THE VALUE OF INFORMATION IN ZERO-SUM GAMES

OLIVIER GOSSNER AND JEAN-FRANÇOIS MERTENS

Abstract. We study the value of information in zero-sum games.

1. Introduction

Blackwell. Extension to n players doesn't hold. Zero-sum games natural generalization.

- A decision maker against an adversary nature is modelled by a zero-sum game
- Zero-sum games have a value (minmax) theorem, which eases comparative statics among information structures 3/ A consequence of the minmax Theorem is that more information can only be beneficial in zero-sum games.

This opens a series of questions.

2. Model and Results

Information schemes and games. For any Hausdorff space X, $\mathcal{B}^X$ denotes its borel σ-field. The set of states of nature K is a Hausdorff space endowed with $\mathcal{B}^K$.

Definition 1. *An information scheme* $\mathfrak{E} = (E, \mathcal{E}, (\mathcal{E}_i), P, \kappa_E)$ *is given by*

- *a probability space* $(E, \mathcal{E}, P)$
- *two sub* σ*-algebras* $\mathcal{E}_1$ *and* $\mathcal{E}_2$ *of* $\mathcal{E}$*.*
- *a* P*-measurable map* $\kappa\colon (E, \mathcal{E}, P) \to (K, \mathcal{B}^K)$ *s.t.* $P \circ \kappa^{-1}$ *is tight.*

Definition 2. *A pay-off function* g *is a bounded continuous map* $g\colon A_1 \times A_2 \times K \to \mathbb{R}$*, where* A_1 *and* A_2 *are player 1 and player 2's compact action spaces.*

K is the "parameter space" of the statisticians, and g is the equivalent of the decision problem. A "state of the world" in E describes players' information on K, but also their whole hierarchies of beliefs on K, as well as potential correlated information they may receive.

$[g, \mathfrak{E}]$ denotes the extended (two person zero-sum) game in which initial information of the players and the true state of nature are initially generated by $\mathfrak{E}$, next players choose actions in A_1 and A_2, and finally pay-offs (to player 1) are determined by g. We rely on the following version of the $\mathsf{min\,max}$ Theorem. From [**?** , 4.3 p. 133] — cf. also below, after prop. 23:

Theorem 3. *The game* $[g, \mathfrak{E}]$ *has a value, denoted* $v_g(\mathfrak{E})$*, and there are optimal strategies.*

The formalism allows to vary separately the information scheme and the game. It allows us to study how changes in the information structure affects values accross games.

Date: June, 2020.

2.1. **Ordering of information schemes.** We now introduce basic relationships between information schemes.

First, *d*ecreasing i's information:

$\mathfrak{E}\ \mathbf{D}_i\ \mathfrak{E}'$ $(i = 1, 2)$ when $\mathfrak{E} = \mathfrak{E}'$ except for $\mathcal{E}_i' \subseteq \mathcal{E}_i$.

The decrease may be immaterial, hence an equivalence (*s*ufficiency):

$\mathfrak{E}\ \mathbf{S}_i\ \mathfrak{E}'$ $(i = 1, 2)$ when $\mathfrak{E}\ \mathbf{D}_i\ \mathfrak{E}'$ and $P(A|\mathcal{E}_i)$ is $\mathcal{E}_i'$-measurable $\forall A \in \mathcal{E}_j \vee \kappa^{-1}(\mathcal{B}^K)$ $(j \neq i)$.

A *d*ecrease in the σ-algebra on states of the world is immaterial, hence we consider it as an equivalence:

$\mathfrak{E}\ \mathbf{D}\ \mathfrak{E}'$ when $\mathfrak{E} = \mathfrak{E}'$ except for $\mathcal{E}' \subseteq \mathcal{E}$

Finally, *i*nclusion of an information scheme into another is the equivalence in which a zero probability common knowledge event is deleted:

$\mathfrak{E}\ \mathbf{I}\ \mathfrak{E}'$ when $E \in \mathcal{E}_1' \cap \mathcal{E}_2'$ with $P'(E) = 1$ and $\mathfrak{E} = \mathfrak{E}'_{|E}$

For two binary relations U and V, we denote UV and U^{-1} the constructed binary relations:

$$X\,UV\,Y \iff \exists Z, X\,U\,Z \text{ and } Z\,V\,Y$$
$$X\,U^{-1}Y \iff Y\,U\,X$$

We are interested in those relations between information schemes that (weakly) improve player 1's situation in 2-person zero-sum games. We thus consider chains of relations that consist of increasing player 1's information, decreasing player 2's information, and of equivalences.

Definition 4. *Let $\preccurlyeq$ be the relation between information schemes induced by any finite sequence of* $\mathbf{I}$, $\mathbf{I}^{-1}$, $\mathbf{D}$, $\mathbf{D}^{-1}$, $\mathbf{D}_1^{-1}$, $\mathbf{D}_2$, $\mathbf{S}_1$, *and* $\mathbf{S}_2^{-1}$.

The following relation means that there exists a *f*aithful relation from $\mathfrak{E}$ to $\mathfrak{F}$ (see [?], also thm ...).

Definition 5. *We let $\mathfrak{E}\ \mathbf{F}\ \mathfrak{F}$ when there is a commutative diagram*

$$(1)\qquad \begin{array}{ccc} \mathfrak{E} & \xrightarrow{d_1} & \mathfrak{E}_1 \\ \downarrow{\scriptstyle d_2} & & \downarrow{\scriptstyle s_2} \\ \mathfrak{E}_2 & \xrightarrow{s_1} & \mathfrak{F} \end{array}$$

where d_i is a $\mathbf{D}_i$ map and s_i a $\mathbf{S}_i$ map. (I.e., $\mathbf{F} = \mathbf{D}_1\mathbf{S}_2 \cap \mathbf{D}_2\mathbf{S}_1$.)

Since $\mathbf{D}_1\mathbf{S}_2$ weakly worsens player 1's situation, while $\mathbf{D}_2\mathbf{S}_1$ weakly improves it, $\mathbf{F}$ neither improves nor weakens player 1's situation. It can therefore be considered as an equivalence.

Based on these relations, we now consider chains consisting of decreasing only one of the player's information as well as equivalences.

Definition 6. *Let $\preccurlyeq_1$ be defined as $\preccurlyeq$ except that $\mathbf{D}_2$ (and $\mathbf{S}_1$) cannot be used, but $\mathbf{F}$ can. Similarly, let $\preccurlyeq_2$ be defined as $\preccurlyeq$ except that $\mathbf{D}_1^{-1}$ (and $\mathbf{S}_2^{-1}$) cannot be used, but $\mathbf{F}^{-1}$ can. Finally, $\sim$ is defined as $\preccurlyeq$ except that $\mathbf{D}_1^{-1}$, $\mathbf{S}_1$, $\mathbf{D}_2$ and $\mathbf{S}_2^{-1}$ cannot be used, but $\mathbf{F}$ and $\mathbf{F}^{-1}$ can.*

Points 1 – 5 in the Theorem below relate those orderings, while point 6 expresses monotonicity of the value w.r.t. information. Note that point 1 follows straight from the definitions.

Theorem 7. (1) *$\preccurlyeq$, $\preccurlyeq_1$, $\preccurlyeq_2$ and $\sim$ are transitive and reflexive, and $\sim$ is symmetric.*

(2) *Those relations are represented by:*
 (a) $\preccurlyeq \;=\; \mathbf{D}^{-1}\mathbf{F}^{-1}\mathbf{IFDD}^{-1}\mathbf{F}^{-1}\mathbf{I}^{-1}\mathbf{D}_2\mathbf{D}_1^{-1}\mathbf{IFDD}^{-1}\mathbf{F}^{-1}\mathbf{I}^{-1}\mathbf{FD}$
 (b) $\preccurlyeq_1 \;=\; \mathbf{D}^{-1}\mathbf{F}^{-1}\mathbf{IFDD}^{-1}\mathbf{F}^{-1}\mathbf{I}^{-1}\mathbf{D}_1^{-1}\mathbf{IFDD}^{-1}\mathbf{F}^{-1}\mathbf{I}^{-1}\mathbf{FD}$
 (c) $\preccurlyeq_2 \;=\; \mathbf{D}^{-1}\mathbf{F}^{-1}\mathbf{IFDD}^{-1}\mathbf{F}^{-1}\mathbf{I}^{-1}\mathbf{D}_2\mathbf{IFDD}^{-1}\mathbf{F}^{-1}\mathbf{I}^{-1}\mathbf{FD}$
 (d) $\sim \;=\; \mathbf{IFDD}^{-1}\mathbf{F}^{-1}\mathbf{I}^{-1}\mathbf{IFDD}^{-1}\mathbf{F}^{-1}\mathbf{I}^{-1}\mathbf{FD}$

(3) $\preccurlyeq_1$ *and* $\preccurlyeq_2$ *commute.*

(4) $\preccurlyeq \;=\; \preccurlyeq_1\preccurlyeq_2$

(5) $\sim$ *is the equivalence relation induced by any of the three orders, i.e.* $\mathfrak{E} \sim \mathfrak{F}$ *iff* $\mathfrak{E} \preccurlyeq \mathfrak{F}$ *and* $\mathfrak{E} \succcurlyeq \mathfrak{F}$.

(6) *All functions* $v_g(\cdot)$ *are* $\preccurlyeq$*-monotone.*

In order to characterize the equivalence classes for $\preccurlyeq$, we need a few additional tools, introduced in the next section.

Universal beliefs spaces, and consistent priors. We recall a couple of definitions and properties from [**?** , ch. III].

For any Hausdorff space X, $\Delta(X)$ denotes the (Hausdorff) space of tight probability measures on $\mathcal{B}^X$.

A *beliefs system* [**?** , comment 1.1 p. 110] is pair of Hausdorff spaces (Σ_1, Σ_2) together with continuous maps $\sigma_i \colon \Sigma_i \to \Delta(K \times \Sigma_{-i})$ $(i = 1, 2)$.

The *universal belief space* (*ibid.*) is a beliefs system $(\Theta_i, \theta_i)_{1,2}$ such that for every beliefs system $(\Sigma_i, \sigma_i)_{1,2}$, there exist continuous maps $(\phi_i)_{1,2}$ such that the following diagrams commute:

$$\begin{array}{ccc} \Sigma_i & \xrightarrow{\ \sigma_i\ } & \Delta(K \times \Sigma_{-i}) \\ \downarrow \phi_i & & id. \downarrow \quad \downarrow \phi_{-i} \\ \Theta_i & \xrightarrow{\ \theta_i\ } & \Delta(K \times \Theta_{-i}) \end{array}$$

Then, the ϕ_i are unique.

The universal beliefs space always exists, the θ_i are then homeomorphisms, and Θ_i is "unique" [**?** , thm. 1.1 p. 107].

Let $\Omega = K \times \Theta_1 \times \Theta_2$ be right name for this space ?, and for $P \in \Delta(\Omega)$, denote by P_i its marginal on Θ_i.

P is *consistent* iff $P(B) = \int \theta_i(B) P_i(d\theta_i)$ $(i = 1, 2)$ $\forall B \in \mathcal{B}^\Omega$ [**?** , def. 2.1 p. 119].

Denote by Π the space of consistent priors. Π is closed and convex [**?** , thm. 2.3 p. 120].

Any $Q \in \Delta(\Omega)$ is identified with the information scheme $\mathfrak{E}_Q = (\Omega, \mathcal{B}^\Omega, Q, (\mathcal{T}_i)_{1,2}, \kappa)$, where $\mathcal{T}_i$ is the σ-field spanned by $\mathcal{B}^{\Theta_i}$, and $\kappa \colon \Omega \to K$ is the projection map. For $Q \in \Pi$ such information schemes are called canonical.

Further, for any information scheme $\mathfrak{E}$, there exists a unique corresponding $P_\mathfrak{E} \in \Pi$, with a "unique" map from E to Ω [**?** , thm. 2.5 p. 122].

For $\mu \in \Delta(\Theta_1)$, define Q_μ by $Q_\mu(d\theta_1, d\theta_2, dk) = \theta_1(d\theta_2, dk)\mu(d\theta_1)$, and let $\Delta_b(\Theta_1)$ be the set of $\mu \in \Delta(\Theta_1)$ such that the marginal of Q_μ on K is tight. For $\mu \in \Delta(\Theta_1)$, P_μ represents the consistent prior $P_{\mathfrak{E}_{Q_\mu}}$ [**?** , def. 4.9 p. 135]. Define similarly $\Delta_b(\Theta_2)$ and P_ν for $\nu \in \Delta_b(\Theta_2)$.

Theorem 8. Π *is the set of equivalence classes of* $\preccurlyeq$*. i.e.,* $\mathfrak{E}_1 \sim \mathfrak{E}_2$ *iff* $P_{\mathfrak{E}_1} = P_{\mathfrak{E}_2}$.

Corollary 9. $\preccurlyeq$ *and the* $\preccurlyeq_i$ *induce orders (i.e., anti-symmetric) on* Π.

Corollary 10. $\forall g$*,* v_g *becomes a function on* Π *(i.e.,* $v_g(\mathfrak{E}) = v_g(P_\mathfrak{E})$ $\forall \mathfrak{E}$*).*

The functions v_g *are continuous and affine on* Π.

Proof. The first sentence follows from theorem 7, point 6 and corollary 9. For the second, cf. [**?** , prop. 4.4 p. 133]. □

Theorem 11. *For K completely regular, the functions v_g (with finite action sets) span the topology on Π.*

Analytic characterizations.

Theorem 12. *For P and Q in Π, $P \preccurlyeq Q$ (resp. $P \preccurlyeq_1 Q$, $P \preccurlyeq_2 Q$) iff there exists $R \in \Delta(\Omega \times \Omega')$ (where $\Omega' = K' \times \Theta'_1 \times \Theta'_2$ is a copy of Ω) with P and Q as first and second marginals and s.t.:*

(1) *the support of R is contained in the diagonal of $K \times K'$*
(2) *Ω and Θ'_2 are conditionally independent given Θ_2*
(3) *Ω' and Θ_1 are conditionally independent given Θ'_1*
(4) *for $\preccurlyeq_1$: Ω' and Θ_2 are conditionally independent given Θ'_2*
for $\preccurlyeq_2$: Ω and Θ'_1 are conditionally independent given Θ_1

Corollary 13. *Assume K completely regular. Then the graphs of $\preccurlyeq$, $\preccurlyeq_1$ and $\preccurlyeq_2$ in $\Delta(\Omega) \times \Delta(\Omega)$ are closed and convex.*

Corollary 14. *For K completely regular, any monotone net in Π for $\preccurlyeq$, $\preccurlyeq_i$, and the opposite orders converges, and the monotonicity is preserved at the limit.*

Remark that in the 1-person situation, and if K were a separable metric space, the martingale convergence theorem would imply that, when the player is faced with an increasing or a decreasing sequence of σ-fields of observations (on the same sample space), his posteriors would converge a.s. to the limiting posteriors. Here our observations are not on the same sample space, so there would be no meaning even to a convergence in probability: the convergence has to be weakened to a convergence in distribution: all the sharpness of the martingale convergence theorem is gone. But given that, we obtain a generalization to 2 players, with any of the orders, with arbitrary nets instead of sequences, and arbitrary completely regular spaces.

Corollary also shows that our definition of the orders using finite chains of basic relations was right: infinite chains would not yield anything more.

The following Theorem shows how $\preccurlyeq$ on canonical information structures can be obtained by degrading the information of player 1, then augmenting that of player 2.

Theorem 15. *$P \preccurlyeq P'$ iff $\exists Q \in \Pi$ s.t.*

(a) *there exists a transition probability ρ from Θ_2 to Θ'_2 s.t., with $\mathfrak{E}$ the information scheme on $\Omega \times \Theta'_2$ with $P \otimes \rho$ where player 2 is informed (only) of θ'_2, $P_{\mathfrak{E}} = Q$. And then ρ can be chosen such that the above "unique map" from $\mathfrak{E}$ to Ω induces the identity on Θ_2, more precisely, define g P_1 measurable from Θ_1 to $(\Theta_1, \mathcal{B}^{\Theta_1})$ such that P_1-a.e. $g(\theta_1)[d\theta_2, dk] = \int_{\tilde{\Theta}_2} \rho(d\theta_2 | d\tilde{\theta}_2) \theta_1(d\tilde{\theta}_2, dk)$, then if h denotes the "unique" map $\forall B \in \mathcal{B}^{\Theta_1} \otimes \mathcal{B}^{\Theta_2} \otimes \mathcal{B}^K$, $h^{-1}(B) = (h \times \mathbb{I}_{\Theta_2} \times \mathbb{I}_K)^{-1}(B)$ $P \otimes \rho$ a.e., where $\mathbb{I}_{\Theta_2}$ is the identity map from Θ'_2 to Θ_2.*

and dually:

(b) *there exists a transition probability ρ' from Θ_1 to Θ'_1 s.t., with $\mathfrak{E}'$ the information scheme on $\Omega \times \Theta'_1$ with $P' \otimes \rho'$ where player 1 is informed (only) of θ'_1, $P_{\mathfrak{E}'} = Q$ and $\int \rho' dP'_1 = Q_1$. And then ...*

or equivalently:

(a') *there exists $\mu \in \Delta(\Theta_2)$ such that $\int \phi d\mu \leq \int \phi dP_2$ for every convex l.s.c. function ϕ on Θ_2 which is bounded below, and such that $P_\mu = Q$.*[1]

[1]The first condition ensures that $\mu \in \Delta_b(\Theta_2)$, and in particular P_μ is well defined

(b') *there exists* $\nu \in \Delta(\Theta_1)$ *such that* $\int \psi d\nu \leq \int \psi dP_1'$ *for every convex l.s.c. function* ψ *on* Θ_1 *which is bounded below, and such that* $P_\nu = Q$.

Corollary 16. *A function* v *on* Π *is* $(\preccurlyeq)$-monotone *iff it is*

(a) monotone w.r.t. 2: $\forall P \in \Pi, \forall \mu \in \Delta(\Theta_2), \int \phi d\mu \leq \int \phi dP_2 \,\forall \phi$ *convex l.s.c. on* $\Delta(\Theta_2) \Rightarrow v(P) \leq v(P_\mu)$

(b) *and similarly,* monotone w.r.t. 1: $\forall P \in \Pi, \forall \nu \in \Delta(\Theta_1), \int \phi d\nu \leq \int \phi dP_1 \,\forall \phi$ *convex l.s.c. on* $\Delta(\Theta_1) \Rightarrow v(P_\nu) \leq v(P)$

Monotonicity w.r.t. 1 [resp. 2] corresponds to the classical convexity (w.r.t. 2) [resp. concavity w.r.t. 1] found in repeated games with incomplete information. And they strengthen the tentative generalizations of concavity and convexity at the end of [**?** , ch. III].

3. Proofs

3.1. **Functorial aspects of consistent priors.** For a *"paving"* $\mathcal{P}$ (a set of subsets of a set X), $\mathcal{P}_\sigma$, $\mathcal{P}_\delta$ and $\mathcal{P}_c$ denote the pavings consisting respectively of the countable unions, the countable intersections and the complements of elements of $\mathcal{P}$; $\mathcal{P}_{\sigma\delta} = (\mathcal{P}_\sigma)_\delta$, and so on. If X is a topological space, $\mathcal{Z}$ denotes the paving of zero sets, i.e., sets $f^{-1}(0)$ for f real-valued and continuous, and $\mathcal{K}$ that of compact subsets.

For beliefs spaces, functorial properties were obtained in [**?** , thm. 1.2 p. 111]. We will need analogous properties for consistent priors of point 3 there (point 1 is dealt with in cor. 2.4 p. 120 *loc. cit.*, and, as to point 2, consistent priors are only defined on the universal beliefs space).

Proposition 17. *Assume* K_1, K_2 *Hausdorff. For* $f\colon K_1 \to K_2$ *continuous, let* $\Omega(f) \stackrel{def}{=} f \times \Theta_1(f) \times \Theta_2(f)\colon \Omega(K_1) \to \Omega(K_2)$, *and* $\Pi(f) \stackrel{def}{=} \Delta(\Omega(f))_{|\Pi(K_1)}$.

(1) *The transpose* $\Omega^*(f)$ *of* $\Omega(f)$ *embeds* $C(\Omega(K_2))$ *into* $C(\Omega(K_1))$, *as Banach algebras, and commutes with the* θ_i *as operators from* $C(\Omega(K.))$ *to itself.*

(2) *For* K *compact,* $C(\Omega)$ *is the smallest closed algebra* A *containing* $C(K)$ *and s.t.* $\theta_i(f) \in A$ $(i = 1, 2)$ $\forall f \in A$.

For any information scheme $\mathfrak{E}$ *about* K_1, *let* $f \circ \mathfrak{E}$ *be the information scheme about* K_2 *obtained by replacing* κ_1 *in* $\mathfrak{E}$ *by* $f \circ \kappa_1$, *and* $T(f)\colon \mathfrak{E} \to f \circ \mathfrak{E}$ *be the identity on* E.

(3) $\Pi(f)\colon \Pi(K_1) \to \Pi(K_2)$ *is continuous, and the following diagram commutes (a.e.), the maps* ϕ_i *being as in [**?** , thm. 2.5 p. 122], — so* $\Pi(f)(P_{\mathfrak{E}}) = P_{f\circ\mathfrak{E}}$:

$$\begin{array}{ccc} \mathfrak{E} & \xrightarrow{T(f)} & f \circ \mathfrak{E} \\ \big\downarrow \phi_1 & & \big\downarrow \phi_2 \\ P_{\mathfrak{E}} & \xrightarrow{\Omega(f)} & P_{f\circ\mathfrak{E}} \end{array}$$

(4) *If* $f\colon K_1 \to K_2$ *is one-to-one, or an inclusion (of a closed subset, of a* $\mathcal{Z}$*-subset, of a* $\mathcal{Z}_{c\delta}$*-subset), so is* $\Pi(f)\colon \Pi(K_1) \to \Pi(K_2)$. *In case of an inclusion,* $K_1 \subseteq K_2$, *one has more precisely* $\Pi(K_1) = \Pi(K_2) \cap \Delta(\Omega(K_1)) = \{\, Q \in \Pi(K_2) \mid (Q \circ \kappa_2^{-1})(K_1) = 1 \,\}$.

(5) *For* K_1 K*-analytic, if* $f\colon K_1 \to K_2$ *is onto, so is* $\Pi(f)$.

Proof. 1: Follows immediately from property (P) [thm. 1.1.1 p. 107 in **?**].

2: Suffices by Stone-Weierstrass to show that A separates points, and hence that it separates points of all Ω_n's ($= \Omega$ with the hierarchies of beliefs truncated at level n [cf. **?** , thm. 1.1.3 p. 108]). This follows by induction: it holds by definition for $\Omega_{-1} = K$, and for the induction step, when knowing that we have all continuous functions f on Ω_n (Stone-Weierstrass as above), those f will separate points of $\Delta(\Omega_n)$, hence in particular the $\theta_i(f)$ separate points of $\Theta_{i,n+1}$.

3: Continuity of $\Pi(f)$ follows by definition from that of $\Omega(f)$ [property (P) in thm. 1.1.1 p. 107 in **?**]; and that its values are contained in $\Pi(K_2)$ follows from the last statement, applied to the canonical information schemes. As to that one, by [**?** , thm. 2.5.2 p. 122], suffices to show that $\Omega(f)\circ\phi_1$ satisfies the requirements on a ϕ_2 in point 1 of that theorem, $T(f)$ being the identity on E; continuity of $\Omega(f)$ ensures the measurability and point 1b, point 1a is by definition of $\Omega(f)$ and of ϕ_1, while the left hand member in point 1c equals $\theta^i(\phi_1(e))\big([\Omega(f)]^{-1}(B)\big)$ by property (P) *loc. cit.* and the right hand member equals $P\big(\phi_1^{-1}[(\Omega(f))^{-1}(B)] \mid \mathcal{E}_i\big)(e)$ (by definition), hence equality follows from point 1c *loc. cit.* for ϕ_1.

4: By the corresponding result for $\Omega(f)$ [**?** , thm. 1.2.3a p. 111], 9.b.1 and 9.b.2 [**?** , p. 428] imply our conclusions, except:

(a) for the "more precisely", remains to show that $Q\in\Pi(K_2), (Q\circ\kappa_2^{-1})(K_1) = 1 \Rightarrow Q\in\Pi(K_1)$, since the inclusions from left to right are now clear. Fix then $B\subseteq K_1$ in $\mathcal{B}^{K_2}$ with $Q(\kappa_2^{-1}(B)) = 1$; using [**?** , thm. 1.2.3b p. 111] with B as K_1, we conclude that inductively A_n^i is borel and $Q(A_n^i) = 1$, hence $Q(\Omega(K_1)) = 1$ since $\Omega(B)\subseteq\Omega(K_1)$ (point 3a *loc. cit.*), so $Q\in\Pi(K_2)\cap\Delta(\Omega(K_1))$. $Q\in\Pi(K_1)$ follows now straight from the definition of consistency.

(b) for the inclusion of a closed subset, of a $\mathcal{Z}$-subset, of a $\mathcal{Z}_{c\delta}$-subset, our conclusions are that $\Delta(\Omega(K_1))$ is such a subset of $\Delta(\Omega(K_2))$; the equality $\Pi(K_1) = \Pi(K_2)\cap\Delta(\Omega(K_1))$ implies then the result.

5: For $Q\in\Pi(K_2)$, choose $\mu\in\Delta(K_1)$ s.t. $f(\mu)=\kappa_2(Q)$, using 9.b.3 in [**?** , p. 428]. Let ν denote the image of μ on the graph of f, and use [**?** , II.1Ex.16c p. 75] to get a corresponding transition probability ρ from K_2 to K_1, i.e., $\rho\colon K_2\to\Delta(K_1)$ is such that the inverse image of every borel set is $\kappa_2(Q)$-measurable, the induced probability on $\mathcal{B}^{\Delta(K_1)}$ is tight, and $\nu(B) = \int\rho(B|x)(\kappa_2(Q))(dx)$ for all B in the product of the borel σ-fields.

$Q\otimes\rho$ defines a tight distribution on $\Omega(K_2)\times K_1$; with the projection to K_1, this defines an information scheme $\mathfrak{E}$ about K_1: let $P = P_{\mathfrak{E}}$, and $Q' = \Pi(f)(P)$. We must show that $Q' = Q$. By (3), $Q' = P_{f\circ\mathfrak{E}}$, and $f\circ\mathfrak{E}$ is the (canonical) information scheme Q (about K_2), followed by the transition ρ to K_1 and then f from K_1 to a copy K_2' of K_2, and where the "state of nature" is generated from the coordinate in K_2'. There is no loss to extend this to a tight distribution on $\Omega(K_2)\times K_1\times K_2'$. We claim that this distribution is carried by the diagonal in $K_2\times K_2'$. To prove this, suffices to take a pair of disjoint open sets O and O' in K_2, and to prove that $(\kappa_2(Q)\otimes(f\circ\rho))(O\times O') = 0$. The left hand member equals $\int\rho(O\times f^{-1}(O')|x)(\kappa_2(Q))(dx)$, i.e., $\nu(O\times f^{-1}(O'))$. Since ν is by definition carried by the graph of f, we get indeed 0. The intermediate factor K_1 (as well as the factor K_2) can be forgotten for computing the associated consistent prior since it affects neither the true state of nature nor the information of the players. Thus our distribution on $K_2\times\Theta_1(K_2)\times\Theta_2(K_2)\times K_2'$ has Q as marginal on the first 3 factors, and is carried by the diagonal in $K_2\times K_2'$: its marginal on the last 3 factors is also Q. So $Q' = P_{f\circ\mathfrak{E}}$ is the canonical distribution associated to $\mathfrak{E}_Q$: $Q' = Q$. □

Corollary 18.

(1) *For K_1 compact, if f is a quotient map, so are $\Omega(f)$ and $\Pi(f)$.*
(2) *For K compact, and for a sequence f_n of continuous functions on Ω, there is a metrisable quotient $\bar{K}$ of K s.t. the f_n factor through the map $\Omega\to\Omega(\bar{K})$.*

Proof. 1: Since any continuous map from a compact space onto a Hausdorff space is a quotient map, this follows from point 5 of prop. 17.

2: Consider, for each finite subset α of $C(K)$, the smallest algebra A_α containing α and s.t. $\theta_i(f) \in A_\alpha$ $(i = 1, 2)$ $\forall f \in A_\alpha$. By prop. 17.2, $\bigcup_\alpha A_\alpha$ is dense in $C(\Omega)$. We obtain thus a sequence α_k, s.t. all f_n are in the closure of $\bigcup_k A_{\alpha_k}$. Then the closed algebra C_0 spanned by $\bigcup_k \alpha_k$ and the constants defines the metrisable quotient $\bar{K}$, with quotient map ϕ. The image of $C(\Omega(\bar{K}))$ by $\Omega^*(\phi)$ contains all f_n, by prop. 17.1, since all operations to construct them from elements of C_0 (algebra-, $\theta_i(\cdot)$, limits) are preserved by $\Omega^*(\phi)$. □

Lemma 19. *Let $\mathfrak{E} \preccurlyeq \mathfrak{F}$ be two information schemes about K. Assume $f : K' \to K$ is continuous, and either an inclusion, with $P_\mathfrak{E}(K') = 1$, or bijective, with K' K-analytic. Then $\mathfrak{E}' = f^{-1} \circ \mathfrak{E}$ and $\mathfrak{F}' = f^{-1} \circ \mathfrak{F}$ (with the obvious meaning — cf. prop. 17) are well-defined information schemes about K', with $\mathfrak{E}' \preccurlyeq \mathfrak{F}'$, and $P_{\mathfrak{E}'} = P_{\mathfrak{F}'}$ iff $P_\mathfrak{E} = P_\mathfrak{F}$.*

Proof. The schemes are well-defined: first, also $P_\mathfrak{F}(K') = 1$, since $\mathfrak{E} \preccurlyeq \mathfrak{F}$ implies they induce the same distribution on K. Next, for the inclusion, the definition assumes that $E' = \kappa_E^{-1}(f(K'))$, and use prop. 17.4 and 17.3. For the bijection, suffices clearly to show that f^{-1} is universally measurable (à la Lusin). Now, by 9.b.3 in [**?** , p. 428], every $\mu \in \Delta(K)$ is the image by $\Delta(f)$ of some $\mu' \in \Delta(K')$. Then with C a compact subset of K' with large μ'-measure, $f(C)$ is compact in K with large μ-measure, and f^{-1} is continuous on it: f^{-1} is μ-Lusin measurable.

$P_{\mathfrak{E}'} = P_{\mathfrak{F}'}$ if $P_\mathfrak{E} = P_\mathfrak{F}$ by prop. 17.4, and conversely, because $\mathfrak{E} = f \circ \mathfrak{E}'$ and $\mathfrak{F} = f \circ \mathfrak{F}'$ (up to null sets in case of an inclusion, i.e., $f \circ \mathfrak{E}'$ $\mathbf{I}\mathbf{S}_1\mathbf{S}_2\mathbf{D}$ $\mathfrak{E}$, where the last 3 operations only remove null sets.).

$\mathfrak{E}' \preccurlyeq \mathfrak{F}'$: since $f \circ \mathfrak{E}'$ $\mathbf{I}\mathbf{S}_1\mathbf{S}_2\mathbf{D}$ $\mathfrak{E}$ which are equivalences, we can assume that $\kappa_E(E) \subseteq f(K')$, so now the schemes are strictly well-defined, and $\mathfrak{E} = f \circ \mathfrak{E}'$. By the same argument, all $\mathbf{I}, \mathbf{I}^{-1}, \mathbf{D}, \mathbf{D}^{-1}, \mathbf{D}_1^{-1}, \mathbf{D}_2, \mathbf{S}_1, \mathbf{S}_2^{-1}$ in the chain are still such operations when viewed as operating between the corresponding schemes about K' (recall that sufficiency, being defined in terms of conditional expectations, is unaffected by null sets). □

3.1.1. *A modification of canonical information schemes.* Assume P is a canonical information scheme, then replacing $\mathcal{B}^\Omega$ by $\mathcal{P} = \mathcal{B}^K \times \mathcal{B}^{\Theta_1 \times \Theta_2}$ leads to a modified canonical information scheme $\mathfrak{C}_P^m$ such that P $\mathbf{D}$ $\mathfrak{C}_P^m$.

Lemma 20. *For any information scheme $\mathfrak{E}$, there is a modified map $\phi^m : E \to \Omega$ having the same properties as ϕ in [**?** , thm. 2.5 p. 122], except that $\mathcal{B}^\Omega$ is replaced by $\mathcal{P}$, and $\phi^m = (\kappa, \phi_1, \phi_2)$, where ϕ_i is $\mathcal{E}_i \vee \mathcal{N}$-measurable to $\mathcal{B}^{\Theta_i}$ — $\mathcal{N}$ denoting the σ-field of all negligible subsets of $(E, \mathcal{E}, P)$.*

Remark 21. So ϕ^m induces the modified canonical information scheme $\mathfrak{C}^m_{P_\mathfrak{E}}$.

Proof. [**?** , rem. 2.8 p. 126] proves the statement, except that the measurability of ϕ_i is only obtained (from [**?** , thm. 2.5.1c p. 122]) as "$\phi_i(B)$ is $\mathcal{E}_i \vee \mathcal{N}$-measurable for every $B \in \mathcal{P}$". To conclude from this to our statement, observe that (by tightness) it suffices to prove that $\phi_1^{-1}(C)$ is $\mathcal{E}_i \vee \mathcal{N}$-measurable for any compact set $C \subseteq \Theta_1$. Let then C' be a compact subset disjoint from C: by the same argument, suffices to show that there exists a borel subset $B \supseteq C$ with $B \cap C' = \emptyset$ s.t. $\phi_1^{-1}(B)$ is $\mathcal{E}_i \vee \mathcal{N}$-measurable. Suffices thus to prove that every $x \neq y$ there is such a borel set B which is a neighbourhood of x and whose complement is one of y. Observe that the proof that $\Delta(X)$ is Hausdorff for X Hausdorff rests on the fact that, given $\mu_1 \neq \mu_2 \in \Delta(X)$, there exist disjoint open sets O_1 and O_2, and $\alpha_i \in \mathbb{R}, \alpha_1 + \alpha_2 > 1$ s.t. $\mu_i(O_i) > \alpha_i$, so with $V_i = \{\, \mu \in \Delta(X) \mid \mu(O_i) > \alpha_i \,\}$, V_1 and V_2 are disjoint open sets in $\Delta(X)$ containing resp. μ_1 and μ_2. When taking $(\mu_1, \mu_2) = (x, y)$, the set V_1 becomes our desired set B. □

3.2. **Values.**

Proposition 22. *Let Σ_1, Σ_2 be compact convex spaces, and $G_n = ((\Sigma_i)_i, g_n)$ and $G = ((\Sigma^i)_i, g)$ be zero-sum games (complete information). Assume all pay-off functions g_n, g are separately continuous in both arguments, quasi-concave in the first and quasi-convex in the second, and such that $(g_n)_n$ converges uniformly to g. Then:*

(1) *Values $V(G_n)$, $V(G)$ exist as well as optimal strategies in the corresponding games, and $\lim V(G_n) = V(G)$;*
(2) *If $\sigma_n^i \in \Sigma_i$ is (ε-)optimal for player i in G_n and if $\lim \sigma_n^i = \sigma^i$, then σ^i is (ε-)optimal in G.*

Proof. The existence of values and optimal strategies follow by Sion's theorem [e.g., **?** , theorem 1.6 p. 4]. For point 2, by the uniform convergence, and monotonicity, increasing a bit ε allows to assume that $g_n = g\,\forall n$. Then $\forall \tau$, $g(\sigma_n, \tau) \geq V(G) - \varepsilon\,\forall n \Rightarrow g(\sigma, \tau) \geq V(G) - \varepsilon$ by the separate continuity. □

We will apply the above result via:

Proposition 23. *For a game with incomplete information, endow each player i's strategy space Σ_i — the set of transition probabilities from $(E, \mathcal{E}_i)$ to his action space A_i — with the "weak" topology, i.e., the weakest making continuous all integrals of products of an integrable function on $(E, \mathcal{E}_i)$ with a continuous function on A_i.*

Each Σ_i is then compact convex in a locally convex space, and metrisable if A_i is so and $(E, \mathcal{E}, P)$ is separable; and the pay-off, separately continuous.

Proof. Cf. first paragraph of the proof of **?** , prop. 4.3 p. 133]. The metrisability conclusion is then obvious, from the existence of a countable set of continuous functions that separates points. □

Proof of thm. 3 and of thm. 7.6. Thm. 3 is immediate from the above. Thm. 7.6 follows then from the monoticity of values w.r.t. information, except for $\mathbf{S}_i$ which is also a classic argument, cf. e.g. the proof of prop. 4.5 p. 134 in [**?**]. □

Proposition 24. $\mathfrak{E} \sim P_{\mathfrak{E}}$

Proof. Let ϕ^m be the modified map from $\mathfrak{E} = (E, \mathcal{E}, (\mathcal{E}_i), \kappa_E, P)$ to $\mathfrak{E}^m_{P_{\mathfrak{E}}} = (\Omega, \mathcal{P}, (\mathcal{B}^{\Theta_i}), \text{proj}_K, P_{\mathfrak{E}})$ as in lemma 20.

Let $\mathfrak{E}^c$ be $\mathfrak{E}$ in which all σ-algebras are completed by elements of P probability zero — hence $\mathfrak{E}\;\mathbf{D}^{-1}\mathbf{S}_1\mathbf{S}_2\;\mathfrak{E}^c$ —, and let $\mathfrak{F}$ be obtained from $\mathfrak{E}^c$ by replacing $\mathcal{E}_i^c$ by $\mathcal{F}_i = \phi^{m,-1}(\mathcal{B}^{\Theta_i})$. It follows from lemma 20 that $\mathcal{F}_i \subseteq \mathcal{E}_i^c$ and that $\mathcal{F}_i$ is a sufficient statistic for $\mathcal{E}_i^c$ on $\mathcal{F}_j \vee \kappa_E^{-1}(\mathcal{K})$ $(j \neq i)$, hence $\mathfrak{E}^c\;\mathbf{F}\;\mathfrak{F}$.

First assume $E \cap \Omega = \emptyset$. Let $G = E \cup \Omega$, endowed with $\mathcal{G} = \mathcal{E}^c \vee \mathcal{P}$, $\mathcal{G}_i = \mathcal{F}_i \vee \mathcal{B}^{\Theta_i}$, and $\kappa_G = \kappa_E \vee \text{proj}_K$. Considering P and $P_{\mathfrak{E}}$ as probabilities on $\mathcal{G}$, this defines $\mathfrak{G}_P = (G, \mathcal{G}, (\mathcal{G}_i), \kappa_G, P)$ and $\mathfrak{G}_{P_{\mathfrak{E}}} = (G, \mathcal{G}, (\mathcal{G}_i), \kappa_G, P_{\mathfrak{E}})$ such that $\mathfrak{F}\;\mathbf{I}\;\mathfrak{G}_P$ and $\mathfrak{E}^m_{P_{\mathfrak{E}}}\;\mathbf{I}\;\mathfrak{G}_{P_{\mathfrak{E}}}$.

We now connect $\mathfrak{G}_P$ to $\mathfrak{G}_{P_{\mathfrak{E}}}$. First, decrease $\mathcal{G}_i$ to $\mathcal{G}'_i$ spanned by the sets $\phi^{m,-1}(B) \cup B$, $B \in \mathcal{B}^{\Theta_i}$, and next $\mathcal{G}$ to $\mathcal{G}'$ spanned by the sets $\phi^{m,-1}(B) \cup B$, $B \in \mathcal{P}$. Note that κ_G is $\mathcal{G}'$-measurable since ϕ^m is the modified map, and that P and $P_{\mathfrak{E}}$ coincide on $\mathcal{G}'$ ($P_{\mathfrak{E}}$ being the image of P by ϕ^m), hence denote the resulting scheme by $\mathfrak{G}_0$. Now, each element of $\mathcal{G}_i$ differs of an element of $\mathcal{G}'_i$ by an element of $\mathcal{G}_i$ of P-probability 0, and by an element of $\mathcal{G}_i$ of $P_{\mathfrak{E}}$-probability 0. Hence $\mathfrak{G}_P\;\mathbf{S}_1\mathbf{S}_2\mathbf{D}\;\mathfrak{G}_0$, and $\mathfrak{G}_0\;\mathbf{D}^{-1}\mathbf{S}_1^{-1}\mathbf{S}_2^{-1}\;\mathfrak{G}_{P_{\mathfrak{E}}}$.

If $E \cap \Omega \neq \emptyset$, let $P'_{\mathfrak{E}}$ be a copy of $P_{\mathfrak{E}}$ over a space Ω' s.t. $\Omega' \cap \Omega = \Omega' \cap E = \emptyset$. The previous construction shows that $E \sim P'_{\mathfrak{E}} \sim P_{\mathfrak{E}}$. □

3.3. **Topological properties of strategy spaces.**

Lemma 25. *Let (g_n) converge to g_∞ in $(L_\infty, \sigma(L_\infty, L_1))$. There exists a sequence of convex combinations of (g_n), $g'_k = \sum_n \alpha_{k,n} g_n$, such that:*

- *$\alpha_{n,k}$ goes to infinity, i.e. for every n, $\lim_{k\to\infty} \alpha_{n,k} = 0$.*
- *g'_n converges to g_∞ P a.s.*

Proof. Let D be the set of convex combinations of $(g_n)_n$, and let $\bar{D}$ be the closure of D for the Mackey topology $\tau(L_\infty, L_1)$. Since $\bar{D}$ is $\tau(L_\infty, L_1)$ closed it is also $\sigma(L_\infty, L_1)$ closed. Hence $g_\infty \in \bar{D}$. Recall that on bounded subsets of L_∞, $\tau(L_\infty, L_1)$ coincides with the topology of convergence in probability (and also with the L_2 and L_1 topologies). Since $\bar{D}$ is bounded in L_∞, the result follows by Egorov's theorem. □

Proposition 26. *For X compact metric, there exists a borel map h from $X^{\mathbb{N}}$ to X such that $h((x_n)_n) = x$ whenever x_n converges to x.*

Proof. First notice that the set C of converging sequences of elements of $[0,1]$ is a borel subset of $[0,1]^{\mathbb{N}}$. Indeed, for $m \in \mathbb{N}$ and $\varepsilon > 0$, let $F_{m,\varepsilon}$ be the set of sequences $(y_n)_n \in [0,1]^{\mathbb{N}}$ such that there exists $n, n' > n$ with $|y_n - y_{n'}| > \varepsilon$. $F_{m,\varepsilon}$ is open in $[0,1]^{\mathbb{N}}$, and C is simply the complement of $\cup_{k\in\mathbb{N}^*} \cap_{m\in\mathbb{N}} F_{m,1/k}$. Let $(\phi_i)_{i\in\mathbb{N}}$ be a sequence of continuous functions from X to $[0,1]$ that separates points. A sequence $(x_n)_n \in X^{\mathbb{N}}$ converges if and only if for every i, $(\phi_i(x_n))_n \in C$. Hence the subset $D \subseteq X^{\mathbb{N}}$ of converging sequences is borel. The mapping h from D to X that associates its limit to every converging sequence is borel if and only if for every continuous function ϕ from X to $[0,1]$, $\phi \circ h$ is borel. Since $(\phi \circ h)((x_n)_n) = \lim_n \phi(x_n)$, it is enough to prove that the map l from C to $[0,1]$ such that $l((y_n)_n) = \lim_n y_n$ is borel. This last point comes from the fact that l is the limit of the sequence of measurable "projection" maps $p_i \colon C \to [0,1]$ defined by $p_i((y_n)_n) = y_i$. □

Proposition 27. *Fix an information scheme $\mathfrak{E} = (E, \mathcal{E}, (\mathcal{E}_i)_i, P, \kappa_E)$, compact metric convex sets A_i, and $(\gamma_n)_n = ((A_i)_i, g_n)_n$ and $\gamma = ((A_i)_i, g)$ s.t. all pay-off functions $(g_n)_n$ and g are separately continuous from $A_1 \times A_2 \times K$ to $\mathbb{R}$, and concave in the first argument, and s.t. $(g_n)_n$ converges uniformly to g. For any sequence of pure optimal strategies $(\sigma_{1,n})_n$ in $\Gamma(\gamma_n, \mathfrak{E})$, there exists a borel map $f_1 \colon (A_1)^{\mathbb{N}} \to A_1$ such that the strategy σ_1 defined by $\sigma_1(e) = f_1(\sigma_{1,n}(e))$ is optimal in $\Gamma(\gamma, \mathfrak{E})$.*

Proof. By prop. 22 and 23, extract a subsequence along which $\sigma_{1,n}$ converges, say to σ_1 (note that the sub-σ-field of $\mathcal{E}_1$ spanned by the $\sigma_{1,n}$ is separable); σ_1 is then optimal in $\Gamma(\gamma, \mathfrak{E})$. So there is a sequence of convex combinations of the unit masses at the $\sigma_{1,n}$ s.t. those combinations converge, for a.e. e, weakly in $\Delta(A_1)$ to $\sigma_1(e)$: let g_n be a dense sequence of continuous functions on A_1, and take by lemma 25 for each n_0 a convex combination of $(\delta_{\sigma_{1,n}})_{n\geq n_0}$ which is, in $L_1(E, \mathbb{R}^{n_0})$, 2^{-n_0} close to σ_1 on each of $g_1 \dots g_{n_0}$; then a.s. those convex combinations converge to σ_1 on a dense set of continuous functions, hence weakly. Apply now prop. 26; since the sequence of convex combinations of point masses is clearly a borel function on $A_1^{\mathbb{N}}$, $\sigma_1(e)$ is a borel function of $(\sigma_{1,n}(e))_n$. Map finally each $\sigma_1(e)$ to its barycentre; this is clearly borel, and still yields an optimal strategy by the concavity of the pay-off. Let f_1 be the composition of those two borel maps. □

3.4. **Required information in a game.**

Definition 28. *Let γ be a game with compact metric action spaces and continuous pay-off function, and $\mathfrak{E} = (E, \mathcal{E}, (\mathcal{E}_i), P, \kappa_E)$ be an information scheme. The sub-sigma-field $\mathcal{F}_i$ of $\mathcal{E}_i$ is required for player i in a game γ with action spaces (A_j)*

when for any optimal (behavioral) strategy σ_i of player i in γ extended by $\mathfrak{E}$, there exists a map F_i from A_i to the set of probability measures on $(E, \mathcal{F}_i)$ such that:

- *$\forall X \in \mathcal{F}_i$, the map $a_i \mapsto F_i(X)(a_i)$ is measurable.*
- *$\forall X \in \mathcal{F}_i$, $(F_i \circ \sigma_i)(X)(e) = 1_X$ a.s.*

where by definition, $(F_i \circ \sigma_i)(X)(e) = \int_{A_i} F_i(X)(a_i) d\sigma_i(a_i)(e)$.

Lemma 29. *Let K be a separable metric space and γ a game with compact metric A_i's and continuous pay-off function. There exists a game γ' with compact metric A'_i's and continuous pay-off function s.t., $\forall i$:*

- *In γ', i has a unique best reply for each belief on $K \times A'_{3-i}$.*
- *For any information scheme $\mathfrak{E} = (E, \mathcal{E}, (\mathcal{E}_i)_i, P, \kappa_E)$ and sub σ-fields $\mathcal{F}_i$ of $\mathcal{E}_i$ such that $\mathcal{F}_i$ is required for player i in γ extended by $\mathfrak{E}$, $\mathcal{F}_i$ is also required in γ' extended by $\mathfrak{E}$.*

In particular, γ' extended by $\mathfrak{E}$ has P-a.s. unique optimal strategies.

Proof. First, consider the game $\bar{\gamma}$ with action spaces $S_i = \Delta(A_i)$ (with the weak* topology) and pay-off function $\bar{g}$ defined by $\bar{g}(s_1, s_2, k) = E_{s_1,s_2} g(a_1, a_2, k)$. Choose strictly concave functions[2] f_i on S_i, and define a sequence of "perturbed" games $\bar{\gamma}_n$ with action spaces S_i and pay-off functions defined by $\bar{g}_n(s_1, s_2, k) = \bar{g}(s_1, s_2, k) + \frac{1}{n}(f_1(s_1) - f_2(s_2))$. For each belief on $K \times S_{3-i}$, each player i has an unique best response in $\bar{\gamma}_n$. Hence in $\bar{\gamma}_n$ extended by $\mathfrak{E}$, each player i has a pure and P-a.s. unique optimal strategy $\sigma_{i,n}$. We finally define a game $\gamma' = ((A'_i)_i, g')$ in which the sequence of games $(\bar{\gamma}_n)_n$ is played simultaneously: $A'_i = (S_i)^{\mathbb{N}}$ and $g'((s_{1,n})_n, (s_{2,n})_n, k) = \sum_n 2^{-n} g(s_1, s_2, k)$. The A'_i are compact metric (for the product topology) and g' is continuous. The unique optimal strategy for player i in γ' is σ_i defined by $\sigma_i(e) = (\sigma_{i,n}(e))_n$.

We now prove that $\mathcal{F}_i$ is required in γ' extended by $\mathfrak{E}$. Since the sequence of pay-off functions (g_n) converges uniformly to g, proposition 27 provides a measurable map f_i from $A'_i = S_i^{\mathbb{N}}$ to $S_i = \Delta(A_i)$, thus a transition probability from A'_i to A_i such that σ_i defined by $\sigma_i(e) = f_i((\sigma_{i,n}(e))_n)$ is optimal in γ extended by $\mathfrak{E}$. Now, $\mathcal{F}_i$ being required in γ, let F_i from A_i to $(E, \mathcal{F}_i)$ be as in definition 28. Then $F_i \circ f$ is the required transition probability from A'_i to $(E, \mathcal{F}_i)$. □

3.4.1. *One person decision problems.* In the one player case, a (canonical) information scheme I is called a (canonical) statistical experiment,and a game d, decision problem—with value $val(I, d)$. An experiment I_1 is said less informative than another I_2, denoted $I_1 \le I_2$, if for any decision problem d with finite action set D and continuous pay-off function on $K \times D$, $val(I_1, d) \le val(I_2, d)$.

Lemma 30. *If K is compact metric and $I_1 \le I_2$, then for every decision problem d on K with a Blackwell space of actions D and borel nonnegative pay-off function on $K \times D$, $val(I_1, d) \le val(I_2, d)$.*

Proof. First we assume D compact metric, and show that $val(I_1, d) \le val(I_2, d)$ for a decision problem d with a continuous pay-off function g on $K \times D$. The map f that associates to any belief $x \in \Delta(K)$ the maximum expected pay-off under x,

$$f(x) = \sup_{a \in D} \mathbf{E}_x g(k, a)$$

is continuous and convex on $\Delta(K)$. Thus there exists a sequence $a_n \in D$ s.t. $f_n(x) = \sup_{i \le n} \mathbf{E}_x g(k, a_i)$ converges uniformly to f. The restriction d_n of d to $\{a_1 \dots a_n\}$ is a decision problem with finite decision set and continuous pay-off, so

[2] For instance, consider a dense sequence $(\phi_n)_n$ of continuous linear functionals on $\Delta(A_i)$ (which is compact metric), and let $f_i(x) = \sum_n 2^{-n} \|\phi_n\|^{-2} (\phi_n(x))^2$.

$val(I_1, d_n) \le val(I_2, d_n)$. Moreover, if μ_j ($j \in \{1, 2\}$) is the marginal on Θ (the set of types of the player) of the canonical statistical experiment associated to I_j, one has

$$val(I_j, d) = \int_{\Delta(K)} f d\mu_j$$

and similarly for d_n. Hence in the limit $val(I_1, d) \le val(I_2, d)$.

Now, by a theorem of Blackwell, Cartier, Fell and Meyer, [e.g., **?** , remark 1.34 p. 78] there exists a family $(T_x)_{x\in\Delta(K)}$ of probability measures on $\Delta(K)$ such that each T_x has barycentre x, for every borel set B of $\Delta(K)$ the map $x \to T_x(B)$ is borel, and $\mu_2(B) = \int_{x\in\Delta(K)} T_x(B)d\mu_1$. Let d be a decision problem on K with Blackwell space of actions D and borel nonnegative pay-off function on $K \times D$, and let f be the convex universally measurable function defined as before on $\Delta(K)$: then $val(I_j, d) = \int_{\Delta(K)} f d\mu_j$, because there exist universally measurable strategies which are uniformly ε-optimal, D being Blackwell. Thus $\int_{\Delta(K)} f d\mu_1 \le \int_{\Delta(K)} f d\mu_2$ by Jensen's inequality. Hence the result. □

Lemma 31. *Given K compact metric, there exists a decision problem d with compact metric action space and continuous pay-off function such that for every canonical experiment I, the borel σ-field on $\Delta(X)$ is required in d extended by I. More precisely, there exists a continuous function F from A to $\Delta(K)$ s.t. for any canonical experiment I, and any optimal decision function σ in d extended by I, $F(\sigma(x)) = x$ a.s.*

Proof. Fix two arbitrary decisions a' and a''. Let D be the set of continuous functions from $K \times \{a', a''\}$ to $[0, 1]$, and D' a countable dense subset of D, for the uniform topology. For any pair of beliefs x_1, x_2 on K, there exists two continuous functions $g(a', .)$ and $g(a'', .)$ on K s.t. $\mathbf{E}_{x_1} g(a', .) > \mathbf{E}_{x_1} g(a'', .)$ and $\mathbf{E}_{x_2} g(a'', .) > \mathbf{E}_{x_2} g(a', .)$, hence there exist $d \in D'$ s.t. the optimal actions in d given x_1 and x_2 differ. Let $(d_l)_{l\in\mathbb{N}^*}$ be an enumeration of D', and d the decision problem with action set $A = \{a', a''\}^{\mathbb{N}}$ and pay-off function $d_0((a_l)_{l>0}, k) = \sum_{l>0} 2^{-l} d_l(a_l, k)$. d_0 is jointly continuous, so the expected pay-off is jointly continuous on $A \times \Delta(K)$. Thus, with $R(x)$ be the set of best responses to $x \in \Delta(K)$, $x \mapsto R(x)$ is u.s.c. on $\Delta(K)$; in particular, $R(\Delta(K))$ is compact. Notice that $R(x_1) \cap R(x_2) = \emptyset$ for $x_1 \ne x_2$. Define $F: R(\Delta(K)) \to \Delta(K)$ as $F(a) = x$ whenever $a \in R(x)$. Then F is a map with closed graph (R u.s.c.) between compact spaces, hence continuous on its domain $R(\Delta(K))$. So the "more precisely" clause is established, and hence the borel σ-field on $\Delta(K)$ is required in $\Gamma(I, d)$. □

Proposition 32. *Given a compact metric K, there exists a decision problem d_0 with action space $\Delta(K)$ such that for any belief x on K, the only optimal action in d_0 is x.*

Proof. Take the decision problem d of lemma 31, and apply lemma 29 to it: in the new decision problem d_0, with action set A', there is a unique optimal action for each belief on K, so this is a continuous function $f: \Delta(K) \to A'$. f is one to one because the borel σ-field on $\Delta(K)$ is required in d_0 for any canonical experiment. We have thus still all those properties when reducing the action set to $A'' = f(\Delta(K))$. But now f is a homeomorphism between A'' and $\Delta(K)$. □

Lemma 33. *For a decision problem d_0 as in prop. 32, for any pair I_1, I_2 of statistical experiments on K with $I_1 \le I_2$, either I_1 and I_2 are associated to the same canonical experiment, or $val(I_1, d_0) < val(I_2, d_0)$.*

Proof. By lemma 30, $val(I_1, d_0) \leq val(I_2, d_0)$, with equality when I_1 and I_2 have the same canonical experiment. Let $I_1 \leq I_2$ be canonical, i.e. I_i is represented by a probability μ_i over $X_i = \Delta(K)$. I_i also defines a probability measure P_i over $X_i \times K$. Since $I_1 \leq I_2$, there exists a transition probability Q from X_2 to X_1 s.t. P_1 is the marginal on $X_1 \times K$ of the law induced by P_2 and Q on $X_1 \times X_2 \times K$. If $val(I_1, d_0) = val(I_2, d_0)$, following Q and playing optimally given x_1 is an optimal strategy in d_0 extended by I_2. By uniqueness of the optimal action in d_0, it follows that $Q(x_2, .)$ is the Dirac mass at x_2, μ_2 a.s. Hence $\mu_1 = \mu_2$. □

3.4.2. *Main lemma of this part.*

Lemma 34. *Assume K is compact metric. For any game γ with compact metric action spaces and continuous pay-off there exists a game γ' with compact metric action spaces and continuous pay-off such that, for any information scheme $\mathfrak{E} = (E, \mathcal{E}, (\mathcal{E}_i)_{i \in I}, P, \kappa_E)$, if $\mathcal{F}_i \subseteq \mathcal{E}_i$ is required for some player i in γ, then the sub-σ-field $\mathcal{F}_{3-i}$ of $\mathcal{E}_{3-i}$ generated by his opponent's beliefs on $K \times (E, \mathcal{F}_i)$ is required for this opponent in γ'.*

Proof. Let $\gamma = (A, B, g)$. We shall fix $i = 2$ throughout the proof, constructing thus in fact a game γ'_2; γ' will be the game $\gamma'_1 \times \gamma'_2$ where both are played in parallel. Using lemma 29, we can assume that in γ, player 2 has a unique best response for each belief on $K \times A$. Hence player 2's optimal strategy is unique, and pure, thus the unit mass at some point $b(e)$, where $b\colon E \to B$ is $\mathcal{E}_2$-measurable. For the space of states of nature $K \times B$, let d be a "separating" decision problem for player one (in the sense of prop. 32) with action space $X = \Delta(K \times B)$.

Given $n > 0$, let γ'_n be the game with action spaces $A \times X$ and B, and with pay-off $g'_n(a, x, b, k) = g(a, b, k) + \frac{1}{n} d(k, b, x)$. Finally, γ' is the game with action spaces $A' = (A \times X)^{\mathbb{N}}$ and $B' = B^{\mathbb{N}}$ and pay-off $g'((a_n, x_n), (b_n), k) = \sum_n 2^{-n} . g'_n(a_n, x_n, b_n, k)$.

Claim 35. Let τ be any optimal strategy of 2 in γ'. Then b_n converges in $P \times \tau$-probability to $b(e)$.

Proof. By the uniqueness of 2's optimal strategy in γ, prop. 22 and 23 imply the convergence of τ_n to $\delta_{b(.)}$ in the sense that, for any continuous function f on B, $\tau_n(f)$ converges $\sigma(L_\infty, L_1)$ to $f(b(e))$.

Let then B_1 and B_2 be 2 disjoint closed sets in B. There is a continuous function f from B to $[0, 1]$ which equals 1 on B_1 and 0 on B_2. Let $Y = b^{-1}(B_2)$: the weak convergence implies that the integral on Y of $\tau_n(f) - f(b(e))$ tends to 0. But since $f(b(e)) = 0$ and $\tau_n(f) \geq \tau_n(B_1)$, this implies that $\mu_n(B_1 \times B_2)$ tends to zero, with μ_n the probability on $B \times B$ induced by (τ_n, τ_0). By compactness of B, there exists for every neighborhood U of the diagonal in B^2 a finite number of pairs of disjoint closed sets B_1 et B_2 s.t. the products $B_1 \times B_2$ cover $\complement U$: $\mu_n(U)$ tends to 1, for every such neighborhood, i.e., $d(x, y)$ tends in probability to 0 under μ_n, and hence, being bounded, its integral tends to zero: $E \int_B d(b(e), y) \tau_n(dy|e) \to 0$. □

Claim 36. Let τ be optimal for 2 in γ'. Then $(P \times \tau)(k, b_n | \mathcal{E}_1)$ converges weakly to $(P \times \tau)(k, b(e) | \mathcal{E}_1)$ in P-probability.

Proof. Since $K \times B$ is compact metric, the conditional probabilities exist. Suffices to prove that from any subsequence we can extract a further subsequence along which the conclusion holds. Since $b_n \to b(e)$ in probability (claim 35), extract an a.s. convergent subsequence (Egorov). Now for any continuous function f on $K \times B$, $E\big(f(k, b_n) \big| \mathcal{E}_1\big) \to E\big(f(k, b(e)) \big| \mathcal{E}_1\big)$ a.s. (dominated convergence). □

Claim 37. Given a pair of optimal strategies (σ, τ), there exists $G\colon A' \to \Delta(K \times B)$ such that $G(a_n, x_n) = Q(k, b(e) | \mathcal{E}_1)$ a.s.

Proof. Use $x_n \overset{\text{a.s.}}{=} Q(k, b_n|\mathcal{E}_1)$ (prop. 32), claim 36, Egorov, and prop. 26. □

3.5. $P_{\mathfrak{E}}$ depends only on the $\preccurlyeq$ indifference class of $\mathfrak{E}$.

Proposition 38. *Assume K is a separable metric space. There exists a game γ with compact metric action spaces and continuous pay-offs s.t. each player has a unique best reply for any prior on K and his opponent's action, and s.t., for any consistent prior P, there are pure optimal strategies which are borel isomorphisms from Θ_i to its image.*

Proof. The separable metric spaces are the subspaces of compact metric spaces; by [? , theorem 1.2.3a p. 111] and prop. 17.4, topological inclusion is preserved when going to the universal beliefs spaces, so we can assume K compact metric, the borel isomorphism property being preserved by restriction to a subspace. Use then lemma 34 inductively, starting with $\mathcal{F}_i = \{\emptyset, \Omega\}$ and γ_0 a game with singleton action sets. Let γ_n be the game obtained at the n^{th} stage of the induction. If $\mathcal{T}_i^n$ denotes the sub-σ-field of $\mathcal{T}_i$ spanned by the first n levels of the hierarchy of types, then $\mathcal{T}_i^n$ is by construction required for i in γ_n, for any consistent prior P.

Let $\gamma_\infty = \prod_n \gamma_n$ be the game where all γ_n are played in parallel (and pay-offs summed after multiplication by suitable weights): $\mathcal{T}_i$ is required in γ_∞ for each i and any consistent prior P. Using lemma 29 yields now further the uniqueness of best replies, and so the existence of pure, borel optimal strategies (i using at each θ_i the unique best reply against some fixed borel optimal strategy of j). With M the sup norm of the game, replace then A_i by its disjoint union with a copy Θ_i' of Θ_i, defining i's pay-off as $-M-1$ when he plays in Θ_i' and his opponent not, and as 0 when both do: our previous conclusions are unaffected. Given a pure borel optimal strategy $a_i(\theta_i)$, there is, $\mathcal{T}_i$ being required, a borel map $f_i \colon A_i \to \Theta_i$ s.t. $f_i \circ a_i$ is a.e. the identity (separability of σ-fields): $N = \{\, \theta_i \mid (f_i \circ a_i)(\theta_i) \neq \theta_i \text{ or } a_i(\theta_i) \in \Theta_i' \,\}$ is a negligible borel set, and a_i injective outside. Redefine then a_i on N by $a_i(\theta_i) = \theta_i \in \Theta_i'$: it is still a borel pure optimal strategy, and is one to one, hence [? , 5.e p. 425] a borel isomorphism with its image, A_i and Θ_i being compact metric. □

Remark 39. 1) Uniqueness of the best reply implies its continuity on $\Delta(K \times A_{-i})$, and the a.s. uniqueness of optimal strategies.
2) So, for any consistent P, the distribution $\bar{P}$ (with marginal $\bar{P}_i$ on A_i) on $\Omega \times A_1 \times A_2$ induced by P and optimal strategies depends only on P.
3) $P \mapsto \bar{P}_i$ is injective: for $P \neq P'$, their marginals on Θ_i also differ; a pure optimal strategy in $\frac{1}{2}P + \frac{1}{2}P'$ which is a borel isomorphism with its image is optimal in P and P', so by the borel isomorphism, $\bar{P}_i \neq \bar{P}_i'$.

Lemma 40. *Assume K compact metric. For consistent priors $P \neq P'$, $v_g(P) \neq v_g(P')$ for some pay-off function with finite action spaces g.*

Proof. Else $v_g(P) = v_g(P')$ for any pay-off function g: if e.g. $v_g(P) > v_g(P')$ for some g, this relation is preserved when replacing the action spaces by a sufficiently fine finite discretisation (use first [? , prop. 4.3 p. 133] for player II in $\gamma(P')$, next for player I in $\gamma(P)$). Let then $G = ((A_i), g)$ be as in prop. 38; for a continuous function h on A_1 and $\epsilon > 0$ let G_ϵ be the game with action spaces (A_i) and pay-off function $g_\epsilon(a_1, a_2, k) = g(a_1, a_2, k) + \epsilon h(a_1)$. The Mills derivative $\mathsf{lim}_{\epsilon \to 0}(V(G_\epsilon, P) - V(G, P))/\epsilon$ is then also equal at P and P', and [cf. ? , I.1Ex.6 p. 11] equals $\int h d\bar{P}_1$ (rem. 2 above): $\forall h, \int h d\bar{P}_1 = \int h d\bar{P}_1'$, contradicting rem. 3 above. □

Proof of theorem 11. As in prop. 38, suffices to deal with the case of compact K, since the completely regular spaces are the subspaces of compact spaces. And then the space of consistent priors is also compact [cf. **?** , cor. 2.4 p. 120], so suffices to show that the $V(G, P)$ separate points.

$P_1 \neq P_2$ implies there is a continuous function on Ω whose integral differs under P_1 and P_2. Hence by cor. 18.2, for some metrisable quotient $\bar{K}$ of K, still $P_1 \neq P_2$ for the induced priors (cf. prop. 17.3) on $\Omega(\bar{K})$. Apply now lemma 40: there is a game G with finite action spaces and with pay-offs continuous on $\bar{K}$, for which $Val(G, P_1) \neq Val(G, P_2)$. □

Proposition 41. *$P_{\mathfrak{E}}$ depends only on the $\preccurlyeq$ indifference class of $\mathfrak{E}$.*

Proof. Assume $\mathfrak{E} \preccurlyeq \mathfrak{E}' \preccurlyeq \mathfrak{E}$, and $P_{\mathfrak{E}} \neq P_{\mathfrak{E}'}$. Fix (tightness) a sequence K_i of disjoint compact subsets of K s.t., with $K_\infty = \bigcup_i K_i$, $P_{\mathfrak{E}}(K_\infty) = P_{\mathfrak{E}'}(K_\infty) = 1$. By lemma 19, $\mathfrak{E}$ and $\mathfrak{E}'$ can be viewed as schemes over K_∞, and still $\mathfrak{E} \preccurlyeq \mathfrak{E}' \preccurlyeq \mathfrak{E}$ when viewed as schemes over K_∞, and $P_{\mathfrak{E}} \neq P_{\mathfrak{E}'} \in \Pi(K_\infty)$. Thus we can assume $K = \bigcup_i K_i$.

Let now L denote the space K, with each K_i as additional open subset: the map $f : L \to K$ is bijective and continuous, and L is completely regular, being locally compact, and K-analytic, being a $\mathcal{K}_\sigma$. So by lemma 19, we still have $\mathfrak{E} \preccurlyeq \mathfrak{E}' \preccurlyeq \mathfrak{E}$ when viewed as schemes over L, and $P_{\mathfrak{E}} \neq P_{\mathfrak{E}'}$, thus contradicting thm. 11 (by thm. 7.6 and prop. 24). □

Proof of Theorem 8. Use prop. 41 and prop. 24. □

Remark 42. Cor. 9 and 10 follow now. Thm. 7.5 and 7.6 are also established: for 7.5, from prop. 41 and prop. 24, because $\sim \subseteq \preccurlyeq_i \subseteq \preccurlyeq$ $(i = 1, 2)$; and 7.6 was done after prop. 23. Thus remain to be proved before theorem 12 only points 2 to 4 of thm. 7.

3.6. **Comparison of canonical information structures.**

Proof of Theorem 12. We first claim that there exists such R iff there exist $\mathfrak{E}$ and $\mathfrak{F}$ with $P_{\mathfrak{E}} = P$, $P_{\mathfrak{F}} = Q$, and $\mathfrak{E}\ \mathbf{D}_1^{-1}\mathbf{D}_2\ \mathfrak{F}$ (resp. $\mathfrak{E}\ \mathbf{D}_1^{-1}\ \mathfrak{F}$, $\mathfrak{E}\ \mathbf{D}_2\ \mathfrak{F}$). Assume such R exists, and let $\tilde{E} = \{ (\omega, \omega') \in \Omega \times \Omega' \mid k = k' \}$ endowed with R, and let $\tilde{\mathcal{E}}_i = \mathcal{B}^{\Theta_i} \times \mathcal{B}^{\Theta'_i}$ $(i = 1, 2)$. $\mathfrak{E}$ and $\mathfrak{F}$ are the same except that $\mathcal{E}_1 = \mathcal{B}^{\Theta_1}$ and $\mathcal{F}_2 = \mathcal{B}^{\Theta'_2}$. It is then clear that $\mathfrak{E}\ \mathbf{D}_1^{-1}\mathbf{D}_2\ \mathfrak{F}$, and $P_{\mathfrak{E}} = P$ by 2 and $P_{\mathfrak{F}} = Q$ by 3. For $\preccurlyeq_1$, the argument is the same except that already $P_{\tilde{\mathfrak{E}}} = Q$ since $\tilde{\mathfrak{E}}\ \mathbf{F}\ Q$ by 3 and 4. Dually for $\preccurlyeq_2$.

Assume now such $\mathfrak{E}$, $\mathfrak{F}$, where w.l.o.g. the σ-fields contain all null sets. Thus, $\mathcal{F}_2 \subseteq \mathcal{E}_2$ and $\mathcal{E}_1 \subseteq \mathcal{F}_1 \subseteq \mathcal{E}$ on $(E, \mathcal{E}, P)$. Let $\phi_{\mathfrak{E}}$ and $\phi_{\mathfrak{F}}$ from $(E, \mathcal{E}, P)$ to Ω and to Ω' be the modified maps of lemma 20 corresponding to $\mathfrak{E}$ and $\mathfrak{F}$. Let $\phi = (\phi_{\mathfrak{E}}, \phi_{\mathfrak{F}})$ from $(E, \mathcal{E}, P)$ to $\Omega \times \Omega'$. Then, ϕ is by definition measurable to $\mathcal{B}^K \times \mathcal{B}^{\Theta_1 \times \Theta_2} \times \mathcal{B}^{K'} \times \mathcal{B}^{\Theta'_1 \times \Theta'_2}$, hence induces a probability measure R on the product of those four σ-fields. The marginals of R on Ω and on Ω' are the restrictions to the corresponding σ-fields of P and Q respectively, in particular tight, hence R has a unique tight extension to $\mathcal{B}^{\Omega \times \Omega'}$, and this has P and Q as marginals on Ω and Ω'. Every product of disjoint open sets in K and K' (times $\Theta_1 \times \Theta_2 \times \Theta'_1 \times \Theta'_2$) is R-negligible. Hence point 1.

$\theta'_2 = \theta'_2(\phi_{\mathfrak{F}}(e))$ is (lemma 20) $\mathcal{F}_2$-, hence $\mathcal{E}_2$-measurable, thus $\mathcal{E}_2$'s independence of Ω given Θ_2 implies 2. Point 3 is dual.

Similarly for point 4. Hence our first claim.

We have thus proved the existence of such R when $E\ \mathbf{D}_1^{-1}\ F$ or $E\ \mathbf{D}_2\ F$. Recall that $P_{\mathfrak{E}} = P_{\mathfrak{F}}$ whenever E and F are related by $\mathbf{I}$, $\mathbf{D}$, or $\mathbf{F}$. Hence the existence of a composition of such R when $P \preccurlyeq Q$, $P \preccurlyeq_1 Q$, or $P \preccurlyeq_2 Q$. Remains thus only to show that the relation of P and Q to be related by such an R is transitive (in each of the three cases).

Assume P, P' on Ω and Ω' are related by R and P', P'' on Ω' and Ω'' by R'. Let $\rho(d\omega|\omega')$ and $\rho'(d\omega''|\omega')$ be the conditionals defined by R and R' [**?** , II.1Ex.16c

p. 75], and define a tight distribution $\tilde{R}$ on $\Omega \times \Omega' \times \Omega''$ by its marginal P' on Ω' and by having the product $\rho \otimes \rho'$ as conditional on $\mathcal{B}^{\Omega} \times \mathcal{B}^{\Omega''}$ given Ω'. $\tilde{R}$ has R and R' as marginals on $\Omega \times \Omega'$ and on $\Omega' \times \Omega''$ resp., hence its support is contained in the diagonal of $K \times K' \times K''$, and $\Omega \times \Omega'$ and $\Omega' \times \Omega''$ are conditionally independent under $\tilde{R}$ given Ω'.

For $X \in \mathcal{B}^{\Theta''_2}$, by the conditional independence of $\Omega \times \Omega'$ and $\Omega' \times \Omega''$ given Ω', $\tilde{R}(X|\Omega \times \Omega') = \tilde{R}(X|\Omega')$, which equals $\tilde{R}(X|\Theta'_2)$ by point 2 for the marginal R' of $\tilde{R}$. Taking now conditional expectations given Ω yields: $\tilde{R}(X|\Omega) = \mathsf{E}(\tilde{R}(X|\Theta'_2)|\Omega)$, $= \mathsf{E}(\tilde{R}(X|\Theta'_2)|\Theta_2)$ by point 2 for the marginal R of $\tilde{R}$, hence $= \mathsf{E}(\tilde{R}(X|\Omega \times \Omega')|\Theta_2) = \tilde{R}(X|\Theta_2)$. Hence point 2 for (the marginal on $\Omega \times \Omega''$ of) $\tilde{R}$, and point 3 is dual. Remains to deal with point 4, e.g. for $\preccurlyeq_2$: this is the same argument, with subscripts 1 instead of 2, replacing just (twice) the use of (2) by that of (4). □

End of proof of thm. 7. If P and Q are related by R as in prop. 12, we proved above that then $P \sim \mathfrak{E}\ \mathbf{D}_1^{-1}\mathbf{D}_2\ \mathfrak{F} \sim Q$, hence point 4, and clearly $\mathbf{D}_1^{-1}\mathbf{D}_2 = \mathbf{D}_2\mathbf{D}_1^{-1}$, hence point 3.

By remark 42, remains thus only to deal with point 2. We start with a lemma:

Lemma 43. *Let $\mathfrak{E} \sim \mathfrak{F}$, be such that $E \cap F = \emptyset$ and all σ-algebras $\mathcal{E}$, $\mathcal{F}$, $\mathcal{E}^i$, $\mathcal{F}^i$ are complete ($i = 1, 2$). Then $\mathfrak{E}\ \mathbf{IFDD}^{-1}\mathbf{F}^{-1}\mathbf{I}^{-1}\ \mathfrak{F}$.*

Proof of lemma 43. The inclusions embed $\mathfrak{E}$ and $\mathfrak{F}$ into $\mathfrak{G}$, the union $\mathfrak{E} \cup \mathfrak{F}$, endowed resp. with P_E and P_F. (The σ-fields and the map κ on $\mathfrak{G}$ are the obvious ones.) Let ϕ_E^m and ϕ_F^m denote the maps from lemma 20 for $\mathfrak{E}$ and $\mathfrak{F}$ resp., and use them to define ϕ on $\mathfrak{G}$. Define then $\mathcal{G}'$ and $\mathcal{G}'_i$ ($i = 1, 2$) as the inverse images by ϕ of $\mathcal{P}$ and of $\mathcal{B}^{\Theta_i}$ resp. It follows from lemma 20 that $\mathcal{G}' \subseteq \mathcal{G}$ and $\mathcal{G}'_i \subseteq \mathcal{G}_i$, that κ is $\mathcal{G}'$-measurable, and that decreasing both $\mathcal{G}_i$ on $\mathfrak{G}$ (with P_E or P_F) to $\mathcal{G}'_i$ is $\mathbf{F}$. Decreasing then also $\mathcal{G}$ to $\mathcal{G}'$ is $\mathbf{D}$. And on $\mathcal{G}'$, P_E and P_F coincide, because it is the probability distribution induced by the common canonical prior $P_{\mathfrak{E}} = P_{\mathfrak{F}}$ (thm. 8). □

Proof of thm. 7 2d. Let $\mathfrak{E} \sim \mathfrak{F}$. Take a copy $\mathfrak{E}'$ of $\mathfrak{E}$ s.t. E' is disjoint from both E and F, and add all null sets to $\mathcal{E}'$ and to $\mathcal{E}'_i$ ($i = 1, 2$). Let also $\mathfrak{F}'$ denote $\mathfrak{F}$ where all null sets have been added to $\mathcal{F}'$ and to $\mathcal{F}'_i$ ($i = 1, 2$). Clearly $\mathfrak{F}'\ \mathbf{FD}\ \mathfrak{F}$, and $\mathfrak{E}'\ \mathbf{IFDD}^{-1}\mathbf{F}^{-1}\mathbf{I}^{-1}\ \mathfrak{F}'$ by lemma 43. To prove $\mathfrak{E}\ \mathbf{IFDD}^{-1}\mathbf{F}^{-1}\mathbf{I}^{-1}\ \mathfrak{E}'$, argue as in the proof of lemma 43, except that now $\mathcal{G}'$ and $\mathcal{G}'_i$ ($i = 1, 2$) are the σ-fields $\{\, B \cup B' \mid B \in \mathcal{E} \text{ (resp. } \mathcal{E}_i) \,\}$, where B' is the copy of B in E'. □

Proof of thm. 7 2a, 2b and 2c. Let $\mathfrak{E} \preccurlyeq_1 \mathfrak{F}$ (or $\mathfrak{E} \preccurlyeq_1 \mathfrak{F}$, or $\mathfrak{E} \preccurlyeq_1 \mathfrak{F}$), and $R \in \Delta(\Omega \times \Omega')$ as in prop. 12. The set $\tilde{G} = \{\, (\omega, \omega') \in \Omega \times \Omega' \mid k = k' \,\}$ has R-outer measure 1, since any set in $\mathcal{B}^{\Omega \times \Omega'}$ disjoint from G has R-measure 0. Let $\mathcal{G}$ be the trace of $\mathcal{B}^{\Omega \times \Omega'}$ on G, and R the induced probability measure on it. Endow G with κ to K defined the obvious way, and with the trace σ-algebras $\mathcal{G}_E^i$ and $\mathcal{G}_F^i$ of $\mathcal{B}^{\Theta_i}$ and $\mathcal{B}^{\Theta'_i}$ ($i = 1, 2$). Complete $\mathcal{G}$, $\mathcal{G}_E^i$ and $\mathcal{G}_F^i$ ($i = 1, 2$) on G by all R-null sets, and still denote them the same way. Assume $G \cap E = G \cap F = \emptyset$, or take a copy of G with this property. This defines two information structures $\mathfrak{G}_E = (G, \mathcal{G}, (\mathcal{G}_E^i)_{i=1,2}, R, \kappa)$ and $\mathfrak{G}_F = (G, \mathcal{G}, (\mathcal{G}_F^i)_{i=1,2}, R, \kappa)$ such that $\mathfrak{E} \sim \mathfrak{G}_E$ and $\mathfrak{G}_F \sim \mathfrak{F}$.

Hence from lemma. 43 $\mathfrak{E}\ \mathbf{D}^{-1}\mathbf{F}^{-1}\mathbf{IFDD}^{-1}\mathbf{F}^{-1}\mathbf{I}^{-1}\ \mathfrak{G}_E$ and $\mathfrak{G}_F\ \mathbf{IFDD}^{-1}\mathbf{F}^{-1}\mathbf{I}^{-1}\ \mathbf{F}\ \mathbf{D}\ F$.

If $\mathfrak{E} \preccurlyeq_1 \mathfrak{F}$, point 2 of prop. 12 implies that adding $\mathcal{G}_F^2$ to $\mathcal{G}_E^2$ in $\mathfrak{G}_E$ is $\mathbf{S}_2^{-1}$. Adding then $\mathcal{G}_F^1$ is $\mathbf{D}_1^{-1}$, and by points 3 and 4, going from $((\mathcal{G}_E^1 \vee \mathcal{G}_F^1), (\mathcal{G}_E^2 \vee \mathcal{G}_F^2))$ to $(\mathcal{G}_F^1, \mathcal{G}_F^2)$ is $\mathbf{F}$: $\mathfrak{G}_E\ \mathbf{S}_2^{-1}\mathbf{D}_1^{-1}\mathbf{F}\ \mathfrak{G}_F$. And similarly, $\mathfrak{G}_E\ \mathbf{F}^{-1}\mathbf{D}_2\mathbf{S}_1\ \mathfrak{G}_F$ if $\mathfrak{E} \preccurlyeq_2 \mathfrak{F}$, and $\mathfrak{G}_E\ \mathbf{S}_2^{-1}\mathbf{D}_2\mathbf{D}_1^{-1}\mathbf{S}_1\ \mathfrak{G}_F$ if $\mathfrak{E} \preccurlyeq \mathfrak{F}$.

To complete the proof, notice that **FIF** $=$ **IF**, and in particular $\mathbf{S}_i$**IF** $=$ **IF** $(i = 1, 2)$. □

This ends the proof of thm. 7 — and of everything up to theorem 12 included. □

3.7. Vector orderings.

Lemma 44. *Given an inclusion $K_1 \subseteq K_2$, the inclusion $\Pi(K_1) \subseteq \Pi(K_2)$ (prop. 17.4) is order-preserving for any of the orders $\preccurlyeq, \preccurlyeq_1$ and $\preccurlyeq_2$.*

Proof. Obvious (cf. also lemma 19). □

Corollary 45. *Assume K completely regular (this is not needed for the convexity part of the conclusions).*

1. *The subset of R's (with arbitrary marginals) in $\Delta(\Omega\times\Omega')$ satisfying theorem 12 is closed and convex.*
2. *The graphs of $\preccurlyeq$, $\preccurlyeq_1$ and $\preccurlyeq_2$ in $\Delta(\Omega) \times \Delta(\Omega)$ are closed and convex.*

Remark 46. Point 2 proves cor. 13.

Proof. 1. Our set equals $\{R \in \Delta(\Omega \times \Omega^1) \mid$ marginals belong to Π, (1), (2), (3), (4)$\}$. Now condition (1) is equivalent to $R(O \times O') = 0$ for any pair of disjoint open sets O and O' in K, so determines a closed convex subset. Also Π is closed and convex in $\Delta(\Omega)$, as mentioned before, so since the map from $\Delta(\Omega \times \Omega')$ to $\Delta(\Omega)$ is affine and continuous, the condition that the marginals belong to Π determines also a closed, convex subset.

Finally, the conditional independence conditions can be rewritten as

$$E[\varphi(\omega)\psi(\theta_2, \theta'_2)] = E[\theta_2(\varphi)\psi(\theta_2, \theta'_2)] \tag{2}$$

$$E[\varphi(\omega')\psi(\theta_1, \theta'_1)] = E[(\theta'_1(\varphi)\psi(\theta_1, \theta'_1)] \tag{3}$$

$$(\preccurlyeq_1)\colon E[\varphi(\omega')\psi(\theta_2, \theta'_2)] = E[\theta'_2(\varphi)\psi(\theta_2, \theta'_2)] \tag{4}$$

for bounded borel functions φ and ψ (bounded continuous functions in the completely regular case). Since for φ bounded (resp., continuous), $\theta_i(\varphi)$ is so too, it follows that conditions (2), (3) and (4) are of the form that a family of bounded borel (resp., continuous) affine functions of R vanishes — so, the set of solutions is convex (and closed).

2. The convexity part follows immediately from 1. The closedness also does — first in the compact case, since a continuous image of a compact set is compact, next in the completely regular case, by lemma 44. □

Lemma 47. Π *is a simplex, i.e., the linear subspace it spans of the space of signed measures on Ω is a (complete) sublattice.*

Proof. Taking marginals on Θ_i commutes with the lattice operations, since all measures have the same conditionals on $K \times \prod_{j\neq i} \Theta_j$. □

Corollary 48.
1. *The sets $C = \{\lambda(P - Q) \mid P \succcurlyeq Q, \lambda \geq 1\}$, and C_i similarly defined with $\succcurlyeq_i$, are pointed convex cones.*
2. *$P \succcurlyeq Q \Leftrightarrow P - Q \in C,\ \forall P, Q \in \Pi$ (i.e., $\succcurlyeq$ is a "vector ordering") iff*

   $$\alpha P + (1-\alpha)R \succcurlyeq \alpha Q + (1-\alpha)R \Rightarrow P \succcurlyeq Q,\ \forall P, Q, R \in \Pi, \forall\alpha \in]0, 1[$$

   and similarly for $\succcurlyeq_i$ and C_i.
3. *For K completely regular, the cone C or C_i is closed if the corresponding order satisfies the conditions sub 2.*

Proof. 1. C is a cone, since for $0 \leq \lambda < 1$ and $P \succcurlyeq Q$, $\lambda(P-Q) = (\lambda P + (1-\lambda)R) - (\lambda Q + (1-\lambda)R)$ ($R \in \Pi$ arbitrary), which belongs to C by convexity of the graph of $\preccurlyeq$ (cor. 45). C is a convex cone since $\lambda(P-Q)+\lambda'(P'-Q') = (\lambda+\lambda')\left[\left(\frac{\lambda}{\lambda+\lambda'}P + \frac{\lambda'}{\lambda+\lambda'}P'\right) - \left(\frac{\lambda}{\lambda+\lambda'}Q + \frac{\lambda'}{\lambda+\lambda'}Q'\right)\right]$, which again belongs to C by convexity of the graph. And C is pointed because $\preccurlyeq$ is an order (anti-symmetric).

2. The condition is clearly necessary. So assume it holds, and consider $P', Q' \in \Pi$ with $P'-Q' = \lambda(P-Q)$ and $P \succcurlyeq Q$: we have to show that $P' \succcurlyeq Q'$. Let $R' = P' \wedge Q'$, $r = \|R'\|$, $R'' = \frac{1}{r}R'$, $P'' = \frac{1}{1-r}(P'-R')$, $Q'' = \frac{1}{1-r}(Q'-R')$ (and say $R'' \in \Pi$ arbitrary if $r = 0$, and assume w.l.o.g. that $r < 1$). Then P'', Q'' and R'' belong to Π by lemma 47, and $P' = (1-r)P'' + rR''$, $Q' = (1-r)Q'' + rR''$ — so, by convexity of the graph, it suffices to prove that $P'' \succcurlyeq Q''$. Since $P'' - Q'' = \frac{\lambda}{1-r}(P-Q)$, we are reduced to the initial problem, but where P' and Q' are in addition mutually singular. Then $\lambda P \geq P' - Q'$ and $\lambda P \geq 0$ imply, by the mutual singularity, that $\lambda P \geq P'$, so $\lambda \geq 1$ and $\lambda P - P' = \lambda Q - Q' = (\lambda-1)R$ with $R \in \Pi$. I.e., $P = \frac{1}{\lambda}P' + \frac{\lambda-1}{\lambda}R$, and $Q = \frac{1}{\lambda}Q' + \frac{\lambda-1}{\lambda}R$, hence by one condition, $P \succcurlyeq Q$ does indeed imply $P' \succcurlyeq Q'$.

3. Consider first the case where K is compact. C being convex, to prove that it is weak-closed it suffices to show that its intersection with every closed ball is so, using e.g. cor. 22.7 in [**?**]. Let thus $R_\alpha = \lambda_\alpha(P_\alpha - Q_\alpha)$ be a bounded net in C, converging say to R (in the space of signed measures on Ω). By our condition sub 2 and lemma 47, we can remove the common part of P_α and Q_α, and still preserve their order (after renormalizing): i.e., we can assume that P_α and Q_α are mutually singular. Since $\|P_\alpha - Q_\alpha\| = 2$, it follows that λ_α is bounded. Fix an ultrafilter on α, and let $P_\alpha \to P$, $Q_\alpha \to Q$, $\lambda_\alpha \to \lambda$ according to this ultrafilter (compactness, ...). Then, by closedness of the graph, $P \succcurlyeq Q$, hence indeed $R = \lambda(P-Q) \in C$. Consider finally the completely regular case: embedding K into its Stone-Ĉech compactification $\widetilde{K}$, the result follows by lemma 44 from the previous case. □

Remark 49. Points 2 and 3 of the corollary imply, by the separation theorem, that $\preccurlyeq$ (or $\preccurlyeq_i$) is a vector ordering iff it is generated by the monotone continuous affine functionals (i.e., $P \succcurlyeq Q$ iff $\varphi(P) \geq \varphi(Q)$ for every $\preccurlyeq$-monotone continuous affine functional φ on Π).

Lemma 50. *A subset of* Π *is tight iff the set of its marginals on* K *is so.*

Proof. We prove the lemma even in the I-person case. The set of marginals of a tight set is obviously always tight. For the converse, the set of marginals being tight means there exists a l.s.c. function $\varphi_0 \colon K \to \mathbb{R}_+$ such that $\varphi_0 \geq 1$, $\{x \in K \mid \varphi_0(x) \leq L\}$ is compact $\forall L \in \mathbb{R}$, and $\exists M \in \mathbb{R}$: $\int \varphi_0 dP \leq M, \forall P$ in our subset S. h Let then inductively $\psi_n^i = \theta_i(\varphi_n) \forall i \in I$, $\varphi_{n+1} = \varphi_n + \frac{1}{2^n}\frac{1}{\#I}\sum_{i\in I}\psi_n^i$: $S \subseteq \Pi$ implies $\int \psi_n^i dP = \int \varphi_n dP$ $\forall P \in S$, hence $\int \varphi_{n+1} dP = (1+\frac{1}{2^n})\int \varphi_n dP$, so $\int \varphi_n dP \leq M\Pi_{k=1}^n(1+\frac{1}{2^k}) \leq eM$. Also inductively, each ψ_n^i and hence each φ_n is l.s.c., so, with $\varphi = \lim_n \varphi_n$, we get that $\varphi \colon \Omega \to \mathbb{R}$ is l.s.c., ≥ 1, satisfies (monotone convergence) $\int \varphi dP \leq e.M$ $\forall P \in S$, and finally, $\forall L, \{\omega \mid \varphi(\omega) \leq L\}$ is compact: using [**?** , thm. 1.1.3 p. 108]), and observing that, by induction, φ_{n+1} depends only on $\omega_n = ((\theta_{i,n})_{i\in I}, k)$, one gets inductively over n, first that $K_{n,L}^0 = \{\omega_{n-1} = ((\theta_{i,n-1})_{i\in I}, k) \mid \varphi_n(\omega_{n-1}) \leq L\}$ is compact $\forall L$, hence that $K_{i,n,L} = \{\theta \in \Theta_{i,n} \mid \theta(\varphi_n) \leq L\}$ is compact $\forall(i,L)$ by Prohorov's criterion, and thus $K_{n+1,L}^0$ is compact, being a closed subset (l.s.c. of φ_{n+1}) of the product of compact sets $K_{n,L}^0 \times \Pi_{i\in I}K_{i,n,2^n(\#I)L}$. Therefore $\{\omega \mid \varphi(\omega) \leq L\}$, being a closed subset (projective limit, and lower-semi-continuity of φ) of the product of compact sets $K_{0,L}^0 \times \Pi_{n=0}^\infty \Pi_{i\in I}K_{i,n,2^n(\#I)L}$, is also compact. Thus S is tight. □

Proof of corollary 2.1. Since $P \preccurlyeq Q$ implies that P and Q have the same marginal on K (e.g. by thm. 12.1), all P_α, in the monotone net have the same marginal on K. Hence, by the above lemma, the P_α are tight — thus, by Prohorov, relatively compact in $\Delta(\Omega)$. So the net has limit points. Let P be any such limit points: by closedness of the graph, $\forall\alpha_0, \forall\alpha \geq \alpha_0$, $P_\alpha \geq P_{\alpha_0}$ goes to the limit and implies $P \in \lim P_\alpha \geq P_{\alpha_0}$ (implying $P \in \Pi$): so $P \geq P_\alpha\ \forall\alpha$. Then if P' is another limit point, we also have $P' \geq P_\alpha$, $\forall\alpha$ — hence, going to the limit over α (closedness of the graph again), $P' \geq P$. Thus dually $P \geq P'$ also, and hence (anti-symmetry) $P = P'$: the limit point is unique. Together with relative compactness of the net, this implies that the net converges. □

3.8. **Barycentres.**

Lemma 51. *Let $X = \Delta(Y)$ with Y Hausdorff, and $\mu \in \Delta(X)$. The σ-additive measure $\bar\mu$ defined by $\bar\mu(B) = \int x(B)\mu(dx)$ for $B \in \mathcal{B}^Y$ is τ-smooth.*

Proof. If $0 \leq f_\alpha \to f$ is an increasing net of l.s.c. functions on Y, $\bar\mu(f_\alpha) \nearrow \bar\mu(f)$. □

Definition 52. *$\precsim$ is the partial order on $\Delta(X)$ defined by $\mu \precsim \nu$ iff $\mu(f) \leq \nu(f)$ for all f convex l.s.c. bounded from below.*

Lemma 53. *Let $X = \Delta(Y)$ with Y Hausdorff. The following definitions of the barycentre $\bar\mu$ of $\mu \in \Delta(X)$ are equivalent:*

(1) *$\bar\mu \in X$ is such that $\delta_{\bar\mu} \precsim \mu$.*
(2) *$\bar\mu$ defined as a σ-additive measure by $\bar\mu(B) = \int x(B)\mu(dx)$ for $B \in \mathcal{B}^Y$ belongs to $\Delta(Y)$.*

Proof. (2) ⇒ (1) ("Jensen's lemma"). When μ has a compact support on which the restriction of ϕ is continuous, approximating it by probability measures with finite support yields the result. (Recall that the topology on X is defined as the weakest topology for which the measures of open sets are l.s.c. functions and that X is a Hausdorff space itself under this topology. This implies indeed that if the μ_α converge to μ, then their barycentres converge to the barycentre of μ. Since ϕ is continuous on the support of μ, $\int \phi d\mu_\alpha \to \int \phi d\mu$. On the other hand ϕ l.s.c. implies $\liminf \phi(\bar\mu_\alpha) \geq \phi(\bar\mu)$.)

In general, let K_n be a sequence of disjoint compact sets that exhaust μ and such that the restriction of ϕ to each K_n is continuous (use Lusin's theorem). Let μ_n be the normalized restriction of μ to K_n, and $\alpha_n = \mu(K_n)$. Note that $\bar\mu_n$ is in X because $\bar\mu = \sum \alpha_n \bar\mu_n$ as σ-additive barycentres, hence the sum being tight, each of the summands is also tight. For each n, $\int \phi d\mu_n \geq \phi(\bar\mu_n)$, and since each of the members is bounded below, this inequality extends to the sum: $\int \phi d\mu \geq \sum \alpha_n \phi(\bar\mu_n)$. To prove that $\sum \alpha_n \phi(\bar\mu_n) \geq \phi(\bar\mu)$, remains thus only to prove the result for the case of a measure μ with countable support.

We now view α as a measure with countable support on X. The measures $\beta_n = \frac{1}{\sum_1^n \alpha_i} \sum_1^n \alpha_i \delta_{x_i}$ converge in norm to α, therefore the barycentres $\bar\beta_n$ of β_n converge to the barycentre $\bar\alpha$ of α. To see that $\sum_n \alpha_n \phi(x_n) \geq \phi(\bar\alpha)$, remark that for each m, $\sum_1^m \alpha_n \phi(x_n) \geq (\sum_1^m \alpha_n)\phi(\bar\beta_m)$, that $\sum_1^m \alpha_n \phi(x_n) \to \sum_n \alpha_n \phi(x_n)$ by boundedness below of ϕ, and $\liminf(\sum_1^m \alpha_n)\phi(\bar\beta_m) \geq \phi(\bar\alpha)$ since ϕ is l.s.c.

(1) ⇒ (2). Let $\bar\mu$ be as in (1), and $\tilde\mu$ be defined by $\tilde\mu(B) = \int x(B)\mu(dx)$ for $B \in \mathcal{B}^Y$. Then, for h l.s.c. and bounded from below on Y, f given by $f(x) = x(h)$ is convex l.s.c. and bounded from below, so $\bar\mu(h) = f(\bar\mu) \leq \int f(x)\mu(dx) = \int x(h)\mu(dx) = \tilde\mu(h)$. Hence $\bar\mu(O) \leq \tilde\mu(O)$ for every open set O, so $\bar\mu(B) \leq \tilde\mu(B)$ for all $B \in \mathcal{B}^Y$ since $\tilde\mu$ is τ-smooth. Hence $\tilde\mu = \bar\mu$. □

Lemma 54. *$\mu \in \Delta(X)$ has a barycentre if and only if $\bar\mu$ is carried by a K_σ.*

Proof. Immediate since any τ-smooth measure on a K_σ is tight. □

Lemma 55. *Assume $\mu \precsim \nu$. Then μ has a barycentre if and only if ν has one, and both coincide.*

Proof. If μ has barycentre $\bar{\mu}$, then $\delta_{\bar{\mu}} \precsim \mu \precsim \nu$ so that ν has barycentre $\bar{\mu}$. Assume ν has a barycentre $\bar{\nu}$ and let $\bar{\mu}(B) = \int x(B)\mu(dx)$. For h bounded l.s.c. on Y, define f convex, bounded l.s.c. on X by $f(x) = x(h)$. Then $\bar{\mu}(h) = \mu(f) \le \nu(f) = \bar{\nu}(h)$. For h bounded borel, let (h_n) be a decreasing sequence of l.s.c. functions $\ge h$ s.t. $\bar{\nu}(h_n)$ converges to $\bar{\nu}(h)$: $\bar{\mu}(h) \le \bar{\mu}(h_n) \le \bar{\nu}(h_n)$. Hence $\bar{\mu}(h) \le \bar{\nu}(h)$ for all bounded borel h, thus $\bar{\mu} = \bar{\nu}$. So $\bar{\mu} \in \Delta(Y)$. □

3.9. **Cartier's Theorem.**

Proposition 56. *Assume either μ or ν have a barycentre. Then $\mu \precsim \nu$ if and only if there exists $P \in \Delta(X \times X)$ that has μ and ν as marginals and such that for every bounded borel function h on Y, $E_P(x_2(h)|x_1) = x_1(h)$ P_1-a.s.*

Proof. We first show that it suffices to prove the proposition assuming both μ and ν have a barycentre. In the direction where P has to be constructed, use lemma 55. In the other direction, note that $\int x_1(B)\mu(dx_1) = \int E_P(x_2(B)|x_1)\mu(dx_1) = E_P x_2(B) = \int x_2(B)\nu(dx_2)$, and apply lemma 53 (2) .

Assume first Y is compact. For the "if" part use Jensen's theorem. For the "only if" part, theorem 35 p. 288 of [**?**] yields a measure θ on D_0 with barycentre μ, ν, where D_0 is the set of pairs $(\delta_x, \eta) \in \Delta(X)^2$ such that $\delta_x \precsim \eta$. Define now P by $\int h(x,y)dP = \int h(x,y)\eta(dy)\theta(dx, d\eta)$. Obviously P has (μ, ν) as marginals. For f affine and continuous, $E_P(f(y)|\eta, x) = \int f(y)\eta(dy) = f(x)$ P a.s. and thus $E_P(f(y)|x) = f(x)$ P_1 a.s. This holds when $f(\mu) = \mu(h)$ for h continuous, and by taking the limit for h Baire. This generalizes to h borel, since any such function is the sum of a Baire function and one which is negligible for both μ and ν.

We extend the proposition from Y compact to locally compact. Let Y be locally compact, and Y' its Alexandroff compactification, $X' = \Delta(Y')$. Since Y is borel (open) in Y', X and $\Delta(X)$ are borel in X' and $\Delta(X')$. It suffices to show that each l.s.c. convex bounded below f on X is the restriction of such a map on X'; this is because every convex l.s.c. function on X is a sup of integrals of bounded continuous functions on Y that converge at infinity.

We extend the proposition from Y locally compact to countable disjoint unions of compact sets. Let thus $Y = \cup_n K_n$, where (K_n) is a family of disjoint compact sets, and let Y' be Y endowed with the topology with as open sets those whose intersection with each K_n is open in K_n. Y' is locally compact. Since the topology on K_n is unchanged, and since the K_n are borel both in Y and in Y', the borel sets and the tight measures on Y and Y' are the same, i.e. X' is X endowed with a stronger topology. As Y' is K-analytic so are X' and $\Delta(X')$, and the continuous canonical injection from $\Delta(X')$ to $\Delta(X)$ is onto, cf. [**?** , 9.b.3 p. 428] and [**?** , 9.c p. 429]: $\Delta(X')$ is a reinforced topology on $\Delta(X)$, and so is $\Delta(X' \times X')$ on $\Delta(X \times X)$.

Remains thus only to show that the order on measures is unchanged, i.e., that if $\mu(f) \le \nu(f)$ for all convex l.s.c. f, bounded below on X, the same holds on X'. Since X' is completely regular (locally compact), such an f on X' is a sup of integrals of bounded continuous functions. And since μ and ν are tight, integrals go to the limit along increasing nets of l.s.c. functions. Suffices thus to consider $f(x) = \mathsf{max}_{i=1\ldots n} x(\varphi_i)$, where the φ_i are bounded continuous functions on Y'.

Let now $M = \mathsf{sup}_{i,y} \varphi_i(y)$, and $\varphi_i^k = \varphi_i$ on K_l for $l \le k$, and $= M$ for $l > k$ and let $f^k = \mathsf{max}_{i=1\ldots n} x(\varphi_i^k)$. Each φ_i^k is l.s.c. on Y hence f^k is convex l.s.c. bounded below on X. Hence the inequality for the f^k, so for f by monotone convergence.

We now prove the general case. Let $\bar{\mu}$ and $\bar{\nu}$ be the barycentres of μ and ν, and (K_n) be a sequence of disjoint compact sets in Y that exhaust $\bar{\mu}+\bar{\nu}$. Let $Y' = \cup_n K_n$ and $X' = \Delta(Y')$. Note that Y' is a borel subspace of Y hence X' is a borel subspace of X and, by the same argument, $\Delta(X')$ is a borel subspace of $\Delta(X)$.

For one direction, assume $P \in \Delta(X \times X)$ having the stated properties and observe first since the marginals μ, ν of P belong to $\Delta(X')$, $P \in \Delta(X' \times X')$, and has the stated properties relative to X'. Hence that remains to show that $\mu \precsim \nu$ relative to X whenever the same holds relative to X'. This is because restrictions to X' of convex l.s.c. functions on X have the same properties on X'.

In the other direction, given $\mu \precsim \nu$ on X, we first want to prove the same relation holds on X'. Since $\bar{\mu}, \bar{\nu} \in X'$ it follows that $\mu, \nu \in \Delta(X')$. To prove that $\mu \precsim \nu$ on X' let φ be bounded from below l.s.c. convex on X', let $\bar{\varphi} = \varphi$ on X' and $\bar{\varphi} = +\infty$ on $X - X'$, let $\hat{\varphi}(x) = \liminf_{y \to x} \bar{\varphi}(y)$ on X. $\hat{\varphi}$ is clearly l.s.c. bounded below (X is Hausdorff), and $\hat{\varphi}_{|X'} = \varphi$ follows from φ l.s.c. on X'. Remains the convexity: obviously $\bar{\varphi}$ is convex. Let $x_1, x_2 \in X$ and $0 < \beta < 1$, $x_{1,\alpha} \to x_1$ and $x_{2,\alpha} \to x_2$ s.t. $\bar{\varphi}(x_{i,\alpha}) \to \hat{\varphi}(x_i)$, and let $z = \beta x_1 + (1-\beta)x_2$, $z_\alpha = \beta x_{1,\alpha} + (1-\beta)x_{2,\alpha}$. For U open in Y, $\liminf z_\alpha \geq z(U)$ follows from $\liminf x_{i,\alpha} \geq x_i(U)$, hence by definition of the weak topology $z_\alpha \to z$. Since $\hat{\varphi}$ is l.s.c. and convex, $\hat{\varphi}(z) \leq \liminf \bar{\varphi}(z_\alpha) \leq \beta\hat{\varphi}(x_1) + (1-\beta)\hat{\varphi}(x_2)$. So, we proved that every convex l.s.c. bounded from below map on X' is the restriction of a such map on X, and the converse is straightforward. This shows $\mu \precsim \nu$ on X'. The proposition on X' yields $P \in \Delta(X' \times X')$ with the desired properties. Since $X' \times X'$ is a subspace of $X \times X$, P has the desired properties in $\Delta(X \times X)$. □

3.10. **Proof of theorem 15.** Under (a), we have $P \preccurlyeq_2 \mathfrak{E}$, and since $\mathfrak{E} \sim P_{\mathfrak{E}} = Q$ we get indeed $P \preccurlyeq_2 Q$. Similarly (b) yields $Q \preccurlyeq_1 P'$, hence the "if" part.

In the other direction, start from the distribution R in thm. 12 (with P' as Q). Let $\mathfrak{E} = (\Omega \times \Omega', R, \Theta_1, \Theta_2')$ (with the borel sets, and the obvious map to K).

Let $Q \in \Delta(\Omega'')$ be the canonical information structure associated to $(\Theta_1 \times \Theta_2' \times K, R, \Theta_1, \Theta_2')$, and ϕ the corresponding canonical map, ϕ is also canonical from $\mathfrak{E}$ to Q since the properties to be checked [**?** , thm. 2.5.1 p. 122] are the same. R and ϕ induce a (tight) probability R' on $\Omega \times \Omega' \times \Omega''$ (carried by "the diagonal of $K \times K \times K$").

For $B'' \in \mathcal{B}^{\Theta_2''}$, $\phi^{-1}(B'')$ differs from some $B' \in \mathcal{B}^{\Theta_2'}$ by a null set [**?** , thm. 2.5.1 p. 122]. Since the conditional probability of B' given Ω is Θ_2-measurable by thm. 12.2, the conditional probability of B'' given Ω is so too: Θ_2'' and Ω are conditionally independent given Θ_2. Hence, if $\rho(\cdot|\cdot)$ is a regular conditional probability on Θ_2'' given Θ_2 (tightness), then ρ is also a regular conditional probability on Θ_2'' given Ω. In particular, P and ρ induce the correct probability on $\Omega \times \Theta_2''$. Hence 15 (a), and 15 (b) is dual.

Remains to show that 15 (a) is equivalent to 15 (a') (and hence also 15 (b) to 15 (b')). Under 15 (a), let $\nu(\theta_2')(d\theta_2)$ be a regular conditional probability on Θ_2 given θ_2' under $P \otimes \rho$, in the sense of [**?** , II.1Ex.16c p. 75]. Let $\pi(\theta_2')(d\omega) = \theta_2(d\theta_1, dk)\nu(\theta_2')(d\theta_2)$. Note that, by continuity of θ_2, for any open set O in Ω, $\theta_2(O)$ is l.s.c. in θ_2 (i.e., θ_2 is also a continuous map to $\Delta(\Omega)$). Therefore, for any borel set B in Ω, $\theta_2(B)$ is borel measurable, and hence $\pi(\theta_2')(B)$ is well defined, and borel measurable. It follows then immediately that π is a borel transition probability from Θ_2' to Ω. Further, consider now an increasing net O_α of open sets in Ω, with union O. The $\theta_2(O_\alpha)$ form then, as argued above, an increasing net of l.s.c. functions, and converge pointwise to $\theta_2(O)$ by regularity of θ_2. So, by regularity of $\nu(\theta_2')$, $\pi(\theta_2')(O_\alpha)$ increases pointwise to $\pi(\theta_2')(O)$: each $\pi(\theta_2')$ is "τ-smooth", so to prove its tightness, remains only to show it is carried by a K_σ. Note that, under $P \otimes \rho$, θ'_2 and ω are independent given θ_2, so for B borel in Ω, $Prob(B|\theta_2, \theta'_2) = P(B|\theta_2) = \theta_2(B)$ (consistency of P). Thus $Prob(B|\theta'_2) = \int \theta_2(B)\nu(\theta'_2)(d\theta_2) = \pi(\theta'_2)(B)$: π is the conditional probability on Ω given θ'_2 under $P \otimes \rho$. Therefore, let B be a K_σ in Ω with $P(B) = 1$: one must also have $\pi(\theta'_2)(B) = 1$ a.e., so, redefining $\nu(\theta'_2)$ on the exceptional set, we get now that each $\pi(\theta'_2)$ is tight. Let then $\bar{\nu}(\theta'_2)$ denote the marginal of $\pi(\theta'_2)$ on $\Theta_1 \times K$: it is tight too, hence in Θ_2 (by its homeomorphism with $\Delta(\Theta_1 \times K)$), and is the barycentre of $\nu(\theta'_2)$. Thus each $\nu(\theta'_2) \in \Delta(\Theta_2)$ indeed has a barycentre $\bar{\nu}(\theta'_2)$ in Θ_2.

We now show that the map $\bar{\nu}\colon \Theta'_2 \to \Theta_2$ is, under $P \otimes \rho$, borel-measurable, and induces a tight distribution $\mu \in \Delta(\Theta_2)$ on the borel sets of Θ_2. Observe that the map from $\nu(\theta'_2) \in \Delta(\Theta_2)$ to $\pi(\theta'_2) \in \Delta(\Omega)$ is continuous (this is just on the range of ν, since elsewhere the values might not even belong to $\Delta(\Omega)$), by the continuity of θ_2 (argument as above). And the map from $\pi(\theta'_2)$ to its marginal $\bar{\nu}(\theta'_2)$ is clearly continuous. So the borel measurability of ν to $\Delta(\Theta_2)$, and the tightness of the induced distribution on $\mathcal{B}^{\Delta(\Theta_2)}$, are preserved by composition with those continuous maps.

For ϕ on Θ_2 convex l.s.c. and bounded below, we apply lemma 53.1, with Θ_2 $(= \Delta(K \times \Theta_1))$ for X, and obtain $\int \phi(\theta_2)\nu(\theta'_2)(d\theta_2) \geq \phi(\bar{\nu}(\theta'_2))$. Both sides of the inequality are borel-measurable w.r.t. θ'_2, by our measurability properties for ν and $\bar{\nu}$; since they are also bounded below, we can integrate the inequality w.r.t. θ'_2. The repeated integral in the left hand member becomes then just $\int \phi(\theta_2)P(d\theta_2)$, since ϕ is P-integrable — and hence $P \otimes \rho$-integrable with the same integral. And by definition of μ, the right hand side becomes just $\int \phi d\mu$: our inequality is established.

Remains thus only to prove that $P_\mu = Q$. By definition, $P_\mu = P_{\mathfrak{E}_\mu}$, where $\mathfrak{E}_\mu$ equals Ω endowed with $\theta_2(d\theta_1, dk)\mu(d\theta_2)$. And $Q = P_{\mathfrak{E}}$ (where player 2 is informed only of θ'_2). Now in $\mathfrak{E}$, $\bar{\nu}(\theta'_2)$, being the posterior of 2 on $\Theta_1 \times K$, is a sufficient statistic for 2, so $P_{\mathfrak{E}} = P_{\mathfrak{E}'}$, where $\mathfrak{E}'$ equals $\mathfrak{E}$ except that player 2 is only informed of $\bar{\nu}(\theta'_2)$. Now the joint distribution under $\mathfrak{E}'$ of $(\theta_1, \bar{\nu}(\theta'_2), k)$ equals $\theta_2(d\theta_1, dk)\mu(d\theta_2)$, thus $\mu \in \Delta_b(\Theta_1)$, P_μ is well defined, and $P_{\mathfrak{E}'} = P_{\mathfrak{E}_\mu}$, and hence our equality.

To prove that (a') implies (a), observe that P_2 has a barycentre: the marginal of P on $\Theta_1 \times K$. So, by lemma 55, μ also has a barycentre, and in particular $\mu \in \Delta_b(\Theta_2)$: P_μ is well defined. And proposition 56 yields $R \in \Delta(\Theta_2 \times \Theta'_2)$, with P_2 and μ as respective marginals, such that $E(\theta_2(h)|\theta'_2) = \theta'_2(h)$ μ a.e. for every h borel bounded on $\Theta_1 \times K$. Let ρ be the conditional under R on Θ'_2 given Θ_2. We know that $P_\mu = Q$ and need to prove that $P_{\mathfrak{E}} = Q$, where $\mathfrak{E}$ is the information scheme on $(\Omega \times \Theta'_2, P \otimes \rho)$ where player 2 observes θ'_2 only. Let now $\bar{P}$ denote $R \otimes \theta_2$: since $R \in \Delta(\Theta_2 \times \Theta'_2)$ and θ_2 is continuous from Θ_2 to $\Delta(\Theta_1 \times K)$, $\bar{P}$ is τ-smooth on $\mathcal{B}^{\Omega \times \Theta'_2}$, with $\int h(\omega, \theta'_2)d\bar{P} = \mathsf{E}_R \int h(\omega, \theta'_2)\theta_2(d\omega)$ $\forall h \geq 0$ borel on $\Omega \times \Theta'_2$ — in particular, $\mathcal{B}^\Omega$ and $\mathcal{B}^{\Theta_2 \times \Theta'_2}$ are conditionally independent given Θ_2 under $\bar{P}$. Observe finally that $\bar{P}$ has P as marginal on Ω since the marginal of R on Θ_2 is P_2, and hence $\bar{P} \in \Delta(\Omega \times \Theta'_2)$, being τ-smooth and having tight marginals P on Ω and R on $\Theta_2 \times \Theta'_2$. By the conditional independence, ρ is also the conditional probability on Θ'_2 given Ω under $\bar{P}$.

So $P \otimes \rho$ is well defined on $\mathcal{B}^\Omega \otimes \mathcal{B}^{\Theta'_2}$ and is the restriction of $\bar{P}$ to that σ-field. Thus $\mathfrak{E}$ is equivalent $(\mathbf{D}^{-1})$ to the information scheme $(\Omega \times \Theta'_2, \bar{P})$ in which player 2 only observes θ'_2. Let $\tilde{P}$ be the marginal of $\bar{P}$ on $\tilde{\Omega} = \Theta_1 \times \Theta'_2 \times K$. $\tilde{P} \in \Delta(\tilde{\Omega})$ and the marginal of $\tilde{P}$ on $\tilde{\Theta}_2$ is μ, the marginal of R on Θ'_2. Now $\mathfrak{E}$ becomes equivalent $(\mathbf{D})$ to $(\tilde{\Omega}, \tilde{P})$. Remains to show that $(\tilde{\Omega}, \tilde{P})$ is also the information scheme $\mathfrak{E}_\mu$ induced by μ.

I.e., that $\forall B \in \mathcal{B}^\Omega$, $\tilde{P}(B) = \int \theta_2(B)\mu(d\theta_2)$. Let P' denote the right-hand member. Since $\mu \in \Delta(\Theta_2)$ and since θ_2 is continuous from Θ_2 to $\Delta(\Omega)$, P' is τ-smooth on $\mathcal{B}^\Omega$. For $B = B_1 \times B_2$ with $B_1 \in \mathcal{B}^{\Theta_1 \times K}$ and $B_2 \in \mathcal{B}^{\Theta_2}$ this means:

$E_{\bar{P}}[\mathbb{I}_{B_1}|\theta_2'] = \theta_2'(B_1)$. The left hand equals $E_{\bar{P}}[E_{\bar{P}}[\mathbb{I}_{B_1}|\theta_2, \theta_2']|\theta_2']$. Since by the conditional independence above $E_{\bar{P}}[\mathbb{I}_{B_1}|\theta_2, \theta_2'] = \bar{P}(B_1|\theta_2)$, $= P(B_1|\theta_2)$ P being the marginal on Ω, $= \theta_2(B_1)$ since $P \in \Pi$, the left hand member equals $\mathsf{E}_R[\theta_2(B_1)|\theta_2']$, R being the marginal on $\Theta_2 \times \Theta_2'$, $= \theta_2'(B_1)$ by the property of R. This proves the particular case. Thus P' is τ-smooth on $\mathcal{B}^\Omega$, $\tilde{P} \in \Delta(\Omega)$, and $P'(B_1 \times B_2) = \tilde{P}(B_1 \times B_2)$ for all $B = B_1 \times B_2$ with $B_1 \in \mathcal{B}^{\Theta_1 \times K}$ and $B_2 \in \mathcal{B}^{\Theta_2}$. This extends immediately to finite unions of such sets, since every such finite union can be re-written as a disjoint finite union. In particular, $P'(B) = \tilde{P}(B)$ whenever B is a basic open set (i.e., a finite union of products of an open set in $\Theta_1 \times K$ and an open set in Θ_2). Hence, by τ-smoothness, this extends to every open B, and then to every B borel. $\square$

References