MATHEMATICS OF OPERATIONS RESEARCH
Vol. 35, No. 4, November 2010, pp. 851–863
ISSN 0364-765X | EISSN 1526-5471 | 10 | 3504 | 0851

DOI 10.1287/moor.1100.0467

Evaluating Information in Zero-Sum Games with Incomplete Information on Both Sides

Bernard De Meyer
Paris School of Economics-Université de Paris, Maison des Sciences Economiques, 75013 Paris, France, demeyer@univ-paris1.fr

Ehud Lehrer
School of Mathematical Sciences, Tel Aviv University, 69978 Tel Aviv, Israel, lehrer@post.tau.ac.il, www.math.tau.ac.il/~lehrer

Dinah Rosenberg
HEC Paris, 78351 Jouy-en-Josas, France, rosenberg@hec.fr

We study zero-sum games with incomplete information and analyze the impact that the information players receive has on the payoffs. It turns out that the functions that measure the value of information share two properties. The first is Blackwell monotonicity, which means that each player gains from knowing more. The second is concavity on the space of conditional probabilities. We prove that any function satisfying these two properties is the value function of a zero-sum game.

Key words: value-of-information function; zero-sum game; game with incomplete information; Blackwell monotonicity
MSC2000 subject classification: Primary: 91A05; secondary: 91A28
OR/MS subject classification: Primary: noncooperative games; secondary: statistics, information, convexity
History: Received May 13, 2009; revised April 29, 2010. Published online in *Articles in Advance* November 10, 2010.

1. Introduction. In strategic interactions some relevant aspects of the environment might be imperfectly known to players. However, 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.

We chose to analyze this question in the set-up 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 [4] 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 [12]. It is well known that this property does not extend to nonzero-sum games, and various attempts have been made to understand when it might be extended (see, for instance, Hirshleifer [13], Bassan et al. [3], Kamien et al. [14], Neyman [19], Lehrer et al. [17]).

The second approach is to study the impact of an information structure on the outcome of a specific interaction. A typical question in this line is whether the outcome depends in any particular and discernable way on the information. Because 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 value-of-information function that 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 [10] treated deterministic information structures and characterized those functions that are value-of-information functions of *one-player* decision problems. Blackwell's [4] 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 [2].

Lehrer and Rosenberg [16] studied the nature of value-of-information functions in two cases: (a) one-sided deterministic information: the game depends on a state of nature that 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 that is equally known to both players. In case (a), the more refined the partition the higher the value. Lehrer and Rosenberg [16] showed that *any* function, defined over partitions, that 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 [1], Mertens and Zamir [18]).

Formally, the information structure in a two-player interaction is characterized by a pair of probability transitions (ρ_1, ρ_2) where ρ_i is a probability transition between the set of types of player i and his set of messages. Each player is informed only of his message, which therefore endows him with some partial information on his own type (and no information at all on his opponent's type). Then a zero-sum game, whose payoffs depend on the players' types, is played. The value of this game depends on the information structure, (ρ_1, ρ_2). 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 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 if 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. Obviously, one can never observe the complete relation between the outcome of an interaction and all possible information structures. However, it is plausible to assume that different information structures are used in different times in the same or similar interaction. This might happen, for instance, due to information acquisition. An outside observer who knows the relevant information structure and gets to know the outcomes of the interaction will then be able to tell whether his finding is consistent with the Bayesian model.

One may wonder how an outside observer could know the information structure without knowing the payoffs. The information structure determines the knowledge the players have about the state of nature and in this respect is "objective." The payoffs, on the other hand, are expressed in vN-M utility terms and are in nature "subjective." The outside observer might have an access to the objective ingredient of the game and not the subjective ones.

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 [11]). The proof method uses duality technique inspired by Fenchel conjugate of convex functions (see Rockafellar [20]). The duality technique has been introduced by De Meyer [6]. For any game with incomplete information on one side, De Meyer defined a dual game in which the value is the Fenchel conjugate of the value of the initial game. The subject of dual games has been widely investigated (see De Meyer and Rosenberg [9], De Meyer and Marino [7], De Meyer and Moussa-Saley [8], Laraki [15], Sorin [21]).

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 we 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. A zero-sum game with incomplete information is described by a pair of type-sets, K for player 1 and L for player 2, action spaces A and B for player 1 and 2, respectively, and a payoff function g: $A \times B \times K \times L \to \mathbb{R}$.

The sets K and L are assumed to be finite, and types $(k, l) \in K \times L$ are randomly chosen with a given independent distribution $\mathbf{p} \otimes \mathbf{q}$. Moreover, 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 a measurable space $(M_i, \mathscr{M}_i)$. The signal is randomly chosen according to the distribution $\rho_i = \rho_i(h)$, where $h = k, l$, depending on i. ρ_1 (resp. ρ_2) is thus a transition probability from K (resp. L) to $(M_1, \mathscr{M}_1)$ (resp. $(M_2, \mathscr{M}_2)$). $I_i := (M_i, \mathscr{M}_i, \rho_i)$ is called the information structure of player i.

2.2. Strategies and payoffs. The action spaces might be infinite, and we therefore have to assume that they are endowed with σ-algebras $\mathscr{A}$ and $\mathscr{B}$.[1] Upon receiving the messages player 1 takes an action $a \in A$, and player 2 takes an action $b \in B$. Players are allowed to make mixed moves, and a strategy of player 1 is thus a transition probability σ: $M_1 \to A$ while a strategy of player 2 is a transition probability τ: $M_2 \to B$. In other words, for all m_1, $\sigma(m_1)$ is a probability on $(A, \mathscr{A})$ according to which player 1 chooses his action if he receives the message m_1. Furthermore, for all $A' \in \mathscr{A}$, the map $m_1 \to \sigma(m_1)[A']$ is measurable. A similar condition applies to τ.

When the types are (k, l) and the actions are (a, b), player 2 pays player 1 $g(a, b, k, l)$. For all k, l, the payoff map $(a, b) \to g(a, b, k, l) \in \mathbb{R}$ is assumed to be $\mathscr{A} \otimes \mathscr{B}$-measurable. A strategy σ (resp. τ) induces, together with $\mathbf{p}, \rho_1$ (resp. $\mathbf{q}, \rho_2$), 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 associated with a pair of strategies (σ, τ) is the expected payoff, namely $g(\sigma, \tau) := \mathrm{E}_{\pi_{\sigma, \tau}}[g(a, b, k, l)]$. However, because 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 the payoffs it induces are bounded from below for any strategy of player 2. Formally, σ is *admissible* if there exists $m \in \mathbb{R}$ such that, $\forall\, l, b$, $\mathrm{E}_{\pi_\sigma}[|g(k, a, l, b))|] < \infty$, and $\mathrm{E}_{\pi_\sigma}[g(k, a, l, b)] \geq m$. Admissible strategies of player 2 are defined in a similar fashion. The sets of admissible strategies are denoted Σ_{ad} and T_{ad}. Note that $g(\sigma, \tau)$ is well defined if either $\sigma \in \Sigma_{ad}$ or $\tau \in T_{ad}$.

A game Γ is represented by $\Gamma = \langle A, \mathscr{A}, B, \mathscr{B}, g\rangle$ with the above measurability assumptions.

2.3. Standard information structures and value of information functions. 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 $P_k(m_1)$. Because the message m_1 is random,[2] $P(m_1) := (P_k(m_1))_{k\in K} \in \Delta(K)$ is a random variable, and the information structure thus induces (ex ante) a distribution over $\Delta(K)$. We refer to this distribution as the *law of the information structure* and denote it by $[I_1]$. Denote by $\Delta_p(\Delta(K))$ the set of all probability measures μ over $\Delta(K)$ whose average value is $\mathbf{p}$ (i.e., $\mathrm{E}_\mu[P] = \mathbf{p}$). Note that $[I_1]$ belongs to this set.

Different values of m_1 could lead to the same conditional probability $P(m_1)$. There is thus apparently more information embedded in m_1 than in $P(m_1)$ alone. However, this additional information is irrelevant for predicting k, because k and m_1 are independent given $P(m_1)$. Therefore, after computing $P(m_1)$, player 1 might forget the value of m_1 and play in a game Γ a strategy that just depends on $P(m_1)$. Indeed, consider an arbitrary strategy σ and the probability Π_σ it induces jointly with $\mathbf{p}, \rho_1$ on $M_1 \times A \times K$. Denote $\bar{\sigma}(P)$ the conditional probability on A given $P(m_1) = P$.[3] Denote next $\bar{\sigma}$ the strategy that consists in playing $\bar{\sigma}(P(m_1))$ upon receiving

[1] We further make the technical hypothesis that $(A, \mathscr{A})$ and $(B, \mathscr{B})$ are standard Borel. A probability space $(A, \mathscr{A})$ is standard Borel if there is a one-to-one bimeasurable map between $(A, \mathscr{A})$ and a Borelian subset of $\mathbb{R}$. This hypothesis ensures 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 then be defined as a measurable map from $([0, 1] \times M_1, \mathscr{B}_{[0,1]} \otimes \mathscr{M}_1)$ to $(A, \mathscr{A})$, where $\mathscr{B}_{[0,1]}$ 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 $\sigma(u, m_1)$.

[2] For any measurable set M, we denote by $\Delta(M)$ the set of probability distributions over M.

[3] Because $(A, \mathscr{A})$ is Borel standard, such a transition probability from $\Delta(K)$ to A exists.

the message m_1. It appears that the marginals π_σ and $\pi_{\bar\sigma}$ of Π_σ and $\Pi_{\bar\sigma}$ on $K \times A$ coincide.[4] Because the payoff in Γ depends just on π_σ, we infer that for every strategy τ of player 2, $g(\sigma,\tau) = g(\bar\sigma,\tau)$.

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 $m'_1 := P(m_1)$ when the message under I_1 is 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, because 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 is be denoted Q.

In a game Γ, for each strategy σ of player 1 under the information structure I_1, the above defined strategy $\bar\sigma$ can be used under information structure I'_1 because it depends only on $P(m_1) = m'_1$. Because we showed that σ and $\bar\sigma$ yield the same expected payoff in Γ, it implies that the highest payoff that player 1 can guarantee in Γ can also be achieved under the standard information structure I'_1. A similar argument can be used for player 2. This implies that the values of Γ and of the game with standard information structure I'_1, I'_2 are the same. From this point on, we therefore focus on games with standard information structures, understanding that the results on the values of such games apply to general games with incomplete information

2.4. Value-of-information function. The game with the information structures 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

$$V_\Gamma(I_1, I_2) = \sup_{\sigma\in\Sigma_{ad}} \inf_{\tau\in T} g(\sigma,\tau) = \inf_{\tau\in T_{ad}} \sup_{\sigma\in\Sigma} g(\sigma,\tau).$$

Our goal is to study the relationship between the value of the game and the information structures. We showed in the above subsection that if (I'_1, I'_2) is the standard information structure associated to (I_1, I_2) then, whenever the value exists, we have $V_\Gamma(I_1, I_2) = V_\Gamma(I'_1, I'_2)$. This means that the value $V(I_1, I_2)$ depends only on the law induced by the information structure on $\Delta(K)$. We can therefore define a value-of-information function as a function over such laws.

Definition 2.1. A function $v\colon \Delta_p(\Delta(K)) \times \Delta_q(\Delta(L)) \to \mathbb{R}$ is a value-of-information function if there is a game Γ such that for any standard information structure (I_1, I_2) the value $V_\Gamma(I_1, I_2)$ exists and $V(I_1, I_2) = v([I_1],[I_2])$.

The purpose of this paper is to characterize the value-of-information functions when we let information structures vary, while A, B, K, L, and g remain fixed throughout.

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 $\Delta_p(\Delta(K))$. We 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 $E(X_1 \mid X_2) = X_2$. A similar definition holds for distributions in $\Delta_q(\Delta(L))$.

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

Blackwell [4] 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.

We use this notion here to define Blackwell monotonicity in our context.

[4] To see this, observe that once (k, m_1) has been chosen by nature, a is selected by the lottery $\sigma(m_1)$, that is independently of k. So, for an arbitrary function T of (a,k): $E_{\Pi_\sigma}[T \mid m_1] = \sum_{k_0} P_{k_0}(m_1) E_{\Pi_\sigma}[T(a,k_0) \mid m_1]$. It follows that

$$\begin{aligned} E_{\Pi_\sigma}[T \mid P(m_1)] &= E_{\Pi_\sigma}[E_{\Pi_\sigma}[T \mid m_1] \mid P(m_1)] \\ &= \sum_{k_0} P_{k_0}(m_1) E_{\Pi_\sigma}[T(a,k_0) \mid P(m_1)] \\ &= \sum_{k_0} P_{k_0}(m_1) E_{\bar\sigma(P(m_1))}[T(a,k_0)] \\ &= E_{\Pi_{\bar\sigma}}[T \mid P(m_1)], \end{aligned}$$

and thus $E_{\Pi_\sigma}[T] = E_{\Pi_{\bar\sigma}}[T]$ for any function T of (a,k). This indicates that $\pi_\sigma = \pi_{\bar\sigma}$.

Definition 3.2. A function v: $\Delta_p(\Delta(K)) \times \Delta_q(\Delta(L)) \to \mathbb{R}$ is *Blackwell monotonic* if
(i) $\forall\,\nu$, the map $\mu \to v(\mu, \nu)$ is increasing with respect to $\succeq$,
(ii) $\forall\,\mu$, the map $\nu \to v(\mu, \nu)$ is decreasing with respect to $\succeq$.

Another property of function that often has been used in the context of repeated games with incomplete information is concavity-convexity (see Aumann and Maschler [1]).

Definition 3.3. A function v: $\Delta_p(\Delta(K)) \times \Delta_q(\Delta(L)) \to \mathbb{R}$ is *concave-convex, semicontinuous* if
(i) $\forall\,\nu$, the map $\mu \to v(\mu, \nu)$ is concave, upper semicontinuous (usc) with respect to the weak topology on $\Delta_p(\Delta(K))$,
(ii) $\forall\,\mu$, the map $\nu \to v(\mu, \nu)$ is convex, lower semicontinuous (lsc) with respect to the weak topology on $\Delta_q(\Delta(L))$.

Note that we have here two notions, Blackwell monotonicity and concavity-convexity. Both reflect in different ways the fact that information is valuable in zero-sum games.

Let v be a function on $\Delta_p(\Delta(K)) \times \Delta_q(\Delta(L))$ that is supposed to be a value of information function. Concavity with respect to μ compares two situations. In the first, player 2 knows that player 1's information is given by $\lambda\mu + (1-\lambda)\mu'$, meaning that with probability λ the law of the conditional probability is μ, and with probability $1-\lambda$ it is μ', but player 2 does not know the outcome of the lottery. In this situation the value of the underlying game is $v(\lambda\mu + (1-\lambda)\mu', \nu)$. In the second situation, player 2 knows the outcome of the lottery so that the value of the game is $\lambda v(\mu, \nu) + (1-\lambda) v(\mu', \nu)$. Player 2 is better informed in the second situation, and hence the value is smaller. This is the concavity condition. Concavity therefore expresses the fact that having more information about the other player's signal is valuable.

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 $\mathrm{E}[X_1 \mid X_2] = X_2$ 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, and the value of the game is therefore higher.[5] Therefore, Blackwell monotonicity expresses the fact that having some more precise information on one's own type is valuable.

The properties of concavity-convexity and Blackwell monotonicity are logically independent. To illustrate this fact consider the function f defined on $\Delta_p(\Delta(K))$ as $f(\mu) = -\mathrm{E}_\mu[\|p - \mathbf{p}\|]$. This function is linear in μ and, in particular, concave. Suppose now that μ_2 is a Dirac mass at $\mathbf{p}$ and $\mu_1 \in \Delta_p(\Delta(K))$ puts no weight at $\mathbf{p}$. Clearly, μ_1 is more informative than μ_2 and at the same time $f(\mu_1) < 0 = f(\mu_2)$. It implies that f is not Blackwell monotonic.

3.2. Main theorem. The main result of this paper, proved in the next sections, is as follows.

Theorem 3.1. *The map v: $\Delta_p(\Delta(K)) \times \Delta_q(\Delta(L)) \to \mathbb{R}$ is a value-of-information function iff it is concave-convex, semicontinuous, and Blackwell-monotonic.*

Blackwell proved that monotonicity is a necessary condition for a function to be the value of a one-player optimization problem. Because 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 [16] proved that this is also the case in games with incomplete information on one side with deterministic information.

As in Lehrer and Rosenberg [16], Theorem 3.1 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, one implication of Theorem 3.1 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 [16]. Admittedly, such violations can hardly be observed in reality because it is very rare to observe outcomes of the same zero-sum interactions under various (if not all) possible information structures.

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

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

When ρ_1 and ρ_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 therefore be modeled as a pair of partitions $\mathcal{P}$, $\mathcal{Q}$, respectively, of K and L, with the interpretation that if the true state is k, l, then 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. 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. We denote by $\mathcal{K}$ and $\mathcal{L}$, respectively, the sets of partitions of K and L.

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 $\mathcal{P}$. Blackwell monotonicity in this case is translated to the following. 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 3.2. *Let v be a function on pairs of partitions of K and L. It is a value-of-information function iff it is Blackwell monotonic. Moreover, it can be then obtained as the value of a finite game* (*i.e., A and B are finite*).

This theorem differs from Theorem 3.1 in two respects. First, the games are finite. This finiteness can be derived from Theorem 3.1, taking into account that there are only finitely many partitions. Second, the condition of concavity-convexity of the function is missing here. The reason is that any function over partitions that is Blackwell monotonic can be extended to a function over laws of posterior probabilities that is concave-convex, semicontinuous, and Blackwell monotonic (this result is proved in §10). Thus, Theorem 3.2 derives from Theorem 3.1 and extends (Lehrer and Rosenberg [16]) 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 set-up 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 $\mathrm{E}_\mu[\psi(P)]\geq f(\mu)$ for every law μ on $\Delta(K)$.

The second part of the theorem characterizes functions that have the concave-convex property and are also Blackwell monotonic. It gives a clearer sense of the difference between both notions. For a function to be Blackwell monotonic in addition, 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 is convex.

To formally state the result, we need a few notations. Ψ^0 and Ψ^1 denote, respectively, the sets of continuous and continuously differentiable functions from $\Delta(K)$ to $\mathbb{R}$. The set of convex functions in Ψ^i is denoted $\Psi^{i,\,\mathrm{vex}}$, $i=0,1$. Similarly, Φ^0 and Φ^1 denote, respectively, the sets of continuous and continuously differentiable functions from $\Delta(L)$ to $\mathbb{R}$. $\Phi^{i,\,\mathrm{cav}}$ is the set of concave functions in Φ^i.

For a given function $\psi\in\Psi^0$ and a law μ in $\Delta_p(\Delta(K))$, we denote by $\tilde{\psi}(\mu)$ the expectation of ψ with respect to μ (i.e., $\int\psi(p)\,d\mu(p)$). Similarly, for $\nu\in\Delta_q(\Delta(L))$ and $\phi\in\Phi^0$, $\tilde{\phi}(\nu)$ denotes $\int\phi(q)\,d\nu(q)$.

For $f\colon\Delta_p(\Delta(K))\to\mathbb{R}$, we define Ψ^i_f (resp. $\Psi^{i,\,\mathrm{vex}}_f$) to be the set of functions $\psi\in\Psi^i$ (resp. $\psi\in\Psi^{i,\,\mathrm{vex}}$) such that $\forall\,\mu\in\Delta_p(\Delta(K))$, $f(\mu)\leq\tilde{\psi}(\mu)$. Similarly, for $f\colon\Delta_q(\Delta(L))\to\mathbb{R}$, Φ^i_f, and $\Phi^{i,\,\mathrm{cav}}_f$ denotes the set of $\phi\in\Phi^i$, (resp. $\Phi^{i,\,\mathrm{cav}}$) such that $\forall\,\nu\in\Delta_q(\Delta(L))$, $\tilde{\phi}(\nu)\leq f(\nu)$.

THEOREM 4.1. (i) *If a function $f\colon\Delta_p(\Delta(K))\to\mathbb{R}$ is concave and usc with respect to the weak topology, then $\forall\,\mu$, $f(\mu)=\inf_{\psi\in\Psi^0_f}\tilde{\psi}(\mu)$.*

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

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

This theorem will be proved in §7.

5. The conditions of Theorem 3.1 are necessary. In this section we show that any value-of-information function is concave-convex and Blackwell monotonic.

THEOREM 5.1. *If v is a value-of-information function, then v is concave-convex, semicontinuous, and Blackwell monotonic.*

PROOF. Indeed, one can compute the expected payoff $g(\sigma, \tau)$ given a pair of strategies (σ, τ) with $\tau \in T_{ad}$. This is

$$g(\sigma, \tau) = \mathrm{E}_{[I_1]}\left[\sum_k P_k \, \mathrm{E}_{\sigma(P),\, \pi_\tau}[g(a, b, k, l)]\right],$$

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

$$\sup_\sigma g(\sigma, \tau) = \mathrm{E}_{[I_1]}\left[\sup_{\sigma(P)} \sum_k P_k \, \mathrm{E}_{\sigma(P),\, \pi_\tau}[g(a, b, k, l)]\right] = \mathrm{E}_{[I_1]}[\psi(P)],$$

where $\psi(P) := \sup_{a \in A} \sum_k P_k \, \mathrm{E}_{\pi_\tau}[g(a, b, k, l)]$. As supremum of linear functions of P, ψ is convex lsc on $\Delta(K)$. Note that because τ is admissible, ψ is bounded from above on $\Delta(K)$ and ϕ is thus continuous on $\Delta(K)$.

$\sup_\sigma g(\sigma, \tau)$ is thus the expectation of a convex function in P, and 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, because ψ is continuous, we also proved that $\sup_\sigma g(\sigma, \tau)$ is linear and continuous in $[I_1]$ in the weak topology on $\Delta_p(\Delta(K))$. As the infimum of linear weakly continuous function of $[I_1]$, $v([I_1], [I_2])$ is concave and usc for the weak topology on $\Delta_p(\Delta(K))$. $\square$

6. Sketch of the proof of Theorem 3.1. For an arbitrary concave-convex, Blackwell monotonic, and semicontinuous function v, we construct a game Γ such that $V_\Gamma(I_1, I_2) = v([I_1], [I_2]),\ \forall I_1, I_2$. We fix the function v and the laws μ and ν on $\Delta_p(\Delta(K))$ and $\Delta_q(\Delta(L))$.

The construction is based on two observations. The first is an implication of the structure theorem. Assume that player 2 is asked to choose a function $\psi \in \Psi^{1,\,\mathrm{vex}}_{v(\cdot,\,\nu)}$ (interpreted as a payoff function to player 1) with the restriction that when the signal of player 1 is P, the payoff is $\psi(P)$. The value of the corresponding game is, by the structure theorem, $\inf_{\psi \in \Psi^{1,\,\mathrm{vex}}_{v}(\cdot,\,\nu)} \tilde{\psi}(\nu) = v(\mu, \nu)$.

The second observation is needed to construct a game that induces $\psi(P)$ as a payoff for any given convex function ψ. For any convex and differentiable function ψ and any point $\tilde{p} \in \Delta(K)$, one can consider the following function $H_\psi(P, \tilde{p})$ of $P, \tilde{p} \in \Delta(K)$ that is the affine function of P that is tangent to ψ at point $\tilde{p}$:

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

At fixed P, this function is maximal in $\tilde{p}$ when $P = \tilde{p}$ and its maximum is $\psi(P)$. In other words, because ψ is convex, it is the envelope of all linear functions that are below it.

This suggests the following construction. Once informed of P, player 1 selects $\tilde{p} \in \Delta(K)$. Player 2, after being informed of Q, selects $\psi \in \Psi^{1,\,\mathrm{vex}}_{v(\cdot,\,\nu)}$. The payoff is

$$\psi(\tilde{p}) + \left\langle \nabla\psi(\tilde{p}), P - \tilde{p} \right\rangle.$$

The game defined so far is not a well-defined incomplete information game simply because the set of player 2's strategies depends on ν. To overcome this problem, we note that if a function ψ belongs to $\Psi^{1,\,\mathrm{vex}}$, then for some large enough α (interpreted as a cost), $\psi + \alpha$ belongs to $\Psi^{1,\,\mathrm{vex}}_{v(\cdot,\,\nu)}$. This means that $\forall \mu \in \Delta_p(\Delta(K)),\ \tilde{\psi}(\mu) + \alpha \geq v(\mu, \nu)$. The cost α needs to be larger than

$$w(\psi, \nu) := \sup_{\tilde{\mu} \in \Delta_p(\Delta(K))} v(\tilde{\mu}, \nu) - \tilde{\psi}(\tilde{\mu}). \tag{2}$$

The idea is to let player 2 choose any function $\psi \in \Psi^{1,\,\mathrm{vex}}$ and to use $\psi + w(\psi, \nu)$ in the previous game. Note that w is the Fenchel transform of v. The payoff in this new game is thus

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

Note that the first two terms in this payoff function are affine in P, which is the necessary property for a function to be the payoff function of a game where player 1 is incompletely informed.

We show now that $w(\psi, \nu)$ can be expressed as the value of a game where player 2 is incompletely informed. The dual version of the structure theorem states that $v(\tilde{\mu}, \nu) = \sup_{\phi \in \Phi^{1,\text{cav}}_{v(\tilde{\mu}, \cdot)}} \tilde{\phi}(\nu)$. Introducing this in Equation (2), we get

$$w(\psi, \nu) = \sup_{(\tilde{\mu}, \phi) \in \mathcal{F}} \tilde{\phi}(\nu) - \tilde{\psi}(\tilde{\mu}),$$

where $\mathcal{F}$ is the set of pairs $(\tilde{\mu}, \phi)$ such that $\tilde{\mu} \in \Delta_p(\Delta(K))$ and $\phi \in \Phi^{1,\text{cav}}_{v(\tilde{\mu}, \cdot)}$. Because ϕ is concave, $\phi(Q) = \inf_{\tilde{q} \in \Delta(L)} \phi(\tilde{q}) + \langle \nabla\phi(\tilde{q}), Q - \tilde{q} \rangle$. Therefore, $w(\psi, \nu)$ is the value of the game where, once informed of P, player 1 selects $(\tilde{\mu}, \phi) \in \mathcal{F}$. On the other hand, after observing Q, player 2 chooses $\tilde{q} \in \Delta(L)$, and the payoff is $\phi(\tilde{q}) + \langle \nabla\phi(\tilde{q}), Q - \tilde{q} \rangle - \tilde{\psi}(\tilde{\mu})$.

Combining the two games together, we obtain the following game. Upon receiving their respective messages P and Q, player 1 selects $(\tilde{p}, (\tilde{\mu}, \phi)) \in \Delta(K) \times \mathcal{F}$, and player 2 selects $\tilde{q} \in \Delta(L)$ and $\psi \in \Psi^{1,\text{vex}}$. The payoff is then

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

An alternative way to describe this game is the following. $v(\mu, \nu)$ is the value of the game where, after being informed, player 2 selects $\tilde{q} \in \Delta(L)$, player 1 selects $\phi \in \Phi^{1,\text{cav}}_{v(\mu, \cdot)}$, and the payoff is $\phi(\tilde{q}) + \langle \nabla\phi(\tilde{q}), Q - \tilde{q} \rangle$. The action space $\Phi^{1,\text{cav}}_{v(\mu, \cdot)}$ of player 1 in this game depends on μ. This problem is solved as we let player 1 choose a fictitious information structure $\tilde{\mu}$ and use it to choose $\phi \in \Phi^{1,\text{cav}}_{v(\tilde{\mu}, \cdot)}$. However, to verify that the optimal choice of player 1 is indeed $\tilde{\mu} = \mu$, we impose a cost $\tilde{\psi}$ associated to each choice of $\tilde{\mu}$.

This construction is similar to the dual game introduced by De Meyer [5, 6]. He proposed to introduce, for each incomplete information game where the information on the state is a parameter of the game, a dual game in which the information becomes a choice of the informed player and the parameter of the game is the cost of choosing a particular information structure. De Meyer [5, 6] proved that the value of such a game is the Fenchel transform of the value of the original game. He also defined the bidual game in which the uninformed player chooses a cost function first and then the corresponding dual game is played. The Fenchel duality equality proves that the value of the bidual game is identical to the value of the dual game. Here we use a similar construction. Player 2 chooses a cost function $\psi \in \Psi^{1,\text{vex}}$, and player 1 chooses an information law $\tilde{\mu} \in \Delta_p(\Delta(K))$. Simultaneously, both players play the previous game: player 1 chooses $\phi \in \Phi^{1,\text{cav}}_{v(\tilde{\mu}, \cdot)}$ and $\tilde{p} \in \Delta(K)$, and player 2 chooses $\tilde{q} \in \Delta(L)$.

The payoff is

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

Note that concavity of ϕ and convexity of ψ imply that the payoff is maximum when $\tilde{p} = P$ and minimum when $\tilde{q} = Q$, and this minimax payoff is $\tilde{\phi}(\nu) + \tilde{\psi}(\mu) - \tilde{\psi}(\tilde{\mu})$. If $\tilde{\mu} \neq \mu$, then there is a choice of ψ that can make the payoff arbitrarily small so that the optimal choice of player 1 is $\tilde{\mu} = \mu$. The payoff is then $\tilde{\phi}(\nu)$, which is by definition at most $v(\tilde{\mu}, \nu) = v(\mu, \nu)$.

7. Proof of the structure Theorem 4.1. For $f\colon \Delta_p(\Delta(K)) \to \mathbb{R}$, we define Ψ^0_f (resp. Ψ^1_f) as the set of continuous (resp. continuously differentiable) functions $\psi\colon \Delta(K) \to \mathbb{R}$ such that $\forall \mu \in \Delta_p(\Delta(K))$, $f(\mu) \le \tilde{\psi}(\mu)$.

For $i = 1, 0$, $\Psi^{i,\text{vex}}_f$ denotes the set of convex functions $\psi \in \Psi^i_f$. Recall Theorem 4.1.

THEOREM 7.1. (i) *If a function $f\colon \Delta_p(\Delta(K)) \to \mathbb{R}$ is concave usc with respect to the weak topology, then* $\forall \mu$, $f(\mu) = \inf_{\psi \in \Psi^0_f} \tilde{\psi}(\mu)$.

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

(iii) *In this case, we also have* $\forall \mu$, $f(\mu) = \inf_{\psi \in \Psi^{1,\text{vex}}_f} \tilde{\psi}(\mu)$.

PROOF. We prove the first claim. Let f be a concave weakly usc function on $\Delta_p(\Delta(K))$. Because $\Delta_p(\Delta(K))$ is weakly compact, f is bounded from above so that $\Psi^0_f \neq \varnothing$. The definition of Ψ^0_f implies that $\forall \mu, f(\mu) \le \inf_{\psi \in \Psi^0_f} \tilde{\psi}(\mu)$.

As for the converse inequality, for a given $\bar{\mu} \in \Delta_p(\Delta(K))$, let t be such that $t > f(\bar{\mu})$. The set $A := \{\mu \mid f(\mu) < t\}$ is then an open set in the weak topology, and it contains $\bar{\mu}$. Therefore, there exists a positive ϵ and a finite family $\{\psi_1, \ldots, \psi_n\}$ of continuous functions such that $B := \{\mu \mid \forall i\colon \tilde{\psi}_i(\mu) < \tilde{\psi}_i(\bar{\mu}) + \epsilon\}$ is included in A. In other words, if we define $L(\mu, s) \in \mathbb{R}^{n+1}$ as

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

then $\forall\mu$, $\forall s\geq 0$: $L(\mu,s)\notin\,]0,\infty[^{n+1}$. Because f is concave and $\tilde{\psi}_i$ is linear, the set $H:=\{L(\mu,s)\mid\mu\in\Delta_p(\Delta(K)), s\geq 0\}$ 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 $\forall\mu$, $\forall s\geq 0$:

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

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 $\mu=\bar{\mu}$ would then be strictly positive. This is a contradiction.

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

We now prove Claim 2. Let f be also Blackwell increasing. Let ψ be in Ψ_f^0 and let $\phi:=\text{vex}(\psi)$ be the convexifiaction of ψ. That is, ϕ is the largest lsc convex function dominated by ψ. Because for all x in $\Delta(K)$ there is a probability distribution μ_x on $\Delta(K)$ such that $\mathrm{E}_{\mu_x}[p]=x$ and that $\phi(x)=\mathrm{E}_{\mu_x}[\psi(p)]$, for any random variable X with probability distribution μ in $\Delta_p(\Delta(K))$, 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 $\mu'\succeq\mu$. Because $\psi\in\Psi_f$ and f is Blackwell increasing, we infer that $\tilde{\phi}(\mu)=\mathrm{E}[\phi(X)]=\mathrm{E}[\mathrm{E}[\psi(Y)\mid X]]=\mathrm{E}[\psi(Y)]=\tilde{\psi}(\mu')\geq f(\mu')\geq f(\mu)$. This is true for all μ. Thus, we have actually proved that if $\psi\in\Psi_f^0$, then $\text{vex}(\psi)=\phi\in\Psi_f^0$. Therefore, $f(\mu)\leq\inf_{\psi\in\Psi_f^{0,\text{vex}}}\tilde{\psi}(\mu)\leq\inf_{\psi\in\Psi_f^0}\tilde{\psi}(\mu)=f(\mu)$.

Claim 3 follows from the fact that a convex continuous function ψ on $\Delta(K)$ can be uniformly approximated by a function $\psi_2\in\Psi^{1,\text{vex}}$, 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 $O_g:=\{p\in\Delta(K)\colon g(p)>\psi(p)-\epsilon/2\}$ indexed by $g\in\mathcal{G}$ is thus an open covering of the compact set $\Delta(K)$. It therefore contains a finite subcovering $\mathcal{G}'$, and we conclude that $\psi_1(p):=\max\{g(p)+\epsilon/2\colon g\in\mathcal{G}'\}$ 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 functionals 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 $\rho>0$, $\psi_2(x):=\mathrm{E}[\psi_1(x+\rho Z)]$, where Z is a random vector whose components are independent standard normal random variables. Then ψ_2 is continuously differentiable because of the smoothing property of the convolution operator. ψ_2 is also convex because for any $\lambda\in[0,1]$, $x,x'\in\mathbb{R}^K$,

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

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

8. The duality between v and w.

The Fenchel transform w of v was defined in Equation (2) as

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

The next theorem states that v is the Fenchel transform of w.

Theorem 8.1 (Fenchel Duality Equation).

$$\forall\mu,\nu,\; v(\mu,\nu)=\inf_{\psi\in\Psi^{1,\text{vex}}}w(\psi,\nu)+\tilde{\psi}(\mu).$$

Proof. From the definition of w in Equation (2), we have 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 $\forall\mu,\nu$, $\inf_{\psi\in\Psi^{1,\text{vex}}}w(\psi,\nu)+\tilde{\psi}(\mu)\geq v(\mu,\nu)$.

On the other hand, if $\psi\in\Psi^{1,\text{vex}}_{v(\cdot,\nu)}$, 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 that

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

The theorem is thus proved. □

9. Proof of Theorem 3.1

9.1. Definition of Γ. The game Γ is defined as follows. The action space A for player 1 is $\Delta(K)\times\mathcal{F}$, where $\mathcal{F}$ is the set of pairs $(\tilde{\mu},\phi)$ such that $\tilde{\mu}\in\Delta_p(\Delta(K))$ and $\phi\in\Phi^{1,\text{cav}}_{v(\tilde{\mu},\cdot)}$. The action space B for player 2 is $\Psi^{1,\text{vex}}\times\Delta(L)$. In state (k,l), if $a=(\tilde{p},\tilde{\mu},\phi)\in A$ and $b=(\psi,\tilde{q})$, the payoff $g(k,l,a,b)$ is

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

We need only 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, $\psi_Q\in\Psi^{1,\text{vex}}$ and $\tilde{q}_Q\in\Delta(L)$.

The expected payoff with respect to (P,Q) is then

$$\sum_{k,l}P_kQ_lg(a_P,b_Q,k,\ell)=\psi_Q(\tilde{p}_P)+\langle\nabla\psi_Q(\tilde{p}_P),P-\tilde{p}_P\rangle-\tilde{\psi}_Q(\mu_P)+\phi_P(\tilde{q}_Q)+\langle\nabla\phi_P(\tilde{q}_Q),Q-\tilde{q}_Q\rangle.$$

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 $\Gamma([I_1],[I_2])$. $\tilde{p}_P:=P$, $\tilde{\mu}_P:=[I_1]$, and $\phi_P:=\phi$ for an arbitrary $\phi\in\Phi^{1,\text{cav}}_{v([I_1],\cdot)}$.

The expected payoff with respect to (P,Q) is

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

Because P and Q are independent, and the law of P is $[I_1]$, the expectation of the first term with respect to Q is $\mathrm{E}_P[\psi_Q(P)]=\tilde{\psi}_Q([I_1])$. When taking expectations, the first two terms cancel each other. Furthermore, because 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 $\mathrm{E}_Q[\phi(Q)]=\tilde{\phi}([I_2])$.

Because $\phi\in\Phi^{1,\text{cav}}_{v([I_1],\cdot)}$ is arbitrary, we conclude that player 1 can guarantee $\sup_{\phi\in\Phi^{1,\text{cav}}_{v([I_1],\cdot)}}\tilde{\phi}([I_2])$, which is equal to $v([I_1],[I_2])$ (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 $\psi\in\Psi^{1,\text{vex}}$ leads to the following expression for the expected payoff with respect to (P,Q):

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

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

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

Because $(\phi_P;\mu_P)\in\mathcal{F}$, we have $\phi_P\in\Phi^{1,\text{cav}}_{v(\mu_P,\cdot)}$, and thus $\tilde{\phi}_P([I_2])\leq v(\mu_P,[I_2])$. The payoff of player 1 with respect to P, is less than

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

According to Equation (2), $w(\psi,[I_2])$ is greater than or equal to the last two terms in this expression, and player 1's expected payoff is thus at most

$$\tilde{\psi}([I_1])+w(\psi,[I_2]).$$

Because the function $\psi\in\Psi^{1,\text{vex}}$ in the above strategy of player 2 is arbitrary, player 2 can guarantee $\inf_{\psi\in\Psi^{1,\text{vex}}}\tilde{\psi}([I_1])+w(\psi,[I_2])$. According to Theorem 8.1, this is exactly $v([I_1],[I_2])$, as desired.

10. Finite partitions. In this section we prove that any function $V(\mathcal{P},\mathcal{Q})$ of partitions that is monotonic, with respect to refinements, is equal to $V(\mathcal{P},\mathcal{Q})=v([\mathcal{P}],[\mathcal{Q}])$, where v is a concave-convex, Blackwell monotonic, semicontinuous function on $\Delta_p(\Delta(K))\times\Delta_q(\Delta(L))$. A given V, defined on finitely many points of $\Delta_p(\Delta(K))\times\Delta_q(\Delta(L))$, will be extended to the whole set.

For any nonempty subset C of K, denote P_C the conditional probability on K given C. That is, $P_{C,k}:=(\mathbf{p}_k/\mathbf{p}(C))\mathbb{1}_{k\in C}$. Note that if $C\neq C'$, then $P_C\neq P_{C'}$. Let S denote $S:=\{P_C\colon\varnothing\neq C\subseteq K\}$. S is thus a subset of $\Delta(K)$. Any partition $\mathcal{P}$ induces a law $[\mathcal{P}]$: a distribution over S, which assigns to P_C the probability $\mathbf{p}(C)\mathbb{1}_{C\in\mathcal{P}}$. The law $[\mathcal{P}]$ belongs thus to the set $\Delta_p(S)$ of probability distributions on S with expectation $\mathbf{p}$.

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

Proof. Suppose that $r, q \in \Delta_p(S)$ are such that $\alpha q + (1-\alpha)r = [\mathcal{P}]$ for some $\alpha \in (0,1)$. In particular, the support of both q and r is $\{P_C; C \in \mathcal{P}\}$. The expectations of q and r are equal to $\mathbf{p}$. It implies that for every $C \in \mathcal{P}$, $q(P_C) = r(P_C) = \mathbf{p}(C)$, meaning that $q = r = [\mathcal{P}]$. □

A *decomposition of a measure* $\mu \in \Delta_p(\Delta(K))$ is a pair $(\lambda, \mu_\bullet)$ such that

(i) λ is a probability on partitions: $\lambda \in \Delta(\mathcal{K})$.

(ii) $\mu_\bullet$ is a map from $\mathcal{K}$ to $\Delta_p(\Delta(K))$ such that $\sum_{\mathcal{P}\in\mathcal{K}} \lambda(\mathcal{P})\mu_{\mathcal{P}} = \mu$.

(iii) $\forall \mathcal{P} \in \mathcal{K}$, $[\mathcal{P}] \preceq \mu_{\mathcal{P}}$.

Notice that there always exists a decomposition of a measure μ. Indeed, because the trivial partition $\{K\}$ is Blackwell dominated by any measure $\mu \in \Delta_p(\Delta(K))$, 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 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 $\nu \in \Delta_q(\Delta(L))$ is defined analogously as follows.

(i) ξ is a probability on partitions: $\xi \in \Delta(\mathcal{L})$.

(ii) $\nu_\bullet$ is a map from $\mathcal{L}$ to $\Delta_q(\Delta(L))$ such that $\sum_{\mathcal{Q}\in\mathcal{L}} \xi(\mathcal{Q})\nu_{\mathcal{Q}} = \nu$.

(iii) $\forall \mathcal{Q} \in \mathcal{L}$: $[\mathcal{Q}] \preceq \nu_{\mathcal{Q}}$.

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 $V: \mathcal{K} \times \mathcal{L} \to \mathbb{R}$, 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 10.1.

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

Proof. Because the payoff is bilinear in (λ, ξ), we just have to prove that $\Lambda(\mu)$ and $\Xi(\nu)$ are convex compact sets. Thus let $\lambda, \lambda^i \in \Lambda(\mu)$, $i = 0, 1$ 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 a convex combination of two points in $\Delta(\mathcal{K})$, λ belongs to $\Delta(\mathcal{K})$. Next, we prove in Lemma 10.2 that the set $\{\mu \in \Delta_p(\Delta(K)): \mu \succeq [\mathcal{P}]\}$ 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. Because $\Delta_p(\Delta(K))$ 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 a limit that we denote $\mu_{\mathcal{P}}$.

To conclude that $\lambda \in \Lambda(\mu)$, we have to prove that $(\lambda, \mu_\bullet)$ is a decomposition of μ. Clearly, λ belongs to $\Delta(\mathcal{K})$, because 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 10.2 that the set $\{\mu \in \Delta_p(\Delta(K)): \mu \succeq [\mathcal{P}]\}$ is closed in the weak topology. Because for every n, $\mu^n_{\mathcal{P}}$ belongs to this set, we also have $\mu_{\mathcal{P}} \succeq [\mathcal{P}]$, and $(\lambda, \mu_\bullet)$ is thus a decomposition of μ as required. □

Lemma 10.2. *For $\mu' \in \Delta_p(\Delta(K))$, let $R_{\mu'}$ denote $R_{\mu'} := \{\mu \in \Delta_p(\Delta(K)): \mu \succeq \mu'\}$. Then $R_{\mu'}$ is a convex subset of $\Delta_p(\Delta(K))$, which is closed in the weak topology.*

Proof. Let $\mu^i \in R_{\mu'}$, $r^i \geq 0$, $i = 0, 1$ be such that $1 = \sum_i r^i$, $\mu = \sum_i r^i \mu^i$. Because $\mu^i \succeq \mu'$, there exists a random vector (X^i_1, X^i_2) satisfying $\mathrm{E}[X^i_2 \mid X^i_1] = X^i_1$, where X^i_1 is μ'-distributed and X^i_2 is μ^i-distributed. We may assume that these two random vectors are independent and 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 $X_j := X^U_j$ satisfies the property that X_1 is μ'-distributed, and X_2 is μ-distributed. Furthermore,

$$\mathrm{E}[X_2 \mid U, X^0_1, X^1_1] = \mathrm{E}[UX^1_2 + (1-U)X^0_2 \mid U, X^0_1, X^1_1]$$

$$
\begin{aligned}
&= U\,\mathrm{E}[X_2^1 \mid U, X_1^0, X_1^1] + (1-U)\,\mathrm{E}[X_2^0 \mid U, X_1^0, X_1^1]\\
&= UX_1^1 + (1-U)X_1^0\\
&= X_1.
\end{aligned}
$$

Therefore, $\mathrm{E}[X_2 \mid X_1] = X_1$, 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 $\tilde{\psi}(\mu) = \lim_{n\to\infty} \tilde{\psi}(\mu^n) \geq \tilde{\psi}(\mu')$. Because this holds for all continuous convex function ψ, we infer that $\mu \succeq \mu'$, and the proof is complete. □

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

Theorem 10.2. *v is a Blackwell monotonic, concave-convex, semicontinuous function on $(\Delta_p(\Delta(K)) \times \Delta_q(\Delta(L)))$, such that*

$$\forall \mathcal{P} \in \mathcal{K}, \quad \forall \mathcal{Q} \in \mathcal{L}\colon\ v([\mathcal{P}], [\mathcal{Q}]) = V(\mathcal{P}, \mathcal{Q}). \tag{3}$$

Proof. (1) We start by proving Equation (3). To do so, we first argue that if $\lambda \in \Lambda([\mathcal{P}'])$, then λ can assign a positive probability only to partitions $\mathcal{P}$ that are less informative than $\mathcal{P}'$ (i.e., $\mathcal{P}'$ is a refinement of $\mathcal{P}$). Indeed, let $(\lambda, \mu_\bullet)$ be a decomposition of $[\mathcal{P}']$. Then $[\mathcal{P}'] = \sum_{\mathcal{P}\in\mathcal{K}} \lambda(\mathcal{P})\mu_{\mathcal{P}}$. Because S is the support of $[\mathcal{P}']$, we infer that $\mu_{\mathcal{P}} \in \Delta_p(S)$ for all $\mathcal{P}$, such that $\lambda(\mathcal{P}) > 0$. We demonstrate in Lemma 10.1 that the law of any partition $\mathcal{P}'$ is an extreme point of $\Delta_p(S)$. Therefore, for all $\mathcal{P}$, such that $\lambda(\mathcal{P}) > 0$, $\mu_{\mathcal{P}} = [\mathcal{P}'] \succeq [\mathcal{P}]$.

A similar argument shows that if $\xi \in \Xi([\mathcal{Q}'])$, then $\forall \mathcal{Q}$ satisfying $\xi(\mathcal{Q}) > 0$, and the following holds: $[\mathcal{Q}'] \succeq [\mathcal{Q}]$. Let λ^* denote the Dirac measure on $\mathcal{P}'$. Then $\lambda^* \in \Lambda(\mathcal{P}')$, because $(\lambda^*, \mu_\bullet^*)$ is decomposition of $[\mathcal{P}']$, where $\mu_\bullet^*$ is defined as $\mu_{\mathcal{P}}^* := [\mathcal{P}]$. Therefore,

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

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

(2) We next prove that $\mu \mapsto v(\mu, \nu)$ is concave. Let μ be a convex combination of μ^i, $i = 0, 1$: $\mu = \alpha_1\mu^1 + \alpha_0\mu^0$, with $\alpha_i \geq 0$ and $\alpha_1 + \alpha_0 = 1$. Let $\lambda^i \in \Lambda(\mu^i)$ be such that

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

We then argue that $\lambda := \sum_i \alpha_i\lambda^i$ belongs to $\Lambda(\mu)$. Because $\lambda^i \in \Lambda(\mu^i)$, there exists a decomposition $(\lambda^i, \mu_\bullet^i)$ of μ^i. Define then $\mu_\bullet$ by $\mu_{\mathcal{P}} := (\sum_i \alpha_i\lambda^i(\mathcal{P})\mu_{\mathcal{P}}^i)/\lambda(\mathcal{P})$ if $\lambda(\mathcal{P}) > 0$, $\mu_{\mathcal{P}} := [\mathcal{P}]$ otherwise. $(\lambda, \mu_\bullet)$ turns out to be a decomposition of μ, and thus $\lambda \in \Lambda(\mu)$ as announced. Indeed, because $\Delta(\mathcal{K})$ is convex, $\lambda \in \Delta(\mathcal{K})$. Next $\sum_{\mathcal{P}\in\mathcal{K}} \lambda(\mathcal{P})\mu_{\mathcal{P}} = \sum_i \alpha_i(\sum_{\mathcal{P}\in\mathcal{K}} \lambda^i(\mathcal{P})\mu_{\mathcal{P}}^i) = \mu$. Finally, $\mu_{\mathcal{P}} \succeq [\mathcal{P}]$ because the set $\{\mu \in \Delta_p(\Delta(K))\colon \mu \succeq [\mathcal{P}]\}$ is convex, as proved in Lemma 10.1.

The concavity of $\mu \mapsto v(\mu, \nu)$ then follows easily:

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

(3) That $\mu \to v(\mu, \nu)$ is Blackwell increasing is a straightforward consequence from the definition of v and the following monotonicity property of $\Lambda(\mu)$ that we now prove: if $\mu \preceq \mu'$, then $\Lambda(\mu) \subset \Lambda(\mu')$.

Indeed, if $\mu \preceq \mu'$ and $\lambda \in \Lambda(\mu)$, then there exists a decomposition $(\lambda, \mu_\bullet)$ of μ and a random vector (X, X') with respective marginals μ and μ', satisfying $\mathrm{E}[X' \mid X] = X$. The outcome of the following procedure is then a μ'-distributed random variable. Select first $\mathcal{P} \in \mathcal{K}$ according to λ, then select X according to $\mu_{\mathcal{P}}$, and finally select X' with the lottery δ_X (where δ_x denotes the conditional law of $X' \mid X = x$).

In this way we get 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 selected with probability δ_X. Because $[\mathcal{P}] \preceq \mu_{\mathcal{P}} \preceq \mu'_{\mathcal{P}}$, we infer that $(\lambda, \mu'_\bullet)$ is a decomposition of μ', and thus $\lambda \in \Lambda(\mu')$, as desired.

(4) We next argue that $\mu \mapsto v(\mu, \nu)$ is lsc. We show that the set $H := \{\mu\colon v(\mu, \nu) \geq a\}$ is closed in the weak topology on $\Delta_p(\Delta(K))$. 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 $v(\mu^n, \nu) = \inf_{\xi\in\Xi(\nu)} \sum_{\mathcal{P},\mathcal{Q}} \lambda^n(\mathcal{P})\xi(\mathcal{Q})V(\mathcal{P}, \mathcal{Q}) \geq a$. 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 $\sum_{\mathcal{P}\in\mathcal{K}} \lambda^n(\mathcal{P})\mu^n_{\mathcal{P}} = \mu^n \to \mu = \sum_{\mathcal{P}\in\mathcal{K}} \lambda(\mathcal{P})\mu_{\mathcal{P}}$. Because $\{\mu \in \Delta_p(\Delta(K))\colon \mu \succeq [\mathcal{P}]\}$ is closed in the weak topology (see Lemma 10.2), 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 $\xi \in \Xi(\nu)$,

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

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

(5) To conclude the proof, notice that Theorem 10.1 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. A concluding remark. 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.

Acknowledgment. The authors thank two anonymous referees and an associate editor for their precise and very helpful comments.

References

[1] Aumann, R. J., M. Maschler, with the collaboration of R. Stearns. 1995. *Repeated Games with Incomplete Information*. MIT Press, Cambridge, MA.
[2] Azrieli, Y., E. Lehrer. 2008. The value of a stochastic information structure. *Games Econom. Behav.* **63**(2) 679–693.
[3] Bassan, B., O. Gossner, M. Scarsini, S. Zamir. 2003. Positive value of information in games. *Internat. J. Game Theory* **32**(1) 17–31.
[4] Blackwell, D. 1953. Equivalent comparison of experiments. *Ann. Math. Statist.* **24**(2) 265–272.
[5] De Meyer, B. 1996. Repeated games and partial differential equations. *Math. Oper. Res.* **21**(1) 209–236.
[6] De Meyer, B. 1996. Repeated games, duality, and the central limit theorem. *Math. Oper. Res.* **21**(1) 237–251.
[7] De Meyer, B., A. Marino. 2005. Duality and optimal strategies in the finitely repeated zero-sum games with incomplete information on both sides. Mimeo.
[8] De Meyer, B., H. Moussa-Saley. 2003. On the origin of the Brownian motion in finance. *Internat. J. Game Theory* **31**(2) 285–319.
[9] De Meyer, B., D. Rosenberg. 1999. "Cav *u*" and the dual game. *Math. Oper. Res.* **24**(3) 619–626.
[10] Gilboa, I., E. Lehrer. 1991. The value of information—An axiomatic approach. *J. Math. Econom.* **20**(5) 443–459.
[11] Gossner, O. 2010. Ability and knowledge. *Games Econom. Behavior* **69** 95–105.
[12] Gossner, O., J.-F. Mertens. 2001. The value of information in zero-sum games. Mimeo.
[13] Hirshleifer, J. 1971. The private and social value of information and the reward to inventive activity. *Amer. Econom. Rev.* **61**(4) 561–574.
[14] Kamien, M., Y. Taumann, S. Zamir. 1990. On the value of information in a strategic conflict. *Games Econom. Behav.* **2**(2) 129–153.
[15] Laraki, R. 2002. Repeated games with lack of information on one side: The dual differential approach. *Math. Oper. Res.* **27**(2) 419–440.
[16] Lehrer, E., D. Rosenberg. 2006. What restrictions do Bayesian games impose on the value of information? *J. Math. Econom.* **42**(3) 343–357.
[17] Lehrer, E., D. Rosenberg, E. Shmaya. 2010. Signalling and mediation in games with common interest. *Games Econom. Behav.* **68**(2) 670–682.
[18] Mertens, J. F., S. Zamir. 1971. The value of two player zero-sum repeated games with lack of information on both sides. *Internat. J. Game Theory* **1** 39–64.
[19] Neyman, A. 1991. The positive value of information. *Games Econom. Behav.* **3** 350–355.
[20] Rockafellar, R. T. 1970. *Convex Analysis*. Princeton University Press, Princeton, NJ.
[21] Sorin, S. 2002. *A First Course in Zero-Sum Repeated Games*. Springer, New York.