THREE ESSAYS ON EPISTEMIC GAME THEORY

by

Nicodemo De Vito

A Thesis submitted to the faculty of
University of Siena
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

Department of Economics

University of Siena

April 2010

ABSTRACT

The main subject of this dissertation is the theory of type structures in epistemic game theory. The three, closely linked chapters are an investigation to the fundamental question of the representation of belief hierarchies in strategic environments.

On the Existence of a Universal Type Space

Abstract

In this paper we present a generalization of Brandenburger and Dekel (1993)'s construction of the universal type space assuming that (i) the underlying space of nature states is a Hausdorff topological space, and (ii) players' hierarchical beliefs are Radon probability measures. Our construction scheme is compared to the contruction already given in Mertens and Zamir (1985), and we show that the two approaches lead to an equivalent epistemic characterization of the universal type space.

JEL Classification: C70, D80, D82.

Keywords: Hierachies of Beliefs, Type spaces, Epistemic Game Theory.

Ambiguous Types and Possibility Structures

Abstract

In this paper, we present the construction of a space of states of the world and the corresponding type structure describing all the uncertainty facing each player in a strategic environment. The construction employs a hierarchy of multiple beliefs, rather than a single belief as in the standard Harsanyi's model. Under very mild restrictions on the players' beliefs and parameter spaces, we provide a sufficiently general framework which includes Ahn's (2007) space of hierarchies of ambiguous beliefs as a special case, and we identify epistemically the universal Harsanyi's type structure as a substructure. We also show the existence of a universal, complete possibility structure for any space of basic uncertainty, and discuss connections among different existing constructions of type structures.

JEL Classification: C70, D80, D82.

Keywords: Hierarchies of beliefs, ambiguity, type structures.

Which Hierarchies of Beliefs Belong to the Universal Type Structure?

Abstract

This paper introduces the notion of strong coherence for hierarchies of beliefs in games, and shows that, in a measure-theoretic setup, the universal type structure constructed in Heifetz and Samet (1998) consists of all hierarchies of beliefs displaying strong coherence and common certainty of strong coherence. Additionally, it establishes the existence of a canonical measure-theoretic isomorphism, and shows that strong coherence corresponds to the classical notion of coherence used in topological settings as in Brandenburger and Dekel (1993).

JEL Classification: C70, D80, D82.

Keywords: Hierarchies of Beliefs, Type structures, Epistemic game theory

Contents

Chapter 1

On the existence of a universal type space

1.1 Introduction

Games with incomplete information describe strategic environments where some payoff-relevant states are not common knowledge among the players. Harsanyi [28] points out that, in principle, the Bayesian analysis of every game with incomplete information should require a model in which each player is endowed with an infinite belief hierarchy: a belief about the payoff-relevant states (i.e., a first order belief), a belief about other players' beliefs about the payoff-relevant states (i.e., a second order belief), and so on. Taking this observation into account, Harsanyi introduces the notion of *type spaces*, the elements of which are simply called *types*, as a parsimonious model to represent the belief hierarchies. Each element of a player's type space is associated with a probability measure defined over the product of the payoff-relevant state space and the type spaces of her opponents. As is shown in [30], such a probability measure summarizes all the mutual certainties and uncertainties of the players, and each belief hierarchy induced by a player's type is *coherent*, in the sense that the marginals of higher-level beliefs coincide with the corresponding lower-level beliefs. However, Harsanyi's proposal of using type spaces as a model of interactive beliefs in games leaves open the opposite question: how can a type space be built from coherent belief hierarchies?

Two schemes coexist in game theory literature for modelling belief hierarchies as types.

The first construction scheme, introduced by Mertens and Zamir [44] (MZ), uses the concept of projective limit for showing that every coherent belief hierarchy of a player admits a unique limit extension to a probability measure over the payoff-relevant state space and the other players' coherent hierarchies. It turns out that the space T_{MZ} of all coherent belief hierarchies is a *universal* type space. That is, T_{MZ} is a type space on its own right, and every conceivable type space can be embedded in T_{MZ} in a manner preserving the structure of higher order beliefs. The second construction scheme, developed by Brandenburger and Dekel [17] (BD), relies instead on identifying the universal type space with the largest subset T_{BD} of belief hierarchies which, in their terminology, satisfies common knowledge of coherence. Both kinds of constructions are based on topological assumptions on the payoff-relevant state space, namely, compact Hausdorff in MZ and Polish (i.e., complete separable metric) in BD.

In this paper, we extend the BD-like construction of the universal type space to more general, topological cases. Specifically, we show that the analysis can be carried out under the following assumptions: (i) the payoff-relevant state space is Hausdorff, and (ii) higher order beliefs are described by Radon probability measures. Two reasons motivate this study. First, as we shall point out formally in Section 1.3, the existence of a BD universal type space makes implicit use of a compactness assumption for the payoff-relevant state space. Since BD require only that the underlying space be Polish, this poses the question as to whether a universal type space can be built from hierarchies of beliefs without any reference to compactness. This question is not new. Mertens, Sorin and Zamir [43] and Heifetz [29] have already provided a generalization of the MZ approach under various other topological assumptions. However, the main disadvantage of their hierarchic construction via projective limit is that it is not explicit about the epistemic characterization of the universal type space, namely, common knowledge of coherence. This leads to the second motivation of this paper, that is, how MZ and BD construction schemes relate to each other.

Our main result (Proposition 2) fills this gaps and sheds light on the relationship between the aforementioned approaches. The strategy of the proof (see Section 1.4) is the following. Since, as we shall carefully explain, the BD approach does not suffice alone to formally prove the existence of a universal type space in absence of compactness (that is, T_{BD} may be empty), we start by defining an appropriate, projective sequence of sets of coherent belief hierarchies.

We show that this sequence has a non-empty projective limit, by appealing to a well known theorem of Bourbaki ([15]). Finally, a simple proof by induction shows that this projective limit set satisfies common knowledge of coherence, and it is maximal in the sense that every other set of hierarchies displaying common knowledge of coherence is contained in it.

Viewed from a slightly different angle, the main contribution of this paper establishes the formal equivalence of the spaces T_{MZ} and T_{BD}.[1] Indeed, the method of the proof of Proposition 2 relies on an adaptation of the one in MZ, with the remarkable difference that our method allows for a wider variety of structural assumptions on the parameter space and probability measures. This result has two important implications. First, it shows that the construction scheme by projective sequences of probability spaces, although not explicit about the notion of common knowledge of coherence, is necessary to formally prove the existence of a universal type space. Second, the generality of the proof allows to widen some BD-like constructions of "set theoretic" type spaces to a large extent. For instance, a more general version of the universal type space for possibility models in Mariotti, Meier and Piccione ([42]) can be provided along the main lines of the proof scheme outlined here, as we will show in Chapter II.

This paper is organized as follows. Section 1.2 contains some preliminary definitions and mathematical results. Section 1.3 presents the construction of the universal space. The proof of the main result (Proposition 2) is the subject of Section 1.4. Finally, Section 1.5 concludes with some remarks concernig the relationship between our framework and related constructions of types. Omitted proofs are collected in the appendix, which contains further matematical definitions and results needed for Section 1.4.

1.2 Preliminaries and notation

For any non-empty topological space X, let $\mathcal{B}(X)$ denote its Borel σ-field and $\mathcal{P}(X)$ the set of all Borel σ-additive probability measures on $\mathcal{B}(X)$. A *Radon probability measure* on X is a measure $\mu \in \mathcal{P}(X)$, such that for every $A \in \mathcal{B}(X)$ and every $\epsilon > 0$, there exists a compact set $K \subseteq A$ such that $\mu(A \backslash K) < \epsilon$. Denote by $\Delta(X)$ the space of Radon probability measures on X. If D is a non-empty, Borel subset of X, define $\Delta(D) = \{\mu \in \Delta(X) \,|\, \mu(D) = 1\}$.

[1] To the best of our knowledge, the current paper should be the first one which presents a proof on the epistemic equivalence between the spaces T_{MZ} and T_{BD}. As is shown in Section 1.4, this is far from being immediate.

The set $\Delta(X)$ is endowed with the *narrow topology*, which is defined as the coarsest topology on $\Delta(X)$ for which all the maps $\mu \to \int_X f d\mu$ from $\Delta(X)$ into $\mathbb{R}$ are lower semi-continuous, as f varies in the set of all bounded, lower semi-continuous functions on X.[2] As a subbasic system of neighborhoods, this topology assigns to each $\mu_0 \in \Delta(X)$ the sets of the form

$$V(\mu_0; \epsilon, f) = \left\{ \mu \in \Delta(X) \,|\, \int_X f d\mu \geq \int_X f d\mu_0 - \epsilon, \; f \text{ is lower semi-continuous, } \epsilon > 0 \right\}.$$

For a non-empty set $O \subseteq X$, let $\mathbf{1}_O$ denote the characteristic function on O, i.e., $\mathbf{1}_O(x) = 1$ if $x \in O$, and $\mathbf{1}_O(x) = 0$ otherwise. Since every characteristic function on an open subset of X is lower semi-continuous, the narrow topology on $\Delta(X)$ can be defined by means of the base generated by all sets of the form

$$V(\mu_0; \epsilon, \mathbf{1}_O) = \{\mu \in \Delta(X) \,|\, \mu(O) \geq \mu_0(O) - \epsilon, \; O \subseteq X \text{ open, } \epsilon > 0\}.$$

For an overview of the Radon measures and the narrow topology, see Topsoe [57] and Schwartz [54]. The following lemma collects some properties we shall repeatedly make use of.

Lemma 1 *Let X be a non-empty topological space, and D a non-empty subset of X. The following statements hold true.*

1. *If X is Hausdorff, so is $\Delta(X)$.*

2. *$\Delta(X)$ is compact if and only if X is compact.*

3. *If X is Polish, then $\Delta(X) = \mathcal{P}(X)$. Furthermore, $\Delta(X)$ is also Polish.*

4. *If X is Hausdorff and D is closed, then $\Delta(D)$ is closed.*

Proof. For Claim 1, see [54], Proposition 2, p. 371, or [57], Theorem 11.2. Claim 2 is a consequence of the Hahn Banach Theorem and Riesz Representation Theorem; see [54], p. 379, and the notes to §11 in Topsoe [57], p. 76. For the first part of Claim 3, see [12], Theorem 7.1.7., while for the second part, see [2], Theorem 15.15. To prove the fourth claim, remark

[2] The narrow topology was first introduced by Topsoe [57] under the name "weak topology". The terminology adopted in this Chapter is that of Schwartz [54]. In the language of probability theory, the term "weak topology" usually refers to the weak* topology as defined in the current Chapter.

that the narrow topology is also the weakest topology on $\Delta(X)$ for which the map $\mu \mapsto \int_X h d\mu$ is upper semicontinuous whenever h is upper semicontinuous. Since D is closed, $\mathbf{1}_D$ is upper semicontinuous. Thus by definition of the narrow topology, the map $\mu \mapsto \int_X \mathbf{1}_D d\mu = \mu(D)$ is upper semicontinuous and therefore $\Delta(D)$ is narrowly closed. Trivially $\Delta(D)$ is non-empty and convex. ■

According to Lemma 1, if X is a non-empty Hausdorff topological space, then $\Delta(X)$ is non-empty: since any singleton is compact in a Hausdorff space, $\Delta(X)$ contains all the probability measures concentrated on a finite or countable number of points. Let $\delta : X \to \Delta(X)$ be the function which maps each $x \in X$ to its Dirac measure $\delta(x)$. Clearly, $\delta(X) \subseteq \Delta(X)$ and the function δ is a homeomorphism onto its image.

For any non-empty, topological space X, the *weak* topology* on $\Delta(X)$ is the coarsest topology for which the map $\mu \to \int_X \phi d\mu$ from $\Delta(X)$ into $\mathbb{R}$ is continuous, as ϕ varies in the set of all bounded, continuous functions on X. By definition, the weak* topology is coarser than the narrow topology on $\Delta(X)$, and it is not be appropriate for an arbitrary Hausdorff space X since there might not exist sufficiently many continuous functions on it. Indeed, if $\Delta(X)$ is endowed with the weak* topology, Lemma 1.(1) does not hold; see appendix 1.6.2 for an example. The following result identifies the topological structure of the base space X for which the narrow topology and the weak* topology on $\Delta(X)$ are equivalent (for a proof, see [57], Theorem 8.1).

Lemma 2 *If X is a completely regular space (e.g., compact Hausdorff or Polish), the narrow topology on $\Delta(X)$ coincides with the weak* topology on $\Delta(X)$.*

Let X and Y be arbitrary measure spaces, and for any measurable $f : X \to Y$, let $\mathcal{L}_f : \Delta(X) \to \mathcal{P}(Y)$ denote the image measure on Y induced by f, defined by $\mathcal{L}_f(\mu)[E] = \mu\left(f^{-1}(E)\right)$ for any $\mu \in \Delta(X)$ and any Borel set $E \subseteq Y$. It is easy to check that $\mathcal{L}_f$ is well defined, i.e., $\mathcal{L}_f(\mu) \in \mathcal{P}(Y)$ for any $\mu \in \Delta(X)$. However, we have the following result:

Lemma 3 *Let X, Y be Hausdorff topological spaces and $f : X \to Y$ be continuous. Then $\mathcal{L}_f$ is a continuous function from $\Delta(X)$ into $\Delta(Y)$. Furthermore,*

(i) if f is injective, so is $\mathcal{L}_f$;

(ii) if f is surjective and open, then $\mathcal{L}_f$ is surjective;

Proof. We first show that $\mathcal{L}_f(\mu) \in \Delta(Y)$ for any $\mu \in \Delta(X)$. Let F be a Borel subset of Y. By continuity of f, $f^{-1}(F)$ is a Borel subset of X. Thus, for any $\varepsilon > 0$ there exists a compact $C \subseteq f^{-1}(F)$ such that $\mu\left(f^{-1}(F)\setminus C\right) < \varepsilon$. Since f is continuous, the set $K = f(C)$ is compact, hence $C \subseteq f^{-1}(K)$. Using the fact that $f^{-1}(F \setminus K) = \left(f^{-1}(F)\right) \setminus \left(f^{-1}(K)\right) \subseteq \left(f^{-1}(F)\right) \setminus C$, we get

$$
\begin{aligned}
\mathcal{L}_f(\mu)[F\backslash K] &= \mu\left(f^{-1}(F\backslash K)\right) \\
&\leq \mu\left(f^{-1}(F)\backslash C\right) < \varepsilon.
\end{aligned}
$$

To show continuity, let h be a bounded, lower semi-continuous function on Y. Clearly, the composition $h \circ f$ is a bounded, lower semi-continuous function on X. Assume that $\mu \in \Delta(X)$ belongs to the (sub-basic) neighborhood $V(\mu_0; \epsilon; h \circ f)$ of $\mu_0 \in \Delta(X)$. Then $\int_X (h \circ f)\, d\mu \geq \int_X (h \circ f)\, d\mu_0 - \epsilon$, and by the Change of Variables Theorem ([2], Theorem 13.46) $\int_Y h d\mathcal{L}_f(\mu) \geq \int_Y h d\mathcal{L}_f(\mu_0) - \epsilon$, hence $\mathcal{L}_f(\mu)$ belongs to the neighborhood $V(\mathcal{L}_f(\mu_0); \epsilon; h)$ of $\mathcal{L}_f(\mu_0) \in \Delta(Y)$.

(i) Let f be injective, and suppose $\mathcal{L}_f(\mu_1) = \mathcal{L}_f(\mu_2)$ for μ_1, $\mu_2 \in \Delta(X)$. Then, for every compact set $K \subseteq X$,

$$
\begin{aligned}
\mu_1(K) &= \mathcal{L}_f(\mu_1)[f(K)] \\
&= \mathcal{L}_f(\mu_2)[f(K)] \\
&= \mu_2(K),
\end{aligned}
$$

where the first equality follows from the fact that f is an injection. The result $\mu_1 = \mu_2$ follows from the definition of Radon probability measure.

(ii) Since f is surjective, there exists (by the Axiom of Choice) an injective, right inverse function $g : Y \to X$ such that $f \circ g = Id_Y$, where Id_Y denotes the identity on Y. We claim that g is continuous. First remark f is an open, continuous surjection: this implies that each subset O of Y is open if and only if $f^{-1}(O)$ is an open subset of X. Let A be an non-empty, open subset of X. Thus $A = f^{-1}(B)$ for some non-empty $B \subseteq Y$, B open. To prove that g is continuous it suffices to show that $g^{-1}(A)$ is open. But this follows, since

$$\begin{aligned} g^{-1}(A) &= \left(g^{-1} \circ f^{-1}\right)(B) \\ &= (f \circ g)^{-1}(B) \\ &= (Id_Y)^{-1}(B) \\ &= B. \end{aligned}$$

It follows from the previous steps of the proof that the function $\mathcal{L}_g : \Delta(Y) \to \Delta(X)$ is well defined, continuous and injective. For an arbitrary $\upsilon \in \Delta(Y)$, define $\mu = \mathcal{L}_g(\upsilon)$. The function $\mu \mapsto \mathcal{L}_f(\mu)$ is such that

$$\mathcal{L}_f(\mu)[A] = \mu\left(f^{-1}(A)\right) = \upsilon\left(g^{-1}\left(f^{-1}(A)\right)\right) = \upsilon(A),$$

for all $A \in \mathcal{B}(Y)$. In other words, $\mathcal{L}_f(\mu) = \upsilon$. Since υ is arbitrary, the conclusion follows. ■

Given a countable product space $\prod_{n\in\mathbb{N}} X_n$ where $\mathbb{N}$ is the set of natural numbers, we write Proj_{X_m} as the canonical projection from $\prod_{n\in\mathbb{N}} X_n$ to X_m, for each $m \in \mathbb{N}$. For any $l, m \in \mathbb{N}$ satisfying $l \leq m$, we write $\mathrm{Pr}_{l,m}$ as the coordinate projection from $\prod_{j=1}^{m} X_j$ into $\prod_{j=1}^{l} X_j$. We consider any product, finite or countable, of topological spaces as a topological space with the product topology. If X and Y are arbitrary Hausforff topological spaces, then the marginal measure of $\mu \in \Delta(X \times Y)$ on X is defined as $\mathrm{marg}_X \mu = \mathcal{L}_{\mathrm{Proj}_X}(\mu)$, which is, by Lemma 3.(ii), a continuous function from $\Delta(X \times Y)$ onto $\Delta(X)$.

1.3 The model

Fix a two-player set I; given a player $i \in I$, we denote by $-i$ the other player in I. Let S be a non-empty space describing payoff-relevant states that each player is uncertain about. The space S is called parameter space. The following assumption will be mantained throughout the paper.

Assumption 1 S is a Hausdorff topological space.

Since each player $i \in I$ is uncertain about the realization of the payoff-relevant state, she must have a prior (probability distribution) on the parameter space S; such a prior is called first order belief. However, a first order belief does not exaust all the uncertainty faced by each player: player i realises that player $-i$ has a first order belief on S as well, and this prior is unknown to her. Consequently, player i's second order belief is a probability distribution over S and the space of $-i$'s first order beliefs. Continuing in this fashion, each player is completely characterized by an infinite hierarchy of beliefs.

We impose the following restriction on belief hierarchies of the players.

Assumption 2 Players' hierarchical beliefs are described by Radon probability measures.

Formally, for each $i \in I$ define inductively the sequence of spaces $\left\{X_n^i\right\}_{n=0}^{\infty}$ by

$$X_0^i = S, \tag{1.3.1}$$

$$X_{n+1}^i = X_n^{-i} \times \Delta\left(X_n^{-i}\right);\ n \geq 0. \tag{1.3.2}$$

An element $t_{n+1}^i \in \Delta\left(X_n^i\right)$ is a $(n+1)$-order belief; one can easily show that, according to our notation,

$$X_{n+1}^i = X_0^i \times \prod_{l=0}^{n} \Delta\left(X_l^{-i}\right).$$

The set of all possible, infinite belief hierarchies for player i is $T_0^i = \prod_{n=0}^{\infty} \Delta\left(X_n^i\right)$. A *type* is an infinite belief hierarchy, and is denoted by $t^i = \left(t_1^i, t_2^i, ...\right)$. Lemma 1 implies that for all $n \geq 0$, X_{n+1} and $\Delta\left(X_n^i\right)$ are Hausdorff spaces. It follows that T_0^i, endowed with the product topology, is also Hausdorff. An analogous conclusion holds if the parameter space S is assumed to be Polish or compact.

Definition 1 *A type $t^i \in T_0^i$ is coherent if and only if, for any $n \geq 1$,*

$$\text{marg}_{X_{n-1}^{-i}}(t_{n+1}^i) = t_n^i.$$

The coherence condition simply requires that different level of beliefs assign the same probability to the same event. The space of all coherent types for player $i \in I$ is denoted by

T_1^i.

Note that since there is a common uncertainty space S, the sets T_0^i and T_1^i are the copies of the same sets T_0 and T_1, respectively. Therefore, from now on, we drop the superscript i (or $-i$) for notational simplicity where no confusion results.

Lemma 4 *T_1 is a closed subset of T_0.*

Proof. Let $(t^\alpha) = (t_1^\alpha, t_2^\alpha, ...)$ be a net in T_1 converging in the product topology to $t = (t_1, t_2, ...)$, that is, $t_n^\alpha \to t_n$ with respect to the narrow topology, for any $n \geq 1$. We have to show that $t \in T_1$, i.e., $\text{marg}_{X_{n-1}}(t_{n+1}) = t_n$ for any $n \geq 1$. By Lemma 1.(3), $\text{marg}_{X_{n-1}} = \mathcal{L}_{\text{Proj}_{X_{n-1}}}$ is a continuous function, hence for all $n \geq 1$, $t_n^\alpha \to t_n$ implies $\text{marg}_{X_{n-1}}(t_{n+1}^\alpha) \to \text{marg}_{X_{n-1}}(t_{n+1})$, which proves the claim. ■

The following proposition generalizes its counterpart in BD under weaker topological assumptions. It establishes that each coherent type of a player can be equivalently described as a single belief about the parameter space and the space of (coherent or not) belief hierarchies of her opponent.

Proposition 1 *There exists a homeomorphism $f : T_1 \to \Delta(S \times T_0)$ such that*

$$\forall n \geq 0,\ \text{marg}_{X_n}\left[f\left((t_1, t_2, ...)\right)\right] = \text{Proj}_{\Delta(X_n)}\left((t_1, t_2, ...)\right).$$

This result is an immediate consequence of the following lemma, whose proof is contained in the appendix (see also Theorem 2).

Lemma 5 *Let $\{Z_n\}_{n=0}^{\infty}$ be a collection of Hausdorff spaces, and let*

$$D = \left\{ (\mu_1, \mu_2, ...) \middle| \begin{array}{c} \mu_n \in \Delta\left(Z_0 \times ... \times Z_{n-1}\right),\ \forall n \geq 1,\ \textit{and} \\ \text{marg}_{Z_0 \times ... \times Z_{n-2}} \mu_n = \mu_{n-1},\ \forall n \geq 2 \end{array} \right\}. \tag{1.3.3}$$

Then there exists a homeomorphism $f : D \to \Delta\left(\prod_{n=0}^{\infty} Z_n\right)$ such that

$$\forall n \geq 1,\ \text{marg}_{Z_0 \times ... \times Z_{n-2}}\left[f\left((\mu_1, \mu_2, ...)\right)\right] = \mu_{n-1}.$$

Proof of Proposition 1. Let $Z_0 = X_0 = S$, $\forall n \geq 1$, $Z_n = \Delta(X_{n-1})$. By Lemma 1, each Z_n is Hausdorff, and so is $\Delta\left(\prod_{n=0}^{\infty} Z_n\right) = \Delta(S \times T_0)$ with the product topology. The space T_1 corresponds to the space D in (1.3.3). The result follows immediately from Lemma 5. ■

Although the coherence condition rules out the possibility that higher order beliefs contradict each other, it does not exauste all the uncertainty faced by each player. Proposition 1 states that player i's type induces a belief over $-i$'s types, but it does not exclude hierarchies such that i believes that $-i$'s beliefs are not coherent (recall that $f(t) \in \Delta(S \times T_0)$). That is, i's type does not induce a belief over $-i$'s beliefs over i's type. The intended interpretation is the following: if player i's belief hierarchies are represented by $f(t)$, then she is coherent, but she does not know that her opponent is. It follows that a type completely encodes all possible beliefs (including beliefs over beliefs over types) if it satisfies *common knowledge of coherence*.[3]

Formally, we say that player i, endowed with type t, knows event $E \in \mathcal{B}(S \times T_0)$ if $f(t) \in \Delta(E)$. Define inductively the following sets:

$$\begin{aligned} T_{k+1} &= \{t \in T_1 \,|\, f(t) \in \Delta(S \times T_k)\}, \; k \geq 1, \\ T &= \cap_{k \geq 1} T_k. \end{aligned}$$

The set $T \times T$ is the space of pair of types which display coherence and common knowledge of coherence.

It is easy to check (use Lemma 1 and Lemma 4, and then proceed by induction) that $\{T_k\}_{k \geq 1} = \left\{f^{-1}(\Delta(S \times T_{k-1}))\right\}_{k \geq 1}$ is a sequence of non-empty, nested closed sets. This raises the question: is T non-empty? In general, a decreasing sequence of non-empty sets may have an empty intersection. But in the particular case this sequence of sets is in a compact space, the intersection is non-empty.[4] Neverthless, the following result shows that T is non-empty even if the compactness assumption does not hold, and T is a universal type space, i.e., a Harsanyi type space which is both terminal and complete (Siniscalchi, [55]). That is, T embeds all other type spaces as a belief-closed subsets, and at same time, every subjective belief on T is associated

[3] In general, the term "knowledge" refers to justified true belief that comes from logical deduction. It is therefore conceptually different from probability one belief ("certainty"). For expositional ease, we adopt the same terminology as in BD.

[4] A similar drawback is also faced by Mariotti, Meier and Piccione [42] in their BD-like construction of a universal possibility model. This problem motivates their choice of assuming a compact parameter space.

with a type. The product space $\Omega = S \times T \times T$ is called *Universal Belief Space*, and it is the space of all possible states of the world.

Proposition 2 *T is non-empty. Moreover, the restriction of f to T induces a homeomorphism $g : T \to \Delta(S \times T)$.*

A simple proof of Proposition 2 could be provided along the lines of that one in BD, *assuming that T is non-empty.* In this case, simply observe

$$\begin{aligned} T &= \cap_{k\geq 1} f^{-1}\left(\Delta\left(S \times T_{k-1}\right)\right) \\ &= f^{-1}\left(\cap_{k\geq 1}\Delta\left(S \times T_{k-1}\right)\right), \end{aligned}$$

and since $\cap_{k\geq 1}\Delta\left(S \times T_{k-1}\right) = \Delta\left(S \times \left(\cap_{k\geq 1}T_{k-1}\right)\right)$,[5] and f is onto, then $f(T) = \Delta(S \times T)$, which implies the existence of the desired homeomorphism. Clearly, without the compactness assumption on S, this proof is far from being complete. It follows that a general proof of Proposition 2 must be carried out in a completely different fashion, as shown in the following section.

1.4 Proof of Proposition 2

For each player $i \in I$, let $\Theta_0^i = S$, $\Upsilon_1^i = \Delta\left(\Theta_0^i\right)$, and for all $n \geq 1$,

$$\begin{aligned} \Theta_n^i &= \Theta_0^i \times \Upsilon_n^{-i}; \\ \Upsilon_{n+1}^i &= \left\{\left(t_1^i, ..., t_n^i, t_{n+1}^i\right) \in \Upsilon_n^i \times \Delta\left(\Theta_n^i\right) : \mathrm{marg}_{\Theta_{n-1}^i} t_{n+1}^i = t_n^i\right\}. \end{aligned}$$

The space Θ_n^i is player i's domain of uncertainty of level $n+1$: it consists of the parameter space and what player $-i$ believes about the payoff-relevant state, what player $-i$ believes about what player i believes about the payoff-relevant state,..., and so on, up lo level n. The set Υ_{n+1}^i consists of $(n+1)$-tuples of coherent beliefs over $\Theta_0^i, ..., \Theta_n^i$. Furthermore, not only

[5] If $\mu \in \cap_{k\geq 1}\Delta\left(S \times T_{k-1}\right)$, then $\mu\left(S \times T_{k-1}\right) = \mu\left(S \times \left(\cap_{k\geq 1}T_{k-1}\right)\right) = 1$, because a countable intersection of probability one events as probability one, hence $\mu \in \Delta\left(S \times \left(\cap_{k\geq 1}T_{k-1}\right)\right)$. Conversely, suppose that $\mu \in \Delta\left(S \times \left(\cap_{k\geq 1}T_{k-1}\right)\right)$. Since $\left(\cap_{k\geq 1}T_{k-1}\right) \subseteq T_{k-1}$, it follows that $\Delta\left(S \times \left(\cap_{k\geq 1}T_{k-1}\right)\right) \subseteq \Delta\left(S \times T_{k-1}\right)$, thus $\mu \in \Delta\left(S \times T_{k-1}\right)$ for each $k \geq 1$, i.e., $\mu \in \cap_{k\geq 1}\Delta\left(S \times T_{k-1}\right)$.

that each i's beliefs are coherent but she also considers only coherent beliefs of $-i$ (only those are in support of her beliefs). This implies that, compared to the system of spaces defined by (1.3.1)-(1.3.2), the two sequences of spaces, $\left\{\Upsilon_n^i\right\}$ and $\left\{\Theta_n^i\right\}$, are such that

$$\Theta_n^i \subseteq X_n^i,\ \Upsilon_{n+1}^i \subseteq \prod_{l=0}^{n} \Delta\left(X_l^i\right),\ \forall n \geq 0.$$

For each $i \in I$, $n \geq 1$, let $\pi_{n,n+1}^i : \Upsilon_{n+1}^i \rightarrow \Upsilon_n^i$ denote the projection on the factor spaces of the sequence $\left\{\Upsilon_n^i\right\}$. The projection $\sigma_{n-1,n}^i : \Theta_n^i \rightarrow \Theta_{n-1}^i$ satisfies

$$\sigma_{n-1,n}^i = \begin{cases} \Pr_{0,1}^i & n = 1 \\ \left(Id_{\Theta_0^i}; \pi_{n-1,n}^{-i}\right) & n \geq 2 \end{cases}.$$

Clearly, $\sigma_{n-1,n}^i$ is the restriction of $\Pr_{n-1,n}^i : X_n^i \rightarrow X_{n-1}^i$ to the subspace Θ_n^i. The following Lemma implies that this restriction is also onto Θ_{n-1}^i.[6]

Lemma 6 *For all $n \geq 1$, $\pi_{n,n+1}^i : \Upsilon_{n+1}^i \rightarrow \Upsilon_n^i$ is onto.*

Proof. For every $\left(t_1^i, ..., t_n^i\right) \in \Upsilon_n^i$ we have to find a $t_{n+1}^i \in \Delta\left(\Theta_n^i\right)$ such that $\left(t_1^i, ..., t_n^i, t_{n+1}^i\right) \in \Upsilon_{n+1}^i$, i.e., $\text{marg}_{\Theta_{n-1}^i} t_{n+1}^i = \mathcal{L}_{\sigma_{n-1,n}^i}\left(t_{n+1}^i\right) = t_n^i$. That is, any coherent n-level hierarchy can be extended to a coherent $(n+1)$-level hierarchy.

For $n = 1$, define $\Psi_1^i : \Theta_0^i \rightarrow \Theta_1^i = \Theta_0^i \times \Delta\left(\Theta_0^{-i}\right)$ by

$$\Psi_1^i(s) = \left(Id_{\Theta_0^i}(s), \delta^{-i}(s)\right)$$

for $s \in S$. The function Ψ_1^i is continuous, hence, by Lemma 3, $\mathcal{L}_{\Psi_1^i} : \Delta\left(\Theta_0^i\right) \rightarrow \Delta\left(\Theta_1^i\right)$ is also continuous and well defined. For each $t_1^i \in \Delta\left(\Theta_0^i\right)$, $\mathcal{L}_{\Psi_1^i}\left(t_1^i\right)$ represents player i's second order belief according to which she thinks: (i) the parameter $s \in S$ is governed by the probability distribution t_1^i, and (ii) player $-i$ is certain that some $s \in A$ occurred, for some $A \in \mathcal{B}(S)$.

[6] Mertens and Zamir ([44], Proposition 2.10) provide a related result for belief spaces (cf. [44], Definition 2.2). Their method of the proof relies on the Hahn-Banach and Riesz representation theorems, which cannot work in the present case as the compactness property for the spaces under consideration is lacking. Our proof applies a technique which generalizes the one used by Heifetz ([29])

Indeed,

$$\begin{aligned}
\operatorname{marg}_{\Theta_0^i}\mathcal{L}_{\Psi_1^i}\left(t_1^i\right)(B) &= \mathcal{L}_{F_1^i}\left(t_1^i\right)\left(\left(\Pr_{0,1}^i\right)^{-1}(B)\right) \\
&= t_1^i\left(\left(\Psi_1^i\right)^{-1}\circ\left(\Pr_{0,1}^i\right)^{-1}(B)\right) \\
&= t_1^i\left(\left(\Pr_{0,1}^i\circ\Psi_1^i\right)^{-1}(B)\right) \\
&= t_1^i\left(Id_{\Theta_0^i}(B)\right) \\
&= t_1^i(B),
\end{aligned}$$

for all $B \in \mathcal{B}(S)$, as required. Hence, $\pi_{1,2}^i$ is onto.

Suppose that, for each $i \in I$, for $k = 1, ..., n-1$, we have already found a continuous function $\Psi_k^i : \Theta_{k-1}^i \to \Theta_k^i$ such that $\operatorname{marg}_{\Theta_{k-1}^i}\mathcal{L}_{\Psi_k^i}\left(t_k^i\right) = t_k^i$, which implies that $\pi_{k,k+1}^i$ and $\sigma_{k,k+1}^i$ are onto. Define $\Psi_{k+1}^i : \Theta_k^i \to \Theta_{k+1}^i$ by

$$\Psi_{k+1}^i\left(s, \left(t_1^{-i}, ..., t_k^{-i}\right)\right) = \left(s, \left(t_1^{-i}, ..., t_k^{-i}, \mathcal{L}_{\Psi_k^{-i}}(t_k^{-i})\right)\right),$$

which is continuous since, by the induction hypothesis, Ψ_k^{-i} is continuous and hence, by Lemma 3, $\mathcal{L}_{\Psi_k^{-i}}$ is also continuous. It is easy to check that $\sigma_{k,k+1}^i \circ \Psi_{k+1}^i = Id_{\Theta_k^i}$. So the measure $\mathcal{L}_{\Psi_{k+1}^i}\left(t_{k+1}^i\right) \in \Delta\left(\Theta_{k+1}^i\right)$ is the required extension of t_{k+1}^i, since for all $E \in \mathcal{B}\left(\Theta_k^i\right)$,

$$\begin{aligned}
\operatorname{marg}_{\Theta_k^i}\mathcal{L}_{\Psi_{k+1}^i}\left(t_{k+1}^i\right)(E) &= \mathcal{L}_{\Psi_{k+1}^i}\left(t_{k+1}^i\right)\left(\left(\sigma_{k,k+1}^i\right)^{-1}(E)\right) \\
&= t_{k+1}^i\left(\left(\Psi_{k+1}^i\right)^{-1}\circ\left(\sigma_{k,k+1}^i\right)^{-1}(E)\right) \\
&= t_{k+1}^i(E).
\end{aligned}$$

■

The families $\left\{\Upsilon_n^i, \pi_{n,n+1}^i\right\}_{n\geq 1}$ and $\left\{\Theta_{n-1}^i, \sigma_{n-1,n}^i\right\}_{n\geq 1}$ are projective sequences of non-empty Hausdorff spaces and surjective bonding maps (in fact, projections); hence, by Theorem 1 in Appendix 6.1, their projective limits, $\varprojlim\left\{\Upsilon_n^i, \pi_{n,n+1}^i\right\}_{n\geq 1}$ and $\varprojlim\left\{\Theta_{n-1}^i, \sigma_{n-1,n}^i\right\}_{n\geq 1}$, are non-empty. Next, define

$$\Upsilon^i = \left\{\left(t_1^i, t_2^i, ...\right) \in T_0^i \,\middle|\, \left(t_1^i, ..., t_n^i\right) \in \Upsilon_n^i,\ \forall n \geq 1\right\},$$
$$\Theta^i = S \times \Upsilon^{-i}.$$

By Lemma 4, Υ^i is a closed subset of T_1^i; furthermore both Υ^i and Θ^i can be identified as the projective limits of the sequences $\left\{\Upsilon_n^i, \pi_{n,n+1}^i\right\}_{n\geq 1}$ and $\left\{\Theta_{n-1}^i, \sigma_{n-1,n}^i\right\}_{n\geq 1}$, respectively, as stated in the following

Claim 1 *Υ^i and Θ^i are homeomorphic to $\varprojlim \left\{\Upsilon_n^i, \pi_{n,n+1}^i\right\}_{n\geq 1}$ and $\varprojlim \left\{\Theta_{n-1}^i, \sigma_{n-1,n}^i\right\}_{n\geq 1}$, respectively.*

The proof of Claim 1 is contained in Appendix 1.6.4.

Note that there is a common uncertainty space S, hence the sets Θ^i and Υ^i are the copies of the same sets Θ and Υ, respectively. As in Section 1.3, if no confusion may arise, we omit the superscript i (or $-i$). We show now how the set Υ can be identified with T.

First notice that, for every $(t_1, t_2, ...) \in \Upsilon$ the structure $\left(\{\Theta_n\}_{n\geq 1}, \{t_n\}_{n\geq 1}, \{\sigma_{n,n+1}\}_{n\geq 1}\right)$ is a projective sequence of Radon probability measures such that, for all $n \geq 1$,

$$\mathcal{L}_{\sigma_{n,n+1}}\left(t_{n+1}\right) = t_n.$$

Denote by $\overline{\sigma}_n$ the projection from Θ to Θ_n. According to Theorem 2 in Appendix 6.1, for every $(t_1, t_2, ...) \in \Upsilon$ there exists a unique Radon probability measure $t \in \Delta(\Theta) = \Delta(S \times \Upsilon)$ such that, for all $n \geq 1$,

$$\mathcal{L}_{\overline{\sigma}_n}(t) = t_n.$$

In other words, there exists a bijective map $g : \Upsilon \to \Delta(S \times \Upsilon)$ such that, for every hierarchy $(t_1, t_2, ...) \in \Upsilon$,

$$\mathrm{marg}_{\Theta_{n-1}}\left(g\left(t_1, t_2, ...\right)\right) = t_n.$$

The map g can be taken as the restriction of $f : T_1 \to \Delta(S \times T_0)$ to Υ, since f satisfies the required properties. Indeed, we show that $\Upsilon = T$. First remark that Υ is closed in T_1 by coherence of beliefs (Lemma 4), hence, by Lemma 1, $\Delta(S \times \Upsilon)$ is a closed subset of $\Delta(S \times T_1)$.

The homeomorphism f of Proposition 1 yields $f(\Upsilon) \subseteq \Delta(S \times T_0)$; however, $f(\Upsilon) = g(\Upsilon) = \Delta(S \times \Upsilon) \subseteq \Delta(S \times T_1)$, hence $\Upsilon \subseteq f^{-1}(\Delta(S \times T_1)) = T_2$. Assume, by way of induction, that $\Upsilon \subseteq T_k$. Since $f(\Upsilon) = \Delta(S \times \Upsilon) \subseteq \Delta(S \times T_k)$, we get $\Upsilon \subseteq f^{-1}(\Delta(S \times T_k)) = T_{k+1}$. It follows that $\Upsilon \subseteq T_k$, for all $k \geq 1$, hence $\Upsilon \subseteq T = \cap_{k\geq 1} T_k$. Since the sequence $\{T_k\}_{k\geq 1}$ is decreasing, $\Upsilon = T$, as required.

1.5 Concluding remarks

1.5.1 On the role of Assumptions 1 and 2

The main technical tool for the construction of a universal type space from coherent belief hierarchies is Lemma 5. Assumptions 1 and 2 are crucial for this result to hold. Assumption 2 is automatically satisfied in BD by Lemma 1.(3), while is not stated directly in MZ, where the players' beliefs are described by general Borel probability measures. Specifically, the MZ framework is based on the *incorrect* claim that $\mathcal{P}(X)$ is compact whenever X is compact ([44], p. 357). A counterexample given by Dieudonné (see [12], Example 7.1.3) shows that even on a compact Hausdorff space a Borel probability measure may fail to be Radon; and Lemma 1.(2) states that, for X to be compact, the compactness of $\Delta(X)$ is necessary. This implies that Assumption 2 cannot be dispensed with in MZ framework.

1.5.2 Topologies on the space of probability measures

Blau [8] introduced another topology on the space of positive measures (not necessarily Radon) on a given set which is stronger than the narrow topology. Let X be a topological space. Let M_X be the space of all positive measures on X. Blau's topology on M_X is the topology generated by sub-basic neighborhoods of the form

$$\mathcal{O}(\mu_0; O, \epsilon) := \{\mu \in M_X | \mu(O) > \mu_0(O) - \epsilon \text{ and } |\mu(X) - \mu_0(X)| < \epsilon\},$$

where O is an open subset of X and $\epsilon > 0$. Consider the space M_X^r of finite, positive Radon measures equipped with Blau's topology (clearly, $M_X^r \subseteq M_X$). The main result proved by Blau is that if X is *normal*, the weak* topology on M_X^r coincides with Blau's topology. This result is stated in [29] in terms of the inherited topological structure on $\Delta(X)$. It is worth noting that, even if Blau's topology is finer than the narrow topology, their subspace topologies on $\Delta(X) \subseteq M_X^r$ coincide. Theorem 5* in Heifetz [29] should be replaced by Lemma 2 in Section 2, which establishes the correct equivalence between the weak* topology (and hence the narrow topology) and Blau's topology in terms of the weaker assumption of complete regularity for the topological space X.[7]

1.5.3 Borel structures on $\Delta(X)$

The crucial feature of the narrow topology is that it leads to inheritance of the Hausdorff property from X to $\Delta(X)$. The first justification for the choice of the narrow topology is pragmatic, that is, it meets the conditions of Lemma 5, which delivers our main result. However, an additional perspective on the narrow topology may be provided by considering the Borel structure it generates on the space $\Delta(X)$. For this, we recall the notion of belief operators (cf. [46]): for an event (measurable set) $E \subseteq S \times T^{-i}$, let $B_i^p(E)$ be the event "player i of type t^i believes that the probability of E is at least p", i.e.,

$$B_i^p(E) = \left\{ \left(s, t^{-i}\right) \in S \times T^{-i} : g\left(t^i\right)(E) \geq p \right\}.$$

As each set of the form $B_i^p(E)$ represents the object of beliefs of player $-i$, it must be measurable. To satisfy this requirement, the space $\Delta(X)$ must be endowed with the σ-field $\Sigma_{\Delta(X)}$ generated by all sets of the form

$$b^p(E) = \{\mu \in \Delta(X) : \mu(E) \geq p\}$$

where E is a (Borel) measurable subset of X and $0 \leq p \leq 1$. Lemma 24 in Chapter III states that, if X is a Hausdorff space and $\Delta(X)$ is endowed with the narrow topology, then $\Sigma_{\Delta(X)} = \mathcal{B}(\Delta(X))$. This result does *not* hold if $\Delta(X)$ is endowed with the weak* topology: in this case, the example in Appendix 6.2. shows that $\mathcal{B}(\Delta(X))$ might be the trivial σ-field, hence $\mathcal{B}(\Delta(X)) \subseteq \Sigma_{\Delta(X)}$. In this sense, the narrow topology on $\Delta(X)$ is the "right" topology

[7] Recall that a normal topological space is also completely regular, but not vice versa.

for every hierarchic construction of types under the assumptions formulated in this paper.

1.5.4 The measure-theoretic case

The Radon property of Borel probability measures ensures the σ-additivity of the limit measure in projective sequences of measure spaces. Heifetz and Samet [31] construct a parameter space S, satisfying Assumption 1, and a coherent hierarchy of beliefs of a player that has no σ-additive coherent extension over S and the coherent hierarchies of her opponent. This counterexample is based on a classical result of Andersen and Jessen (see [12], Example 7.7.3), which shows that the Radon property of Borel probability measures cannot be omitted in generalized versions of Kolmogorov's extension theorem (like Theorem 2 in the Appendix). In the general case in which the parameter space S is measurable, Heifetz and Samet [30] offer a construction of the universal type space in a different fashion. They distiguish the universal type space from the set of coherent hierarchies, and, by their counterexample in [31], they conclude that the former is in general a proper subset of the latter. In Chapter III, we shall provide a full characterization of the hierarchies of beliefs containted in Heifetz and Samet's type space. Such hierarchies satisfy a simple refinement of the notion of coherence, which we call *strong coherence*. Proposition 13 in Chapter III shows that the coherent hierarchies of beliefs considered in the current Chapter are also strongly coherent.

1.6 Appendix

1.6.1 Summary of projective limit theory

In this appendix, we provide some of the background definitions and results from the theory of projective limit spaces, that are necessary to present the proof of Lemma 5 and the results in Section 4. For a more thorough treatment see [22] or [50]. As it is costumary, we denote by $\mathbb{N}$ the set of natural numbers.

A *projective sequence* is a family $\{Y_n, f_{m,n}\}_{m,n\in\mathbb{N}}$ of spaces Y_n and functions $f_{m,n} : Y_n \to Y_m$ such that:

- for each $n \in \mathbb{N}$, Y_n is non-empty topological space;

- for any $m, n \in \mathbb{N}$ satisfying $m \leq n$, $f_{m,n}$ is continuous;

- $f_{m,p} = f_{m,n} \circ f_{n,p}$ for any $m, n, p \in \mathbb{N}$ satisfying $m \leq n \leq p$, and $f_{n,n} = Id_{Y_n}$ for every $n \in \mathbb{N}$.

The spaces Y_n are called *coordinate* (or *factor*) *spaces* and the maps $f_{m,n}$ are called *bonding maps*. Given a projective sequence $\{Y_n, f_{m,n}\}_{m,n\in\mathbb{N}}$, the *projective limit of* $\{Y_n, f_{m,n}\}_{m,n\in\mathbb{N}}$, denoted by $\varprojlim \{Y_n, f_{m,n}\}_{m,n\in\mathbb{N}}$, is defined as

$$\varprojlim \{Y_n, f_{m,n}\}_{m,n\in\mathbb{N}} = \left\{ \{y_n\} \in \prod_{n\in\mathbb{N}} Y_n \,|y_m = f_{m,n}(y_n)\,, \text{ for each } m, n \in \mathbb{N} \text{ s.t. } m \leq n \right\}.$$

For each $l \in \mathbb{N}$, the map $\overline{f}_l : \varprojlim \{Y_n, f_{m,n}\}_{m,n\in\mathbb{N}} \to Y_l$ is the restriction of the projection map $\mathrm{Proj}_{Y_l} : \prod_{n\in\mathbb{N}} Y_n \to Y_l$ to $\varprojlim \{Y_n, f_{m,n}\}_{m,n\in\mathbb{N}}$. Clearly, for any $m, n \in \mathbb{N}$ such that $m \leq n$, the maps $\overline{f}_n$ and $\overline{f}_m$ satisfy the equality $\overline{f}_m = f_{m,n} \circ \overline{f}_n$.

The projective limit $\varprojlim \{Y_n, f_{m,n}\}_{m,n\in\mathbb{N}}$ inherits the subspace topology as a subset of the product $\prod_{n\in\mathbb{N}} Y_n$. It is known (see [22], Proposition 2.5.1) that if $\{Y_n, f_{m,n}\}_{m,n\in\mathbb{N}}$ is a projective sequence of Hausdorff spaces Y_n then its projective limit, $\varprojlim \{Y_n, f_{m,n}\}_{m,n\in\mathbb{N}}$, is a closed subset of the Cartesian product $\prod_{n\in\mathbb{N}} Y_n$.

The next Theorem, a special case of Proposition 5, p. 198, of Bourbaki [15], provides a sufficient condition for a projective limit to be non-empty.

Theorem 1 *Let $\{Y_n, f_{m,n}\}_{m,n\in\mathbb{N}}$ be a projective sequence of topological spaces Y_n and* surjective *bonding maps $f_{m,n}$. Then, if $Y = \varprojlim \{Y_n, f_{m,n}\}_{m,n\in\mathbb{N}}$, the map $\overline{f}_n : Y \to Y_n$ is surjective for each $n \in \mathbb{N}$, and Y is non-empty provided that none of the Y_n's is empty.*

As a corollary of this theorem, it can be proved that if each bonding map $f_{m,n}$ is the coordinate projection, then the projective limit space, $\varprojlim \{Y_n, f_{m,n}\}_{m,n\in\mathbb{N}}$, can be identified (homeomorphically) with the Cartesian product $\prod_{n\in\mathbb{N}} Y_n$.

For each $n \in \mathbb{N}$, let μ_n be a Radon probability measure on the Hausdorff space Y_n. The family $\{Y_n, f_{m,n}, \mu_n\}_{m,n\in\mathbb{N}}$ is a *projective sequence of Radon probability measures* if

- $\{Y_n, f_{m,n}\}_{m,n\in\mathbb{N}}$ is a projective sequence of Hausdorff spaces Y_n;

- $\mu_m = \mathcal{L}_{f_{m,n}}(\mu_n)$, for any $m, n \in \mathbb{N}$ such that $m \leq n$.

The structure $\{Y, \overline{f}_n, \mu\}_{n\in\mathbb{N}}$ is called *measure projective limit* of the projective sequence of Radon probability measures $\{Y_n, f_{m,n}, \mu_n\}_{m,n\in\mathbb{N}}$ if

- $Y = \varprojlim \{Y_n, f_{m,n}\}_{m,n\in\mathbb{N}}$, and $\overline{f}_m : Y \to Y_m$ satisfies $\overline{f}_m = f_{m,n} \circ \overline{f}_n$ for any $m, n \in \mathbb{N}$ such that $m \leq n$.

- μ is a Radon probability measure (called *projective limit measure*) on Y such that

$$\mathcal{L}_{\overline{f}_n}(\mu) = \mu_n \text{ for any } n \in \mathbb{N}.$$

The next theorem, a generalization of Kolmogorov's extension theorem (see [14], pag.53-54), will be used extensively in the proofs of Lemma 1 and Proposition 2.

Theorem 2 *Let $\{Y_n, f_{m,n}, \mu_n\}_{m,n\in\mathbb{N}}$ be a projective sequence of Radon probability measures. Then the measure projective limit $\{Y, \overline{f}_n, \mu\}_{n\in\mathbb{N}}$ exists and μ is a unique Radon probability measure.*

1.6.2 A counterexample to Lemma 1 with the weak* topology

Here we provide an example which shows that the weak* topology can be *strictly* weaker the the narrow topology on $\Delta(X)$, even if X is a Hausdorff topological space.

Let $X = \{(x, y) : y \geq 0,\ x, y \in \mathbb{Q}\}$, and endow it with the irrational slope topology (see [56], Example 75). Thus X is a Hausdorff, non-regular topological space and is also a Lusin space. Every real valued continuous function on X is constant, so the weak* star topology on $\Delta(X)$ is simply the trivial topology $\{\emptyset, \Delta(X)\}$, hence non-Hausdorff. On the other hand, by Lemma 1.(1) $\Delta(X)$ is Hausdorff if it is endowed with the narrow topology. Thus, the weak* topology is strictly weaker than the narrow topology on $\Delta(X)$.

1.6.3 Proof of Lemma 5

For pedagological purposes, it is worth denoting inductively the sequence $\{Z_n\}_{n=0}^{\infty}$ of Hausdorff spaces as follows: $\Omega_1 = Z_0$ and for all $m \geq 1$, $\Omega_{m+1} = \prod_{i=0}^{m} Z_i = \Omega_m \times Z_m$. For any $n \geq m$, denote by $\mathrm{Pr}_{m,n}$ the coordinate projection from Ω_n onto Ω_m, i.e., $\mathrm{Pr}_{m,n}(z_0, ..., z_{m-1}, ..., z_{n-1}) = (z_0, ..., z_{m-1})$. The set D in (1.3.3) equivalently can be expressed as

$$D = \left\{ (v_1, v_2, ...) \in \prod_{i=1}^{\infty} \Delta(\Omega_i) \,\middle|\, v_{i+1} \circ \mathrm{Pr}_{i,i+1}^{-1} = v_i, \; \forall i \geq 1 \right\}.$$

For all $i \geq 1$, denote by $\rho_i : D \to \Delta(\Omega_i)$ the coordinate projections, i.e., $\rho_i(v_1, v_2, ...) = v_i$. It follows from the construction that for every $(v_1, v_2, ...) \in D$, $\left(\{\Omega_n\}, \{v_n\}, \{\mathrm{Pr}_{m,n}\}_{m \leq n}, \right)_{m,n \in \mathbb{N}}$ is a projective sequence of Radon probability measures. According to the definition, the projective limit of $\{\Omega_n\}_{n \in \mathbb{N}}$ subject to the projections $\{\mathrm{Pr}_{m,n}\}_{m \leq n}$ is given by

$$\varprojlim \Omega_n = \left\{ ((z_0), (z_0, z_1), ...) \in \prod_{n \in \mathbb{N}} \Omega_n \,\middle|\, \begin{array}{c} (z_0, ..., z_{n-1}) = \\ \mathrm{Pr}_{n,n+1}((z_0, ..., z_{n-1}), z_n), \; \forall n \in \mathbb{N} \end{array} \right\},$$

thus $\mathrm{Pr}_m : \varprojlim \Omega_n \to \Omega_m$ is the corresponding projection on the coordinate Ω_m. Since each $\mathrm{Pr}_{m,n}$, $m \leq n$, is a coordinate projection, it follows from standard arguments (see [49] pp. 116-117 or [50], p. 416) that $\varprojlim \Omega_n$ is non-empty and it can be identified (homeomorphically) with the Cartesian product $\Omega = \prod_{i=0}^{\infty} Z_i$, so that there is no need to distinguish between these two spaces.

Having done these preparations we proceed now to prove the existence of a homeomorphism $f : D \to \Delta(\Omega)$, where $\Delta(\Omega)$ is endowed with the narrow topology and D inherits the subspace topology as a subset of the product $\prod_{i=0}^{\infty} \Delta(\Omega_i)$.

Let $h : \Delta(\Omega) \to D$ be the map which associates with every $v \in \Delta(\Omega)$ a point $(v_1, v_2, ...)$ in D, namely $v \mapsto [h_n(v)]_{n \in \mathbb{N}} = [\mathcal{L}_{\mathrm{Pr}_n}(v)]_{n \in \mathbb{N}}$. It follows from Theorem 1 that to every array $(v_1, v_2, ...) \in D$ there corresponds a unique Radon probability measure $v \in \Delta(\Omega)$ such that for all $n \in \mathbb{N}$ the marginal of v on Ω_n is v_n. Thus h is bijective. We now show that h is an open map, i.e., if U is open in $\Delta(\Omega)$, then $h(U)$ is open in D. Since h is injective, it suffices to show $h(U)$ is open whenever U is a subbasic open set. Pick $v_0 \in \Delta(\Omega)$. A subbasic neighborhood of v_0 is a set of the form

$$V(v_0; \epsilon, \mathbf{1}_C) = \{ v \in \Delta(\Omega) \, | v(C) > v_0(C) - \epsilon \},$$

where C is an open subset of Ω and $\epsilon > 0$. Every open set $C \subseteq \Omega = \prod_{n=0}^{\infty} Z_n$ has the form $C = \cup_{n=0}^{\infty} \mathrm{Pr}_n^{-1}(G_n)$ where $G_n \subseteq \Omega_n = \prod_{i=0}^{n-1} Z_i$ is open for all n and $\mathrm{Pr}_n^{-1}(G_n)$ increases (see, e.g., [52]). Hence there exists a $l \in \mathbb{N} \cup \{0\}$ such that $v_0 \left(C \backslash \mathrm{Pr}_l^{-1}(G_l)\right) < \frac{\epsilon}{2}$. The subbasic set

$$W\left(v_0; \epsilon, \mathbf{1}_{\mathrm{Pr}_l^{-1}(G_l)}\right) = \left\{v \in \Delta(\Omega) \,|\, v\left(\mathrm{Pr}_l^{-1}(G_l)\right) > v_0\left(\mathrm{Pr}_l^{-1}(G_l)\right) - \epsilon\right\}$$

is contained in $V(v_0; \epsilon, \mathbf{1}_C)$, since for every $v \in W\left(v_0; \epsilon, \mathbf{1}_{\mathrm{Pr}_l^{-1}(G_l)}\right)$

$$\begin{aligned} v_0(C) - v(C) &= \left[v_0(C) - v_0\left(\mathrm{Pr}_l^{-1}(G_l)\right)\right] + \left[v_0\left(\mathrm{Pr}_l^{-1}(G_l)\right) - v\left(\mathrm{Pr}_l^{-1}(G_l)\right)\right] \\ &\quad + \left[v\left(\mathrm{Pr}_l^{-1}(G_l)\right) - v(C)\right] \\ &< \frac{\epsilon}{2} + \epsilon - \frac{\epsilon}{2} = \epsilon. \end{aligned}$$

Consequently it suffices to show $h\left(W\left(v_0; \epsilon, \mathbf{1}_{\mathrm{Pr}_l^{-1}(G_l)}\right)\right)$ is open in D. To this end, note that $h_l(v)(G_l) = \mathcal{L}_{\mathrm{Pr}_l}(v)(G_l)$ and

$$W_l(h_l(v_0); \epsilon, \mathbf{1}_{G_l}) = \left\{\mathcal{L}_{\mathrm{Pr}_l}(v) \in \Delta(\Omega_l) \,|\, \mathcal{L}_{\mathrm{Pr}_l}(v)(G_l) > \mathcal{L}_{\mathrm{Pr}_l}(v_0)(G_l) - \epsilon\right\}$$

is a subbasic open neighborhood of $\mathcal{L}_{\mathrm{Pr}_l}(v_0) \in \Delta(\Omega_l)$. This implies

$$h\left(W\left(v_0; \epsilon, \mathbf{1}_{\mathrm{Pr}_l^{-1}(G_l)}\right)\right) = \rho_n^{-1}\left(W_l(h_l(v_0); \epsilon, \mathbf{1}_{G_l})\right)$$

is a subbasic open neighborhood of $\left(\mathcal{L}_{\mathrm{Pr}_n}(v_0)\right)_{n \in \mathbb{N}} \in D$. Thus h is an open bijection, hence a homeomorphism. To conclude the proof, define $f = h^{-1}$.

1.6.4 Proof of Claim 1

The proof follows the ideas of Rao [49], pp. 177-118.

By definition, $\varprojlim \left\{\Upsilon_n^i, \pi_{n,n+1}^i\right\}_{n \geq 1}$ is a closed subset of $\prod_{n \geq 1} \Upsilon_n^i$ and $\pi_m^i : \varprojlim \left\{\Upsilon_n^i, \pi_{n,n+1}^i\right\}_{n \geq 1} \to \Upsilon_m^i$ is the restriction of the natural projection $\mathrm{Proj}_{\Upsilon_m^i} : \prod_{n \geq 1} \Upsilon_n^i \to \Upsilon_m^i$ to $\varprojlim \left\{\Upsilon_n^i, \pi_{n,n+1}^i\right\}_{n \geq 1}$. Clearly, for any $m, n \in \mathbb{N}$ such that $m \leq n$, the maps π_m^i and π_n^i satisfy the equality $\pi_m^i = \pi_{m,n}^i \circ \pi_n^i$.

Let $g_n^i : \Upsilon^i \to \Upsilon_n^i$ be the restriction to Υ^i of the projection map from $\prod_{l=0}^{\infty} \Delta\left(X_l^i\right)$ to

$\prod_{l=0}^{n-1} \Delta\left(X_l^i\right)$. Then for $m, n \in \mathbb{N}$ satisfying $m \leq n$, we have $g_m^i = \pi_{m,n}^i \circ g_n^i$ since $g_{m,n}^i = \pi_{m,n}^i$. Now we wish to find a function $\widetilde{v} : \Upsilon^i \to \varprojlim \left\{P_n^i, \pi_{n,n+1}^i\right\}_{n\geq 1}$ such that $\pi_n^i \circ \widetilde{v} = g_n^i$.

Since $g_n^i\left(t_1^i, t_2^i, ...\right) = \left(t_1^i, ..., t_n^i\right)$, we need to show that $\left(\pi_n^i \circ v\right)\left(t_1^i, t_2^i, ...\right) = \left(t_1^i, ..., t_n^i\right)$. Let $\widetilde{v} : \Upsilon^i \to \varprojlim \left\{\Upsilon_n^i, \pi_{n,n+1}^i\right\}_{n\geq 1}$ be the map defined as $\widetilde{v}(\cdot) = \left(\pi_1^i(\cdot), \pi_2^i(\cdot), ...\right)$. Since each $\pi_n^i(\cdot)$ is continuous and onto (according to Lemma 6), so is $\widetilde{v}$. It remains to show that $\widetilde{v}$ is injective. Take $t_*^i, t_{**}^i \in \Upsilon^i$ such that $\widetilde{v}\left(t_*^i\right) = \widetilde{v}\left(t_{**}^i\right)$. It is immediate to check that $t_*^i = t_{**}^i$ in that each finite subset of t_*^i is the same as that of t_{**}^i (i.e., all the coordinates of t_*^i are equal to those of t_{**}^i).

To prove the second statement, first remark that $\prod_{n\geq 1} \Theta_n^i = S^{\mathbb{N}} \times \prod_{n\geq 1} \Upsilon_n^i$, where $S^{\mathbb{N}}$ denotes the set of all sequences on S. Denote by $\theta^i = \left(\theta_1^i, \theta_2^i, ...\right)$ each element of $\prod_{n\geq 1} \Theta_n^i$. By definition

$$\varprojlim \left\{\Theta_{n-1}^i, \sigma_{n-1,n}^i\right\}_{n\geq 1} = \left\{ \theta^i \in S^{\mathbb{N}} \times \prod_{n\geq 1} \Upsilon_n^i \;\middle|\; \begin{array}{c} \sigma_{n,n+1}^i\left(\theta_{n+1}^i\right) = \left(Id_{\Theta_1^i}; \pi_{n-1,n}^{-i}\right)\left(\theta_{n+1}^i\right) \\ = \left(s; \left(t_1^{-i}, ..., t_n^{-i}\right)\right) = \theta_n^i \end{array} \right\},$$

therefore $\varprojlim \left\{\Theta_{n-1}^i, \sigma_{n-1,n}^i\right\}_{n\geq 1} = Diag\left(S^{\mathbb{N}}\right) \times \varprojlim \left\{\Upsilon_n^{-i}, \pi_{n,n+1}^i\right\}_{n\geq 1}$, where $Diag\left(S^{\mathbb{N}}\right)$ is the set of all constant sequences on S. Since $Diag\left(S^{\mathbb{N}}\right)$ is homeomorphic to S, the result follows from the homeomorphism between $\varprojlim \left\{\Upsilon_n^{-i}, \pi_{n,n+1}^i\right\}_{n\geq 1}$ and Υ^{-i}.

Chapter 2

Ambiguous types and possibility structures

2.1 Introduction

Type structures, initially introduced by Harsanyi [28] and eventually formalized by Mertens and Zamir [44], are the predominant models to describe interactive uncertainty in strategic environments. The benefit of type structures is that they allow the modeller to describe the agents' hierarchies of beliefs by a *single* probability measure, without sacrificing the generality of game theoretic analysis. The present paper is concerned with the problem of representing interactive beliefs in games where each player is endowed with *multiple priors* over the set of payoff-relevant states, instead of a unique probability measure as in the standard Harsanyi framework.

The relevance of this problem arises naturally in strategic settings where each agent faces *ambiguity* about the unknown features of the environment and\or the other agents' choices. In the context of individual decision making, the standard Bayesian framework is inappropriate to distinguish between risky scenarios, where probabilities of all payoff-relevant events are available to the decision maker, and *ambiguous* or *uncertain* scenarios, where information may be too imprecise to be summarized by a single probability measure. The Ellsberg paradox and related experimental findings provide evidence that agents may be averse to situations where a single probability measure over the states of the world is not objectively known: a decision maker prefers a risky bet (a bet on events with known probabilities) to an ambiguous bet (a bet on events with unknown probabilities), and his preferences cannot be be rationalized by a single probabilistic belief. Motivated by this view, there is now a large literature on individual decision-making under ambiguity, starting with the multiple-prior models of Gilboa and Schmeidler [26] and Bewley [10] and the non-additive probability model of Schmeidler [53]. In the multiple-prior models, agents use a set of probability measures rather than just a single probability measure to determine the expected utility from actions.

If the distinction between risk and ambiguity is meaningful in a single person decision making, it must also be meaningful in interactive decision environments, which involve the delicate issue of hierarchies of beliefs. Consider agents who are uncertain about important parameters of the environment. According to Harsanyi [28] approach, the source of uncertainty can be summarized by an underlying state space, called the *space of nature states* or *parameter space*, each element of which can be thought as a complete description of the players' payoffs, strategy sets and the like. (Such a parameter space would correspond to the set of states of the world in a single person decision making.) The Harsanyi approach to uncertainty in games via type structures provides a parsimonious representation of infinite hierarchies of beliefs within the classical Bayesian framework of individual decision making. Each element of an agent's type structure, called simply *type*, is associated with a single subjective belief over the product of the parameter space and the types of her opponents. The crucial aspect of a type structure is that the probability measure associated with a type encodes an *implicit* description of the agents' hierarchies of beliefs. Indeed, the circularity feature of a type structure makes it possible to specify, if necessary, the mutual beliefs of each agent about nature, about the other agents' beliefs about nature, and so on, in a recursive way. Moreover, the hierarchy described by a type is coherent, in the sense that the marginals of higher-order beliefs coincide with the corresponding lower-order beliefs. The formalism of type structures suffers no loss of generality, in that, as showed by Mertens and Zamir [44], the space of all coherent hierarchies of beliefs arising from a fixed parameter space yields a universal type structure; that is, this structure has the same circularity feature of a type structure, and more importantly, every other type structure can be embedded into it in a unique belief-preserving way.

Nerverthless, the advantage of Harsanyi's analysis of incomplete information games should not make us overlook some potential drawbacks of this approach. First, the notion of Harsanyi type assumes that the uncertainty concerning payoff-relevant issues is reduced to risk, in the sense that individual preferences are based on beliefs that are representable by a single probability measure; Machina and Schmeidler ([38], [39]) provide a thoughtful characterization of this class of preferences and refer to them as *probabilistically sophisticated*. As already remarked, such an assumption cannot capture some aspects of decision-making which can be attributed ambiguity of the problem, e.g., the Ellsberg paradox. Second, and more importantly, this notion of type requires more than just one single level of probabilistic sophistication. Even accepting the postulate that each player has a precise evaluation (a single subjective belief) of his own payoffs, Harsanyi's notion of type requires that each player has a precise evaluation of the other player's evaluation of payoffs, a precise evaluation of the other player's evaluation of his evaluation of payoffs, and so forth. It might be well that an agent is probabilistically sophisticated on payoff-relevant parameters, but unable to summarize her beliefs on the other agents' beliefs to a single probability. In other words, even though the uncertainty concerning payoff-relevant issues were reduced to risk, the same procedure could be demanding for the uncertainty on beliefs about how the other agents may think about each other.

To recap: if ambiguous beliefs are plausible in decision problems for a single agent, we believe they are eminently more plausible in interactive settings with multiple agents. Following this obervation, Ahn [1] builds a universal, set-theoretic type structure which allows for hierarchies of ambiguity. Under the assumption that the parameter space is a compact metric space, he constructs a model of interactive beliefs where each player is allowed to have a compact set of multiple priors on all payoff-relevant events. In turn, each player is also allowed to have multiple beliefs about the (possibly multiple) priors of the other player, and so on. He shows if the players share common certainty of the internal consistency of their levels of ambiguous beliefs, then a type for a player completely specifies her compact streams of beliefs on the other's type. A natural question arises as to whether the assumptions (i) compact parameter space and (ii) compact sets of multiple belief hierarchies, can be dispensed with. Despite its technical content, this question is important for some reasons.

In the context of standard type structures, the existence of a universal structure was first proved by Mertens and Zamir [44] under the assumption that the parameter space is a compact

Hausdorff space and all involved functions are continuous. Other versions of universal type structures are provided by Brandenburger and Dekel [17], Heifetz [29] and Mertens, Sorin and Zamir [43], under a variety of structural assumptions. The common feature of all these constructions is *completeness*, a concept introduced by Brandenburger [16]: in words, a model is complete if any belief about nature states and the hierarchies of the other players corresponds to a hierarchy of beliefs, so that there is no loss of generality in representing player's types by hierarchies. However, Brandenburger [16] shows that, except in degenerate cases, any purely set-theoretic type model of beliefs is necessarily incomplete in that it is always possible to construct a belief that is not held by any type.

A result in a similar vein is obtained by Brandenburger and Keisler [18] in a quite general model theoretic framework. They show that if a first order language is used to describe all the relevant "events" then there cannot exist a complete structure. This negative result suggests that a natural strategy to obtain a complete representation of agents' interactive beliefs in set theoretic models is to impose restrictions on the "richness" of the language used by the agents. In this sense, the choice of a topology on the structure of the model turns out to be crucial to decribe and represent such a language. Put differently, the main question of the current paper is the following: which are the minimal topological restrictions on (sets of) beliefs to get a complete, universal type structure?

In Chapter I we have shown the existence of a universal type structure *a là* Mertens and Zamir under general topological assumptions which covers all the aforementioned results. Specifically, we show that the analysis can be carried out under the following assumptions: (i) the payoff-relevant state space is Hausdorff, and (ii) higher order beliefs are Radon probability measures. In the current paper, we adopt a slight relaxation of assumption (ii), by assuming that higher order beliefs are represented by a compact set of Radon probability measures. We then show that there exists a complete, universal structure which allows for hierarchies of ambiguous beliefs, thus including the results in [1] and Chapter I as special cases.

Our model relates to the existence of a universal possibility structure. A possibility structure is a set-theoretic model of interactive beliefs in which a player knows either that a particular statement is true or false, or knows that it may be true or false (hence, possible) but assigns no probability to it. In a possibility structure, the beliefs of a player are represented by a possibility set consisting of all states regarded possible. Mariotti, Meier and Piccione [42] (MMP, herafter) provide a Brandenburger-Dekel type foundations for possibility structures that identifies an agent's type with a hierarchy of possibility sets. We show (Section 2.5) that a universal possibility structure can be identified as a special case of our set-up. The existence of this possibility structure does not require the assumption of compactness for the parameter space as in MMP, hence our framework provides the most general (that is, weakest) restrictions on the agents' language that must be imposed to obtain a universal, complete possibility structure. We discuss this aspect in detail in Section 2.5.1, and we show how this generalization can be achieved in Section 2.5.2.

We have already mentioned and discussed the literature more immediately related to the present paper. Other relevant contributions on the foundations of games with ambiguity include [20],[23] and [33]. A discussion of some of these papers is deferred to the concluding section.

This paper is organized as follows. Section 2.2 contains some preliminary definitions and mathematical results which are necessary to present the construction of the universal type structure. Section 2.3 presents the construction of the space of ambiguous belief hierarchies. Section 2.4 discusses the implicit representation of type structures and relates the latter to the space constructed in Section 2.3. Possibility structures are the subject of Section 2.5. Finally, Section 2.6 concludes. Omitted proofs are collected in the appendix, which contains further matematical definitions and results needed for Section 2.5.

2.2 Mathematical preliminaries and notation

2.2.1 Topologies for the space of Radon probability measures

For any non-empty topological space X, let $\mathcal{B}(X)$ denote its Borel σ-field and $\mathcal{P}(X)$ the set of all Borel σ-additive probability measures on $\mathcal{B}(X)$. A *Radon probability measure* on X is a measure $\mu \in \mathcal{P}(X)$, such that for every $A \in \mathcal{B}(X)$ and every $\epsilon > 0$, there exists a compact set $K \subseteq A$ such that $\mu(A \backslash K) < \epsilon$. Denote by $\Delta(X)$ the space of Radon probability measures on X. If D is a non-empty, Borel subset of X, define $\Delta(D) = \{\mu \in \Delta(X) \,|\, \mu(D) = 1\}$.

The set $\Delta(X)$ is endowed with the *narrow topology*, which is defined as the coarsest topology on $\Delta(X)$ for which all the maps $\mu \to \int_X f d\mu$ from $\Delta(X)$ into $\mathbb{R}$ are lower semi-continuous, as

f varies in the set of all bounded, lower semi-continuous functions on X. As a subbasic system of neighborhoods, this topology assigns to each $\mu_0 \in \Delta(X)$ the sets of the form

$$V(\mu_0; \epsilon, f) = \left\{ \mu \in \Delta(X) \mid \int_X f d\mu \geq \int_X f d\mu_0 - \epsilon,\ f \text{ is lower semi-continuous, } \epsilon > 0 \right\}.$$

It is known (cf. [57], Theorem 8.1 and Lemma 2 in Chapter I) that if X is a *completely regular space* (e.g., compact Hausdorff or Polish), the narrow topology on $\Delta(X)$ coincides with the usual *weak* topology* on $\Delta(X)$.[1] For an overview of the Radon measures and the narrow topology, see Topsoe [57] and Schwartz [54]. The following lemma, whose proof can be found in Chapter I, collects some properties of the narrow topology we shall repeatedly make use of.

Lemma 7 *Let X be a non-empty topological space, and D a non-empty subset of X. The following statements hold true.*

1. *If X is Hausdorff, so is $\Delta(X)$.*
2. *$\Delta(X)$ is compact if and only if X is compact.*
3. *If X is Polish, then $\Delta(X) = \mathcal{P}(X)$. Furthermore, $\Delta(X)$ is also Polish.*
4. *If X is Hausdorff and D is closed, then $\Delta(D)$ is closed.*

Let $\delta : X \to \Delta(X)$ be the function which maps each $x \in X$ to its Dirac measure $\delta(x)$. Thus $\delta(X) \subseteq \Delta(X)$ denotes the space of all Dirac probability measures on X. It is easy to check that δ is a homeomorphism on its image.

2.2.2 Image measures and their properties

Let X and Y be arbitrary measure spaces, and for any measurable $f : X \to Y$, let $\mathcal{L}_f : \Delta(X) \to \mathcal{P}(Y)$ denote the image measure on Y induced by f, defined by $\mathcal{L}_f(\mu)[E] = \mu\left(f^{-1}(E)\right)$ for any $\mu \in \Delta(X)$ and any Borel set $E \subseteq Y$. An easy check shows that $\mathcal{L}_f$ is well defined, i.e., $\mathcal{L}_f(\mu) \in \mathcal{P}(Y)$ for any $\mu \in \Delta(X)$. However, we have the following result:

[1] For any non-empty, topological space X, the *weak* topology* on $\Delta(X)$ is the coarsest topology for which the map $\mu \to \int_X \phi d\mu$ from $\Delta(X)$ into $\mathbb{R}$ is continuous, as ϕ varies in the set of all bounded, continuous functions on X.

Lemma 8 *Let X, Y be Hausdorff topological spaces and $f : X \to Y$ be continuous. Then $\mathcal{L}_f$ is a continuous function from $\Delta(X)$ into $\Delta(Y)$. Furthermore,*

(i) if f is injective, so is $\mathcal{L}_f$;

(ii) if f is surjective and open, then $\mathcal{L}_f$ is surjective;

(iii) if f is surjective, then $\mathcal{L}_f \circ \delta = \delta \circ f$;

(iv) if Z is a Hausdorff space and $g : Y \to Z$ a continuous map, then $\mathcal{L}_{g \circ f} = \mathcal{L}_g \circ \mathcal{L}_f$.

Proof. We need to prove only point (iii), since (iv) is trivial and the remaining part is proved in Chapter I. Let A be a Borel subset of Y. By continuity of f, the set $f^{-1}(A)$ is measurable in X. Now

$$\begin{aligned} \delta(y)[A] &= \delta(y)\left[f\left(f^{-1}(A)\right)\right] \\ &= (\delta \circ f)(x)\left[f^{-1}(A)\right] \\ &= \begin{cases} 1 & x \in f^{-1}(A) \\ 0 & x \notin f^{-1}(A) \end{cases}, \end{aligned}$$

where the first equality follows from the surjectivity of f. On the other hand

$$\begin{aligned} \mathcal{L}_f(\delta(x))[A] &= (\delta)(x)\left[f^{-1}(A)\right] \\ &= \begin{cases} 1 & x \in f^{-1}(A) \\ 0 & x \notin f^{-1}(A) \end{cases}, \end{aligned}$$

which proves the claim. ■

Unless otherwise stated, for a given countable product space $\prod_{n\in\mathbb{N}} X_n$ where $\mathbb{N}$ is the set of natural numbers, we write Proj_{X_m} as the canonical projection from $\prod_{n\in\mathbb{N}} X_n$ to X_m, for each $m \in \mathbb{N}$. For any $l, m \in \mathbb{N}$ satisfying $l \leq m$, we write $\mathrm{Pr}_{l,m}$ as the coordinate projection from $\prod_{j=1}^{m} X_j$ into $\prod_{j=1}^{l} X_j$. We consider any product, finite or countable, of topological spaces as a topological space with the product topology. If X and Y are arbitrary Hausforff topological spaces, then the marginal measure of $\mu \in \Delta(X \times Y)$ on X is defined as $\mathrm{marg}_X \mu = \mathcal{L}_{\mathrm{Proj}_X}(\mu)$, which is, by the Lemma 8, a continuous and surjective function from $\Delta(X \times Y)$ onto $\Delta(X)$.

2.2.3 Hyperspaces of sets

For any topological space X, $\mathcal{K}X$ refers to the space of compact, non-empty subsets of X. The space $\mathcal{K}X$ is endowed with the *Vietoris topology* τ_V, which is defined as follows. Given $E \subseteq X$, we define these two subsets of $\mathcal{K}X$:

$$E^- = \{A \in \mathcal{K}X | A \cap E \neq \varnothing\},$$

$$E^+ = \{A \in \mathcal{K}X | A \subseteq E\}.$$

The Vietoris topology τ_V on $\mathcal{K}X$ is the topology generated by subbasic sets of form V^- and W^+, where V and W are both open subsets of X. An equivalent formulation of the Vietoris topology on $\mathcal{K}X$ is as follows (see [7], Lemma 1.2). Let $(U_i)_{i=1,\ldots,n}$ be a finite collection of open sets of X. The Vietoris topology τ_V on $\mathcal{K}(X)$ is generated by *basic* neighborhoods of the following form:

$$\langle U_1, ..., U_n \rangle = \{A \in \mathcal{K}X | A \subseteq \cup_{i=1}^n U_i \text{ and } A \cap U_i \neq \varnothing, \, i = 1, ..., n\}.$$

That is, the collections of the form $\langle U_1, ..., U_n \rangle$, with $U_1, ..., U_n$ open in X, form a basis for the Vietoris topology τ_V on $\mathcal{K}X$.

Using just the definitions, it is easy to see that a net $\{A_\alpha\}$ on $\mathcal{K}X$ converges to $A \in \mathcal{K}X$ with respect to τ_V, written $A_\alpha \overset{\tau_V}{\to} A$, if and only the following are satisfied:

(a) (lower Vietoris convergence V^-) $A_\alpha \overset{V^-}{\to} A$ if for every open $U \subseteq X$, $U \cap A \neq \varnothing$ implies that $U \cap A_\alpha \neq \varnothing$ eventually.

(b) (upper Vietoris convergence V^+) $A_\alpha \overset{V^+}{\to} A$ if for every open $U \subseteq X$, $A \subseteq U$ implies that $A_\alpha \subseteq U$ eventually.

For future reference, the following lemma collects some useful properties of the Vietoris topology on $\mathcal{K}X$ (for a proof, see [45] or [9]).

Lemma 9 *Let X be a non-empty topological space. The following statements hold true.*

1. *X is Hausdorff if and only if $\mathcal{K}X$ is Hausdorff.*

2. *X is (completely) regular if and only if $\mathcal{K}X$ is (completely) regular.*

3. *X is (locally) compact if and only if $\mathcal{K}X$ is (locally) compact.*

4. *X is metrizable if and only if $\mathcal{K}X$ is metrizable.*

It is worth noting that if X is a compact metric space, the Vietoris topology on $\mathcal{K}X$ coincides with the topology induced by the Hausdorff distance.

Since in the following we require that all topological spaces be Hausdorff, for each X there is a natural injection $e : X \to \mathcal{K}X$ taking $x \in X$ to its singleton $\{x\}$. Clearly e is a topological embedding, and in the metric case an isometry. Observe that the Vietoris convergence on $\mathcal{K}X$ is compatible with the Hausdorff topology on X, that is, $x_\alpha \to x$ if and only if $\{x_\alpha\} \overset{\tau_V}{\to} \{x\}$ (i.e., $e(x_\alpha) \to e(x)$).

For a continuous function $f : X \to Y$, define its extension $f^{\mathcal{K}} : \mathcal{K}X \to \mathcal{K}Y$ by $f^{\mathcal{K}}(A) = f(A) = \{f(x)|x \in A\}$.

Lemma 10 *Suppose that X and Y are Hausdorff topological spaces and $f : X \to Y$ is continuous. Let Proj_X denote the natural projection from the product space $X \times Y$ onto X. Then the following statements hold true:*

1. *the induced map $f^{\mathcal{K}} : \mathcal{K}X \to \mathcal{K}Y$ is continuous;*

2. *if f is surjective, then $f^{\mathcal{K}}$ is surjective;*

3. *if f is injective, then $f^{\mathcal{K}}$ is injective;*

4. *the induced map $\mathrm{Proj}_X^{\mathcal{K}} : \mathcal{K}(X \times Y) \to \mathcal{K}X$ is continuous, surjective and open.*

5. *if Z is a Hausdorff space and $g : Y \to Z$ a continuous map, then $(f \circ g)^{\mathcal{K}} = f^{\mathcal{K}} \circ g^{\mathcal{K}}$.*

The proof of the Lemma 10 is available in Appendix 2, while Section 2.8.2 contains some remarks showing how Lemma 10.(4) (which turns out to be the main mathematical contribution of this paper) provides a general result which covers analogous cases.

2.3 Hierarchic construction of ambiguous beliefs

Fix a two-player set I;[2] given a player $i \in I$, we denote by $-i$ the other player in I. Let S be a non-empty space describing payoff-relevant states that each player is uncertain about. The space S is called *parameter space*. The following assumption will be mantained throughout the paper.

Assumption 1 S is a Hausdorff topological space.

Since each player $i \in I$ faces ambiguity or risk about the realization of the payoff-relevant state, she is endowed with a set of priors (probability distributions) on the parameter space S; such multiple priors are called first-order (ambiguous) beliefs. However, first-order beliefs do not exaust all the uncertainty faced by each player: player i realises that player $-i$ has at least one first-order belief on S as well, and this set of priors is unknown to her. Thus, player i's second-order beliefs are represented by a set of probability distributions over S and the space of $-i$'s first-order beliefs. Continuing in this fashion, each player is completely characterized by an infinite hierarchy of (ambiguous) beliefs.

We impose the following restriction on belief hierarchies of the players.

Assumption 2 Players' hierarchical beliefs are described by *compact* sets of Radon probability measures.

Formally, for each $i \in I$ define inductively the sequence of spaces $\left\{X_n^i\right\}_{n=0}^{\infty}$ by

$$X_0^i = S, \tag{2.3.1}$$

$$X_{n+1}^i = X_n^{-i} \times \mathcal{K}\Delta\left(X_n^{-i}\right);\, n \geq 0. \tag{2.3.2}$$

[2] The analysis can be trivially extended to more than two players.

An element $h_{n+1}^i = \left(A_1^i, A_2^i, ..., A_{n+1}^i\right) \in \mathcal{K}\Delta\left(X_n^i\right)$ is a $(n+1)$-order ambiguous belief; one can easily show that, according to our notation,

$$X_{n+1}^i = X_0^i \times \prod_{l=0}^{n} \mathcal{K}\Delta\left(X_l^{-i}\right).$$

The set of all possible, infinite hierarchies of ambiguous beliefs for player i is $H_0^i = \prod_{n=0}^{\infty} \mathcal{K}\Delta\left(X_n^i\right)$. A hierarchy of beliefs $\overline{h}^i = \left(\overline{A}_1^i, \overline{A}_2^i, ...\right)$ is *unambiguous* or *probabilistically sophisticated* if $\overline{A}_n^i$ is a singleton for every n. The space of unambiguous belief hierarchies is denoted $\overline{H}_0^i$. The space H_0^i is endowed with the product topology, thus, according to Lemma 7 and Lemma 9, is a Hausdorff space. An analogous conclusion holds if the parameter space S is assumed to be Polish or compact.

Recall that for any $\mu_{n+1}^i \in \Delta\left(X_n^i\right) = \Delta\left(X_{n-1}^{-i} \times \mathcal{K}\Delta\left(X_{n-1}^{-i}\right)\right)$, the mapping $\operatorname{marg}_{X_{n-1}^{-i}} : \Delta\left(X_n^i\right) \to \Delta\left(X_{n-1}^{-i}\right)$ is defined as $\operatorname{marg}_{X_{n-1}^{-i}} \mu_{n+1}^i = \mathcal{L}_{\operatorname{Proj}_{X_{n-1}^{-i}}}\left(\mu_{n+1}^i\right)$. Thus, its extension $\operatorname{marg}_{X_{n-1}^{-i}}^{\mathcal{K}} : \mathcal{K}\Delta\left(X_n^i\right) \to \mathcal{K}\Delta\left(X_{n-1}^{-i}\right)$ is defined by

$$\operatorname{marg}_{X_{n-1}^{-i}}^{\mathcal{K}}(A_{n+1}^i) = \left\{\operatorname{marg}_{X_{n-1}^{-i}} \mu_{n+1}^i | \mu_{n+1}^i \in A_{n+1}^i,\ A_{n+1}^i \in \mathcal{K}\Delta\left(X_n^i\right)\right\}.$$

Since $\operatorname{Proj}_{X_{n-1}^{-i}}$ is continuous and open (by definition of product topology), it follows from Lemma 8 and Lemma 10 that $\operatorname{marg}_{X_{n-1}^{-i}}^{\mathcal{K}}$ is surjective and continuous, for all $n \geq 1$.

Definition 2 *A hierarchy of beliefs $h^i = \left(A_1^i, A_2^i, ...\right) \in H_0^i$ is coherent if and only if, for any $n \geq 1$,*

$$\operatorname{marg}_{X_{n-1}^{-i}}^{\mathcal{K}}(A_{n+1}^i) = A_n^i.$$

This definition of coherence is a simple generalization of the notion of coherence as in [44] or [17]; both notions coincide if each A_n^i is a singleton. As pointed out by Ahn [1], this definition of coherence encodes two forms of consistency between different levels of hierarchical beliefs. The first form of consistency - called "forward consistency" - is implied by the set containement $\operatorname{marg}_{X_{n-1}^{-i}}^{\mathcal{K}}(A_{n+1}^i) \subseteq A_n^i$, which states that player i's higher order beliefs can be supported only by her lower level beliefs. The other form of consistency - called "backward consistency" - is implied by the set containement $\operatorname{marg}_{X_{n-1}^{-i}}^{\mathcal{K}}(A_{n+1}^i) \supseteq A_n^i$, which states that only elements in A_n^i can be at service of some higher order beliefs.

Note that since there is a common uncertainty space S, the sets H_0^i and $\overline{H}_0^i$ are the copies of the same sets H_0 and $\overline{H}_0$, respectively. Therefore, from now on, we drop the superscript i (or $-i$) for notational simplicity where no confusion results.

The space of all coherent hierarchies of beliefs is denoted by H_1. The space of unambiguos, coherent hierarchies of beliefs is $\overline{H}_1 = H_1 \cap \overline{H}_0 = \{(A_1, A_2, ...) \in H_1 : |A_n| = 1, \forall n \geq 1\}$.

Lemma 11 *H_1 and $\overline{H_1}$ are closed subsets of H_0. Moreover, H_1 is homeomorphic to $\mathcal{K}\overline{H}_1$.*

Proof. Consider the set of ambiguous hierarchies for which the $(n+1)$-order beliefs are coherent, namely

$$H_{1,n} = \left\{ (A_1, A_2, ...) \in H_0 \left| \text{marg}_{X_{n-1}}^{\mathcal{K}}(A_{n+1}) = A_n \right. \right\}.$$

Since $H_1 = \cap_{n=1}^{\infty} H_{1,n}$, it is enough to show that $H_{1,n}$ is closed for any $n \geq 1$. Let $\{h^k\} = (A_1^k, A_2^k, ...)$ be a net in $H_{1,n}$ which converges to $h^* = (A_1^*, A_2^*, ...)$. Since $H_{1,n}$ inherits the product topology on H_0, this implies that $h^k \to h^*$ if and only if $A_{n+1}^k \xrightarrow{\tau_V} A_{n+1}^*$, for all $n \geq 0$. We claim that $h^* \in T_{1,n}$, i.e., $\text{marg}_{X_{n-1}}^{\mathcal{K}}(A_{n+1}^*) = A_n^*$. To this end, observe that the functions $\text{marg}_{X_{n-1}}^{\mathcal{K}} : \mathcal{K}\Delta(X_n) \to \mathcal{K}\Delta(X_{n-1})$ and $\text{marg}_{\mathcal{K}\Delta(X_{n-1})}^{\mathcal{K}} : \mathcal{K}\Delta(X_n) \to \mathcal{K}\Delta(\mathcal{K}\Delta(X_{n-1}))$ are continuous in the product topology according to Lemma 8 and Lemma 10. This implies $\text{marg}_{X_{n-1}}^{\mathcal{K}}(A_{n+1}^k) \to \text{marg}_{X_{n-1}}^{\mathcal{K}}(A_{n+1}^*)$, for all $n \geq 0$. By construction, $\text{marg}_{X_{n-1}}^{\mathcal{K}}(A_{n+1}^k) = A_n^k$ for every k. Moreover, $A_n^k \xrightarrow{\tau_V} A_n^*$ by definition of product topology. Hence, $\text{marg}_{X_{n-1}}^{\mathcal{K}}(A_{n+1}^*) = A_n^*$.

To show that $\overline{H}_1$ is closed, recall that $\overline{H}_1 = H_1 \cap \overline{H}_0$, thus it suffices to show that $\overline{H}_0$ is closed. Consider the set of hierarchies whose the n-order beliefs are unambiguous, namely

$$\overline{H}_{0,n} = \{(A_1, A_2, ...) \in H_0 : |A_n| = 1\}.$$

Since $\overline{H}_0 = \cap_{n=1}^{\infty} \overline{H}_{0,n}$, it only remains to prove that $\overline{H}_{0,n}$ is closed for all $n \geq 1$. To this end, let $\{h^k\} = (A_1^k, A_2^k, ...)$ be a net in $\overline{H}_{0,n}$ which converges to $h^* = (A_1^*, A_2^*, ...)$. This means that the net $\{A_n^k\}$ converges, in the Vietoris topology, to A_n^*. To conclude the proof, we have to show that A_n^* is a singleton. Suppose, to the contrary, that $|A_n^*| > 1$. The net $\{A_n^k\}$ can be identified, by the topological embedding $e : \Delta(X_{n-1}) \to \mathcal{K}\Delta(X_{n-1})$, as a net in $\Delta(X_{n-1})$. Thus $\{e^{-1}(A_n^k)\}$ converges to $e^{-1}(A_n^*)$, i.e., the net $\{e^{-1}(A_n^k)\}$ has more that one limit. This contradicts the Hausdorff property of the space $\Delta(X_{n-1})$. Therefore, A_n^* must be a singleton.

The existence of a homeomorphism between H_1 and $\mathcal{K}\overline{H}_1$ was already proved in [1], Proposition 1, by using a compactness argument (the finite intersection property of a appropriately controlled family of closed subsets). Here we provide a more general proof. The space $\overline{H}_1$ can be naturally regarded as a subset of $\prod_{n=0}^{\infty}\Delta(X_n)$ under the identification $(\mu_n)_{n\geq 1}\rightarrow\left\{(\mu_n)_{n\geq 1}\right\}$. We slightly abuse notation by denoting $\mathrm{Proj}_{\Delta(X_{n-1})}$ as the continuous projection from $\overline{H}_1$ onto $\Delta(X_{n-1})$, and $\mathrm{Proj}^{\mathcal{K}}_{\Delta(X_{n-1})}$ as the corresponding extension. Let $G=(G_n)_{n\geq 1}:\mathcal{K}\overline{H}_1\rightarrow H_1$ be the map defined by

$$(G_n(K))_{n\geq 1}=\left(\mathrm{Proj}^{\mathcal{K}}_{\Delta(X_{n-1})}(K)\right)_{n\geq 1}$$

where the n-th component satisfies

$$\begin{aligned}G_n(K) &= \left\{\mathrm{Proj}_{\Delta(X_{n-1})}\left(\overline{h}\right)\left|\overline{h}\in K\right.\right\}\\ &= \left\{\mathrm{marg}_{X_{n-1}}\mathrm{Proj}_{\Delta(X_n)}\left(\overline{h}\right)\left|\overline{h}\in K\right.\right\}\\ &= \mathrm{marg}^{\mathcal{K}}_{X_{n-1}}G_{n+1}(K)\end{aligned}$$

by coherence of beliefs. The map G is an open surjection: since $\mathrm{Proj}_{\Delta(X_{n-1})}$ is open and onto (by definition of product topology), so is $\mathrm{Proj}^{\mathcal{K}}_{\Delta(X_{n-1})}$ (Lemma 10, claims (2) and (4)). We now show that G is injective. Let K and $K^{'}$ be two compact subsets of $\overline{H}_1$ such that $G(K)=G\left(K^{'}\right)$. Fix an arbitrary $\overline{h}=(\mu_n)_{n\geq 1}\in K$. Thus, for all $n\geq 1$, $G_n\left(\left\{\overline{h}\right\}\right)=\mathrm{Proj}_{\Delta(X_{n-1})}\left(\overline{h}\right)=\mu_n$, and since by assumption $G(K)=G\left(K^{'}\right)$, we have that $G_n(K)=G_n\left(K^{'}\right)$, for all $n\geq 1$, which implies $\mu_n\in G_n\left(K^{'}\right)$. The set $G_n^{-1}(\mu_n)=\left\{\left(\mu_1^{'},\mu_2^{'},...\right)\in\overline{H}_1\left|\mu_n^{'}=\mu_n\right.\right\}$ is closed (by continuity of G_n) and the set $K_n=K^{'}\cap G_n^{-1}(\mu_n)$ is non-empty (and also compact). Clearly $\cap_{n\geq 1}G_n^{-1}(\mu_n)=\left\{\overline{h}\right\}$, and if $\mu_n^{'}=\mu_n$, then by coherence $\mu_m^{'}=\mu_m$ for all $m\leq n$. So $\left(G_n^{-1}(\mu_n)\right)_{n\geq 1}$ is a decreasing sequence of closed subsets of $\overline{H}_1$, and $(K_n)_{n\geq 1}$ is a family of nested, closed subsets of $K^{'}$with non-empty intersection: $\cap_{n\geq 1}K_n=K^{'}\cap\left[\cap_{n\geq 1}G_n^{-1}(\mu_n)\right]=\left\{\overline{h}\right\}$. Then $(\mu_n)_{n\geq 1}\in K^{'}$. This proves $K\subseteq K^{'}$. The proof of the reverse implication, $K\supseteq K^{'}$, is analogous. So $K=K^{'}$, which proves that G is injective. In summary, G is an open bijection, hence a homeomorphism. ■

The next lemma, a general version of Lemma 1 in [17], provides the main technical tool for proving the existence of a canonical homeomorphism. It is a consequence of Theorem 4 in

Appendix 1, and the reader is referred to Lemma 5 in Chapter I for a proof.

Lemma 12 *Let* $\{Z_n\}_{n=0}^{\infty}$ *be a collection of Hausdorff spaces, and let*

$$D = \left\{ (\mu_1, \mu_2, ...) \left| \begin{array}{c} \mu_n \in \Delta\left(Z_0 \times ... \times Z_{n-1}\right),\ \forall n \geq 1,\ \text{and} \\ \text{marg}_{Z_0 \times ... \times Z_{n-2}} \mu_n = \mu_{n-1},\ \forall n \geq 2 \end{array} \right. \right\}. \tag{2.3.3}$$

Then there exists a homeomorphism $f : D \to \Delta\left(\prod_{n=0}^{\infty} Z_n\right)$ *such that*

$$\forall n \geq 1,\ \text{marg}_{Z_0 \times ... \times Z_{n-2}} \left[f\left((\mu_1, \mu_2, ...)\right)\right] = \mu_{n-1}.$$

The next proposition states that a coherent hierarchy of beliefs for a player is equivalent to a compact set of beliefs over the parameter space and the (not necessarily coherent) hierarchy of beliefs of her opponent.

Proposition 3 *There exists a homeomorphism* $f : H_1 \to \mathcal{K}\Delta\left(S \times H_0\right)$ *such that*

$$\forall n \geq 0,\ \text{marg}_{X_n}^{\mathcal{K}} \left[f\left((A_1, A_2, ...)\right)\right] = \text{Proj}_{\mathcal{K}\Delta(X_n)} \left((A_1, A_2, ...)\right).$$

Proof. Set $Z_0 = X_0$ and, for any $n \geq 1$, $Z_n = \mathcal{K}\Delta\left(X_{n-1}\right)$. By Lemma 7 and Lemma 9, each Z_n is Hausdorff, and so is $\prod_{n=0}^{\infty} Z_n$; hence the narrow topology on $\Delta\left(\prod_{n=0}^{\infty} Z_n\right)$ is also Hausdorff. By construction, $Z_0 \times ... \times Z_n = X_n$ and $\prod_{n=0}^{\infty} Z_n = S \times H_0$, while the set of unambiguous coherent hierarchies of beliefs $\overline{H}_1$ correponds to the set D in (2.3.3). Thus, by Lemma 12, there is a homeomorphism $\overline{f} : \overline{H}_1 \to \Delta\left(S \times H_0\right)$. Apply Lemma 10 to $\overline{f}^{\mathcal{K}} : \mathcal{K}\overline{H}_1 \to \mathcal{K}\Delta\left(S \times H_0\right)$, so that $\overline{f}^{\mathcal{K}}$ is a homeomorphism. To conclude the proof, observe that, according to Lemma 11, H_1 is homeomorphic to $\mathcal{K}\overline{H}_1$ through the map $G^{-1} : H_1 \to \mathcal{K}\overline{H}_1$. The composition $f = \overline{f}^{\mathcal{K}} \circ G^{-1}$ is the required homeomorphism. ■

The homeomorphism just described implies that player i's coherent hierarchy determines his set of beliefs over player $-i$'s hierarchies of beliefs. However, even if player i's hierarchy of ambiguos beliefs $h \in H_1$ is coherent (in the sense of Definition 1), some elements of $f(h)$ may assign positive probability to sets of incoherent ambiguous hierarchies of the other player, that is, player i may believe it is possible that player $-i$'s ambiguous hierarchy is not coherent. This precludes the possibility for player i's coherent hierarchy to determine i's beliefs over $-i$'s beliefs over i's hierarchy of beliefs, and hence, inductively, all possible (ambiguous) beliefs. Thus, following [17], we close the model by imposing *common certainty of coherence*.[3]

Formally, we say that player i, endowed with an ambiguous hierarchy h, is certain of (a measurable) event $E \subseteq S \times H_0$ if $f(h) \subseteq \Delta(E)$, that is, all his possible unambiguous beliefs belonging to the set $f(h)$ put probability one on E. Common certainty of coherence is imposed by defining inductively the sets:

$$\begin{aligned} H_{k+1} &= \{h \in H_1 \,|\, f(h) \subseteq \Delta(S \times H_k)\}, \, k \geq 1, \\ H_\infty &= \cap_{k \geq 1} H_k. \end{aligned}$$

$H_\infty \times H_\infty$ is the set of pairs of types satisfying common certainty of coherence, in the sense that each player believes that the other player's belief hierarchy is coherent, believes that the other player believes that her belief hierarchy is coherent, and so on. In the same fashion, we define *common certainty of coherence and probabilistic sophistication* by defining inductively the sets:

$$\begin{aligned} \overline{H}^c_{k+1} &= \left\{h \in \overline{H}_1 \,\middle|\, f(h) \in \Delta\left(S \times \overline{H}^c_k\right)\right\}, \, k \geq 1, \\ \overline{H}^c_\infty &= \cap_{k \geq 1} \overline{H}^c_k. \end{aligned}$$

(Observe that $f(h)$ is a singleton, i.e. $f(h) \in \Delta(S \times H_0)$, if and only if h is probabilistically sophisticated. So the sets $\overline{H}^c_{k+1}$ and $\overline{H}^c_\infty$ are well-defined). Similarly, $\overline{H}^c_\infty \times \overline{H}^c_\infty$ is interpreted as the set of pairs of players' hierarchies of beliefs such that player i's hierarchy: is probabilistically sofisticated; is certain that player $-i$'s hierarchy is probabilistically sofisticated; is certain that player $-i$'s hierarchy is certain player i's hierarchy is probabilistically sophisticated; and so on. $\overline{H}^c_\infty$ is naturally regarded as a subset of H_∞ under the identification $\overline{h} \mapsto \{\overline{h}\}$. The relation between hierarchies of beliefs satisfying common certainty of coherence and common certainty of probabilistic sofistication is stated in the next lemma, which is an immediate consequence of Lemma 11.

[3] Brandenburger and Dekel [17] use the term "common knowledge of coherence". In general, the term "knowledge" refers to justified true belief that comes from logical deduction. It is therefore conceptually different from probability one belief ("certainty").

Lemma 13 *$\overline{H}_\infty^c$ and $\overline{H}_\infty$ are non-empty, closed subsets of H_1 and $\overline{H}_1$, respectively.*

Proof. First remark that, as f is a homeomorphism, $H_k = f^{-1}(\mathcal{K}\Delta(S \times H_{k-1}))$ for each $k \geq 1$, hence by construction $\{H_k\}_{k\geq 1}$ is a family of non-empty and nested sets. We now show by induction that H_k is closed, for any $k \geq 1$. The statement is true for $k = 1$: H_1 is closed by Lemma 11. Next suppose that the statement holds true for H_k, $k \geq 1$. Thus $\Delta(S \times H_k)$ is closed according to Lemma 7.4, and Lemma 2.2 of [45] implies that $\mathcal{K}\Delta(S \times H_k)$ is also closed. Using the fact that f is a homeomorphism, we conclude that $H_{k+1} = f^{-1}(\mathcal{K}\Delta(S \times H_k))$ is closed. So $H_\infty = \cap_{k\geq 1} H_k$ is a closed subset of H_1. The proof for $\overline{H}_\infty^c$ is similar, and therefore omitted. Moreover, $\overline{H}_\infty^c$ is non-empty: this is shown in Chapter I, Proposition 2. Since $\overline{H}_\infty^c$ can be regarded as a subset of H_∞, it follows that H_∞ is also non-empty. ■

We can now state the main result of this section.

Proposition 4 *The restriction of f to H_∞ induces a homeomorphism $g : H_\infty \to \mathcal{K}\Delta(S \times H_\infty)$.*

Proof. Since the restriction of the homeomorphism f to H_∞ is hereditarily continuous, injective and open, it remains to show that $f(H_\infty) = \mathcal{K}\Delta(S \times H_\infty)$. This can be accomplished by showing that the functions $\overline{f}$ and G^{-1} in the proof of Proposition 3 are onto, when restricted to H_∞ and $\overline{H}_\infty$, respectively. This involves only some minor changes of the proof in [1], so we omit the details. ■

Corollary 1 *The restriction of f to $\overline{H}_\infty^c$ induces a homeomorphism $\overline{g} : \overline{H}_\infty^c \to \Delta\left(S \times \overline{H}_\infty^c\right)$.*

Proof. Simply observe

$$\begin{aligned}
\overline{H}_\infty^c &= \cap_{k\geq 1} f^{-1}\left(\Delta\left(S \times \overline{H}_{k-1}^c\right)\right) \\
&= f^{-1}\left(\cap_{k\geq 1}\Delta\left(S \times \overline{H}_{k-1}^c\right)\right) \\
&= f^{-1}\left(\Delta\left(S \times \cap_{k\geq 1}\overline{H}_{k-1}^c\right)\right) \\
&= f^{-1}\left(\Delta\left(S \times \overline{H}_\infty^c\right)\right),
\end{aligned}$$

where the third equality follows from the fact that a countable intersection of probability one events has probability one (cf. footnote 5 in Chapter I); since f is onto, then $f\left(\overline{H}_\infty^c\right) = \Delta\left(S \times \overline{H}_\infty^c\right)$. Finally, observe $f = \overline{f}$ on $\overline{H}_\infty^c$. ■

2.4 Type structures

In this section, to simplify the notation, we adopt the following conventions. Given a player $i \in I = \{1, 2\}$, we denote by j the other player in I. Define $I_0 = I \cup \{0\}$, where "0" stands for "nature". Hence, for each $i \in I$, we refer to $-i$ as the set $\{0, j\}$.

Definition 3 *An S-based type structure is a structure $\mathcal{T} = (S, T_i, g_i)_{i \in I}$ such that for each $i = 1, 2$, T_i is a Hausdorff space and g_i is a continuous function*

$$g_i : T_i \to \mathcal{K}\Delta\left(S \times T_j\right).$$

Members of T_1, T_2 are called types. *The set $S \times T_1 \times T_2$, whose members are called* states *(*of the world*), is called* belief space. *An S-based type structure $\mathcal{T}$ is* complete *if g_i is onto for each $i = 1, 2$.*

Definition 3 is based on a standard epistemic definition (cf. [30]). Every type structure provides an *implicit* description of belief hierarchies: with each type $t_i \in T_i$ for a player is associated a compact set of Radon measures on the parameter space and types for the other player. The difference from the standard definition is the use of a set of measures rather than one measure.

It is clear from Definition 3 and Proposition 4 that the structure $\mathcal{T}_u = (S, H_\infty, H_\infty, g, g)$ is a complete, symmetric S-based type structure, by setting $T_i = H_\infty$ and $g_i = g$, for each $i \in I$. Some features distinguish $\mathcal{T}_u$ from any other type structure. First, $\mathcal{T}_u$ is selected from explicit hierarchies in accordance with the notion of coherence (Definition 2), as it is shown in Section 3. Second, $\mathcal{T}_u$ is universal, in the sense that if $\mathcal{T} = (S, T_i, g_i)_{i \in I}$ is another S-based type structure, then each T_i can be embedded in H_∞ in a way preserving the implicit description of higher-order (ambiguous) beliefs.

The primary focus of this section is to show that $\mathcal{T}_u$ is, in fact, a *universal type structure*. This will be done in two steps. First, in subsection 2.4.1, we will specify how types induces hierarchies of (ambiguous) beliefs by showing that, for a given S-based type structure $\mathcal{T} = (S, T_i, g_i)_{i \in I}$, it is possible to associate with every $t_i \in T_i$ a point in the space H_0. We shall use a procedure analogous to that in [30] and [6] that allows for hierarchies of ambiguity; our construction scheme parallels their development and many mathematical steps are appropriately adapted.[4] Next, in subsection 2.4.2, we will formulate an appropriate notion of embedding for type spaces, and show that any type space (satisfying certain technical regularity assumption) corresponds to a belief-closed subset of $\mathcal{T}_u$ (precise definition will be given below).

2.4.1 From types to belief hierarchies

For a given S-based type structure $\mathcal{T} = (S, T_i, g_i)_{i \in I}$, we define an *i-description map* $\varphi_i : T_i \to H_0$ associating with each $t_i \in T_i$ a corresponding ambiguous hierarchy, for each $i \in I$. The hierarchy $\varphi_i (t_i) = \left(\varphi_i^1 (t_i), \varphi_i^2 (t_i), ...\right)$ is called the *i-description* of t_i. Each i-decription map is defined inductively:

- (base step: $n = 1$) For each $i \in I$, $t_i \in T_i$, define $\varphi_i^1 = \mathcal{L}^{\mathcal{K}}_{\text{Proj}_S} \circ g_i : T_i \to \mathcal{K}\Delta (S \times T_j) \to \mathcal{K}\Delta (S)$ by

$$\varphi_i^1 (t_i) = \mathcal{L}^{\mathcal{K}}_{\text{Proj}_S} [g_i (t_i)] = \text{marg}^{\mathcal{K}}_S [g_i (t_i)] .$$

  For each $i \in I$, $t_j \in T_j$, $s \in S$, set $\psi^0_{-i} = Id_S$ and define $\psi^1_{-i} : S \times T_j \to X_1 = S \times \mathcal{K}\Delta (S)$ by

$$\psi^1_{-i} (s, t_j) = \left(\psi^0_{-i} (s), \varphi^1_j (t_j)\right) = \left(s, \varphi^1_j (t_j)\right) .$$

- (inductive step: $n + 1$, $n \geq 1$) Suppose we have already defined the functions $\varphi_i^n : T_i \to \mathcal{K}\Delta (X_{n-1})$ and $\psi^n_{-i} : S \times T_j \to X_n = X_{n-1} \times \mathcal{K}\Delta (X_{n-1})$. For each $i \in I$, $t_i \in T_i$, $E_n \in \mathcal{B}(X_n)$, define $\varphi_i^{n+1} = \mathcal{L}^{\mathcal{K}}_{\psi^n_{-i}} \circ g_i : T_i \to \mathcal{K}\Delta (S \times T_j) \to \mathcal{K}\Delta (X_n)$ as

$$\begin{aligned} \varphi_i^{n+1} (t_i) [E_n] &= \mathcal{L}^{\mathcal{K}}_{\psi^n_{-i}} (g_i (t_i)) [E_n] \\ &= g_i (t_i) \left[\left(\psi^n_{-i}\right)^{-1} (E_n)\right] . \end{aligned}$$

[4] Ahn [1] borrows from MMP an indirect procedure to show that his type structure is universal. Such a procedure depends crucially on the compactness assumption for the parameter space. In contrast, the standard procedure adopted in the current paper is more general. It would be important future work to analyze how these two procedures are related, as the MMP indirect procedure provides a way to compare H_∞ to the infinite consumption problems of Gul and Pesendorfer [27], as it is shown in [1].

For each $i \in I$, $t_j \in T_j$, $s \in S$, define $\psi_{-i}^{n+1} : S \times T_j \to X_{n+1} = X_n \times \mathcal{K}\Delta(X_n)$ as

$$\psi_{-i}^{n+1} = \left(\psi_{-i}^{n}, \varphi_j^{n+1}\right),$$

that is, $\psi_{-i}^{n+1}(s, t_j) = \left(\psi_{-i}^{n}(s, t_j), \varphi_j^{n+1}(t_j)\right) = \left(s, \varphi_j^{1}(t_j), ..., \varphi_j^{n}(t_j), \varphi_j^{n+1}(t_j)\right)$.

An easy check (use Lemma 8 and Lemma 10 in the base step, and then proceed by induction) shows that each φ_i is a continuous function.

2.4.2 Type morphisms and universality

Given an S-based type structure $\mathcal{T} = (S, T_i, g_i)_{i\in I}$, we address the question on whether each set T_i can be embedded in H_∞, the set of all conceivable ambiguous hierarchies satisfying coherence and common certainty of coherence. The building block we use to construct an appropriate notion of embedding for type structures is a continuous function $m_{-i} : S \times T_j \to S \times T_j'$ which induces the continuous map $\mathcal{L}_{m_{-i}}^{\mathcal{K}} : \mathcal{K}\Delta(S \times T_j) \to \mathcal{K}\Delta\left(S \times T_j'\right)$.

Definition 4 *Let $\mathcal{T} = (S, T_i, g_i)_{i\in I}$ and $\mathcal{T}' = (S, T_i', g_i')_{i\in I}$ be two S-based type structures. The function $m = (m_0, m_1, m_2) : S \times T_1 \times T_2 \to S \times T_1' \times T_2'$ is called* type morphism *if*

1. *$m_0 = Id_S$,*
2. *for each $i \in I$, $m_i : T_i \to T_i'$ is a continuous function such that $g_i' \circ m_i = \mathcal{L}_{m_{-i}}^{\mathcal{K}} \circ g_i$, where $m_{-i} = (m_0, m_j) : S \times T_j \to S \times T_j'$.*

The morphism is a type isomorphism *if m is a homeomorphism. Say $\mathcal{T}$ can be* embedded *into $\mathcal{T}'$ if there is a type morphism from $\mathcal{T}$ to $\mathcal{T}'$. Say $\mathcal{T}$ and $\mathcal{T}'$ are* isomorphic *if there is a type isomorphism between them.*

Condition (2) in the definition of type morphism expresses consistency between the function $m_i : T_i \to T_i'$ and the induced function $\mathcal{L}_{m_{-i}}^{\mathcal{K}} : \mathcal{K}\Delta(S \times T_j) \to \mathcal{K}\Delta\left(S \times T_j'\right)$. The gain the intuition, let t_i be an arbitrarily fixed type in T_i. The function m_i maps t_i to some $t_i' \in T_i'$, thus $g_i'(t_i')$ defines a compact set of Radon probability measures on $S \times T_j'$. On the other hand, the function g_i maps t_i to some compact set C of $\Delta(S \times T_j)$, hence $\mathcal{L}_{m_{-i}}^{\mathcal{K}}(C)$ defines a compact set of Radon probability measures on $S \times T'_j$. Thus, Condition (2) states that $g'_i(t'_i)$ and $\mathcal{L}^{\mathcal{K}}_{m_{-i}}(C)$ should coincide in that both are generated by the same type $t_i \in T_i$. In words, a type morphism preserves the implicit description of belief hierarchies for the players.

Another important property of type morphisms is that they preserve the explicit description of ambiguous belief hierarchies. Lemma 17 in Appendix 3 shows that if $\mathcal{T}$ and $\mathcal{T}'$ are S-based type structures such that $\mathcal{T}$ can be embedded into $\mathcal{T}'$, then $\varphi_i(T_i) \subseteq \varphi_i(T'_i)$ for every player i; that is, every S-based belief hierarchy that is generated by some type in $\mathcal{T}$ is also generated by some type in $\mathcal{T}'$. This formalizes the idea of viewing type morphisms as a manner to relate types in one structure to types in a wider structure.[5] We characterize this intepretation of type morphism for ambiguous types by restricting type structures to be non-redundant (cf. [44]). A type structure is non-redundant if any two distinct types induce distinct ambiguous hierarchies.[6] Formally:

Definition 5 *An S-based type structure $\mathcal{T} = (S, T_i, g_i)_{i\in I}$ is* non-redundant *if, for each $i \in I$, the i-description map φ_i is injective.*

It is evident from this definition that the structure $\mathcal{T}_u = (S, H_\infty, H_\infty, g, g)$ is non-redundant, as each i-description map turns out to be the identity. The following result gives a characterization of non-redundant type structures in terms of the measurable structure induced by type morphisms.

Lemma 14 *Fix S-based, non-redundant type structures $\mathcal{T} = (S, T_i, g_i)_{i\in I}$ and $\mathcal{T}' = (S, T'_i, g'_i)_{i\in I}$. Let $m = (m_0, m_1, m_2)$ be a type morphism from $\mathcal{T}$ and $\mathcal{T}'$. Then m is injective. Moreover, if m^{-1} is Borel measurable, then the product space $S \times m_1(T_1) \times m_2(T_2)$ forms a* belief-closed subspace *of $S \times T'_1 \times T'_2$, i.e.,*

$$\forall i \in I,\ \forall t'_i \in m_i(T_i) : g'_i(t'_i) \in \mathcal{K}\Delta(S \times m_j(T_j)).$$

[5] An alternative notion of embedding for type structures is that of *hierarchy morphism.* Unlike type morphisms, a hierarchy morphism makes explicit reference to belief hierarchies. A type morphism is a hierarchy morphism, but not vice versa. For more on this, see Friedenberg and Meier ([25]).

[6] Mertens and Zamir ([44], Definition 2.4 and Proposition 2.5) formulate the non-redundancy condition in terms of a separation condition which implies the property here. According to their formulation, a type structure $\mathcal{T} = (S, T_i, g_i)_{i\in I}$ is non-redundant if the σ-field of each T_i separates the points. One can easily show (cf. [36]) that the our definition of non-redundant type is a direct implication of Mertens and Zamir's definition.

The above lemma says that, if $\mathcal{T}$ can be embedded uniquely into $\mathcal{T}'$ by the type morphism m, we can essentially regard $\mathcal{T}$ as a (measurable) substructure of $\mathcal{T}'$. The belief space $S \times m_1(T_1) \times m_2(T_2)$ is associated with the type structure $\mathcal{T}'' = (S, m_i(T_i), \kappa_i)_{i \in I}$, where $\kappa_i : m_i(T_i) \to \mathcal{K}\Delta(S \times m_j(T_j))$ is such that $\kappa_i(t_i')[E] = g_i'(t_i')[E]$ for each $E \in \mathcal{B}(S \times m_j(T_j))$, $i \in I$.[7] We call $\mathcal{T}''$ the substructure of $\mathcal{T}'$ induced by $\mathcal{T}$.

Definition 6 *An S-based type structure $\mathcal{T}' = (S, T_i', g_i')_{i \in I}$ is* universal *if for every other S-based type structure $\mathcal{T} = (S, T_i, g_i)_{i \in I}$ there is a unique type morphism from $\mathcal{T}'$ to $\mathcal{T}$. In this case, the set $S \times T_1' \times T_2'$ is called* universal belief space*.*

Of course, any two universal type structures are isomorphic.[8]

We state now the main result of this section.

Proposition 5 *Let $\mathcal{T} = (S, T_i, g_i)_{i \in I}$ be an arbitrary S-based type structure, and, for each $i \in I$, let $\varphi_i : T_i \to H_0$ be an i-description map. Then, for each $i \in I$,*

1. *$\varphi_i(T_i) \subseteq H_\infty$,*

2. *$\varphi = (Id_S, \varphi_1, \varphi_2)$ is the unique type morphism from $\mathcal{T}$ to $\mathcal{T}_u$.*

Thus $\mathcal{T}_u$ is the unique universal type structure (up to type isomorphism).

We conclude this section with some remarks concerning the implications of Proposition 5. Clearly, any belief-closed subspace of $S \times H_\infty \times H_\infty$ is associated with a non-redundant S-based type structure. The reverse claim, however, does not hold. To see why, let $\mathcal{T} = (S, T_i, g_i)_{i \in I}$ be an arbitrary, non-redundant S-based type structure. Lemma 14 yields that the product space $S \times \varphi_1(T_1) \times \varphi_2(T_2)$ forms a belief-closed subspace of $S \times H_\infty \times H_\infty$ whenever the canonical type morphism φ has a measurable inverse. This last condition is satisfied if both S and each T_i are compact (cf. [44]) or Polish (cf . [24]). In the most general set-up (i.e., Hausdorff topological spaces), a non-redundant type space $\mathcal{T}$ may fail to identify a belief-closed subspace

[7] Note that, since m^{-1} is measurable, $m_j(T_j)$ is a Borel subset of T_j. Thus, every Borel subset of $m_j(T_j)$ is also a Borel subset of T_j. This implies that the map κ_i is well defined.

[8] Within the framework of category theory, S-based type spaces for player set I, as objects, and type morphisms, as morphisms, form a category. The "universal type space" is a terminal object in the category of type spaces.

of $S \times H_\infty \times H_\infty$ in that $\varphi_i(T_i)$ may not be a Borel subset of H_∞. This issue is discussed formally in Appendix 3, and Lemma 19 furnishes a sufficient condition on the topological structures of $\mathcal{T}$ which generalizes the aforementioned results. Moreover, in Appendix 3 we characterize the universal belief space $S \times H_\infty \times H_\infty$ as the union of all its possible belief-closed subspaces. This provides a rationale for constructing $\mathcal{T}_u$ by taking the union of all families of smallest type structures.

2.5 Possibility structures and ambiguity

In this section, we will be highlighting the relationship between type and possibility structures. First, we shall briefly review the basic notions and properties of possibility structures we will be using. Next, we will present the construction of a complete possibility structure as a (strict) substructure of $\mathcal{T}_u$.

2.5.1 Complete possibility structures

Possibility structures are models of interactive beliefs which specify, for each player, a set of states and a set-theoretic belief consisting of all states regarded possible. These structures are related to partitioned representations of knowledge (cf. [4]). Formally, let S^i be the underlying parameter space. Given a set X, let $\mathcal{N}(X)$ denote a family of non-empty subsets of X (in general, the power set). The following definition is borrowed from [16].

Definition 7 *An S-based* possibility structure *is a structure $\mathcal{P} = \left(S, T^i, v_i\right)_{i \in I}$ such that for each $i = 1, 2$, T^i is a non-empty space and v_i is a mapping*

$$v_i : T^i \to \mathcal{N}\left(S \times T^{-i}\right).$$

Members of T^i are called (possibility) types. For each $i = 1, 2$, v_i is a possibility mapping *which assigns to each type $t^i \in T^i$ a set-theoretic belief $v_i\left(t^i\right) \subseteq S \times T^{-i}$ about the basic uncertainty parameter and the type of player $-i$. The subset $v_i\left(t^i\right)$ is called the* possibility set *of type t^i of player i. The structure $\mathcal{P}$ is called* complete *if both v_i and v_{-i} are onto.*

In a complete possibility structure, for every possibility set of player i, there is a type of player $-i$ associated with that set, and vice versa. Brandenburger [16] shows that, generically (i.e., when $\mathcal{N}$ is the power set), a complete S-based possibility structure does not exist. As Brandenburger carefully explains, the existence of a complete structure calls for topological assumptions that attend to rule out certain kinds of possibility sets which restrict the cardinality of $\mathcal{N}$.[9] In light of this non-existence result, Mariotti, Meier and Piccione [42] (MMP, hereafter) consider *compact possibility structures* consisting of: (i) a common parameter space S that is compact Hausdorff; (ii) possibility sets that are non-empty compact (namely, $\mathcal{N} = \mathcal{K}$), (iii) a possibility mapping which is continuous. MMP show that, under these topological restrictions on the parameter space and possibility sets, a universal compact possibility structure exists. Formally, by defining $X_0^{MMP} = S$ and $X_{n+1}^{MMP} = X_n^{MMP} \times \mathcal{K}X_n^{MMP}$, the space of MMP-hierarchies is $\Theta_0^{MMP} = \prod_{n=0}^{\infty} \mathcal{K}X_n^{MMP}$. An MMP-hierarchy $\left(A_1^{MMP}, A_2^{MMP}, ...\right)$ is MMP-coherent if, for each $n \geq 1$,

$$A_n^{MMP} = \mathrm{Proj}_{X_{n-1}^{MMP}} A_{n+1}^{MMP}.$$

MMP provide a Brandenburger-Dekel type foundations for compact continuous possibility structures by showing the existence of a non-empty, compact subset of MMP-hierarchies, Θ_∞^{MMP}, which is homeomorphic to $\mathcal{K}\left(S \times \Theta_\infty^{MMP}\right)$.

Observe that, since $\delta : X \to \delta(X)$ is a homeomorphism, the induced map $\delta^{\mathcal{K}} : \mathcal{K}X \to \mathcal{K}\delta(X)$ is also a homeomorphim by Lemma 10. Each MMP-hierarchy $\left(A_1^{MMP}, A_2^{MMP}, ...\right)$ can thus be identified as an ambiguous hierarchy $\left(\delta^{\mathcal{K}}\left(A_1^{MMP}\right), \delta^{\mathcal{K}}\left(A_2^{MMP}\right), ...\right)$. Moreover, by Lemma 8.(iii), $\mathrm{marg}_X \circ \delta = \delta \circ \mathrm{Proj}_X$.[10] MMP-coherency condition can be translated as follows: the belief hiearchy $\left(\delta^{\mathcal{K}}\left(A_1^{MMP}\right), \delta^{\mathcal{K}}\left(A_2^{MMP}\right), ...\right)$ is coherent if, for each $n \geq 1$,

$$\mathrm{marg}_{X_{n-1}^{MMP}}^{\mathcal{K}}\left(\delta^{\mathcal{K}}\left(A_{n+1}^{MMP}\right)\right) = \delta^{\mathcal{K}}\left(A_n^{MMP}\right). \tag{2.5.1}$$

[9] Brandenburger and Keisler [18] provide a microfoudation of this result in terms of first-order language.

[10] Observe that if K is a compact subset of X, then

$$\mathrm{marg}_X\left(\delta(K)\right) = \delta\left(\mathrm{Proj}_X^{-1}(K)\right)$$

is compact, because Proj is a proper map (see [13], Theoreme 1, p.115).

Denote by v the canonical homeomorphism from Θ_∞^{MMP} to $\mathcal{K}\left(S \times \Theta_\infty^{MMP}\right)$. Observe that if S is compact Hausdorff, so is H_∞.[11] The universal space Θ_∞^{MMP} can thus be identified as a compact subset of H_∞ by the embedding map $\delta^{\mathcal{K}} \circ v : \Theta_\infty^{MMP} \to \mathcal{K}\left(S \times \Theta_\infty^{MMP}\right) \to \mathcal{K}\Delta\left(S \times H_\infty\right)$.

It should be noted that the existence of the universal space Θ_∞^{MMP} depends crucially on the compactness of S. ([42], Proposition 2). In the next subsection we shall prove that, under the same assumptions as in Section 2.2, a complete, universal possibility structure does exist, hence strengthening MMPs findings. The result is not proved by adapting the method of the proof of MMP. Rather, we use the above described "embedding procedure" which identifies each MMP-hierarchy with an ambiguous hierarchy of the space H_∞. The crucial point is to show that, in absence of the compactness assumption for the parameter space, the universal space of MMP-like hierarchies is non-empty. This will be accomplished by using a construction scheme, based on the notion of projective sequence of measures, which is a slight adaptation of the one in Chapter I.

2.5.2 The universal space of complete possibility structures

As in Section 2.3, let S -the parameter space- be a non-empty Hausdorff space and $I = \{1, 2\}$ the set of players. The space of Dirac probability measures on some space X^i for player $i \in I$ is denoted $\delta\left(X^i\right)$. Observe that $\delta^{\mathcal{K}}\left(\mathcal{K}X^i\right)$ is homeomorphic to $\mathcal{K}\delta\left(X^i\right)$; hence, for each $A^i \subseteq X^i$, A^i compact, we slight abuse notation by regarding $\delta^{\mathcal{K}}\left(A^i\right)$ as a subset of $\mathcal{K}\delta\left(X^i\right)$. We define inductively two sequences of spaces, $\left\{\Theta_{n-1}^i\right\}_{n\geq 1}$ and $\left\{P_n^i\right\}_{n\geq 1}$, as follows.

For each player $i \in I$, let $\Theta_0^i = S$, $P_1^i = \mathcal{K}\delta\left(\Theta_0^i\right)$,, and for all $n \geq 1$,

$$\begin{aligned} \Theta_n^i &= \Theta_0^i \times P_n^{-i}; \\ P_{n+1}^i &= \left\{ \begin{array}{c} \left(\delta^{\mathcal{K}}\left(A_1^i\right), ..., \delta^{\mathcal{K}}\left(A_n^i\right), \delta^{\mathcal{K}}\left(A_{n+1}^i\right)\right) \in P_n^i \times \mathcal{K}\delta\left(\Theta_n^i\right) : \\ \operatorname{marg}^{\mathcal{K}}_{\Theta_{n-1}^i}\left(\delta^{\mathcal{K}}\left(A_{n+1}^i\right)\right) = \delta^{\mathcal{K}}\left(A_n^i\right) \end{array} \right\}. \end{aligned}$$

The space Θ_n^i is player i's *domain of possibility* of level $n+1$: it consists of the parameter space and the states of nature player $-i$ considers possible, what player $-i$ "thinks" about the states of natures player i considers possible,..., and so on, up lo level n. The compact set of

[11] This follows from the finite intersection property of the family of compact sets $\{H_k\}_{k\geq 1}$.

Dirac measures $\delta^{\mathcal{K}}\left(A^i_{n+1}\right)$ is player i's possibility set on the space Θ^i_n, while P^i_{n+1} is the set of lower-order possibility sets which are coherent with $\delta^{\mathcal{K}}\left(A^i_{n+1}\right)$. It should be remarked that not only player i's possibility sets are coherent but she also considers only coherent possibility sets of player $-i$.[12] This implies that, compared to the system of spaces defined by (2.3.1)-(2.3.2), the two sequences of spaces, $\left\{P^i_n\right\}$ and $\left\{\Theta^i_n\right\}$, are such that

$$\begin{aligned} \Theta^i_n &\subseteq X^i_n, \\ P^i_{n+1} &\subseteq \prod_{l=0}^{n} \mathcal{K}\delta\left(X^i_l\right) \subseteq \prod_{l=0}^{n} \mathcal{K}\Delta\left(X^i_l\right). \end{aligned}$$

For each $i \in I$, $n \geq 1$, let $\pi^i_{n,n+1} : P^i_{n+1} \to P^i_n$ denote the projection on the factor spaces of the sequence $\left\{P^i_n\right\}$. The projection $\sigma^i_{n-1,n} : \Theta^i_n \to \Theta^i_{n-1}$ satisfies

$$\sigma^i_{n-1,n} = \begin{cases} \mathrm{Pr}_{0,1} & n = 1 \\ \left(Id_{\Theta^i_0}; \pi^{-i}_{n-1,n}\right) & n \geq 2 \end{cases}.$$

Clearly, $\sigma^i_{n-1,n}$ is the restriction of $\mathrm{Pr}_{n-1,n} : X^i_n \to X^i_{n-1}$ to the subspace Θ^i_n. The following Lemma, which extends Lemma 6 in Chapter I, implies that this restriction is also onto Θ^i_{n-1}.

Lemma 15 *For all $n \geq 1$, $\pi^i_{n,n+1} : P^i_{n+1} \to P^i_n$ is onto.*

The families $\left\{P^i_n, \pi^i_{n,n+1}\right\}_{n\geq 1}$ and $\left\{\Theta^i_{n-1}, \sigma^i_{n-1,n}\right\}_{n\geq 1}$ are projective sequences of non-empty Hausdorff spaces and surjective bonding maps (in fact, projections); hence, by Theorem 1 in Appendix 1, their projective limits, $\varprojlim \left\{P^i_n, \pi^i_{n,n+1}\right\}_{n\geq 1}$ and $\varprojlim \left\{\Theta^i_{n-1}, \sigma^i_{n-1,n}\right\}_{n\geq 1}$, are non-empty. Now, define

$$\begin{aligned} P^i &= \left\{\left(\delta^{\mathcal{K}}\left(A^i_1\right), \delta^{\mathcal{K}}\left(A^i_2\right), ...\right) \in \prod_{l=0}^{\infty} \mathcal{K}\delta\left(X^i_l\right) \middle| \left(\delta^{\mathcal{K}}\left(A^i_1\right), ..., \delta^{\mathcal{K}}\left(A^i_n\right)\right) \in P^i_n, \forall n \geq 1\right\}, \\ \Theta^i &= S \times P^{-i}. \end{aligned}$$

[12] In the current construction, we impose coherence at all levels of the hierarchies. In section 2.3 we adopted the alternative construction where hierarchies are initially unrestricted and coherence is imposed eventually. As it is shown in Chapter I, these two construction schemes are equivalent in terms of the epistemic charactherization (i.e., common certainty of coherence) of the universal type structure. However, the present construction is more appropriate to formally prove the existence of a universal structure in the non-compact case.

An easy check shows that P^i is a closed subset of H_1^i, and both P^i and Θ^i can be identified as the projective limits of the sequences $\left\{P_n^i, \pi_{n,n+1}^i\right\}_{n\geq 1}$ and $\left\{\Theta_{n-1}^i, \sigma_{n-1,n}^i\right\}_{n\geq 1}$, respectively, as stated in the following

Claim 2 *P^i and Θ^i are homeomorphic to $\varprojlim \left\{P_n^i, \pi_{n,n+1}^i\right\}_{n\geq 1}$ and $\varprojlim \left\{\Theta_{n-1}^i, \sigma_{n-1,n}^i\right\}_{n\geq 1}$, respectively.*

The proof of this claim can be provided by adapting the main lines of the proof of Claim 1 in Chapter I, Appendix 6.3. We omit the details here.

Note that there is a common uncertainty space S, hence the sets Θ^i and P^i are the copies of the same sets Θ and P, respectively. As in Section 2.3, if no confusion may arise, we omit the superscript i (or $-i$). The space $S \times P$ is the terminal domain of possibility for each player, as stated in the next result.

Proposition 6 *P is a non-empty, closed subset of H_∞. Moreover, the restriction of g to P induces a homeomorphism $\widetilde{g} : P \to \mathcal{K}\delta\left(S \times P\right)$.*

As MMP pointed out, neither the Mertens and Zamir type space nor the MMP possibility model can be embedded into the other. The construction outlined above shows that the space $P \cap \overline{H}_\infty^c$ is non-empty: in particular, it contains all the "perfect information" types, i.e., all the hierarchies of beliefs associated with a game with perfect information.

Define v as the composite map $\widetilde{g} \circ \left(\delta^{\mathcal{K}}\right)^{-1} : P \to \mathcal{K}\delta\left(S \times P\right) \to \mathcal{K}\left(S \times P\right)$. Finally, the structure $\mathcal{P}_u = (S, P, P, v, v)$ can be shown to be a complete, universal possibility structure by arguments analogous to those in Section 2.4.

2.6 Discussion

This section discusses some technical and conceptual aspects of the paper.

2.6.1 On hypertopologies

Lemma 12 plays the central technical role in the construction of the universal structure $\mathcal{T}_u$ in Section 2.3. Since it requires that all topologies be Hausdorff, we use the narrow topology on the space $\Delta(X)$ and then the Vietoris topology on the correponding hyperspace $\mathcal{K}(\Delta(X))$ to obtain the Hausdorff separation property, according to Lemma 8 and Lemma 10. A comprehensive discussion of the separation properties of the narrow topology on $\Delta(X)$ and its role in the construction of a universal structure is given in Chapter I.

A natural concern regards the choice of the "right" hypertopology on the space $\mathcal{K}(\Delta(X))$ which should be at least Hausdorff in order to apply Lemma 12. As discussed comprehensively in [9], in general the Vietoris topology is too much strong for some applications. Among the class of topologies on a given hyperspace, the *Fell topology*, called also topology of closed convergence, is defined as the topology τ_F on $\mathcal{K}(X)$ generated by subbasic sets of the form V^- where V is an open subset of X, and $(K^c)^+$ where K is a compact subset of X, and K^c means the complement of K. If compact subsets in the definition are replaced by closed subsets (i.e., K is closed, so K^c is open), we obtain the Vietoris topology τ_V. Since the family of compact sets is smaller than the family of closed sets (in a Hausdorff space), it holds always that the Fell topology is weaker than the Vietoris topology, and they coincide if the underlying space is compact. However, in non-compact cases, the Fell topology fails the Hausdorff separation property for $\mathcal{K}(X)$ unless the underlying space X is assumed to be *locally* compact (see [2], Lemma 3.92).[13] In our framework, this means that the space $\mathcal{K}(\Delta(X))$, endowed with the Fell topology, is Hausdorff whenever $\Delta(X)$ is locally compact. However, as it is shown in [37], Theorem 2, $\Delta(X)$ is locally compact if and only if X is compact[14] - hence $\Delta(X)$ is compact, according to Lemma 7. This means that the Fell topology is not useful for our porpouses, since it works only when X is compact - and in this last case it coincides with the Vietoris topology. This motivates our choice of the Vietoris topology as a *default* topology for the hyperspaces under considerations.

[13] Lemma 3.92 in [2] is stated for the space $\mathcal{C}(X)$ of closed subsets of the topological space X. The result easily extends to the hyperspace $\mathcal{K}(X)$.

[14] The result in [37] is stated for completely regular, Hausdorff topological spaces and for the set of tight signed measures, endowed with the weak* topology. The assumption of complete regularity for the topological space X can be dispensed with by assuming that the space $\mathcal{M}(X)$ of tight signed measures is endowed with the narrow topology. In this case, $\Delta(X)$ is a strict subset of $\mathcal{M}(X)$, in that every Radon measure is tight (cf. [11]).

2.6.2 On the compactness assumption

Ambiguous beliefs are modelled here as compact sets of Radon probability measures. But the only property of compact sets of measures that we used in the proof is the Hausdorff separation property of hyperspaces of the form $\mathcal{K}(\Delta(X))$ (Lemma 9.1; see also Appendix 2.8.2). Redefine the set $\mathcal{K}(\Delta(X))$ to be the space of all *closed*, non-empty subsets of Radon probability measures defined on the Borel σ-field of the *regular* Hausdorff space X. According to Lemma 7 and Theorem 4.9 in [45], the space $\mathcal{K}(\Delta(X))$ is Hausdorff. The previous sections can be read now as a proof of the existence a universal type structure when ambiguous beliefs are represented by closed sets of Radon probability measures.

This result has some important implications for the class of possibility structures, specifically for the representation of possibility sets as closed subsets of Hausdorff spaces. First, this family of complete possibility structures include that one in MMP, as it imposes less restrictions on the parameter space and possibility sets: recall the any compact Hausdorff space is also regular, and that compact sets are closed. The assumption of complete regularity for the parameter space is not overly restrictive, in that it includes important cases (e.g., Polish parameter spaces) commonly used in applications. Moreover, a possibility structure based on closed sets leads to the following characterization of possibility sets. Let $\mathcal{P} = (S, P_i, v_i)_{i \in I}$ be any complete possibility structure. The restriction for possibility sets of each player $i \in I$ to be closed subsets of $S \times P_i$, is equivalent to consider the quotient space $2^{S \times P_i} / \sim$, where any two sets $X, Y \in 2^{S \times P_i}$ are $\sim$-equivalent if and only if they have the same closure (cf. MMP, p. 308).

This immediately leads to the related question whether it is possible to obtain a universal possibility structure in the general case where possibility sets are not restricted to be closed.[15] Let $\mathcal{O}X$ denote the space of open subsets of X distinct from X itself, and endow $\mathcal{O}X$ with the topology generated by the complements of the sets in $\mathcal{K}X$. This implies that $\mathcal{O}X$ is homeomorphic to $\mathcal{K}X$.[16] The construction outlined in Section 2.5 yields the existence of a space $\widetilde{P}$ and a homeomorphism $\widetilde{v} : \widetilde{P} \to \mathcal{O}\left(S \times \widetilde{P}\right)$. A universal, complete possibility structure based on open sets cannot exist for the following reason. By definition, a possibility structure based

[15] This question was also addressed by MMP. Their (negative) answer depends crucially on the compactness of possibility sets.

[16] The Vietoris topology on $\mathcal{O}X$ is generated by all subsets of the form $\{O \in \mathcal{O}X | F \subseteq O\}$ and $\{O \in \mathcal{O}X | O \cup F \neq X\}$ for $F \in \mathcal{K}X$. Observe that $X \notin \mathcal{O}X$, because $\emptyset \notin \mathcal{K}X$.

on open sets must not include the empty set in the class of possibility sets. In this last case, a space of the form $\mathcal{O}X\backslash\emptyset$ is no longer homeomorphic to $\mathcal{K}X$;[17] consequently, a hierarchy of possibility open sets cannot belong to the space $\widetilde{P}$, as this space is homeomorphic to $\mathcal{O}\left(S\times\widetilde{P}\right)$, which instead contains the empty set. Rather, following the interpretation of MMP, the set $\mathcal{O}\left(S\times\widetilde{P}\right)$ should be considered as the space of subsets regarded as *impossible* by a player, and the *impossibility structure* $\widetilde{\mathcal{P}}=\left(S,\widetilde{P},\widetilde{P},\widetilde{v},\widetilde{v}\right)$ as the mirror image of $\mathcal{P}_u$.

2.6.3 Hierarchies of conditional beliefs

Models of hierarchical beliefs are particularly important for the epistemic analysis of solution concepts in games. In the context of extensive-form games, Battigalli and Siniscalchi [6] extend the notion of Mertens and Zamir's belief hierarchy by constructing a universal structure whose elements are collections of conditional beliefs (mathematically, conditional probability systems - see [51]). Specifically, each level of a belief hiearchy is represented by a probability measure conditional to a non-empty collection of sets representing the events that are observable by a player. Under this interpretation, an infinite hierarchy of conditional beliefs encodes an agent's disposition to believe conditional on every "relevant hypothesis". Battigalli and Siniscalchi use this notion of hierarchies of conditional beliefs to provide an epistemic analysis of forward induction in dynamic games. The analysis outlined in the previous sections can be extended to such a setting. It is easy to check that Ahn's construction of a universal space of conditional, ambiguous belief systems ([1], Section 4.1) can be appropriately adapted to show that it continues to hold under Assumptions 1 and 2 in Section 2.3. It is worth noting that, since separability plays no role in our framework (recall that both in [1] and in [6], the parameter space is assumed to be Polish), the collection of conditioning events is not required to be at most countable.[18] In this way, it is possible to provide, by means of the "embedding procedure" outlined in Section 2.5, a generalization of MMP's universal structure of conditional possibility sets ([42], Section 5).

[17] In MMP's framework ([42], footnote 3), $\mathcal{O}X\backslash\emptyset$ is not compact.

[18] Let $\mathcal{B}$ the non-empty collection of events in S wich are obervable by a player. The only assumption needed in our framework is that $\mathcal{B}$ must be a collection of non-empty clopen (i.e., both closed and open) sets of S such that $S\in\mathcal{B}$ (cf. [42]).

2.6.4 Hierarchies of preferences

Assuming that the space of nature states S is compact Hausdorff, Epstein and Wang [23] (EW, hereafter) construct a space T^{EW} whose elements are hierarchies of *preferences*, rather than beliefs. They prove that the compact space T^{EW} is homeomorphic to $\mathcal{P}\left(S \times T^{EW}\right)$, the compact set of all *regular* preferences over acts defined on $S \times T^{EW}$; Chen [20] shows the universality of the type structure associated with T^{EW}.

The EW set-up is general enough to include several existing constructions of type structures. The standard Mertens and Zamir's type space T^{MZ} can be viewed as a special case of EW, where preferences conform to the expected utility theory. The crucial aspect of this identification is provided by Theorem 4.2 in EW, according to which if S is compact Hausdorff, then $\Delta(S)$ is homeomorphic to a closed (and compact) subspace of regular preferences $\mathcal{P}^{MZ}(S)$ under the expected utility functional. The space $\mathcal{P}^{MZ}(S)$ is a model of preferences, and by Theorem 6.1(a) in [23], it can be derived a space T^{MZ} homeomorphic to $\mathcal{P}\left(S \times T^{MZ}\right)$ and hence to $\Delta\left(S \times T^{MZ}\right)$. An analogous argument shows that also the space Θ_{∞}^{MMP} can be regarded as a subset of T^{EW}. This follows again from Theorem 4.2 in [23], by identifying the compact set S with its set of Dirac measures $\delta(S)$, which is homeomorphic to the space of regular preferences $\mathcal{P}^{MMT}(S)$. These positive results should suggest that also Ahn's type structure could be identified with a model of preferences.[19] The problem is that Theorem 4.2 in [23] is not valid for the space $\mathcal{K}\Delta(S)$. However, as pointed out by Chen ([20], Proposition 1), the space $\mathcal{V}\Delta(S) \subseteq \mathcal{K}\Delta(S)$ of compact *convex* subsets of Radon probability measures can be identified with a model of preferences and related to the space T^{EW}.

We do not know how our findings relate to these existing embedding results, in that a version of the EW type structure along the directions outlined in this paper (namely, under Assumptions 1 and 2 in Section 2.3) is still lacking.[20] Addressing this issue clearly requires some technical reformulation of the EW approach, namely the choice of the "right" topology on the space $\mathcal{P}(S)$ of all regular preferences so to be compatible with Vietoris topology on $\mathcal{K}\Delta(S)$, as well as a generalization of Theorem 4.2. in [23]. Therefore, this task is best left as

[19] As it is shown in [20], Appendix A.4, the space T^{EW} cannot be embedded into Ahn's space of ambiguous hierarchies, and hence in H_{∞} as well.

[20] Chen [20] proposes a similar line of research for explaining some puzzling features (like his example in Section 4) arising from the existing, partial embedding results.

a possible topic for further research.

2.7 Appendix 1: Summary of projective limit theory

In this appendix, we provide some of the background definitions and results from the theory of projective limit spaces, that are necessary to present the results in Section 2.4. For a more thorough treatment see [22] or [50]. As it is costumary, we denote by $\mathbb{N}$ the set of natural numbers.

A *projective sequence* is a family $\{Y_n, f_{m,n}\}_{m,n\in\mathbb{N}}$ of spaces Y_n and functions $f_{m,n} : Y_n \to Y_m$ such that:

- for each $n \in \mathbb{N}$, Y_n is non-empty topological space;
- for any $m, n \in \mathbb{N}$ satisfying $m \leq n$, $f_{m,n}$ is continuous;
- $f_{m,p} = f_{m,n} \circ f_{n,p}$ for any $m, n, p \in \mathbb{N}$ satisfying $m \leq n \leq p$, and $f_{n,n} = Id_{Y_n}$ for every $n \in \mathbb{N}$.

The spaces Y_n are called *coordinate* (or *factor*) *spaces* and the maps $f_{m,n}$ are called *bonding maps*. Given a projective sequence $\{Y_n, f_{m,n}\}_{m,n\in\mathbb{N}}$, the *projective limit of* $\{Y_n, f_{m,n}\}_{m,n\in\mathbb{N}}$, denoted by $\varprojlim \{Y_n, f_{m,n}\}_{m,n\in\mathbb{N}}$, is defined as

$$\varprojlim \{Y_n, f_{m,n}\}_{m,n\in\mathbb{N}} = \left\{ \{y_n\} \in \prod_{n\in\mathbb{N}} Y_n \,|y_m = f_{m,n}(y_n)\,, \text{ for each } m, n \in \mathbb{N} \text{ s.t. } m \leq n \right\}.$$

For each $l \in \mathbb{N}$, the map $\overline{f}_l : \varprojlim \{Y_n, f_{m,n}\}_{m,n\in\mathbb{N}} \to Y_l$ is the restriction of the projection map $\text{Proj}_{Y_l} : \prod_{n\in\mathbb{N}} Y_n \to Y_l$ to $\varprojlim \{Y_n, f_{m,n}\}_{m,n\in\mathbb{N}}$. Clearly, for any $m, n \in \mathbb{N}$ such that $m \leq n$, the maps $\overline{f}_n$ and $\overline{f}_m$ satisfy the equality $\overline{f}_m = f_{m,n} \circ \overline{f}_n$.

The projective limit $\varprojlim \{Y_n, f_{m,n}\}_{m,n\in\mathbb{N}}$ inherits the subspace topology as a subset of the product $\prod_{n\in\mathbb{N}} Y_n$. It is known (see [22], Proposition 2.5.1) that if $\{Y_n, f_{m,n}\}_{m,n\in\mathbb{N}}$ is a projective sequence of Hausdorff spaces Y_n then its projective limit, $\varprojlim \{Y_n, f_{m,n}\}_{m,n\in\mathbb{N}}$, is a closed subset of the Cartesian product $\prod_{n\in\mathbb{N}} Y_n$.

The next Theorem, a special case of Proposition 5, p. 198, of Bourbaki [15], provides a sufficient condition for a projective limit to be non-empty.

Theorem 3 *Let* $\{Y_n, f_{m,n}\}_{m,n\in\mathbb{N}}$ *be a projective sequence of topological spaces* Y_n *and* surjective *bonding maps* $f_{m,n}$. *Then, if* $Y = \varprojlim \{Y_n, f_{m,n}\}_{m,n\in\mathbb{N}}$, *the map* $\overline{f}_n : Y \to Y_n$ *is surjective for each* $n \in \mathbb{N}$, *and* Y *is non-empty provided that none of the* Y_n*'s is empty.*

As a corollary of this theorem, it can be proved that if each bonding map $f_{m,n}$ is the coordinate projection, then the projective limit space, $\varprojlim \{Y_n, f_{m,n}\}_{m,n\in\mathbb{N}}$, can be identified (homeomorphically) with the Cartesian product $\prod_{n\in\mathbb{N}} Y_n$.

For each $n \in \mathbb{N}$, let μ_n be a Radon probability measure on the Hausdorff space Y_n. The family $\{Y_n, f_{m,n}, \mu_n\}_{m,n\in\mathbb{N}}$ is a *projective sequence of Radon probability measures* if

- $\{Y_n, f_{m,n}\}_{m,n\in\mathbb{N}}$ is a projective sequence of Hausdorff spaces Y_n;
- $\mu_m = \mathcal{L}_{f_{m,n}}(\mu_n)$, for any $m, n \in \mathbb{N}$ such that $m \leq n$.

The structure $\{Y, \overline{f}_n, \mu\}_{n\in\mathbb{N}}$ is called *measure projective limit* of the projective sequence of Radon probability measures $\{Y_n, f_{m,n}, \mu_n\}_{m,n\in\mathbb{N}}$ if

- $Y = \varprojlim \{Y_n, f_{m,n}\}_{m,n\in\mathbb{N}}$, and $\overline{f}_m : Y \to Y_m$ satisfies $\overline{f}_m = f_{m,n} \circ \overline{f}_n$ for any $m, n \in \mathbb{N}$ such that $m \leq n$.
- μ is a Radon probability measure on Y such that

$$\mathcal{L}_{\overline{f}_n}(\mu) = \mu_n \text{ for any } n \in \mathbb{N}.$$

The next theorem, a generalization of Kolmogorov's extension theorem (see [15], p.53-54), will be used extensively in the proof of Proposition 6.

Theorem 4 *Let* $\{Y_n, f_{m,n}, \mu_n\}_{m,n\in\mathbb{N}}$ *be a projective sequence of Radon probability measures. Then the measure projective limit* $\{Y, \overline{f}_n, \mu\}_{n\in\mathbb{N}}$ *exists and* μ *is a unique Radon probability measure.*

2.8 Appendix 2: Proofs for Section 2.2

2.8.1 Proof of Lemma 10

(1) To show continuity of $f^{\mathcal{K}}$, let $\langle U_1, ..., U_n \rangle$ be a basic open set in $\mathcal{K}Y$, that contains $f^{\mathcal{K}}(A)$. Then $\left\langle f^{-1}(U_1), ..., f^{-1}(U_n) \right\rangle$ is a basic open set in $\mathcal{K}X$ containing A such that if K is in $\left\langle f^{-1}(U_1), ..., f^{-1}(U_n) \right\rangle$, then $f^{\mathcal{K}}(K)$ is in $\langle U_1, ..., U_n \rangle$.

(2) Let B be a non-empty, compact subset of Y. We show that there exists $A \in \mathcal{K}X$ such that $f^{\mathcal{K}}(A) = B$. Since f is surjective, then $f(A) = B$ for some non-empty $A \subseteq X$, and $A \subseteq f^{-1}(B)$. Let $\{O_\lambda\}_{\lambda \in \Lambda}$ be an open cover of B. Continuity of f implies that $f^{-1}(\cup_{\lambda \in \Lambda} O_\lambda) = \cup_{\lambda \in \Lambda} f^{-1}(O_\lambda)$ is an open set which covers A. The set B can be covered by a finite subcover $\{O_\lambda\}_{\lambda \in \Lambda^*}$, where Λ^* is a finite index set. Therefore $\left\{f^{-1}(O_\lambda)\right\}_{\lambda \in \Lambda^*}$ is a finite subcover of A.

(3) Let A, $B \in \mathcal{K}X$ such that $A \cap B = \varnothing$. By continuity of f, $f(A)$ and $f(B)$ are compact. Since f is injective, $f(x_1) \neq f(x_1)$ for any $x_1 \in A$, $x_2 \in B$, i.e., $f(A) \cap f(B) = \varnothing$. But this means $f^{\mathcal{K}}(A) \neq f^{\mathcal{K}}(B)$.

(4) Statements (1) and (2) imply that $\mathrm{Proj}_X^{\mathcal{K}}$ is continuous and onto. Recall that Proj_X is an open map. Openess of the induced mapping $\mathrm{Proj}_X^{\mathcal{K}}$ is a consequence of the following lemma, which extends Lemma 4.1 of [32] to arbitrary topological spaces and Vietoris convergence.

Lemma 16 *Let X, Y be topological spaces and $f : X \to Y$ be continuous and onto. Then f is open if and only if $f^{-1}(y_\alpha) \xrightarrow{\tau_V} f^{-1}(y)$ for every net (y_α) in Y such that $y_\alpha \to y$.*

Proof. (Sufficiency) Let U be an open subset of X. We need to show that $f(U)$ is open in Y. This will be accomplished by showing that $Y \backslash f(U)$ is closed. Let (y_α) be a net in $Y \backslash f(U)$ such that $y_\alpha \to y$. If $y \in f(U)$, then $f^{-1}(y) \cap U \neq \varnothing$. Since $f^{-1}(y_\alpha) \xrightarrow{\tau_V} f^{-1}(y)$, it follows that $f^{-1}(y_\alpha) \cap U \neq \varnothing$ eventually, which implies $f\left(f^{-1}(y_\alpha) \cap U\right) \subseteq \{y_\alpha\} \cap f(U) \neq \varnothing$ eventually, i.e., $y_\alpha \in f(U)$ eventually, a contradiction. Therefore $y \in Y \backslash f(U)$, that is, $Y \backslash f(U)$ is closed.

(Necessity) Let (y_α) be a net in Y such that $y_\alpha \to y$. By definition, there exists an open set U in Y such that $\{y\} \subseteq U$, hence $\{y_\alpha\} \subseteq U$ eventually. Clearly $f^{-1}(y) \subseteq f^{-1}(U)$, and $f^{-1}(U)$ is open by the assumption of f to be continuous, while $f^{-1}(y)$ is non-empty, because f is onto. Consequently $f^{-1}(y_\alpha) \subseteq f^{-1}(U)$ eventually. This establishes Upper Vietoris convergence. Let us show that $f^{-1}(y_\alpha) \xrightarrow{V^-} f^{-1}(y)$. Let $x \in f^{-1}(y)$ and let V be an open neighborhood of x.

Since f is open, $f(V)$ is an open neighborhood of $y = f(x)$, hence $y_\alpha \in f(V)$ eventually, which implies that $f^{-1}(y_\alpha) \cap V \neq \varnothing$ eventually. ■

Let (B_α) be a net in $\mathcal{K}X$ such that $B_\alpha \overset{\tau_V}{\to} B$. According to the Lemma 16, it remains to show that $\left(\mathrm{Proj}_X^{\mathcal{K}}\right)^{-1}(B_\alpha) \overset{\tau_V}{\to} \left(\mathrm{Proj}_X^{\mathcal{K}}\right)^{-1}(B)$ eventually. But this follows from the definition of product topology.

(5) See Lemma 2, (iv), in [42].

2.8.2 Some remarks on Lemma 10

The statements (1)-(3) and (5) in Lemma 10 continue to hold even in the case $\mathcal{K}X$ is the hyperspace of all closed subsets of X. In this last case, the proof of Lemma 10.(2) is even simpler. Let B a closed subset of Y. The surjectivity of f implies the existence of a subset A of X such that $f(A) = B$. The set A is closed since, again by the surjectivity and continuity of f, $f(A) = f\left(f^{-1}(B)\right) = B$. Thus $f^{\mathcal{K}}$ is surjective.

The openess of the map $\mathrm{Proj}_X^{\mathcal{K}} : \mathcal{K}(X \times Y) \to \mathcal{K}X$ established in Lemma 10.(4) plays a crucial role in our framework; e.g. in Lemma 11, which drives the basic result in Proposition 3.

A special case of Lemma 10.(4) (and Lemma 16 as well) is the following. Let X and Y be *continua*, i.e., compact connected metric spaces. The set $\mathcal{K}X$ (resp. $\mathcal{K}Y$) is the hyperspace of all closed (hence compact) subsets of X (resp. Y), and is endowed with the topology metrized by the Hausdorff distance; in this case the notion of Vietoris convergence corresponds to the classical notion of convergence in the sense of Kuratowski (see [2], Definition 3.80).

Let $\mathcal{C}X$ (resp. $\mathcal{C}Y$) denote the hyperspace of all subcontinua in X (resp. Y). Theorem 4.3 in [32] shows that f is an open map if and only if $f^{\mathcal{K}}$ is open, while this is not true for the induced map $f^{\mathcal{C}} : \mathcal{C}X \to \mathcal{C}Y$ between subcontinua; specifically, if $f^{\mathcal{C}}$ is open, so is f, but the reverse implication is not true (cf. the example in [32], p. 244). Lemma 16 does not hold in this case, so the induced map $\mathrm{Proj}_X^{\mathcal{C}} : \mathcal{C}(X \times Y) \to \mathcal{C}Y$ is not necessarily open. Consequently, our framework would not possibly work if multiple priors were represented by subcontinua (e.g. convex sets) of metrizable spaces of probability measures. The following argument shows that this is not the case.

A result by Charatonik et al. ([19], Corollary 19) shows that if Y is a continuum, and $\mathrm{Proj}_{[0,1]} : [0,1] \times Y \to [0,1]$ is the natural projection, then the induced map $\mathrm{Proj}_{[0,1]}^{\mathcal{C}}$ is open.

To understand the implications of this result, recall that an *arc* is a set S such that for any two points s_1 and s_2 in S there is a homeomorphism $h : [0,1] \to S$ such that $h(0) = s_1$ and $h(1) = s_2$. (equivalently, one can say that h is an arc and that S is arcwise-connected.) Hence, Corollary 19 in [19] holds for every arc (or arcwise-connected set) S. In Hausdorff spaces, any convex set is also arcwise-connected; thus if X and Y are *convex continua*, it follows that the induced map $\mathrm{Proj}_X^{\mathcal{C}} : \mathcal{C}(X \times Y) \to \mathcal{C}Y$ is open. Our framework deals with convex spaces of this sort (the spaces of Radon probability measures), thus Lemma 10.(4) holds good even though the hyperspaces under consideration are represented by subcontinua.

2.9 Appendix 3: Proofs for Section 2.4

2.9.1 Proof of Lemma 14.

We begin by showing that type morphisms preserve hierarchy description maps.

Lemma 17 *Fix S-based type spaces $\mathcal{T} = (S, T_i, g_i)_{i \in I}$ and $\mathcal{T}' = \left(S, T_i', g_i'\right)_{i \in I}$, and let m be a type morphism from $\mathcal{T}$ to $\mathcal{T}'$. Then, for each $n \geq 1$, $t_i \in T_i$, $i \in I_0$,*

$$\varphi_i^n(t_i) = \varphi_i^n(m_i(t_i)).$$

Proof. By induction on n. Fix an arbitrary $t_i \in T_i$.

($n = 1$): By definition

$$\begin{aligned} \varphi_i^1(t_i) &= \mathrm{marg}_S^{\mathcal{K}}[g_i(t_i)], \\ \varphi_i^1(m_i(t_i)) &= \mathrm{marg}_S^{\mathcal{K}}\left[g_i'(m_i(t_i))\right]. \end{aligned}$$

Thus we have to show

$$g_i'(m_i(t_i))\left[E_0 \times T_j'\right] = g_i(t_i)[E_0 \times T_j],$$

for any $E_0 \in \mathcal{B}(S)$. Since m is a type morphism, we get

$$\begin{aligned} g_i'(m_i(t_i))\left[E_0 \times T_j'\right] &= \mathcal{L}_{m_{-i}}^{\mathcal{K}} g_i(t_i)\left[E_0 \times T_j'\right] \\ &= g_i(t_i)\left[m_{-i}^{-1}\left(E_0 \times T_j'\right)\right] \\ &= g_i(t_i)\left[\left\{(s,t_j) : (s, m_j(t_j)) \in E_0 \times T_j'\right\}\right] \\ &= g_i(t_i)\left[E_0 \times T_j\right], \end{aligned}$$

as required. This implies $\psi_{-i}^1(s,t_j) = \psi_{-i}^1(s, m_j(t_j))$, for each $s \in S$, $t_j \in T_j$.

($n \geq 2$): Suppose that $\varphi_i^n(t_i) = \varphi_i^n(m_i(t_i))$ holds for $n \geq 2$, which implies that $\psi_{-i}^n = \psi_{-i}^n \circ m_{-i}$. Pick $E_n \in \mathcal{B}(X_n)$. We get

$$\begin{aligned} \varphi_i^{n+1}(t_i)[E_n] &= \mathcal{L}_{\psi_{-i}^n}^{\mathcal{K}} g_i(t_i)[E_n] \\ &= \mathcal{L}_{\psi_{-i}^n \circ m_{-i}}^{\mathcal{K}} g_i(t_i)[E_n] \\ &= \mathcal{L}_{\psi_{-i}^n}^{\mathcal{K}}\left(\mathcal{L}_{m_{-i}}^{\mathcal{K}} g_i(t_i)\right)[E_n] \\ &= \mathcal{L}_{\psi_{-i}^n}^{\mathcal{K}}\left(g_i'(m_i(t_i))\right)[E_n] \\ &= \varphi_i^{n+1}(m_i(t_i))[E_n], \end{aligned}$$

where the first equality is by definition of φ_i^{n+1}, the second equality follows from the implication of the induction hypothesis, the third equality is implied by Lemma 8, the fourth equality follows from the property that m is a type morphism, and the last equality is again by definition of φ_i^{n+1}. This proves that the statement is true for $n+1$, concluding the proof. ■

Corollary 2 *If both $\mathcal{T}$ and $\mathcal{T}'$ are non redundant S-based type spaces, then m is injective.*

Proof. Suppose, by way of contradiction, that m is not injective. Then, for each $i \in I$, there exist t_i, $t_i^* \in T_i$, $t_i \neq t_i^*$, such that $m_i(t_i) = m_i(t_i^*)$. Since $\mathcal{T}'$ is non-redundant, it follows that $\varphi_i(m_i(t_i)) = \varphi_i(m_i(t_i^*))$, and thus, by the above Lemma, $\varphi_i(t_i) = \varphi_i(t_i^*)$. This contradicts the non-redundancy of $\mathcal{T}$. ■

Lemma 18 *Fix non-redundant, S-based type spaces $\mathcal{T} = (S, T_i, g_i)_{i \in I}$ and $\mathcal{T}' = \left(S, T_i', g_i'\right)_{i \in I}$. Let m be a type morphism from $\mathcal{T}$ to $\mathcal{T}'$ such that m^{-1} is measurable. Then, for every $t_i \in T_i$,*

$E \in \mathcal{B}(S \times T_j)$,

$$g_i(t_i)[E] = g_i^{'}(m_i(t_i))[m_{-i}(E)]. \tag{2.9.1}$$

Proof. Observe

$$\begin{aligned} g_i^{'}(m_i(t_i))[m_{-i}(E)] &= \left(\mathcal{L}^{\mathcal{K}}_{m_{-i}} \circ g_i\right)(t_i)[m_{-i}(E)] \\ &= g_i(t_i)\left[m_{-i}^{-1}(m_{-i}(E))\right] \\ &= g_i(t_i)[E], \end{aligned}$$

where in the first two equalities we use the definition of type morphism and in the third we use the fact the m is injective (Corollary 2). Note that (2.9.1) is well defined, since the measurability of m^{-1} implies that $m_{-i}(E) \in \mathcal{B}\left(S \times T_j^{'}\right)$, for every $E \in \mathcal{B}(S \times T_j)$. ■

The proof of Lemma 14 follows immediately from Lemma 18.

2.9.2 Proof of Proposition 5

The proof is divided in three main steps. In the first step, we show that for each $t_i \in T_i$, the corresponding *i-description* $\varphi_i(t_i)$ belongs to H_∞, the collection of infinite hierarchies of ambiguous beliefs satisfying coherence and common certainty of coherence. In the second step, we show that the map $\varphi = (Id_S, \varphi_1, \varphi_2)$ is a type morphism. In the third step, we show that this type morphism is unique. In all the three cases, the proof is by induction.

First step: $\varphi_i(T_i) \subseteq H_\infty$.

By definition of i-description, $\varphi_i(T_i) \subseteq H_0$. We use induction to prove $\varphi_i(T_i) \subseteq H_\infty$.

(*Base step*): To show $\varphi_i(T_i) \subseteq H_1$, we need to verify that for all $t_i \in T_i$, $n \geq 1$,

$$\text{marg}^{\mathcal{K}}_{X_{n-1}}\left(\varphi_i^{n+1}(t_i)\right) = \varphi_i^n(t_i),$$

that is,

$$\mathcal{L}^{\mathcal{K}}_{\text{Pr}_{n-1,n}}\left(\varphi_i^{n+1}(t_i)\right) = \varphi_i^n(t_i). \tag{2.9.2}$$

(recall the maps $\text{Pr}_{n-1,n} : X_n = X_{n-1} \times \mathcal{K}\Delta(X_{n-1}) \rightarrow X_{n-1}$ and $\varphi_i^{n+1} : T_i \rightarrow \mathcal{K}\Delta(X_n) =$

$\mathcal{K}\Delta\left(X_{n-1}\times\mathcal{K}\Delta\left(X_{n-1}\right)\right)$. To this end, pick any $E_{n-1}\in\mathcal{B}\left(X_{n-1}\right)$. Then

$$
\begin{aligned}
\mathcal{L}^{\mathcal{K}}_{\mathrm{Pr}_{n-1,n}}\left(\varphi_i^{n+1}\left(t_i\right)\right)\left[E_{n-1}\right] &= \varphi_i^{n+1}\left(t_i\right)\left[\mathrm{Pr}^{-1}_{n-1,n}\left(E_{n-1}\right)\right]\\
&= \varphi_i^{n+1}\left(t_i\right)\left[E_{n-1}\times\mathcal{K}\Delta\left(X_{n-1}\right)\right]\\
&= \mathcal{L}^{\mathcal{K}}_{\psi^n_{-i}}\left(g_i\left(t_i\right)\right)\left[E_{n-1}\times\mathcal{K}\Delta\left(X_{n-1}\right)\right]\\
&= g_i\left(t_i\right)\left[\left(\psi^n_{-i}\right)^{-1}\left(E_{n-1}\times\mathcal{K}\Delta\left(X_{n-1}\right)\right)\right]\\
&= g_i\left(t_i\right)\left[\left\{\left(s,t_j\right):\psi^n_{-i}\left(s,t_j\right)\in E_{n-1}\times\mathcal{K}\Delta\left(X_{n-1}\right)\right\}\right]\\
&= g_i\left(t_i\right)\left[\left\{\left(s,t_j\right):\psi^{n-1}_{-i}\left(s,t_j\right)\in E_{n-1},\ \varphi^n_j\left(t_j\right)\in\mathcal{K}\Delta\left(X_{n-1}\right)\right\}\right]\\
&= g_i\left(t_i\right)\left[\left\{\left(s,t_j\right):\psi^{n-1}_{-i}\left(s,t_j\right)\in E_{n-1}\right\}\right]\\
&= g_i\left(t_i\right)\left[\left(\psi^{n-1}_{-i}\right)^{-1}\left(E_{n-1}\right)\right]\\
&= \mathcal{L}^{\mathcal{K}}_{\psi^{n-1}_{-i}}\left(g_i\left(t_i\right)\right)\left[E_{n-1}\right]\\
&= \varphi^n_i\left(t_i\right)\left[E_{n-1}\right],
\end{aligned}
$$

where the sixt equality follows from the definition of ψ^n_{-i}, while the seventh equality follows from the definition of $\varphi^n_j:T_j\to\mathcal{K}\Delta\left(X_{n-1}\right)$. Thus, (2.9.2) is proved.

To prove the inductive step, we need the following

Claim 3 *Let f be the homeomorphism of Proposition 3. Then*

$$f\circ\varphi_i=\mathcal{L}^{\mathcal{K}}_{\varphi_{-i}}\circ g_i, \tag{2.9.3}$$

where $\varphi_{-i}=\left(Id_S,\varphi_j\right)$.

Proof. Take $E_n\in\mathcal{B}\left(X_n\right)=\mathcal{B}\left(X_{n-1}\times\mathcal{K}\Delta\left(X_{n-1}\right)\right)$. Clearly, the set $\mathrm{Proj}^{-1}_{X_n}\left(E_n\right)=E_n\times\prod_{p\geq n}\mathcal{K}\Delta\left(X_p\right)$ is a cylinder of $S\times H_0=S\times\prod_{n=0}^{\infty}\mathcal{K}\Delta\left(X_n\right)$ with base E_n (in X_n). We denote by $\mathcal{A}$ the family of measurable cylinders on $S\times H_0$, which generates the product σ-field $\sigma\left(\mathcal{A}\right)$. As usual, $\mathcal{B}\left(S\times H_0\right)$ denotes the σ-field generated by the product topology on $S\times H_0$.

Now, for each $t_i \in T_i$,

$$\begin{aligned}
f(\varphi_i(t_i))\left[\mathrm{Proj}_{X_n}^{-1}(E_n)\right] &= \mathcal{L}^{\mathcal{K}}_{\mathrm{Proj}_{X_n}}(f(\varphi_i(t_i)))[E_n] \\
&= \varphi_i^{n+1}(t_i)[E_n] \\
&= \mathcal{L}^{\mathcal{K}}_{\psi^n_{-i}}(g_i(t_i))[E_n] \\
&= g_i(t_i)\left[\left(\psi^n_{-i}\right)^{-1}(E_n)\right] \\
&= g_i(t_i)\left[\left\{(s,t_j) : \left(s, \varphi_j^1(t_j), \ldots, \varphi_j^n(t_j)\right) \in E_n\right\}\right] \\
&= g_i(t_i)\left[\left\{(s,t_j) : \left(s, \varphi_j(t_j)\right) \in \mathrm{Proj}_{X_n}^{-1}(E_n)\right\}\right] \\
&= g_i(t_i)\left[\left(\varphi_{-i}\right)^{-1}\left(\mathrm{Proj}_{X_n}^{-1}(E_n)\right)\right] \\
&= \left(\mathcal{L}^{\mathcal{K}}_{\varphi_{-i}} \circ g_i\right)(t_i)\left[\mathrm{Proj}_{X_n}^{-1}(E_n)\right],
\end{aligned}$$

where the second equality follows from (2.9.2) and the fact that f is a homeomorphism, the third equality is by definition, and the sixt equality follows from the definition of cylinder set. Hence (2.9.3) holds true for every cylinder on $S \times H_0$.

To show that (2.9.3) holds for every set in $\sigma(\mathcal{A})$, we check that all the conditions of Dynkin's π-λ theorem (see [2], Lemma 4.10) are satisfied. Let $\mathcal{F} \subseteq \sigma(\mathcal{A})$ be the family of sets E for which (2.9.3) holds.

- $S \times H_0 \in \mathcal{F}$: trivially

$$\begin{aligned}
\mathcal{L}^{\mathcal{K}}_{\psi^n_{-i}}(g_i(t_i))[S \times H_0] &= g_i(t_i)\left[\left(\varphi_{-i}\right)^{-1}(S \times H_0)\right] \\
&= g_i(t_i)\left[\left(Id_S, \varphi_j\right)^{-1}(S \times H_0)\right] \\
&= g_i(t_i)[S \times T_j] \\
&= f(\varphi_i(t_i))[S \times T_j] \\
&= 1.
\end{aligned}$$

- $(A \in \mathcal{F}) \Rightarrow (A^c \in \mathcal{F})$: Let $E \in \mathcal{F}$. So

$$\begin{aligned} f(\varphi_i(t_i))[(S \times H_0) \backslash E] &= 1 - f(\varphi_i(t_i))[E] \\ &= 1 - \left(\mathcal{L}^{\mathcal{K}}_{\varphi_{-i}} \circ g_i\right)(t_i)[(S \times H_0) \backslash E] \\ &= \left(\mathcal{L}^{\mathcal{K}}_{\varphi_{-i}} \circ g_i\right)(t_i)[E^c]. \end{aligned}$$

- $(\{E_n\}_{n \geq 1} \in \mathcal{F}$, E_n pairwise disjoint $\forall n \geq 1) \Rightarrow (\cup_{n \geq 1} E_n \in \mathcal{F})$:

$$\begin{aligned} f(\varphi_i(t_i))[\cup_{n \geq 1} E_n] &= \sum_{n \geq 1} f(\varphi_i(t_i))[E_n] \\ &= \sum_{n \geq 1} \left(\mathcal{L}^{\mathcal{K}}_{\varphi_{-i}} \circ g_i\right)(t_i)[E_n] \\ &= \left(\mathcal{L}^{\mathcal{K}}_{\varphi_{-i}} \circ g_i\right)(t_i)[\cup_{n \geq 1} E_n], \end{aligned}$$

  hence $(\cup_{n \geq 1} E_n) \in \mathcal{F}$.

Thus $\mathcal{F}$ is a λ-system. Since by assumption $\mathcal{A} \subseteq \mathcal{F}$, it follows from the π-λ theorem that $\mathcal{F} = \sigma(\mathcal{A})$.

Finally, remark that the product σ-field $\sigma(\mathcal{A})$ is in general smaller than the σ-field $\mathcal{B}(S \times H_0)$ (they coincide if each factor space has a countable base for the topology - see [2], Theorem 4.44). Thus, we invoke Theorem 2 in [52] which states the existence of a uniquely determined Radon probability measure on $\mathcal{B}(S \times H_0)$ extending any product measure on $\sigma(\mathcal{A})$. This concludes the proof of the claim. ■

(*Inductive step*): Recall that $\varphi_i(t_i) \in H_k$, $k \geq 2$, if and only if $f(\varphi_i(t_i)) \in \mathcal{K}\Delta(S \times H_{k-1})$, for each $t_i \in T_i$. Suppose that $\varphi_j(T_j) \subseteq H_{k-1}$. Then, for each $t_i \in T_i$:

$$\begin{aligned} f(\varphi_i(t_i))[S \times H_{k-1}] &= \left(\mathcal{L}^{\mathcal{K}}_{\varphi_{-i}} \circ g_i\right)(t_i)[S \times H_{k-1}] \\ &= g_i(t_i)\left[(\varphi_{-i})^{-1}(S \times H_{k-1})\right] \\ &= g_i(t_i)\left[\{(s, t_j) : \varphi_j(t_j) \in H_{k-1}\}\right] \\ &= g_i(t_i)[S \times T_j] \\ &= 1, \end{aligned}$$

where the first equality follows from Claim 3 and the fourth from the induction hypothesis. Thus $f(\varphi_i(t_i)) \in \mathcal{K}\Delta(S \times H_{k-1})$, as required.

Second step: $\varphi = (Id_S, \varphi_1, \varphi_2)$ is a type morphism.

First, we show that φ is continuous. Since Id_S is continuous, we need to show - by induction - that $\varphi_i = (\varphi_i^1, \varphi_i^2, ...)$ is continuous, for each $i \in I$. By definition, $\varphi_i^1 = \mathcal{L}^{\mathcal{K}}_{\text{Proj}_S} \circ g_i$, where g_i is continuous by assumption, and $\mathcal{L}^{\mathcal{K}}_{\text{Proj}_S}$ is continuous by Lemma 8 and Lemma 10. Hence φ_i^1 is continuous, for each $i \in I$. Now assume, by way of induction, that for $i \in I$, $k = 1, ...n$, φ_i^k is continuous. This implies that $\psi^n_{-i} = \left(s, \varphi_j^1, ..., \varphi_j^n\right)$ is also continuous. Then, by Lemma 8 and Lemma 10, $\mathcal{L}^{\mathcal{K}}_{\varphi_{-i}}$ is continuous and thus $\varphi_i^{n+1} = \mathcal{L}^{\mathcal{K}}_{\varphi_{-i}} \circ g_i$ is also continuous. Finally, note that, since $\varphi_i(T_i) \subseteq H_\infty$ for each $i \in I$ (as proved in the first step), it follows from Proposition 4 that equation (2.9.3) can be written as

$$g \circ \varphi_i = \mathcal{L}^{\mathcal{K}}_{\varphi_{-i}} \circ g_i, \tag{2.9.4}$$

which implies that condition (2) in Definition 4 holds. Hence $\varphi = (Id_S, \varphi_1, \varphi_2)$ is a type morphism, as required.

Third step: Uniqueness of the type morphism φ.

Suppose that $\phi = (Id_S, \phi_1, \phi_2)$ is a type morphism from $\mathcal{T}$ to $\mathcal{T}_u$. We have to show $\phi = \varphi$. Since $g \circ \phi_i = \mathcal{L}^{\mathcal{K}}_{\varphi_{-i}} \circ g_i$, and g is invertible (being a homeomorphism), then $\varphi_i = g^{-1} \circ \mathcal{L}^{\mathcal{K}}_{\varphi_{-i}} \circ g_i$. Hence we can write the $(n+1)$-th component of $\phi_i : T_i \to H_\infty$ as

$$\begin{aligned} \phi_i^{n+1}(t_i) &= g^{-1}\left(\mathcal{L}^{\mathcal{K}}_{\phi^n_{-i}}(g_i(t_i))\right) \\ &= \text{marg}^{\mathcal{K}}_{X_n}\left(\mathcal{L}^{\mathcal{K}}_{\phi^n_{-i}}(g_i(t_i))\right) \end{aligned}$$

by coherence of beliefs. Thus it suffices to show that for each $n \geq 0$, $i \in I$, $t_i \in T_i$,

$$\varphi_i^{n+1}(t_i) = \text{marg}^{\mathcal{K}}_{X_n}\left(\mathcal{L}^{\mathcal{K}}_{\phi^n_{-i}}(g_i(t_i))\right).$$

The proof is by induction.

(*Base step*): Let A_0 be a Borel subset of $X_0 = S$. Then

$$\begin{aligned}
\operatorname{marg}^{\mathcal{K}}_{X_o}\left(\mathcal{L}^{\mathcal{K}}_{\phi^0_{-i}}\left(g_i\left(t_i\right)\right)\right)\left[A_0\right] &= \operatorname{marg}^{\mathcal{K}}_{S}\left(\mathcal{L}^{\mathcal{K}}_{Id_S}\left(g_i\left(t_i\right)\right)\right)\left[A_0\right] \\
&= \operatorname{marg}^{\mathcal{K}}_{S}\left(g_i\left(t_i\right)\right)\left[Id_S^{-1}\left(A_0\right)\right] \\
&= \operatorname{marg}^{\mathcal{K}}_{S}\left(g_i\left(t_i\right)\right)\left[A_0\right] \\
&= \varphi^1_i\left(t_i\right)\left[A_0\right],
\end{aligned}$$

which proves the claim for $n = 0$.

(*Inductive step*): Suppose that the statement holds true for $n = 0, ..., k-1$. This implies

$$\begin{aligned}
\left(s, \phi^1_j\left(t_j\right), ..., \phi^k_j\left(t_j\right)\right) &= \left(s, \operatorname{marg}^{\mathcal{K}}_{X_o}\left(\mathcal{L}^{\mathcal{K}}_{\phi^0_{-i}}\left(g_i\left(t_i\right)\right)\right), ..., \operatorname{marg}^{\mathcal{K}}_{X_{k-1}}\left(\mathcal{L}^{\mathcal{K}}_{\phi^{k-1}_{-i}}\left(g_i\left(t_i\right)\right)\right)\right) \\
&= \psi^k_{-i}\left(s, t_j\right),
\end{aligned}$$

for any $s \in S$, $t_i \in T_i$, $t_j \in T_j$. Pick $A_k \in \mathcal{B}\left(X_k\right)$; the set $\operatorname{Proj}^{-1}_{X_k}\left(A_k\right) = A_k \times \prod_{l \geq k} \mathcal{K}\Delta\left(X_l\right)$ is a cylinder of $S \times H_0$ with base A_k (in X_k). Then

$$\begin{aligned}
\operatorname{marg}^{\mathcal{K}}_{X_k}\left(\mathcal{L}^{\mathcal{K}}_{\phi^k_{-i}}\left(g_i\left(t_i\right)\right)\right)\left[A_k\right] &= \mathcal{L}^{\mathcal{K}}_{\phi^k_{-i}}\left(g_i\left(t_i\right)\right)\left[\operatorname{Proj}^{-1}_{X_k}\left(A_k\right)\right] \\
&= g_i\left(t_i\right)\left[\left(\phi^k_{-i}\right)^{-1}\left(\operatorname{Proj}^{-1}_{X_k}\left(A_k\right)\right)\right] \\
&= g_i\left(t_i\right)\left[\left\{\left(s, t_j\right) : \left(s, \phi_j\left(t_j\right)\right) \in \operatorname{Proj}^{-1}_{X_k}\left(A_k\right)\right\}\right] \\
&= g_i\left(t_i\right)\left[\left\{\left(s, t_j\right) : \left(s, \phi^1_j\left(t_j\right), ..., \phi^k_j\left(t_j\right)\right) \in A_k\right\}\right] \\
&= g_i\left(t_i\right)\left[\left\{\left(s, t_j\right) : \psi^k_{-i}\left(s, t_j\right) \in A_k\right\}\right] \\
&= g_i\left(t_i\right)\left[\left(\psi^k_{-i}\right)^{-1}\left(A_k\right)\right] \\
&= \varphi^{k+1}_i\left(t_i\right)\left[A_k\right],
\end{aligned}$$

where the third equality follows from the definition of cylinder, and the fifth equality follows from the induction hypothesis. This concludes the proof of the third step.

2.9.3 Belief closed subspaces.

Definition 8 *For each $i \in I$, let T_i be a Borel subset of H_1. The product $S \times T_1 \times T_2$ is a belief-closed subspace of $S \times H_1 \times H_1$ if*

$$f(h_i) \in \mathcal{K}\Delta(S \times T_j),$$

for each $h_i \in T_i$, $i, j \in I$.

Lemma 19 *Let $\mathcal{T} = (S, T_i, g_i)_{i \in I}$ be a non-redundant S-based type space such that, for each $i \in I$, S and T_i are Lusin spaces. Then $S \times \varphi_1(T_1) \times \varphi_2(T_2)$ is a belief-closed subspace of $S \times H_1 \times H_1$.*

Proof. According to Definition 8, we need to show that, for each $i \in I$, $\varphi_i(T_i)$ is a Borel subset of H_1. (Observe that in this case H_1 is Lusin, being so S.) Since T_i is Lusin, the claim follows from [11], Corolloary 2.3.9. ■

Lemma 13 encompasses two classical results on non-redundant type spaces, namely

- S, T_1 and T_2 are compact Hausdorff space (cf. [44])
- S, T_1 and T_2 are complete, separable metric (Polish) spaces (cf. [24]).

The Claim follows from the fact that both compact Hausdorff spaces and Polish spaces are Lusin spaces.[21] Observe that Lemma 14 provides a result which is stronger than what the definition of belief-closed subspace requires. In the general case, a non redundant type space $\mathcal{T} = (S, T_i, g_i)_{i \in I}$ may fail to identify a belief-closed subspace of $S \times H_1 \times H_1$, as $\varphi_i(T_i)$ is not necessarily a Borel subset of H_1. This motivates the following definition.

Definition 9 *For each $i \in I$, let T_i be a subset of H_1. The product $S \times T_1 \times T_2$ is an* pseudo belief-closed subspace *of $S \times H_1 \times H_1$ if there exist $T_1^* \subseteq T_1$ and $T_2^* \subseteq T_2$ such that T_1^*, $T_2^* \in \mathcal{B}(H_1)$ and*

$$f(h_i) \in \mathcal{K}\Delta\left(S \times T_j^*\right),$$

for each $h_i \in T_i$, $i, j \in I$.

[21] In Mertens and Zamir's case, one obtains a stronger result: $\varphi_i(T_i)$ is compact in H_1.

Clearly, any belief-closed subspace of $S \times H_1 \times H_1$ is also "pseudo" belief-closed.

If $\left(S \times T_1^\theta \times T_2^\theta\right)_{\theta \in \Theta}$ is any family of belief closed subspaces of $S \times H_1 \times H_1$, then

$$\cup_{\theta \in \Theta} \left(S \times T_1^\theta \times T_2^\theta\right) = S \times \left(\cup_{\theta \in \Theta} T_1^\theta\right) \times \left(\cup_{\theta \in \Theta} T_2^\theta\right)$$

is "pseudo" belief-closed, in the sense that for each $h_i \in \cup_{\theta \in \Theta} T_i^\theta$, there is $E \in \mathcal{B}\left(\cup_{\theta \in \Theta} T_j^\theta\right)$ (namely the appropriate T_j^θ) such that $f(h_i)[S \times E] = 1$ (here the qualifier "pseudo" refers to the possibility that $\cup_{\theta \in \Theta} T_j^\theta$ is not a Borel subset of H_1).

For each $i \in I$, define $H^* = \cup_{\theta \in \Theta} T_i^\theta$. Clearly, the set $S \times H^* \times H^*$ is "pseudo" belief-closed. However, we have the following result:

Proposition 7 $H^* = H_\infty$.

Proof. Since $S \times H_\infty \times H_\infty$ is belief-closed, it is also "pseudo" belief-closed, hence $H^* \supseteq H_\infty$. In the other direction, we show $S \times T_1 \times T_2 \subseteq S \times H_k \times H_k$, for each $k \geq 1$. The statement is vacously true for $k = 1$. Suppose that is true for $k \geq 2$. This implies that each $h_i \in T_i \subseteq H_k$ is such that $f(h_i) \in \mathcal{K}\Delta(S \times T_j)$, where $T_j \subseteq H_k$. Hence each $h_i \in T_i$ satisfies $f(h_i) \in \mathcal{K}\Delta(S \times H_k)$, that is, $h_i \in H_{k+1}$. Thus $H^* = \cup_{\theta \in \Theta} T_i^\theta \subseteq H_\infty$. ■

2.10 Appendix 4: Proofs for Section 2.5

2.10.1 Proof of Lemma 15

To prove the statement, for every $\left(\delta^{\mathcal{K}}\left(A_1^i\right), ..., \delta^{\mathcal{K}}\left(A_n^i\right)\right) \in P_n^i$, $n \geq 1$, we have to find $\delta^{\mathcal{K}}\left(A_{n+1}^i\right) \in \mathcal{K}\delta\left(\Theta_n^i\right)$ such that

$$\mathrm{marg}_{\Theta_{n-1}^i}^{\mathcal{K}} \delta^{\mathcal{K}}\left(A_{n+1}^i\right) = \delta^{\mathcal{K}}\left(A_n^i\right).$$

Toward this end, we show the existence of a continuous function $\Psi_n^i : \Theta_n^i \to \Theta_{n+1}^i$ such that $\sigma_{n,n+1}^i \circ \Psi_n^i = Id_{\Theta_n^i}$, for all $n \geq 0$. The proof is by induction.

(Base step) Let B_1^{-i} be a compact subset of $\Theta_0^{-i} = S$ and define the map $\Psi_0^i : \Theta_0^i \to \Theta_1^i = \Theta_0^i \times P_1^{-i}$ by

$$\Psi_0^i(s) = \left(s, \delta^{\mathcal{K}}\left(B_1^{-i}\right)\right), \ s \in S.$$

Ψ_0^i is continuous, as each of its components is continuous, and clearly $\sigma_{0,1}^i \circ \Psi_1^i = Id_{\Theta_0^i}$. The set

$$\begin{aligned}\delta^{\mathcal{K}}\left[\left(\Psi_0^i\right)^{\mathcal{K}}\left(A_1^i\right)\right] &= \delta\left\{\Psi_0^i\left(s\right)\left|s \in A_1^i \subseteq S,\, A_1^i \text{ compact}\right.\right\} \\ &= \delta\left\{\left(s, \delta^{\mathcal{K}}\left(B_1^{-i}\right)\right)\left|s \in A_1^i \subseteq S,\, A_1^i \text{ compact}\right.\right\}\end{aligned}$$

is a set of Dirac measures on Θ_1^i according to which player i is certain that: (1) the realized state of nature is in A_1^i, and (2) player $-i$ is certain that the realized state of nature is in B_1^{-i}.

(Induction step) Suppose that there exists a continuous function $\Psi_{n-1}^i : \Theta_{n-1}^i \to \Theta_n^i$ such that $\sigma_{n-1,n}^i \circ \Psi_{n-1}^i = Id_{\Theta_{n-1}^i}$. Define the function $\left(\delta^i \circ \Psi_{n-1}^i\right)^{\mathcal{K}} : \mathcal{K}\Theta_{n-1}^i \to \mathcal{K}\delta\left(\Theta_n^i\right)$ by

$$\left(\delta^i \circ \Psi_{n-1}^i\right)^{\mathcal{K}}\left(A_n^i\right) = \delta\left(\Psi_{n-1}^i\left(A_n^i\right)\right),$$

$A_n^i \in \mathcal{K}\Theta_{n-1}^i$. This function is clearly continuous. By the induction hypothesis $\sigma_{n-1,n}^i \circ \Psi_{n-1}^i = Id_{\Theta_{n-1}^i}$, thus

$$\begin{aligned}\operatorname{marg}_{\Theta_{n-1}^i}^{\mathcal{K}}\left(\delta^{\mathcal{K}}\left[\left(\Psi_n^i\right)^{\mathcal{K}}\left(A_n^i\right)\right]\right) &= \left(\delta^{\mathcal{K}} \circ \left(\sigma_{n-1,n}^i\right)^{\mathcal{K}}\right)\left[\left(\Psi_{n-1}^i\right)^{\mathcal{K}}\left(A_n^i\right)\right] \\ &= \delta^{\mathcal{K}}\left(\left(\sigma_{n-1,n}^i\right)^{\mathcal{K}} \circ \left(\Psi_{n-1}^i\right)^{\mathcal{K}}\right)\left[A_n^i\right] \\ &= \delta^{\mathcal{K}}\left(\sigma_{n-1,n}^i \circ \Psi_{n-1}^i\right)^{\mathcal{K}}\left[A_n^i\right] \\ &= \delta^{\mathcal{K}}\left(Id_{\Theta_{n-1}^i}^{\mathcal{K}}\right)\left[A_n^i\right] \\ &= \delta^{\mathcal{K}}\left(A_n^i\right),\end{aligned}$$

where the first equality follows from Lemma 8. (iii) and the third from Lemma 10. Finally, define $\Psi_n^i : \Theta_n^i \to \Theta_{n+1}^i$ by

$$\begin{aligned}&\Psi_n^i\left(s, \left(\delta^{\mathcal{K}}\left(B_1^{-i}\right), ..., \delta^{\mathcal{K}}\left(B_n^{-i}\right)\right)\right) \\ =\ &\left(s, \left(\delta^{\mathcal{K}}\left(B_1^{-i}\right), ..., \delta^{\mathcal{K}}\left(B_n^{-i}\right), \delta^{\mathcal{K}}\left[\Psi_{n-1}^{-i}\left(B_n^{-i}\right)\right]\right)\right),\end{aligned}$$

where $B_l^{-i} \in \mathcal{K}\delta\left(\Theta_{l-1}^{-i}\right)$, $l = 1, ..., n$. The map Ψ_n^i is continuous and such that $\sigma_{n,n+1}^i \circ$

$\Psi_n^i = Id_{\Theta_n^i}$.

2.10.2 Proof of Proposition 6

To economize in notation, we denote by $\left(\delta_1^{\mathcal{K}}, \delta_2^{\mathcal{K}}, ...\right)$ each element in P, where $\delta_n^{\mathcal{K}} = \delta^{\mathcal{K}}(A_n)$, A_n compact in $\delta(\Theta_n)$, for all $n \geq 1$. Note that, for every $\left(\delta_1^{\mathcal{K}}, \delta_2^{\mathcal{K}}, ...\right) \in P$ the structure $\left(\{\Theta_n\}_{n\geq 1}, \{\delta_n^{\mathcal{K}}\}_{n\geq 1}, \{\sigma_{n,n+1}^{\mathcal{K}}\}_{n\geq 1}\right)$ is a projective sequence of compact sets of Dirac (hence Radon) probability measures such that, for all $n \geq 1$,

$$\begin{aligned}\left(\delta_{n+1} \circ \sigma_{n,n+1}^{-1}\right)^{\mathcal{K}} &= \delta^{\mathcal{K}} \circ \sigma_{n,n+1}^{\mathcal{K}} \\ &= \delta_n^{\mathcal{K}},\end{aligned}$$

where the first equality follows from Lemma 8. Denote by $\overline{\sigma}_n$ the projection from Θ to Θ_n. According to Theorem 4, for every $\left(\delta_1^{\mathcal{K}}, \delta_2^{\mathcal{K}}, ...\right) \in P$ there exists a unique, compact set of Radon (in fact, Dirac) probability measures $\delta_*^{\mathcal{K}} \in \mathcal{K}\delta(\Theta) = \mathcal{K}\delta(S \times P)$ such that, for all $n \geq 1$,

$$\begin{aligned}\left(\delta_* \circ \overline{\sigma}_n^{-1}\right)^{\mathcal{K}} &= \delta_*^{\mathcal{K}} \circ \overline{\sigma}_n^{\mathcal{K}} \\ &= \delta_n^{\mathcal{K}}.\end{aligned}$$

That is, there exists a bijective map $\widetilde{g} : P \to \mathcal{K}\delta(S \times P)$ such that, for every hierarchy $\left(\delta^{\mathcal{K}}(A_1), \delta^{\mathcal{K}}(A_2), ...\right) \in P$,

$$\mathrm{marg}_{\Theta_{n-1}}^{\mathcal{K}}\left(\widetilde{g}\left(\left(\delta^{\mathcal{K}}(A_1), \delta^{\mathcal{K}}(A_2), ...\right)\right)\right) = \delta^{\mathcal{K}}(A_n).$$

We show that the map $\widetilde{g}$ is the restriction of $g : H_\infty \to \mathcal{K}\Delta(S \times H_\infty)$ to P. This will be accomplished by showing that P is closed in H_∞. First remark that P is closed in H_1 by coherence of beliefs, thus $\mathcal{K}\delta(S \times P) \subseteq \mathcal{K}\Delta(S \times H_1)$. The homeomorphism f of Proposition 3 implies that $f(P) \subseteq \mathcal{K}\Delta(S \times H_0)$; however, Theorem 3 and the argument above establish that $f(P) = \widetilde{g}(P) = \mathcal{K}\delta(S \times P) \subseteq \mathcal{K}\Delta(S \times H_1)$, hence $P \subseteq f^{-1}(\mathcal{K}\Delta(S \times H_1)) = H_2$. An easy check by induction shows that if $P \subseteq H_{k-1}$ then $P \subseteq H_k$, for every $k \geq 1$. Therefore P is closed in $H_\infty = \cap_{k\geq 1} H_k$, and the restriction of g to P induces the homeomorphism $\widetilde{g} : P \to \mathcal{K}\delta(S \times P)$.

Chapter 3

Which hierarchies of beliefs belong to the universal type structure?

3.1 Introduction

Hierarchies of beliefs are of basic importance in strategic interactions, but their explicit description has limited appeal mainly because it does not provide a suitable model for game theoretic analysis. Motivated by this view, Harsanyi [28] introduced the notions of *type* and *type structures* as a convenient modelling device used to describe interactive uncertainty in games. Suppose the players are uncertain about which events in some set Θ of states hold. The set Θ is called *parameter space*, and it specifies the values of objective parameters - like payoffs - that do not depend on the players' uncertainties. A Θ-based type structure comprises: (1) a (measurable) set of types, called *type space*, each element of which encapsulates all strategically relevant aspects of a player's information about Θ; and (2) for each player's type, a subjective belief (probability measure) about the underlying parameter space Θ and the type spaces of the other players. This simple structure provides an implicit representation about players' uncertainty, in the sense that it does not describe hierarchies of beliefs directly. However, it is possible to associate with the subjective belief of each type, an explicit hierarchy of belief: that is, a player's belief over Θ (i.e., first-order belief), a player's belief over Θ and other players' first-order beliefs (i.e., second-order belief), and so on. Such hierarchies of beliefs are called *description of types*. Harsanyi's approach via type structures carries no loss of generality in that, as showed by Heifetz and Samet [30], the set of all descriptions of types turns out to be a well defined type space; the associated type structure is also *universal*, in the sense that every other type structure can be embedded into it in a way which preserves the measurable structure of belief hierarchies.

In this paper we ask: which hierarchies of beliefs are descriptions of types? This question summarizes important problems and open issues arising from the literature on the universal type structure. The remainder of this Introduction motivates in more detail.

The distinctive feature of type structures is that the induced descriptions of types are *coherent*, in the sense that the marginals of higher-order beliefs equal the corresponding lower-order beliefs. This notion of coherence formalizes the idea that beliefs at different levels do not contradict one another, and it plays an instrumental role in the first topological constructions of universal type structures, like in [44] and [17]. Assuming that the parameter space is a compact Hausdorff topological space, Mertens and Zamir ([44]) show that the set of all coherent hierarchies of beliefs corresponds to the space of all descriptions of types. Under slightly different topological assumptions, Brandenburger and Dekel [17] offer a concise epistemic characterization of the universal type structure: the set of all description of types is identified as the maximal subset of hierarchies of beliefs displaying coherence and common certainty of coherence. Any other type structure imposes further epistemic restrictions on hierarchies of beliefs, and this is common certainty among the players. Therefore, coherence and common certainty of coherence can be viewed as the weakest epistemic restriction imposed on players' hierarchies of beliefs, and the universal type structure as the largest type structure which "epistemically contains" all possible type structures.

As in all subsequent works in topological setups ([29],[43]), the strategy of the proof for the existence of a universal type structure remains the same: it starts with belief hierarchies satisfying this notion of coherence and obtaines the space of all descriptions of types as on output. This goes by means of the following proof scheme:

1. topologize all the relevant domains of uncertainty of belief hierarchies to yield continuous functions between them,

2. use the notion of coherence on belief hierarchies to obtain a projective sequence of prob-

ability measure spaces,

3. show, by means of (generalized versions of) the Kolmogorov Extension Theorem, that this projective sequence of probability measures admits a unique σ-additive extension, and finally

4. establish the existence of a canonical measure-theoretic isomorphism (specifically, a homeomorphism) between the space of all coherent belief hierarchies of a player and the set of all single beliefs over the space of nature states and coherent belief hierarchies of the other players.

The use of a topology on the relevant domains of uncertainty is crucial for obtaining a unique σ-additive extension, in that the Kolmogorov Extension Theorem requires that all probability measures involved be Radon, that is, Borel probability measures which can be approximated from inside by compact sets.[1] By dropping this last assumption, Heifetz and Samet ([31]) provide an example - in the topological setup - of a coherent hierarchy of beliefs which does not admit a σ-additive extension, and hence it cannot be made a type in some type structure. This result has some important interdipendent consequences, both at technical and conceptual levels.

At technical level, the result in [31] implies that the proof scheme outlined above cannot be replicated in the general measure-theoretic framework, where all the relevant spaces and functions are assumed to be purely measurable and no restrictions are imposed on beliefs. This is so because every coherent belief hierarchy defines unambiguously an additive set function (with total mass 1) which cannot be extended to a σ-additive one unless some special "pseudo topological" assumptions on the probability measures are made.[2] As such, Heifetz and Samet ([30]) prove the existence a universal type structure in a completely different fashion. They consider type structures as primitives, and show that the space of all induced descriptions of types is a type space. The σ-additive extension of the set function associated with an arbitrary

[1] Section 5 of the current chapter provides a formal definition of Radon probability measures and some of their properties. For an overview, see [54].

[2] Compact measures, introduced by Marczewski ([41]), are an instance of "pseudo topological" measures. They play a crucial role in abstract formulations of Kolmogorov Extension Theorem, like in [2], Theorem 15.26. In topological settings, every Radon measure is compact.

description of some type is obtained by using the probability measure of some type structure which give rise to this description. Thus, this strategy of proof is the opposite of the one adopted in the topological setup, but is more general, as it does not impose any restriction on beliefs. However, unlike to the topological framework, it leaves open the question of the existence of a canonical isomorphism.

The counterexample in [31], coupled with the existence result of a universal type structure in the measurable setup, implies a somewhat surprising result: the set of all descriptions of types is always a subset of the space of coherent belief hierarchies. This stands in sharp contrast to the topological characterization of the universal type structure. As such, it raises the following conceptual issue. The adoption of the notion of coherence as the natural, minimal requirement to describe all the relevant states of affairs entails that a theoretic analysis of a specific game via type structures may exclude some strategically relevant hierarchies. If coherent hierarchies of beliefs are what really matter, then we must be assured that any hierarchy we might wish to model can be captured in some type structure. It might well be that in some games a coherent hierarchy of beliefs (like that in [31]) is strategically relevant, yet this hierarchy cannot be expressed as type. If this were the case, using type structures could prevent a correct analysis of the problem at hand, hence the Harsanyi framework may not be sufficiently general to model any incomplete information scenario.

Still, an appropriate, stronger definition of coherence on belief hierarchies has to be provided, as the universal type structure cannot be expressed in terms of coherence alone. An acceptable refinement of the notion of coherence should be at least strong enough to charactherize all the descriptions of types in the measure theoretic framework, but at the same time quite general to include the topological constructions of types.

In this paper, we put forward a notion of coherence which satisfies the above requirements. We call it *strong coherence*, as opposed to the foregoing notion of coherence, which will be referred to as *weak coherence*. Both concepts involve a connection between different levels of hierarchies of beliefs. To understand the difference, recall that weak coherence requires that any event (measurable set) E in the space of n-order beliefs must have the same (marginal) probability in any higher order beliefs. If we think of events as those sets of states the players can describe, weak coherence can be seen as a restriction on the hierarchies about only those sets that can be the objects of players' reasoning. The notion of strong coherence, discussed in detail in Section 3.3.1, concerns a restriction on belief hierarchies about all possible, not necessarily measurable, subsets of the relevant domains of uncertainties. Intuitively, the additional restriction corresponds to the interpretation that sets representing what players can not describe in the space of n-order beliefs must be of the "same size" (expressed by means of the outer probability measure) in any higher order domain of uncertainty. In other words, a player cannot describe in any high order domain some state of affairs which are not expressible in lower order domains.

Much work in the current paper is devoted to establishing that strongly coherent hierarchies represent the hierarchies of beliefs contained in the universal type structure. To do this, we adapt the proof scheme outlined above to the our measure-theoretic setting. Here, the problem of the σ-additivity of the set function associated with a strongly coherent hierarchy is solved by using a recent result obtained by Pintèr [47], who shows that a projective sequence of probability spaces, under an additional assumption called ϵ-completeness, admits a unique limit extension on the projective limit space. Indeed, our notion of strong coherence comprises both the classical notion of self consistency of probability measures and the concept of ϵ-completeness.[3] Our main result (Theorem 5) establishes the existence of a measure-theoretic isomorphism between the space of all strongly coherent hierarchies of a player and the set of all single beliefs over the space of nature states and other players' strongly coherent hierarchies. This reconciles the topological and measure theoretic approaches to type structures in two main aspects. First, it is relatively straightforward to show (Proposition 12) that each type structure can be embedded in this space. Thus, from the viewpoint of the existence of a universal type structure, our characterization of strongly coherent hierarchies supplements the strategy of the proof in [30].[4] Second, we show that in all mentioned topological works strong coherence is equivalent ot weak coherence (Proposition 13), hence the isomorphism in Theorem 5 reduces to the classical canonical homeomorphism. This means that, contrarily to what stated in [30], the universal,

[3] In the current chapter, we adopt the term strong in place of ϵ-complete to denote a special class of projective sequences of probability spaces. See the Appendix.

[4] As Heifetz and Samet ([30], p.341) note: "(...) the standard way of proving the existence of a universal space, namely by showing that the space of all coherent description is the desired space, could not possibly work in the general measure-theoretic framework". Our results show that the standard way works successfully if only strongly coherent hierarchies are considered.

measurable type structure does not differ from the topological one, as they both represent the one and the same typology of hierarchies. As a matter of fact, the notion of strong coherence offers an intuitive explanation on why the hierarchy of beliefs in [31] (which turns out to be weakly but not strongly coherent) cannot be expressed as a type. We discuss this further in Section 3.6, and conclude that strong - and not weak - coherence is the natural requirement to impose on hierarchies of beliefs in order to describe all possible states of the world.

The remainder of this paper is organized as follows. Section 3.2 introduces the basic notations and terminology. Hierarchies of beliefs and type structures are the subject of Section 3.3. Section 3.4 states and prove the main result, and Section 3.5 relates it to topological type structures. Finally, Section 3.6 concludes. The Appendix contains some definitions and auxiliary results which are needed for the proofs of Section 3.4.1.

3.2 Preliminaries

In this section, we recall a few notions of (probability) measure theory to establish the terminology and the notations used in the paper.

Let X be a measurable space with a σ-field Σ_X, the elements of which are called *events*. Let $\Delta(X)$ denote the space of all σ-additive probability measures on Σ_X. The space $\Delta(X)$ is endowed with the σ-field Σ_Δ which is generated by all sets of the form

$$b^p(E) = \{\mu \in \Delta(X) : \mu(E) \geq p\}$$

where $E \in \Sigma_X$ and $0 \leq p \leq 1$. Note that the σ-field Σ_Δ is the restriction to $\Delta(X)$ of the σ-field generated by the Borel cylinders in $[0,1]^{\Sigma_X}$ (i.e., the σ-field generated by maps $\mu \mapsto \mu(E)$, for all $E \in \Sigma_X$).

For $\mu \in \Delta(X)$, we denote by μ^* the outer measure induced by μ. That is, $\mu^*(A)$ for $A \subseteq X$ is defined by $\mu^*(A) = \inf\{\mu(E) : E \in \Sigma_X,\ A \subseteq E\}$.

Unless otherwise stated, given a countable product space $\prod_{n\in\mathbb{N}} X_n$ where $\mathbb{N}$ is the set of natural numbers, we write Pr_m as the canonical projection from $\prod_{n\in\mathbb{N}} X_n$ to X_m, for each $m \in \mathbb{N}$. For any $l, m \in \mathbb{N}$ satisfying $l \leq m$, we write $\mathrm{Pr}_{l,m}$ as the coordinate projection from $\prod_{j=1}^{m} X_j$ onto $\prod_{j=1}^{l} X_j$. We consider any product, finite or countable, of measurable spaces with the product σ-field and any subspace of a measurable space with the relative σ-field.

Given a probabiliy measure $\mu \in \Delta(X)$ and a measurable map $f : X \to Y$, we denote by $\mu \circ f^{-1}$ the function $\mu \circ f^{-1} : \Delta(X) \to \Delta(Y)$. That is, $\mu \circ f^{-1}$ is the image measure of μ under f defined by $\mu \circ f^{-1}(A) = \mu\left(f^{-1}(A)\right)$ for all measurable subset A of Y. An easy check shows that the map $\mu \circ f^{-1}$ is also measurable.[5] Corrispondingly, we denote by $\mu^* \circ f^{-1}$ the outer measure induced by $\mu \circ f^{-1}$.

For a measurable space X, we write Id_X as the identity map on X, that is, $Id_X(x) = x$ for all $x \in X$.

3.3 Hierarchies of beliefs and measurable type structures

Throughout, we fix a two-player set I;[6] given a player $i \in I$, we denote by j the other player in I. Both players share a common space S, called *parameter space* or *space of nature states*, representing all the relevant objective parameters of the game. The space S is a measurable space endowed with the σ-field Σ_S. The sets in Σ_S reflect the subsets in S the players can describe, i.e., Σ_S reflects the players' language about S. Define $I_0 = I \cup \{0\}$, where "0" stands for "nature". Hence, for each $i \in I$, we refer to $-i$ as the set $\{0, j\}$.

In the following, we consider two different approaches to model higher order beliefs about S, the explicit approach by hierarchies (Section 3.3.1) and the Harsanyi implicit approach by type structures (Section 3.3.2).

[5] See, for instance, [36], Lemma 4.

[6] The generalization to three or more players is immediate.

3.3.1 Weakly and strongly coherent hierarchies of beliefs

For each $i \in I$, let $\left\{X_n^i\right\}_{n=0}^{\infty}$ be a sequence of spaces defined inductively as follows:

$$\begin{aligned} X_0^i &= S \\ X_1^i &= X_0^j \times \Delta\left(X_0^j\right) \\ X_2^i &= X_1^j \times \Delta\left(X_1^j\right) \\ &= X_0^j \times \Delta\left(X_0^j\right) \times \Delta\left(X_1^j\right) \\ &\quad ... \\ X_{n+1}^i &= X_n^j \times \Delta\left(X_n^j\right) \\ &= X_0^j \times \prod_{l=0}^{n} \Delta\left(X_l^j\right) \\ &\quad ... \end{aligned}$$

Each space X_n^i represents the domain of uncertainty of player i's higher order beliefs. A probability measure μ_{n+1}^i over X_n^i is called a $(n+1)$-order belief. For each $i \in I$, set $H^i = \prod_{n=0}^{\infty} \Delta\left(X_n^i\right)$.[7] The infinite product H^i is the space of all possible, infinite belief hierarchies for player i, with generic element $h^i = \left(\mu_1^i, \mu_2^i, ...\right)$.

Definition 10 (Weak Coherence) *A belief hierarchy $h^i \in H^i$ is weakly coherent if and only if, for all events $E_{n-1} \subseteq X_{n-1}^i$, $n = 1, 2, ...$,*

$$\mu_{n+1}^i \circ \left(Pr_{n-1,n}^i\right)^{-1} \left(E_{n-1}\right) = \mu_n^i \left(E_{n-1}\right). \tag{3.3.1}$$

This notion of weak coherence (called simply coherence in the relevant literature) states that beliefs of different order assign the same probability to the same *event* (measurable set). Thereby, it concerns only with the players' language. The following proposition states an important property of weakly coherent beliefs in terms of all (not necessarily measurable) subsets of the domains of uncertainty. It links the notion of weak coherence to that one of strong coherence which is introduced below.

[7] Note that since there is a common parameter space S, the sets H^i and H^j are the copies of the same set H. We do not drop the superscript i (or j) as it makes the results in Section 3.4 more transparent.

Proposition 8 *If $h^i \in H^i$ is weakly coherent, then for all $E_{n-1} \subseteq X^i_{n-1}$, $n = 1, 2, ...,$*

$$\left(\mu^i_{n+1}\right)^* \circ \left(Pr^i_{n-1,n}\right)^{-1}\left(E_{n-1}\right) \leq \left(\mu^i_n\right)^*\left(E_{n-1}\right). \tag{3.3.2}$$

Proof. Let E_n be an arbitrary subset of X^i_n. By the surjectivity of the projection mappings, there exists $E_{n-1} \subseteq X^i_{n-1}$ such that $E_n = \left(\mathrm{Pr}^i_{n-1,n}\right)^{-1}\left(E_{n-1}\right)$. By Carathéodory's Theorem ([50], Theorem II.9, p.39-40), there exists a measurable set A_{n-1} in X^i_{n-1} such that $E_{n-1} \subseteq A_{n-1}$ and $\left(\mu^i_n\right)^*\left(E_{n-1}\right) = \mu^i_n\left(A_{n-1}\right)$. Since $\left(\mathrm{Pr}^i_{n-1,n}\right)^{-1}\left(E_{n-1}\right) \subseteq \left(\mathrm{Pr}^i_{n-1,n}\right)^{-1}\left(A_{n-1}\right)$, we get

$$\begin{aligned}
\left(\mu^i_{n+1}\right)^*\left[\left(\mathrm{Pr}^i_{n-1,n}\right)^{-1}\left(E_{n-1}\right)\right] &\leq \left(\mu^i_{n+1}\right)^*\left[\left(\mathrm{Pr}^i_{n-1,n}\right)^{-1}\left(A_{n-1}\right)\right] \\
&= \mu^i_{n+1}\left[\left(\mathrm{Pr}^i_{n-1,n}\right)^{-1}\left(A_{n-1}\right)\right] \\
&= \mu^i_n\left(A_{n-1}\right) \\
&= \left(\mu^i_n\right)^*\left(E_{n-1}\right),
\end{aligned}$$

where the first equality follows from the measurability of $\left(\mathrm{Pr}^i_{n-1,n}\right)^{-1}\left(A_{n-1}\right)$, and the second equality follows from the weak coherence of belief hierarchies. ■

Definition 11 (Strong Coherence) *A belief hierarchy $h^i \in H^i$ is strongly coherent if and only if, for all $E_{n-1} \subseteq X^i_{n-1}$, $n = 1, 2, ...,$*

$$\left(\mu^i_{n+1}\right)^* \circ \left(Pr^i_{n-1,n}\right)^{-1}\left(E_{n-1}\right) = \left(\mu^i_n\right)^*\left(E_{n-1}\right). \tag{3.3.3}$$

Accordingly, we say that the $(n+1)$-level belief μ^i_{n+1} is weakly (resp. strongly) coherent with respect to μ^i_n if (3.3.1) (resp. (3.3.3)) holds.

Of course, strong coherence implies weak coherence. A closer look at the proof of Proposition 8 delivers the difference between the two concepts. If the set $E_n \subseteq X^i_n$ is non-measurable, so is $E_{n-1} \subseteq X^i_{n-1}$. In this case, strong coherence yields $\left(\mu^i_{n+1}\right)^*\left(E_n\right) = \left(\mu^i_n\right)^*\left(E_{n-1}\right)$; that is, the family of non-measurable subsets in some order domain of uncertainty should be of the "same size" along all order domains of uncertainty. If measurable sets reflect the players' language, strong coherence additionally requires that what players *cannot* describe in some order domain of uncertainty should not be describable in any higher order domain. By contrast, weak coherence involves only measurable sets, i.e., the objects of reasoning of the players. It might be well that the inequality in (3.3.2) could be strict. We discuss this further in Section 6.3 (see also Example 1 in the Appendix), while in Section 5 we state a condition under which weak coherence implies strong coherence.

Remark 1 *A reinterpretation of strong coherence, which can be seen as a "mirror image" of the one already given, is as follows. Let* $E_n \subseteq X_n^i$ *be an arbitrary event. By a slight adaptation of the proof of Proposition 8, is is easy to check that, if strong coherence holds,* $\mu_{n+1}^i (E_n) = \mu_{n+1}^i \left[\left(Pr_{n-1,n}^i \right)^{-1} (A_{n-1}) \right]$*, that is,*

$$\mu_{n+1}^i \left[\left(\left(Pr_{n-1,n}^i \right)^{-1} (A_{n-1}) \right) \triangle E_n \right] = 0$$

for some A_{n-1} *measurable in* X_{n-1}^i*. (here* $\triangle$ *stands for the symmetric difference). In other words, the class of all events in the space of* $(n+1)$*-level beliefs differs from the family of the events induced by the space of* n*-level beliefs only up to* μ_{n+1}^i*-null sets. Under this view, strong coherence imposes the requirement that the only language the players can use in some order domain of uncertainty is essentially the one inherited from the lower order domains.*

3.3.2 Type structures and type morphisms

The following definition, which formalizes Harsanyi's implicit approach to hierarchies of beliefs, is borrowed from [30].

Definition 12 (Type Structure) *An* S*-based type structure is a structure* $\mathcal{T} = \langle S, (T_i)_{i \in I}, (m_i)_{i \in I} \rangle$*, where*

1. *for each* $i \in I$*,* T_i *is a measurable space;*
2. *for each* $i \in I$*,* m_i *is a measurable function* $m_i : T_i \to \Delta (S \times T_j)$*.*

The space T_i is called type space, and m_i specifies the belief of each type over nature and other players' types.[8]

[8] Observe that some authors ([6], [30], [31]) use the terminology "type space" for what is called "type structure" here.

The notion of type morphism captures the idea that a type structure $\mathcal{T}$ is "contained in" another type structure $\mathcal{T}_*$ if $\mathcal{T}$ can be embedded into $\mathcal{T}_*$ in a way which preserves the beliefs associated with types. Formally:

Definition 13 *Let $\mathcal{T} = \langle S, (T_i)_{i \in I}, (m_i)_{i \in I} \rangle$ and $\mathcal{T}' = \langle S, (T'_i)_{i \in I}, (m'_i)_{i \in I} \rangle$ be two S-based type structures. The function $\varphi = (\varphi_0, \varphi_1, \varphi_2) : S \times T_i \times T_j \to S \times T'_i \times T'_j$ is called* type morphism *(from $\mathcal{T}$ to $\mathcal{T}'$) if*

1. *$\varphi_0 = Id_S$,*
2. *for each $i \in I$, $t_i \in T_i$, $\varphi_i : T_i \to T'_i$ is a measurable function such that $m'_i(\varphi_i(t_i)) = m_i(t_i) \circ \varphi_{-i}^{-1}$, where $\varphi_{-i} = (\varphi_0, \varphi_j) : S \times T_j \to S \times T'_j$.*

The morphism is a type isomorphism *if φ is a measure theoretic isomorphism. Say $\mathcal{T}$ and $\mathcal{T}'$ are* isomorphic *if there is a type isomorphism between them.*

Condition (2) in the definition of type morphism expresses consistency between the function $\varphi_i : T_i \to T'_i$ and the induced function $\mu \circ \varphi_{-i}^{-1} : \Delta(S \times T_j) \to \Delta\left(S \times T'_j\right)$. To gain the intuition, let t_i be an arbitrarily fixed type in T_i. The function φ_i maps t_i to some $t'_i \in T'_i$, hence $m'_i(t'_i)$ defines a probability measure on $S \times T'_j$. On the other hand, we can use the function m_i to obtain from t_i a probability measure ν of $\Delta(S \times T_j)$, and then use the map $v \circ \varphi_{-i}^{-1}$ to get $v' \in \Delta(S \times T_j)$. Thus, condition (2) states that $m'_i(t'_i)$ and v' coincide as they are both generated by the same type $t_i \in T_i$. In words, a type morphism preserves the implicit description of belief hierarchies for the players.

Definition 14 *An S-based type structure $\mathcal{T}_U = \langle S, (T^*_i)_{i \in I}, (m^*_i)_{i \in I} \rangle$ is universal if for every S-based type structure $\mathcal{T} = \langle S, (T_i)_{i \in I}, (m_i)_{i \in I} \rangle$ there is a unique type morphism from $\mathcal{T}$ to $\mathcal{T}_U$.*

The space T^*_i is called *universal type space* for player i.

Remark 2 *S-based type structures, as objects, and type morphisms, as morphisms, form a category. The S-based universal type structure is the terminal object in the category of S-based type structures.*

Proposition 9 *If it exists, the S-based universal type structure is unique up to type isomorphism.*

Proof. The terminal object of any category, if it exists, is unique up to isomorphism (cf. [5], Proposition 2.8). ■

3.4 Main Result

This section is devoted to showing the main result of the paper.

Theorem 5 *For each player $i \in I$, there exists a non-empty set $\overline{H}^i \subseteq H^i$ and a measure-theoretic isomorphism $f^i : \overline{H}^i \to \Delta\left(S \times \overline{H}^j\right)$ such that*

$$\forall n \geq 1,\ f^i\left(\left(\mu_1^i, \mu_2^i, ...\right)\right) \circ \left(Pr_{n-1,n}^i\right)^{-1} = \mu_n^i.$$

The space $\overline{H}^i \times \overline{H}^j$ is the set of all paires of belief hierarchies satisfying strong coherence and common certainty of strong coherence. The type structure $\mathcal{T}_U = \langle S, \left(\overline{H}^i\right)_{i\in I}, \left(f^i\right)_{i\in I}\rangle$ is the unique universal type structure.

Section 3.4.1 provides a measure-theoretic construction of the type structure $\langle S, \left(\overline{H}^i\right)_{i\in I}, \left(f^i\right)_{i\in I}\rangle$ from strongly coherent hierarchies of beliefs, while Section 3.4.2 shows that this type structure is universal.

Here we sketch the main lines of the proof. In Section 3.4.1, we begin by redefining the notions of domain of uncertainties for belief hierarchies. Unlike to the sequence of spaces $\left\{X_n^i\right\}_{n=0}^{\infty}$ introduced in Section 3.3.1, these new domains of uncertainties require that, at any order of belief hierchies, player i restricts the class of possible events to those for which strong coherence of player j's hierarchies holds. This defines a projective sequence of belief hierarchies displaying strong coherence and common certainty of strong coherence. This step is the measure theoretic analogue to the standard construction scheme used in papers like [44]. We first show that the projective limit of this sequence is non-empty (Proposition 10). Next, we establish the existence of a measure-theoretic isomorphism f^i. This last step is novel to the literature, as it relates the notion of the resulting measure-theoretic isomorphism to the canonical homeomorphism of the topological setup. The proof for the universality applies techniques which are similar to the ones used in [30]. Remarks 4 and 5 in Section 3.4.2 outline the differences.

3.4.1 Type structure containing strongly coherent belief hierarchies

This section inductively constructs the set of all hierarchies of beliefs that are strongly coherent, assign probability 1 to strongly coherent hierarchies of beliefs, and so on. For the relevant definitions concerning projective sequences of spaces, see the Appendix.

For each player $i \in I$, let $\Theta_0^i = S$, $\overline{H}_1^i = \Delta\left(\Theta_0^i\right)$, and for all $n \geq 1$,

$$\begin{aligned} \Theta_n^i &= \Theta_0^i \times \overline{H}_n^j; \\ \overline{H}_{n+1}^i &= \left\{\left(\mu_1^i, ..., \mu_n^i, \mu_{n+1}^i\right) \in \overline{H}_n^i \times \Delta\left(\Theta_n^i\right) : \mu_{n+1}^i \text{ is strongly coherent w.r.to } \mu_n^i\right\}. \end{aligned}$$

The interpretation of the sequences $\left\{\Theta_n^i\right\}$ and $\left\{\overline{H}_{n+1}^i\right\}$ is as follows.

The space Θ_n^i is player i's domain of uncertainty of order $n+1$: it consists of the parameter space and what player j believes about the nature state, what player j believes about what player i believes about the nature state,..., and so on, up lo level n. The space $\overline{H}_{n+1}^i$ is the set of $(n+1)$-tuples of strongly coherent beliefs over $\Theta_0^i, ..., \Theta_n^i$. Notice that not only each i's beliefs are strongly coherent but she also considers only strongly coherent beliefs of j - only those are in support of her beliefs. This implies that, compared to the sequence of spaces $\left\{X_n^i\right\}_{n=0}^{\infty}$ introduced in Section 3.3.1, the sequences $\left\{\Theta_n^i\right\}$ and $\left\{\overline{H}_{n+1}^i\right\}$ are such that

$$\Theta_n^i \subseteq X_n^i,\, \overline{H}_{n+1}^i \subseteq \prod_{l=0}^{n} \Delta\left(X_l^i\right),\ \forall n \geq 0.$$

For each $i \in I$, $n \geq 1$, let $\pi_{n,n+1}^i : \overline{H}_{n+1}^i \to \overline{H}_n^i$ denote the projection on the factor spaces of the sequence $\left\{\overline{H}_n^i\right\}$. The projection $\rho_{n-1,n}^i : \Theta_n^i \to \Theta_{n-1}^i$ satisfies

$$\rho_{n-1,n}^i = \begin{cases} \Pr_{0,1} & n = 1 \\ \left(Id_{\Theta_0^i}; \pi_{n-1,n}^j\right) & n \geq 2 \end{cases}.$$

Clearly, $\rho_{n-1,n}^i$ is the restriction of $\Pr_{n-1,n} : X_n^i \to X_{n-1}^i$ to the subspace Θ_n^i, and hence is a measurable function. If $\left(\mu_1^i, ..., \mu_n^i, \mu_{n+1}^i\right) \in \overline{H}_{n+1}^i$ then μ_{n+1}^i is strongly coherent w.r.to μ_n^i, and formally -

$$\left(\mu_{n+1}^i\right)^* \left[\left(\rho_{n-1,n}^i\right)^{-1} \left(E_{n-1}\right)\right] = \left(\mu_n^i\right)^* \left(E_{n-1}\right),$$

for all $E_{n-1} \subseteq \Theta_{n-1}^i$.

The families $\left\{\overline{H}_n^i, \pi_{n,n+1}^i\right\}_{n\geq 1}$ and $\left\{\Theta_{n-1}^i, \rho_{n-1,n}^i\right\}_{n\geq 1}$ are strong projective sequences of non-empty measurable spaces, and let $\varprojlim \overline{H}_n^i$ and $\varprojlim \Theta_n^i$ be their projective limits. Further, let $\pi_n^i : \varprojlim \overline{H}_n^i \to \overline{H}_n^i$ and $\rho_n^i : \varprojlim \Theta_n^i \to \Theta_n^i$ be the corresponding projections.

Define also

$$\begin{aligned} \overline{H}^i &= \left\{\left(\mu_1^i, \mu_2^i, ...\right) \in H^i \,\middle|\, \left(\mu_1^i, ..., \mu_n^i\right) \in \overline{H}_n^i, \forall n \geq 1\right\}, \\ \Theta^i &= S \times \overline{H}^j. \end{aligned}$$

The following result states that both $\varprojlim \overline{H}_n^i$ and $\varprojlim \Theta_n^i$ can be isomorphically identified with $\overline{H}^i$ and Θ^i, respectively, so that there is no need to distinguish between them. (The proof that $\overline{H}^i$ is a measurable subset of H^i is notationally cumbersome and not crucial for the construction. It is therefore relegated to the Appendix.)

Proposition 10 $\varprojlim \overline{H}_n^i$ *and* $\varprojlim \Theta_n^i$ *are non-empty and they are measure theoretic isomorphic to* $\overline{H}^i$ *and* Θ^i*, respectively.*

Proof. Observe that $\rho_n^i = \left(Id_{\Theta_0^i}; \pi_n^j\right)$, thus it is not hard to verify that $\varprojlim \overline{H}_n^i$ is measure theoretic isomorphic to $\Theta_0^i \times \varprojlim \overline{H}_n^j = S \times \varprojlim \overline{H}_n^j$. So it remains to prove that $\varprojlim \overline{H}_n^i$ is non-empty and measure theoretic isomorphic to $\overline{H}^i$.

First we claim that $\pi_{1,n}^i\left(\overline{H}_n^i\right) \supseteq \pi_{1,n+1}^i\left(\overline{H}_{n+1}^i\right)$ for all $n \geq 1$. Indeed, by definition $\pi_{n,n+1}^i\left(\overline{H}_{n+1}^i\right) \subseteq \overline{H}_n^i$ and $\pi_{1,n+1}^i = \pi_{1,n}^i \circ \pi_{n,n+1}^i$, hence $\pi_{1,n+1}^i\left(\overline{H}_{n+1}^i\right) = \pi_{1,n}^i\left(\pi_{n,n+1}^i\left(\overline{H}_{n+1}^i\right)\right) \subseteq \pi_{1,n}^i\left(\overline{H}_n^i\right)$. From this claim, it follows that the sequence $\left\{V_n^i\right\}_{n\geq 1} \subseteq \overline{H}_1^i$ defined by $V_n^i = \overline{H}_1^i \backslash \pi_{1,n}^i\left(\overline{H}_n^i\right)$ is such that $V_n^i \subseteq V_{n+1}^i$, for all $n \geq 1$. (Observe that the sets V_n^i are not necessarily measurable).

Consider now

$$\begin{aligned}\left(\pi_{1,n}^{i}\right)^{-1}\left(V_{n}^{i}\right) &= \left(\pi_{1,n}^{i}\right)^{-1}\left(\overline{H}_{1}^{i}\setminus\pi_{1,n}^{i}\left(\overline{H}_{n}^{i}\right)\right)\\ &= \left(\pi_{1,n}^{i}\right)^{-1}\left(\overline{H}_{1}^{i}\right)\setminus\left(\pi_{1,n}^{i}\right)^{-1}\left(\pi_{1,n}^{i}\left(\overline{H}_{n}^{i}\right)\right).\end{aligned}$$

By definition $\overline{H}_{n}^{i}\subseteq\left(\pi_{1,n}^{i}\right)^{-1}\left(\pi_{1,n}^{i}\left(\overline{H}_{n}^{i}\right)\right)$, and since $\pi_{1,n}^{i}\left(\overline{H}_{n}^{i}\right)\subseteq\overline{H}_{1}^{i}$ we have that

$$\left(\pi_{1,n}^{i}\right)^{-1}\left(\pi_{1,n}^{i}\left(\overline{H}_{n}^{i}\right)\right)\subseteq\left(\pi_{1,n}^{i}\right)^{-1}\left(\overline{H}_{1}^{i}\right).$$

Still, by definition $\overline{H}_{n}^{i}=\left(\pi_{1,n}^{i}\right)^{-1}\left(\overline{H}_{1}^{i}\right)$, thus $\left(\pi_{1,n}^{i}\right)^{-1}\left(\pi_{1,n}^{i}\left(\overline{H}_{n}^{i}\right)\right)=\left(\pi_{1,n}^{i}\right)^{-1}\left(\overline{H}_{1}^{i}\right)$. That is, $\left(\pi_{1,n}^{i}\right)^{-1}\left(V_{n}^{i}\right)=\emptyset$.

Let E_0 be an arbitrary measurable subset of $\Theta_{0}^{i}=S$. By strong coherence

$$\begin{aligned}\left(\mu_{n+1}^{j}\right)^{*}\left[\left(\rho_{0,n+1}^{i}\right)^{-1}\left(E_0\times V_{n}^{i}\right)\right] &= \left(\mu_{n+1}^{j}\right)^{*}\left[\left(Id_{\Theta_{0}^{i}};\pi_{n-1,n}^{j}\right)^{-1}\left(E_0\times V_{n}^{i}\right)\right]\\ &= \left(\mu_{n+1}^{j}\right)^{*}\left[E_0\times\left(\pi_{1,n}^{i}\right)^{-1}\left(V_{n}^{i}\right)\right]\\ &= \left(\mu_{1}^{j}\right)^{*}\left(E_0\times\emptyset\right)\\ &= 0,\end{aligned}$$

for all $n\geq 1$, hence $\left(\mu_{n+1}^{j}\right)^{*}\left(E_0\times\left(\cup_{n\geq 1}\left(\pi_{1,n}^{i}\right)^{-1}\left(V_{n}^{i}\right)\right)\right)=0$, which implies $\left(\pi_{1,n}^{i}\right)^{-1}\left(\overline{H}_{1}^{i}\setminus\left(\cup_{n\geq 1}^{i}V_{n}^{i}\right)\right)\neq\emptyset$, for all $n\geq 1$. But this means that for all $\mu_{1}^{i}\in\overline{H}_{1}^{i}\setminus\left(\cup_{n\geq 1}V_{n}^{i}\right)$ there exists $h^{i}\in\varprojlim\overline{H}_{n}^{i}$ such that $\mu_{1}^{i}=\pi_{1}^{i}\left(h^{i}\right)$, thereby establishing the non-emptiness of $\varprojlim\overline{H}_{n}^{i}$.

The proof that $\varprojlim\overline{H}_{n}^{i}$ is measure theoretic isomporphic to $\overline{H}^{i}$ is almost the same as that one provided in the topological setup by Claim 1 in Chapther I. It involves only some necessary minor changes (replace the word "homeomorphism" there with "measure-theoretic isomorphism"), so we omit the details. ■

Now, to each $h^{i}=\left(\mu_{1}^{i},\mu_{2}^{i},...\right)\in\overline{H}^{i}$ there corresponds a strong projective sequence of probability spaces

$$\left(\left(\Theta_{n}^{i},\mu_{n+1}^{i}\right),\rho_{m,n}^{i},\mathbb{N}\right)_{m\leq n}=\left(\left(S\times\overline{H}_{n}^{j},\mu_{n+1}^{i}\right),\rho_{m,n}^{i},\mathbb{N}\right)_{m\leq n}$$

where the $(n+1)$-order belief μ_{n+1}^{i} is defined on the σ-field over $\Theta_{n}^{i}=S\times\overline{H}_{n}^{j}$ which is the one ineherited from the product σ-field over $H^i = \prod_{n=0}^{\infty} \Delta\left(X_n^i\right)$.

This projective sequence has a limit, as stated in the following

Lemma 20 *The projective sequence* $\left(\left(\Theta_n^i, \mu_{n+1}^i\right), \rho_{m,n}^i, \mathbb{N}\right)_{m \leq n}$ *admits a unique projective limit probability measure* μ^i *over* Θ^i *such that*

$$\mu^i \circ \left(\rho_n^i\right)^{-1} = \mu_{n+1}^i, \text{ for every } n \geq 0. \tag{3.4.1}$$

Thus there exists a bijective map between $\overline{H}^i$ *and* $\Delta\left(\Theta^i\right)$.

Proof. As is shown in [50] (Lemma IV.1 and p.417), the family $\mathcal{A}_{\Theta^i}$ of sets of the form

$$\left\{\left(\rho_n^i\right)^{-1}(E_n) : n \geq 0,\, E_n \text{ is measurable in } \Theta_n^i\right\}$$

is a field which generates the σ-field on Θ^i; and by weak coherence, there exists a unique additive set function μ^i with total mass 1 on $\mathcal{A}_{\Theta^i}$ for which (3.4.1) holds. Therefore, there is an onto map between $\overline{H}^i$ and the space of additive measures on Θ^i. This map is also injective. To see this, let E_n and E_{n-1} be events in Θ_n^i and Θ_{n-1}^i, respectively, for which $\left(\rho_n^i\right)^{-1}(E_n) = \left(\rho_n^i\right)^{-1}(E_{n-1})$. Then any set function satisfying (3.4.1) yields $\mu_{n+1}^i(E_n) = \mu_n^i(E_{n-1})$. Finally, strong coherence implies, by Theorem 6 in the Appendix, that μ^i is σ-additive on $\mathcal{A}_{\Theta^i}$, hence it can be uniquely extended to the σ-field generated by $\mathcal{A}_{\Theta^i}$. ■

We are ready to prove the existence of a canonical measure theoretic isomorphism.

Proposition 11 $\overline{H}^i$ *is measure-theoretic isomorphic to* $\Delta\left(S \times \overline{H}^j\right)$.

In the proof of Proposition 11 we shall make use of the following two lemmas.

Lemma 21 *Let* $(Y_n)_{n\in\mathbb{N}}$ *be a countable collection of measurable spaces and, for each* $n \in \mathbb{N}$, *fix functions* $f_n : X \to Y_n$, *where* X *is a measurable space. Define* $f : X \to \prod_{n\in\mathbb{N}} Y_n$ *by*

$$f(x) = (f_1(x), f_2(x), ...)$$

for all $x \in X$. *Then* f *is measurable if and only if each* f_n *is measurable.*

Proof. By an easy adaptation of Lemma 4.49 in [2]. ■

Lemma 22 *Let $\mathcal{F}$ be a π-system of sets which generates the σ-field Σ_X on a measurable space X. Then the σ-field $\mathcal{F}_\Delta$ is the same as Σ_Δ.*

Proof. See [58], Lemma 3.6. ■

Proof of Proposition 11. Recall that $\overline{H}^i$ is endowed with the relative σ-field of the product $H^i = \prod_{n=0}^{\infty} \Delta\left(X_n^i\right)$, while $\Delta\left(S \times \overline{H}^j\right)$ is endowed with the σ-field generated by $b^p(E) = \left\{\mu^i \in \Delta\left(S \times \overline{H}^j\right) : \mu^i(E) \geq p\right\}$, for each event $E \subseteq S \times \overline{H}^j$ and $p \in [0,1]$.

Let $g^i : \Delta\left(S \times \overline{H}^j\right) \to \overline{H}^i$ be the map defined by

$$\mu^i \longmapsto \left[g_n^i\left(\mu^i\right)\right]_{n\geq 1} = \left[\mu^i \circ \left(\rho_{n-1}^i\right)^{-1}\right]_{n\geq 1}.$$

(Note that $g_n^i\left(\mu^i\right) \in \overline{H}_n^i$, $n \geq 1$.) Lemma 20 implies that g^i is bijective. Each g_n^i is measurable, thus, by Lemma 21, g^i is measurable. It remains to show that the inverse of g^i, which is denoted by f^i, is measurable. By Corollary 4.24 in [2], it suffices to prove the claim for the system of generators of $\Delta\left(S \times \overline{H}^j\right)$. Let E be a measurable subset of $\Theta_k^i = S \times \overline{H}_k^j$. Consider the set

$$\begin{aligned}\left(f^i\right)^{-1}\left(\beta^p\left(\left(\rho_k^i\right)^{-1}(E)\right)\right) &= \left\{h^i \,\middle|\, f\left(h^i\right)\left(\left(\rho_k^i\right)^{-1}(E)\right) \geq p\right\} \\ &= \left\{h^i \,\middle|\, \mu_{k+1}^i(E) \geq p\right\},\end{aligned} \tag{3.4.2}$$

where the second equality follows from the definition of f^i. The set in (3.4.2) is measurable, as it is a cylinder set with base in $\overline{H}_{k+1}^i$. But the family of cylinder sets $\left\{\left(\rho_k^i\right)^{-1}(E)\right\}$ is a field (hence a π-system) which generates the σ-field over $S \times \overline{H}^j$. By Lemma 22, the family $\beta^p\left(\left(\rho_k^i\right)^{-1}(E)\right)$ generates the σ-field over $\Delta\left(S \times \overline{H}^j\right)$. Thus f^i is measurable. ■

Finally, according to Definition 12, $\mathcal{T}_U = \langle S, \left(\overline{H}^i\right)_{i\in I}, \left(f^i\right)_{i\in I}\rangle$ is a well defined type structure.

Remark 3 *In Chapter I, it is shown that $\varprojlim \overline{H}_n^i$ is non-empty by proving that each map $\pi_{n,n+1}^i : \overline{H}_{n+1}^i \to \overline{H}_n^i$ is onto. This method of the proof depends crucially on topological assumptions, and cannot be replicated in the present framework. The proof used in the current paper follows the ideas of Pintér ([47], Theorem 3.2.(1)).*

3.4.2 From types to belief hierarchies

In this section, we relate S-based type structures to the type structure $\mathcal{T}_U$. To this end, we need to specify how types induce hierarchies of beliefs. The belief hierarchies are derived inductively.

For a given type structure $\langle S, (T_i)_{i\in I}, (m_i)_{i\in I}\rangle$ it is possible to construct a natural map, called *hierarchy description map*, which unfolds the higher order beliefs of each player $i \in I$, by assigning to each $t_i \in T_i$ a strongly coherent hierarchy of beliefs in $\overline{H}^i$. For each $i \in I$, let $\widetilde{h}_0^{-i} : S \times T_j \to S$ be the projection mapping, and for all $k \geq 1$, let $\widetilde{h}_k^0 : S \to S$ be the identity on $S = \Theta_0^j$. The first order descrition map for player i, $\widetilde{h}_1^i : T_i \to \overline{H}_1^i$, is defined by

$$\widetilde{h}_1^i(t_i) = m_i(t_i) \circ \left(\widetilde{h}_0^{-i}\right)^{-1}.$$

We will define the $(k+1)$th orders descrition map. First let $\widetilde{h}_k^{-i} : S \times T_j \to S \times \overline{H}_k^j$ be the unique function, defined by $\widetilde{h}_k^{-i} = \left(\widetilde{h}_k^0, \widetilde{h}_k^j\right)$, which satisfies, for all $k \geq 1$,

$$\widetilde{h}_{k-1}^{-i} = \rho_{k-1,k}^i \circ \widetilde{h}_k^{-i}. \tag{3.4.3}$$

(Recall that $\rho_{k-1,k}^i$ denotes the projection from $\Theta_k^i = S \times \overline{H}_k^j$ to $\Theta_{k-1}^i = S \times \overline{H}_{k-1}^j$). That is, for each couple $(s, t_j) \in S \times T_j$, the expression $\widetilde{h}_k^{-i}(s, t_j) = \left(\widetilde{h}_k^0(s), \widetilde{h}_k^j(t_j)\right)$ describes the nature state s and the beliefs up to order k for $t_j \in T_j$. Inductively, define $\widetilde{h}_{k+1}^i : T_i \to \overline{H}_{k+1}^i$ by

$$\begin{aligned}
\widetilde{h}_{k+1}^i(t_i) &= \left(\widetilde{h}_k^i(t_i), m_i(t_i) \circ \left(\widetilde{h}_k^{-i}\right)^{-1}\right) \\
&= \left(m_i(t_i) \circ \left(\widetilde{h}_0^{-i}\right)^{-1}, ..., m_i(t_i) \circ \left(\widetilde{h}_{k-1}^{-i}\right)^{-1}, m_i(t_i) \circ \left(\widetilde{h}_k^{-i}\right)^{-1}\right).
\end{aligned}$$

Finally, for each $i \in I$, define $\widetilde{h}^i : T_i \to \overline{H}^i$ by $\widetilde{h}^i(t_i) = \left(\widetilde{h}_1^i(t_i), \widetilde{h}_2^1(t_i), ...\right)$. Thus $\widetilde{h}^i(t_i)$ is the *descrition of type* $t_i \in T_i$. Each type induces strongly coherent hierarchies of beliefs, as stated in the following Lemma.

Lemma 23 *The hierarchy decription map* $\widetilde{h}^i : T_i \to \overline{H}^i$ *is well defined and measurable.*

Proof. For a give type $t_i \in T_i$, we need to verify that $\widetilde{h}_{k+1}^i(t_i) \in \overline{H}_{k+1}^i$ for each $k \geq 0$, i.e.,

$$\left(m_i(t_i) \circ \left(\widetilde{h}_k^{-i}\right)^{-1}\right)^* \circ \left(\rho_{k-1,k}^i\right)^{-1}(A) = \left(m_i(t_i) \circ \left(\widetilde{h}_{k-1}^{-i}\right)^{-1}\right)^*(A),$$

for all $A \subseteq \Theta^i_{k-1}$. That is,

$$\left(\widetilde{h}^{-i}_k\right)^{-1} \circ \left(\rho^i_{k-1,k}\right)^{-1} = \left(\widetilde{h}^{-i}_{k-1}\right)^{-1}.$$

But this follows directly from (3.4.3).

We now check the measurability. By definition, $\widetilde{h}^i_1(t_i) = m_i(t_i) \circ \left(\widetilde{h}^{-i}_0\right)^{-1}$ is a composition of measurable functions, and, by Lemma 21, $\widetilde{h}^{-i}_1 = \left(\widetilde{h}^0_1, \widetilde{h}^j_1\right)$ is measurable. Suppose, by way of induction, that $\widetilde{h}^i_l$ and $\widetilde{h}^{-i}_l$ are measurable, $i \in I$ and $l = 1, ..., k$. Consider $l = k+1$. By definition, $\widetilde{h}^i_{k+1}(t_i) = \left(\widetilde{h}^i_k(t_i), m_i(t_i) \circ \left(\widetilde{h}^{-i}_k\right)^{-1}\right)$, hence, by Lemma 21 and the induction hypothesis, $\widetilde{h}^i_{k+1}$ is measurable. A repeated application of Lemma 21 yields the measurability of both $\widetilde{h}^{-i}_{k+1} = \left(\widetilde{h}^0_{k+1}, \widetilde{h}^j_{k+1}\right)$ and $\widetilde{h}^i(t_i) = \left(\widetilde{h}^i_1(t_i), \widetilde{h}^1_2(t_i), ...\right)$. ■

We make use of this last result in the following

Proposition 12 *For every type structure* $\mathcal{T} = \langle S, (T_i)_{i\in I}, (m_i)_{i\in I}\rangle$, *the map* $\widetilde{h} = \left(Id_S, \widetilde{h}^i, \widetilde{h}^j\right)$: $S \times T_i \times T_j \to S \times \overline{H}^i \times \overline{H}^j$ *is the unique type morphism from* $\mathcal{T}$ *to* $\mathcal{T}_U = \langle S, \left(\overline{H}^i\right)_{i\in I}, (f_i)_{i\in I}\rangle$.

Proof. By Lemma 23, $\widetilde{h}^i$ is measurable, $i \in I$, hence the measurability of $\widetilde{h}$ follows from Lemma 21. We now check Condition (2) in Definition 12, i.e., for each $i \in I$,

$$f^i\left(\widetilde{h}^i(t_i)\right) = m_i(t_i) \circ \left(Id_S; \widetilde{h}^j\right)^{-1}.$$

But this follows, since, by definition, $\widetilde{h}^i(t_i) = \left(m_i(t_i) \circ \left(\widetilde{h}^{-i}_0\right)^{-1}, m_i(t_i) \circ \left(\widetilde{h}^{-i}_1\right)^{-1}, ...\right)$, and f^i is a measure theoretic isomorphism according to Proposition 11. To show the uniqueness of $\widetilde{h}$, suppose that $\phi = \left(Id_S, \phi^i, \phi^j\right)$ is a type morphism from $\mathcal{T}$ to $\mathcal{T}_U$. It needs to check that $\widetilde{h} = \phi$, i.e., for each $k \geq 0$, the $(k+1)$th orders description map $\phi^i_{k+1}(t_i)$ equals $\widetilde{h}^i_{k+1}(t_i)$, $i \in I$, $t_i \in T_i$. To this end, recall that, by definition of type morphism, $f^i\left(\phi^i(t_i)\right) = m_i(t_i) \circ \left(Id_S; \phi^j\right)^{-1}$ and, being f^i injective, we get $\phi^i(t_i) = g^i \circ m_i(t_i) \circ \left(\phi^{-i}\right)^{-1}$, where g^i is the inverse of f^i as defined in the proof of Proposition 11. Thus we can write the $(k+1)$th element of $\phi^i_{k+1}(t_i)$ as

$$\begin{aligned} g^i_{k+1} \circ m_i(t_i) \circ \left(\phi^{-i}\right)^{-1} &= m_i(t_i) \circ \left(\phi^{-i}\right)^{-1} \circ \left(\rho^i_k\right)^{-1} \\ &= m_i(t_i) \circ \left(\rho^i_k \circ \phi^{-i}\right)^{-1}, \end{aligned}$$

and we have to show that it equals the corresponding $(k+1)$th element of $\widetilde{h}^i_{k+1}(t_i)$, namely for each $k \geq 0$,

$$m_i(t_i) \circ \left(\rho^i_k \circ \phi^{-i}\right)^{-1} = m_i(t_i) \circ \left(\widetilde{h}^{-i}_k\right)^{-1}.$$

We prove by induction that $\rho^i_k \circ \phi^{-i} = \widetilde{h}^{-i}_k$, for each $k \geq 0$. This is trivially true for $k = 0$: in this case $\widetilde{h}^{-i}_0 : S \times T_j \to S$ is the projection mapping, thus for each $(s, t_j) \in S \times T_j$,

$$\begin{aligned}\left(\rho^i_0 \circ \phi^{-i}\right)(s, t_j) &= \rho^i_0\left(s, \phi^{-i}(t_j)\right)) \\ &= \widetilde{h}^{-i}_0(s, t_j) \\ &= s.\end{aligned}$$

Assume that the claim holds true for $l = 1, ..., k$. Consider $l = k+1$, and observe that $\rho^i_k \circ \phi^{-i} = \rho^i_{k,k+1} \circ \rho^i_{k+1} \circ \phi^{-i}$. Since by (3.4.3), $\widetilde{h}^{-i}_k = \rho^i_{k,k+1} \circ \widetilde{h}^{-i}_{k+1}$, it follows from the induction hypothesis that

$$\begin{aligned}\rho^i_k \circ \phi^{-i} &= \rho^i_{k,k+1} \circ \rho^i_{k+1} \circ \phi^{-i} \\ &= \rho^i_{k,k+1} \circ \widetilde{h}^{-i}_{k+1},\end{aligned}$$

thus $\rho^i_{k+1} \circ \phi^{-i} = \widetilde{h}^{-i}_{k+1}$, as required. ■

We conclude this section with two remarks concerning the relationship of the result in this section with those in [30].

Remark 4 *The way we provided an inductive construction of the hierarchy description maps is the same as that in Heifetz and Samet ([30], [31]). Neverthless, it should be noted that* $\widetilde{h}_i : T_i \to \overline{H}^i$ *is not equivalent to Heifetz and Samet's decription map, but they coincide only on the first order beliefs. To see why, recall that in [31]*

- *each* $\overline{H}^i_k$ *corresponds to the kth order space of weakly coherent belief hierarchies - the case* $\overline{H}^i_1$ *is the same for both kinds of hierarchies*

- $\rho^i_{k-1,k}$ *corresponds to the natural projection between* Θ^i_k *and* Θ^i_{k-1}*, where each* Θ^i_k *equals* $\Theta^i_0 \times \overline{H}^j_k$ *- hence only weakly coherent hierarchies are considered by the players.*

In this last case, the map $\widetilde{h}_i : T_i \to \overline{H}^i$ *(which must satisfy equation (3.4.3) to be well defined) describes weakly - not necessarily strongly - coherent hierarchies of beliefs. Viewed from this angle, the map* $\widetilde{h}_i$ *as defined in the current paper can be thought of as a "stronger" version of the hierarchy descrition map in [31].*

Remark 5 *Two important properties of type morphisms and description maps are the following: (P1) Type morphims preserve description maps; and (P2) the map* $\widetilde{h}_i : \overline{H}^i \to \overline{H}^i$ *is the identity. Specifically, both properties state that type morphisms not only preserve the implicit descrition of types (Definition 4.(2)) but also their explicit description, thus providing a link between the two prevalent approaches to model interactive uncertainty in games. The proofs of P1 and P2 are exactly the same as in [30] (Proposition 5.1 and Lemma 5.4, respectively), and they are crucial to show the existence of a universal type structure in Heifetz and Samet's framework. In constrast, our proof of Proposition 12 relies on the fact that the map* f^i *is a measure theoretic isomorphism, which is not proved in [30].*

3.4.3 Proof of the main result

Theorem 5 follows as a corollary of the results presented in Sections 3.4.1-3.4.2.

Proof of Theorem 5. The existence of $\overline{H}^i$ follows from Proposition 10, and f^i is a measure-theoretic isomorphim by Proposition 11. $\mathcal{T}_U$ is universal by Proposition 12, and unique according to Proposition 9. ■

3.5 Comparison to the topological framework

In this section, we relate the results of Section 3.4.1 to the topological construction of a type structure consisting of all belief hierarchies displaying *weak* coherence and common certainty of *weak* coherence. For that, we adopt the framework developed in Chapter I, which provides a generalization of earlier works ([17],[29],[43],[44]).

We find it to convenient to prepare the comparison by establishing some general results about the topological structure of the relevant domains of uncertainty.

For any non-empty topological space X, let $\mathcal{B}(X)$ denote its Borel σ-field. A *Radon probability measure* on X is a Borel probability measure μ such that for every $A \in \mathcal{B}(X)$ and every

$\epsilon > 0$, there exists a compact set $K \subseteq A$ such that $\mu(A \backslash K) < \epsilon$. Denote by $\Delta_R(X)$ the space of all Radon probability measures on $\mathcal{B}(X)$.

The space $\Delta_R(X)$ is endowed with the *narrow topology*. This is the weakest topology for which all the maps $\mu \to \int_X f d\mu$ are lower (resp. upper) semi-continuous, as f varies in the set of all bounded, lower (resp. upper) semi-continuous functions on X. As a subbasic system of neighborhoods, the narrow topology assigns to each $\mu_0 \in \Delta_R(X)$ the sets of the form

$$V(\mu_0; \epsilon) = \{\mu \in \Delta(X) \,|\, \mu(O) \geq \mu_0(O) - \epsilon,\, O \subseteq X \text{ open},\, \epsilon > 0\}.$$

It is known (see [57], Theorem 11.2) that if X is Hausdorff, so is $\Delta_R(X)$. Specifically, if X is completely regular (e.g., compact Hausdorff or Polish), then the narrow topology on $\Delta_R(X)$ coincides with the usual weak* topology ([57], Theorem 8.1).[9] Denote $\mathcal{B}_{\Delta_R}$ by the Borel σ-field generated by the narrow topology on $\Delta_R(X)$. The following lemma singles out a basic fact which relates the topological structure of belief hierarchies to the measurable structure Σ_Δ adopted in the previous sections.[10]

Lemma 24 *If X is a Hausdorff topological space, then $\Sigma_{\Delta_R} = \mathcal{B}_{\Delta_R}$.*

Proof. Denote by Ξ the family of all closed subsets of X. By Lemma 22, sets of form

$$b^p(E) = \{\mu \in \Delta(X) : \mu(E) \geq p\},\, E \in \Xi,\, p \in [0,1], \tag{3.5.1}$$

generates the σ-field Σ_{Δ_R}. But sets as in (3.5.1) constitute a system of generators for $\mathcal{B}_{\Delta_R}$. Indeed, if E is closed subset of X, then the characteristic function $\mathbf{1}_E$ is upper semicontinuous. By definition of narrow topology, the map $\mu \to \int_X \mathbf{1}_E d\mu = \mu(E)$ is upper semi-continuous,

[9] For any non-empty, topological space X, the *weak* topology* on $\Delta(X)$ is the coarsest topology for which the map $\mu \to \int_X \phi d\mu$ from $\Delta(X)$ into $\mathbb{R}$ is continuous, as ϕ varies in the set of all bounded, continuous functions on X. By definition, the weak* topology is coarser than the narrow topology on $\Delta(X)$. The narrow topology was first introduced by Topsoe [57] under the name "weak topology". The terminology adopted in this paper is that of Schwartz [54]. In the language of probability theory, the term "weak topology" usually refers to the weak* topology as defined in the current paper.

[10] Theorem 17.24 in [34] presents a related result for *all* Borel probability measures on a *separable metric* space X. This Theorem coincides with Lemma 5 if X is assumed to be Souslin, as every Borel probability measure on X is Radon (Theorem 7.4.3 in [12]). Note that Theorem 17.24 in [34] employs the weak* topology on the space of probability measures, which is equivalent to the narrow topology since X is completely regular ([57], Theorem 8.1).

hence, $\left\{\mu \in \Delta(X) : \mu(E) = \int_X \mathbf{1}_E d\mu \geq p\right\}$ is closed in $\Delta_R(X)$ for each $p \in [0,1]$. The Borel σ-field $\mathcal{B}_{\Delta_R}$ is exactly the smallest σ-field which contains the closed sets as in (3.5.1). Thus $\Sigma_{\Delta_R} = \mathcal{B}_{\Delta_R}$. ■

We present now a special version of the measure theoretic setup in Section 3.3. The restrictions we impose on the parameter space S and player's beliefs are topological. Formally:

Assumption 1 *S is a Hausdorff topological space.*

Assumption 2 *Players' hierarchical beliefs are described by Radon probability measures.*

In light of Assumption 2, as we consider the set of *all* beliefs over a domain of uncertainty X_n^i, we set $\Delta_R\left(X_n^i\right) = \Delta\left(X_n^i\right)$ as in Section 3.3.1.

For each $i \in I$, let $\left\{X_n^i\right\}_{n=0}^{\infty}$ be the sequence of spaces defined inductively in Section 3.3.1. Here, each product space is endowed with the product topology, hence each domain of uncertainty $X_n^i = X_{n-1}^j \times \Delta\left(X_{n-1}^j\right)$ is Hausdorff. A hierarchy of beliefs $h^i = \left(\mu_1^i, \mu_2^i, ...\right)$ is an element of the Hausdorff space $H^i = \prod_{n=0}^{\infty} \Delta\left(X_n^i\right)$.

The role of Assumption 1 is to guarantee, by Lemma 24, that the Borel structure of each domain of uncertainty satisfies the requirements imposed on the players' language as in the measure theoretic framework. The following result states that, if Assumption 2 holds, weak coherence is equivalent to strong coherence for hierarchies of beliefs, so there is no need to distinguish between the two concepts.

Proposition 13 *Let Assumption 2 hold. Then a belief hierarchy $h^i \in H^i$ is strongly coherent if and only if it is weakly coherent.*

Proof. A strongly coherent hierarchy is weakly coherent by definition. Conversely, fix a weakly coherent hierarchy $h^i = \left(\mu_1^i, \mu_2^i, ...\right) \in H^i$, and let F_{n-1} be an arbitrary subset of X_{n-1}^i. By Theorem II.9 in [50], there exists a measurable set $E_{n-1} \subseteq X_{n-1}^i$ such that $F_{n-1} \subseteq E_{n-1}$ and $\left(\mu_n^i\right)^*(F_{n-1}) = \mu_n^i(E_{n-1})$. Since μ_n^i is a Radon measure, for each $\epsilon > 0$ there exists a compact set $K \subseteq E_{n-1}$ such that $\mu_n^i(E_{n-1} \setminus K) < \epsilon$, that is, $\mu_n^i(E_{n-1}) < \mu_n^i(K) + \epsilon$. Observe that $(F_{n-1} \setminus K) \cup K \subseteq E_{n-1}$, hence

$$(\mu_n^i)^*(F_{n-1} \setminus K) \leq \mu_n^i(E_{n-1}) - \mu_n^i(K) < \epsilon,$$

which implies

$$(\mu_n^i)^*(F_{n-1}) < \mu_n^i(K) + \epsilon. \tag{3.5.2}$$

Observe also that $(K \setminus F_{n-1}) \cup F_{n-1} \subseteq E_{n-1}$, hence

$$(\mu_n^i)^*(K \setminus F_{n-1}) + (\mu_n^i)^*(F_{n-1}) \leq \mu_n^i(E_{n-1}),$$

which implies

$$(\mu_n^i)^*(K \setminus F_{n-1}) = 0. \tag{3.5.3}$$

According to Proposition 8, (3.5.3) yields $(\mu_{n+1}^i)^* \circ \left(\mathrm{Pr}_{n-1,n}^i\right)^{-1}(K \setminus F_{n-1}) = 0$. Rewrite K as a union of disjoint sets: $K = (K \setminus F_{n-1}) \cup (F_{n-1} \cap K)$. We get

$$\begin{aligned}\mu_{n+1}^i \circ \left(\mathrm{Pr}_{n-1,n}^i\right)^{-1}(K) &= (\mu_{n+1}^i)^* \circ \left(\mathrm{Pr}_{n-1,n}^i\right)^{-1}(K \setminus F_{n-1}) \\ &\quad + (\mu_{n+1}^i)^* \circ \left(\mathrm{Pr}_{n-1,n}^i\right)^{-1}(F_{n-1} \cap K),\end{aligned}$$

so by weak coherence on K and by (3.5.3),

$$\begin{aligned}\mu_n^i(K) &= (\mu_n^i)^*(F_{n-1} \cap K) \\ &= (\mu_{n+1}^i)^* \circ \left(\mathrm{Pr}_{n-1,n}^i\right)^{-1}(F_{n-1} \cap K).\end{aligned}$$

Using the fact that $\left(\mathrm{Pr}_{n-1,n}^i\right)^{-1}(F_{n-1} \cap K) \subseteq \left(\mathrm{Pr}_{n-1,n}^i\right)^{-1}(F_{n-1})$, we obtain

$$\begin{aligned}\mu_n^i(K) &\leq (\mu_{n+1}^i)^* \circ \left(\mathrm{Pr}_{n-1,n}^i\right)^{-1}(F_{n-1}) \\ &\leq (\mu_n^i)^*(F_{n-1}) \\ &< \mu_n^i(K) + \epsilon,\end{aligned}$$

where the second and third inequalities follow from Proposition 8 and from (3.5.2), respectively. By the arbitrariness of $\epsilon > 0$, we conclude that

$$(\mu_{n+1}^i)^* \circ \left(\mathrm{Pr}_{n-1,n}^i\right)^{-1} (F_{n-1}) = (\mu_n^i)^* (F_{n-1}),$$

thereby establishing the strong coherence for $h^i \in H^i$. ■

The next proposition turns out to be a special case of Theorem 5. As immediate consequence, the subsequent corollary derives two well known results on universal type spaces: the first is due to Mertens and Zamir ([44], the case of compact parameter space), while the other is due to Brandenburger and Dekel ([17], the case of S being Polish).[11]

Proposition 14 *Under Assumptions 1 and 2, there exists, for each $i \in I$, a non-emtpy set $\overline{H}^i \subseteq H^i$ which is homeomorphic to $\Delta\left(S \times \overline{H}^j\right)$ under the map f^i as in Theorem 5. The space $\overline{H}^i \times \overline{H}^j$ is the set of all paires of belief hierarchies satisfying weak coherence and common certainty of weak coherence.*

Proof. Using Lemma 24 and Proposition 13, the proof is the same as in Section 3.4.1. Additionally, f^i is a bijective open map (hence a homeomorphism) according to Lemma 5 in Chapter I. ■

Corollary 3 *If S is compact Hausdorff or Polish, so are $\overline{H}^i$ and $\Delta\left(S \times \overline{H}^j\right)$, $i \in I$. Furthermore, if S is Polish, Assumption 2 is automatically satisfied.*

Proof. S is compact if and only if $\Delta(S)$ is compact (see the notes to §11 in Topsoe [57], p. 76). Inductively, each X_n^i is compact Hausdorff, and, since the projective limit is closed in the product topology, $\overline{H}^i$ is compact Hausdorff as well. The case of S being a Polish space is analogous, as $\Delta(S)$ is also Polish ([2], Theorem 15.15). In this last case, observe that each Borel probability measure on a Polish space is Radon ([12], Theorem 7.4.3), so Assumption 2 is automatically satisfied. ■

[11] It is worth noting that Assumption 2 is not stated explicitly in [44], where all beliefs are represented by general Borel probability measures. This is evidently a mistake: see Chapter 1, Section 1.5.1, for a discussion on why Assumption 2 cannot be omitted even in case of compact parameter spaces.

3.6 Discussion

3.6.1 Strong coherence and common certainty of strong coherence

The modelling strategy which was used in Section 3.4.1 to obtain the structure $\mathcal{T}_U$ is complementary to the one following the approch of Brandenburger and Dekel ([17]) to weak coherence. This alternative strategy constructs the space of all possible belief hierarchies (i.e., the space H^i in Section 3.3.1) and then it imposes strong coherence and common certainty of strong coherence. By contrast, in our construction scheme strong coherence is imposed at any level of the hierarchy, by the introduction of the domains of uncertainties Θ_n^i. In a topological framework, Proposition 2 in Chapter I shows that both approaches are equivalent - they yield two homeomorphic type structures. With some necessary minor changes, it is easy to show that this result also holds in the measure-theoretic framework (i.e., the two structures are measure-theoretic isomorphic). This alternative strategy has the advantage of explicitly formalizing a well defined notion of common certainty of (strong) coherence for belief hierarchies; however, it has the basic drawback that a formal proof for the existence of the relevant set of belief hierarchies - namely, $\varprojlim \overline{H}_n^i$ is non-empty - cannot be provided, unless some special topological assumptions are made (see Chapter I for details).

3.6.2 Topological type structures and universality

How general is the topological framework? We have seen (Proposition 13) that, from the point of view of representing all strongly coherent hierarchies, the topological setup does not impose too much limitative restrictions on the parameter space S and belief hierarchies. For instance, finite sets of nature states are commonly used in applications, hence, by endowing such parameter spaces with the discrete topology, Assumptions 1 and 2 are automatically satisfied. This may be comforting because while it is problematic to assume some particular structure on the parameter space (as it is nothing more than an artificial modelling construct), there may be sometimes good reasons to assume structure on the physical world S.

Neverthless, a potential concern with the use of topological type structures relates to the issue of universality. Typically, the topological setup imposes the following assumptions on type structures and type morphisms (cf. Definitions 12 and 13):

- for a given type structure $\mathcal{T} = \langle S, (T_i)_{i \in I}, (m_i)_{i \in I} \rangle$, the type map $m_i : T_i \to \Delta(S \times T_j)$ is continuous and $\Delta(S \times T_j)$ is endowed with the narrow topology;

- every type morphism is a continuous function.

With these assumptions, the class of S-based topological type structures form a category (cf. Remark 2), which we denote $\mathcal{C}_{NT}^{S}$, whose terminal object corresponds to the topological type structure constructed in Section 3.5. Indeed, going through the proof of the universality property of $\mathcal{T}_U$ in Section 3.4.2, it is easy to check that the same proof, with measurability conditions replaced by continuity, proves that the type structure established in Proposition 14 is universal. By contrast, a very recent paper by Pintér [48] shows that the type structure $\mathcal{T}_U$, if topologized, cannot be the terminal object in the category of topological type structures. This negative result does not contradict our positive one, rather it deepens our understanding on how limitative the topological restrictions on type structures are.

The category $\mathcal{C}_{NT}^{S}$ contains type structures in which the spaces of probability measures are endowed with the narrow topology. By virtue of Lemma 24, the narrow topology gives rise to a Borel measurable structure which reflects the basic sentences of players' language (encoded by the σ-field Σ_{Δ}) that a type structure should capture, i.e., events like "player i believes with probability at least p that an event occurs".[12] But, as is shown in [48], the narrow topology is not the weakest topology whose induced Borel structure coincides with Σ_{Δ} - in general, there is no weakest topology among all the topologies satisfying this property. As such, Pintér's result is stated for the category $\mathcal{C}_{\tau^*}^{S}$ of topological type structures, where the subscript τ^* stands for any topology on the spaces of probability measures whose Borel σ-fields are compatible with Σ_{Δ}. Clearly, the category $\mathcal{C}_{\tau^*}^{S}$ is broader than $\mathcal{C}_{NT}^{S}$, hence the existence result of a universal topological type structure given here must be considered as an exception, rather than the rule. Thus, although our result is not at odds with the non-existence result in [48], we leave open the question as to whether restricting attention to topological type structures in $\mathcal{C}_{NT}^{S}$ may affect game theoretic analisys.

[12] Such statements can be formalized by means ofthe so-called "belief operators", introduced by Monderer and Samet ([46]) as a general version of the standard knowledge operators (cf. [4]).

3.6.3 Weak coherence and Universal Belief Spaces

Heifetz and Samet [31] consider weak coherence as the natural, minimal requirement that has to be imposed on belief hierarchies to obtain a well defined notion of "state of the world". Specifically, denoting by C^i the (projective limit) set of player i's weakly coherent belief hierarchies, Heifetz and Samet define the space of all states of worlds - the *Universal Belief Space* - by the product $S \times C^i \times C^j$. Other authors ([29], [44], [43]) also define Universal Belief Spaces in a similar fashion. However, weakly coherent hierarchies may fail to describe types, as it is shown in [31], and confirmed by our main result. In light of this, Heifetz and Samet conclude that "the Harsanyi implicit approach does not exhaust all the states of affairs that can be described explicitly in the general measure-theoretic case"([31], p. 476).

In our view, such a definition of Universal Belief Space should be firmly reformulated. In fact, we propose to rely on strong coherence to obtain an acceptable notion of state of the world, so that Universal belief space should be expressed by the product $S \times \overline{H}^i \times \overline{H}^j$. The reason behind this can be elucitated by looking closely at the hierarchy of beliefs constructed in [31], which is, up to now, the only instance of a weakly, but not strongly, coherent belief hierarchy that we know.

Technically, the hierarchy in Heifetz and Samet's paper ([31]) is a refinement of the Anderssen and Jessen's ([3]) construction of a weak projective sequence of probability spaces having no limit.[13] According to our results, it is exactly the lack of the strong coherence property that causes the failure of this hierarchy to represent a possible type. Informally, an event reflects the players' language, thus a minimal requirement that has to be imposed in order to construct a belief hierarchy as a type is that any higher order domain of uncertainty should contain at least as many events as those in any lower order domain. However, if strong coherence is precluded, going through the levels of the hierarchy, one always finds a strictly larger space of events, and

[13] The construction of the belief hierarchy in [31] is quite involved. Here, it is noteworthy that the projective sequence (μ_n), where each μ_n is defined as the product measure

$$\mu_n = \prod_{1 \leq k < n} v_{0,k} \times \prod_{0 \leq k < \infty} v_{k,k+n}$$

(see Heifetz and Samet's paper for the relevant notation and definitions), is weak but not strong, as $(v_{m,n})_{n \geq m}$ corresponds to the sequence (v_n) described in Example 1 in the Appendix. Consequently, the hierarchy $(\kappa_n)_{n \geq 1}$ in [31], where $\kappa_n = \mu_n \circ (g_i^n)^{-1}$ and g_i^n is an appropriate embedding, is weakly but not strongly coherent, and has no σ-additive extension.

hence this enlargment process would never reach an end. The example in [31] reflects a situation for which the players' ability to talk about things becomes more and more powerful when the levels of the hierarchy increase, and this precludes the possibility of describing exhaustively the epistemic attitudes of the players.

As argued by Aumann [4], a model which is not common certain among the players cannot provide a complete representation of all possible states of the world, as a state, by its definition, should include all the relevants aspects that are object of uncertainty. In this sense, the hierarchy described by Heifetz and Samet [31] cannot be part of a commonly certain model, and strong coherence is the "right" requirement that belief hierarchies must satisfy to provide a full description of any incomplete information scenario.

3.7 Appendix

3.7.1 Projective sequences of probability spaces: definitions and results

For the convenience of the reader and to fix some notation that can vary from author to author, we provide some of the background definitions and results from the theory of projective systems of measure spaces. Here we restrict attention to projective sequences, i.e., the index set is $\mathbb{N}$. For a more thorough treatment see [22] or [50].

Definition 15 *A projective sequence of spaces is a structure* $\{Y_n, f_{m,n}\}_{m,n\in\mathbb{N}}$ *of spaces* Y_n *and functions* $f_{m,n} : Y_n \to Y_m$ *such that:*

- *for each* $n \in \mathbb{N}$, Y_n *is non-empty;*
- $f_{m,p} = f_{m,n} \circ f_{n,p}$ *for any* $m, n, p \in \mathbb{N}$ *satisfying* $m \leq n \leq p$, *and* $f_{n,n} = Id_{Y_n}$ *for every* $n \in \mathbb{N}$.

The spaces Y_n *are called coordinate (or factor) spaces and the maps* $f_{m,n}$ *are called bonding maps.*

Definition 16 *Let* $\{Y_n, f_{m,n}\}_{m,n\in\mathbb{N}}$ *be a projective sequence of spaces. The set*

$$\varprojlim Y_n = \left\{ \{y_n\} \in \prod_{n\in\mathbb{N}} Y_n \,|y_m = f_{m,n}(y_n)\,,\ \textit{for each } m,n \in \mathbb{N} \textit{ s.t. } m \leq n \right\}.,$$

is called the projective limit of $\{Y_n, f_{m,n}\}_{m,n\in\mathbb{N}}$. *For each* $l \in \mathbb{N}$, *the map* $\overline{f}_l : \varprojlim Y_n \to Y_l$ *is the restriction of the projection map* $Pr_l : \prod_{n\in\mathbb{N}} Y_n \to Y_l$ *to* $\varprojlim Y_n$.

Clearly, for any $m, n \in \mathbb{N}$ such that $m \leq n$, the maps $\overline{f}_n$ and $\overline{f}_m$ satisfy the equality $\overline{f}_m = f_{m,n} \circ \overline{f}_n$. Specifically, if each bonding map $f_{m,n}$ is the coordinate projection, then the projective limit space $\varprojlim Y_n$ can be identified (isomorphically) with the Cartesian product $\prod_{n\in\mathbb{N}} Y_n$.

For each $n \in \mathbb{N}$, let μ_n be a probability measure on the measure space (Y_n, Σ_n).

Definition 17 *The structure* $\{(Y_n, \Sigma_n, \mu_n), f_{m,n}\}_{m,n\in\mathbb{N}}$ *is called* weak projective sequence of probability spaces *if*

(1) *For all* $n \in \mathbb{N}$, (Y_n, Σ_n, μ_n) *is a probability space;*

(2) $\{(Y_n, \Sigma_n), f_{m,n}\}_{m,n\in\mathbb{N}}$ *is a projective sequence of measurable spaces* (Y_n, Σ_n) *and* $f_{m,n}$ *is* Σ_m*-measurable;*

(3) $\mu_m(A) = \mu_n\left(f_{m,n}^{-1}(A)\right)$, *for all* $m, n \in \mathbb{N}$ *such that* $m \leq n$, $A \in \Sigma_m$.

The structure $\{(Y_n, \Sigma_n, \mu_n), f_{m,n}\}_{m,n\in\mathbb{N}}$ *is called* strong projective sequence of probability spaces *if condition (3) is replaced by*

(3)' $\mu_m^*(A) = \mu_n^*\left(f_{m,n}^{-1}(A)\right)$, for all $m, n \in \mathbb{N}$ such that $m \leq n$, $A \subseteq Y_m$.

Thus, the notion of "strongness" for projective sequence of probability spaces makes no assumptions about the existence of a topology, but instead requires a special connection between the μ_ns. The following is an example of a weak projective sequence of probability spaces which is not strong. It is based on [12], Example 7.7.3.

Example 1 *Let* (Y, Σ, λ) *be a probability space, where* Y *is a separable metric space and* λ *is an atomless probability measure on the Borel* σ*-field* Σ. *There exists a sequence of non-measurable sets* $Y_n \subseteq Y$ *such that* $Y_{n+1} \subseteq Y_n$, $\lambda^*(Y_n) = 1$, *and* $\cap_{n\geq 1} Y_n = \emptyset$ *(cf. Exercise 9.12.85 in [12]). The Borel* σ*-field* Σ_n *on each* Y_n *is defined by* $\Sigma_n = Y_n \cap \Sigma$. *Let* λ_n *be the restriction of* λ *on* Σ_n, *i.e.,* $\lambda_n(E) = \lambda^*(E)$ *for each* $E \in \Sigma_n$. *By Theorem 3.3.6 in [21], the measure* λ_n *is* σ*-additive, so that* $(Y_n, \Sigma_n, \lambda_n)$ *is a well defined probability space, each* Y_n *being a separable metric space.*

For each $n \geq 1$, define the map $\gamma_n : Y_n \to \prod_{l=1}^{n} Y_l$ by

$$\gamma_n(y) = \underbrace{(y, ..., y)}_{n \text{ times}}$$

and let μ_n be a probability measure on the product σ-field $\bigotimes_{l=1}^{n} \Sigma_l$ such that $v_n = \lambda_n \circ \gamma_n^{-1}$. Thus, v_n is concentrated on the diadonal

$$D_n = \left\{ (y_1, ..., y_n) \in \prod_{l=1}^{n} Y_l \,| y_1 = y_2 = ... = y_n \right\},$$

i.e., $v_n(D_n) = 1$(observe that $D_n \in \bigotimes_{l=1}^{n} \Sigma_l$ as each Y_n is a separable metric space). The structure

$$\left\{ \left(\prod_{l=1}^{n} Y_l, \bigotimes_{l=1}^{n} \Sigma_l, v_n \right), f_{m,n} \right\}_{m,n \in \mathbb{N}}$$

is a weak projective sequence of probability spaces, where $f_{m,n} : \prod_{l=1}^{n} Y_l \to \prod_{l=1}^{m} Y_l$ stands for the coordinate projection. Indeed, for each $A \in \bigotimes_{l=1}^{m} \Sigma_l$,

$$\begin{aligned} v_n\left(f_{m,n}^{-1}(A)\right) &= \lambda_n \circ \gamma_n^{-1} \circ f_{m,n}^{-1}(A) \\ &= \lambda_n \circ (f_{m,n} \circ \gamma_n)^{-1}(A) \\ &= \lambda_n \circ \gamma_m^{-1}(A) \\ &= v_m(A), \end{aligned}$$

as required by Condition (3) in Definition 17. This projective sequence is not strong. To see this, consider the set $B \subseteq Y_1$ with $B = Y_1 \setminus Y_2$. Clearly, $B \notin \Sigma_1$ but $v_1^*(B) = 1$.[14] *Note that $f_{1,2}^{-1}(B) = (Y_1 \setminus Y_2) \times Y_2$, and $D_2 \cap f_{1,2}^{-1}(B) = \emptyset$, i.e., $f_{1,2}^{-1}(B) \subseteq (Y_1 \times Y_2) \setminus D_2$ with $v_2((Y_1 \times Y_2) \setminus D_2) = 0$. This implies*

$$0 = v_2^*\left(f_{1,2}^{-1}(B)\right) < v_1^*(B) = 1.$$

---

[14] *See Theorem 3.3.4. in Dudley [21]. Observe that this theorem is stated for the measure space $([0,1], \mathcal{B}([0,1]), v)$, where v stands for the Lebesgue measure on the Borel σ-field $\mathcal{B}([0,1])$, but it holds good in arbitrary separable metric spaces - see again Exercise 9.12.85 in [12].*

Definition 18 *The probability space* (Y, Σ, μ) *is called* measure projective limit *of the weak projective sequence of probability spaces* $\{(Y_n, \Sigma_n, \mu_n), f_{m,n}\}_{m,n\in\mathbb{N}}$ *if*

- $Y = \varprojlim Y_n \neq \emptyset$,
- $\Sigma = \sigma\left(\cup_{n\in\mathbb{N}} \overline{f}_n^{-1}(\Sigma_n)\right)$, *i.e.* Σ *is the coarsest* σ*-field for which each projection map* Pr_n *is measurable,*
- μ *is a probability measure on* Σ *such that*

$$\mu\left(\overline{f}_n^{-1}(A)\right) = \mu_n(A) \text{ for all } A \in \Sigma_n,\ n \in \mathbb{N}.$$

Additionally, if $\{(Y_n, \Sigma_n, \mu_n), f_{m,n}\}_{m,n\in\mathbb{N}}$ *is a strong projective sequence, then the probability measure* μ *satisfies*

$$\mu^*\left(\overline{f}_n^{-1}(A)\right) = \mu_n^*(A) \text{ for all } A \subseteq Y_n,\ n \in \mathbb{N}.$$

Here is the formal statement of the extension result established by Pinter ([47], Theorem 3.2.) (slightly reformulated here concerning notation and terminology). It is used in the proof of Lemma 20.

Theorem 6 *Let* $\{(Y_n, \Sigma_n, \mu_n), f_{m,n}\}_{m,n\in\mathbb{N}}$ *be a strong projective sequence of probability spaces. Then the measure projective limit* (Y, Σ, μ) *exists and is unique.*

It is noteworthy that Theorem 6 provides only a sufficient condition for the existence and uniqueness of a measure projective limit. For this, see the example in [47]. Finally, observe that the weak projective sequence of probability spaces in Example 1 has no limit (cf. [12], Example 7.7.3.).

The proof of Theorem 6 reduces to an adaptation of the one in [47], which we report here for the sake of completeness.

Proof. Let μ be the unique additive set function on the field $\mathcal{A} = \cup_{n\in\mathbb{N}} \overline{f}_n^{-1}(\Sigma_n)$ satisfying $\mu\left(\overline{f}_n^{-1}(A)\right) = \mu_n(A)$ for all $A \in \Sigma_n$, $n \in \mathbb{N}$. (Such a set function always exists, see e.g. [50], p.417) We need to show that μ is σ-additivite on $\mathcal{A}$, so that it can be extended uniquely on $\sigma(\mathcal{A})$ by the standard Charatheodory procedure. To this end, it suffices to prove that if $\{B_k\}_{k\in\mathbb{N}}$ is a decreasing sequence of sets in $\mathcal{A}$ such that $B_k \downarrow \emptyset$, then $\lim_{k\to\infty} \mu(B_k) = 0$. (see Exercise I.7 in [50].)

Since each B_k belong to $\mathcal{A}$, there is $k_n \in \mathbb{N}$ such that $B_k = \overline{f}_{k_n}^{-1}(A_{k_n})$ for some $A_{k_n} \in \Sigma_{k_n}$, and $\cap_{k_n\in\mathbb{N}} \overline{f}_{k_n}^{-1}(A_{k_n}) = \emptyset$. Thus we can consider, without loss of generality, a sequence $\{B_n\}_{n\in\mathbb{N}}$ such that for all $n \in \mathbb{N}$,

- $B_n = \overline{f}_n^{-1}(A_n)$ for some $A_n \in \Sigma_n$,
- $\overline{f}_n^{-1}(A_n) \supseteq \overline{f}_{n+1}^{-1}(A_{n+1})$,
- $\cap_{n\in\mathbb{N}} \overline{f}_n^{-1}(A_n) = \emptyset$.

We show that

$$\lim_{n\to\infty} \mu(B_n) = \lim_{n\to\infty} \mu\left(\overline{f}_n^{-1}(A_n)\right) = \lim_{n\to\infty} \mu_n(A_n) = 0.$$

For that, we need the following three claims.

Claim 4 *$f_{n,n+1}^{-1}(A_n) \supseteq A_{n+1}$, for all $n \in \mathbb{N}$.*

Proof. *Recall that $\overline{f}_n = f_{n,n+1} \circ \overline{f}_{n+1}$ and $\overline{f}_n^{-1}(A_n) \supseteq \overline{f}_{n+1}^{-1}(A_{n+1})$, so*

$$\begin{aligned}\left(f_{n,n+1} \circ \overline{f}_{n+1}\right)^{-1}(A_n) &= \overline{f}_{n+1}^{-1}\left(f_{n,n+1}^{-1}(A_n)\right) \\ &\supseteq \overline{f}_{n+1}^{-1}(A_{n+1}).\end{aligned}$$

Since $\overline{f}_{n+1}$ is surjective, $\overline{f}_{n+1} \circ \overline{f}_{n+1}^{-1}$ turns out to be the identity map, thus

$$f_{n,n+1}^{-1}(A_n) \supseteq A_{n+1},$$

as required. ■

Now let $\{L_n\}_{n\in\mathbb{N}}$ be a sequence of subsets of A_0 defined as follows:

$$L_n = \left\{ y \in A_0 \,\middle|\, f_{0,n}^{-1}(y) \subseteq \left(f_{0,n}^{-1}(A_0)\right) \diagdown A_n \right\}.$$

Note that each L_n is not necessarily measurable.

Claim 5 *$L_n \subseteq L_{n+1}$, for all $n \in \mathbb{N}$, and $L_n \uparrow A_0$.*

Proof. *Pick $y \in L_n$. By definition of L_n, $f_{0,n}^{-1}(y)$ is not a subset of A_n. By Claim 1 $f_{n,n+1}^{-1}(A_n) \supseteq A_{n+1}$, so*

$$\begin{aligned} f_{0,n+1}^{-1}(y) &= (f_{0,n} \circ f_{n,n+1})^{-1}(y) \\ &= f_{n,n+1}^{-1}\left(f_{0,n}^{-1}(y)\right) \end{aligned}$$

is not a subset of A_{n+1}, that is, $y \in L_{n+1}$. Thus $L_n \uparrow A_0$, i.e., $\cup_{n\in\mathbb{N}} L_n = A_0$. ■

Claim 6 *$f_{0,n}^{-1}(L_n) \subseteq \left(f_{0,n}^{-1}(A_0)\right) \setminus A_n$, for all $n \in \mathbb{N}$.*

Proof. *If $y \in L_n$, then $f_{0,n}^{-1}(y) \subseteq \left(f_{0,n}^{-1}(A_0)\right) \setminus A_n$, hence $f_{0,n}^{-1}(L_n) \subseteq \left(f_{0,n}^{-1}(A_0)\right) \setminus A_n$.* ■

From Claim 6, it follows that

$$\begin{aligned} \mu_n^*\left(f_{0,n}^{-1}(L_n)\right) &\leq \mu_n^*\left(\left(f_{0,n}^{-1}(A_0)\right) \setminus A_n\right) \\ &= \mu_n\left(\left(f_{0,n}^{-1}(A_0)\right) \setminus A_n\right) \\ &\leq \mu_n\left(f_{0,n}^{-1}(A_0)\right) - \mu_n(A_n) \\ &= \mu_0(A_0) - \mu_n(A_n), \end{aligned} \tag{$\bigstar$}$$

where the second equality follows from the measurability of the set $\left(f_{0,n}^{-1}(A_0)\right) \setminus A_n$, and the fourth equality follows from the definition of weak projective sequence.

Since the projective sequence is strong, from ($\bigstar$) we get

$$\mu_0^*(L_n) \leq \mu_0(A_0) - \mu_n(A_n). \tag{$\bigstar\bigstar$}$$

Moreover, by Claim 5, $L_n \uparrow A_0$, thus

$$\mu_0^*(\cup_{n\in\mathbb{N}} L_n) = \lim_{n\to\infty} \mu_0^*(L_n) = \mu_0(A_0). \tag{$\bigstar\bigstar\bigstar$}$$

It follows from ($\bigstar\bigstar$) and ($\bigstar\bigstar\bigstar$) that

$$\mu_0(A_0) \leq \mu_0(A_0) - \lim_{n\to\infty} \mu_n(A_n),$$

which is satisfied if and only if

$$\begin{aligned}\lim_{n\to\infty}\mu_n(A_n) &= \lim_{n\to\infty}\mu\left(\overline{f}_n^{-1}(A_n)\right)\\ &= \lim_{n\to\infty}\mu(B_n)=0.\end{aligned}$$

This proves the σ-additivity of μ. ■

3.7.2 Proof of the measurability of $\overline{H}^i$

Here we show that $\overline{H}^i$ is a measurable subset of H^i. To this end, it is enough to show that $\underleftarrow{\lim}\overline{H}^i_n$ is a measurable subset of $\prod_{n=1}^{\infty}\overline{H}^i_n$. The result then will follow from the measure-theoretic isomorphism between $\underleftarrow{\lim}\overline{H}^i_n$ and $\overline{H}^i$.

We need the following auxiliary definition.

Definition 19 *The measurable space (X,Σ_X), or the σ-field Σ_X, is σ-separative if there exists a countable family $\mathcal{A}\subseteq\Sigma_X$ such that for every $x,y\in X$ with $x\neq y$ there exists $B\in\mathcal{A}$ with $x\in B$ and $y\notin B$.*

σ-separativity is a strengthening of the definition of *separitivity* introduced by Liu [36].[15] The following Lemma collects three properties of σ-separativity which will be used in the course of the proof.[16]

Lemma 25 *(Properties of σ-separitivity)*

1. *The measurable space $(\Delta(X),\Sigma_\Delta)$ is σ-separative.*

2. *The product of any σ-separative measurable spaces is σ-separative.*

3. *Let the function $\psi:(X,\Sigma_X)\to(Y,\Sigma_Y)$ be measurable. If (Y,Σ_Y) is σ-separative, then the graph*

$$G(\psi)=\{(x,y)\in X\times Y\,|y=\psi(x)\}$$

belongs to the product σ-field $\Sigma_X\otimes\Sigma_Y$.

[15] The notion of σ-separativity is also known under the name of "almost separability" in the literature of projective systems of measure spaces (e.g., [40]).

[16] Claims 1 and 2 in Lemma 25 are an easy adaptation of Lemma 1 in [36], while Claim 3 generalizes Theorem 4.45 in [2].

Proof. (1) Recall that $[0,1]^{\Sigma_X}$ denotes the space of all set functions λ on Σ such that $\lambda(X) = 1$, and is endowed with the product σ-field $\mathcal{M}$, that is, the coarsest σ-field for which all evaluation maps $e_A : [0,1]^{\Sigma_X} \to [0,1]$ defined as $e_A(\lambda) = \lambda(A)$ are measurable, for every $A \in \Sigma_X$. The Borel σ-field $\mathcal{B}([0,1])$ of the unit interval is the coarsest σ-field which contains the countable family $\mathcal{B}_0 = \{[p,1] \subseteq [0,1] : p \in \mathbb{Q}\}$, i.e., $\mathcal{B}([0,1]) = \sigma(\mathcal{B}_0)$. Let $\mathcal{M}_0 = \cup_{A\in\Sigma_X} e_A^{-1}(\sigma(\mathcal{B}_0))$. Then $\mathcal{M}_0$ is a field on $[0,1]^{\Sigma_X}$, and $\mathcal{M} = \sigma(\mathcal{M}_0)$. Note that $\Sigma_\Delta = \mathcal{M} \cap \Delta(X)$; thus, it is enough to show that $\mathcal{M}$ is σ-separative. Consider two set functions μ_1, $\mu_2 \in [0,1]^{\Sigma_X}$ such that $\mu_1 \neq \mu_2$.Thus there exists $A \in \Sigma_X$ such that $\mu_1(A) \geq p > \mu_2(A)$, for $p \in \mathbb{Q} \cap [0,1]$, hence $\mu_1 \in e_A^{-1}([p,1])$ and $\mu_2 \notin e_A^{-1}([p,1])$, where $e_A^{-1}([p,1])$ belongs to $e_A^{-1}(\sigma(\mathcal{B}_0))$, a countable subfamily of $\mathcal{M}$.

(2) Let Ω be an arbitrary index set and X_ω be σ-separative for each $\omega \in \Omega$. Let $X = \prod_{\omega\in\Omega} X_\omega$ be endowed with the product σ-field, i.e., the coarsest σ-field for which the canonical projections Pr_ω are measurable. Let $x = (x_\omega)_{\omega\in\Omega}$ and $x' = (x'_\omega)_{\omega\in\Omega}$ be two different points in X. So, there is $\omega^* \in \Omega$ such $x_{\omega^*} \neq x'_{\omega^*}$. The σ-field $\Sigma_{X_{\omega^*}}$ is σ-separative, so there exists a countable family $\mathcal{E} \subseteq \Sigma_{X_{\omega^*}}$ such that $x_{\omega^*} \in B$ and $x'_{\omega^*} \notin B$, for some $B \in \mathcal{E}$. Thus, $x \in \mathrm{Pr}_{\omega^*}^{-1}(B)$ and $x' \notin \mathrm{Pr}_{\omega^*}^{-1}(B)$, and $\mathrm{Pr}_{\omega^*}^{-1}(B)$ belongs to $\mathrm{Pr}_{\omega^*}^{-1}(\mathcal{E})$, which is a countable subfamily of the product σ-field on X.

(3) Let $\{E_n\}_{n\in\mathbb{N}}$ be a countable family in Σ_Y such that for $y_1, y_2 \in Y$ with $y_1 \neq y_2$ there exists $E \in \{E_n\}_{n\in\mathbb{N}}$ with $y_1 \in E$ and $y_2 \notin E$. Denote by $\mathbf{1}_{E_n}$ the characteristic function on each E_n. Let ς be the real valued function on $(X \times Y, \Sigma_X \otimes \Sigma_Y)$ defined by

$$\varsigma(x,y) = \sum_{n=1}^{\infty} \left|\mathbf{1}_{E_n}(\psi(x)) - \mathbf{1}_{E_n}(y)\right|.$$

Clearly, the function ς is measurable, hence $\varsigma^{-1}(0) \in \Sigma_X \otimes \Sigma_Y$. But if $(x,y) \in \varsigma^{-1}(0)$ then $\mathbf{1}_{E_n}(\psi(x)) = \mathbf{1}_{E_n}(y)$ for all $n \in \mathbb{N}$, i.e., $\psi(x) = y$. Therefore $\varsigma^{-1}(0) = G(\psi)$, so $G(\psi) \in \Sigma_X \otimes \Sigma_Y$. ■

We are ready to prove that $\underleftarrow{\lim}\overline{H}_n^i$ is a measurable subset of $\prod_{n=1}^{\infty} \overline{H}_n^i$. Observe that each $\overline{H}_n^i$ is a subset of $\prod_{l=0}^{n-1} \Delta\left(X_l^i\right)$, which is σ-separative by Lemma 25. Therefore both $\overline{H}_n^i$ and $\prod_{n=1}^{\infty} \overline{H}_n^i$ are σ-separative.

Denote by $\widehat{h}^i_n$ a generic element of $\overline{H}^i_n$, i.e., $\widehat{h}^i_n = (\mu^i_1, ..., \mu^i_{n-1}, \mu^i_n)$, and let

$$\widehat{H}^i_n = \bigcap_{k=1}^{n} \left\{ \left(\widehat{h}^i_1, ..., \widehat{h}^i_{n-1}, \widehat{h}^i_n\right) \in \prod_{l=1}^{n} \overline{H}^i_l \left| \widehat{h}^i_k = \pi^i_{k,n}\left(\widehat{h}^i_n\right) \right. \right\}.$$

By Lemma 25.3, the set

$$G\left(\pi^i_{k,n}\right) = \left\{ \left(\widehat{h}^i_k, \widehat{h}^i_n\right) \in \overline{H}^i_k \times \overline{H}^i_n \left| \widehat{h}^i_k = \pi^i_{k,n}\left(\widehat{h}^i_n\right) \right. \right\}$$

is $\Sigma_{\overline{H}^i_k} \otimes \Sigma_{\overline{H}^i_n}$-measurable, hence the set

$$G\left(\pi^i_{k,n}\right) \times \prod_{l=1}^{k-1} \overline{H}^i_l \times \prod_{l=k+1}^{n-1} \overline{H}^i_l = \left\{ \left(\widehat{h}^i_1, ..., \widehat{h}^i_{n-1}, \widehat{h}^i_n\right) \in \prod_{l=1}^{n} \overline{H}^i_l \left| \widehat{h}^i_k = \pi^i_{k,n}\left(\widehat{h}^i_n\right) \right. \right\}$$

is a cylinder in $\otimes_{l=1}^{n} \Sigma_{\overline{H}^i_l}$. It follows that also the set

$$\widehat{H}^i_n = \bigcap_{k=1}^{n} \left(G\left(\pi^i_{k,n}\right) \times \prod_{l=1}^{k-1} \overline{H}^i_l \times \prod_{l=k+1}^{n-1} \overline{H}^i_l \right)$$

belongs to $\otimes_{l=1}^{n} \Sigma_{\overline{H}^i_l}$.

Now, let $\overline{p}^i_n : \prod_{l=1}^{\infty} \overline{H}^i_n \to \prod_{l=1}^{n} \overline{H}^i_n$ be the canonical projection. (Hence, π^i_n is the restriction of $\overline{p}^i_n$ to $\underleftarrow{\lim} \overline{H}^i_n$.) It is easy to check that the sequence $\left\{ \left(\overline{p}^i_n\right)^{-1} \left(\widehat{H}^i_n\right) \right\}_{n\in\mathbb{N}}$ is decreasing, where

$$\left(\overline{p}^i_n\right)^{-1} \left(\widehat{H}^i_n\right) = \left\{ \left(\widehat{h}^i_1, \widehat{h}^i_2, ...\right) \in \prod_{l=1}^{\infty} \overline{H}^i_l \left| \widehat{h}^i_k = \pi^i_{k,n}\left(\widehat{h}^i_n\right), \ k = 1, ..., n \right. \right\}.$$

By definition,

$$\begin{aligned} \underleftarrow{\lim} \overline{H}^i_n &= \bigcap_{n=1}^{\infty} \left(\overline{p}^i_n\right)^{-1} \left(\widehat{H}^i_n\right) \\ &= \left\{ \left(\widehat{h}^i_1, \widehat{h}^i_2, ...\right) \in \prod_{l=1}^{\infty} \overline{H}^i_l \left| \widehat{h}^i_k = \pi^i_{k,n}\left(\widehat{h}^i_n\right), \ \forall n \geq k \right. \right\}, \end{aligned}$$

therefore $\underleftarrow{\lim} \overline{H}^i_n \in \otimes_{l=1}^{\infty} \Sigma_{\overline{H}^i_l}$.

Bibliography

[1] Ahn, D. S., "Hierarchies of Ambiguous Beliefs", *Journal of Economic Theory*, 2007, **136**, 286-301.

[2] Aliprantis, C. D. and K. C. Border, *Infinite Dimensional Analysis*, Springer-Verlag, Berlin, 1999.

[3] Andersen, S. E. and B. Jessen, "On the Introduction of Measures in Infinite Product Sets", *Danske Videnskabernes Selskab Matematisk-Fysiske Meddelelser*, 1948, **4**, 3-7.

[4] Aumann, R.J., "Agreeing to disagree", *The Annals of Statistics*, 1976, **4**, 1236-1239.

[5] Awodey, S., *Category Theory*, Oxford Logic Guides, Oxford University Press, 2006.

[6] Battigalli, P. and M. Siniscalchi, "Hierarchies of Conditional Beliefs and Interactive Epistemology in Dynamic Games", *Journal of Economic Theory*, 1999, **88**, 188-230.

[7] Bankston, P., "Clopen Sets in Hyperspaces", *Proceedings of the American Mathematical Society*, 1976, **54**, 298-302.

[8] Blau, J. H., "The Space of Measures on a Given Set", *Fundamenta Mathematicae*, 1951, **38**, 23-34.

[9] Beer, G., *Topologies on Closed and Closed Convex Sets*, 1993, Kluwer Academic Publishers.

[10] Bewley, T., "Knightian Decision Theory: Part I", *Cowles Foundation Discussion Paper* 807, 1986.

[11] Bogachev, V. I., "Measures on Topological Spaces", *Journal of Mathematical Sciences*, 1998, **91**, 3033-3156.

[12] Bogachev, V., *Measure Theory*, Springer-Verlag, Berlin, 2006.

[13] Bourbaki N., *Elements de mathematique, Topologie generale*, 1965, Hermann, Paris.

[14] Bourbaki N., *Eléments de Mathématque, Intégration, Livre VI, Chapitre IX, Intégration sur Les Espaces Topologiques Séparés*, Hermann, 1969.

[15] Bourbaki, N., *Elements of Mathematics. Theory of Sets*, Springer-Verlag, Berlin, 2004.

[16] Brandenburger, A., "On the Existence of a 'Complete' Possibility Structure", In Basili, M., Dimitri, N., Gilboa, I. (Eds.), *Cognitive Processes and Economic Behavior*, 2003 Routledge, 30-34.

[17] Brandenburger, A. and E. Dekel, "Hierarchies of Beliefs and Common Knowledge", *Journal of Economic Theory*, 1993, **59**, 189-198.

[18] Brandenburger, A., and H. J. Keisler, "An Impossibility Theorem on Beliefs in Games", *Studia Logica*, 2006, **84**, 211-240.

[19] Charatonik J.J., Charatonik W.J. and A. Illanes, "Openess of Induced Mappings", *Topology and its Applications*, 1999, **98**, 67-80.

[20] Chen, Y., "Universality of the Epstein-Wang Type Structure", *Games and Economic Behavior*, forthcoming.

[21] Dudley, R. M., *Real Analysis and Probability*, Cambridge University Press, New York, 2004.

[22] Engelking, R., *General Topology*, Heldermann Verlag, 1989.

[23] Epstein, L., and T. Wang, "Beliefs about Beliefs Without Probabilities", *Econometrica*, 1996, **64**, 1343-1373.

[24] Friedenberg, A., "When Do Type Structures Contain All Hierarchies of Beliefs?", *Games and Economic Behavior*, forthcoming.

[25] Friedenberg, A. and M. Meier, "On the Relationship Between Hierarchy and Type Morpshims", *Economic Theory*, forthcoming.

[26] Gilboa, I. and D. Schmeidler, "Maxmin Expected Utility with Non-unique Prior", *Journal of Mathematical Economics*, 1989, **18**, 141-153.

[27] Gul, F. and W. Pesendorfer, "Self-control and the Theory of Consumption", *Econometrica*, 2004, **72**, 119-158.

[28] Harsanyi, J. C., "Games with Incomplete Information Played by "Bayesian" Players", Parts I, II, and III, *Management Science*, 1967-68, **14**, 159-182, 320-334, 486-502.

[29] Heifetz, A., "The Bayesian Formulation of Incomplete Infomation - the Non-Compact Case", *International Journal of Game Theory*, 1993, **21**, 329-338.

[30] Heifetz, A. and D. Samet, "Topology-free Typology of Beliefs", *Journal of Economic Theory*, 1998, **82**, 324-341.

[31] Heifetz, A. and D. Samet, "Coherent Beliefs are not Always Types", *Journal of Mathematical Economics*, 1999, **32**, 475-488.

[32] Hosokawa, H., "Induced mappings on hyperspaces", *Tsukuba Journal of Mathematics*, 1997, **21**, 239–250.

[33] Kaji, A. and T. Ui, "Incomplete Information Games with Multiple Priors", *Japanese Economic Review*, 2005, **56**, 332-351.

[34] Kechris, A. S., *Classical Descriptive Set Theory*, Springer-Verlag, New York, 1995.

[35] Knight, F., *Risk, Uncertainty, and Profit*, 1921, Houghton Mifflin, New York.

[36] Liu, Q., "On Redundant Types and Bayesian Formulation of Incomplete Information", *Journal of Economic Theory*, 2009, **44**, 2115-2145.

[37] Luther, N. Y., "Locally Compact Spaces of Measures", *Proceedings of the American Mathematical Society*, 1970, **25**, 541-547.

[38] Machina, M. and D. Schmeidler, "A More Robust Definition of Subjective Probability", *Econometrica*, 1992, **60**, 745-780.

[39] Machina, M. and D. Schmeidler, "Bayes Without Bernoulli: Conditions for Probabilistically Sophisticated Choice", *Journal of Economic Theory*, 1995, **67**, 106-28.

[40] Mallory D. J. and M. Sion, "Limits of Inverse Systems of Measures", *Annales de l'Institut Fourier*, 1971, **21**, 25-57.

[41] Marczewski, E., "On Compact Measures", *Fundamenta Mathematicae*, 1953, **40**, 113-124.

[42] Mariotti, T., Meier, M. and M. Piccione, "Hierarchies of Beliefs for Compact Possibility Models", *Journal of Mathematical Economics*, 2005, **41**, 303–324.

[43] Mertens, J. F., Sorin S. and S. Zamir, *Repeated Games: Part A: Background Material*, CORE Discussion Paper No. 9420, Université Catolique de Louvain, 1994.

[44] Mertens, J. F. and S. Zamir, "Formulation of Bayesian Analysis for Games with Incomplete Information", *International Journal of Game Theory*, 1985, **14**, 1-29.

[45] Michael, E., "Topologies on Spaces of Subsets", *Transactions of the American Mathematical Society*, 1951, **71**, 152-182.

[46] Monderer, D. and D. Samet, "Approximating Common Knowledge with Common Beliefs", *Games and Economic Behavior*, 1989, **1**, 170-190.

[47] Pintér, M., "The Existence of an Inverse Limit of an Inverse System of Measure Spaces - A Purely Measurable Case", *Acta Mathematica Hungarica*, 2010, **126**, 65-77.

[48] Pintér, M., "The Non-Existence of a Universal Topological Type Space", *Journal of Mathematical Economics*, 2010, **46**, 223-229.

[49] Rao, M. M., *Foundation of stochastic analysis*, Academic Press, New York, 1981

[50] Rao, M. M., *Measure Theory and Integration*, CRC Press, 2004.

[51] Renyi, A., "On a new axiomatic theory of probability", *Acta Mathematica Academiae Scientiarum Hungaricae*, 1955, **6**, 285-335.

[52] Ressel, P., "Some Continuity and Measurability Results on Spaces of Measures", *Mathematica Scandinavica*, 1977, **40**, 69-78.

[53] Schmeidler, D., "Subjective Probability and Expected Utility Without Additivity", *Econometrica*, 1989, **57**, 571-587.

[54] Schwartz, L., *Radon Measures on Arbitrary Topological Spaces and Cylindrical Measures*, Oxford University Press, London, 1973.

[55] Siniscalchi, M., "Epistemic Game Theory: Beliefs and Types", *The New Palgrave Dictionary of Economics*, edited by L. Blume and S. Durlauf, McMillan, forthcoming.

[56] Steen, L.Y. and Seebach, J.A., *Counterexamples in Topology*, Springer-Verlag, Berlin, 1978.

[57] Topsoe, F., *Topology and Measure*, Springer-Verlag, Berlin, 1970.

[58] Viglizzo, I. D., *Coalgebras on Measurable Spaces*, Ph.D. thesis, Indiana University, 2005.