prove the next lemma that is the crux point of this section.

Lemma 10.1 *For any partition $\mathcal{P}$ of K, the law $[\mathcal{P}]$ is an extreme point of $\Delta_{\mathbf{p}}(S)$.*

Proof: Let μ be an element of $\Delta_{\mathbf{p}}(S)$. Define $\phi(C) := \mu(P_C)/\mathbf{p}(\mathbf{C})$. Then for all $k \in K$: $\mathbf{p}_k = \sum_{C\in\mathcal{C}} \mu(P_C)P_{C,k} = \sum_{C\in\mathcal{C}} \phi(C)\mathbb{1}_{k\in C}\mathbf{p}_k$, and thus

$$\forall k \in K : \sum_{C\in\mathcal{C}} \phi(C)\mathbb{1}_{k\in C} = 1. \tag{5}$$

Since $\forall C : \phi(C) \geq 0$, we infer that $\phi(C) \leq 1$ (This follows from (5) with $k \in C$).

A map ϕ from $\mathcal{C}$ to $[0,1]$ satisfying (5) will be called a Φ-*system*. Note that if ϕ is a Φ-system, then μ, defined by $\mu(P_C) := \phi(C).\mathbf{p}(\mathbf{C})$, belongs to $\Delta_{\mathbf{p}}(S)$. The map $\mu \to \phi$ is thus a linear bijection between $\Delta_{\mathbf{p}}(S)$ and the set of Φ-systems. To get our claim, we have to prove that the Φ-system ϕ corresponding to $[\mathcal{P}]$ is an extreme point in the set of Φ-systems. From (4), we obtain $\phi(C) = \mathbb{1}_{C\in\mathcal{P}}$. The corresponding ϕ is thus $\{0,1\}$-valued and therefore cannot be written as a convex combination of Φ systems different from itself. ■

A *decomposition of a measure* $\mu \in \Delta_{\mathbf{p}}(\Delta(K))$ is a pair $(\lambda, \mu_\bullet)$ such that

1. λ is a probability on partitions: $\lambda \in \Delta(\mathcal{K})$.
2. $\mu_\bullet$ is a map from $\mathcal{K}$ to $\Delta_{\mathbf{p}}(\Delta(K))$ such that $\sum_{\mathcal{P}\in\mathcal{K}} \lambda(\mathcal{P})\mu_{\mathcal{P}} = \mu$.
3. $\forall \mathcal{P} \in \mathcal{K} : [\mathcal{P}] \preceq \mu_{\mathcal{P}}$

Notice that there always exists a decomposition of a measure μ. Indeed, since the trivial partition $\{K\}$ is Blackwell dominated by any measure $\mu \in \Delta_{\mathbf{p}}(\Delta(K))$, the Dirac measure λ on $\{K\}$ joint with $\mu_\bullet$ defined by $\mu_{\mathcal{P}} := [\mathcal{P}]$ if $\mathcal{P} \neq \{K\}$, and $\mu_{\{K\}} := \mu$, is a decomposition of μ. We will denote by $\Lambda(\mu)$ the set of $\lambda \in \Delta(\mathcal{K})$ such that there exists a decomposition $(\lambda, \mu_\bullet)$ of μ.

A decomposition $(\xi, \nu_\bullet)$ of a measure $\nu \in \Delta_{\mathbf{q}}(\Delta(L))$ is defined analogously as follows.

1. ξ is a probability on partitions: $\xi \in \Delta(\mathcal{L})$.
2. $\nu_\bullet$ is a map from $\mathcal{L}$ to $\Delta_{\mathbf{q}}(\Delta(L))$ such that $\sum_{\mathcal{Q}\in\mathcal{L}} \xi(\mathcal{Q})\nu_{\mathcal{Q}} = \nu$.
3. $\forall \mathcal{Q} \in \mathcal{L} : [\mathcal{Q}] \preceq \nu_{\mathcal{Q}}$

We also define $\Xi(\nu)$ as the set of $\xi \in \Delta(\mathcal{L})$ such that there exists a decomposition $(\xi, \nu_\bullet)$ of ν.

For a Balckwell monotonic function $V : \mathcal{K} \times \mathcal{L} \to \mathbb{R}$, we may then consider the following function v.

$$v(\mu,\nu) := \sup_{\lambda\in\Lambda(\mu)} \inf_{\xi\in\Xi(\nu)} \sum_{\mathcal{P}\in\mathcal{K}} \sum_{\mathcal{Q}\in\mathcal{L}} \lambda(\mathcal{P})\xi(\mathcal{Q})V(\mathcal{P},\mathcal{Q}).$$

The next theorem states that $v(\mu,\nu)$ can be viewed as the value of a game where both players have optimal strategies.

Theorem 10.2

$$\begin{aligned} v(\mu,\nu) &= \max_{\lambda\in\Lambda(\mu)} \min_{\xi\in\Xi(\nu)} \sum_{\mathcal{P}\in\mathcal{K}} \sum_{\mathcal{Q}\in\mathcal{L}} \lambda(\mathcal{P})\xi(\mathcal{Q})V(\mathcal{P},\mathcal{Q}) \\ &= \min_{\xi\in\Xi(\nu)} \max_{\lambda\in\Lambda(\mu)} \sum_{\mathcal{P}\in\mathcal{K}} \sum_{\mathcal{Q}\in\mathcal{L}} \lambda(\mathcal{P})\xi(\mathcal{Q})V(\mathcal{P},\mathcal{Q}). \end{aligned}$$

Proof: Since the payoff is bilinear in (λ,ξ), we just have to prove that $\Lambda(\mu)$ and $\Xi(\nu)$ are convex compact sets. Let thus $\lambda,\lambda^i \in \Lambda(\mu)$, $i=0,1$ and let $r^i \geq 0$ be such that $1=\sum_i r^i$ and $\lambda=\sum_i r^i\lambda^i$. Then there exists $\mu^i_\bullet$ such that $(\lambda^i,\mu^i_\bullet)$ is a decomposition of μ. Defining thus $\mu_{\mathcal{P}} := (\sum_i r^i\lambda^i(\mathcal{P})\mu^i_{\mathcal{P}})/\lambda(\mathcal{P})$ if $\lambda(\mathcal{P})>0$ and $\mu_{\mathcal{P}} := [\mathcal{P}]$ otherwise, we infer that $(\lambda,\mu_\bullet)$ is also a decomposition of μ. Indeed, as convex combination of two points in $\Delta(\mathcal{K})$, λ belongs to $\Delta(\mathcal{K})$. Next, we prove in lemma 10.3 that the set $\{\mu\in\Delta_{\mathbf{p}}(\Delta(K)) : \mu \succeq [\mathcal{P}]\}$ is convex. Therefore, as a convex combination of two points in this set, $\mu_{\mathcal{P}}$ also belongs to it and thus $\mu_{\mathcal{P}} \succeq [\mathcal{P}]$, whenever $\lambda(\mathcal{P})>0$. We also have $\mu_{\mathcal{P}}=[\mathcal{P}]$ if $\lambda(\mathcal{P})=0$. Therefore, $(\lambda,\mu_\bullet)$ is a decomposition of μ and λ belongs to $\Lambda(\mu)$, which proves that $\Lambda(\mu)$ is a convex set.

We next prove that it is also closed. Let $\{\lambda^n\}_{n\in\mathbb{N}}$ be a sequence of points in $\Lambda(\mu)$ that converges to λ and let $\mu^n_\bullet$ be such that $(\lambda^n,\mu^n_\bullet)$ is a decomposition of μ^n. Since $\Delta_{\mathbf{p}}(\Delta(K))$ is a compact set in the weak topology, we may assume (by considering a subsequence if needed) that $\forall\mathcal{P}\in\mathcal{K} : \{\mu^n_{\mathcal{P}}\}_{n\in\mathbb{N}}$ converges weakly to $\mu_{\mathcal{P}}$.

To conclude that $\lambda\in\Lambda(\mu)$ we have to prove that $(\lambda,\mu_\bullet)$ is a decomposition of μ. Clearly, λ belongs to $\Delta(\mathcal{K})$, since this set is closed. It is also obvious that $\sum_{\mathcal{P}}\lambda(\mathcal{P})\mu_{\mathcal{P}} = \lim_{n\to\infty}\sum_{\mathcal{P}}\lambda^n(\mathcal{P})\mu^n_{\mathcal{P}} = \mu$. Finally, we prove in Lemma 10.3 that the set $\{\mu\in\Delta_{\mathbf{p}}(\Delta(K)) : \mu\succeq[\mathcal{P}]\}$ is closed in the weak topology. Since $\mu^n_{\mathcal{P}}$ belongs to this set, for every n we have $\mu_{\mathcal{P}}\succeq[\mathcal{P}]$ and $(\lambda,\mu_\bullet)$ is a decomposition of μ as required. ■

Lemma 10.3 *For $\mu'\in\Delta_{\mathbf{p}}(\Delta(K))$, let $R_{\mu'}$ denote $R_{\mu'} := \{\mu\in\Delta_{\mathbf{p}}(\Delta(K)) : \mu\succeq\mu'\}$. Then, R_μ is a convex subset of $\Delta_{\mathbf{p}}(\Delta(K))$, which is closed in the weak topology.*

Proof: Let $\mu^i\in R_{\mu'}$, $r^i\geq 0$, $i=0,1$ be such that $1=\sum_i r^i$, $\mu=\sum_i r^i\mu^i$. Since $\mu^i\succeq\mu'$, there exists a random vector (X^i_1,X^i_2) satisfying $E[X^i_2|X^i_1]=X^i_1$ where X^i_1 is μ'-distributed and X^i_2 is μ^i-distributed. We may assume that these two random vectors are defined on the same probability space, which also contains an independent random variable U taking the value i with probability $r^i, (i=0,1)$. The random vector (X_1,X_2) defined as $X^i_j := X^U_j$ satisfies the property that X_1 is μ'-distributed and X_2 is μ-distributed. Furthermore, $E[X_2|X_1]=X_1$, and thus $\mu\succeq\mu'$ as desired.

If $\{\mu^n\}_{n\in\mathbb{N}}$ is a sequence in $R_{\mu'}$ that weakly converges to μ, then for all continuous convex function ψ, $\tilde{\psi}(\mu^n) \geq \tilde{\psi}(\mu')$. Due to the weak convergence of $\{\mu^n\}_{n\in\mathbb{N}}$, we get $\tilde{\psi}(\mu) = \lim_{n\to\infty} \tilde{\psi}(\mu^n) \geq \tilde{\psi}(\mu')$. Since this holds for all continuous convex function ψ, we infer that $\mu \succeq \mu'$ as announced. ■

We are now ready to prove the main result of this section:

Theorem 10.4 *v is a Blackwell monotonic concave convex semicontinuous function on $(\Delta_{\mathbf{p}}(\Delta(K)) \times \Delta_{\mathbf{q}}(\Delta(L)))$ such that*

$$\forall \mathcal{P} \in \mathcal{K}, \forall \mathcal{Q} \in \mathcal{L} : v([\mathcal{P}], [\mathcal{Q}]) = V(\mathcal{P}, \mathcal{Q}). \tag{6}$$

Proof: We first prove Eq. (6). We demonstrate in Lemma 10.1 that the law of any partition $\mathcal{P}'$ is an extreme point of $\Delta_{\mathbf{p}}(S)$. Therefore, if $(\lambda, \mu_\bullet)$ is a decomposition of $[\mathcal{P}']$, then for all $\mathcal{P}$ such that $\lambda(\mathcal{P}) > 0$: $\mu_{\mathcal{P}} = [\mathcal{P}'] \succeq [\mathcal{P}]$. A similar argument indicates that if $(\xi, \nu_\bullet)$ is a decomposition of $[\mathcal{Q}']$, then $\forall \mathcal{Q}$ such that $\xi(\mathcal{Q}) > 0$, $\nu_{\mathcal{Q}} = [\mathcal{Q}'] \succeq [\mathcal{Q}]$. Therefore, if $(\lambda^*, \mu^*_\bullet)$ denotes the decomposition of $[\mathcal{P}']$ defined by λ^* the dirac measure on $\mathcal{P}'$ and $\mu^*_{\mathcal{P}} := [\mathcal{P}]$, we get:

$$\begin{array}{rcl} v([\mathcal{P}'], [\mathcal{Q}']) & \geq & \inf_{(\xi,\nu_\bullet)} \sum_{\mathcal{P}\in\mathcal{K}} \sum_{\mathcal{Q}\in\mathcal{L}} \lambda^*(\mathcal{P})\xi(\mathcal{Q})V(\mathcal{P}, \mathcal{Q}) \\ & = & \inf_{(\xi,\nu_\bullet)} \sum_{\mathcal{Q}\in\mathcal{L}} \xi(\mathcal{Q})V(\mathcal{P}', \mathcal{Q}) \\ & \geq & V(\mathcal{P}', \mathcal{Q}'), \end{array}$$

since $\mathcal{Q} \to V(\mathcal{P}, \mathcal{Q})$ is decreasing with respect to the order of refinement and $[\mathcal{Q}'] \succeq [\mathcal{Q}]$, ξ-almost surely. A dual argument indicates that $v([\mathcal{P}'], [\mathcal{Q}']) \leq V(\mathcal{P}', \mathcal{Q}')$.

We next prove that $\mu \to v(\mu, \nu)$ is concave. Let μ be a convex combination of μ^i, $i = 0, 1$: $\mu = r_1\mu^1 + r_0\mu^0$, with $r_i \geq 0$ with $r_1 + r_0 = 1$. Let $(\lambda^i, \mu^i_\bullet)$ be a decomposition of μ^i, $i = 0, 1$ such that

$$v(\mu^i, \nu) = \inf_{\xi,\in\Xi(\nu)} \sum_{\mathcal{P}\in\mathcal{K}} \sum_{\mathcal{Q}\in\mathcal{L}} \lambda^i(\mathcal{P})\xi(\mathcal{Q})V(\mathcal{P}, \mathcal{Q}).$$

Then, $(\lambda, \mu_\bullet)$ is a decomposition of μ, where $\lambda := \sum_i r_i\lambda^i$ and $\mu_\bullet$ is defined by $\mu_{\mathcal{P}} := (\sum_i r_i\lambda^i(\mathcal{P})\mu^i_{\mathcal{P}})/\lambda(\mathcal{P})$ if $\lambda(\mathcal{P}) > 0$, $\mu_{\mathcal{P}} := [\mathcal{P}]$ otherwise. Indeed, since $\Delta(\mathcal{K})$ is convex, $\lambda \in \Delta(\mathcal{K})$. Next $\sum_{\mathcal{P}\in\mathcal{K}} \lambda(\mathcal{P})\mu_{\mathcal{P}} = \sum_i r_i \left(\sum_{\mathcal{P}\in\mathcal{K}} \lambda^i(\mathcal{P})\mu^1_{\mathcal{P}}\right) = \mu$. Finally $\mu_{\mathcal{P}} \succeq [\mathcal{P}]$ since the set $\{\mu \in \Delta_{\mathbf{p}}(\Delta(K)) : \mu \succeq [\mathcal{P}]\}$ is convex, as proved in Lemma 10.1.

As a consequence,

$$\begin{array}{rcl} v(\mu,\nu) & \geq & \inf_{\xi,\in\Xi(\nu)} \sum_{\mathcal{P},\mathcal{Q}} \lambda(\mathcal{P})\xi(\mathcal{Q})V(\mathcal{P},\mathcal{Q}) \\ & = & \inf_{\xi,\in\Xi(\nu)} \sum_i r_i \left(\sum_{\mathcal{P},\mathcal{Q}} \lambda^1(\mathcal{P})\xi(\mathcal{Q})V(\mathcal{P},\mathcal{Q})\right) \\ & \geq & \sum_i r_i \left(\inf_{\xi,\in\Xi(\nu)} \sum_{\mathcal{P},\mathcal{Q}} \lambda^1(\mathcal{P})\xi(\mathcal{Q})V(\mathcal{P},\mathcal{Q})\right) \\ & = & \sum_i r_i v(\mu^i,\nu). \end{array}$$

We now prove that $\mu \to v(\mu,\nu)$ is Blackwell increasing. If $\mu \preceq \mu'$ and if $(\lambda,\mu_\bullet)$ is a decomposition of μ, if (X,X') is a random vector with respective marginals μ and μ', with $E[X'|X] = X$, if δ_x denotes the conditional law of $X'|X = x$, then one way to engender a μ' distributed random variable X' is as follows: Select first $\mathcal{P} \in \mathcal{K}$ with the lottery λ. Select then X with the lottery $\mu_{\mathcal{P}}$. Finally select X' with the lottery δ_X. We get in this way that $\mu' = \sum_{t\in T} \lambda(\mathcal{P})\mu'_{\mathcal{P}}$ where $\mu'_{\mathcal{P}}$ denotes the law of X' when X is selected with probability $\mu_{\mathcal{P}}$ and X' is then selected with probability δ_X. Since $[\mathcal{P}] \preceq \mu_{\mathcal{P}} \preceq \mu'_{\mathcal{P}}$, we infer that $(\lambda,\mu'_\bullet)$ is a decomposition of μ'. Therefore $v(\mu',\nu) \geq \inf_{\xi\in\Xi(\nu)} \sum_{\mathcal{P},\mathcal{Q}} \lambda(\mathcal{P})\xi(\mathcal{Q})V(\mathcal{P},\mathcal{Q})$. Since this holds for all decomposition $(\lambda,\mu_\bullet)$ of μ, we conclude as announced that $v(\mu',\nu) \geq v(\mu,\nu)$.

We next argue that $\mu \to v(\mu,\nu)$ is lsc. It needs to be shown that the set $H := \{\mu : v(\mu,\nu) \geq a\}$ is closed in the weak topology on $\Delta_{\mathbf{p}}(\Delta(K))$. Let $\{\mu^n\}$ be a sequence in H that weakly converges to μ, and let $(\lambda^n,\mu^n_\bullet)$ be a decomposition of μ^n such that $v(\mu^n,\nu) = \inf_{(\xi,\nu_\bullet)} \sum_{\mathcal{P},\mathcal{Q}} \lambda^n(\mathcal{P})\xi(\mathcal{Q})V(\mathcal{P},\mathcal{Q}) \geq a$. Considering a subsequence, we may assume that $\lambda^n \to \lambda$ and $\forall \mathcal{P} \in \mathcal{K} : \mu^n_{\mathcal{P}} \to \mu_{\mathcal{P}}$, as n goes to ∞. Then, $\sum_{\mathcal{P}\in\mathcal{K}} \lambda^n(\mathcal{P})\mu^n_{\mathcal{P}} = \mu^n \to \mu = \sum_{\mathcal{P}\in\mathcal{K}} \lambda(\mathcal{P})\mu_{\mathcal{P}}$. Since $\{\mu \in \Delta_{\mathbf{p}}(\Delta(K)) : \mu \succeq [\mathcal{P}]\}$ is closed in the weak topology (see Lemma 10.3), we conclude that $(\lambda,\mu_\bullet)$ is a decomposition of μ. Therefore $v(\mu,\nu) \geq \inf_{\xi\in\Xi(\nu)} \sum_{\mathcal{P},\mathcal{Q}} \lambda(\mathcal{P})\xi(\mathcal{Q})V(\mathcal{P},\mathcal{Q})$. Now, for all decompositions $(\xi,\nu_\bullet)$ of ν:

$$\sum_{\mathcal{P},\mathcal{Q}} \lambda(\mathcal{P})\xi(\mathcal{Q})V(\mathcal{P},\mathcal{Q}) = \lim_{n\to\infty} \sum_{\mathcal{P},\mathcal{Q}} \lambda^n(\mathcal{P})\xi(\mathcal{Q})V(\mathcal{P},\mathcal{Q}) \geq v(\mu^n,\nu) \geq a.$$

Therefore, $v(\mu,\nu) \geq a$ and $\mu \in H$, as desired.

To conclude the proof notice that Theorem 10.2 allows us to interchange sup and inf in the definition of $v(\mu,\nu)$. The argument presented above can be adapted to prove that $\nu \to v(\mu,\nu)$ is convex, usc and Blackwell decreasing and to complete the proof. ■

11 Concluding remarks

The payoff functions constructed above need not be bounded. It is clear however that if we restrict attention to bounded games, the value function is still concave-convex and Blackwell monotonic. It is easy to prove that it is also Lipschitz. The inverse direction is not straightforward.

We leave open the question of what happens when the components of the state on which the players have some information fail to be independent. Phrased differently, suppose the state is chosen from a set K, then each player gets a state-specific signal. In the lack of independence, signals are correlated. Thus, given a signal a player can obtain some information about the signal of his opponent. In this situation the notion of monotonicity is unclear, and the duality method is not well understood.

12 References

Aumann, R. J. and M. Maschler, with the collaboration of R. Stearns (1995), *Repeated games with incomplete information*, M.I.T. Press.

Azrieli, Y. and E. Lehrer (2004), "The value of a stochastic information structure," to appear in *Games and Economic Behavior*.

Bassan, B., O. Gossner, M. Scarsini and S. Zamir (2003), "Positive value of information in games," *International Journal of Game Theory*, **32**, 17-31.

Blackwell, D. (1953), "Equivalent Comparison of Experiments," *Ann. Math. Stat.*, **24**, 265-272.

De Meyer, B., (1996), "Repeated games and partial differential equations," *Mathematics of Operations Research*, **21**, 209-236.

De Meyer, B., and A. Marino, (2005), "Duality and optimal strategies in the finitely repeated zero-sum games with incomplete information on both sides," *Mathematics of Operations Research*.

De Meyer, B., and Moussa-Saley, (2003), "On the origin of the Brownian motion in finance," *International Journal of Game Theory*, **31**, 285-319.

De Meyer, B. and D. Rosenberg, (1999), "Cav u and the dual game," *Mathematics of Operations Research*, **24**, 619-626.

Gilboa, I. and E. Lehrer (1991), "The Value of Information - An Axiomatic Approach," *Journal of Mathematical Economics*, **20**, 443-459.

Gossner, O. (2006), "Ability and knowledge," mimeo.

Gossner, O. and J.-F. Mertens (2001),"The Value of Information in Zero-Sum Games," mimeo.

Hirshleifer, J. (1971), "The private and social value of information and the reward to inventive activity," *American Economic Review*, **61**, 561-574.

Kamien, M., Y. Taumann and S. Zamir (1990), "On the value of information in a strategic conflict," *Games and Economic Behavior*, **2**, 129-153.

Laraki, R. (2002), "Repeated games with lack of information on one side: the dual differential approach," *Mathematics of Operations Research*, **27**, 419-440.

Lehrer, E. and D. Rosenberg (2006), "What restrictions do Bayesian games impose on the value of information?" *Journal of Mathematical Economics*, **42**, 343-357.

Lehrer, E., D. Rosenberg and E. Shmaya (2006) "Signalling and mediation in games with common interest," mimeo.

Mertens, J.F. and S. Zamir (1971), "The value of two player zero-sum repeated games with lack of information on both sides," *International Journal of Game Theory*, **1**, 39-64.

Mertens, J.F. and S. Zamir (1980), "Minmax and Maxmin of repeated games with incomplete information," *International Journal of Game Theory*, **9**, 201-215.

Neyman, A. (1991) "The positive value of information," *Games and Economic Behavior*, **3**, 350-355.

Rockaffellar, R.T. (1970), *Convex Analysis*, Princeton, New Jersey, Princeton University Press.

Sorin, S. (2002), *A first course in zero-sum repeated games*, Springer.