A comment on "Admissibility and Assumption"*

Emiliano Catonini† Nicodemo De Vito‡

December 2017 - PRELIMINARY AND INCOMPLETE

Abstract

Our recent research (see [17]) emphasizes the importance of hierarchies of lexicographic beliefs in finite, static games, showing how it affects the epistemic justification of Iterated Admissibility. Here, we discuss in detail the properties of hierarchies of lexicographic beliefs in terms of their minimal representation. This allows us to elucidate our formulation of the canonical type structure, and to compare it to the construction due to Lee ([40]). We also show that the results in [17] do not hinge on the presence of redundancies in the representation of Lexicographic Expected Utility preferences.

Keywords: lexicographic probability systems, hierarchies of beliefs, hierarchies of preferences, lexicographic type structures, universality, type morphisms, iterated admissibility.

*Working paper. Financial support from European Research Council (STRATEMOTIONS–GA 324219) is gratefully acknowledged.

†Higher School of Economics, Moscow, emiliano.catonini@gmail.com

‡Bocconi University, nicodemo.devito@unibocconi.it

1 Introduction

Lexicographic Probability Systems (henceforth, LPS) have been used in recent papers to provide epistemic foundations of solution concepts in games. An LPS is a finite sequence $(\mu_1, ..., \mu_n)$ of probabilities (called also "theories") over a relevant space of uncertainty, and the decision maker uses them lexicographically to determine her preferences. To clarify, the decision maker starts by evaluating her optimal choice according to the first theory μ_1; if μ_1 leads to more than one optimal choice, then the decision maker uses μ_2 to break ties, and so on. So, LPS's generalize usual probabilistic beliefs, as they are representations of Lexicographic Expected Utility preferences, which include standard Subjective Expected Utility preferences as special case (see [7] for an axiomatic derivation). The advantage of LPS's is that they can be used to formalize the idea that a decision maker deems an event "infinitely more likely than" another, without ruling out the possibility that the latter occurs.

Brandenburger, Friedenberg and Keisler ([15],henceforth, BFK) introduced the formalism of lexicographic type structures as a device to model players' interactive beliefs in a strategic setting: Formally, for each player there is a set of types; each type is associated with an LPS over the product of the space of primitive uncertainty and the set of the opponents' types. Lexicographic type structures have been proven extremely useful for epistemic analyses of Iterated Admissibility (i.e., iterated weak dominance) in finite, static games with complete information, e.g., BFK, Dekel, Friedenberg and Siniscalchi ([20]) and Keisler and Lee ([37]).

A lexicographic type structure provides only an *implicit* way to describe the belief hierarchies of the players: The above mentioned papers do not construct a canonical, universal type structure into which any other type structure can be mapped in a unique belief-preserving way. By contrast, in this paper we adopt an *explicit* perspective, in which hierarchies of LPS's are described and discussed in detail: Our aim is to provide an in-depth analysis on the hierarchical approach adopted by two different contributions to this field of research, namely the works of Lee ([38],[40]) and Catonini and De Vito ([17],[18]). Next, we show how hierarchies of LPS's are fundamental for an epistemic characterization of Iterated Admissibility in finite games.

Before describing our results in more detail we survey briefly the background.

Hierarchies of LPS's and type structures. In our related work ([18]), we establish the foundation of lexicographic type structures and hierarchies of LPS's in the same way as papers such as [42] and [14] do for standard (i.e., probabilistic) type structures. Specifically, it is shown the set of all hierarchies of LPS's satisfying *coherence and common full belief of coherence* can be endowed with the internal structure of a lexicographic type structure. Such "canonical" type structure turns out to be *universal* in the sense that each type structure can be uniquely mapped into it by a map—called *type morphism*—which preserves the beliefs associated with types.

The notion of coherence for hierarchies of LPS's is a simple generalization of the standard one for hierarchies of probabilistic beliefs; it simply requires that the marginals of the higher-order LPS's coincide with the corresponding lower-order LPS's. There, the marginalization operation is taken pointwise: e.g., the marginal of a length-n LPS is simply defined as the length-n sequence of the marginals of the component measures. Consequently, each coherent hierarchy of LPS's consists of an infinite sequence of LPS's of the same (finite) length. The notion of full belief is the LPS-based analogue of the notion of certainty for probabilistic beliefs.

In [17], we make use of the universality property of the canonical type structure to provide an epistemic characterization of Iterated Admissibility in finite games. We do this by weakening BFK's concept of *Assumption*, which is essentially an "infinitely more likely than" relation between uncertain events, to that of *Cautious Belief.* The weakening is crucial for the existence of a "(Cautious) Rationality and Common Cautious Belief of (Cautious) Rationality" event that justifies Iterated Admissibility, because in any complete and continuous type structure (e.g., the canonical one) the epistemic notion of Rationality and Common Assumption of Rationality (RCAR) introduced by BFK has no bite—that is, there is no state of the world consistent with RCAR (see also [20]).[1]

Preference redundancy and minimal LPS's. The preference-based axiomatization of LPS's, given in [7], highlights the fact that multiple LPS's can represent the same Lexicographic Expected Utility preference relation. One source of redundant LPS-based representations of lexicographic preferences is *non-minimality*: For instance, LPS's (μ_1, μ_1, μ_2) and (μ_1, μ_2) clearly represent the one and the same preference relation. Therefore, as pointed out by Lee [38], for every hierarchy of lexicographic beliefs there are infinitely many distinct hierarchies representing the same *hierarchy of lexicographic preferences*. In other words, the costruction of the canonical space of hierarchies in [18] results in a type structure containing redundant (in terms of preferences) types.

By restricting attention to minimal LPS's, i.e., minimal length representation of lexicographic preferences, Lee ([40]) constructs a space of hierarchies of LPS's satisfying a coherence requirement which is weaker than the one in [18]. Lee's notion of coherence allows a $(k+1)$-order belief to be a longer LPS than k-order belief, but it preserves coherence of *preferences* represented by the beliefs. Such approach is therefore closely related to the preference hierarchical approach of Epstein and Wang [27] and Di Tillio [24].

Lee provides a "bottom-up" construction (*à la* Mertens and Zamir [42]) of the space of hierarchies of minimal beliefs in which some hierarchies cannot be generated by any type structure. The reason why this occurs is that, while the length of all k-order beliefs is finite for all $k \in \mathbb{N}$, it may grow indefinitely as $k \to \infty$. Consequently, there are hierarchies that cannot be summarized by a single LPS, which must necessarily have a finite length. Lee uses this fact to provide an epistemic justification of Iterated Admissibility under BFK's notion of Assumption, and the presence of hierarchies which cannot be generated by types is crucial to obtain a non-empty set of states consistent with RCAR.[2]

Main questions and our contribution. Clearly, there are overlaps and similarities between the approaches outlined above. At the same time, they offer distinct ways of analysis. This raises some interesting questions.

The first question is how the canonical type structure for lexicographic beliefs constructed in [18] relates to the corresponding type structures with minimal beliefs in [40]. We address this issue as follows. First, by selecting only the hierarchies with an upper bound on the length of all finite-order beliefs, we show that a construction of a "canonical" type structure for hierarchies of minimal beliefs is possible, along the lines outlined in [18]. The canonical space of hierarchies constructed in this way turns out to be *behaviorally equivalent* to the canonical space of hierarchies of LPS's in [18]. This is so because Lee's notion of coherence preserves coherence of *preferences* exactly in the same way as the notion of coherence in [18] does. This version of the canonical type structure satisfies a terminality property analogous to that in [18]. Indeed, under an appropriate notion of hierarchy morphism, every type structure can be mapped into it in a way that preserves the hierarchies of minimal beliefs.

It is noteworthy that the canonical type structure of minimal LPS's is *proper* subset of the space of hierarchies constructed in [40]. As we will formally elaborate in Section 7, in order to

[1] BFK's notion of rationality is stronger than the one we adopt in [17] (and also in the current paper), as it includes a full-support requirement on top of lexicographic expected utility maximization. Cautious Rationality features an additional "marginal full-support requirement" (i.e. full support of the marginal LPS on opponent's strategies) and thus is still weaker than BFK's rationality. However, this weakening is not crucial for our characterization results.

[2] The epistemic notion of RCAR in [40] is, in a sense, more general than that in BFK, as it is stated in the space of hierarchies (cf. [40, Section 5]).

meaningfully define the notion of full belief of coherence (in Lee's sense), we need to restrict attention only to those hierarchies with an upper bound on the length of all finite-order beliefs.

The second question we address is whether the epistemic notion of Cautious Rationality and Common Cautious Belief of Cautious Rationality—as elaborated in our work ([17])—still characterizes Iterated Admissibility in the new framework. We provide an affirmative answer, by showing that the results in [17] do *not* hinge on the presence of redundancies in the representation of Lexicographic Expected Utility preferences. Specifically, we show that an analogue of main result in [17] also holds for this version of the canonical type structure. This highlights the fact the epistemic notion of Cautious Rationality and Common Cautious Belief of Cautious Rationality depends only on hierarchies of preferences, not on their LPS-based representations.

Structure of the paper. The remainder of this paper is organized as follows. Section 2 contains some definitions and the basic notation that will be used throughout. LPS's and ther minimal representations of lexicographic preference relations are introduced in Section 3. Section 4 shows how a canonical type structures for hierarchies of lexicographic minimal beliefs can be constructed. We also provide a terminality result for this type structure. Section 6 shows how the analysis in [17] to the epistemic characterization of iterated admissibility can be easily adapted in this new epistemic framework. Section 7 concludes with a discussion on some conceptual and technical aspects of the paper. Proofs omitted from the main text are collected in the Appendix.

2 Preliminaries and notation

We begin with some definitions and the basic notation that will be used throughout the paper.[3] A measurable space is a pair (X, Σ_X), where X is a set and Σ_X is a σ-field, the elements of which are called *events*. When it is clear from the context which σ-field on X we are considering, we suppress reference to Σ_X and simply write X to denote a measurable space. All the sets considered in this paper are assumed to be metrizable topological spaces, and they are endowed with the Borel σ-field. A *Polish* space is a topological space which is homeomorphic to a complete, separable metrizable space. A *Lusin* space is a topological space which is the continuous, injective image of a complete, separable metrizable space.[4] Clearly, a Polish space is also Lusin. Every metrizable Lusin space is measure-theoretic isomorphic to a Borel subset of some Polish space.

If $(X_n)_{n\in\mathbb{N}}$ is a countable collection of pairwise disjoint topological spaces, then the set $X = \cup_{n\in\mathbb{N}} X_n$ is endowed with the *direct sum topology*.[5] The set X is metrizable Lusin (resp. Polish) provided each X_n is metrizable Lusin (resp. Polish).

We consider any product, finite or countable, of topological spaces as a topological space with the product topology. As such, a countable product of metrizable Lusin (resp. Polish) spaces is also metrizable Lusin (resp. Polish). Furthermore, given topological spaces X and Y, we denote by Proj_X the canonical projection from $X \times Y$ onto X; in view of our assumption, the map Proj_X is continuous and open (i.e., the image of each open set in $X \times Y$ is an open set

[3] A more detailed presentation of the following concepts, as well as related mathematical results, can be found in [8], [26], [47], [50], [53]. In the remainder of the paper, we shall make use of the results mentioned in this section, sometimes without referring to them explicitly.

[4] If X is a Lusin topological space, and Σ_X is the corresponding Borel σ-field, then the measurable space (X, Σ_X) is *Standard Borel* (cf. [19]).

[5] The assumption that the spaces X_n are pairwise disjoint is without any loss of generality, since they can be replaced by a homeomorphic copy, if needed (see [26, p.75]).

in X under the map Proj_X). Finally, for a measurable space X, we denote by Id_X the identity map on X, that is, $\mathrm{Id}_X(x) = x$ for all $x \in X$.

3 Lexicographic beliefs and lexicographic preferences

3.1 Lexicographic probability systems

Given a topological space X, we denote by $\mathcal{M}(X)$ the set of Borel probability measures on X. The set $\mathcal{M}(X)$ is endowed with the *weak**-topology. Then, if X is metrizable Lusin (resp. Polish), then $\mathcal{M}(X)$ is also metrizable Lusin (resp. Polish).

We denote by $\mathcal{N}(X)$ (resp. $\mathcal{N}_n(X)$) the set of all finite (resp. length-n) sequences of Borel probability measures on X, that is,

$$\begin{aligned}\mathcal{N}(X) &= \cup_{n\in\mathbb{N}}\mathcal{N}_n(X)\\ &= \cup_{n\in\mathbb{N}}(\mathcal{M}(X))^n.\end{aligned}$$

This means that if $\overline{\mu} \in \mathcal{N}(X)$, then there is some $n \in \mathbb{N}$ such that $\overline{\mu} = (\mu_1, ..., \mu_n)$. Call each $\overline{\mu} = (\mu_1, ..., \mu_n) \in \mathcal{N}(X)$ a **lexicographic probability system** (LPS). The **length** of the LPS $\overline{\mu} \in \mathcal{N}(X)$ is denoted by $\ell(\overline{\mu}) \in \mathbb{N}$.

Each $\overline{\mu} = (\mu_1, ..., \mu_n) \in \mathcal{N}(X)$ is called **lexicographic probability system** (LPS). We say that $\overline{\mu}$ is a **mutually singular** LPS or a **lexicographic conditional probability system** (LCPS) if there are Borel sets $E_1, ..., E_n$ in X such that, for every $l \leq n$, $\mu_l(E_l) = 1$ and $\mu_l(E_m) = 0$ for $m \neq l$. Write $\mathcal{L}(X)$ (resp. $\mathcal{L}_n(X)$) for the set of LCPS's (resp. length-n LCPS's). Both topological spaces $\mathcal{N}(X)$ and $\mathcal{L}(X)$ metrizable Lusin provided X is metrizable Lusin. In particular, if X is Polish, so are $\mathcal{N}(X)$ and $\mathcal{L}(X)$.[6]

For every Borel probability measure μ on a topological space X, the support of μ, denoted by $\mathrm{Supp}\mu$, is the smallest closed subset of X such that $\mu(\mathrm{Supp}\mu) = 1$. The support of an LPS $\overline{\mu} = (\mu_1, ..., \mu_n) \in \mathcal{N}(X)$ is thus defined as $\mathrm{Supp}\overline{\mu} = \cup_{l\leq n}\mathrm{Supp}\mu_l$. So, an LPS $\overline{\mu} = (\mu_1, ..., \mu_n) \in \mathcal{N}(X)$ is of **full-support** if $\cup_{l\leq n}\mathrm{Supp}\mu_l = X$. We write $\mathcal{N}_n^+(X)$ for the set of all full-support, length-n LPS's and $\mathcal{N}^+(X)$ (resp. $\mathcal{L}^+(X)$) for the set of full-support LPS's (resp. full-support LCPS's).

Suppose we are given topological spaces X and Y, and a Borel map $f : X \to Y$. The map $\widetilde{f} : \mathcal{M}(X) \to \mathcal{M}(Y)$, defined by

$$\widetilde{f}(\mu)(E) = \mu\left(f^{-1}(E)\right), \, \mu \in \mathcal{M}(X), \, E \in \Sigma_Y,$$

is called the image (or pushforward) measure map of f. For each $n \in \mathbb{N}$, the map $\widehat{f}_{(n)} : \mathcal{N}_n(X) \to \mathcal{N}_n(Y)$ is defined by

$$(\mu_1, ..., \mu_n) \mapsto \widehat{f}_{(n)}((\mu_1, ..., \mu_n)) = \left(\widetilde{f}(\mu_k)\right)_{k\leq n}.$$

Thus the map $\widehat{f} : \mathcal{N}(X) \to \mathcal{N}(Y)$ defined by

$$\widehat{f}(\overline{\mu}) = \widehat{f}_{(n)}(\overline{\mu}), \, \overline{\mu} \in \mathcal{N}_n(X),$$

is called the **image LPS map of** f. In other words, the map $\widehat{f}$ is the combination of the functions $\left(\widehat{f}_{(n)}\right)_{n\in\mathbb{N}}$, and it is Borel measurable.[7] In particular, if X and Y are metrizable Lusin

[6] We refer the reader to our companion paper [17] for a proof of those results.

[7] For details and proofs related to Borel measurability and continuity of the involved maps, we refer the reader to [17].

spaces, then the marginal measure of $\mu \in \mathcal{M}(X \times Y)$ on X is defined as $\text{marg}_X \mu = \widetilde{\text{Proj}}_X(\mu)$. Consequently, the marginal of $\overline{\mu} \in \mathcal{N}(X \times Y)$ on X is defined as $\overline{\text{marg}}_X \overline{\mu} = \widehat{\text{Proj}}_X(\overline{\mu})$, and $\widehat{\text{Proj}}_X : \mathcal{N}(X \times Y) \to \mathcal{N}(X)$ is a continuous and surjective map.

3.2 Lexicographic preferences

In what follows, we fix a metrizable Lusin space X.

Definition 1 *An* ***act*** *defined on X is a Borel measurable map $f : X \to [0,1]$. The set of all acts on X is denoted by* $\text{ACT}(X)$.

Definition 2 *Let $\overline{\mu} = (\mu_1, ..., \mu_n) \in \mathcal{N}(X)$. The preference relation $\succsim^{\overline{\mu}}$ on $\text{ACT}(X)$ is defined as follows. For all $f, g \in \text{ACT}(X)$,*

$$f \succsim^{\overline{\mu}} g \iff \left(\int_X f d\mu_1, ..., \int_X f d\mu_n \right) \geq_L \left(\int_X g d\mu_1, ..., \int_X g d\mu_n \right),$$

where $\geq_L$ stands for the usual lexicographic order.

A preference relation $\succsim$ on $\text{ACT}(X)$ is called a ***lexicographic expected utility (LEU) preference relation*** *if there exists $\overline{\mu} \in \mathcal{N}(X)$ such that $\succsim = \succsim^{\overline{\mu}}$. In such a case, we say that $\overline{\mu}$ is a* ***LEU preference representation****.*

We also introduce an equivalence relation on $\mathcal{N}(X)$.

Definition 3 *Let $\overline{\mu}, \overline{\nu} \in \mathcal{N}(X)$. We say that $\overline{\mu}$ and $\overline{\nu}$* ***represent the same preferences****, and we write $\overline{\mu} \cong \overline{\nu}$, if $\succsim^{\overline{\mu}} = \succsim^{\overline{\nu}}$. Furthermore, we say that $\overline{\mu}$ is a* ***minimal length LEU preference representation*** *if it is the shortest LPS that represents $\succsim^{\overline{\mu}}$, that is, for all $\overline{\nu} \in \mathcal{N}(X)$,*

$$\overline{\mu} \cong \overline{\nu} \Longrightarrow \ell(\overline{\mu}) \leq \ell(\overline{\nu}).$$

The set of all minimal length LEU preference representations is denoted by $\underline{\mathcal{N}}(X)$.

Remark 1 *The set $\underline{\mathcal{N}}(X)$ is open in $\mathcal{N}(X)$ ([40, Lemma 3.2]).*

The following result shows that for LCPS's each member of the partition induced by the equivalence relation $\cong$ is a singleton (cf. [7, Theorem 5.3]).

Proposition 1 *Let $\overline{\mu}, \overline{\nu} \in \mathcal{L}(X)$. Thus $\overline{\mu} \cong \overline{\nu}$ if and only if $\overline{\mu} = \overline{\nu}$.*

An immediate implication of this fact is given by the following

Corollary 1 $\mathcal{L}(X) \subseteq \underline{\mathcal{N}}(X)$.

We now formally define the concept of pushforward LEU preference relation. Fix a Borel measurable map $\varphi : X \to Y$ between Lusin spaces X and Y. Given a preference relation $\succsim$ on ACT(X), the pushforward preference relation $\succsim^{\varphi}$ on ACT(Y) induced by φ is defined as follows: For all $f, g \in$ACT(Y),

$$f \succsim^{\varphi} g \iff f \circ \varphi \succsim g \circ \varphi.$$

In particular, given a product space $X \times Y$, the pushforward preference relation $\succsim^{\mathrm{Proj}_Y}$ on ACT(Y) induced by the canonical projection $\mathrm{Proj}_Y : X \times Y \to Y$, i.e.,

$$f \succsim^{\mathrm{Proj}_Y} g \iff f \circ \mathrm{Proj}_Y \succsim g \circ \mathrm{Proj}_Y, \forall f, g \in \mathrm{ACT}(Y),$$

is called the marginal preference relation.[8]

Clearly the LPS $\overline{\nu} = (\nu_1, ..., \nu_n) \in \mathcal{N}(Y)$ represents the LEU preference relation $\succsim^{\varphi}$ on ACT(Y) induced by $\varphi : X \to Y$ if and only if there exists $\overline{\mu} = (\mu_1, ..., \mu_n) \in \mathcal{N}(X)$ such that $\widehat{\varphi}(\overline{\mu}) = \overline{\nu}$ and

$$\left(\int_X (f \circ \varphi)\, d\mu_l\right)_{l=1}^{n} \geq_L \left(\int_X (g \circ \varphi)\, d\mu_l\right)_{l=1}^{n},$$

for all $f, g \in \mathrm{ACT}(Y)$.[9] It turns out that

$$\overline{\mu} \cong \overline{\nu} \Longrightarrow \widehat{\varphi}(\overline{\mu}) \cong \widehat{\varphi}(\overline{\nu}), \forall \overline{\mu}, \overline{\nu} \in \mathcal{N}(X).$$

However, we are interested in minimal length representations of pushforward LEU preference relations. To this end, we need to define a map $\mathbf{m}_X : \mathcal{N}(X) \to \underline{\mathcal{N}}(X)$ which minimizes the length of an LPS while preserving the representation of preferences. Given an LPS $\overline{\mu} = (\mu_1, ..., \mu_n)$ and $l \leq n$, let $\overline{\mu}_{\leq l} = (\mu_1, ..., \mu_l)$ and $\overline{\mu}_{-l} = \left(\mu_1, ..., \mu_{l-1}, \mu_{l+1}, ..., \mu_n\right)$. The map $\mathbf{m}_X$ is defined as follows:

$$\mathbf{m}_X(\overline{\mu}) = \begin{cases} \overline{\mu}, & \text{if } \overline{\mu} \in \underline{\mathcal{N}}(X), \\ \mathbf{m}_X\left(\overline{\mu}_{-l}\right), & \text{if } \exists l > 1 \text{ s.t. } \overline{\mu}_{\leq l-1} \in \underline{\mathcal{N}}(X) \text{ and } \overline{\mu}_{\leq l} \notin \underline{\mathcal{N}}(X). \end{cases}$$

To clarify, suppose that μ_l is the first probability measure in the LPS $\overline{\mu} = (\mu_1, ..., \mu_n)$ that can be written as convex combination of the preceeding measures (so that it contains redundant information about preferences). Then μ_l can be removed from $\overline{\mu}$ without changing preferences. The minimization $\mathbf{m}_X(\overline{\mu})$ of $\overline{\mu}$ is thus defined by a recursive procedure which deletes such redundant measures from $\overline{\mu}$ in the order they appear until none remain. See [40] for examples.

As shown in [40, Lemma A.4], the map $\mathbf{m}_X : \mathcal{N}(X) \to \underline{\mathcal{N}}(X)$ is a Borel class 2 map (i.e., the inverse image of every open set is a G_δ-set). It is trivial to check that the map $\mathbf{m}_X$ is onto. We also note the following fact.

Lemma 1 *Let X be a finite space. For each $\overline{\mu} \in \mathcal{N}(X)$, it holds that*

$$\mathrm{Supp}\overline{\mu} = \mathrm{Supp}\mathbf{m}_X(\overline{\mu}).$$

Thus $\mathbf{m}_X(\mathcal{N}^+(X)) = \underline{\mathcal{N}}^+(X)$.

[8] Put differently, $\succsim^{\mathrm{Proj}_Y}$ is the restriction of $\succsim$ to outcomes that are contingent on Y but they are constant along the X-dimension; each act $f \in \mathrm{ACT}(Y)$ is identified with the act $\overline{f} \in \mathrm{ACT}(X \times Y)$ defined as

$$\overline{f}((x, y)) = (f \circ \mathrm{Proj}_Y)((x, y)) = f(y), \forall (x, y) \in X \times Y.$$

[9] To see this, note that

$$\int_Y f d\nu_l = \int_Y f d\widetilde{\varphi}(\mu_l) = \int_X (f \circ \varphi)\, d\mu_l,$$

where the second equality follows from the Change of Variable Theorem ([1, Theorem 13.46]).

We now introduce the notion of minimal-LPS pushforward map.

0

Definition 4 *Fix a Borel measurable map $\varphi : X \to Y$ between Lusin spaces X and Y. Let $\overline{\mu} = (\mu_1, ..., \mu_n) \in \mathcal{N}(X)$. The map*

$$\mathbf{m}_Y \circ \widehat{\varphi} : \mathcal{N}(X) \to \underline{\mathcal{N}}(Y)$$

*is called the **minimal-LPS pushforward map under** φ, and $\mathbf{m}_Y(\widehat{\varphi}(\overline{\mu}))$ is called the **minimal-image LPS** representing $\succsim^{\varphi}$ on* $\mathrm{ACT}(Y)$.

*Given a product space $X \times Y$, the **minimal-marginal map** $\mathrm{margm}_Y : \mathcal{N}(X \times Y) \to \underline{\mathcal{N}}(Y)$ is defined as*

$$\mathrm{margm}_Y = \mathbf{m}_Y \circ \overline{\mathrm{marg}}_Y,$$

*and $\mathrm{margm}_Y(\overline{\mu})$ is called the **minimal-LPS marginal** of $\overline{\mu} \in \mathcal{N}(X \times Y)$ on Y.*

Note the following property of the minimal-LPS pushforward map under $\varphi : X \to Y$.

$$\mathbf{m}_Y(\widehat{\varphi}(\overline{\mu})) = \mathbf{m}_Y(\widehat{\varphi}(\overline{\nu})) \Longrightarrow \widehat{\varphi}(\overline{\mu}) \cong \widehat{\varphi}(\overline{\nu}), \forall \overline{\mu}, \overline{\nu} \in \mathcal{N}(X).$$

The reverse implication is not true, as the following simple example shows.

Example 1 *Let $\nu, \mu \in \mathcal{M}(X)$ such that $\nu \neq \mu$, and let $\varphi : X \to X$ be the identity map. Consider LPS's $\overline{\mu}_1 = (\nu, \mu)$ and $\overline{\mu}_2 = \left(\nu, \frac{1}{2}\nu + \frac{1}{2}\mu\right)$. It clearly holds that*

$$\widehat{\varphi}(\overline{\mu}_1) = \overline{\mu}_1 \cong \overline{\mu}_2 = \widehat{\varphi}(\overline{\mu}_2),$$

but

$$\mathbf{m}_X(\widehat{\varphi}(\overline{\mu}_1)) = \overline{\mu}_1 \neq \overline{\mu}_2 = \mathbf{m}_X(\widehat{\varphi}(\overline{\mu}_2)).$$

Clearly, the LPS pushforward map $\widehat{\varphi}$ preserves preference-equivalence between LPS's. Thus, we point out:

Remark 2 *Fix a Borel map $\varphi : X \to Y$ between Lusin spaces X and Y. The following property holds true:*

$$\mathbf{m}_Y \circ \widehat{\varphi} = \mathbf{m}_Y \circ \widehat{\varphi} \circ \mathbf{m}_X.$$

Of course, an analogous property holds for the minimal-marginal map $\mathrm{margm}_Y : \mathcal{N}(X \times Y) \to \underline{\mathcal{N}}(Y)$. In view of the following result, the map margm_Y is measurable and onto.

Lemma 2 *Fix a Borel measurable map $\varphi : X \to Y$ between Lusin spaces X and Y. Thus the map $\mathbf{m}_Y \circ \widehat{\varphi} : \mathcal{N}(X) \to \underline{\mathcal{N}}(Y)$ is Borel measurable. It is onto provided φ is onto.*

Remark 3 *Note that if $\varphi : X \to Y$ is continuous, then the induced map $\mathbf{m}_Y \circ \widehat{\varphi}$ is not necessarily continuous.*

Lemma 3 *Fix a Borel measurable map $\varphi : X \to Y$ between Lusin spaces X and Y, and let $\overline{\mu} \in \mathcal{N}(X)$. The following statements hold true.*

(1) *If $\widehat{\varphi}(\overline{\mu}) \in \underline{\mathcal{N}}(Y)$, then $\overline{\mu} \in \underline{\mathcal{N}}(X)$.*

(2) *If $\widehat{\varphi}(\overline{\mu}) \in \mathcal{L}(Y)$, then $\overline{\mu} \in \mathcal{L}(X)$.*

We finally record, for future reference, the following simple fact:

Lemma 4 *Fix a Borel measurable map $\varphi : X \to Y$ between Lusin spaces X and Y, and let $\overline{\mu} \in \mathcal{N}(X)$. Thus*

$$\ell(\mathbf{m}_Y(\widehat{\varphi}(\overline{\mu}))) \leq \ell(\widehat{\varphi}(\mathbf{m}_X(\overline{\mu}))) = \ell(\mathbf{m}_X(\overline{\mu})) \leq \ell(\overline{\mu}) = \ell(\widehat{\varphi}(\overline{\mu})).$$

4 Higher order uncertainty in games

4.1 Hierarchies of lexicographic minimal beliefs

Fix a two-players set I;[10] given a player $i \in I$, we denote by $-i$ the other player in I. For each $i \in I$, let S_{-i} be a non-empty space—called *space of primitive uncertainty*—describing aspects of the strategic interaction that player i is uncertain about. Throughout this paper, S_{-i} will represent player $-i$'s strategy set: Player i does not know which strategy player $-i$ is going to choose. Other interpretations are also possible; for instance, S_{-i} may include player $-i$'s set of payoff functions, among which the true one is not known to player i. We assume that each S_i is a metrizable Lusin space, with $|S_i| \geq 2$.

We now construct the space of hierarchies of minimal lexicographic beliefs for each player. Formally, for each $i \in I$ define inductively the sequence of spaces $(\underline{X}_i^k)_{k\geq 0}$ as

$$\underline{X}_i^0 = S_{-i}, \tag{4.1}$$

$$\underline{X}_i^{k+1} = \underline{X}_i^k \times \underline{\mathcal{N}}\left(\underline{X}_{-i}^k\right); \; k \geq 0. \tag{4.2}$$

An element $h_i^{k+1} = \left(\overline{\mu}_i^1, \overline{\mu}_i^2, ..., \overline{\mu}_i^{k+1}\right)$ is a $(k+1)$-order lexicographic minimal belief hierarchy, where $\overline{\mu}_i^k = \left(\mu_i^{k,1}, ..., \mu_i^{k,n}\right) \in \underline{\mathcal{N}}\left(\underline{X}_i^{k-1}\right)$ denotes i's k-order minimum length LPS representing the k-order preference relation. It is easily seen that, according to our notation,

$$\underline{X}_i^{k+1} = \underline{X}_i^0 \times \prod_{l=0}^{k} \underline{\mathcal{N}}\left(\underline{X}_{-i}^l\right).$$

The set of all possible, infinite hierarchies of lexicographic minimal beliefs for player i is $\underline{H}_i^0 = \prod_{k=0}^{\infty} \underline{\mathcal{N}}\left(\underline{X}_i^k\right)$. The space $\underline{H}_i^0$ is a metrizable Lusin space—in particular, $\underline{H}_i^0$ is Polish provided each primitive space of uncertainty S_i is Polish.

As for the case of LPS's, we introduce the concept of length for hierarchies of minimal beliefs, which will turn out to be important for the construction of a canonical type structure.

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

Definition 5 *Fix a hierarchy* $h_i = \left(\overline{\mu}_i^1, \overline{\mu}_i^2, ...\right) \in \underline{H}_i^0$. *The* ***length of the hierarchy*** h_i, *denoted by* $\overline{\ell}(h_i)$, *is defined as follows.*

$$\overline{\ell}(h_i) = \sup \left\{ \ell\left(\overline{\mu}_i^k\right) | k \geq 1 \right\}.$$

The hierarchy $h_i \in \underline{H}_i^0$ *is* $\overline{\ell}$***-bounded*** *if* $\overline{\ell}(h_i) < \infty$.

As usual in the literature, we restrict attention to hierarchies of minimal beliefs which satisfy a coherence requirement. This notion of coherence says that higher order minimal beliefs cannot contradict lower order minimal beliefs. Formally:

Definition 6 *A hierarchy* $h_i = \left(\overline{\mu}_i^1, \overline{\mu}_i^2, ...\right) \in \underline{H}_i^0$ *is coherent if, for each* $k \geq 1$,

$$\mathrm{margm}_{\underline{X}_i^{k-1}} \overline{\mu}_i^{k+1} = \overline{\mu}_i^k.$$

Note that, for hierarchies of subjective expected utility (SEU) preferences, the notion of coherence in Definition 6 reduces to the standard one (cf. [42] or [14]); in such a case, a coherent hierarchy h_i is $\overline{\ell}$-bounded, namely $\overline{\ell}(h_i) = 1$.

We record the properties of coherent and $\overline{\ell}$-bounded hierachies in the following Proposition:

Proposition 2 *Fix a coherent hierarchy* $h_i = \left(\overline{\mu}_i^1, \overline{\mu}_i^2, ...\right) \in \underline{H}_i^0$. *Thus*

$$\ell\left(\overline{\mu}_i^k\right) \leq \ell\left(\overline{\mu}_i^{k+1}\right), \forall k \geq 1.$$

Additionally, if h_i *is* $\overline{\ell}$*-bounded, then there exists* $k' \geq 1$ *such that*

$$\ell\left(\overline{\mu}_i^k\right) = \ell\left(\overline{\mu}_i^{k+1}\right), \forall k \geq k'.$$

(So $\overline{\ell}(h_i) = \ell\left(\overline{\mu}_i^k\right)$ *for all* $k \geq k'$*.)*

For each player $i \in I$, the set of all coherent and $\overline{\ell}$-bounded hierarchies is denoted by $\underline{H}_i^1$.

Lemma 5 *For each* $i \in I$, *the set* $\underline{H}_i^1$ *is Borel (so Lusin) in* $\underline{H}_i^0$. *In particular, if each* S_i *is Polish, then* $\underline{H}_i^1$ *is a Polish subspace of* $\underline{H}_i^0$.

4.2 The canonical space of hierarchies

In this Section, we construct the *canonical spaces of hierarchies of minimal beliefs*, that is, the spaces of all collectively coherent hierarchies of minimal LPS's that can be endowed with the structure of a lexicographic type structure (formally defined in the next Section). Our construction follows quite closely the *top-down* construction of the canonical space given in [14]. An alternative, bottom-up construction (*à la* Mertens and Zamir [42]) is provided in [40].

The following result, which is the analog to [14, Proposition 1], is the building block of our analysis.

Proposition 3 *For each $i \in I$, there exists a Borel isomorphism $f_i : \underline{H}_i^1 \to \underline{\mathcal{N}}\left(S_{-i} \times \underline{H}_{-i}^0\right)$ such that*

$$\operatorname{margm}_{\underline{X}_i^{k-1}} f_i\left(\left(\overline{\mu}_i^1, \overline{\mu}_i^2, ...,\right)\right) = \overline{\mu}_i^k, \quad \forall k \geq 1.$$

The Proof of Proposition **3** makes use of the following Lemma (proved in the Appendix), which is essentially a version of the Kolmogorov Extension Theorem for LPS's (cf. [14, Lemma 1]).

Lemma 6 *Fix a countable collection of metrizable Lusin spaces $(W_l)_{l\geq 0}$, and, for each $k \geq 0$, let $Z_k = \Pi_{l=0}^k W_l$ and $Z = \Pi_{l=0}^\infty W_l$.*

1. *Let $(\overline{\nu}^k)_{k\geq 1}$ be a collection of LPS's on Z_k's satisfying the following consistency condition:*

$$\overline{\operatorname{marg}}_{Z_{k-1}} \overline{\nu}^{k+1} = \overline{\nu}^k, \ \forall k \geq 1. \tag{4.3}$$

   Thus there exists a unique LPS $\overline{\nu}$ on Z such that

$$\overline{\operatorname{marg}}_{Z_{k-1}} \overline{\nu} = \overline{\nu}^k, \ \forall k \geq 1.$$

   We refer to $\overline{\nu}$ as the LPS extension of $(\overline{\nu}^k)_{k\geq 1}$.

2. *Let $(\overline{\rho}^k)_{k\geq 1}$ be another collection of LPS's on Z_k's satisfying the consistency condition as in (4.3), and let $\overline{\rho}$ be its LPS extension. Thus*

$$\overline{\rho} \cong \overline{\nu} \Longleftrightarrow \overline{\rho}^k \cong \overline{\nu}^k, \ \forall k \geq 1.$$

3. *The LPS extension $\overline{\nu}$ of $(\overline{\nu}^k)_{k\geq 1}$ is of minimal-length if and only if there exists $k' \geq 1$ such that $\overline{\nu}^{k'+1} \in \underline{\mathcal{N}}\left(Z_{k'}\right)$.*

Proof of Proposition 3: Note that, for each $i \in I$, the set $S_{-i} \times \underline{H}_{-i}^0$ can be written as

$$S_{-i} \times \underline{H}_{-i}^0 = \underline{X}_i^{k-1} \times \prod_{l=k-1}^{\infty} \underline{\mathcal{N}}\left(\underline{X}_{-i}^l\right).$$

For each $i \in I$, let $\Phi_i : \underline{\mathcal{N}}\left(S_{-i} \times \underline{H}_{-i}^0\right) \to \underline{H}_i^1$ be the map defined by

$$\overline{\mu}_i \longmapsto \left(\Phi_i^k\left(\overline{\mu}_i\right)\right)_{k\geq 1} = \left(\operatorname{margm}_{\underline{X}_i^{k-1}} \overline{\mu}_i\right)_{k\geq 1}.$$

We show that Φ_i is a Borel measurable bijection, which implies that Φ_i is a Borel isomorphism from $\underline{\mathcal{N}}\left(S_{-i} \times \underline{H}_{-i}^{0}\right)$ onto $\underline{H}_i^1$ by virtue of Souslin Theorem (see, e.g., [19, Proposition 8.6.2]). By this, the result follows, since the function $f_i = \Phi_i^{-1}$ satisfies the required properties. Measurability of Φ_i is obvious, in that each Φ_i^k is measurable by Lemma 2.

In order to show that Φ_i is a surjection, fix an arbitrary $h_i = \left(\overline{\mu}_i^1, \overline{\mu}_i^2, ...\right) \in \underline{H}_i^1$. Thus there exists $k' \geq 1$ such that $\overline{\ell}(h_i) = \ell\left(\overline{\mu}_i^k\right)$ for all $k \geq k'$. Let

$$\overline{\nu}_i^k = \overline{\mathrm{marg}}_{\underline{X}_i^k} \overline{\mu}_i^{k'}, \forall k < k'.$$

It follows from the definition of $\mathrm{margm}_{\underline{X}_i^{k-1}}$ and the coherence condition that

$$\mathbf{m}_{\underline{X}_i^{k-1}}\left(\overline{\nu}_i^k\right) = \overline{\mu}_i^k, \forall k < k'.$$

By Lemma 6.1, to the array of LPS's $\left(\overline{\nu}_i^1, ..., \overline{\nu}_i^{k'-1}, \overline{\mu}_i^{k'}, ...\right)$ there corresponds a unique $\overline{\nu}_i \in \mathcal{N}\left(S_{-i} \times \underline{H}_{-i}^{0}\right)$ such that

$$\begin{aligned} \overline{\mathrm{marg}}_{\underline{X}_i^{k-1}} \overline{\nu}_i &= \overline{\mu}_i^k, \forall k \geq k'. \\ \overline{\mathrm{marg}}_{\underline{X}_i^{k-1}} \overline{\nu}_i &= \overline{\nu}_i^k, \forall k < k'. \end{aligned}$$

By Lemma 6.3, it follows that $\overline{\nu}_i \in \underline{\mathcal{N}}\left(S_{-i} \times \underline{H}_{-i}^{0}\right)$, specifically $\ell(\overline{\nu}_i) = \overline{\ell}(h_i)$, and

$$\mathrm{margm}_{\underline{X}_i^{k-1}} \overline{\nu}_i = \overline{\mu}_i^k, \forall k \geq 1.$$

Morever, uniqueness of $\overline{\nu}_i$ implies that Φ_i is an injection, as required. ■

Note that the condition of $\overline{\ell}$-boundedness for coherent hierarchies is essential for Proposition 3 to hold. Without this requirement, a coherent hierarchy could not be summarized by a single LPS, which must necessarily have a finite length. As matter of fact, Lee ([40, Theorem 4.2]) exhibits an example of a coherent hierarchy which fails to be represented by an LPS—such hierarchy is not $\overline{\ell}$-bounded. Note also that, differently from the literature on hierarchies of beliefs, the map f_i is not necessarily a homeomorphism, in that, as already remarked, each map $\mathbf{m}_{\underline{X}_i^{k-1}}$ is not continuous.

We now consider the case in which there is *common full belief of coherence*. As in [18], we say that player i, endowed with a coherent and $\overline{\ell}$-bounded hierarchy h_i, **fully believes** an event $E \subseteq S_{-i} \times \underline{H}_{-i}^{0}$ if $f_i(h_i)(E) = \overline{1}$, where $\overline{1}$ denotes a finite sequences of 1s; that is, every probability measure of the LPS $f_i(h_i) \in \underline{\mathcal{N}}\left(S_{-i} \times \underline{H}_{-i}^{0}\right)$ assigns probability 1 to E. We thus say that a hierarchy h_i is *consistent with full belief of coherence* if

- $h_i \in \underline{H}_i^1$, and
- $f_i(h_i)\left(S_{-i} \times \underline{H}_{-i}^{1}\right) = \overline{1}$.

Common full belief of coherence is imposed by defining inductively, for each $i \in I$, the following sets:

$$\begin{aligned} \underline{H}_i^{l+1} &= \left\{h_i \in \underline{H}_i^1 \,\middle|\, f_i(h_i)\left(S_{-i} \times \underline{H}_{-i}^{l}\right) = \overline{1}\right\}, l \geq 1, \\ \underline{H}_i &= \cap_{l \geq 1} \underline{H}_i^l. \end{aligned}$$

The set $\Pi_{i \in I} \underline{H}_i$ is naturally interpreted as the maximal set of players' hierarchies that are *consistent with common full belief* of coherence.

Proposition 4 *The restriction of f_i to $\underline{H}_i$ induces a Borel isomorphism β_i^{m} from $\underline{H}_i$ onto $\underline{\mathcal{N}}\left(S_{-i} \times \underline{H}_{-i}\right)$.*

The following simple result is needed in the proof of Proposition 4.

Lemma 7 *Let E be a Borel subset of a metrizable Lusin space X. Thus the set $\left\{\overline{\mu} \in \underline{\mathcal{N}}(X) \left| \overline{\mu}(E) = \overline{1}\right.\right\}$ is homeomorphic to $\underline{\mathcal{N}}(E)$.*

Proof of Proposition 4: It is easily seen that

$$\underline{H}_i = \left\{h_i \in \underline{H}_i^1 \left| f_i(h_i)\left(S_{-i} \times \underline{H}_{-i}\right) = \overline{1}\right.\right\}.$$

Indeed, if $h_i \in \underline{H}_i^1$, and $f_i(h_i)\left(S_{-i} \times \underline{H}_{-i}\right) = \overline{1}$, then clearly $h_i \in \underline{H}_i = \cap_{l \geq 1} \underline{H}_i^l$. On the other hand, if $h_i \in \underline{H}_i$, then, by σ-additivity of LPS's, it follows that

$$\begin{aligned} f_i(h_i)\left(S_{-i} \times \underline{H}_{-i}\right) &= f_i(h_i)\left(S_{-i} \times \cap_{l \geq 1} \underline{H}_i^l\right) \\ &= \lim_{l \to \infty} f_i(h_i)\left(S_{-i} \times \underline{H}_{-i}^l\right) \\ &= \overline{1}. \end{aligned}$$

Note that each $\underline{H}_i^l$ is Borel in H_i^1, and an analogous conclusion holds for $\underline{H}_i$. It follows from Lemma 7 that $f_i(\underline{H}_i)$ is isomorphic to $\underline{\mathcal{N}}\left(S_{-i} \times \underline{H}_{-i}\right)$. This shows the existence of a Borel isomorphism β_i^{m} from $\underline{H}_i$ onto $\underline{\mathcal{N}}\left(S_{-i} \times \underline{H}_{-i}\right)$. ■

Herafter, we shall refer to the set $\underline{H} = \Pi_{i \in I} \underline{H}_i$ as the *canonical space of hierarchies of minimal beliefs.*

5 Lexicographic type structures

The following definition is a natural generalization of the standard definition of epistemic type structure with beliefs represented by probability measures, i.e., length-1 LPS (cf. [33]).

Definition 7 *An $(S_i)_{i \in I}$-**based lexicographic type structure** is a structure $\mathcal{T} = \langle S_i, T_i, \beta_i \rangle_{i \in I}$, where*

1. *for each $i \in I$, T_i is a metrizable Lusin space;*
2. *for each $i \in I$, the function $\beta_i : T_i \to \mathcal{N}(S_{-i} \times T_{-i})$ is measurable.*

*We call each space T_i **type space** and we call each β_i **belief map**. Members of type spaces, viz. $t_i \in T_i$, are called **types**. Say $t_i \in T_i$ is a **minimal type** if $\beta_i(t_i) \in \underline{\mathcal{N}}(S_{-i} \times T_{-i})$. Say $t_i \in T_i$ is a **mutually singular** type if $\beta_i(t_i) \in \mathcal{L}(S_{-i} \times T_{-i})$. Say $t_i \in T_i$ is a **full-support** type if $\beta_i(t_i) \in \mathcal{N}^+(S_{-i} \times T_{-i})$. Each element $(s_i, t_i)_{i \in I} \in S \times T$ is called **state (of the world)**.*

The canonical space oh hierarchies $\underline{H} = \Pi_{i \in I} \underline{H}_i$ constructed in the previous section can be endowed with the internal structure of an $(S_i)_{i \in I}$-based type structure. In what follows, we denote by $\mathcal{T}^{\mathrm{mc}} = \langle S_i, \underline{H}_i, \beta_i^{\mathrm{m}} \rangle_{i \in I}$ the type structure associated with $\underline{H}$, and we refer to it as the *minimal canonical type structure.*

Definition 8 *An $(S_i)_{i\in I}$-based lexicographic type structure $\mathcal{T} = \langle S_i, T_i, \beta_i \rangle_{i\in I}$ is*

- ***minimal*** *if each type is minimal;*
- ***LCPS type structure*** *if each type is mutually singular;*
- ***minimal belief-complete*** *if, for each $i \in I$, $\beta_i(T_i) \supseteq \underline{\mathcal{N}}(S_{-i} \times T_{-i})$;*
- ***belief-complete*** *if, for each $i \in I$, $\beta_i(T_i) = \mathcal{N}(S_{-i} \times T_{-i})$.*

The idea of belief-completeness was introduced by Brandenburger ([13]) and adapted to the present context. Note that each type space in a belief-complete type structure has the cardinality of the continuum. It is also notewhorty that

(a) an LCPS type structure is minimal, and

(b) a belief-complete type structure is minimal belief-complete.

The reverse implication in (b) is clearly not true, as type structure $\mathcal{T}^{\mathrm{mc}}$ shows.

5.1 From types to hierarchies

Given a type structure $\mathcal{T} = \langle S_i, T_i, \beta_i \rangle_{i\in I}$, we define for each player $i \in I$ a *m-hierarchy description map* $\mathrm{th}_i : T_i \to \underline{H}_i^0$ associating with each $t_i \in T_i$ a corresponding hierarchy of minimal beliefs. Such map is defined by an inductive procedure which is a slight variant of the one used for belief-hierarchy description maps (cf. [4] and [18]):

- (base step: $k = 1$) For each $i \in I$, let $\mathrm{tx}_{-i}^0 : S_{-i} \times T_{-i} \to \underline{X}_i^0$ be the natural projection. Then, for each $i \in I$, $t_i \in T_i$, the first-order m-hierarchy description map $\mathrm{th}_i^1 : T_i \to \underline{\mathcal{N}}\left(\underline{X}_i^0\right)$ is defined as

$$\begin{aligned} \mathrm{th}_i^1(t_i) &= \mathbf{m}_{\underline{X}_i^0}\left(\widehat{\mathrm{tx}_{-i}^0}(\beta_i(t_i))\right) \\ &= \mathrm{margm}_{S_{-i}}(\beta_i(t_i)). \end{aligned}$$

  For each $i \in I$, define the map $\mathrm{tx}_{-i}^1 : S_{-i} \times T_{-i} \to \underline{X}_i^1 = S_{-i} \times \underline{\mathcal{N}}(S_i)$ as

$$\mathrm{tx}_{-i}^1 = \left(\mathrm{Id}_{S_{-i}}, \mathrm{th}_{-i}^1\right).$$

- (inductive step: $k+1$, $k \geq 1$) Suppose we have already defined, for each $i \in I$, the functions $\mathrm{th}_i^k : T_i \to \underline{\mathcal{N}}\left(\underline{X}_i^{k-1}\right)$ and $\mathrm{tx}_{-i}^k : S_{-i} \times T_{-i} \to \underline{X}_i^k = \underline{X}_i^{k-1} \times \underline{\mathcal{N}}\left(\underline{X}_{-i}^{k-1}\right)$. For each $i \in I$, $t_i \in T_i$, define $\mathrm{th}_i^{k+1} : T_i \to \underline{\mathcal{N}}\left(\underline{X}_i^k\right)$ as

$$\mathrm{th}_i^{k+1}(t_i) = \mathbf{m}_{\underline{X}_i^k}\left(\widehat{\mathrm{tx}_{-i}^k}(\beta_i(t_i))\right);$$

  consequently, the map $\mathrm{tx}_{-i}^{k+1} : S_{-i} \times T_{-i} \to \underline{X}_i^{k+1}$ is defined as

$$\mathrm{tx}_{-i}^{k+1} = \left(\mathrm{tx}_{-i}^k, \mathrm{th}_{-i}^{k+1}\right),$$

  so that $\mathrm{tx}_{-i}^{k+1} = \left(\mathrm{Id}_{S_{-i}}, \mathrm{th}_{-i}^1, ..., \mathrm{th}_{-i}^k, \mathrm{th}_{-i}^{k+1}\right)$.

Thus, for each $i \in I$, the map $\mathrm{th}_i : T_i \to \underline{H}_i^0$ is given by

$$\mathrm{th}_i(t_i) = \left(\mathbf{m}_{\underline{X}_i^{k-1}}\left(\widehat{\mathrm{tx}_{-i}^{k-1}}(\beta_i(t_i))\right)\right)_{k\geq 1}, \; t_i \in T_i,$$

and the map $\mathrm{tx}_{-i} : S_{-i} \times T_{-i} \to S_{-i} \times \underline{H}_{-i}^0$ is defined in a natural manner as $\mathrm{tx}_{-i} = \left(\mathrm{Id}_{S_{-i}}, \mathrm{th}_{-i}\right)$. An easy induction argument shows that both th_i and tx_{-i} are Borel measurable.

5.2 Type morphisms and hierarchy morphisms

Having specified how types induces hierarchies, we now introduce the relevant notions of belief-preserving maps between type structures. In the definitions that follow, we let $T = \Pi_{i\in I} T_i$.

Definition 9 *Fix type structures $\mathcal{T} = \langle S_i, T_i, \beta_i \rangle_{i\in I}$ and $\mathcal{T}' = \langle S_i, T_i', \beta_i' \rangle_{i\in I}$, and, for each $i \in I$, let $\varphi_i : T_i \to T_i'$ be a Borel measurable map.*

1. *The map $(\varphi_i)_{i\in I} : T \to T'$ is called **belief-type morphism (from $\mathcal{T}$ to $\mathcal{T}'$)** if, for each $i \in I$,*
$$\beta_i' \circ \varphi_i = \left(\widehat{\mathrm{Id}_{S_{-i}}, \varphi_{-i}}\right) \circ \beta_i,$$
where $\left(\widehat{\mathrm{Id}_{S_{-i}}, \varphi_{-i}}\right) : \mathcal{N}(S_{-i} \times T_{-i}) \to \mathcal{N}\left(S_{-i} \times T_{-i}'\right)$ is the LPS pushforward map under $\left(\mathrm{Id}_{S_{-i}}, \varphi_{-i}\right) : S_{-i} \times T_{-i} \to S_{-i} \times T_{-i}'$. The map $(\varphi_i)_{i\in I}$ is a belief-type isomophism if it is a Borel isomorphim.

2. *The map $(\varphi_i)_{i\in I} : T \to T'$ is a **minimal belief-hierarchy morphism (from $\mathcal{T}$ to $\mathcal{T}'$)** if*
$$\mathrm{th}_i(t_i) = \mathrm{th}_i'(\varphi_i(t_i)), \; \forall t_i \in T_i, \; \forall i \in I.$$

The notion of belief-type morphism, which is a simple generalization of the related concept for standard type structures (cf. [33]), asks for the belief maps to be preserved. That is, the following diagram commutes:

$$\begin{array}{ccc} T_i & \xrightarrow{\beta_i} & \mathcal{N}(S_{-i} \times T_{-i}) \\ \big\downarrow \varphi_i & & \big\downarrow (\widehat{Id_{S_{-i}}, \varphi_{-i}}) \\ T_i' & \xrightarrow{\beta_i'} & \mathcal{N}(S_{-i} \times T_{-i}') \end{array} \tag{5.1}$$

By constrast, the notion of minimal belief-hierarchy morphism captures the idea that a type structure $\mathcal{T}$ is "contained in" another type structure $\mathcal{T}'$ if $\mathcal{T}$ can be mapped into $\mathcal{T}'$ in a way which preserves the hierarchies of minimal beliefs induced by types.

As one should expect, the important property of belief-type morphisms is that they preserve the explicit description of minimal belief hierarchies.

Proposition 5 *Fix type structures $\mathcal{T} = \langle S_i, T_i, \beta_i \rangle_{i\in I}$ and $\mathcal{T}' = \langle S_i, T_i', \beta_i' \rangle_{i\in I}$, and let $(\varphi_i)_{i\in I} : T \to T'$ be a measurable map. Thus if $(\varphi_i)_{i\in I}$ is a belief-type morphism, then $(\varphi_i)_{i\in I}$ is a minimal belief-hierarchy morphism.*

We illustrate the differences between those two notions of belief-preserving maps by means of the following simple example.

Example 2 *We first show an example on the existence of a minimal belief-hierarchy morphism between two type structures, despite the fact that there is no belief-type morphism between them. Let* $\mathcal{T} = \langle S_i, T_i, \beta_i \rangle_{i \in I}$ *and* $\mathcal{T}' = \langle S_i, T_i', \beta_i' \rangle_{i \in I}$ *be type structures such that, for each* $i \in I = \{1, 2\}$,

$$\begin{aligned} T_i &= \{t_i\}, T_i' = \{t_i'\}, \\ \beta_i(t_i) &= (\nu, \nu) \in \mathcal{N}^2(S_{-i} \times T_{-i}), \\ \beta_i'(t_i') &= (\widehat{\mathrm{Id}_{S_{-i}}, \varphi_{-i}})(\nu) \in \mathcal{M}(S_{-i} \times T_{-i}'), \end{aligned}$$

where, for each $i \in I$, $\varphi_i : T_i \to T_i'$ *is the map defined as* $\varphi_i(t_i) = t_i'$. *Thus* $(\varphi_1, \varphi_2) : T_1 \times T_2 \to T_1' \times T_2'$ *is a minimal belief-hierarchy morphism from* $\mathcal{T}$ *to* $\mathcal{T}'$. *There is no belief-type morphism between* $\mathcal{T}$ *and* $\mathcal{T}'$.

On the other hand, let $\mathcal{T}^* = \langle S_i, T_i^*, \beta_i^* \rangle_{i \in I}$ *and* $\mathcal{T}^{**} = \langle S_i, T_i^{**}, \beta_i^{**} \rangle_{i \in I}$ *be type structures such that, for each* $i \in I = \{1, 2\}$,

$$\begin{aligned} T_i^* &= \{t_i^*\}, T_i^{**} = \{t_i^{**}, z_i^{**}\}, \\ \beta_i^*(t_i^*) &= (\nu), \\ \beta_i^{**}(t_i^{**}) &= (\widehat{\mathrm{Id}_{S_{-i}}, \psi_{-i}})(\nu) \\ \beta_i^{**}(z_i^{**}) &= \left((\widehat{\mathrm{Id}_{S_{-i}}, \phi_{-i}})(\nu), (\widehat{\mathrm{Id}_{S_{-i}}, \phi_{-i}})(\nu)\right), \end{aligned}$$

where, for each $i \in I = \{1, 2\}$,

$$\begin{aligned} \psi_i(t_i^*) &= t_i^{**}, \\ \phi_i(t_i^*) &= z_i^{**}. \end{aligned}$$

It is easily seen that both $(\psi_1, \psi_2) : T_1^* \times T_2^* \to T_1^{**} \times T_2^{**}$ *and* $(\phi_1, \phi_2) : T_1^* \times T_2^* \to T_1^{**} \times T_2^{**}$ *are minimal belief-hierarchy morphisms from* $\mathcal{T}^*$ *to* $\mathcal{T}^{**}$. *The map* (ψ_1, ψ_2) *is also a belief-type morphism.*

All the type structures in Example 2 share the particular feature that the lexicographic beliefs associated with types induce the one and the same minimal belief hierarchy. This leads to the introduction of an important class of type structures, namely type structures satisfying a *non-redundancy* condition. A type structure is minimal-belief-non-redundant if any two distinct types induce distinct minimal belief hierarchies. Formally:

Definition 10 *Fix a type structure* $\mathcal{T} = \langle S_i, T_i, \beta_i \rangle_{i \in I}$.

1. *Say* $\mathcal{T}$ *is* ***minimal-belief-non-redundant*** *(or simply* ***non-redundant****) if, for each* $i \in I$, *the m-hierarchy description map* $\mathrm{th}_i : T_i \to \underline{H}_i^0$ *is injective.*

2. *We say that* $\mathcal{T}$ *is* ***minimal-belief-redundant*** *(or simply* ***redundant****) if it is not non-redundant.*

The structure $\mathcal{T}^{\mathrm{mc}} = \langle S_i, \underline{H}_i, \beta_i^{\mathrm{m}} \rangle_{i \in I}$ is an instance of a non-redundant type structure. On the other hand, type structure $\mathcal{T}^{**}$ in Example 2 is redundant. Specifically, Example 2 shows that, if there is a minimal-belief hierarchy morphism $(\varphi_i)_{i \in I}$ from type structure $\mathcal{T}$ to type structure $\mathcal{T}'$, then $(\varphi_i)_{i \in I}$ need not be unique. The reason why this can occur is that $\mathcal{T}'$ may be redundant (as $\mathcal{T}^{**}$ in Example 2). The following result establishes this fact formally.

Proposition 6 *Let $(\varphi_i)_{i\in I} : T \to T'$ be a minimal belief-hierarchy morphism from $\mathcal{T} = \langle S_i, T_i, \beta_i \rangle_{i\in I}$ to $\mathcal{T}' = \langle S_i, T_i', \beta_i' \rangle_{i\in I}$. If $\mathcal{T}'$ is non-redundant, then $(\varphi_i)_{i\in I}$ is the unique minimal belief-hierarchy morphism.*

5.3 Terminality

The notion of minimal belief-hierarchy morphism is well-suited for analysing an important property of type structures, namely terminality.

Definition 11 *An $(S_i)_{i\in I}$-based type structure $\mathcal{T}' = \langle S_i, T_i', \beta_i' \rangle_{i\in I}$ is **terminal** if for every other $(S_i)_{i\in I}$-based type structure $\mathcal{T} = \langle S_i, T_i, \beta_i \rangle_{i\in I}$ there is a minimal belief-hierarchy morphism from $\mathcal{T}'$ to $\mathcal{T}$.*

The main result of this section shows that $\mathcal{T}^{\mathrm{mc}} = \langle S_i, \underline{H}_i, \beta_i^{\mathrm{m}} \rangle_{i\in I}$ is not only terminal, but the minimal belief-hierarchy morphism from every type structure $\mathcal{T}' = \langle S_i, T_i', \beta_i' \rangle_{i\in I}$ to $\mathcal{T}^{\mathrm{mc}}$ is also unique, and satisfies a property which is analogous to that of a belief-type morphism.

Theorem 1 *Let $\mathcal{T} = \langle S_i, T_i, \beta_i \rangle_{i\in I}$ be an arbitrary $(S_i)_{i\in I}$-based lexicographic type structure, and, for each $i \in I$, let $\mathrm{th}_i : T_i \to \underline{H}_i^0$ be the m-hierarchy description map. Then, for each $i \in I$,*

1. $\mathrm{th}_i(T_i) \subseteq \underline{H}_i$,
2. $(\mathrm{th}_i)_{i\in I}$ *is the unique minimal belief-hierarchy morphism from $\mathcal{T}$ to $\mathcal{T}^{\mathrm{mc}} = \langle S_i, \underline{H}_i, \beta_i^{\mathrm{m}} \rangle_{i\in I}$. Furthermore, it holds that*

$$\beta_i^{\mathrm{m}} \circ \mathrm{th}_i = \mathbf{m}_{S_{-i}\times \underline{H}_{-i}} \circ \left(\widehat{\mathrm{Id}_{S_{-i}}, \mathrm{th}_{-i}}\right) \circ \beta_i,$$

for each $i \in I$.

6 Iterated Admissibility revisited: epistemic foundations

In what follows, we fix a finite game $G = \langle I, (S_i, u_i)_{i\in I} \rangle$, where I is a two-player set and, for every $i \in I$, S_i is the set of strategies with $|S_i| \geq 2$ and $u_i : S_i \times S_{-i} \to \mathbb{R}$ is the payoff function. Each strategy set S_i is given the obvious topology, i.e., the discrete topology. Define the expected payoff function π_i by extending u_i on $\mathcal{M}(S_i) \times \mathcal{M}(S_{-i})$ in the usual way:

$$\pi_i(\sigma_i, \sigma_{-i}) = \sum_{(s_i, s_{-i})\in S_i\times S_{-i}} \sigma_i(s_i)\sigma_{-i}(s_{-i})u_i(s_i, s_{-i}).$$

6.1 Iterated admissibility

The notion of admissible strategy is standard.

Definition 12 *Fix a set $X_i \times X_{-i} \subseteq S_i \times S_{-i}$. A strategy $s_i \in S_i$ is* ***admissible with respect to*** *$X_i \times X_{-i}$ if and only if there exists $\sigma_{-i} \in \mathcal{M}(S_{-i})$ such that $\mathrm{Supp}\sigma_{-i} = X_{-i}$ and $\pi_i(s_i, \sigma_{-i}) \geq \pi_i(s'_i, \sigma_{-i})$ for every $s'_i \in X_i$. If strategy $s_i \in S_i$ is admissible with respect to $S_i \times S_{-i}$, we simply say that s_i is admissible.*

Remark 4 *Fix a set $X_i \times X_{-i} \subseteq S_i \times S_{-i}$. A strategy $s_i \in S_i$ is* ***weakly dominated with respect to*** *$X_i \times X_{-i}$ if there exists $\sigma_i \in \mathcal{M}(S_i)$ with $\sigma_i(X_i) = 1$ such that $\pi_i(\sigma_i, s_{-i}) \geq \pi_i(s_i, s_{-i})$ for every $s_{-i} \in X_{-i}$ and $\pi_i(\sigma_i, s'_{-i}) > \pi_i(s_i, s'_{-i})$ for some $s'_{-i} \in X_{-i}$. A standard result ([45, Lemma 4]) states that a strategy $s_i \in S_i$ is not weakly dominated with respect to $X_i \times X_{-i}$ if and only if it is admissible with respect to $X_i \times X_{-i}$.*

The set of iteratively admissible strategies (henceforth IA set) is defined inductively.

Definition 13 *For each $i \in I$, set $S_i^0 = S_i$ and for every $m \in \mathbb{N}$, let S_i^m be the set of all $s_i \in S_i^{m-1}$ which are admissible w.r.to $S_i^{m-1} \times S_{-i}^{m-1}$. A strategy $s_i \in S_i^m$ is called m-****admissible****. A strategy $s_i \in S_i^\infty = \cap_{m=0}^\infty S_i^m$ is called* ***iteratively admissible****.*

Note that $S_i^m \supseteq S_i^{m+1} \neq \emptyset$ for all $m \in \mathbb{N}$. Moreover, since each strategy set S_i is finite, there exists $M \in \mathbb{N}$ such that $\prod_{i\in I} S_i^\infty = \prod_{i\in I} S_i^M$. Consequently, the IA set $\prod_{i\in I} S_i^\infty$ is non-empty.

6.2 Rationality and Cautiousness

For any two vectors $x = (x_l)_{l=1}^n, y = (y_l)_{l=1}^n \in \mathbb{R}^n$, we write $x \geq_L y$ if either (1) $x_l = y_l$ for every $l \leq n$, or (2) there exists $m \leq n$ such that $x_m > y_m$ and $x_l = y_l$ for every $l < m$. Append to the game G a type structure $\mathcal{T} = \langle S_i, T_i, \beta_i \rangle_{i\in I}$.

Definition 14 . *A strategy $s_i \in S_i$ is* ***optimal under*** *$\beta_i(t_i) = \left(\mu_i^1, ..., \mu_i^n\right) \in \mathcal{N}(S_{-i} \times T_{-i})$ if*

$$\left(\pi_i(s_i, \mathrm{marg}_{S_{-i}}\mu_i^l)\right)_{l=1}^n \geq_L \left(\pi_i(s'_i, \mathrm{marg}_{S_{-i}}\mu_i^l)\right)_{l=1}^n, \ \forall s'_i \in S_i.$$

We say s_i is a ***lexicographic best reply to*** *$\overline{\mathrm{marg}}_{S_{-i}}\beta_i(t_i)$ on S_{-i} if it is optimal under $\beta_i(t_i)$.*

Definition 15 *A type $t_i \in T_i$ is* ***cautious (in $\mathcal{T}$)*** *if $\overline{\mathrm{marg}}_{S_{-i}}\beta_i(t_i) \in \mathcal{N}^+(S_{-i})$. We denote by C_i the set of all $(s_i, t_i) \in S_i \times T_i$ where t_i is cautious.*

Remark 5 *Note that s_i is a lexicographic best reply to $\overline{\mathrm{marg}}_{S_{-i}}\beta_i(t_i)$ if and only if s_i is a lexicographic best reply to* $\mathrm{margm}_{S_{-i}}\beta_i(t_i) \in \underline{\mathcal{N}}(S_{-i})$. *Moreover, by Lemma 1, it follows that*

$$\overline{\mathrm{marg}}_{S_{-i}}\beta_i(t_i) \in \mathcal{N}^+(S_{-i}) \Longleftrightarrow \mathrm{margm}_{S_{-i}}\beta_i(t_i) \in \underline{\mathcal{N}}^+(S_{-i}).$$

Thus $t_i \in T_i$ is cautious (in $\mathcal{T}$) if and only if $\mathrm{margm}_{S_{-i}}\beta_i(t_i) \in \underline{\mathcal{N}}^+(S_{-i})$. *This entails that, if $\beta_i(t_i) \in \mathcal{N}^+(S_{-i} \times T_{-i})$, then* $\mathrm{margm}_{S_{-i}}\beta_i(t_i) \in \underline{\mathcal{N}}^+(S_{-i})$.

Thus, for strategy-type pairs we define the following notions.

Definition 16 *Fix a strategy-type pair $(s_i, t_i) \in S_i \times T_i$.*

1. *Say (s_i, t_i) is* ***rational (in $\mathcal{T}$)*** *if s_i is optimal under $\beta_i(t_i)$. Let R_i be the set of all rational $(s_i, t_i) \in S_i \times T_i$.*
2. *Say (s_i, t_i) is* ***cautiously rational (in $\mathcal{T}$)*** *if it is rational and $(s_i, t_i) \in C_i$.*

The following result is stated and proved in [17].

Proposition 7 *If strategy-type pair $(s_i, t_i) \in S_i \times T_i$ is cautiously rational, then s_i is admissible.*

In [17] it is shown that cautious rationality has an "invariance" property under belief-type morphisms between type structures. A similar property also holds for the case of minimal belef-hierarchy morphisms.

Lemma 8 *Let $\mathcal{T} = \langle S_i, T_i, \beta_i \rangle_{i\in I}$ and $\mathcal{T}^* = \langle S_i, T_i^*, \beta_i^* \rangle_{i\in I}$ be lexicographic type structures, such that there exists a minimal belief-hierarchy morphism $(\varphi_i)_{i\in I} : T \to T^*$ from $\mathcal{T}$ to $\mathcal{T}^*$. Fix a strategy-type pair $(s_i, t_i) \in S_i \times T_i$. Thus*

(i) *$(s_i, t_i) \in C_i$ if and only $(s_i, \varphi_i(t_i)) \in C_i^*$.*

(ii) *$(s_i, t_i) \in R_i$ if and only $(s_i, \varphi_i(t_i)) \in R_i^*$.*

Corollary 2 *A strategy-type pair (s_i, t_i) is cautiously rational in $\mathcal{T}$ if and only $(s_i, \varphi_i(t_i))$ is cautiously rational in $\mathcal{T}^*$.*

6.3 Cautious Belief

The following LPS-based definition of Cautious Belief is taken from [17].

Definition 17 *Fix a type structure $\mathcal{T} = \langle S_i, T_i, \beta_i \rangle_{i \in I}$ and a non-empty event $E \subseteq S_{-i} \times T_{-i}$. Fix also $t_i \in T_i$ with $\beta_i(t_i) = (\mu^1, ..., \mu^n)$. We say that E is* ***cautiously believed under*** *$\beta_i(t_i)$* ***at level*** *$m \leq n$ if the following conditions hold:*

(i) *$\mu^l(E) = 1$ for all $l \leq m$;*

(ii) *for every elementary cylinder $C = \{s_{-i}\} \times T_{-i}$, if $E \cap C \neq \emptyset$ then $\mu^k(E \cap C) > 0$ for some $k \leq m$.*

We say that E is ***cautiously believed under*** *$\beta_i(t_i)$ if it is cautiously believed under $\beta_i(t_i)$ at some level $m \leq n$.*

We say that $t_i \in T_i$ ***cautiously believes*** *E if E is cautiously believed under $\beta_i(t_i)$.*

The notion of Cautious Belief captures the idea that event E and its payoff-relevant components, viz. $E \cap C \neq \emptyset$, are "infinitely more likely than" its complement. A preference-based foundation of Cautious Belief is given in [17], where a comparison with Assumption ([15]) is provided. It turns out that the notion of Cautious Belief does not depend on the specific LPS $\overline{\mu}$ representing the preference relation $\succsim^{\overline{\mu}}$; that is, if the LPS $\overline{\nu}$ is such that $\succsim^{\overline{\nu}} = \succsim^{\overline{\mu}}$, then an event E is cautiously believed in terms of $\succsim^{\overline{\nu}}$ if and only if it is cautiously believed in terms of $\succsim^{\overline{\mu}}$.[11]

Remark 6 *Fix a type structure $\mathcal{T} = \langle S_i, T_i, \beta_i \rangle_{i \in I}$ and a non-empty event $E \subseteq S_{-i} \times T_{-i}$. Fix also $t_i \in T_i$ with $\beta_i(t_i) = \overline{\mu} = (\mu^1, ..., \mu^n)$. Thus E is cautiously believed under $\overline{\mu}$ if and only if it is cautiously believed under $\overline{\nu} \in \mathcal{N}(S_{-i} \times T_{-i})$ such that $\overline{\mu} \cong \overline{\nu}$, e.g., $\overline{\nu} = \mathbf{m}_{S_{-i} \times T_{-i}}(\overline{\mu})$.*

For each player $i \in I$, let $\mathbf{B}_i^c : \Sigma_{S_{-i} \times T_{-i}} \to \Sigma_{S_i \times T_i}$ be the operator defined by

$$\mathbf{B}_i^c(E_{-i}) = \{(s_i, t_i) \in S_i \times T_i \,|t_i \text{ cautiously believes } E_{-i}\}, \; E_{-i} \in \Sigma_{S_{-i} \times T_{-i}}.$$

It is shown in [17] that the set $\mathbf{B}_i^c(E_{-i})$ is Borel in $S_i \times T_i$ for every event $E_{-i} \subseteq S_{-i} \times T_{-i}$; so the operator $\mathbf{B}_i^c : \Sigma_{S_{-i} \times T_{-i}} \to \Sigma_{S_i \times T_i}$ is well-defined.

The Cautious Belief operator $\mathbf{B}_i^c$ has invariance properties under minimal belief hierarchy-morphisms between type structures which are analogous to the ones of (cautious) rationality (cf. Lemma 8 and Corollary 2).

Lemma 9 *Fix type structures $\mathcal{T} = \langle S_i, T_i, \beta_i \rangle_{i \in I}$ and $\mathcal{T}^* = \langle S_i, T_i^*, \beta_i^* \rangle_{i \in I}$. Let $\mathcal{T}^*$ be minimal, and suppose that there is a minimal belief-hierarchy morphism $(\varphi_i)_{i \in I} : T \to T^*$ from $\mathcal{T}$ to $\mathcal{T}^*$ such that, for each $i \in I$,*

$$\beta_i^* \circ \varphi_i = \mathbf{m}_{S_{-i} \times T_{-i}^*} \circ \left(\widehat{\mathrm{Id}_{S_{-i}}, \varphi_{-i}}\right) \circ \beta_i.$$

Let $E_{-i} \subseteq S_{-i} \times T_{-i}$ and $E_{-i}^* \subseteq S_{-i} \times T_{-i}^*$ be non-empty events satisfying the following conditions:

[11] In fact, Definition 17 can be viewed as a representation result of the notion of Cautious Belief in terms of the LPS's representing the preference relation $\succsim$.

1) $\left(\mathrm{Id}_{S_{-i}}, \varphi_{-i}\right)(E_{-i}) \subseteq E^*_{-i}$;

2) $\mathrm{Proj}_{S_{-i}}(E_{-i}) = \mathrm{Proj}_{S_{-i}}\left(E^*_{-i}\right)$.

Then $\left(\mathrm{Id}_{S_i}, \varphi_i\right)\left(\mathbf{B}^c_i(E_{-i})\right) \subseteq \mathbf{B}^c_i\left(E^*_{-i}\right)$.

Lemma 9 is the analogue of Lemma 3 (combined with Remark 4) in [17]. Specifically, [17, Lemma 6] covers the case in which the map $(\varphi_i)_{i\in I}$ is a belief-type morphism from $\mathcal{T}$ to $\mathcal{T}^*$, and $\mathcal{T}^*$ is not required to be minimal. We will make use of Lemma 9 for the proof of Theorem 3 below.

6.4 Common Cautious Belief of Cautious Rationality and Iterated Admissibility

In what follows, fix a type structure $\mathcal{T} = \langle S_i, T_i, \beta_i\rangle_{i\in I}$ and, for each player $i \in I$, let R^1_i be the set of all cautiously rational strategy-type pairs $(s_i, t_i) \in S_i \times T_i$. For each $m > 1$, define R^m_i inductively by

$$R^{m+1}_i = R^m_i \cap \mathbf{B}^c_i\left(R^m_{-i}\right).$$

We write $R^0_i = S_i \times T_i$ and $R^\infty_i = \cap_{m\in\mathbb{N}} R^m_i$ for each $i \in I$. If $(s_i, t_i)_{i\in I} \in \prod_{i\in I} R^{m+1}_i$, we say that there is **cautious rationality and mth-order cautious belief of cautious rationality** at this state. If $(s_i, t_i)_{i\in I} \in \prod_{i\in I} R^\infty_i$, we say that there is **cautious rationality and common cautious belief of cautious rationality** at this state.

Note that, for each $m > 1$,

$$R^{m+1}_i = R^1_i \cap \left(\cap_{l\leq m}\mathbf{B}^c_i\left(R^l_{-i}\right)\right),$$

and each R^m_i is Borel in $S_i \times T_i$ (see [17]).

We now state the main results.

Theorem 2 *Fix a minimal belief-complete type structure $\mathcal{T} = \langle S_i, T_i, \beta_i\rangle_{i\in I}$. Thus, for each $m \geq 0$,*

$$\prod_{i\in I} \mathrm{Proj}_{S_i}\left(R^m_i\right) = \prod_{i\in I} S^m_i.$$

The proof of Theorem 2 is provided in the Appendix.

Theorem 3 *Fix type structure $\mathcal{T}^{\mathrm{mc}} = \langle S_i, \underline{H}_i, \beta^{\mathrm{m}}_i\rangle_{i\in I}$. Thus $\prod_{i\in I} R^\infty_i \neq \emptyset$ and*

$$\prod_{i\in I} \mathrm{Proj}_{S_i}\left(R^\infty_i\right) = \prod_{i\in I} S^\infty_i.$$

The proof of Theorem 3 makes use of the following result.

Lemma 10 *Fix $\mathcal{T}^{\mathrm{mc}} = \langle S_i, \underline{H}_i, \beta^{\mathrm{m}}_i\rangle_{i\in I}$. There exists a finite, minimal type structure $\mathcal{T}^* = \langle S_i, T^*_i, \beta^*_i\rangle_{i\in I}$ such that, for each $i \in I$ and each $m \geq 1$,*
(i) $\mathrm{Proj}_{S_i}\left(R^{*,m}_i\right) = S^m_i$,
(ii) $\left(\mathrm{Id}_{S_i}, \mathrm{th}^*_i\right)\left(R^{*,m}_i\right) \subseteq R^m_i$.

Proof: In [17], it is shown the existence of a finite LCPS (so minimal) type structure $\mathcal{T}^* = \langle S_i, T_i^*, \beta_i^* \rangle_{i\in I}$ such that, for each player $i \in I$,

$$\mathrm{Proj}_{S_i}\left(R_i^{*,k}\right) = S_i^k,\ \forall k \geq 0, \tag{6.1}$$

and $\mathrm{Proj}_{S_i}\left(R_i^{*,\infty}\right) = S_i^\infty$. We show by induction on k that, for each $i \in I$,

$$\left(\mathrm{Id}_{S_i}, \mathrm{th}_i^*\right)\left(R_i^{*,k}\right) \subseteq R_i^k,\ \forall k \geq 1.$$

($k = 1$) If $(s_i, t_i) \in R_i^{*,1}$, then $(s_i, \mathrm{th}_i^*(t_i)) \in R_i^1$ by Lemma 8.

($k \geq 2$) Suppose that the statement is true for $k-1$. Let $(s_i, t_i) \in R_i^{*,k}$. So $(s_i, t_i) \in R_i^{*,k-1}$ and, by the induction hypothesis, $(s_i, \mathrm{th}_i^*(t_i)) \in R_i^{k-1}$. So we need to show that $(s_i, \mathrm{th}_i(t_i)) \in \mathbf{B}_i^c\left(R_{-i}^{k-1}\right)$; this will imply $(s_i, \mathrm{th}_i(t_i)) \in R_i^k$, as required.

Now note that

(a) $\left(\mathrm{Id}_{S_{-i}}, \mathrm{th}_{-i}^*\right)\left(R_{-i}^{*,k-1}\right) \subseteq R_{-i}^{k-1}$,

(b) $\mathrm{Proj}_{S_{-i}}\left(R_{-i}^{*,k-1}\right) = \mathrm{Proj}_{S_{-i}}\left(R_{-i}^{k-1}\right)$.

Part (a) is the induction hypothesis, while part (b) follows from (6.1) and Theorem 2. Since $(s_i, t_i) \in \mathbf{A}_i\left(R_{-i}^{*,k-1}\right)$, then Lemma 9 yields $(s_i, \mathrm{th}_i^*(t_i)) \in \mathbf{B}_i^c\left(R_{-i}^{k-1}\right)$. ■

Proof of Theorem 3: By Lemma 10, there exists a finite, minimal type structure $\mathcal{T}^* = \langle S_i, T_i^*, \beta_i^* \rangle_{i\in I}$ such that, for each $i \in I$ and each $m \geq 1$,

(a) $\mathrm{Proj}_{S_i}\left(R_i^{*,m}\right) = S_i^m$,

(b) $\left(\mathrm{Id}_{S_i}, \mathrm{th}_i^*\right)\left(R_i^{*,m}\right) \subseteq R_i^m$.

Since $(R_i^{*,m})_{m\in\mathbb{N}}$ is a weakly decreasing sequence of finite sets, there exists $N \in \mathbb{N}$ such that $R_i^{*,N} = R_i^{*,\infty}$. So, it follows from (a) that $\mathrm{Proj}_{S_i}\left(R_i^{*,\infty}\right) = S_i^\infty$. Then, for every $s_i \in S_i^\infty$, there exists $t_i \in T_i^*$ such that $(s_i, t_i) \in R_i^{*,m}$ for all $m \in \mathbb{N}$. It thus follows from (b) that $\left(\mathrm{Id}_{S_i}, \mathrm{th}_i^*\right)((s_i, t_i)) \in R_i^m$ for all $m \in \mathbb{N}$. Hence $\left(\mathrm{Id}_{S_i}, \mathrm{th}_i^*\right)((s_i, t_i)) \in R_i^\infty$. Consequently $S_i^\infty \subseteq \mathrm{Proj}_{S_i}\left(R_i^\infty\right) \neq \emptyset$. By Theorem 2, $\mathrm{Proj}_{S_i}\left(R_i^\infty\right) \subseteq S_i^\infty$. The conclusion follows. ■

7 Discussion

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

7.1 Hierarchies of minimal beliefs: the bottom-up approach

The modelling strategy which was used in Section 4.2 to obtain the structure $\mathcal{T}^{\mathrm{mc}} = \langle S_i, \underline{H}_i, \beta_i^{\mathrm{m}} \rangle_{i\in I}$ follows the "top-down" approch of Brandenburger and Dekel [14] (see also [4]) in which hierarchies are first unrestricted and then the coherence requirement is imposed *a posteriori*. An alternative, "bottom-up" approach consists in imposing coherence directly at all levels of the hierarchies. Such approach, first adopted by Mertens and Zamir ([42]; see also [32]) turns out to be equivalent to the "top-down" approach when players' preferences conform to the SEU criterion and there is common belief of this—both approaches lead to two isomorphic type structures. However, such equivalence breaks down in the lexicographic case. We show this by briefly reviewing the "bottom-up" approach; we then argue that the failure of this equivalence rests on the concept of full belief, whose formulation requires the use of an LPS of finite length.

To make this precise, we are going to define inductively the k-order coherent spaces $\widetilde{H}_i^k$, which will consist of coherent k-tuples of minimal length LPS's over Θ_0^i,Θ_1^i,...,Θ_{n-1}^i (the lower-order coherent spaces of the other player and the space of primitive uncertainty). Moreover, each Θ_i^k will be a subset of $\widetilde{H}_i^{k-1} \times \underline{\mathcal{N}}\left(\Theta_i^{k-1}\right)$.

Formally, for each player $i \in I$, let $\Theta_i^0 = S_{-i}$, $\widetilde{H}_i^1 = \underline{\mathcal{N}}\left(\Theta_i^0\right)$, and for all $k \geq 1$,

$$\begin{aligned} \Theta_i^k &= \Theta_i^0 \times \widetilde{H}_{-i}^k; \\ \widetilde{H}_i^{k+1} &= \left\{\left(\overline{\mu}_i^1, ..., \overline{\mu}_i^k, \overline{\mu}_i^{k+1}\right) \in \widetilde{H}_i^k \times \underline{\mathcal{N}}\left(\Theta_i^k\right) : \mathrm{margm}_{\Theta_i^{k-1}} \overline{\mu}_i^{k+1} = \overline{\mu}_i^k\right\}. \end{aligned}$$

The space Θ_i^k is naturally interpreted as player i's domain of uncertainty of order $k+1$; it consists of the space of primitive uncertainty S_{-i} and the set of k-order minimal belief hierarchy of player $-i$. The space $\widetilde{H}_i^{k+1}$ is the set of $(k+1)$-tuples of coherent lexicographic preferences over $\Theta_i^0, ..., \Theta_i^k$. Note carefully that not only each player i's hierarchies are coherent but she also considers only coherent beliefs of $-i$.

In the limit, the set $\widetilde{H}_i$, as defined below, is the set of all coherent hierarchies of minimal beliefs about S_{-i}, about S_{-i} and player $-i$'s minimal beliefs about S_i, and so on. Formally:

$$\begin{aligned} \widetilde{H}_i &= \left\{\left(\overline{\mu}_i^1, \overline{\mu}_i^2, ...\right) \in \prod_{l=0}^{\infty} \underline{\mathcal{N}}\left(\Theta_i^l\right) \middle| \left(\overline{\mu}_i^1, ..., \overline{\mu}_i^k\right) \in \widetilde{H}_i^k, \forall k \geq 1\right\}, \\ \Theta_i &= S_{-i} \times \widetilde{H}_{-i}. \end{aligned}$$

If each $\overline{\mu}_i^k$ is a probability measure (that is, the SEU criterion holds for each player), then standard results ([42], [32]) show the existence of an isomorphism (in fact, a homeomorphism) $f_i : \widetilde{H}_i \to \mathcal{M}\left(\Theta_i\right)$, which is canonical in the sense already discussed. Yet, in the lexicographic case, $\widetilde{H}_i$ and $\underline{\mathcal{N}}\left(\Theta_i\right)$ are *not* isomorphic. There are two reasons for the occurrence of this failure. First, a hierarchy $h_i = \left(\overline{\mu}_i^1, \overline{\mu}_i^2, ...\right) \in \widetilde{H}_i$ may fail to be summarized by a minimal length LPS $\overline{\mu}_i \in \underline{\mathcal{N}}\left(\Theta_i\right)$ if $\overline{\ell}\left(h_i\right) = \infty$. An instance of such a hierarchy is provided in [38].

Second, even though a hierarchy $h_i \in \widetilde{H}_i$ is $\overline{\ell}$-bounded (so that it admits a limit extension over Θ_i), it may fail to have a type representation. To understand how this issue arises, let $\widehat{H}_i^1 \subseteq \widetilde{H}_i$ be the set of $\overline{\ell}$-bounded hierarchies (which can be shown to be measurable in $\widetilde{H}_i$ by using arguments analogous to those in Lemma 5). By using a version of Lemma 6 for projective systems of LPS's (see [18]), it is possible to show the existence of a canonical Borel isomorphism $\phi_i : \widehat{H}_i^1 \to \underline{\mathcal{N}}\left(S_{-i} \times \widetilde{H}_{-i}\right)$. However, the map ϕ_i does not close the model: Even if player i's hierarchy $h_i \in \widetilde{H}_i$ is coherent and $\overline{\ell}$-bounded, $\phi_i\left(h_i\right)$ may deem possible sets of coherent, but *not* $\overline{\ell}$-bounded hierarchies of the other player $-i$. That is, h_i may not induce a preference over $-i$'s preferences over $\widetilde{H}_i$. Thus, to close the model, one needs to proceed exactly as is the "top-down" construction (Section 4.2), by defining inductively, for each $i \in I$, the following sets:

$$\begin{aligned} \widehat{H}_i^{l+1} &= \left\{h_i \in \widehat{H}_i^1 \middle| \phi_i\left(h_i\right)\left(S_{-i} \times \widehat{H}_{-i}^l\right) = \overline{1}\right\}, \ l \geq 1, \\ \widehat{H}_i &= \cap_{l \geq 1} \widehat{H}_i^l. \end{aligned}$$

This leads to a space of hierarchies $\widehat{H} = \Pi_{i \in I} \widehat{H}_i$ which is isomorphic to the canonical hierarchic space $\underline{H}$—and so to a type structure isomorphic to $\mathcal{T}^{\mathrm{mc}}$.

7.2 Hierarchies of lexicographic preferences

Besides non-minimality, there is another way in which different LPS's could represent the same preference relation. This further source of redundat representation is called *Affine Transformation* in [39]. Consider for instance, LPS's $(\nu, \mu), \left(\nu, \frac{1}{2}\nu + \frac{1}{2}\mu\right) \in \mathcal{N}(X)$; clearly, if $\nu \neq \mu$, then

$(\nu, \mu) \cong \left(\nu, \frac{1}{2}\nu + \frac{1}{2}\mu\right)$, and they are both of minimal length. In other words, the set $\underline{\mathcal{N}}(X)$ contains multiple copies of LEU preference representations.

An immediate consequence of this fact is that the canonical type structure $\mathcal{T}^{\mathrm{mc}}$ is not the most parsimonious representation of hierarchies of LEU preferences. A costruction of a non-redundant, canonical space of hierarchies of LEU preferences is still possible (see [38]), and analogues of Theorem 2 and Theorem 3 also hold for this version of the canonical type structure.[12]

7.3 Assumption(s) and unbounded hierarchies

BFK and Dekel et al. [20] show that their version of RCAR cannot be satisfied in any belief-complete and continuous type structure. The basic idea is that a type in such type structures that mth-order assumes rationality must be mapped to an LPS of length greater than $m+1$. Since every LPS in a type structure has finite length, no type can mth-order assume rationality *for all* m. The idea in [40] of using hierarchies that cannot be types is inspired by this intuition. By contrast, with the notion of Cautious Belief in place of Assumption, Theorem 3 shows that a similar issue does not arise.

8 Appendix

8.1 Proof for Section 3

Proof of Proposition 1: Clearly if $\overline{\mu} = \overline{\nu}$ then $\overline{\mu} \cong \overline{\nu}$. On the other hand, suppose that $\overline{\mu} \neq \overline{\nu}$. We need to show that $\overline{\mu} \ncong \overline{\nu}$. Without loss of generality, suppose that $\ell(\overline{\mu}) \geq \ell(\overline{\nu}) = n$. Let $k \leq n+1$ be the least natural number such that $\mu_l = \nu_l$ for each $l < k$, and either (i) $\mu_k \neq \nu_k$ or (ii) $k = n+1$. By mutual singularity, there exists a Borel set E_k such that $\mu_k(E_k) = 1$ and $\mu_l(E_k) = 0$ for all $l < k$. In case (i), there also exist a Borel set F_k such that $\nu_k(F_k) = 1$ and $\nu_l(F_k) = 0$ for all $l < k$, and a Borel set $B_k \subseteq E_k$ such that $\mu_k(B_k) > q > \nu_k(B_k)$ with $q \in (0,1)$. Let $f, g \in$ACT(X) be defined as follows:

$$\begin{aligned} f &= \mathbf{1}_{B_k}, \\ g &= q \cdot \mathbf{1}_{E_k \cup F_k}. \end{aligned}$$

Since $\nu_l = \mu_l$ and $\mu_l(B_k) \leq \mu_l(E_k) = 0$ for each $l < k$, then

$$\int_X f d\nu_l = \int_X f d\mu_l = \mu_l(B_k) = 0, \ \forall l < k.$$

Since $\mu_l = \nu_l$ and $\mu_l(E_k \cup F_k) \leq \mu_l(E_k) + \mu_l(F_k) = \mu_l(E_k) + \nu_l(F_k) = 0$ for each $l < k$, then

$$\int_X g d\nu_l = \int_X g d\mu_l = q \cdot \mu_l(E_k \cup F_k) = 0, \ \forall l < k.$$

Moreover,

$$\int_X f d\mu_k = \mu_k(B_k) > q = q \cdot \mu_k(E_k \cup F_k) = \int_X g d\mu_k.$$

[12] The construction of this "preference non-redundant" version of the canonical type structure requires minimal changes to the construction of $\mathcal{T}^{\mathrm{mc}}$. Details are available on request.

Thus, in case (ii), $f \succ^{\overline{\mu}} g$ but $f \sim^{\overline{\nu}} g$, hence $\overline{\mu} \not\cong \overline{\nu}$. In case (i),

$$\int_X f d\nu_k = \nu_k(B_k) < q = q \cdot \nu_k(E_k \cup F_k) = \int_X g d\nu_k.$$

Thus, $f \succ^{\overline{\mu}} g$ but $g \succ^{\overline{\nu}} f$, hence $\overline{\mu} \not\cong \overline{\nu}$. ■

Proof of Lemma 1: The set containment $\text{Supp}\mathbf{m}_X(\overline{\mu}) \subseteq \text{Supp}\overline{\mu}$ is obvious. On the other hand, suppose there is $x \in X$ such that $x \notin \text{Supp}\mathbf{m}_X(\overline{\mu})$. (Recall that X is a finite discrete space.) Redefine $\mathbf{m}_X(\overline{\mu}) = (v_1, ..., v_k)$. Thus $v_l(\{x\}) = 0$ for every $l \leq k$. It follows that $\mu(\{x\}) = 0$ for every $\mu \in \mathcal{M}(X)$ such that $\mu = \sum_{l=1}^k \alpha_l v_l$ with $\sum_{l=1}^k \alpha_l = 1$ and $\alpha_l \geq 0$. Therefore $x \notin \text{Supp}\overline{\mu}$. ■

Proof of Lemma 2: Measurability follows from [40, Lemma A.4] and the fact that $\widehat{\varphi} : \mathcal{N}(X) \to \mathcal{N}(Y)$ is a Borel map (see [18, Lemma 4.(1)]). Furthermore, if φ is onto, so is $\widehat{\varphi}$ by [18, Lemma 4.(2)]. So $\mathbf{m}_Y \circ \widehat{\varphi}$ is a composition of onto maps, hence onto. ■

Proof of Lemma 3: (1) Suppose that $\overline{\mu}$ is not of minimal length, that is, $\overline{\mu} \notin \underline{\mathcal{N}}(X)$. Thus there exists $\overline{\nu} \in \underline{\mathcal{N}}(X)$ such that $\overline{\mu} \cong \overline{\nu}$ and $\ell(\overline{\mu}) > \ell(\overline{\nu})$. It follows that $\ell(\widehat{\varphi}(\overline{\mu})) > \ell(\widehat{\varphi}(\overline{\nu}))$ but $\widehat{\varphi}(\overline{\mu}) \cong \widehat{\varphi}(\overline{\nu})$. Hence $\widehat{\varphi}(\overline{\mu}) \notin \underline{\mathcal{N}}(Y)$.

(2) If $\widehat{\varphi}(\overline{\mu}) = (\widetilde{\varphi}(\mu_1), ..., \widetilde{\varphi}(\mu_n)) \in \mathcal{L}(Y)$, then for each $l = 1, ..., n$, there are Borel sets E_l in Y such that $\mu_l(\varphi^{-1}(E_l)) = 1$ and $\mu_l(\varphi^{-1}(E_m)) = 0$ for $l \neq m$. Clearly, the sets $\varphi^{-1}(E_1), ..., \varphi^{-1}(E_n) \in \Sigma_X$ satisfy the required properties of mutual singularity for $\overline{\mu}$. ■

Proof of Lemma 4: Since $\widehat{\varphi}$ preserves preference-equivalence between LPS's, we have $\widehat{\varphi}(\mathbf{m}_X(\overline{\mu})) \cong \widehat{\varphi}(\overline{\mu}) \cong \mathbf{m}_Y(\widehat{\varphi}(\overline{\mu}))$, so $\mathbf{m}_Y(\widehat{\varphi}(\overline{\mu})) \cong \widehat{\varphi}(\mathbf{m}_X(\overline{\mu}))$. Then clearly $\ell(\mathbf{m}_Y(\widehat{\varphi}(\overline{\mu}))) \leq \ell(\widehat{\varphi}(\mathbf{m}_X(\overline{\mu})))$. The other equalities and inequalities are obvious. ■

8.2 Proofs for Section 4

Proof of Lemma 5: Fix some $k \geq 1$. We show that the set

$$\underline{H}_i^{1,k} = \left\{ (\overline{\mu}_i^1, \overline{\mu}_i^2, ...) \in \underline{H}_i^0 \,\middle|\, \text{margm}_{\underline{X}_i^{k-1}} \overline{\mu}_i^{k+1} = \overline{\mu}_i^k \right\}$$

is G_δ in $\underline{H}_i^0$. By this, it will follow that an analogous conclusion holds for $\underline{H}_i^1$, since this set can be written as $\underline{H}_i^1 = \cap_{k \geq 1} \underline{H}_i^{1,k}$, i.e., a countable intersection of G_δ-subsets of $\underline{H}_i^0$, so G_δ in $\underline{H}_i^0$.

To this end, first note that the set

$$\begin{aligned} \underline{T}_i^{1,k} &= \left(\text{margm}_{\underline{X}_i^{k-1}}\right)^{-1} \left(\underline{\mathcal{N}}\left(\underline{X}_i^{k-1}\right)\right) \\ &= \left(\mathbf{m}_{\underline{X}_i^{k-1}} \circ \overline{\text{marg}}_{\underline{X}_i^{k-1}}\right)^{-1} \left(\underline{\mathcal{N}}\left(\underline{X}_i^{k-1}\right)\right) \end{aligned}$$

is G_δ in $\underline{\mathcal{N}}(\underline{X}_i^k)$, as it is the inverse image of the open (hence G_δ) set $\underline{\mathcal{N}}\left(\underline{X}_i^{k-1}\right)$ under a Borel class 2 map (see Remark 3). Next, since

$$\begin{aligned} \underline{H}_i^{1,k} &= \prod_{l=0}^{k-1} \underline{\mathcal{N}}\left(\underline{X}_i^l\right) \times \underline{T}_i^{1,k} \times \prod_{l=k+1}^{\infty} \underline{\mathcal{N}}\left(\underline{X}_i^l\right) \\ &= \text{Proj}_{\underline{\mathcal{N}}(\underline{X}_i^k)}^{-1} \left(\underline{T}_i^{1,k}\right), \end{aligned}$$

and $\mathrm{Proj}_{\underline{\mathcal{N}}\left(\underline{X}_i^k\right)} : \underline{H}_i^0 \to \underline{\mathcal{N}}\left(\underline{X}_i^k\right)$ is continuous, it readily follows that $\underline{H}_i^{1,k}$ is G_δ in $\underline{H}_i^0$.

To conclude the proof, note that if each S_i is Polish, so is $\underline{H}_i^0$, and $\underline{H}_i^1$ is a Polish subspace of $\underline{H}_i^0$, since it is a G_δ-subset. ■

Proof of Lemma 6: Part 1: Let n be the length of each $\overline{\nu}^k$, i.e., $\overline{\nu}^k = \left(\nu_1^k, ..., \nu_n^k\right)$. By the consistency condition (4.3), we have that, for each $l \leq n$,

$$\mathrm{marg}_{Z_{k-1}} \nu_l^{k+1} = \nu_l^k, \; \forall k \geq 1.$$

By The Kolmogorov estension theorem, there exists, for each $l \leq n$, a unique probability measure ν_l on Z such that

$$\mathrm{marg}_{Z_{k-1}} \nu_l = \nu_l^k, \; \forall k \geq 1.$$

The LPS $\overline{\nu} = (\nu_1, ..., \nu_n)$ is thus the unique LPS extension of $(\overline{\nu}^k)_{k\geq 1}$.

Part 2: If $\overline{\rho} \cong \overline{\nu}$, then $\mathrm{marg}_{Z_{k-1}} \overline{\rho} \cong \mathrm{marg}_{Z_{k-1}} \overline{\nu}$ for all $k \geq 1$. By the consistency condition, $\overline{\rho}^k \cong \overline{\nu}^k$ for all $k \geq 1$. On the other hand, let $\overline{\rho}^k \cong \overline{\nu}^k$ for all $k \geq 1$. We thus have $\mathbf{m}_{Z_{k-1}}\left(\overline{\rho}^k\right) \cong \mathbf{m}_{Z_{k-1}}\left(\overline{\nu}^k\right)$ for each $k \geq 1$. There exists $K \geq 1$ such that $\ell\left(\mathbf{m}_{Z_{k-1}}\left(\overline{\rho}^k\right)\right) = \ell\left(\mathbf{m}_{Z_{k-1}}\left(\overline{\nu}^k\right)\right)$ for each $k \geq K$. We now construct two sequences of LPS's, viz. $(\overline{\rho}^{*k})_{k\geq 1}$ and $(\overline{\nu}^{*k})_{k\geq 1}$, as follows: For each $k \geq K$, let

$$\begin{aligned} \overline{\rho}^{*k} &= \mathbf{m}_{Z_{k-1}}\left(\overline{\rho}^k\right), \\ \overline{\nu}^{*k} &= \mathbf{m}_{Z_{k-1}}\left(\overline{\nu}^k\right), \end{aligned}$$

while, for each $k < K$, let

$$\begin{aligned} \overline{\rho}^{*k} &= \mathrm{marg}_{Z_{k-1}} \overline{\rho}^{*K}, \\ \overline{\nu}^{*k} &= \mathrm{marg}_{Z_{k-1}} \overline{\nu}^{*K}. \end{aligned}$$

It follows that $\overline{\rho}^k \cong \overline{\rho}^{*k} \cong \overline{\nu}^{*k} \cong \overline{\nu}^k$ for all $k \geq 1$. We let each $\overline{\rho}^{*k}$ (or equivalently $\overline{\nu}^{*k}$) have length $\ell\left(\overline{\rho}^{*k}\right) = n$, viz. $\overline{\rho}^{*k} = \left(\rho_1^{*k}, ..., \rho_n^{*k}\right)$ and $\overline{\nu}^{*k} = \left(\nu_1^{*k}, ..., \nu_n^{*k}\right)$. Fix some $k \geq 1$. Now, for each $m \leq n$, there exists an m-tuple of non-negative scalars $\alpha_1^m, ..., \alpha_m^m \in [0,1]$ such that $\sum_{l=1}^m \alpha_l^m = 1$ and $\rho_m^{*k} = \sum_{l=1}^m \alpha_l^m \nu_l^{*k}$. It turns out that, for each $m \leq n$, $\sum_{l=1}^m \alpha_l^m \nu_l^*$ is the limit extension of the collection of probabilities $(\rho_m^{*k})_{k\geq 1}$. Indeed,

$$\begin{aligned} \mathrm{marg}_{Z_{k-1}} \sum_{l=1}^m \alpha_l^m \nu_l^* &= \sum_{l=1}^m \alpha_l^m \mathrm{marg}_{Z_{k-1}} \nu_l^* \\ &= \sum_{l=1}^m \alpha_l^m \nu_l^{*k} = \rho_m^{*k}. \end{aligned}$$

By the Kolmogorov theorem, such extension $\rho_m^* = \sum_{l=1}^m \alpha_l^m \nu_l^*$ is unique. Thus $\overline{\rho}^* = (\rho_1^*, ..., \rho_n^*)$ is such that $\overline{\rho}^* \cong \overline{\nu}^*$. The limit extensions $\overline{\rho}$ and $\overline{\nu}$ satisfy $\overline{\rho} \cong \overline{\nu}$, because $\overline{\rho}^* = \mathbf{m}_Z\left(\overline{\rho}\right)$ and $\overline{\nu}^* = \mathbf{m}_Z\left(\overline{\nu}\right)$.

Part 3: The sufficiency part is immediate: Since $\overline{\mathrm{marg}}_{Z_{k'}} \overline{\nu} = \overline{\nu}^{k'+1}$ and $\overline{\nu}^{k'+1} \in \underline{\mathcal{N}}\left(Z_{k'}\right)$, then $\overline{\nu} \in \underline{\mathcal{N}}\left(Z\right)$ by Lemma 3. We prove the necessity part by contraposition. Suppose that $\overline{\nu}^k \notin \underline{\mathcal{N}}\left(Z_{k-1}\right)$ for all $k \geq 1$. Let $\overline{\mu}^k \in \underline{\mathcal{N}}\left(Z_{k-1}\right)$ be a minimal-length LPS equivalent to each $\overline{\nu}^k$, that is, $\overline{\mu}^k \cong \overline{\nu}^k$ for all $k \geq 1$. Using the consistency condition (4.3) we get that, for all $k \geq 1$,

$$\overline{\mu}^k \cong \overline{\nu}^k = \overline{\mathrm{marg}}_{Z_{k-1}} \overline{\nu}^{k+1} \cong \overline{\mathrm{marg}}_{Z_{k-1}} \overline{\mu}^{k+1},$$

so that

$$\overline{\text{marg}}_{Z_{k-1}}\overline{\mu}^{k+1} \cong \overline{\mu}^{k},\ \forall k \geq 1.$$

It follows that, for all $k \geq 1$,

$$\begin{aligned}
\ell\left(\overline{\mu}^{k}\right) &\leq \ell\left(\overline{\text{marg}}_{Z_{k-1}}\overline{\mu}^{k+1}\right) \\
&= \ell\left(\overline{\mu}^{k+1}\right) \\
&< \ell\left(\overline{\nu}^{k+1}\right) \\
&= \ell\left(\overline{\nu}\right),
\end{aligned}$$

where the last equality follows from Part 1. Let $\ell^{*} = \sup\left\{\ell\left(\overline{\mu}^{k}\right)|k \geq 1\right\}$. Since $\ell^{*} < \ell\left(\overline{\nu}\right)$ and the sequence $(\ell\left(\overline{\mu}^{k}\right))_{k\geq 1}$ is non-decreasing, then there exists $k' \geq 1$ such that $\ell\left(\overline{\mu}^{k}\right) = \ell^{*}$ for all $k \geq k'$. Let $(\overline{\rho}^{k})_{k\geq 1}$ be another collection of LPS's on Z_k's satisfying

$$\begin{aligned}
\overline{\rho}^{k} &= \overline{\mu}^{k},\ \forall k \geq k', \\
\overline{\text{marg}}_{Z_{k-2}}\overline{\rho}^{k} &= \overline{\rho}^{k-1},\ \forall k < k'.
\end{aligned}$$

It readily follows that, for all $k \geq 1$, $\overline{\rho}^{k} \cong \overline{\nu}^{k}$ and $\ell\left(\overline{\rho}^{k}\right) < \ell\left(\overline{\nu}\right)$. By Part 2, the LPS extension $\overline{\rho}$ of $(\overline{\rho}^{k})_{k\geq 1}$ satisfies $\overline{\rho} \cong \overline{\nu}$ and $\ell\left(\overline{\rho}\right) < \ell\left(\overline{\nu}\right)$. Hence $\overline{\nu} \notin \underline{\mathcal{N}}\left(Z\right)$, as required. ■

Proof of Lemma 7: By [18, Lemma 9], there exists a homeomorphism, viz. ψ, between $\left\{\overline{\mu} \in \mathcal{N}\left(X\right)\middle|\overline{\mu}\left(E\right) = \overline{1}\right\}$ and $\mathcal{N}\left(E\right)$. Thus, we have

$$\mathbf{m}_{E}\left(\psi\left(\left\{\overline{\mu} \in \mathcal{N}\left(X\right)\middle|\overline{\mu}\left(E\right) = \overline{1}\right\}\right)\right) = \mathbf{m}_{E}\left(\mathcal{N}\left(E\right)\right) = \underline{\mathcal{N}}\left(E\right).$$

Using the fact that $\left(\psi\left(\overline{\mu}\right)\right)_{l}\left(B\right) = \overline{\mu}_{l}(B)$ for each $\overline{\mu} = (\mu_1, ..., \mu_n) \in \mathcal{N}\left(X\right)$, $l \leq n$, and Borel $B \subseteq E$, it is immediate to show that $\left(\psi \circ \mathbf{m}_{X}\right)\left(\overline{\mu}\right) = \left(\mathbf{m}_{E} \circ \psi\right)\left(\overline{\mu}\right)$. Then

$$\begin{aligned}
\mathbf{m}_{E}\left(\psi\left(\left\{\overline{\mu} \in \mathcal{N}\left(X\right)\middle|\overline{\mu}\left(E\right) = \overline{1}\right\}\right)\right) &= \psi\left(\mathbf{m}_{X}\left(\left\{\overline{\mu} \in \mathcal{N}\left(X\right)\middle|\overline{\mu}\left(E\right) = \overline{1}\right\}\right)\right) \\
&= \psi\left(\left\{\overline{\mu} \in \underline{\mathcal{N}}\left(X\right)\middle|\overline{\mu}\left(E\right) = \overline{1}\right\}\right).
\end{aligned}$$

Thus, we can conclude

$$\psi\left(\left\{\overline{\mu} \in \underline{\mathcal{N}}\left(X\right)\middle|\overline{\mu}\left(E\right) = \overline{1}\right\}\right) = \underline{\mathcal{N}}\left(E\right).$$

This implies that the restriction of the homeomorphism ψ to $\left\{\overline{\mu} \in \underline{\mathcal{N}}\left(X\right)\middle|\overline{\mu}\left(E\right) = \overline{1}\right\}$ is a homeomorphism onto $\underline{\mathcal{N}}\left(E\right)$. ■

8.3 Proofs for Section 5

Proof of Proposition 5: We will show that for each $i \in I$ and $t_i \in T_i$,

$$\text{th}_{i}^{k}\left(t_{i}\right) = \text{th}_{i}^{\prime k}\left(\varphi_{i}\left(t_{i}\right)\right),\ \forall k \geq 1;$$

this will imply the thesis. The proof is by induction on k.

($k = 1$) Fix $i \in I$ and $t_i \in T_i$. By definition

$$\begin{aligned}
\text{th}_{i}^{1}\left(t_{i}\right) &= \text{margm}_{S_{-i}}\left(\beta_{i}\left(t_{i}\right)\right), \\
\text{th}_{i}^{\prime 1}\left(\varphi_{i}\left(t_{i}\right)\right) &= \text{margm}_{S_{-i}}\left(\beta_{i}^{\prime}\left(\varphi_{i}\left(t_{i}\right)\right)\right);
\end{aligned}$$

so, it is enough to show that

$$\beta_i'(\varphi_i(t_i))\left(E\times T_{-i}'\right)=\beta_i'(t_i)(E\times T_{-i}),$$

for every Borel $E\subseteq S_{-i}$. But this follows from the fact that $(\varphi_i)_{i\in I}$ is a belief-type morphism; indeed,

$$\begin{aligned}\beta_i'(\varphi_i(t_i))\left(E\times T_{-i}'\right) &= \beta_i(t_i)\left(\left(\mathrm{Id}_{S_{-i}},\varphi_{-i}\right)^{-1}\left(E\times T_{-i}'\right)\right)\\ &= \beta_i(t_i)\left(\left\{(s,t_{-i}):\left(s,\varphi_{-i}(t_{-i})\right)\in E\times T_{-i}'\right\}\right)\\ &= \beta_i(t_i)(E\times T_{-i}),\end{aligned}$$

as required.

($k\geq 2$) Suppose that the statement is true up to k; this implies that $\mathrm{tx}_{-i}^k=\mathrm{tx}_{-i}'^k\circ\left(\mathrm{Id}_{S_{-i}},\varphi_{-i}\right)$ for each $i\in I$.

Fix $i\in I$ and $t_i\in T_i$, and pick any $E_k\in\Sigma_{\underline{X}_i^k}$. We get

$$\begin{aligned}\mathrm{th}_i^{k+1}(t_i)(E_k) &= \mathbf{m}_{\underline{X}_i^k}\left(\widehat{\mathrm{tx}_{-i}^k}(\beta_i(t_i))\right)(E_k)\\ &= \mathbf{m}_{\underline{X}_i^k}\left(\left(\widehat{\mathrm{Id}_{S_{-i}},\varphi_{-i}}\right)(\beta_i(t_i))\right)\left(\mathrm{tx}_{-i}'^k\right)^{-1}(E_k)\\ &= \mathbf{m}_{\underline{X}_i^k}\left(\beta_i'(\varphi_i(t_i))\right)\left(\mathrm{tx}_{-i}'^k\right)^{-1}(E_k)\\ &= \mathbf{m}_{\underline{X}_i^k}\left(\widehat{\mathrm{tx}_{-i}'^k}\left(\beta_i'(\varphi_i(t_i))\right)\right)(E_k)\\ &= \mathrm{th}_i'^{k+1}(\varphi_i(t_i))(E_k)\end{aligned}$$

where the first equality is by definition of th_i^{k+1}, the second equality follows from the implication of the induction hypothesis, the third and the fourth equalities follow from the property that $(\varphi_i)_{i\in I}$ is a belief-type morphism, and the last equality is again by definition of $\mathrm{th}_i'^{k+1}$. This proves that the statement is true for $k+1$, concluding the proof. ■

Proof of Proposition 6: Since $\mathcal{T}'$ is non-redundant, then the map $\mathrm{th}_i':T_i'\longrightarrow\underline{H}_i^0$ is injective. Thus, there exists a unique map $\varphi_i:T_i\longrightarrow T_i'$ such that $\mathrm{th}_i'\circ\varphi_i=\mathrm{th}_i$, namely $\varphi_i=(\mathrm{th}_i')^{-1}\circ\mathrm{th}_i$. ■

Proof of Theorem 1: The proof is divided into two main steps. In the first step, we show that for each $t_i\in T_i$, the corresponding *m-hierarchy description* $\mathrm{th}_i(t_i)$ belongs to $\underline{H}_i$. In the second step, we show that the map $(\mathrm{th}_i)_{i\in I}$ is the unique preference-type morphism. In both cases, the proof is by induction.

First step: $\mathrm{th}_i(T_i)\subseteq\underline{H}_i$. By definition of m-hierarchy description, $\mathrm{th}_i(T_i)\subseteq\underline{H}_i^0$. We use induction to prove $\mathrm{th}_i(T_i)\subseteq\underline{H}_i$.

(*Base step*): We show that $\mathrm{th}_i(T_i)\subseteq\underline{H}_i^1$, so we need to verify that for all $t_i\in T_i$ and all $k\geq 1$,

$$\mathrm{margm}_{\underline{X}_i^{k-1}}\left(\mathrm{th}_i^{k+1}(t_i)\right)=\mathrm{th}_i^k(t_i);$$

to this end, we first show that

$$\overline{\mathrm{marg}}_{\underline{X}_i^{k-1}}\left(\widehat{\mathrm{tx}}_{-i}^k(\beta_i(t_i))\right)=\widehat{\mathrm{tx}}_{-i}^{k-1}(\beta_i(t_i)).\tag{8.1}$$

To see this, pick any event $E_{k-1} \in \Sigma_{\underline{X}_i^{k-1}}$. Thus

$$\begin{aligned}
\overline{\mathrm{marg}}_{\underline{X}_i^{k-1}} \left(\widehat{\mathrm{tx}}_{-i}^{k} \left(\beta_i \left(t_i\right)\right)\right) \left(E_{k-1}\right) &= \widehat{\mathrm{tx}}_{-i}^{k} \left(\beta_i \left(t_i\right)\right) \left(E_{k-1} \times \underline{\mathcal{N}} \left(\underline{X}_{-i}^{k-1}\right)\right) \\
&= \left(\left(\widehat{\mathrm{tx}_{-i}^{k-1}, \mathrm{th}_{-i}^{k}}\right) \left(\beta_i \left(t_i\right)\right)\right) \left(E_{k-1} \times \underline{\mathcal{N}} \left(\underline{X}_{-i}^{k-1}\right)\right) \\
&= \beta_i \left(t_i\right) \left(\left\{\left(s_{-i}, t_{-i}\right) \middle| \left(\mathrm{tx}_{-i}^{k-1}, \mathrm{th}_{-i}^{k}\right) \left(\left(s_{-i}, t_{-i}\right)\right) \in E_{k-1} \times \underline{\mathcal{N}} \left(\underline{X}_{-i}^{k-1}\right) \right\}\right) \\
&= \beta_i \left(t_i\right) \left(\left\{\left(s_{-i}, t_{-i}\right) \middle| \begin{array}{c} \mathrm{tx}_{-i}^{k-1} \left(\left(s_{-i}, t_{-i}\right)\right) \in E_{k-1}, \\ \mathrm{th}_{-i}^{k} \left(\left(s_{-i}, t_{-i}\right)\right) \in \underline{\mathcal{N}} \left(\underline{X}_{-i}^{k-1}\right) \end{array} \right\}\right) \\
&= \beta_i \left(t_i\right) \left(\left\{\left(s_{-i}, t_{-i}\right) \middle| \mathrm{tx}_{-i}^{k-1} \left(\left(s_{-i}, t_{-i}\right)\right) \in E_{k-1} \right\}\right) \\
&= \beta_i \left(t_i\right) \left(\left(\mathrm{tx}_{-i}^{k-1}\right)^{-1} \left(E_{k-1}\right)\right) \\
&= \widehat{\mathrm{tx}}_{-i}^{k-1} \left(\beta_i \left(t_i\right)\right) \left(E_{k-1}\right),
\end{aligned}$$

which shows that Eq. (8.1) holds. So (8.1) implies

$$\begin{aligned}
\mathrm{margm}_{\underline{X}_i^{k-1}} \left(\widehat{\mathrm{tx}}_{-i}^{k} \left(\beta_i \left(t_i\right)\right)\right) &= \mathbf{m}_{\underline{X}_i^{k-1}} \left(\overline{\mathrm{marg}}_{\underline{X}_i^{k-1}} \left(\widehat{\mathrm{tx}}_{-i}^{k} \left(\beta_i \left(t_i\right)\right)\right)\right) \\
&= \mathbf{m}_{\underline{X}_i^{k-1}} \left(\widehat{\mathrm{tx}}_{-i}^{k-1} \left(\beta_i \left(t_i\right)\right)\right) \\
&= \mathrm{th}_i^{k} \left(t_i\right).
\end{aligned}$$

By this, we get

$$\begin{aligned}
\mathrm{margm}_{\underline{X}_i^{k-1}} \left(\mathrm{th}_i^{k+1} \left(t_i\right)\right) &= \mathrm{margm}_{\underline{X}_i^{k-1}} \left(\mathbf{m}_{\underline{X}_i^{k}} \left(\widehat{\mathrm{tx}}_{-i}^{k} \left(\beta_i \left(t_i\right)\right)\right)\right) \\
&= \mathbf{m}_{\underline{X}_i^{k-1}} \left(\overline{\mathrm{marg}}_{\underline{X}_i^{k-1}} \left(\mathbf{m}_{\underline{X}_i^{k}} \left(\widehat{\mathrm{tx}}_{-i}^{k} \left(\beta_i \left(t_i\right)\right)\right)\right)\right) \\
&= \mathbf{m}_{\underline{X}_i^{k-1}} \left(\overline{\mathrm{marg}}_{\underline{X}_i^{k-1}} \left(\widehat{\mathrm{tx}}_{-i}^{k} \left(\beta_i \left(t_i\right)\right)\right)\right) \\
&= \mathrm{margm}_{\underline{X}_i^{k-1}} \left(\widehat{\mathrm{tx}}_{-i}^{k} \left(\beta_i \left(t_i\right)\right)\right) \\
&= \mathrm{th}_i^{k} \left(t_i\right),
\end{aligned}$$

where the third equality follows from Remark 2. This concludes the proof of the base step.

To go on, we need the following

Claim 1 *For each $i \in I$, let f_i be the isomorphism of Proposition 3. Thus, the following diagram commutes:*

$$\begin{tikzcd}
T_i \arrow[r, "\beta_i"] \arrow[d, "\mathrm{th}_i"] & \mathcal{N}\left(S_{-i} \times T_{-i}\right) \arrow[d, "\mathbf{m}_{S_{-i} \times \underline{H}_{-i}^{0}} \circ \left(\widehat{\mathrm{Id}_{S_{-i}}, \mathrm{th}_{-i}}\right)"] \\
\underline{H}_i^{1} \arrow[r, "f_i"] & \underline{\mathcal{N}}\left(S_{-i} \times \underline{H}_{-i}^{0}\right)
\end{tikzcd} \quad . \qquad (8.2)$$

Proof of Claim: We will show that for each $k \geq 0$,

$$\operatorname{margm}_{\underline{X}_i^k}\left(\mathbf{m}_{S_{-i}\times \underline{H}_{-i}^0}\left(\widehat{\left(\mathrm{Id}_{S_{-i}}, \mathrm{th}_{-i}\right)}\left(\beta_i\left(t_i\right)\right)\right)\right) = \mathrm{th}_i^{k+1}\left(t_i\right).$$

Then, by uniqueness of f_i, it must hold that

$$f_i\left(\mathrm{th}_i\left(t_i\right)\right) = \mathbf{m}_{S_{-i}\times \underline{H}_{-i}^0}\left(\widehat{\left(\mathrm{Id}_{S_{-i}}, \mathrm{th}_{-i}\right)}\left(\beta_i\left(t_i\right)\right)\right).$$

We have

$$\begin{aligned}\operatorname{margm}_{\underline{X}_i^k}\left(\mathbf{m}_{S_{-i}\times \underline{H}_{-i}^0}\left(\widehat{\left(\mathrm{Id}_{S_{-i}}, \mathrm{th}_{-i}\right)}\left(\beta_i\left(t_i\right)\right)\right)\right) &= \operatorname{margm}_{\underline{X}_i^k}\left(\mathbf{m}_{S_{-i}\times \underline{H}_{-i}^0}\left(\widehat{\mathrm{tx}}_{-i}\left(\beta_i\left(t_i\right)\right)\right)\right) \\ &= \operatorname{margm}_{\underline{X}_i^k}\left(\widehat{\mathrm{tx}}_{-i}\left(\beta_i\left(t_i\right)\right)\right) \\ &= \mathbf{m}_{\underline{X}_i^k}\left(\overline{\operatorname{marg}}_{\underline{X}_i^k}\left(\widehat{\mathrm{tx}}_{-i}\left(\beta_i\left(t_i\right)\right)\right)\right),\end{aligned}$$

where the second equality follows from Remark 2. On the other hand,

$$\begin{aligned}\mathrm{th}_i^{k+1}\left(t_i\right) &= \mathbf{m}_{\underline{X}_i^k}\left(\mathrm{th}_i^{k+1}\left(t_i\right)\right) \\ &= \mathbf{m}_{\underline{X}_i^k}\left(\widehat{\mathrm{tx}}_{-i}^k\left(\beta_i\left(t_i\right)\right)\right).\end{aligned}$$

Therefore, we need to check that

$$\overline{\operatorname{marg}}_{\underline{X}_i^k}\left(\widehat{\mathrm{tx}}_{-i}\left(\beta_i\left(t_i\right)\right)\right) = \widehat{\mathrm{tx}}_{-i}^k\left(\beta_i\left(t_i\right)\right).$$

Fix $E_k \in \Sigma_{\underline{X}_i^k}$. We have

$$\begin{aligned}\widehat{\mathrm{tx}}_{-i}^k\left(\beta_i\left(t_i\right)\right)\left(F_k\right) &= \beta_i\left(t_i\right)\left(\left(\mathrm{tx}_{-i}^k\right)^{-1}\left(E_k\right)\right) \\ &= \beta_i\left(t_i\right)\left(\left\{\left(s_{-i}, t_{-i}\right)\middle|\left(s_{-i}, \mathrm{th}_{-i}^1\left(t_{-i}\right), \ldots, \mathrm{th}_{-i}^k\left(t_{-i}\right)\right) \in E_k\right\}\right) \\ &= \beta_i\left(t_i\right)\left(\left\{\left(s_{-i}, t_{-i}\right)\middle|\left(s_{-i}, \mathrm{th}_{-i}\left(t_{-i}\right)\right) \in \mathrm{Proj}_{\underline{X}_i^k}^{-1}\left(E_k\right)\right\}\right) \\ &= \beta_i\left(t_i\right)\left(\left(\mathrm{Id}_{S_{-i}}, \mathrm{th}_{-i}\right)^{-1}\left(\mathrm{Proj}_{\underline{X}_i^k}^{-1}\left(E_k\right)\right)\right) \\ &= \widehat{\left(\mathrm{Id}_{S_{-i}}, \mathrm{th}_{-i}\right)}\left(\beta_i\left(t_i\right)\right)\left(\mathrm{Proj}_{\underline{X}_i^k}^{-1}\left(E_k\right)\right) \\ &= \widehat{\mathrm{tx}}_{-i}\left(\beta_i\left(t_i\right)\right)\left(\mathrm{Proj}_{\underline{X}_i^k}^{-1}\left(E_k\right)\right) \\ &= \overline{\operatorname{marg}}_{\underline{X}_i^k}\left(\widehat{\mathrm{tx}_{-i}}\left(\beta_i\left(t_i\right)\right)\right)\left(E_k\right).\end{aligned}$$

■

(*Inductive step*): Recall that $\mathrm{th}_i\left(t_i\right) \in \underline{H}_i^l$, $l \geq 2$, if and only if $f_i\left(\mathrm{th}_i\left(t_i\right)\right)\left(S_{-i} \times \underline{H}_{-i}^{l-1}\right) = \overline{1}$, for each $t_i \in T_i$. Suppose that, for each player $i \in I$, $\mathrm{th}_i\left(T_i\right) \subseteq \underline{H}_i^{l-1}$. Thus, by the commutativity of Diagram (8.2), it holds that, for all $t_i \in T_i$,

$$f_i\left(\mathrm{th}_i\left(t_i\right)\right)\left(S_{-i} \times \underline{H}_{-i}^{l-1}\right) = \mathbf{m}_{S_{-i}\times \underline{H}_{-i}^0}\left(\widehat{\left(\mathrm{Id}_{S_{-i}}, \mathrm{th}_{-i}\right)}\left(\beta_i\left(t_i\right)\right)\right)\left(S_{-i} \times \underline{H}_{-i}^{l-1}\right).$$

We show that $f_i\left(\mathrm{th}_i\left(t_i\right)\right)\left(S_{-i} \times \underline{H}_{-i}^{l-1}\right) = \overline{1}$, which implies $\mathrm{th}_i\left(t_i\right) \subseteq \underline{H}_i^l$. To this end, note that

$$\begin{aligned}\widehat{\left(\mathrm{Id}_{S_{-i}}, \mathrm{th}_{-i}\right)}\left(\beta_i\left(t_i\right)\right)\left(S_{-i} \times \underline{H}_{-i}^{l-1}\right) &= \beta_i\left(t_i\right)\left(\left(\mathrm{Id}_{S_{-i}}, \mathrm{th}_{-i}\right)^{-1}\left(S_{-i} \times \underline{H}_{-i}^{l-1}\right)\right) \\ &= \beta_i\left(t_i\right)\left(\left\{\left(s_{-i}, t_{-i}\right) : \mathrm{th}_{-i}\left(t_{-i}\right) \in \underline{H}_{-i}^{l-1}\right\}\right) \\ &= \beta_i\left(t_i\right)\left(S_{-i} \times T_{-i}\right) \\ &= \overline{1};\end{aligned}$$

it follows that

$$\mathbf{m}_{S_{-i}\times \underline{H}^{0}_{-i}}\left(\left(\widehat{\mathrm{Id}_{S_{-i}},\mathrm{th}_{-i}}\right)\left(\beta_i\left(t_i\right)\right)\right)\left(S_{-i}\times \underline{H}^{l-1}_{-i}\right)=\overline{1}.$$

Thus $f_i\left(\mathrm{th}_i\left(t_i\right)\right)\left(S_{-i}\times \underline{H}^{l-1}_{-i}\right)=\overline{1}$, as required.

Second step: $\left(\mathrm{th}_i\right)_{i\in I}$ *is the unique minimal belief-hierarchy morphism from* $\mathcal{T}$ *to* $\mathcal{T}^{\mathrm{mc}}=\langle S_i,\underline{H}_i,\beta^{\mathrm{m}}_i\rangle_{i\in I}$. Measurability of $\left(\mathrm{th}_i\right)_{i\in I}$ follows from an easy induction argument. Clearly, $\left(\mathrm{th}_i\right)_{i\in I}$ is a minimal belief-hierarchy morphism from $\mathcal{T}$ to $\mathcal{T}^{\mathrm{mc}}$ and note that, since $\mathrm{th}_i\left(T_i\right)\subseteq \underline{H}_i$ for each $i\in I$ (as proved in the first step), and each map β^{m}_i is a Borel isomorphism, it follows from Diagram (8.2) that

$$\beta^{\mathrm{m}}_i\circ \mathrm{th}_i=\mathbf{m}_{S_{-i}\times \underline{H}_{-i}}\circ\left(\widehat{\mathrm{Id}_{S_{-i}},\mathrm{th}_{-i}}\right)\circ\beta_i.$$

Finally, uniqueness of $\left(\mathrm{th}_i\right)_{i\in I}$ follows from Proposition 6, since $\mathcal{T}^{\mathrm{mc}}$ is non-redundant. ■

8.4 Proofs for Section 6

Proof of Lemma 8: Part (i): Fix an arbitrary $\left(s_i,t_i\right)\in S_i\times T_i$. By definition of minimal belief-hierarchy morphism,

$$\mathrm{margm}_{S_{-i}}\beta_i\left(t_i\right)=\mathrm{margm}_{S_{-i}}\beta^{*}_i\left(\varphi_i\left(t_i\right)\right).$$

Therefore $\mathrm{margm}_{S_{-i}}\beta_i\left(t_i\right)\in\mathcal{N}^{+}\left(S_{-i}\right)$ if and only if $\mathrm{margm}_{S_{-i}}\beta^{*}_i\left(\varphi_i\left(t_i\right)\right)\in\mathcal{N}^{+}\left(S_{-i}\right)$. It follows from Remark 5 that $\left(s_i,t_i\right)\in C_i$ if and only $\left(s_i,\varphi_i\left(t_i\right)\right)\in C^{*}_i$.

Part (ii): Using the same argument as for Part (i), the result follows from Remark 5. ■

Proof of Lemma 9: Let $\left(s_i,t_i\right)\in\mathbf{B}^{c}_i\left(E_{-i}\right)$, and set $\varphi_i\left(t_i\right)=t^{*}_i$. We have to show that event E^{*}_{-i} is cautiously believed under $\beta^{*}_i\left(t^{*}_i\right)$, that is, conditions (i) and (ii) of Definition 17 are satisfied. To this end, we first show that E^{*}_{-i} is cautiously believed under $\left(\widehat{\mathrm{Id}_{S_{-i}},\varphi_{-i}}\right)\left(\beta_i\left(t_i\right)\right)\in\mathcal{N}\left(S_{-i}\times T^{*}_{-i}\right)$.

Since event E_{-i} is cautiously believed under $\beta_i\left(t_i\right)=\left(\beta^{1}_i\left(t_i\right),...,\beta^{n}_i\left(t_i\right)\right)$, then there exists $m\leq n$ such that $\beta^{l}_i\left(t_i\right)\left(E_{-i}\right)=1$ for all $l\leq m$. Next note that

$$E_{-i}\subseteq\left(\mathrm{Id}_{S_{-i}},\varphi_{-i}\right)^{-1}\left(\left(\mathrm{Id}_{S_{-i}},\varphi_{-i}\right)\left(E_{-i}\right)\right)\subseteq\left(\mathrm{Id}_{S_{-i}},\varphi_{-i}\right)^{-1}\left(E^{*}_{-i}\right),$$

where the first set containment is obvious, while the second one follows from condition 1). Hence, for all $l\leq n$,

$$\beta^{l}_i\left(t_i\right)\left(E_{-i}\right)\leq\beta^{l}_i\left(t_i\right)\left(\left(\mathrm{Id}_{S_{-i}},\varphi_{-i}\right)^{-1}\left(E^{*}_{-i}\right)\right),$$

which implies $\beta^{l}_i\left(t_i\right)\left(\left(\mathrm{Id}_{S_{-i}},\varphi_{-i}\right)^{-1}\left(E^{*}_{-i}\right)\right)=1$ for all $l\leq m$, so that condition (i) of Definition 17 is satisfied.

To show that Condition (ii) of Definition 17 is also satisfied, we proceed as follows. Consider an elementary cylinder $C=\{s_{-i}\}\times T^{*}_{-i}$ satisfying $E^{*}_{-i}\cap C\neq\emptyset$. It turns out that

$$\begin{aligned}\left(\mathrm{Id}_{S_{-i}},\varphi_{-i}\right)^{-1}\left(C\cap E^{*}_{-i}\right) &= \left(\left(\mathrm{Id}_{S_{-i}},\varphi_{-i}\right)^{-1}\left(C\right)\right)\cap\left(\left(\mathrm{Id}_{S_{-i}},\varphi_{-i}\right)^{-1}\left(E^{*}_{-i}\right)\right)\\ &= \left(\{s_{-i}\}\times\varphi^{-1}_{-i}\left(T^{*}_{-i}\right)\right)\cap\left(\left(\mathrm{Id}_{S_{-i}},\varphi_{-i}\right)^{-1}\left(E^{*}_{-i}\right)\right)\\ &= \left(\{s_{-i}\}\times T_{-i}\right)\cap\left(\left(\mathrm{Id}_{S_{-i}},\varphi_{-i}\right)^{-1}\left(E^{*}_{-i}\right)\right)\\ &\supseteq \left(\{s_{-i}\}\times T_{-i}\right)\cap\left(\left(\mathrm{Id}_{S_{-i}},\varphi_{-i}\right)^{-1}\left(\left(\mathrm{Id}_{S_{-i}},\varphi_{-i}\right)\left(E_{-i}\right)\right)\right)\\ &\supseteq \left(\{s_{-i}\}\times T_{-i}\right)\cap E_{-i},\end{aligned}$$

where the fourth line follows from condition 1) of the Lemma. Since $(\{s_{-i}\} \times T_{-i}) \cap E_{-i} \neq \emptyset$ (by condition 2) of the Lemma) and E_{-i} is cautiously believed under $\beta_i(t_i)$ at level $m \leq n$, then there exists $l \leq m$ such that $\beta_i^l(t_i)((\{s_{-i}\} \times T_{-i}) \cap E_{-i}) > 0$. But

$$\beta_i^l(t_i)((\{s_{-i}\} \times T_{-i}) \cap E_{-i}) \leq \beta_i^l(t_i)\left(\left(\mathrm{Id}_{S_{-i}}, \varphi_{-i}\right)^{-1}\left(C \cap E_{-i}^*\right)\right),$$

and since C is an arbitrary elementary cilynder, this shows that Condition (ii) of Definition 17 is satisfied; hence E_{-i}^* is cautiously believed under $\left(\widehat{\mathrm{Id}_{S_{-i}}, \varphi_{-i}}\right)(\beta_i(t_i)) \in \mathcal{N}\left(S_{-i} \times T_{-i}^*\right)$.

Now note that $\left(\widehat{\mathrm{Id}_{S_{-i}}, \varphi_{-i}}\right)(\beta_i(t_i)) \cong \beta_i^*(t_i^*)$, because, under our hypothesis, $\beta_i^*(t_i^*) = \mathbf{m}_{S_{-i} \times T_{-i}^*}\left(\left(\widehat{\mathrm{Id}_{S_{-i}}, \varphi_{-i}}\right)(\beta_i(t_i))\right)$. Hence E_{-i}^* is also cautiously believed under $\beta_i^*(t_i^*)$ by Remark 6. ■

Proof of Theorem 2: For $m = 0$ the statement is immediate. We show by induction that the statement holds for $m \geq 1$. One direction of the proof makes use of a selection argument; that is, for each $i \in I$ and each $m \geq 0$, there are maps $\rho_i^m : S_i \to T_i$ such that the maps $(\mathrm{Id}_{S_i}, \rho_i^m) : S_i \to S_i \times T_i$ satisfy $(\mathrm{Id}_{S_i}, \rho_i^m)(s_i) = (s_i, \rho_i^m(s_i)) \in R_i^m \setminus R_i^{m+1}$ provided $s_i \in S_i^m$.[13] (Of course, each map ρ_i^m is continuous, since strategy sets are endowed with the discrete topology.) To this end, we first define, for each $i \in I$, the map $\rho_i^0 : S_i \to T_i$ as follows: fix some $t_i^0 \in T_i$ such that $\overline{\mathrm{marg}}_{S_{-i}}\beta_i\left(t_i^0\right) \notin \mathcal{N}^+(S_{-i})$ (such t_i^0 exists by belief-completeness), and let $\rho_i^0(s_i) = t_i^0$ for all $s_i \in S_i$. It follows that $\left(s_i, \rho_i^0(s_i)\right) \in R_i^0 \setminus R_i^1 \neq \emptyset$ for all $s_i \in S_i$, because $\rho_i^0(s_i) \in T_i$ is not cautious.

($m = 1$) Fix $i \in I$. Let $s_i \in \mathrm{Proj}_{S_i}\left(R_i^1\right)$, so that $(s_i, t_i) \in R_i^1$ for some $t_i \in T_i$. By Proposition 7, s_i is admissible, i.e., $s_i \in S_i^1$. This shows that $\mathrm{Proj}_{S_i}\left(R_i^1\right) \subseteq S_i^1$.

Conversely, let $s_i \in S_i^1$. So there is a probability measure $\mu_i^1 \in \mathcal{M}(S_{-i})$, with $\mathrm{Supp}\mu_i^1 = S_{-i}$, such that s_i is a (lexicographic) best reply to μ_i^1. Let $\iota : S_{-i} \to S_{-i} \times T_{-i}$ be the function defined by $\iota(s_{-i}) = \left(s_{-i}, \rho_{-i}^0(s_{-i})\right)$, $s_{-i} \in S_{-i}$, where $\rho_{-i}^0 : S_{-i} \to T_{-i}$ is the (constant) function we previously defined. The map ι is continuous and is such that $\iota(s_{-i}) \in R_{-i}^0 \setminus R_{-i}^1$ for all $s_{-i} \in S_{-i}$. Hence the pushforward measure $\widetilde{\iota}\left(\mu_i^1\right) \in \mathcal{M}(S_{-i} \times T_{-i})$ is well-defined, and satisfies $\widetilde{\iota}\left(\mu_i^1\right)\left(R_{-i}^1\right) = 0$; moreover $\overline{\mathrm{marg}}_{S_{-i}}\widetilde{\iota}\left(\mu_i^1\right) = \mu_i^1 \in \mathcal{N}_1^+(S_{-i})$. By minimal belief-completeness and the fact that μ_i^1 is a length-1 (hence minimal) LPS, there is $t_{s_i}^1 \in T_i$ such that $\beta_i\left(t_{s_i}^1\right) = \widetilde{\iota}\left(\mu_i^1\right)$. Clearly, $\left(s_i, t_{s_i}^1\right) \in R_i^1 \setminus R_i^2$, and this shows that $S_i^1 \subseteq \mathrm{Proj}_{S_i}\left(R_i^1 \setminus R_i^2\right)$, and so $S_i^1 \subseteq \mathrm{Proj}_{S_i}\left(R_i^1\right)$.

By arbitrariness of $i \in I$, it follows that $\prod_{i \in I}\mathrm{Proj}_{S_i}\left(R_i^1\right) = \prod_{i \in I} S_i^1$. We can conclude the proof of the basis step by defining a profile of continuous maps $\left(\rho_i^1 : S_i \to T_i\right)_{i \in I}$ as follows: for each $i \in I$,

$$\rho_i^1(s_i) = \begin{cases} t_{s_i}^1, & \text{if } s_i \in S_i^1, \\ \rho_i^0(s_i), & \text{if } s_i \in S_i \setminus S_i^1. \end{cases}$$

It turns out that $\left(s_i, \rho_i^1(s_i)\right) \in R_i^1 \setminus R_i^2$ whenever $s_i \in S_i^1$, as required.

($m \geq 2$) Suppose that the statement has been shown to hold for all $l = 1, ..., m-1$, and that, for each $i \in I$ and $l \leq m-1$, we have shown the existence of continuous maps $\rho_i^l : S_i \to T_i$ satisfying $\left(s_i, \rho_i^l(s_i)\right) \in R_i^l \setminus R_i^{l+1}$ for all $s_i \in S_i^l$. We show that the statement is true for $l = m$.

Fix a player $i \in I$, and let $s_i \in \mathrm{Proj}_{S_i}(R_i^m)$, so that $(s_i, t_i) \in R_i^m$ for some $t_i \in T_i$. It follows from the definition of R_i^m that $(s_i, t_i) \in R_i^{m-1}$, so, by the induction hypothesis, $s_i \in S_i^m$. Also, R_{-i}^{m-1} is cautiously believed under $\beta_i(t_i) = \left(\mu_i^1, ..., \mu_i^n\right)$ at some level $k \leq n$, hence

$$\cup_{l \leq k}\mathrm{Supp}\mathrm{marg}_{S_{-i}}\mu_i^l = \mathrm{Proj}_{S_{-i}}\left(R_{-i}^{m-1}\right) = S_{-i}^{m-1},$$

[13] This implies that each map $(\mathrm{Id}_{S_i}, \rho_i^m)$ satisfies $\mathrm{Proj}_{S_i} \circ (\mathrm{Id}_{S_i}, \rho_i^m) = \mathrm{Id}_{S_i}$. Put differently, $(\mathrm{Id}_{S_i}, \rho_i^m)$ is a continuous selection of the correspondence $\mathrm{Proj}_{S_i}^{-1} : S_i \to 2^{R_i^m \setminus R_i^{m+1}}$.

where the second equality follows from the induction hypothesis (for the first, see Lemma B.2 in [17]). So we can form a nested convex combination of the measures $\mathrm{marg}_{S_{-i}}\mu_i^l$, for $l = 1, ..., k$, to get a probability measure $\nu_i \in \mathcal{M}(S_{-i})$, with $\mathrm{Supp}\nu_i = S_{-i}^{m-1}$, such that s_i is a best reply to ν_i (see [7, Proposition 1]). Thus, s_i is admissible w.r.to $S_i \times S_{-i}^{m-1}$, and a fortiori w.r.to $S_i^{m-1} \times S_{-i}^{m-1}$. Hence $\mathrm{Proj}_{S_i}(R_i^m) \subseteq S_i^m$.

Conversely, let $s_i \in S_i^m$. By Lemma E.1 in BFK, it follows that, for all $l = 1, ..., m$, there is $\nu_i^l \in \mathcal{M}(S_{-i})$, with $\mathrm{Supp}\nu_i^l = S_{-i}^{l-1}$, for which s_i is a best reply among all strategies in S_i. We first show the existence of an LPS $\overline{\mu}_i = (\mu_i^1, ..., \mu_i^m) \in \mathcal{N}(S_{-i} \times T_{-i})$ such that

(a) $\mathrm{marg}_{S_{-i}}\mu_i^l = \nu_i^{m+1-l}$ for each $l = 1, ..., m$; and

(b) R_{-i}^{m-l} is cautiously believed under $\overline{\mu}_i$ at level l for each $l = 1, ..., m-1$, while R_{-i}^m is not cautiously believed.

To this end, we use the fact that we have already shown the existence of functions $\rho_i^0(\cdot)$,...,$\rho_i^{m-1}(\cdot)$ (induction hypothesis) for each $i \in I$. We construct probability measures $\mu_i^l \in \mathcal{M}(S_{-i} \times T_{-i})$ as follows:

$$\mu_i^l = \left(\widetilde{\mathrm{Id}_{S_{-i}}, \rho_{-i}^{m-l}}\right)\left(\nu_i^{m+1-l}\right), \forall l \in \{1, ..., m\}.$$

Let $\overline{\mu}_i \in \mathcal{N}(S_{-i} \times T_{-i})$ be the concatenation of those measures, i.e., $\overline{\mu}_i = (\mu_i^1, ..., \mu_i^m)$. It readily follows that $\overline{\mu}_i$ satisfies property (a), since $\mathrm{Proj}_{S_{-i}} \circ \left(\mathrm{Id}_{S_{-i}}, \rho_{-i}^{m-l}\right) = \mathrm{Id}_{S_{-i}}$ for all $l \in \{1, ..., m\}$. We now show that also property (b) holds. Using the properties of the functions $\rho_{-i}^{m-l}(\cdot)$ specified above, we get that, for all $l \in \{1, ..., m\}$,

$$\mu_i^l\left(R_{-i}^{m-l} \setminus R_{-i}^{m+1-l}\right) = \nu_i^{m+1-l}\left(\left(\mathrm{Id}_{S_{-i}}, \rho_{-i}^{m-l}\right)^{-1}\left(R_{-i}^{m-l} \setminus R_{-i}^{m+1-l}\right)\right) = \nu_i^{m+1-l}\left(S_{-i}^{m-l}\right) = 1. \tag{8.3}$$

Specifically, for $l = 1$ this yields

$$\begin{aligned} \mu_i^1\left(R_{-i}^{m-1} \setminus R_{-i}^m\right) &= 1, \\ \mu_i^1\left(R_{-i}^m\right) &= 0, \end{aligned}$$

hence condition (i) of Cautious Belief is satisfied for R_{-i}^{m-1}, while R_{-i}^m cannot be cautiously believed under $\overline{\mu}_i$. Moreover, note that

$$\mathrm{Suppmarg}_{S_{-i}}\mu_i^1 = \mathrm{Supp}\nu_i^m = S_{-i}^{m-1}.$$

By the induction hypothesis, $S_{-i}^{m-1} = \mathrm{Proj}_{S_{-i}}R_i^{m-1}$; this implies that Condition (ii) of Cautious Belief holds (see Lemma B.2 in [17]). Analogous arguments show that all the conditions of Cautious Belief are satisfied for R_{-i}^{m-l} at level l for each $l = 2, ..., m-1$. This shows that property (b) is satisfied.

Remark 5 and (a) yield that s_i is optimal under $\mathbf{m}_{S_{-i}\times T_{-i}}(\overline{\mu}_i)$. By (a), $\mathrm{marg}_{S_{-i}}\mu_i^m = \nu_i^1$; hence, by $\mathrm{Supp}\nu_i^1 = S_{-i}^0 = S_{-i}$, $\overline{\mathrm{marg}}_{S_{-i}}\overline{\mu}_i \in \mathcal{N}^+(S_{-i})$. So, by Remark 2 and Remark 5, we have $\overline{\mathrm{marg}}_{S_{-i}}\left(\mathbf{m}_{S_{-i}\times T_{-i}}(\overline{\mu}_i)\right) \in \mathcal{N}^+(S_{-i})$. By Remark 6 and (b), it follows that R_{-i}^l is cautiously believed under $\mathbf{m}_{S_{-i}\times T_{-i}}(\overline{\mu}_i)$ for each $l = 1, ..., m-1$, while R_{-i}^m is not cautiously believed.

It now follows from minimal belief-completeness that there is $t_{s_i}^m \in T_i$ such that $\beta_i\left(t_{s_i}^m\right) = \mathbf{m}_{S_{-i}\times T_{-i}}(\overline{\mu}_i)$; this implies $\left(s_i, t_{s_i}^m\right) \in R_i^m \setminus R_i^{m+1}$, hence $s_i \in \mathrm{Proj}_{S_i}\left(R_i^m \setminus R_i^{m+1}\right)$, and a fortiori $s_i \in \mathrm{Proj}_{S_i}(R_i^m)$.

By arbitrariness of $i \in I$, it follows that $\prod_{i\in I}\mathrm{Proj}_{S_i}(R_i^m) = \prod_{i\in I}S_i^m$. To conclude the proof of the inductive step, it remains to define a profile of continuous maps $(\rho_i^m : S_i \to T_i)_{i\in I}$. This is done by letting, for each $i \in I$,

$$\rho_i^m(s_i) = \begin{cases} t_{s_i}^m, & \text{if } s_i \in S_i^m, \\ \rho_i^0(s_i), & \text{if } s_i \in S_i \setminus S_i^m. \end{cases}$$

Clearly, each map $\rho_i^m : S_i \to T_i$ satisfies $(s_i, \rho_i^m(s_i)) \in R_i^m \setminus R_i^{m+1}$ whenever $s_i \in S_i^m$. ■

References

[1] Aliprantis, C.D., and K.C. Border (1999): *Infinite Dimensional Analysis*. Berlin: Springer Verlag.

[2] Battigalli, P., *Restrizioni Razionali su Sistemi di Probabilità Soggettive e Soluzioni di Giochi ad Informazione Completa*, EGEA, Milan, 1993.

[3] Battigalli, P. and A. Friedenberg, "Context-Dependent Forward Induction Reasoning" IGIER w.p. 351, Bocconi University, 2009.

[4] Battigalli, P., and M. Siniscalchi (1999): "Hierarchies of Conditional Beliefs and Interactive Epistemology in Dynamic Games," *Journal of Economic Theory*, 88, 188-230.

[5] Battigalli, P., and M. Siniscalchi (2002): "Strong Belief and Forward Induction Reasoning," *Journal of Economic Theory*, 106, 356-391.

[6] Blume, L., A. Brandenburger, and E. Dekel (1991): "Lexicographic Probabilities and Choice Under Uncertainty," *Econometrica,* 59, 61-79.

[7] Blume, L., A. Brandenburger, and E. Dekel (1991): "Lexicographic Probabilities and Equilibrium Refinements," *Econometrica,* 59, 81-98.

[8] Bogachev, V. (2006): *Measure Theory*. Berlin: Springer Verlag.

[9] Bogachev, V. and A. V. Kolesnikov, "Open Mappings of Probability Measures and the Skorokhod Representation Theorem", *Theory of Probability and its Applications*, 2001, **46**, 1-21.

[10] Borgers, T. (1994): "Weak Dominance and Approximate Common Knowledge," *Journal of Economic Theory*, 64, 265-276.

[11] Bourbaki N., *Eléments de Mathématque, Intégration, Livre VI, Chapitre IX, Intégration sur Les Espaces Topologiques Séparés*, Hermann, 1969.

[12] Bourbaki, N., *Elements of Mathematics. Theory of Sets*, Springer-Verlag, Berlin, 2004.

[13] Brandenburger, A. (2003): "On the Existence of a 'Complete' Possibility Structure," in *Cognitive Processes and Economic Behavior*, ed. by M. Basili, N. Dimitri and I. Gilboa. New York: Routledge, 30-34.

[14] Brandenburger, A., and E. Dekel (1993): "Hierarchies of Beliefs and Common Knowledge," *Journal of Economic Theory*, 59, 189-198.

[15] Brandenburger, A., A. Friedenberg, and H.J. Keisler (2008): "Admissibility in Games," *Econometrica*, 76, 307-352.

[16] Castaing, C., de Fitte P.R. and M. Valadier, *Young Measures on Topological Spaces: With Applications in Control Theory and Probability Theory*, Springer-Verlag, Berlin, 2004.

[17] Catonini, E. and N. De Vito, "Cautious Belief and Iterated Admissibility", 2018, working paper.

[18] Catonini, E.and N. De Vito, "Hierarchies of Lexicographic Beliefs", 2014, working paper.

[19] Cohn, D.L. (2013): *Measure Theory.* Boston: Birkhauser.

[20] Dekel, E., A. Friedenberg, and M. Siniscalchi (2016): "Lexicographic Beliefs and Assumption," *Journal of Economic Theory*, 163, 955-985.

[21] Dekel, E., and D. Fudenberg (1990): "Rational Behavior with Payoff Uncertainty," *Journal of Economic Theory,* 52, 243-67.

[22] Dellacherie, C., Meyer, P., "Probabilities and potential", *Math. Stud.,* **29**, 1978.

[23] De Vito, N., "Complete Conditional Type Structures", working paper, Bocconi University, 2014.

[24] Di Tillio, A. (2008): "Subjective Expected Utility in Games," *Theoretical Economics*, 3, 287-323

[25] Dugundji, J., *Topology*, Allyn and Bacon Inc., Boston, 1966.

[26] Engelking, R. (1989): *General Topology.* Berlin: Heldermann.

[27] Epstein, L.G., and T. Wang (1996): "Beliefs About Beliefs Without Probabilities," *Econometrica*, 64, 1343-1373.

[28] Fremlin, D.H, *Measure Theory, Volume 4. Topological Measure Spaces*, Torres Fremlin Eds., Colchester, 2003.

[29] Friedenberg, A. (2010): "When Do Type Structures Contain All Hierarchies of Beliefs?," *Games and Economic Behavior*, 68, 108-129.

[30] Friedenberg, A., and H.J. Keisler (2011): "Iterated Dominance Revisited," working paper.

[31] Friedenberg, A., and M. Meier (2011): "On the Relationship Between Hierarchy and Type Morphisms," *Economic Theory*, 46, 377-399.

[32] Heifetz, A., "The Bayesian Formulation of Incomplete Infomation - the Non-Compact Case", *International Journal of Game Theory*, 1993, **21**, 329-338.

[33] Heifetz, A., and D. Samet (1998): "Topology-Free Typology of Beliefs," *Journal of Economic Theory*, 82, 324-341.

[34] Heifetz, A., Meier, M. and Schipper, B., "Comprehensive Rationalizability", working paper, 2013.

[35] Kallenberg, O., *Foundations of Modern Probability*, Springer-Verlag, Berlin, 2002.

[36] Kechris, A. (1995): *Classical Descriptive Set Theory.* Berlin: Springer Verlag.

[37] Keisler, H.J., and B.S. Lee (2015): "Common Assumption of Rationality," working paper, University of Toronto.

[38] Lee, B.S. (2013): "Conditional Beliefs and Higher-Order Preferences," working paper, University of Toronto.

[39] Lee, B. S., "A space of Lexicographic Preferences", 2013, working paper.

[40] Lee, B.S. (2016): “Admissibility and Assumption,” *Journal of Economic Theory*, 163, 42-72.

[41] Liu, Q., "On Redundant Types and Bayesian Formulation of Incomplete Information", *Journal of Economic Theory*, 2009, **44**, 2115-2145.

[42] Mertens, J.F., and S. Zamir (1985): “Formulation of Bayesian Analysis for Games With Incomplete Information,” *International Journal of Game Theory*, 14, 1-29.

[43] Myerson, R., “Multistage games with communication”, *Econometrica*, **54**, 1986, 323-358.

[44] Munkres, J.R., *Topology. Second edition*, Prentice-Hall Inc., N.J., 2000.

[45] Pearce, D. (1984): “Rationalizable Strategic Behavior and the Problem of Perfection,” *Econometrica,* 52, 1029-1050.

[46] Rao, M. M., *Foundation of stochastic analysis*, Academic Press, New York, 1981.

[47] Rao, M. M., *Measure Theory and Integration*, CRC Press, 2004.

[48] Rao, M. M., *Conditional Measures and Applications*, Taylor & Francis Group, 2005.

[49] Rao, K.P.S.B., and M.B. Rao, *Theory of Charges*, Academic Press, New York, 1983.

[50] Schwartz, L., *Radon Measures on Arbitrary Topological Spaces and Cylindrical Measures*, Oxford University Press, London, 1973.

[51] Siniscalchi, M., "Epistemic Game Theory: Beliefs and Types", in Durlauf, S.N., Blume, L.E. (Eds.), *The New Palgrave Dictionary of Economics*, Second Edition, Palgrave Macmillan, 2008.

[52] Srivastava, S.M., *A Course on Borel Sets*, Springer-Verlag, New York, 1998.

[53] Topsoe, F., *Topology and Measure*, Springer-Verlag, Berlin, 1970.

[54] Yeh, J., *Real Analysis: Theory of Measure and Integration*, World Scientific Publishing, Singapore, 2006.