Preprints of the
Max Planck Institute for
Research on Collective Goods
Bonn 2016/17

A Homeomorphism Theorem for the Universal Type Space with the Uniform Weak Topology

Martin F. Hellwig

MAX PLANCK SOCIETY

Preprints of the
Max Planck Institute
for Research on Collective Goods
Bonn 2016/17

A Homeomorphism Theorem for the Universal Type Space with the Uniform Weak Topology

Martin F. Hellwig

November 2016

Max Planck Institute for Research on Collective Goods, Kurt-Schumacher-Str. 10, D-53113 Bonn
http://www.coll.mpg.de

A Homeomorphism Theorem for the Universal Type Space with the Uniform Weak Topology*

Martin F. Hellwig
Max Planck Institute for Research on Collective Goods
Kurt Schumacher-Str. 10
D-53113 Bonn, Germany
hellwig@coll.mpg.de

November 12, 2016

Abstract

Kolmogorov's extension theorem provides a natural mapping from the space of coherent hierarchies of an agent's first-order, second-order, etc. beliefs to the space of probability measures over the exogenous parameters and the other agents' belief hierarchies. Mertens and Zamir (1985) showed that, if the spaces of belief hierarchies are endowed with the product topology, then this mapping is a homeomorphism. This paper shows that this mapping is also a homeomorphism if the spaces of belief hierarchies are endowed with the uniform weak topology of Chen et al. (2010) or the universal strategic topology of Dekel et al. (2006), both of which ensure that strategic behaviour exhibits desirable continuity properties.

Key Words: incomplete information, universal type space, uniform weak topology, uniform strategic topology, homeomorphism theorem.

JEL Classification: C70, C72.

---

*Without implicating her, I thank Alia Gizatulina for drawing my attention to the problem considered here and for providing very helpful comments on previous drafts. I also thank Eduardo Faingold and Siyang Xiong for greatly helping my understanding of Chen et al. (2010, 2012, 2016).

1 Introduction

The universal type space approach to modelling strategic interdependence under incomplete information comes in two versions. In one version, an agent's "type" is given by an infinite hierarchy of beliefs about the exogenous parameters affecting the strategic situation, about the other agents' beliefs about the exogenous parameters, about the other agents' beliefs about other agents' beliefs about the exogenous parameters, and so on. In the other version, an agent's "type" is specified as a probabilistic belief about the exogenous parameters and about the other agents' "types".

Following the informal discussion in Harsanyi (1967/68), both versions were formally analysed by Mertens and Zamir (1985). They showed that the two versions are actually equivalent in the sense that there exists a homeomorphism between the space of an agent's belief hierarchies and the space of the agent's beliefs over exogenous parameter and other agents' belief hierarchies. Their result was subsequently generalized by Brandenburger and Dekel (1993) and Heifetz (1993).[1]

Either version of the universal type space approach has a weakness that the other version does not have.[2] Belief hierarchies do not easily fit into the standard framework of decision theory, where a single probability measure is used to represent an agent's uncertainty. This difficulty disappears if an agent's "type" is represented by a single probability measure. However, the specification of an agent's "type" in terms of beliefs about the other agents' "types" involves an element of circularity.

The equivalence results of Mertens and Zamir (1985), Brandenburger and Dekel (1993), and Heifetz (1993) show that these weaknesses do not matter. In particular, they eliminate the circularity involved in defining "types" in terms of beliefs over (other agents') "types" by showing that each belief hierarchy of an agent defines a unique probability measure over exogenous data and other agents' belief hierarchies and that this probability measure can be used to specify an agent's beliefs about the other agents' "types".

However, these results depend on the chosen topology.[3] Mertens and

[1] Whereas Mertens and Zamir (1985) assumed that the space Θ of exogenous parameters is a compact metric space, Brandenburger and Dekel (1993) allow Θ to be a complete separable metric space. Heifetz (1993) requires Θ to be a Hausdorff space, but he imposes additional regularity conditions on beliefs.

[2] For an extensive discussion, see Heifetz and Samet (1998).

[3] For a criticism, see Heifetz and Samet (1998, 1999). Heifetz and Samet (1999) observe

Zamir (1985) assume that the space Θ of exogenous parameters is a compact metric space. For each k, they endow the space of beliefs of order k with the topology of weak convergence of probability measures (the weak* topology) and use a straightforward induction on k to show that the space of beliefs of order k is also a compact metric space. Given this specification of the spaces of beliefs of different orders, they treat the space U_i of coherent belief hierarchies of agent i as a subset of the product of the spaces of beliefs of different levels in the hierarchy and use Kolmogorov's extension theorem to obtain a one-to-one and onto mapping from U_i to the space $\mathcal{M}(\Theta \times U_{-i})$ of probability measures on the space of exogenous parameters and other agents' belief hierarchies. They show that this mapping is a homeomorphism if the spaces of belief hierarchies of the different agents are endowed with the product topology and if $\mathcal{M}(\Theta \times U_{-i})$ is endowed with the topology of weak convergence of probability measures.

This homeomorphism theorem implies that, in assessing the continuity properties of an agent's behaviour with respect to his "type", it makes no difference whether the agent's "type" is specified in terms of a hierarchy of beliefs or in terms of a single probability measure on the exogenous parameters and the other agents' beliefs.[4] Under standard assumptions, therefore, in either formulation, the dependence of behaviour on types exhibits what Dekel et al. (2006) call the *upper strategic convergence property*, i.e. upper hemi-continuity of behaviour correspondences.

However, the product topology on the space of belief hierarchies does not provide for what Dekel et al. (2006) call the *lower strategic convergence property*, namely the property that the minimal ε for which an agent's choices are strictly ε interim correlated rationalizable should depend continuously on the agent's belief hierarchy. The reason is that, under the product topology, belief hierarchies can be treated as similar even if for some very large k the

that, without a suitable topological structure, Kolmogorov's extension theorem cannot be used, so there may be coherent belief hierarchies that do not correspond to beliefs about the other agents' types. Given this criticism, Heifetz and Samet (1998) propose a purely measure theoretic formulation that focuses on beliefs, rather than coherent belief hierarchies.

[4] For an application, see the discussion of the genericity of the McAfee-Reny condition for the feasibility of full surplus extraction in Gizatulina and Hellwig (2016). The condition, which was first formulated in McAfee and Reny (1992), refers to beliefs as probability measures over the other agents' types. In the context of the universal type space, any such condition on beliefs about other agents' types raises questions about the implications of the analysis for a specification in terms of belief hierarchies.

associated beliefs of order k are very different. The set of ε interim correlated rationalizable actions can be sensitive to differences in beliefs of arbitrarily high orders. Rubinstein's (1989) electronic mail game provides an example.

Dekel et al. (2006) and Chen et al. (2010, 2012) have therefore suggested that spaces of belief hierarchies should be endowed with finer topologies. The *strategic* and *uniform strategic topologies* of Dekel et al. (2006) are specified directly in terms of the desired upper and lower convergence properties of ε interim correlated rationalizable choices in strategic games. The *uniform weak topology* of Chen et al. (2010, 2012) is specified in terms of the belief hierarchies only, without any reference to the desired continuity properties of strategic behaviour.

Like the product topology, the uniform weak topology is derived from the topologies on the different spaces of beliefs representing the different orders in the belief hierarchy. Whereas, for very large k, beliefs of order k do not matter very much under the product topology, the uniform weak topology gives equal weight to all orders of beliefs. As shown by Chen et al. (2010, 2012), this property ensures that the uniform weak topology provides for both the lower and the upper strategic convergence properties of Dekel et al. (2006). Indeed, the uniform weak topology coincides with the uniform strategic topology of Dekel et al. (2006), which requires the continuity properties of strategic behaviour to hold uniformly over all strategic games with the specified exogenous parameters.[5]

The strategic, uniform strategic, and uniform weak topologies are all imposed on the space of belief hierarchies. The question is what these topologies imply for the space of beliefs about exogenous parameters and other agents' belief hierarchies. Does Kolmogorov's extension theorem still provide us with a homeomorphism between the space U_i of belief hierarchies of agent i and the space $\mathcal{M}(\Theta \times U_{-i})$ of agent i's beliefs about exogenous parameters and other agents' belief hierarchies? This paper provides a positive answer to this question for the uniform weak topology of Chen at al. (2010, 2012) and, by implication, the uniform strategic topology of Dekel et al. (2006).

Several issues must be addressed. If U_i is endowed with the uniform weak topology, rather than the product topology, continuity of the mapping from U_i to $\mathcal{M}(\Theta \times U_{-i})$ is preserved, but the inverse of this mapping is discontinuous unless $\mathcal{M}(\Theta \times U_{-i})$ is given a finer topology as well. The

[5] Chen et al. (2010) show that the uniform weak topology is at least as fine as the uniform strategic topoplogy; Chen et al. (2012) show that the two are actually equivalent.

obvious candidate is to endow the belief hierarchies U_j of agents $j \neq i$ with the uniform weak topology rather than the product topology and to endow $\Theta \times U_{-i}$ with the corresponding product topology. Because the uniform weak topology is finer than the product topology, this change enlarges the set of continuous functions on $\Theta \times U_{-i}$. As a result, the topology of weak convergence on the space of measures on $\Theta \times U_{-i}$ also becomes finer, which improves the scope for continuity of the projection from $\mathcal{M}(\Theta \times U_{-i})$ to U_i.

But this change raises questions about the mapping from belief hierarchies to beliefs. Chen et al. (2010, 2016) have shown that, with the uniform weak topology, unlike the product topology, the spaces of belief hierarchies are *not* separable. The Borel σ-algebra on the space of belief hierarchies that is generated by the uniform weak topology is strictly larger than the product σ-algebra (or the Borel σ-algebra generated by the product topology).[6] However, the probability measures that are obtained by using Kolmogorov's extension theorem to map belief hierarchies into beliefs are only defined on the product σ-algebra. A belief hierarchy for agent i does not contain enough information to assign probabilities to those sets that belong to the difference between the Borel σ-algebra generated by the uniform weak topology and the product σ-algebra.

We must either give up on the principle that beliefs about other agents' types should be derived from belief hierarchies, or we must limit beliefs to set functions on the σ-algebra that is obtained by endowing each of the belief hierarchies $U_j, j \neq i$, with the product σ-algebra even though this is not the Borel σ-algebra for the chosen topology on $\Theta \times U_{-i}$. Given this choice, I stick to the principle that beliefs about other agents' beliefs should be derived from belief hierarchies and specify the range of the mapping from belief hierarchies to beliefs as the set of probability measures on the σ-algebra that is obtained by treating $\Theta \times U_{-i}$, $U_{-i} = \prod_{j \pm i} U_j$, and the spaces U_j of belief hierarchies of agents $j \neq i$ as product spaces.

At the same time though, I will refer to the uniform weak topology on U_j, $j \neq i$, rather than the product topology, in defining weak convergence of measures on $\Theta \times U_{-i}$. Thus, I will define convergence of a sequence of measures in terms of convergence of integrals of bounded real-valued functions

[6] Whereas Chen et al. (2010, p. 459) claim that the two σ-algebras coincide, Chen et al. (2016) retract this claim and give an example of an open set in the universal weak topology that is not a Borel set of the product topology, namely the ε-neighbourhood of an analytic set that is not a Borel set.

on $\Theta \times \prod_{j \pm i} U_j$ that are measurable with respect to the product σ-algebra and continuous with respect to the topology on $\Theta \times \prod_{j \pm i} U_j$ that is obtained when the spaces U_j, $j \neq i$, have the uniform weak topology. Given this specification of the space of beliefs about exogenous parameters and other agents' belief hierarchies, I will show that the Kolmogorov mapping from belief hierarchies of agent i to beliefs about exogenous parameters and belief hierarchies of other agents is a homeomorphism.

This homeomorphism theorem presumes the continuum hypothesis. The continuum hypothesis ensures that, even though $\Theta \times U_{-i}$ itself is not separable, every measure on $\Theta \times U_{-i}$ has a separable support. As shown in Billingsley (1968), this property ensures that the topology of weak convergence on $\mathcal{M}(\Theta \times U_{-i})$ can be metrized by the Prohorov distance. The proof of the homeomorphism theorem for the universal type space with the uniform weak topology involves, first, showing that Billingsley's arguments carry over to the pressent case where the Borel σ-algebra generated by the topology on $\Theta \times U_{-i}$ is larger than the σ-algebra on which the measures under consideration are defined and, second, relating the Prohorov metric for the topology on $\mathcal{M}(\Theta \times U_{-i})$ to the metric on U_i that Chen et al. (2010) used to define the uniform weak topology.

In the following, Section 2 introduces notation and some mathematical basics. Section 3 specifies the space of belief hierarchies and the different topologies on this space. Section 4 introduces the Kolmogorov mapping from the space of belief hierarchies to the space of beliefs about exogenous parameters and other agents' belief hierarchies. Section 5 introduces the space $\mathcal{M}^{\infty}(\Theta \times U_{-i}^{uw})$ of probability measures with a domain that is given by the product σ-algebra and a topology that is based on the uniform weak topology for belief hierarchies. Section 6 states and proves the homeomorphism theorem for the universal type space with the uniform weak topology. Section 7 concludes with a discussion of the relation between the "universal" type space and abstract (Harsanyi) type spaces when the former has the uniform weak topology.

2 Notation and Mathematical Preliminaries

I begin with some remarks on notation and topological basics: Given any metric space S, I will write $\mathcal{B}(S)$ for the Borel σ-algebra on S and $\mathcal{M}(S)$ for the space of probability measures on $(S, \mathcal{B}(S))$. If necessary, I will use superscripts to indicate which topology and which metric a given space S is endowed with. For example, if U_i is the space of agent i's belief hierarchies, I will write U_i^π and U_i^{uw} to indicate whether U_i has the product topology or the uniform weak topology. Accordingly, I also distinguish between $\mathcal{M}(U_i^\pi)$ and $\mathcal{M}(U_i^{uw})$, the space of probability measures on $(U_i^\pi, \mathcal{B}(U_i^\pi))$ and the space of probability measures on $(U_i^{uw}, \mathcal{B}(U_i^{uw}))$.

Given a metric space S, I endow the space $\mathcal{M}(S)$ of probability measures on $(S, \mathcal{B}(S))$ with the topology of weak convergence of probability measures. In this topology, a sequence $\{\mu^r\}$ in $\mathcal{M}(S)$ converges to a limit μ if and only if, for every bounded and continuous real-valued function on S, the integrals $\int f(s)\mu^r(ds)$ converge to $\int f(s)\mu(ds)$.[7]

As is well known, if S is a separable metric space, then $\mathcal{M}(S)$ is a separable metric space. The topology on $\mathcal{M}(S)$ can be metrized by the Prohorov metric, which specifies the distance $\rho(\mu, \hat{\mu})$ between two measures μ and $\hat{\mu}$ on S as the infimum of the set of ε such that

$$\mu(B) \leq \hat{\mu}(B^\varepsilon) + \varepsilon \text{ and } \hat{\mu}(B) \leq \mu(B^\varepsilon) + \varepsilon \tag{1}$$

for all sets $B \in \mathcal{B}(S)$ with ε-neighbourhoods $B^\varepsilon \in \mathcal{B}(S)$.[8]

In the preceding paragraph, the assumption that S is separable is not actually needed. As shown in Appendix III of Billingsley (1968), this assumption can be replaced by the requirement that the cardinal of S be non-measurable, i.e., that there exists no atomless probability measury that is defined on the class of all subsets of S. Under the continuum hypothesis, this requirement is fulfilled if the cardinality of S does not exceed that of the continuum.

[7]As is well known, if S is a *separable* metric space, $\mathcal{M}(S)$ can be identified with the space of continuous linear functionals on the space $\mathcal{C}(S)$ of bounded continuous real-valued functions on S, i.e. the dual of $\mathcal{C}(S)$, and the topology of weak convergence coincides with the weak* topology. If S is *not* separable, the dual of $\mathcal{C}(S)$ corresponds to the space $rba(S)$ of regular (finitely) additive set functions on $(S, B(S))$, which is larger than $\mathcal{M}(S)$. The topology of weak convergence on $\mathcal{M}(S)$ is then equivalent to the subspace topology that is induced by the weak* topology on $rba(S)$. See, e.g., Parthasarathy (1967), p. 35.

[8]See, e.g., Theorem 5, p. 238, in Billingsley (1968) or Theorem 11.3.3, p. 395, in Dudley (2002).

The topology of weak convergence on $\mathcal{M}(S)$ depends in two ways on the underlying topology on S. First, the topology on S determines the Borel σ-algebra $\mathcal{B}(S)$ and hence the domain of the measures in $\mathcal{M}(S)$. Second, the topology on S determines the set of bounded and continuous real-valued functions on S. Given two topologies $\mathcal{T}^c, \mathcal{T}^f$ on S such that $\mathcal{T}^f$ is finer than $\mathcal{T}^c$, the Borel σ-algebra $\mathcal{B}(S^f)$ that is generated by $\mathcal{T}^f$ is no smaller than and in some cases strictly larger than the Borel σ-algebra $\mathcal{B}(S^c)$ that is generated by $\mathcal{T}^c$, where the notation S^c, S^f indicates whether S has the topology $\mathcal{T}^c$ or the topology $\mathcal{T}^f$. In addition, the set of bounded and continuous real-valued functions on S^f is larger than the set of bounded continuous functions on S^c. The requirement that integrals of all bounded continuous functions converge is more restrictive when S has the topology $\mathcal{T}^f$ than when S has the topology $\mathcal{T}^c$. The topology of weak convergence on $\mathcal{M}(S^f)$ is therefore finer than the topology of weak convergence on $\mathcal{M}(S^c)$. In the specification of the Prohorov metric, the dependence of the topology on $\mathcal{M}(S)$ on the topology on S is implicit in (i) the fact that the sets B for which (1) must hold depend on the topology and (ii) in the fact that, for any B, the ε-neighbourhood $B^\varepsilon \in \mathcal{B}(S)$ that appears in (1) depends on the metric on S.

3 The Space of Belief Hierarchies, the Product Topology, and the Uniform Weak Topology

Following Chen et al. (2010), I use the procedure of Mertens and Zamir (1985) to construct the universal type space. Suppose that there are I agents. Let Θ be a compact metric space of exogenous parameters that affect the strategic situation. Proceeding inductively, define a sequence of spaces $X^0, X^1, X^2, ...$ by setting

$$X^0 = \Theta, X^1 = X^0 \times \mathcal{M}(X^0)^{I-1} \tag{2}$$

and, for each $k \geq 2$,

$$X^k = \left\{ (\theta, \mu^1, ..., \mu^k) \in X^0 \times \prod_{\ell=1}^{k} \mathcal{M}(X^{\ell-1})^{I-1} : \text{marg}_{X^{\ell-2}} \mu^\ell = \mu^{\ell-1}, \ell = 2, ..., k \right\}. \tag{3}$$

For any k, standard arguments imply that, if X^{k-1} is a compact metric space, then $\mathcal{M}(X^{k-1})$ and $X^0 \times \prod_{\ell=1}^{k} \mathcal{M}(X^{\ell-1})^{I-1}$ are also compact metric spaces and so is X^k, which is a closed subset of $X^0 \times \prod_{\ell=1}^{k} \mathcal{M}(X^{\ell-1})^{I-1}$. The assumption that $X^0 = \Theta$ is a compact metric space thus implies that, for every k, X^k and $\mathcal{M}(X^k)$ are compact metric spaces.[9]

The space of belief hierarchies of player i is defined as:

$$U_i = \left\{ (\mu^k)_{k\geq 1} \in \prod_{k\geq 1} \mathcal{M}(X^{k-1}) : \mathrm{marg}_{X^{k-2}} \mu^\ell = \mu^{\ell-1}, k = 2, 3, ... \right\}.^{10} \tag{4}$$

This space is obviously a subset of the product

$$\bar{U}_i := \prod_{k\geq 1} \mathcal{M}(X^{k-1}). \tag{5}$$

Given the topologies on $\mathcal{M}(X^{k-1})$ and the induced Borel σ-algebras $\mathcal{B}(\mathcal{M}(X^{k-1}))$, for $k = 1, 2, ...$, let

$$\mathcal{B}^\infty(\bar{U}_i) := \prod_{k\geq 1} \mathcal{B}(\mathcal{M}(X^{k-1})) \tag{6}$$

be the corresponding product σ-algebra on $\bar{U}_i$ and

$$\mathcal{B}^\infty(U_i) := \{B \cap U_i | B \in \mathcal{B}^\infty(\bar{U}_i)\} \tag{7}$$

the induced σ-algebra on U_i.

The definition of the product σ-algebra is independent of the topology on U_i, but of course we have $\mathcal{B}^\infty(\bar{U}_i) = \mathcal{B}(\bar{U}_i^\pi)$ and

$$\mathcal{B}^\infty(U_i) = \mathcal{B}(U_i^\pi) \tag{8}$$

where the superscript π indicates that $\bar{U}_i$ and U_i are endowed with the product topology that is induced by the topologies on $\mathcal{M}(X^0)$, $\mathcal{M}(X^1), ...$ The

[9] Parthasarathy (1967), p. 45.

[10] Note that the right-hand side of (3) does not depend on i. The subscript i on the left-hand side is therefore irrelevant, a mere mnemonic device indicating that we are talking about the beliefs of agent i.

product topology on $\bar{U}_i$ and U_i can be metrized, e.g., by the metric

$$\rho_i^\pi((\mu^k)_{k\geq 1}, (\hat{\mu}^k)_{k\geq 1}) := \sum_{k=1}^{\infty} \alpha^k \rho^k(\mu^k, \hat{\mu}^k), \tag{9}$$

where α is some number strictly between 0 and 1 and, for any k and any two measues μ^k and $\hat{\mu}^k$ in $\mathcal{M}(X^{k-1})$, $\rho^k(\mu^k, \hat{\mu}^k)$ is the Prohorov distance between μ^k and $\hat{\mu}^k$, i.e.,

$$\rho^k(\mu^k, \hat{\mu}^k) = \inf\{\delta > 0 |\ \mu^k(B) \leq \hat{\mu}^k(B^\delta) + \delta \text{ and } \hat{\mu}^k(B) \leq \mu^k(B^\delta) + \delta \quad \text{for all } B \in \mathcal{B}(X^{k-1})\}, \tag{10}$$

where B^δ is the δ-neighbourhood of B in $X^{k-1} \subset X^0 \times \prod_{\ell=1}^{k-1} \mathcal{M}(X^{\ell-1})^{I-1}$.

As an alternative to the product topology, Chen et al. (2010) introduced the uniform weak topology. The uniform weak topology on U_i is induced by the metric ρ^{uw} such that

$$\rho_i^{uw}((\mu^k)_{k\geq 1}, (\hat{\mu}^k)_{k\geq 1}) := \sup_k \rho^k(\mu^k, \hat{\mu}^k) \tag{11}$$

for any $(\mu^k)_{k\geq 1}$ and $(\hat{\mu}^k)_{k\geq 1}$ in U_i. I will use the notation U_i^{uw} to indicate that U_i is endowed with the uniform weak topology.

The uniform weak topology on U_i is obviously finer than the product topology. In fact, whereas U_i^π, a closed subset of the product $\bar{U}_i$ of compact metric spaces, is itself a compact metric space, Chen et al. (2010) show that U_i^{uw} is not even separable. Chen et al. (2016) also show that the Borel σ-algebra $\mathcal{B}(U_i^{uw})$ that is induced by the uniform weak topology on U_i is strictly larger than the Borel σ-algebra $\mathcal{B}(U_i^\pi)$ that is induced by the product topology on U_i (and hence strictly larger than the product σ-algebra $\mathcal{B}^\infty(U_i)$).

4 Mapping Belief Hierarchies into Beliefs

I now turn to the relation between the space U_i of belief hierarchies for agent i and the space of probability measures on the product $\Theta \times U_{-i}$, where $U_{-i} := \prod_{j\neq i} U_j$ is the space of vectors of belief hierarchies of the other agents. Along the lines of the preceding discussion, for $j \neq i$, let $\mathcal{B}^\infty(U_j)$ be the product σ-algebra on $\prod_{k\geq 1} \mathcal{B}(\mathcal{M}((X^k))$, let $\mathcal{B}^\infty(U_{-i}) = \prod_{j\neq i} \mathcal{B}^\infty(U_j)$, and let $\mathcal{M}^\infty(\Theta \times U_{-i})$ be the set of probability measures on $(\Theta \times U_{-i}, \mathcal{B}(\Theta) \times \mathcal{B}^\infty(U_{-i}))$. The notation $\mathcal{B}^\infty(U_j), \mathcal{M}^\infty(\Theta \times U_{-i})$ is meant to indicate that both spaces are defined without reference to any topology on the spaces U_j of belief hierarchies.

As was noted by Mertens and Zamir (1985), for every belief hierarchy $(\mu^k)_{k\geq 1} \in U_i$ of agent i, there exists a unique probability measure

$$\mu^\infty = \beta_i((\mu^k)_{k\geq 1}) \in \mathcal{M}^\infty(\Theta \times U_{-i}) \tag{12}$$

such that the marginal distributions on $X^1, X^2, ...$ that are induced by $\mu^\infty = \beta_i((\mu^k)_{k\geq 1})$ are just the measures μ^k, $k = 1, 2, ...$ Thus, for any k and any $B_\Theta \in \mathcal{B}(\Theta)$ and $B_j^\ell \in \mathcal{B}(\mathcal{M}(X^{\ell-1}))$, $j \neq i, \ell = 1, ..., k$,

$$\begin{aligned} &\mu^\infty \left(B_\Theta \times \prod_{j\neq i} [B_j^1 \times ... \times B_j^k \times \mathcal{M}(X^k) \times \mathcal{M}(X^{k+1}) \times ...] \right) \qquad (13) \\ &\qquad = \mu^k \left(B_\Theta \times \prod_{j\neq i} B_j^1 \times ... \times \prod_{j\neq i} B_j^k \right). \end{aligned}$$

The consistency condition $\mathrm{marg}_{X^{k-2}} \mu^\ell = \mu^{\ell-1}, \ell = 1, 2, ...$, ensures that, for any $k' > k$, we have

$$\begin{aligned} &\mu^{k'} \left(B_\Theta \times \prod_{j\neq i} B_j^1 \times ... \times \prod_{j\neq i} B_j^k \times \mathcal{M}(X^k)^{I-1} \times ... \times \mathcal{M}(X^{k'-1})^{I-1} \right) \\ &\qquad = \mu^k \left(B_\Theta \times \prod_{j\neq i} B_j^1 \times ... \times \prod_{j\neq i} B_j^k \right), \end{aligned}$$

so the value of $\mu^\infty \left(B_\Theta \times \prod_{j\neq i} [B_j^1 \times ... \times B_j^k \times \mathcal{M}(X^k) \times \mathcal{M}(X^{k+1}) \times ...] \right)$ is independent of whether we use (13) with the given k, B_Θ, and B_j^ℓ, $j \neq i, \ell = 1, ..., k$, or whether we use (13) with $k' > k, k$, B_Θ, and B_j^ℓ, $j \neq i, \ell = 1, ..., k'$, with $B_j^\ell = \mathcal{M}(X^\ell)$ for $\ell \in \{k, ..., k'\}$.

The set function that is given by condition (13) defines a finitely additive meassure on the algebra of products of the form $B_\Theta \times \prod_{j\neq i}[B_j^1 \times ... \times B_j^k \times$

$\mathcal{M}(X^k) \times \mathcal{M}(X^{k+1}) \times ...]$, $k = 1, 2, ...$, By Kolmogorov's extension theorem, this set function can be uniquely extended to a countably additive measure μ^∞ on $\mathcal{B}(\Theta) \times \mathcal{B}^\infty(U_{-i})$.[11]

The mapping $(\mu^k)_{k\geq 1} \to \beta_i((\mu^k)_{k\geq 1})$ is one-to-one and onto, i.e., for any probability measure $\mu^\infty \in \mathcal{M}^\infty(\Theta \times U_{-i})$, there exists a unique belief hierarchy $(\mu^k)_{k\geq 1} \in U_i$ such that $\mu^\infty = \beta_i((\mu^k)_{k\geq 1})$. To see this, it suffices to note that, for any $\mu^\infty \in \mathcal{M}^\infty(\Theta \times U_{-i})$ and any k, (13) defines a measure μ^k on the algebra of products of the form $B_\Theta \times \prod_{j\neq i} B_j^1 \times ... \times \prod_{j\neq i} B_j^k$ and that this measure can be uniquely extended to a measure on $(X^k, \mathcal{B}(X^k))$.[12]

5 The Space $\mathcal{M}^\infty(\Theta \times U_{-i}^{uw})$

Construction of the mapping $(\mu^k)_{k\geq 1} \to \beta_i((\mu^k)_{k\geq 1})$ from U_i to $\mathcal{M}^\infty(\Theta \times U_{-i})$ in the preceding section does not presume anything about the topology on U_i or on $\Theta \times U_{-i}$. To be sure, in appealing to Kolmogorov's extension theorem, one exploits the topological properties of the underlying factor spaces, i.e. the spaces of beliefs or different orders and their domains, but then, no additional assumption about the topology on the spaces of belief hierarchies is involved.

If the spaces of belief hierarchies are given the product topology, we obviously have $\mathcal{B}(U_j^\pi) = \mathcal{B}^\infty(U_j)$ for all j and therefore $\mathcal{B}(\Theta) \times \mathcal{B}(U_{-i}^\pi) = \mathcal{B}(\Theta) \times \mathcal{B}^\infty(U_{-i})$, where

$$\mathcal{B}(U_{-i}^\pi) := \prod_{j\neq i} \mathcal{B}(U_j^\pi) \tag{14}$$

are the Borel σ-algebras on the product $\prod_{j\neq i} U_j$ when the spaces U_j have the product topology. The sets $\mathcal{M}^\infty(\Theta \times U_{-i})$ and $\mathcal{M}(\Theta \times U_{-i}^\pi)$ then are the same, and we may think of β_i as a mapping from U_i^π to $\mathcal{M}(\Theta \times U_{-i}^\pi)$. Mertens and Zamir (1985) show that, as a mapping from U_i^π to $\mathcal{M}(\Theta \times U_{-i}^\pi)$, β_i is actually a homeomorphism. Their argument relies on the fact that, if the product topology is imposed on $U_j, j \neq i$, as well as U_i, then the open sets of both the domain and the range of β_i are finite-dimensional cylinder sets.

[11] Billingsley (1968), p. 228, Dudley (2002), p. 257.

[12] Halmos (1950), p. 54, Dudley (2002), pp. 89ff.

This argument is not available if the spaces of belief hierarchies are endowed with the uniform weak topology, with open sets that *not* finite-dimensional cylinder sets. With the uniform weak topology, it might seem natural to think of beliefs as elements of the space $\mathcal{M}(\Theta \times U_{-i}^{uw})$, but then, as discussed in the introduction, $\mathcal{M}(\Theta \times U_{-i}^{uw})$ is *not* the range of the mapping $(\mu^k)_{k\geq 1} \to \beta_i((\mu^k)_{k\geq 1})$. Because the spaces $U_j^{uw}, j \neq i$, are non-separable, we have $\mathcal{B}(U_j^{uw}) \supsetneq \mathcal{B}^\infty(U_j)$ for all j, and therefore $\mathcal{B}(\Theta) \times \mathcal{B}^\infty(U_{-i}) \subsetneq \mathcal{B}(\Theta) \times \mathcal{B}(U_{-i}^{uw})$.[13] The measures in $\mathcal{M}(\Theta \times U_{-i}^{uw})$ and in $\mathcal{M}^\infty(\Theta \times U_{-i})$ do not even have the same domains.

If we think about $\mathcal{B}(\Theta) \times \mathcal{B}(U_{-i}^{uw})$ as the space of events about which agent i forms probabilistic beliefs, we must come to terms with the fact that agent i's belief hierarchies do not provide sufficient information to pin down his beliefs for all events that are relevant. Conversely, if we assume that all relevant aspects of agent i's beliefs are captured by the agent's belief hierarchies, we cannot treat the beliefs of agent i as elements of the space $\mathcal{M}(\Theta \times U_{-i}^{uw})$. The agent's belief hierarchies do not allow us to assign probabilities to events in the set $\left[\mathcal{B}(\Theta) \times \mathcal{B}(U_{-i}^{uw})\right] \backslash [\mathcal{B}(\Theta) \times \mathcal{B}^\infty(U_{-i})]$.

Given these considerations and taking the point of view that all strategically relevant aspects of agent i's beliefs are captured by the agent's belief hierarchies, it seems natural to think of the agent's beliefs as elements of the space $\mathcal{M}^\infty(\Theta \times U_{-i})$ of measures on $\mathcal{B}(\Theta) \times \mathcal{B}^\infty(U_{-i})$, i.e. the set of events to which one can assign probabilities on the basis of the belief hierarchies and to have the topology on this space reflect the fact that the spaces U_j, $j \neq i$, have the uniform weak topology. In the following, I will therefore consider beliefs as belonging to the space $\mathcal{M}^\infty(\Theta \times U_{-i}^{uw})$ of probability measures on $\mathcal{B}(\Theta) \times \mathcal{B}^\infty(U_{-i})$ and I will endow this space with the topology of uw-weak convergence, by which I mean convergence of integrals of bounded and continuous real-valued functions on $\Theta \times U_{-i}^{uw}$ that are also measurable with respect to the σ-algebra $\mathcal{B}(\Theta) \times \mathcal{B}^\infty(U_{-i})$. It is easy to see that this topology is equivalent to the topology that is induced by the projection from $\mathcal{M}(\Theta \times U_{-i}^{uw})$ to $\mathcal{M}^\infty(\Theta \times U_{-i})$, i.e. by the mapping that assigns to each measure μ on $(\Theta \times U_{-i}^{uw}, \mathcal{B}(\Theta) \times \mathcal{B}(U_{-i}^{uw}))$ the measure on $(\Theta \times U_{-i}, \mathcal{B}(\Theta) \times \mathcal{B}^\infty(U_{-i}))$ that is given by the restriction of μ to the smaller σ-algebra $\mathcal{B}(\Theta) \times \mathcal{B}^\infty(U_{-i})$.[14]

[13] Chen et al. (2016) give an example of a set that belongs to $\mathcal{B}(U_j^{uw})$ but not to $\mathcal{B}^\infty(U_j)$.

[14] In each case, the basic neighbourhoods around a measure μ are given by sets of the form

$$\left\{\hat{\mu} \middle| \left| \int f_k d\hat{\mu} - \int f_k d\mu \right| < \epsilon, k = 1, ..., n \right\}, \tag{*}$$

Although the domain of the measures in $\mathcal{M}^\infty(\Theta \times U_{-i}^{uw})$ does not coincide with the Borel σ-algebra on $\Theta \times U_{-i}^{uw}$, one can still define a version of the Prohorov metric. The Prohorov distance $p(\mu, \hat{\mu})$ between any two measures μ and $\hat{\mu}$ in $\mathcal{M}^\infty(\Theta \times U_{-i}^{uw})$ is defined the infimum of the set of ε such that (1) holds for all $B \in \mathcal{B}(\Theta) \times \mathcal{B}^\infty(U_{-i})$ with

$$B^\varepsilon := \{(\theta^\varepsilon, u_{-i}^\varepsilon) | \text{ for some } (\theta, u_{-i}) \in B, \max[d(\theta^\varepsilon, \theta), \max_{j \neq i} \rho^{uw}(u_j^\varepsilon, u_j)] < \varepsilon\}. \tag{15}$$

This definition presumes that the set B^ε belongs to the σ-algebra $\mathcal{B}(\Theta) \times \mathcal{B}^\infty(U_{-i})$ so that the terms $\mu(B^\varepsilon)$ and $\hat{\mu}(B^\varepsilon)$ are well defined. The following lemma ensures that this is indeed the case.

Lemma 1 *For any $B \in \mathcal{B}(\Theta) \times \mathcal{B}^\infty(U_{-i})$, let B^ε be given by (15). Then $B^\varepsilon \in \mathcal{B}(\Theta) \times \mathcal{B}^\infty(U_{-i})$.*

Proof. Consider the class $\mathcal{C} \subset \mathcal{B}(\Theta) \times \mathcal{B}^\infty(U_{-i})$ be the class of sets for which the lemma is true. It is easy to see that $\mathcal{C}$ contains the finite-dimensional cylinder sets.

Moreover, $\mathcal{C}$ is closed under countable unions: If $B_r, r = 1, 2, ...,$ is any countable family of sets in $\mathcal{C}$, a pair $(\theta^\varepsilon, u_{-i}^\varepsilon)$ belongs to the ε-neighbourhood of $\cup_r B_r$ if and only if it belongs to B_r^ε for some r. The ε-neighbourhood of $\cup_r B_r$ is therefore equal to the union $\cup_r B_r^\varepsilon$. Since $B_r \in \mathcal{C}$ implies $B_r^\varepsilon \in \mathcal{B}(\Theta) \times \mathcal{B}^\infty(U_{-i})$ and $\mathcal{B}(\Theta) \times \mathcal{B}^\infty(U_{-i})$ is closed under countable unions, it follows that $\cup_r B_r^\varepsilon \in \mathcal{B}(\Theta) \times \mathcal{B}^\infty(U_{-i})$ and hence that $\cup_r B_r \in \mathcal{C}$.

Finally, $\mathcal{C}$ is also closed under countable intersectons: If $B_r, r = 1, 2, ...,$ is any countable family of sets in $\mathcal{C}$, a pair $(\theta^\varepsilon, u_{-i}^\varepsilon)$ belongs to the ε-neighbourhood of $\cap_r B_r$ if and only if it belongs to B_r^ε for all r. The ε-neighbourhood of $\cap_r B_r$ is therefore equal to the intersection $\cap_r B_r^\varepsilon$. Since $B_r \in \mathcal{C}$ implies $B_r^\varepsilon \in \mathcal{B}(\Theta) \times \mathcal{B}^\infty(U_{-i})$ and $\mathcal{B}(\Theta) \times \mathcal{B}^\infty(U_{-i})$ is closed under countable intersections, it follows that $\cup_r B_r^\varepsilon \in \mathcal{B}(\Theta) \times \mathcal{B}^\infty(U_{-i})$ and hence that $\cap_r B_r \in \mathcal{C}$.

Thus, $\mathcal{C} = \mathcal{B}(\Theta) \times \mathcal{B}^\infty(U_{-i})$, and the lemma is proved. ■

Proposition 2 *Assume that all measures in $\mathcal{M}^\infty(\Theta \times U_{-i}^{uw})$ have separable suppports. Then the topology of uw-weak convergence on $\mathcal{M}^\infty(\Theta \times U_{-i}^{uw})$ is metrizable by the Prohorov metric p.*

---

where $\varepsilon > 0$ and $f_1, ..., f_n$ are bounded, continnuous real-valued functions on $\Theta \times U_{-i}^{uw}$. For $\mu \in \mathcal{M}^\infty(\Theta \times U_{-i}^{uw})$, the functions $f_1, ..., f_n$ must in addition be measurable with respect to $\mathcal{B}(\Theta) \times \mathcal{B}^\infty(U_{-i})$. By inspection of (*), it is clear that the open sets on $\mathcal{M}^\infty(\Theta \times U_{-i}^{uw})$ are exactly the projections of the open sets in $\mathcal{M}(\Theta \times U_{-i}^{uw})$.

Proof Sketch. By standard arguments, p is actually a metric on $\mathcal{M}^\infty(\Theta \times U_{-i}^{uw})$. To prove the proposition, it suffices to show that, any measure $\mu \in \mathcal{M}^\infty(\Theta \times U_{-i}^{uw})$, the topology generated by this metric and the topology of uw-weak convergence are equivalent. Under the assumption that the support of μ is separable, this follows from arguments given in the proof of Theorem 5, p. 238, of Billingsley (1968). In contrast to the situation here, Billingsley's theorem and proof concern the case where the measures are defined on the Borel σ-algebra that is generated by the topology on the underlying space. However, with uw-weak convergence defined in terms of integrals of functions that are continuous on $\Theta \times U_{-i}^{uw}$ and measurable with respect to $\mathcal{B}(\Theta) \times \mathcal{B}^\infty(U_{-i})$ and a Prohorov distance defined in terms of the metric on $\Theta \times U_{-i}^{uw}$ and measurable sets in $\mathcal{B}(\Theta) \times \mathcal{B}^\infty(U_{-i})$, the arguments given by Billingsley go through step by step without change. The details are left to the reader. ■

In Proposition 2, the condition that all measures in $\mathcal{M}^\infty(\Theta \times U_{-i}^{uw})$ have separable suppports is needed because the spaces $U_j^{uw}, j \neq i$, and therefore the product $\Theta \times U_{-i}^{uw}$ are non-separable. By the arguments given in Billingsley (1968), separability of the support of a measure μ is necessary as well as sufficient for the equivalence of the topology of uw-weak convergence and the topology induced by the Prohorov metric on $\mathcal{M}^\infty(\Theta \times U_{-i}^{uw})$.

By Theorem 2, p. 235, in Billingsley (1968), the condition that all measures in $\mathcal{M}^\infty(\Theta \times U_{-i}^{uw})$ have separable supports is necessarily satisfied if the set $\Theta \times U_{-i}$ has nonmeasurable cardinal, i.e., if there does not exist any atomless probability measure that is defined on all subsets of $\Theta \times U_{-i}^{uw}$. If the continuum hypothesis is assumed to be true, this condition is satisfied if the cardinality of the set $\Theta \times U_{-i}$ is no greater than that of the continuum.[15] Since $\Theta \times U_{-i}$ is a product of compact, hence separable, metric spaces, this latter condition is satisfied. The condition that all measures in $\mathcal{M}^\infty(\Theta \times U_{-i}^{uw})$ have separable suppports may therefore be taken for granted if the continuum hypothesis is assumed to be true.

[15] The discussion here concerns the so-called "problem of measure". See Billingsley (1968), pp. 233-236, and Dudley (2002), Appendix C.

6 The Homeomorphism Theorem for U_i^{uw} and $\mathcal{M}^\infty(\Theta \times U_{-i}^{uw})$

I now return to the topological properties of the Kolmogorov mapping $(\mu^k)_{k\geq 1} \to \beta_i((\mu^k)_{k\geq 1})$ when the spaces of belief hierarchies have the uniform weak topology rather than the product topology. The following result provides an analogue to the homeomorphism theorem of Mertens and Zamir (1985).

Proposition 3 *Assume that all measures in $\mathcal{M}^\infty(\Theta \times U_{-i}^{uw})$ have separable suppports. Then the mapping β_i defines a homeomorphism between U_i^{uw} and $\mathcal{M}^\infty(\Theta \times U_{-i}^{uw})$.*

Proof. Because the mapping β_i is injective and onto, it suffices to show that, both β_i and its inverse β_i^{-1} are continuous.

I first show that β_i is continuous. Proceeding indirectly, suppose that β_i is *not* continuous. Then there exists a sequence $\left\{(\mu^{kr})_{k\geq 1}\right\}_{r=1}^\infty$ and there exists $(\mu^k)_{k\geq 1}$ such that $(\mu^{kr})_{k\geq 1} \in U_i$ for all r, $(\mu^k)_{k\geq 1} \in U_i$,

$$\lim_{r\to\infty} \rho_i^{uw}((\mu^{kr})_{k\geq 1}, (\mu^k)_{k\geq 1}) = 0, \tag{16}$$

but $\mu^{\infty r} = \beta_i((\mu^{kr})_{k\geq 1})$ does not converge to $\mu^\infty = \beta_i((\mu^k)_{k\geq 1})$. To simplify the notaton, we write $u^r = (\mu^{kr})_{k\geq 1}$, $r = 1, 2, ...$, and $u = (\mu^k)_{k\geq 1}$. Taking subsequences if necessary, we may suppose that, for some $\varepsilon > 0$, the Prohorov distance between $\beta_i(u^r)$ and $\beta_i(u)$ exceeds ε for all r. For each r, therefore, there exists a set $W^r \in \mathcal{B}^\infty(\Theta \times U_{-i})$, with ε-neighbourhood $(W^r)^\varepsilon$, such that either

$$\beta_i(W^r|u^r) > \beta_i((W^r)^\varepsilon|u) + \varepsilon \tag{17}$$

or

$$\beta_i(W^r|u) > \beta_i((W^r)^\varepsilon|u^r) + \varepsilon. \tag{18}$$

For any n and any $j \neq i$, let U_j^n by the projection of U_j to the space $\mathcal{M}(X^0) \times ... \times \mathcal{M}(X^n)$, and let $U_{-i}^n := \prod_{j\neq i} U_j^n$. Further, let W^{nr} be the projection of W^r to $\Theta \times U_{-i}^n$ and let $(W^{nr})^\varepsilon$ be an ε-neighbourhood (in $\Theta \times U_{-i}^n$) of W^{nr}. Let

$$\hat{W}^{nr} = W^{nr} \times \prod_{j\neq i} [\mathcal{M}(X_{-j}^n) \times \mathcal{M}(X_{-j}^{n+1}) \times ...] \tag{19}$$

and

$$\hat{W}^{nr\varepsilon} = (W^{nr})^\varepsilon \times \prod_{j\neq i}[\mathcal{M}(X^n_{-j}) \times \mathcal{M}(X^{n+1}_{-j}) \times ...] \tag{20}$$

be the cylinder sets in $\Theta \times U_{-i}$ that are defined by W^{nr} and $(W^{nr})^\varepsilon$. One easily verifies that the sequences $\{\hat{W}^{nr}\}_{n=1}^\infty$ and $\{\hat{W}^{nr\varepsilon}\}_{n=1}^\infty$ are nonincreasing and that

$$W^r = \bigcap_{n=1}^\infty \hat{W}^{nr} \quad \text{and} \quad (W^r)^\varepsilon = \bigcap_{n=1}^\infty \hat{W}^{nr\varepsilon} \tag{21}$$

for all r. By elementary measure theory,[16] it follows that, for any r and any $\delta > 0$, there exists $N^r(\delta)$ such that for $n > N^r(\delta)$,

$$\beta_i((W^r)^\varepsilon|u) \geq \beta_i(\hat{W}^{nr\varepsilon}|u) - \delta \tag{22}$$

and

$$\beta_i((W^r)^\varepsilon|u^r) \geq \beta_i(\hat{W}^{nr\varepsilon}|u^r) - \delta. \tag{23}$$

Moreover,

$$\beta_i(W^r|u^r) \leq \beta_i(\hat{W}^{nr}|u^r) \tag{24}$$

and

$$\beta_i(W^r|u) \leq \beta_i(\hat{W}^{nr}|u) \tag{25}$$

Upon combining (22) - (25) with (17) and (18), we find that,for all r and $\delta > 0$, there exists $N^r(\delta)$ such that for $n > N^r(\delta)$, either

$$\beta_i(\hat{W}^{nr}|u^r) > \beta_i(\hat{W}^{nr\varepsilon}|u) - \delta + \varepsilon \tag{26}$$

or

$$\beta_i(\hat{W}^{nr}|u) > \beta_i(\hat{W}^{nr\varepsilon}|u^r) - \delta + \varepsilon. \tag{27}$$

By (13) and the definitions of $u^r = (\mu^{kr})_{k\geq 1}$ and $u = (\mu^k)_{k\geq 1}$, we also have

$$\beta_i(\hat{W}^{nr}|u^r) = \mu^{nr}(W^{nr}), \beta_i(\hat{W}^{nr\varepsilon}|u^r) = \mu^{nr}((W^{nr})^\varepsilon) \tag{28}$$

and

$$\beta_i(\hat{W}^{nr}|u) = \mu^n(W^{nr}), \beta_i(\hat{W}^{nr\varepsilon}|u) = \mu^n((W^{nr})^\varepsilon) \tag{29}$$

for all r and n. For any r and and any sufficiently large n. therefore, either

$$\mu^n(W^{nr}) > \mu^{nr}((W^{nr})^\varepsilon) - \delta + \varepsilon \tag{30}$$

[16]See Halmos (1950), p.38.

or

$$\mu^{nr}(W^{nr}) > \mu^n((W^{nr})^\varepsilon) - \delta + \varepsilon. \tag{31a}$$

If $\delta < \varepsilon$, we also have $(W^{nr})^{\varepsilon-\delta} \subset (W^{nr})^\varepsilon$ and $(W^r)^{\varepsilon-\delta} \subset (W^r)\varepsilon$, so one may infer that, for all r and $\delta \in (0,\varepsilon)$, there exists $N^r(\delta)$ such that for any $n > N^r(\delta)$, either

$$\mu^n(W^{nr}) > \mu^{nr}((W^{nr})^{\varepsilon-\delta}) - \delta + \varepsilon \tag{32}$$

or

$$\mu^{nr}(W^{nr}) > \mu^n((W^{nr})^{\varepsilon-\delta}) - \delta + \varepsilon. \tag{33a}$$

But then, for any r, for $n > N^k(\delta)$, the Prohorov distance between the measures μ^{nr} and μ^n is at least $\varepsilon - \delta$.

It follows that for all r, we have

$$\rho^{uw}((\mu^{kr})_{k\geq 1}, (\mu^k)_{k\geq 1}) \geq \varepsilon - \delta,$$

i.e. the distance between $u^r = (\mu^{kr})_{k\geq 1}$ and $u = (\mu^k)_{k\geq 1}$ in the metric for the uniform weak topology is at least $\varepsilon - \delta > 0$. This conclusion contradicts the assumption that the sequence $u^r = (\mu^{kr})_{k\geq 1}$ converges to $u = (\mu^k)_{k\geq 1}$ in the uniform weak topology. The assumption that the map $u_i \to \beta_i(u_i)$ is not continuous has thus led to a contradiction and must be false.

Continuity of the map from $\beta_i \in \mathcal{M}^\infty(\Theta \times U^{uw}_{-i})$ to the associated belief hierarchy is easily obtained by observing that, for any set $W^n \in \mathcal{B}(X^n_i)$ the associated cylinder set

$$\hat{W}^n = W^n \times \prod_{j\neq i}[\mathcal{M}(X^n_{-j}) \times \mathcal{M}(X^{n+1}_{-j}) \times ...] \tag{34}$$

belongs to $\mathcal{B}^\infty(\Theta \times U_{-i})$, and, for any $\varepsilon > 0$, the cylinder set

$$\hat{W}^{n\varepsilon} = (W^n)^\varepsilon \times \prod_{j\neq i}[\mathcal{M}(X^n_{-j}) \times \mathcal{M}(X^{n+1}_{-j}) \times ...] \tag{35a}$$

that is defined by the ε-neighbourhood $(W^n)^\varepsilon$ of W^n in X^n_{-i} is actually an ε-neighbourhood of $\hat{W}^n$ in $\Theta \times U^{uw}_{-i}$. Hence, if the Prohorov distance between two measures μ^∞ and $\hat{\mu}^\infty$ in $\mathcal{M}^\infty(\Theta \times U^{uw}_{-i})$ is less than ε, we must have

$$\mu^\infty(\hat{W}^n) < \hat{\mu}^\infty(\hat{W}^{n\varepsilon}) + \varepsilon \tag{36}$$

and

$$\hat{\mu}^{\infty}(\hat{W}^{n}) < \mu^{\infty}(\hat{W}^{n\varepsilon}) + \varepsilon. \tag{37}$$

By the definition of the marginal distributions, it follows that

$$\mu^{n}(W^{n}) < \hat{\mu}^{n}(W^{n\varepsilon}) + \varepsilon \tag{38}$$

and

$$\hat{\mu}^{n}(W^{n}) < \mu^{n}(W^{n\varepsilon}) + \varepsilon. \tag{39}$$

Since the choice of $W^n \in \mathcal{B}(X^n_{-i})$ was arbitary, it follows that the Prohorov distance between μ^n and $\hat{\mu}^n$ is no greater than ε. Since ε may be taken to be arbitrarily close to the Prohorov disctance between μ^∞ and $\hat{\mu}^\infty$, it follows that the Prohorov distance between μ^n and $\hat{\mu}^n$ is no greater than the Prohorov distance between μ^∞ and $\hat{\mu}^\infty$. Since this latter statement holds for all n, it follows that the supremum of the Prohorov distances between the marginal distributions μ^n and $\hat{\mu}^n$ for $n = 1, 2, ...$, is no greater than the Prohorov distance between μ^∞ and $\hat{\mu}^\infty$. Continuity of the map from measures on $\Theta \times U^{uw}_{-i}$ to belief hierarchies in U^{uw}_i follows immediately. ∎

7 Are U^{uw}_i and $\mathcal{M}^\infty(\Theta \times U^{uw}_{-i})$ "Universal"?

Whereas the homeomorphism theorem presented here fills a gap in our understanding of the universal type space with the uniform weak topology, another gap remains. The term "universal" type space reflects the fact that all abstract (Harsanyi) type spaces can be mapped into this space. In the framework of Mertens and Zamir (1985), the mapping is actually an embedding, i.e., any Harsanyi type space is homeomorphic to a subspace of the universal type space.

However, like the homeomorphism theorem relating U_i and $\mathcal{M}(\Theta \times U_{-i})$, this finding depends on the topologies that are imposed. What then can be said about the relation between abstract (Harsanyi) type spaces and the universal type space if the space of belief hierarchies is given the uniform weak topology?

To be more specific, consider an abstract type space model $\{T_i, \Theta_i, \theta_i, b_i\}_{i=1}^{I}$ such that, for any i, T_i and Θ_i are compact metric spaces, $\theta_i : T_i \to \Theta_i$ is a continuous function showing how the exogenous payoff parameters of agent i depend on the agent's abstract type, and $b_i : T_i \to \mathcal{M}(T_{-i})$ is a continuous function showing how the beliefs of agent i about the other agents' abstract types depend the agent's own type. As is well known from Mertens and Zamir (1985), the mappings θ_i and b_i can be used to construct beliefs about the vector $\theta = (\theta_1, ..., \theta_I)$ of payoff parameters, beliefs about the vector of payoff parameters and the vector of beliefs about payoff parameters, etc. This construction yields a mapping from the abstract type spaces to the space of belief hierarchies that are based on $\Theta = \prod_{j=1}^{I} \Theta_i$. The result of Mertens and Zamir (1985) shows that this mapping is continuous if, for each i, $\mathcal{M}(T_{-i})$ has the topology of weak convergence of probability measures and the spaces of belief hierarchies have the product topology.

However, the mapping from abstract type spaces to belief hierarchies need not be continuous if, for each i, $\mathcal{M}(T_{-i})$ has the topology of weak convergence of probability measures and the spaces of belief hierarchies have the uniform weak topology. For example, let $I = 2$ and consider the following version of Rubinstein's (1989) electronic mail game. For $i = 1, 2$, let $T_i = \{0, \frac{1}{2}, \frac{2}{3}, ..., 1\}$ and assume that T_i has the subspace topology that is induced by the usual topology on the unit interval. Let $\Theta_1 = \{0, 1\}$. $\Theta_2 = \{0\}$, and $\theta_1(t_1) = 0$ if $t_1 = 0$, $\theta_1(t_1) = 1$ if $t_1 > 0$. Specify a belief function b_1 for agent 1 so that $b_1(0) = \delta_0$ and, for $n = 1, 2, ...,$ $b_1(\frac{n}{n+1}) = \frac{1}{2}\delta_{(n-1)/n} + (1 - \frac{1}{2})\delta_{n/(n+1)}$ and $b_1(1) = \delta_1$, where for any $t \in [0, 1]$, δ_t is the degenerate measure that assigns all probability mass to the singleton $\{t\}$. Similarly, specify a belief function b_2 for agent 2 so that, for $n = 0, 1, 2, ...,$ $b_2(\frac{n}{n+1}) = \frac{1}{2}\delta_{n/(n+1)} + (1 - \frac{1}{2})\delta_{(n+1)/(n+2)}$ and $b_2(1) = \delta_1$. Then, for $i = 1, 2$, if $\mathcal{M}(T_{-i})$ has the topology of weak convergence of probability measures, the belief function b_i is continuous. However, the associated mappings from T_1 and T_2 into the $\Theta_1 \times \Theta_2$-based spaces U_1 and U_2 of belief hierarchies with the uniform weak topology are not continuous.

To see this, consider a strategic game in which each agent has a choice between two actions, a_0 and a_1. Suppose that, for each agent, action a_0 always gives the payoff zero, but action a_1 gives the payoff $Y > 0$ if $\theta_1 = 1$ and if the other agent also chooses the action a_1 and the payoff $-X < 0$ otherwise, where $X > Y$. Then, one easily verifies that, for each agent i, if $\varepsilon < \frac{1}{2}(Y - X)$, then for all $t_i \in T_i \backslash \{1\}$, a_0 is the unique ε interim correlated rationalizable action of agent i with the abstract type t_i, but for $t_i = 1$, the action a_1 is interim rationalizable. The lower strategic convergence property of Dekel et al. (2006) fails to hold. However, by Theorem 1 of Chen et al.

(2010), this property would hold if the mappings from T_1 and T_2 into the $\Theta_1 \times \Theta_2$-based spaces U_1 and U_2 of belief hierarchies with the uniform weak topology were continuous.

At this point, the question is what topology on $\mathcal{M}(T_{-i})$, $i = 1, ..., I$, would validate the statement that any abstract type space is homeomorphic to a subset of the associated universal type space when belief hierarchies are given the uniform weak topology (or the uniform strategic topology of Dekel et al. (2006)). A trivial answer would be the coarsest topology under which the mappings from abstract type spaces to belief hierarchies with the uniform weak topology are continuous. But then the question is whether this topology can be characterized in terms of the spaces T_i and $\mathcal{M}(T_{-i})$ themselves without reference to belief hierarchies and the uniform weak topology. An answer to this question would complete the program of recovering the homeomorphism results of Mertens and Zamir (1985) in a setting where beliefs of an arbitrarily high order can make a significant difference to strategic behaviour.

References

[1] Billingsley, P. (1968), *Convergence of Probability Measures*, Wiley, New York.

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

[3] Chen, Y., A. Di Tillio, E. Faingold, and S. Xiong (2010), Uniform Topologies on Types, *Theoretical Economics* 5, 445 -478.

[4] Chen, Y., A. Di Tillio, E. Faingold, and S. Xiong (2012), The Strategic Impact of Higher-Order Beliefs, Cowles Foundation Discussion Paper No. 1875, Yale University, http://ssrn.com/abstract=2142497.

[5] Chen, Y., A. Di Tillio, E. Faingold, and S. Xiong (2016), Corrigendum, mimeo, Yale University et al.

[6] Dekel, E., D. Fudenberg, and S. Morris (2006), Topologies on Types, *Theoretical Economics* 1, 275 - 309.

[7] Dudley, R.M. (2002), *Real Analysis and Probability*, Cambridge Studies in Advanced Mathematics 74, Cambridge University Press, Cambridge, UK.

[8] Gizatulina, A., and M. Hellwig (2016), The Genericity of the McAfee-Reny Condition for Full Surplus Extraction, Preprint 08/2015, revised 2016, Max Planck Institute for Research on Collective Goods, Bonn, Germany, http://www.coll.mpg.de/pdf_dat/2015_08online.pdf .

[9] Halmos, P.R. (1950), *Measure Theory*, Van Nostrand, New York.

[10] Harsanyi, J.C. (1967–68), Games with incomplete information played by 'Bayesian' players, I–III. *Management Science* 14, 159–182, 320–334, 486–502.

[11] Heifetz, A. (1993), The Bayesian Formulation of Incomplete Information – The Non-Compact Case, *International Journal of Game Theory* 21, 329 – 338.

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

[13] Heifetz, A., and D. Samet (1999), Coherent beliefs are not always types, *Journal of Mathematical Economics* 32, 475–488.

[14] McAfee, P., and P. Reny (1992), Correlated Information and Mechanism Design, *Econometrica* 60, 395 - 421.

[15] Mertens, J.F., and S. Zamir (1985), Formulation of Bayesian Analysis for Games with Incomplete Information, *International Journal of Game Theory* 14, 1–29.

[16] Parthasarathy, K. (1967), *Probability Measures on Metric Spaces*, Academic Press, New York and London.

[17] Rubinstein, A. (1989), The electronic mail game: Strategic behavior under 'almost common knowledge', *American Economic Review* 79, 385–391.