Discussion Papers of the
Max Planck Institute for
Research on Collective Goods
2016/17

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

Martin F. Hellwig

MAX PLANCK
SOCIETY

Discussion Papers of the
Max Planck Institute
for Research on Collective Goods 2016/17

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

Martin F. Hellwig

November 2016

This version: July 2022

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 Topology*

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

July 20, 2022

Abstract

The paper shows that the Mertens-Zamir (1985) reconciliation of belief hierarchy and type space models of incomplete information is robust to the requirement that the topology on belief hierarchies reflect the continuity properties of strategic behaviour, taking account of the fact that beliefs of arbitrarily high orders in agents' belief hierarchies can have a significant impact on strategic behaviour. When endowed with one of the finer topologies proposed by Fudenberg et al. (2006) and Chen et al. (2010, 2017), the space of belief hierarchies is still homeomorphic to the space of probability measures (beliefs) over exogenous data and other agents' belief hierarchies. The canonical mapping from nonredundant abstract type spaces with continuous belief functions to the space of belief hierarchies is an embedding if the range of belief functions has the topology of convergence in total variation.

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

JEL Classification: C70, C72.

*I am grateful to Alia Gizatulina for drawing my attention to the problem studied in this paper, to Eduardo Faingold and Siyang Xiong for enlightening me about Chen et al. (2010, 2016, 2017), and to Alan Sultan for enlightening me about his work with George Bachman, in particular, Bachman and Sultan (1980). I am also grateful to several referees for helpful comments, in particular, to referees of an earlier version for alerting me to a serious error.

1 Introduction

The Program of Mertens and Zamir (1985). The theory of games with incomplete information took off when Harsanyi (1967/68) introduced the notion of an agent's "type" as determining the agent's payoff function and the agent's beliefs about exogenous parameters of the game and about the other agents' "types". Previously, there had been a sense that thinking about incomplete information involves an infinite regress, beginning with agents' beliefs about the exogenous parameters and going on with their joint beliefs about the exogenous parameters and the other agents' beliefs about the exogenous parameters, their joint beliefs about the exogenous parameters and the other agents' joint beliefs about the exogenous parameters and yet other agents' beliefs about the exogenous parameters, and so on. Harsanyi's notion of a "type" as determining beliefs represented by a single probability measure over exogenous parameters and over other agents' "types" avoided the hierarchies of beliefs, beliefs about beliefs, etc. and brought games of incomplete information into the domain of standard decision theory and game theory. However, the gain in tractability from the avoidance of the infinite regress came at a cost because the notion of "type" is a black box and the notion that an agent's "type" determines the agent's beliefs about exogenous parameters and about the other agents' "types" involves an element of circularity.[1]

Whereas Harsanyi left open the question how "types" are related to belief hierarchies, Mertens and Zamir (1985) developed a mathematical framework in which the belief hierarchies themselves can be interpreted as "types". In this framework, Kolmogorov's extension theorem implies that any coherent belief hierarchy of an agent defines a unique probability measure over exogenous parameters and other agents' belief hierarchies such that the finite-order beliefs generated by this measure are exactly the finite-order beliefs in the given hierarchy.[2] If belief hierarchies are interpreted as "types", the mapping from an agent's belief hierarchies to probability measures over exogenous parameters and other agents' belief hierarchies that is provided by Kolmogorov's theorem can be interpreted as a mapping from "types" to joint probability measures over exogenous parameters and other agents' "types", just like a belief mapping in Harsanyi's approach. In this formulation, the element of circularity inherent in the notion of "type" is apparent rather than real, an artefact of the infinite dimension of the spaces of belief hierarchies.

[1] For an extensive discussion of the strengths and weakness of thinking in terms of belief hierarchies versus thinking in terms of type spaces, see Heifetz and Samet (1998).

[2] A belief hierarchy is coherent if the beliefs of different orders are mutually compatible.

The analysis of Mertens and Zamir (1985) was subsequently generalized by Brandenburger and Dekel (1993) and Heifetz (1993).

This body of work also shows that, if the beliefs of any order are given the topology of weak convergence of probability measures (the weak* topology) and if belief hierarchies are given the product topology, then the mapping from belief hierarchies to beliefs over exogenous data and other agents' belief hierarchies is a homeomorphism, one-to-one, bi-continuous, and onto. Moreover, for any nonredundant abstract type space with belief functions à la Harsanyi, the natural mapping from abstract types to hierarchies of beliefs of different orders is an embedding, a homeomorphism between the abstract type space and some subspace of the space of belief hierarchies. In some of the literature, the space of belief hierarchies is therefore called the *universal type space*.[3]

Heifetz and Samet (1998, 1999) have criticized this literature for what they consider to be an excessive reliance on topological assumptions. The use of Kolmogorov's extension theorem presumes that the space of belief hierarchies is treated as a product of topological spaces. Heifetz and Samet (1999) show that, without topological assumptions, there may exist coherent belief hierarchies that are not compatible with any probability measure over exogenous parameters and other agents' belief hierarchies. Also, without topological assumptions, Heifetz and Samet (1998) show that with purely measure theoretic tools one can still obtain a type space into which any other type space can be mapped by a belief-preserving function. The elements of this universal type space can be identified with coherent belief hierarchies, but not all coherent belief hierarchies belong to it.

From a measure theoretic perspective, the Heifetz-Samet critique is important because it clarifies the conditions that underlie the interpretation of belief hierarchies as types in the sense of the Harsanyi formalism.[4] From a

[3] See, e.g., Brandenburger and Dekel (1993), Dekel et al. (2006, 2007), Chen at al. (2010, 2017). Mertens and Zamir (1985) refer to the space of belief hierarchies as the *universal belief space*; they use the term universal type space for the space of measures induced by a belief hierarchy. Heifetz (1993) as well as Heifetz and Samet (1998, 1999) also refer to "types" as measures.

[4] However, from the perspective of Gul (1998), the precedence that Heifetz and Samet give to "types" over coherent belief hierarchies would seem to be questionable. Whereas Heifetz and Samet (1998) talk about the universal type space as a set of *states of the world*, Gul (1998) seems to think about an agent's belief hierarchy as a matter concerning the agent's own reasoning. His concerns about the metaphysical assumptions needed to assume the existence of a common prior would also seem to apply to the objectification inherent in thinking about the different agents' reasonings as pertaining to a common space of states of the world. The focus of Dekel et al. (2006, 2007) on interim rationalizable as opposed to equilibrium outcomes can also be interpreted in this individualistic vein.

game theoretic perspective, however, the Heifetz-Samet critique is moot if the topological assumptions that they criticize are needed for game theoretic purposes anyway. One needs topological assumptions to study the continuity properties of agents' strategic behaviours, i.e., of the mappings that relate the choices agents take to the data of their decision problems. The topological assumptions that yield desirable continuity properties of behaviours are also sufficient for the use of Kolmogorov's extension theorem.

The homeomorphism theorems of Mertens and Zamir (1985) and subsequent papers should be seen in this context. These theorems indicate that, for the continuity properties of an agent's strategic behaviour, it does not make a difference whether one thinks about the agent's information in terms of belief hierarchies or in terms of beliefs over exogenous parameters and other agents' belief hierarchies. For any continuity property in one approach there is an equivalent continuity property in the other approach and vice versa. Similarly, it does not make a difference whether one thinks about strategic behaviour in terms of abstract types à la Harsanyi or in terms of the belief hierarchies that are obtained by the natural mapping from abstract types to belief hierarchies.

In the purely measure-theoretic approach of Heifetz and Samet (1998), there is no analogue of these homeomorphism theorems. Without topologies, one cannot study the continuity properties of strategic behaviour.

The Critique of Dekel et al. (2006). Continuity properties of behaviours depend on the topologies that are imposed on the different spaces. The homeomorphism theorems of Mertens and Zamir (1985) and subsequent papers rely on the assumptions that the beliefs of any order have the topology of weak convergence of probability measures and that belief hierarchies have the product topology. The question is whether this specification of topologies provides a good basis for studying continuity. Dekel et al. (2006) assert that the answer to this question is positive if one is only interested in upper hemi-continuity and negative if one is also interested in lower hemi-continuity.

Upper strategic convergence property: A behaviour correspondence is upper hemi-continuous if, for any convergent sequence of exogenous data and any convergent sequence of choices induced by these data, the limit of the sequence of choices belongs to the set of choices that the behaviour correspondence stipulates for the limit of the sequence of exogenous data. For the topologies used by Mertens and Zamir (1985), or for any finer topology, for optimizing choices of an agent, this property follows by standard arguments based on the maximum theorem.[5] In games of complete information, this upper hemi-continuity property of optimizing behaviour translates directly into an upper hemi-continuity property of Nash equilibrium and rationalizable-outcome correspondences. Similarly, Dekel et al. (2006) show, that, in games of incomplete information, for the indicated topology on the space of belief hierarchies, the correspondence mapping belief hierarchies into interim correlated rationalizable outcomes has what they call the upper strategic convergence property (Theorem 2, p. 294). They also show that this property fails to hold for any coarser topology on the space of belief hierarchies (Theorem 1, p. 294).

Lower strategic convergence property: A behaviour correspondence is lower hemi-continuous if, for any convergent sequence of exogenous data and any choice that the behaviour correspondence stipulates for the limit of the sequence of exogenous data, there exists a sequence of stipulated choices associated with the sequence of exogenous data that converges to the choice taken at the limit of the sequence of exogenous data. Unless best responses are everywhere unique, optimizing behaviours will usually not be lower hemi-continuous, nor will equilibrium and rationalizable outcomes.

However, following Fudenberg and Levine (1986), the literature has also considered the lower hemi-continuity properties of solution concepts involving behaviours that are only ε-optimal, so stipulated choices may fall short of the optimum by less than ε. If the ε-optimality conditions are written as strict inequalities, then in games of complete information, lower hemi-continuity holds for ε-optimal choices, as well as ε-Nash equilibrium and ε-rationalizable outcomes.[6]

Given this observation, Dekel et al. (2006) have proposed to add what they call a lower strategic convergence property to the desiderata for a topology on the space of belief hierarchies of an agent. They formulate this lower strategic convergence property in terms of interim correlated ε-rationalizable choices. In Dekel et al. (2007), they show that all types that have the same hierarchies of beliefs also have the same set of interim correlated rationalizable outcomes and conclude that this solution concept characterizes common certainty of rationality in the universal type space.

However, the product topology on the space of belief hierarchies is too coarse for the lower strategic convergence property. The reason is that, under

[5] See Berge (1959), p. 121.

[6] Dekel et al. (2006, 2007) specify ε-best-response conditions as weak inequalities. In this case lower hemi-continuity need not hold, but the lower strategic convergence property implies that the mapping from parameters to the smallest ε for which lower hemi-continuity holds is lower semi-continuous.

the product topology, two belief hierarchies can be similar even though, for very large k the associated beliefs of order k are very different. Therefore the lower strategic convergence fails whenever the set of interim correlated ε-rationalizable outcomes is sensitive to differences in beliefs of arbitrarily high orders. An example of this phenomenon is provided by Rubinstein's (1989) electronic mail game. Ely and Peski (2011) provide a systematic analysis of its occurrence.

Finer Topologies. Because of this shortcoming of the product topology, Dekel et al. (2006) and Chen et al. (2010, 2012) have proposed finer topologies for the space of belief hierarchies. The *strategic topology* of Dekel et al. (2006) is explicitly specified with reference to the desired upper and lower convergence properties of interim correlated ε-rationalizable choices, with open sets defined so that these continuity properties hold over all strategic games with the given exogenous parameter spaces. The *uniform strategic topology* of Dekel et al. (2006) is defined so that these continuity properties hold uniformly over all strategic games with the given exogenous parameter spaces.

Whereas Dekel et al. (2006) specify their topologies with explicit references to the desired continuity properties of strategic behaviour, Chen et al. (2010, 2017) introduce a topology that is defined only in terms of the beliefs of different orders in a hierarchy, without any reference to strategic behaviour. In contrast to the product topology, which is not very sensitive to beliefs of very high orders, their *uniform weak topology* gives equal weight to beliefs of all orders. This property ensures the lower and the upper strategic convergence properties of interim correlated ε-rationalizable choices. Chen et al. (2010, 2017) show that their uniform weak topology is actually equivalent to the uniform strategic topology of Dekel et al. (2006). Given this equivalence, I propose to use the simpler term *uniform topology* for both.

Chen et al. (2017) also show that a weaker form of uniform weak convergence, *uniform weak convergence on frames*, is equivalent to convergence in the strategic topology of Dekel et al. (2006). By restricting the uniformity of convergence to frames, they allow for the fact that, unlike the uniform strategic topology, the strategic topology of Dekel et al. (2006) does not require strategic convergence to be uniform over all games.

Objective of this Paper. With the new topologies of Dekel et al (2006) and Chen et al. (2010, 2017), it is not clear what becomes of the homeomorphism theorems of Mertens and Zamir (1985) and Brandenburger and Dekel (1993). What can be said about the continuity properties of the mapping from the belief hierarchies of an agent to the set of probability measures over exogenous parameters and the belief hierarchies of other agents that is provided by Kolmogorov's extension theorem?

The change in topologies on belief hierarchies affects both the domain and the range of this mapping. If we think about the domain of the mapping, having more open sets works in favour of continuity; if we think about the range of the mapping, having more open sets works against continuity.[7] What can one say about the balance of these countervailing effects?

In fact, these two countervailing effects just cancel each other out: The mapping from the belief hierarchies of an agent to the set of probability measures over exogenous parameters and the belief hierarchies of other agents is a homeomorphism even if the spaces of belief hierarchies have the uniform topology.

Measure Theoretic Issues. Before one can address this issue, however, one must deal with a fundamental difficulty. The Borel σ-algebra that is generated by the uniform topology is strictly finer than the Borel σ-algebra generated by the product topology. Kolmorogorov's extension theorem only provides for the assignment of probabilities to events in the Borel σ-algebra for the product topology. How then should we think about probabilistic beliefs over exogenous data and other agents' types when the space of other agents' types has the uniform topology?

One might think of extending the measure provided by Kolmogorov's theorem to the Borel σ-algebra for the uniform weak topology. However, under the usual set theoretic assumptions, such an extension need not exist. Moreover, if an extension exists, it need not be unique.

The measure provided by Kolmogorov's theorem may be concentrated on an uncountable set that is discrete in the uniform topology and may assign probability zero to singletons. Every subset of this uncountable discrete set belongs to the Borel σ-algebra for the uniform weak topology. An extension of the measure provided by Kolmogorov's theorem from the Borel σ-algebra for the product topology to the Borel σ-algebra for the uniform topology would therefore have to assign a probability to *every* subset of the uncountable discrete set. This is incompatible with a standard axiom for the set theory underlying most probability theory, that the cardinal c of the

[7] The range of the belief mapping consists of measures over exogenous parameters and other agents' belief hierarchies. If the space of these measures has the topology of weak convergence, i.e., convergence of integrals of bounded continuous functions, having more open sets in the space of other agents' belief hierarchies implies that there are also more open sets in the space of such measures.

continuum is not atomlessly measurable.[8]

In view of this problem, I consider the σ-algebras in the spaces of belief hierarchies that are generated by the open balls, rather than all open sets, in the uniform topology. [9] Whereas any open set is an arbitrary union of open spheres, the σ-algebra generated by the open balls only allows for countable unions and is strictly coarser than the σ-algebra generated by the open sets, i.e., the Borel σ-algebra.

For the spaces of belief hierarchies, the σ-algebra that is generated by the open balls actually coincides with the σ-algebra that is generated by the product topology, as well as the σ-algebra in the purely measure theoretic construction of Heifetz and Samet (1998). Therefore, the construction of the mapping from an agent's belief hierarchies to probability measures over exogenous parameters and other agents' belief hierarchies poses no difficulties. Kolmogorov's extension theorem provides for an unambiguous assignment of probabilities to events in the specified σ-algebra and for a well-defined mapping from belief hierarchies of an agent to probability measures over exogenous parameters and belief hierarchies of the other agents.

However, the domain of this mapping now has the uniform topology rather than the product topology. The range of this mapping is given the topology of weak convergence, defined in terms of convergence of integrals of bounded continuous real-valued functions that are also measurable with respect to the specified σ-algebra.[10] The set of continuous functions to be considered depends on the topology, so with the uniform topology on the spaces of other agents' belief hierarchies, the set of such functions is larger than with the product topology. The induced topology of weak convergence is therefore strictly finer than the topology of weak convergence that in induced when the spaces of other agents' belief hierarchies have the product

[8]The question whether c is atomlessly measurable is at the core of what Banach and Kuratowski (1929) call the "probem of measure". Following Ulam (1930), for a long time, the literature formulated the "problem of measure" in terms of the distinction between "measurable" and "nonmeasurable" cardinals; see e.g., Appendix III in Billingsley (1968). Nowadays the distinction "measurable/nonmeasurable" is used for the case where the measures under consideration are two-valued. For the problem of measure as posed in Banach and Kuratowski (1929), the terminology is "atomlessly measurable/not atomlessly measurable". For more on the problem of measure, including its relation to the Continuum Hypothesis, see the end of Section 5 below.

[9]These σ-algebras were introduced by Dudley (1966,1967). He was concerned with the formulation of convergence theorems for stochastic processes exhibiting jump discontinuities. He introduced the σ-algebras that are generated by open balls (in the uniform topology) as an alternative to working with the Borel σ-algebras for the Skorokhod topology, as in Billingsley (1968).

[10]This topology was introduced by Dudley (1966, 1967). See also Pollard (1979).

topology. The greater fineness of the topology on the range exactly matches the greater fineness of the topology on the domain of the Kolmogorov mapping so that the homeomorphism property of the mapping is preserved.

Embedding Abstract Type Spaces in the Universal Type Space. I go on to study what the imposition of the uniform topology on the space of belief hierarchies does to the Mertens-Zamir result that every abstract type space with continuous belief functions can be embedded in the space of belief hierarchies. Two results are obtained. First, not surprisingly, if the ranges of the belief functions in an abstract-type-space model, i.e. probability measures over exogenous data and other agents' types are given the topology of weak convergence, the mappings from abstract type spaces to belief hierarchies are *not* generally continuous when the spaces of belief hierarchies have the uniform topology. Second, if the ranges of belief functions in an abstract-type-space model are given the topology of convergence in total variation, the mappings from abstract type spaces to belief hierarchies are continuous even when the spaces of belief hierarchies have the uniform topology. For nonredundant compact type spaces, these mappings are again embeddings. A remaining open question is whether this finding could also be obtained with a coarser topology on the range of agents' belief functions in an abstract type space model.

Relevance for Genericity Considerations. The homeomorphism theorem for the relation between belief hierachies and probability measures on exogenous parameters and other agents' belief hierarchies matters not only for continuity but also for genericity considerations. An example is provided by the scope for full surplus extraction in mechanism design with correlated values, studied in Gizatulina and Hellwig (2017). The analysis in Gizatulina and Hellwig (2017) focuses on a condition that McAfee and Reny (1992) introduced as being necessary and sufficient for full surplus extraction to be feasible. This condition 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 this condition for belief hierarchies. The homeomorphism theorem of this paper implies that, if the universal type space has the uniform topology, any genericity results that are obtained from the McAfee-Reny condition can be translated into genericity results about the space of belief hierarchies.

Another example concerns the (non-)genericity of common priors. The very notion of a common prior refers to the mappings from agents' types to their probabilistic beliefs about exogenous parameters and other agents' types. A common prior is a prior probability distribution on the space of exogenous parameters and all agents' types under which, for every agent, this mapping is a regular conditional distribution for exogenous parameters and other agents' types given the agent's own type. Given this definition, one would expect the question of genericity or sparseness of models with common priors to be addressed in terms of the space of priors, i.e. probability measures on the space of exogenous parameters and all agents' types.[11]

Chen et al. (2017) instead study this question in terms of belief hierarchies, introducing notions of "common-prior types" and "non-common-prior types" (belief hierarchies) and proving that "non-common-prior types" are generic in the strategic topology. Perhaps the homeomorphism theorem of this paper will provide a basis for linking their result with an analysis of the (non-)genericity of common priors when treated as measures on the space of exogenous parameters and other agents' belief hierarchies.

Outline. Section 2 provides a brief review of the main results on belief hierarchies with the product topology. Section 3 discusses the critique of Dekel et al. (2006) and introduces the uniform topology on belief hierarchies. Section 4 discusses the difficulties involved in using the Borel σ-algebra induced by the uniform topology on belief hierarchies and instead introduces the σ-algebra generated by the open balls in the uniform topology. Section 5 formulates and proves the homeomorphism theorem for the Kolmogorov mapping when belief hierarchies have the uniform topology and beliefs are defined as probability measures on the σ-algebra generated by the open balls in the uniform topology. Section 6 studies the problem of embedding abstract type spaces in the space of belief hierarchies.

2 Belief Hierarchies with the Product Topology: A Review

In constructing the space of belief hierarchies, I follow Brandenburger and Dekel (1993). Their formulation is simpler than that of Mertens and Zamir (1985). It is also more explicit about the role of Kolmogorov's extension theorem.

Let Θ be a non-singleton compact metric space of exogenous parameters.

[11] See for example Hellman and Samet (2012).

Suppose that there are $I = 2$ agents.[12] Proceeding inductively, define spaces $X^0, X^1, X^2, ...$ by setting

$$X^0 = \Theta, \quad X^1 = X^0 \times \mathcal{M}(X^0), \tag{2.1}$$

and, for $k = 2, 3, ...,$

$$X^k = \left\{(\theta, \mu^1, ..., \mu^k) \in X^0 \times \mathcal{M}(X^0) \times ... \times \mathcal{M}(X^{k-1}) | \operatorname{marg}_{X^{k-2}}\mu^k = \mu^{k-1}\right\}, \tag{2.2}$$

where $\operatorname{marg}_{X^{k-2}}\mu^k$ is the marginal distribution on X^{k-2} that is induced by μ^k and, for any $\ell < k$, $\mathcal{M}(X^\ell)$ is the space of probability measures on the Borel σ-algebra $\mathcal{B}(X^\ell)$ on X^ℓ.

Throughout the paper, I write $\mathcal{B}(X)$ for the Borel σ-algebra on a metric space X and $\mathcal{M}(X)$ for the space of probability measures on $(X, \mathcal{B}(X))$, endowed with the topology of weak convergence (the weak* topology). In this topology, a sequence $\{\mu^r\}$ in $\mathcal{M}(X)$ converges to a limit $\mu \in \mathcal{M}(X)$ if and only if, for every bounded continuous real-valued function on X, the integrals $\int f(x)d\mu^r(x)$ converge to $\int f(x)d\mu(x)$. Because the continuity of a real-valued function f on X depends on the topology on X, the notion of weak convergence of probability measures on $(X, \mathcal{B}(X))$ also depends on the topology on X.

If X is a compact metric space and $\mathcal{M}(X)$ has the topology of weak convergence, then $\mathcal{M}(X)$ is also a compact metric space.[13] By a straightforward induction therefore, the assumption that $X^0 = \Theta$ is a compact metric space implies that, for any $k > 0$, $\mathcal{M}(X^{k-1})$ and X^k are also compact metric spaces.

The spaces $X^0, X^1, X^2, ...$ contain the objects about which agent i forms beliefs of orders $1, 2, 3, ...,$ namely the exogenous parameters and the lower-order beliefs of agent $j \neq i$. The condition $\operatorname{marg}_{X^{k-2}}\mu^k = \mu^{k-1}$ in (2.2) indicates that agent j's beliefs are coherent in the sense that, for each k, agent j's beliefs of orders k and $k-1$ order beliefs assign the same joint distribution to the exogenous parameter and the belief of order $k-2$ of agent i.

A *belief hierarchy* for player $i \in \{1, 2\}$ is defined as a sequence $\{\mu_i^k\}_{k=1}^\infty$of measures such that $\mu_i^k \in \mathcal{M}(X^{k-1})$ for $k = 1, 2, ...$ and

$$\operatorname{marg}_{X^{k-2}}\mu_i^k = \mu_i^{k-1} \tag{2.3}$$

[12] For general I, the notation would be more complicated, but the analysis would be the same.

[13] See, e.g., Parthasarathy (1967), p. 45.

for $k = 2, 3, ...$ Condition (2.3) ensures that player i's own beliefs of different orders are themselves coherent, like the beliefs of agent j that agent i considers possible.

For $i \in \{1, 2\}$, let H_i be the set of all belief hierarchies of player i. By construction, this set is the same for both agents. Even so, it will be useful to write H_i and H_j to indicate whose belief hierarchies are meant. These sets are subsets of the infinite product

$$\bar{H} := \prod_{k \geq 1} \mathcal{M}(X^{k-1}). \tag{2.4}$$

The sets H_i and H_j are interrelated. For any k, let π_k be the projection from $\bar{H}$ to $\mathcal{M}(X^{k-1})$ and let $\pi^k = (\pi_1, ...\pi_k)$ be the projection from $\bar{H}$ to $\mathcal{M}(X^0) \times ... \times \mathcal{M}(X^{k-1})$. Then

$$\pi_k(H_i) = \pi_j(H_j) = \mathcal{M}(X^{k-1}), \tag{2.5}$$

and one can rewrite (2.2) in the form

$$X^k = \Theta \times \pi^k(H_j). \tag{2.6}$$

Upon combining these two equations, with k replaced by $k-1$ in (2.6), one obtains

$$\pi_k(H_i) = \mathcal{M}(\Theta \times \pi^{k-1}(H_j)) \tag{2.7}$$

for $k = 1, 2, ...$ A belief hierarchy for agent i can be interpreted as a family of measures on the projections from the space $\Theta \times H_j$ of exogenous parameters and belief hierarchies of agent $j \neq i$ to the finite-dimensional products $\Theta \times \mathcal{M}(X^0) \times ... \times \mathcal{M}(X^{k-1})$.

Proposition 2.1 *Let $\bar{H}$ have the product topology and let H_1, H_2 have the associated subspace topologies. Then, for $i \in \{1, 2\}$ and every belief hierarchy $\{\mu_i^k\}_{k=1}^{\infty}$ in H_i, there exists a unique probability measure*

$$\mu_i^{\infty} = \beta(\{\mu_i^k\}_{k=1}^{\infty}) \in \mathcal{M}(\Theta \times H_j) \tag{2.8}$$

such that, for any k, the marginal distribution on $X^{k-1} = \Theta \times \pi^{k-1}(H_j)$ that is induced by μ_i^{∞} is equal to the k^{th} order belief μ_i^k in the hierarchy $\{\mu_i^k\}_{k=1}^{\infty}$.

Proof. The proposition is implied by Proposition 2 of Brandenburger and Dekel (1993). I nevertheless give a proof in order to provide a better starting point for the subsequent discussion. First, for any k, the measure μ_i^k in the belief hierarchy can be used to define a measure μ_i^∞ on the "cylinder" σ-algebra

$$\mathcal{B}(\Theta \times \pi^{k-1}(H_j)) \times \{\emptyset, \mathcal{M}(X^k)\} \times \{\emptyset, \mathcal{M}(X^{k+1})\} \times ... \qquad (2.9)$$

by setting

$$\mu_i^\infty(B_k \times \mathcal{M}(X^k) \times \mathcal{M}(X^{k+1}) \times ...) = \mu_i^k(B_k) \qquad (2.10)$$

for any $B_k \in \mathcal{B}(\Theta \times \pi^{k-1}(H_j))$. The coherence condition (2.3) ensures that the specifications (2.10) that are given for different k are mutually compatible. For example, for B_k taking the form $B_{k-1} \times \mathcal{M}(X^{k-1})$, (2.10) and (2.3) yield

$$\mu_i^\infty(B_{k-1} \times \mathcal{M}(X^k) \times \mathcal{M}(X^{k+1}) \times ...) = \mu_i^k(B_k) = \mu_i^{k-1}(B_{k-1}),$$

which is just (2.10) with k replaced by $k-1$.

As specified, μ_i^∞ is a finitely additive measure on the "cylinder" sets $B_k \in \mathcal{B}(\Theta) \times \mathcal{B}(\mathcal{M}(X^0)) \times ... \times \mathcal{B}(\mathcal{M}(X^{k-1})) \times \mathcal{M}(X^k) \times \mathcal{M}(X^{k+1}) \times ...$, $k = 1, 2, ...$ By Kolmogorov's extension theorem, this measure has a unique extension to a countably additive measure on the product σ-algebra $\mathcal{B}(\Theta \times H_j)$. ■

Proposition 2.1 defines a mapping

$$\{\mu_i^k\}_{k=1}^\infty \longmapsto \beta(\{\mu_i^k\}_{k=1}^\infty) \qquad (2.11)$$

from the space H_i of agent i's belief hierarchies to the space $\mathcal{M}(\Theta \times H_j)$ of probability measures on $(\Theta \times H_j, \mathcal{B}(\Theta \times H_j))$. I call this mapping the *Kolmogorov mapping.*

The Kolmogorov mapping is invertible: For any $\mu_i^\infty \in \mathcal{M}(\Theta \times H_j)$, one has

$$\mu_i^k = \mu_i^\infty \circ (proj_{X^{K-1}})^{-1}, \qquad (2.12)$$

and therefore

$$\{\mu_i^\infty \circ (proj_{X^{k-1}})^{-1}\}_{k=1}^\infty = \beta^{-1}(\mu_i^\infty), \qquad (2.13)$$

where $proj_{X^{K-1}}$ is the projection from $\Theta \times H_j$ to X^{k-1}.[14]

To emphasize that H_i and H_j have the product topology, I write H_i^p and H_j^p rather than H_i and H_j. With this topology, one easily verifies that the mappings β and β^{-1} are both continuous. Thus, one obtains:

[14] The argument in Mertens and Zamir (1985) uses (2.12) to obtain μ_i^∞ as a projective limit of the measures μ_i^k. In this argument, the problem of assigning probabilities to events that do not belong to the algebra of cylinder sets and the recourse to Kolmogorov's theorem for this purpose appear implicitly rather than explicitly.

Proposition 2.2 *The Kolmogorov mapping from H_i^p to $\mathcal{M}(\Theta \times H_j^p)$ is a homeomorphism.*

Propositions 2.1 and 2.2 provide the basis for a reconciliation of belief hierarchy and type space approaches to incomplete information. For two agents, a Θ-based abstract type space model is given by a pair $T_i, i = 1, 2$, of compact metric spaces and a pair $b_i : T_i \to \mathcal{M}(\Theta \times T_j)$, $i = 1, 2$ and $j \neq i$, of continuous belief mappings with the interpretation that, if $t_i \in T_i$ is the "type" of agent i, then $b_i(t_i) \in \mathcal{M}(\Theta \times T_j)$ is the agent's probabilistic belief about the exogenous parameter $\theta \in \Theta$ and the type $t_j \in T_j$ of the other agent.

Propositions 2.1 and 2.2 show that, if we think of belief hierarchies as "types", we have exactly the Harsanyi format, with $\beta(\cdot)$ as a belief function that maps a player's types into beliefs, i.e., probability measures over exogenous parameters and the other player's types.[15] Upon setting $T_i = H_i^p$ and $b_i(\cdot) = \beta(\cdot)$, $i = 1, 2$, one obtains an abstract type space model that differs from the general formulation only in that the belief functions are given by the Kolmogorov mapping. The fact that the Kolmogorov mapping is a homeomorphism implies that the space of belief hierarchies of agent i and the space of probability measures over exogenous parameters and belief hierarchies of agent j are topologically equivalent, so that, for any statement about the continuity properties of functions and correpondences on one space, there is an equivalent statement about continuity properties of functions and correspondences on the other space.

The triple $\Theta \times H_1^p \times H_2^p$ is called the *universal* type space because every nonredundant Θ-based abstract type space can be embedded in it: Given compact metric spaces T_i and continuous belief functions $b_i : T_i \to \mathcal{M}(\Theta \times T_j)$, $i = 1, 2$, for any i and t_i, one obtains a function

$$t_i \longmapsto \{\varphi_i^k(t_i)\}_{k=1}^{\infty} \tag{2.14}$$

from T_i into H_i by setting

$$\varphi_i^1(t_i) = b_i(\cdot|t_i) \circ (\pi_\Theta)^{-1} \tag{2.15}$$

[15] Different papers in the literature assign different meanings to the word "type". The identification of "types" with belief hierarchies follows Brandenburger and Dekel (1993), see also Dekel et al. (2006) and Chen et al. (2010, 2017). In contrast, Mertens and Zamir (1985) and other authors define "types" as measures, such as the measure μ_i^∞ in Proposition 2.1; see also Heifetz (1993) as well as Heifetz and Samet (1998, 1999). Mertens and Zamir (1985) actually refer to spaces of measures and spaces of belief hierarchies as equal "up to BL morphisms".

and, for $k > 1$,

$$\varphi_i^k(t_i) = b_i(\cdot|t_i) \circ (\pi_\Theta \times (\varphi_j^{k-1} \circ \pi_{T_j}))^{-1}, \tag{2.16}$$

where π_Θ is the projection from $\Theta \times T_j$ to Θ and π_{T_j} is the projection from $\Theta \times T_j$ to T_j. To see that, for every t_i, the sequence $\{\varphi_i^k(t_i)\}_{k=1}^\infty$ belongs to H_i, it suffices to observe that, for any $i, j \neq i$, and k, the function

$$(\theta, t_j) \longmapsto \pi_\Theta \times (\varphi_j^{k-1} \circ \pi_{T_j}) = (\theta, \varphi_j^{k-1}(t_j)) \tag{2.17}$$

takes values in the space X^{k-1} that is defined by (2.1) and (2.2). Therefore the function φ_i^k takes values in $\mathcal{M}(X^{k-1})$.

Proposition 2.3 *Given the compact metric spaces T_i and the continuous belief functions $b_i : T_i \to \mathcal{M}(\Theta \times T_j)$, $i = 1, 2, j \neq i$, for $i = 1, 2$, the function $t_i \longmapsto \{\varphi_i^k(t_i)\}_{k=1}^\infty$ that is defined by (2.15) and (2.16) maps T_i continuously into H_i^p. If this function is injective, it is an embedding, i.e., a homeomorphism between T_i and a subspace of H_i^p.*

Proof. The first statement of the proposition is equivalent to the statement that, for any k and any bounded continuus real-valued function h^k on X^{k-1}, the mapping

$$t_i \longmapsto \int_{X^{k-1}} h^k(x^{k-1}) d\varphi_i^k(x^{k-1}|t_i) \tag{2.18}$$

is continuous. By (2.15), for $k = 1$, this is equivalent to the statement that the mapping

$$t_i \longmapsto \int_{X^0} h^1(x^0) d\varphi_i^1(x^0|t_i) = \int_{\Theta \times T_j} h^1(\theta) db_i(\theta, t_j|t_i)$$

is continuous. The latter statement is true because h^1 is continuous and bounded and the belief function $b_i : T_i \to \mathcal{M}(\Theta \times T_j)$ is continuous.

Proceeding by induction on k, consider $\hat{k} > 1$ and suppose that the continuity of the mapping (2.18) has been established for any $k \leq \hat{k} - 1$, for $i = 1, 2$, and any bounded continuus real-valued function h^k on X^{k-1}. By (2.16), the continuity of the mapping (2.18) for $k = \hat{k}$, for $i = 1, 2$, and any bounded continuus real-valued function h^k on X^{k-1} is equivalent to the statement that the mapping

$$t_i \longmapsto \int_{\Theta \times T_j} h^k(\theta, \varphi_j^{k-1}(t_j)) db_i(\theta, t_j|t_i)$$

is continuous. This statement is true because h^k is continuous and bounded, by the induction hypothesis the mapping $t_j \longmapsto \varphi_j^{k-1}(t_j)$ is continuous, and the belief function $b_i : T_i \rightarrow \mathcal{M}(\Theta \times T_j)$ is also continuous. The first statement of the proposition is thereby proved. The second statement follows by standard arguments. ■

3 The Choice of Topology

Propositions 2.1 - 2.3 rely on the topological structure implied by the topology of weak convergence on the spaces of beliefs of different orders and the product topology on the spaces of belief hierarachies. This reliance on the topological structure has been criticized from two sides. On one side, Heifetz and Samet (1998, 1999) have criticized the reliance on topology in the proof of Proposition 2.1 and have proposed a purely measure theoretic approach for constructing a universal type space associated with a space Θ of exogenous parameters. On the other side, Dekel et al. (2006) and Chen et al. (2010, 2017) have argued that the chosen topology is too coarse to capture all the continuity properties of strategic behaviour that we should be interested in. This paper is concerned with implications of the latter critique but in view of several reactions that I have received I will make a remark on the former critique as well.

Whereas Heifetz and Samet (1998, 1999) were concerned with the need for topological assumptions in using Kolmogorov's theorem to prove the existence of a universal type space, later adherents of the purely measure theoretic approach have also dropped the topological questions of Propositions 2.2 and 2.3 from their agenda. Thus, a referee of this paper wrote: "It is not clear why (the) weak(*) topology should be used, definitely it is not true that the weak(*) topology is the weakest topology such that its Borel σ-field includes the epistomologically essential events like a player believes a certain event happens at least (with) a certain probability. For this issue of topology see Pintér (2010)." Pintér (2010, p. 225) claims that "previous papers on type spaces which use topology do so because they want to apply the Kolmogorov extension theorem, not because this is the natural mathematical apparatus to express information. However, even if a certain topology is taken for only mathematical reasons, since the events in the model are the Borel sets of the given topology, the chosen topology describes and represents the information of the considered player." The main result of Pintér (2010) asserts the nonexistence of a universal topological type space; this result is based on the fact that "there is no weakest topology among the topologies whose Borel σ-fields meet" a specified measure theoretic condition.

For someone who has been trained on the work of Arrow, Debreu, and Nash, this is a strange way to think about the choice of a topology. In the tradition of these authors, topologies are chosen with a view to their implications for the continuity properties of certain functions and correspondences that one is interested in, *not* for the information content of the Borel σ-algebras they generate.

Pintér's claim that "previous papers on type spaces which use topology do so because they want to apply the Kolmogorov extension theorem" is counterfactual. In the literature on choice under uncertainty, the weak* topology, or topology of weak convergence of probability measures, is usually chosen because it is the coarsest topology that permits the application of Berge's (1959, p. 121) maximum theorem. Specifically, if A is a space of actions and $u : \Theta \times A \rightarrow \mathbb{R}$ is a bounded continuous function, the topology of weak convergence of probability measures is the coarsest topology on the space $\mathcal{M}(\Theta)$ under which the function $\mu \longmapsto \max_{a\in A} \int_\Theta u(\theta, a)d\mu(a)$ from $\mathcal{M}(\Theta)$ to $\mathbb{R}$ is continuous and the correspondence $\mu \longmapsto \arg\max_{a\in A} \int_\Theta u(\theta, a)d\mu(a)$ from $\mathcal{M}(\Theta)$ to A is upper hemicontinuous and closed-valued.[16] For a game theorist interested in best-response correspondences, this property is important regardless of whether the specified topology "is the natural mathematical apparatus to express information". The usefulness of the chosen topology for the application of Kolmogorov's theorem comes in as a happy coincidence, rather than the main reason for choosing this topology.

Dekel et al (2006) raise the question of what is the appropriate topology for the space of belief hierarchies without reference to Kolmogorov's theorem. They are interested in the continuity properties of the correspondences relating the interim correlated ε-rationalizable choices of an agent to the agent's belief hierarchies. In line with the argument of the preceding paragraph, they show that the topology specified in the preceding section, involving the topology of weak convergence for beliefs of any one order and the associated product topology for the hierarchies of beliefs of all order, is the coarsest topology with what they call the *upper strategic convergence property*, which implies that, in all games with continuous and bounded von Neumann-Morgenstern utility functions u, the correspondences relating the

[16] For a discussion, see, for example, Jordan (1977). A finer topology may be required if Θ is a space of time paths of exogenous variables and A is a space of adapted time paths of actions, and conditional distributions of remaining time paths do not depend continuously on histories; see Hellwig (1996).

interim correlated ε-rationalizable choices of an agent to the agent's belief hierarchies are upper hemi-continuous.[17]

Even so, Dekel et al. (2006) consider the given topology to be unsatisfactory. In addition to upper hemi-continuity, they are also interested in lower hemi-continuity properties of behaviour correspondences. To be sure, in the absence of appropriate convexity assumptions, best responses, i.e. maximizers, are not generally unique, and best-response correspondences are not lower-hemicontinuous. However, correspondences involving the dependence of ε-best responses on the parameters of the optimization problems are usually lower hemi-continuous if the ε-best-response conditions are treated as strict inequalities; this is true for ε-optimal solutions to simple optimization problems, for ε-Nash equilibria, and for ε-rationalizable actions setting involving complete but possibly imperfect information. It is *not* true for interim correlated ε-rationalizable choices under incomplete information if agents' belief hierarchies are given the product topology associated with beliefs of any order having the topology of weak convergence.

Therefore Dekel et al. (2006) propose finer topologies for the space of belief hierarchies. The *strategic topology* of Dekel et al. (2006) is defined in terms of the very upper and lower strategic convergence properties that they consider desirable; in this topology, two belief hierarchies of an agent are close if and only if the associated sets of interim correlated ε-rationalizable actions in all games are close. The *uniform strategic topology* of Dekel et al. (2006) requires the strategic convergence properties to hold uniformly over all games.

In an alternative formulation, Chen et al. (2010, 2017) introduce the *uniform weak topology,* which is defined so that two belief hierarchies of an agent are close if and only if the beliefs of all orders in these hierarchies are uniformly close. Closeness here is defined in terms of the well-known Prohorov metric, which induces the topology of weak convergence on the space of measures in question.

Unlike the topologies proposed by Dekel et al. (2006), the uniform weak topology is defined with reference to the belief hierarchies as such, without any reference to strategic convergence properties. However, Chen at al. (2010, 2017) show that the uniform weak topology is actually equivalent to the uniform strategic topology of Dekel et al. (2006), i.e. two belief hierarchies are close in the uniform strategic topology if and only if they are close in the uniform weak topology.[18] Because of this equivalence, I use the

[17] See Theorems 1, p. 291, and 2, p. 294, in Dekel et al. (2006).

[18] Chen et al. (2017) also show that uniform weak convergence on all frames is equivalent

simpler term *uniform topology* for both.

For a formal treatment, recall that the space H_j of the belief hierarchies of agent j is a subset of the infinite product $\prod_{k\geq 1} \mathcal{M}(X^{k-1})$ of spaces of measures on $X^0, X^1, X^2, ...$ By construction, each of the spaces $X^0, X^1, X^2, ...$ is a compact metric space, and so is each of the spaces $\mathcal{M}(X^0), \mathcal{M}(X^1), \mathcal{M}(X^2), ...$ For any k, the topology of weak convergence on $\mathcal{M}(X^{k-1})$ is metrized by the Prohorov metric, which defines the distance $\rho_k(\mu^k, \hat{\mu}^k)$ between two measures μ^k and $\hat{\mu}^k$ in $\mathcal{M}(X^{k-1})$ as the infimum of the set

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

where B^δ is the δ-neighbourhood of B in X^{k-1}.

Given the Prohorov metrics $\rho_k(\cdot, \cdot)$ on $\mathcal{M}(X^{k-1}), k = 1, 2, ...$, the *uniform metric* $\rho^u(\cdot, \cdot)$ on the infinite product $\prod_{k\geq 1} \mathcal{M}(X^{k-1})$ is given by the formula

$$\rho^u(\{\mu^k\}_{k=1}^\infty, \{\hat{\mu}^k\}_{k=1}^\infty) = \sup_k \rho_k(\mu^k, \hat{\mu}^k). \tag{3.1}$$

The uniform topology is the topology induced by this metric. This topology stands in contrast to the product topology, which is induced by the metric $\rho^p(\cdot, \cdot)$ satisfying

$$\rho^p(\{\mu^k\}_{k=1}^\infty, \{\hat{\mu}^k\}_{k=1}^\infty) = \sum_{k=1}^\infty \alpha^k \rho_k(\mu^k, \hat{\mu}^k),$$

for all $\{\mu^k\}_{k=1}^\infty, \{\hat{\mu}^k\}_{k=1}^\infty$ and some $\alpha \in (0, 1)$. For any k', no matter how large, if $\rho_{k'}(\mu^{k'}, \hat{\mu}^{k'})$ is significantly different from zero, then $\rho^u(\{\mu^k\}_{k=1}^\infty, \{\hat{\mu}^k\}_{k=1}^\infty)$ is also significantly different from zero, but $\rho^p(\{\mu^k\}_{k=1}^\infty, \{\hat{\mu}^k\}_{k=1}^\infty)$ may be small if k' is large.

I will use a superscript u to indicate that a space of belief hierarchies has the uniform topology. Thus, H_i^u and H_j^u are the spaces of belief hierarchies of agents i and j with the uniform topology. The same spaces with the product topology will be denoted as H_i^p and H_j^p. Thus $\mathcal{B}(H_j^u)$ and $\mathcal{B}(\Theta \times H_j^u) = \mathcal{B}(\Theta) \times \mathcal{B}(H_j^u)$ are the Borel σ-algebras on H_j and $\Theta \times H_j$ that are

to convergence in the strategic topology. They prove their equivalence results under the assumption that the space Θ of exogenous parameters and the agents' action sets are finite. In Section 5.3, p. 1451, of Chen et al. (2017), however, they point out that their results carry over to the case where these space are compact metrizable and the payoff functions are continuous and bounded as well as equicontinuous in θ.

induced when H_j has the uniform topology, and $\mathcal{B}(H_j^p)$ and $\mathcal{B}(\Theta \times H_j^p) = \mathcal{B}(\Theta) \times \mathcal{B}(H_j^p)$ are the Borel σ-algebras on H_j and $\Theta \times H_j$ that are induced when H_j has the product σ-algebra, as in the preceding section.

4 The Choice of Event Algebra

Contrary to the presumption of Pintér (2010) cited above, the choice of topology need not prejudge the specification of the set of events to which an agent assigns probabilities. To be sure, we like to work with Borel σ-algebras, but this is very much a matter of mathematical convenience.

In specifying the set of events about which agents form beliefs, Dekel et al. (2006) actually rely on the σ-algebra $\mathcal{B}(\Theta \times H_j^p)$ that is induced by the product topology. Dekel et al. (2007) rely on the σ-algebra involved in the Θ-based universal type space of Heifetz and Samet (1998). Under their assumptions, the two σ-algebras actually coincide, so there is no inconsistency between the treatment in Dekel et al. (2006) and the treatment in Dekel et al. (2007).[19]

There is however a real difference between these two σ-algebras and the Borel σ-algebra $\mathcal{B}(\Theta \times H_j^u) = \mathcal{B}(\Theta) \times \mathcal{B}(H_j^u)$ for the uniform topology on the spaces of belief hierarchies. Whereas H_j^p is separable, H_j^u is not, and the σ-algebra $\mathcal{B}(\Theta \times H_j^u)$ is strictly finer than the σ-algebra $\mathcal{B}(\Theta \times H_j^p)$ used by Dekel et al. (2006).[20]

Because of this difference, any attempt to rely on $\mathcal{B}(\Theta \times H_j^u)$ rather than $\mathcal{B}(\Theta \times H_j^p)$ is fraught with serious difficulties. Measures in the range of the Kolmogorov mapping are defined on $\mathcal{B}(\Theta \times H_j^p)$ rather than $\mathcal{B}(\Theta \times H_j^u)$. How can agent i assign probabilities to events that belong to the difference $\mathcal{B}(\Theta \times H_j^u) \backslash \mathcal{B}(\Theta \times H_j^p)$?

There would be no problem if the measure $\mu_i^\infty = \beta(\{\mu_i^k\}_{k=1}^\infty)$ that is generated by the Kolmogorov mapping β and the belief hierarchy $\{\mu_i^k\}_{k=1}^\infty$ could simply be extended from $\mathcal{B}(\Theta \times H_j^p)$ to $\mathcal{B}(\Theta \times H_j^u)$ and if moreover the ex-

[19] More generally, the Θ-based universal type space of Heifetz and Samet (1998) is a subset of the space of coherent belief hierarchies. If the σ-algebra on Θ is the Borel σ-algebra, Kolmogorov's theorem ensures that any coherent belief hierarchy is an element of the Θ-based universal type space.

[20] Whereas Chen et al. (2010) write that $\mathcal{B}(\Theta \times H_j^u)$ and $\mathcal{B}(\Theta \times H_j^p)$ are equal, Chen et al. (2016) retract this claim, giving an example of a set in $\mathcal{B}(H_j^u) \backslash \mathcal{B}(H_j^p)$. The set in question is an analytic set that has a non-analytic complement and therefore is not in $\mathcal{B}(H_j^p)$, but that can be represented as a countable intersection of sets in $\mathcal{B}(H_j^u)$ and therefore is in $\mathcal{B}(H_j^u)$.

tension was unique.[21] However, this is not the case. Under the set theoretic axioms that are usually made in probability theory, not every measure on $(\Theta \times H_j^p, \mathcal{B}(\Theta \times H_j^p))$ can be extended to a measure on $(\Theta \times H_j^u, \mathcal{B}(\Theta \times H_j^u))$. Moreover, if such an extension exists, it need not be unique.

Remark 4.1 *Assume that the cardinal* $\mathbf{c}$ *of the continuum is not atomlessly measurable. Then there exists a belief hierarchy* $\{\mu_i^k\}_{k=1}^{\infty}$ *such that the measure* $\mu_i^{\infty} = \beta(\{\mu_i^k\}_{k=1}^{\infty})$ *that is defined on* $(\Theta \times H_j^p, \mathcal{B}(\Theta \times H_j^p))$ *cannot be extended to a countably additive measure on* $(\Theta \times H_j^u, \mathcal{B}(\Theta \times H_j^u))$.

Proof. The assumption made implies that, that, for set $Y \subset \Theta \times H_j$, there is no nontrivial countably additive atomless measure that is defined on the set of all subsets of Y. To prove the remark, it suffices to specify a set $Y \subset \Theta \times H_j$ and a belief hierarchy $\{\mu_i^k\}_{k=1}^{\infty}$ such that the measure $\mu_i^{\infty} = \beta(\{\mu_i^k\}_{k=1}^{\infty})$ is concentrated on Y and is atomless and any extension of μ_i^{∞} from $\mathcal{B}(\Theta \times H_j^p)$ to $\mathcal{B}(\Theta \times H_j^u)$ would be defined on all subsets of Y.

I construct Y as follows. Fix two distinct elements θ_1, θ_2 in Θ. Consider the infinite product $\{1,2\}^{\infty}$. For any $x = \{x_k\}_{k=1}^{\infty} \in \{1,2\}^{\infty}$, define $y(x) = \{y_k(x_k)\}_{k=1}^{\infty} \in \Theta \times H_j$ by setting

$$y_1(x_1) = \theta_{x_1}, y_2(x_2) = \delta_{\theta_{x_2}}, y_3(x_3) = \delta_{\delta_{\theta_{x_3}}}, ...,$$

and let Y be the range of . The projection of Y to Θ is the set $\{\theta_1, \theta_2\}$. The projection of Y to the space of agent j's first-order beliefs is the set $\{\delta_\theta | \theta \in \{\theta_1, \theta_2\}\}$ of degenerate distributions that have all mass concentrated at $\theta^1 \in \{\theta_1, \theta_2\}$; agent i considers agent j to have first-order beliefs concentrated at a singleton subset of $\{\theta_1, \theta_2\}$. The projection of Y to the space of agent j's first- and second-order beliefs is the set $\{\delta_{\theta^1} \times \delta_{\delta_{\theta^2}} | \theta^1 \in \{\theta_1, \theta_2\}, \theta^2 \in \{\theta_1, \theta_2\}\}$ i.e., agent i considers agent j to have second-order beliefs concentrated at $\delta_{\delta_{\theta^2}}$ where $\delta_{\delta_{\theta^2}}$ is the degenerate distribution that has all mass concentrated at δ_{θ^2} with θ^2. Proceeding in this way, one obtains a space of belief hierarchies for agent j such that, for any k, the beliefs of order k are characterized by some $\theta^k \in \{\theta_1, \theta_2\}$.

One easily verifies that the mapping $x \longmapsto y(x)$ is a homeomorphism between the infite product $\{1,2\}^{\infty}$ and the subspace Y of $\Theta \times H_j$ is both spaces have the product topology and if both spaces have the uniform topology. The problem of extending a measure from the σ-algebra $\mathcal{B}(\Theta \times H_j^p))$

[21] In a previous version of this paper, I suggested that this was actually the case and that the sets in $\mathcal{B}(\Theta \times H_j^u)$ belonged to the completion of $\mathcal{B}(\Theta \times H_j^p)$ under every measure. The argument I gave, however, was flawed, and the suggestion is false. I am grateful to a referee for pointing out the error.

to the σ-algebra $\mathcal{B}(\Theta \times H_j^u)$) is thus equivalent to the problem of extending a measure from the σ-algebra that is generated by the product topology on $\{1,2\}^\infty$ to the σ-algebra that is generated by the uniform topology on $\{1,2\}^\infty$.

To see that such an extension need not exist, think about the elements of $\{1,2\}^\infty$ as the outcomes of an infinite sequence of tosses of a fair coin. The associated measure on the product σ-algebra on $\{1,2\}^\infty$ is the product measure that assigns probability $\frac{1}{2^k}$ to the set of sequences of the form $\{x_\ell\}_{\ell=0}^\infty$ with $x_\ell = \hat{x}_\ell$ for $\ell = 0, ...k-1$, for any given k and $(\hat{x}_0, ..., \hat{x}_{k-1}) \in \{1,2\}^k$. This measure assigns probability zero to any singleton $\{x\} \subset \{1,2\}^\infty$. Moreover, because any singleton $\{x\} \subset \{1,2\}^\infty$ is open in the uniform topology on $\{1,2\}^\infty$, the σ-algebra that is generated by the uniform topology on $\{1,2\}^\infty$ contains all subsets of $\{1,2\}^\infty$. The existence of an extension of the measure for infinite sequences of outcomes of fair coin tosses from σ-algebra that is generated by the product topology on $\{1,2\}^\infty$ to the σ-algebra that is generated by the uniform topology on $\{1,2\}^\infty$ is therefore incompatible with the axiom that the cardinal of the continuum is not atomlessly measurable.
■

Remark 4.2 *For every belief hierarchy* $\{\mu_i^k\}$, *the measure* $\mu_i^\infty = \beta(\{\mu_i^k\}_{k=1}^\infty)$ *that is defined on* $(\Theta \times H_j^p, \mathcal{B}(\Theta \times H_j^p))$ *can be extended to a finitely additive measure on* $(\Theta \times H_j^u, \mathcal{B}(\Theta \times H_j^u))$, *but for some belief hierarchies, the extension is not unique.*

Proof. Existence of a finitely additive extension follows from Theorem 2.1, p. 390, in Bachman and Sultan (1980). By Theorem 3.1, p. 544, of Bachman and Sultan (1977), the extension is unique for all measures on $(\Theta \times H_j^p, \mathcal{B}(\Theta \times H_j^p))$ *if and only if* $\mathcal{B}(\Theta \times H_j^p)$ separates $\mathcal{B}(\Theta \times H_j^u)$ in the sense that, for any pair A, B of nonintersecting sets in $\mathcal{B}(\Theta \times H_j^u)$, there exists a pair C, D of nonintersecting sets in $\mathcal{B}(\Theta \times H_j^p)$ such that $A \subset C$ and $B \subset D$.

To see that this condition is violated, consider the set Y in the proof of the preceding lemma and the measures $\mu_i^\infty = \beta(\{\mu_i^k\}_{k=1}^\infty)$ that are concentrated on Y. Recall that the mapping $x \longmapsto y(x)$ $\{1,2\}^\infty$ to Y that is a homeomorphism if both $\{1,2\}^\infty$ and Y have the product topology and is also a homeomorphism if both $\{1,2\}^\infty$ and Y have the uniform topology. Clearly the pair A, B of nonintersecting sets in $\mathcal{B}(\Theta \times H_j^u)$ satisfies the separation condition of Bachman and Sultan if and only if there exists a pair C^*, D^* of nonintersecting sets in the product σ-algebra on $\{1,2\}^\infty$ such that the images A^*, B^* of A and B under the inverse of the mapping $x \longmapsto y(x)$ satisfy $A^* \subset C^*$ and $B^* \subset D^*$. Let A, B be such that A^* is the set of all sequences $\{x_\ell\}_{\ell=0}^\infty$ in $\{1,2\}^\infty$ with $x_\ell = 1$ no more than finitely many times and B^* is the set of all sequences $\{x_\ell\}_{\ell=0}^\infty$ in $\{1,2\}^\infty$ with $x_\ell = 2$ no more than finitely many times. Then $A^* \cap B^* = \emptyset$, but $A^* \subset C^*$ and $B^* \subset D^*$ imply $C^* = D^* = \{1,2\}^\infty$, which is incompatible with the separation condition $C^* \cap D^* = \emptyset$. Indeed, it is easy to show that the assigment of probabilities to the specified sets A and B is to some extent arbitrary. ■

Given these difficulties, I see no scope for a general treatment of agent's beliefs about exogenous parameters and other agents' belief hierarchies as measures on $(\Theta \times H_j^u, \mathcal{B}(\Theta \times H_j^u))$ and determined by the agent's own belief hierarchy. As an alternative, I therefore consider measures on the σ-algebras $\mathcal{B}_0(H_j^u)$ and $\mathcal{B}_0(\Theta \times H_j^u) = \mathcal{B}(\Theta) \times \mathcal{B}_0(H_j^u)$ that are generated by the open balls in the uniform topology (rather than the open sets). These σ-algebras were introduced by Dudley (1966, 1967).

For any belief hierarchy $\{\mu^k\}_{k=1}^\infty \in H_j^u$ and any $r > 0$, the open r-ball $B^u(\{\mu^k\}_{k=1}^\infty, r)$ around $\{\mu^k\}_{k=1}^\infty$ in the uniform topology is given as

$$B^u(\{\mu^k\}_{k=1}^\infty, r) = \bigcup_{n=1}^{\infty} \prod_{k=0}^{\infty} B_k(\mu^k, r - \frac{1}{n}), \tag{4.1}$$

where, for each k and $r' > 0$, $B_k(\mu^k, r')$ is the open r'-ball around μ^k, as given by the metric ρ_k on $\mathcal{M}(X^{k-1})$.

Because the ρ^u-open balls form a basis for the uniform topology, the open sets in the uniform topology can be represented as arbitrary unions of ρ^u-open balls. In contrast, $\mathcal{B}_0(H_j^u)$ only allows for *countable* unions of open balls. Therefore, $\mathcal{B}_0(H_j^u)$ and $\mathcal{B}_0(\Theta \times H_j^u)$ are much coarser than $\mathcal{B}(H_j^u)$ and $\mathcal{B}(\Theta \times H_j^u)$. The following lemma shows that, in fact, $\mathcal{B}_0(H_j^u)$ and $\mathcal{B}_0(\Theta \times H_j^u)$ are coextensive with $\mathcal{B}(H_j^p)$ and $\mathcal{B}(\Theta \times H_j^p)$.

Lemma 4.3 *The σ-algebra $\mathcal{B}_0(H_j^u)$ that is generated by the ρ^u-open balls on H_j^u is coextensive with the σ-algebra $\mathcal{B}(H_j^p)$ that is generated by the product topology, i.e., $\mathcal{B}_0(H_j^u) = \mathcal{B}(H_j^p)$. Similarly, $\mathcal{B}_0(\Theta \times H_j^u) = \mathcal{B}(\Theta \times H_j^p)$.*

Proof. I first prove that $\mathcal{B}_0(H_j^u) \subset \mathcal{B}(H_j^p)$. Recall that $\mathcal{B}(H_j^p)$ is the smallest σ-algebra on H_j that contains the cylinder sets, the sets of the form

$$C = B_1 \times ... \times B_\ell \times \mathcal{M}(X^\ell) \times \mathcal{M}(X^{\ell+1}) \times ..., \tag{4.2}$$

for some ℓ and open $B_1, ..., B_\ell$, and that is closed under countable unions and intersections. Since (4.1) can be rewritten in the form

$$B^u(\{\mu^k\}_{k=1}^{\infty}, r) = \bigcap_{\ell=1}^{\infty} \bigcup_{n=1}^{\infty} \left[\prod_{k=1}^{\ell} B_k(\mu^k, r - \frac{1}{n}) \times \prod_{k=\ell+1}^{\infty} \mathcal{M}(X^{k-1}) \right],$$

any ρ^u-open r-ball is a countable intersection of countable unions of cylinder sets and hence an element of $\mathcal{B}(H_j^p)$. Since $\mathcal{B}_0(H_j^u)$ is the smallest σ-algebra on H_j that contains the ρ^u-open balls and that is closed under countable unions and intersections, it follows that any set in $\mathcal{B}_0(H_j^u)$ is also in $\mathcal{B}(H_j^p)$.

To prove that, conversely, $\mathcal{B}(H_j^p) \subset \mathcal{B}_0(H_j^u)$, it suffices to show that any set of the form (4.2), with $B_1, ..., B_\ell$ open, belongs to $\mathcal{B}_0(H_j^u)$. For this purpose, let C be given as in (4.2). For any $N > \ell$, consider the projection

$$\Pi^N(C) = B_1 \times ... \times B_\ell \times \mathcal{M}(X^\ell) \times ... \times \mathcal{M}(X^N) \subset \mathcal{M}(X^0) \times ... \times \mathcal{M}(X^N). \tag{4.3}$$

Observe that, when endowed with the uniform metric

$$\rho^N(\{\mu^k\}_{k=1}^N, \{\hat{\mu}^k\}_{k=1}^M) := \max_{k=1}^N \rho_k(\mu^k, \hat{\mu}^k), \tag{4.4}$$

the finite product $\mathcal{M}(X^0) \times ... \times \mathcal{M}(X^N)$ is separable and has a countable basis $\{z^1, z^2, ...\}$. For $m = 1, 2, ...$ and $r > 0$, let $U^N(k, r)$ be the ρ^N-open r-ball around z^j. Let $M^*(C)$ be the set of pairs (m, r_m) such that $U^N(m, r_m) \subset C$ for some $r_m > 0$. Then, by standard arguments,

$$\Pi^N(C) = \bigcup_{M^*(C)} U^N(m, r_m).$$

Next, fix some $\{\mu^k\}_{k=N+1}^{\infty}$ and, for any $m \in M^*(C)$, define

$$\hat{U}^N(m, r_m) = U^N(m, r_m) \times B_{N+1}(\mu^{N+1}, r_m) \times B_{N+2}(\mu^{N+2}, r_m) \times ...$$

and

$$\bar{U}^N(m, r_m) = \bigcup_{n=1}^{\infty} [\hat{U}^N(m, r_m - \frac{1}{n}) \times B_{N+1}(\mu^{N+1}, r_m - \frac{1}{n}) \times B_{N+2}(\mu^{N+2}, r_m - \frac{1}{n}) \times ...].$$

Then for any $N > \ell$ and any m, $\bar{U}^N(m, r_m)$ is a ρ^u-open ball in H_j and belongs to $\mathcal{B}_0(H_j^u)$. Moreover,

$$C = \bigcup_{N=\ell+1}^{\infty} \bigcup_{m=1}^{\infty} \bar{U}^N(m, r_m).$$

Since $\mathcal{B}_0(H_j^u)$ is closed under countable unions and intersections, it follows that $C \in \mathcal{B}_0(H_j^u)$, and, more generally, $\mathcal{B}(H_j^p) \subset \mathcal{B}_0(H_j^u)$.

The second statement of the proposition follows immediately. ■

Corollary 4.4 *For any belief hierarchy* $\{\mu_i^k\}_{k=1}^\infty \in H_i$, *the unique probability measure* $\mu_i^\infty = \beta(\{\mu_i^k\}_{k=1}^\infty)$ *on* $(\Theta \times H_j^p, \mathcal{B}(\Theta \times H_j^p))$ *that is given by Proposition 2.1 is also a probability measure on* $(\Theta \times H_j^u, \mathcal{B}_0(\Theta \times H_j^u))$.

By working with $\mathcal{B}_0(\Theta \times H_j^u)$ rather than $\mathcal{B}(\Theta \times H_j^u)$, one avoids the dilemma of whether and how the measure $\beta(\{\mu_i^k\}_{k=1}^\infty)$ can be extended from $\mathcal{B}(\Theta \times H_j^p)$ to $\mathcal{B}(\Theta \times H_j^u)$. The Kolmogorov mapping suffices as a belief function, as in the case of the σ-algebra $\mathcal{B}(\Theta \times H_j^p)$ that is induced by the product topology. Moreover, the σ-algebra $\mathcal{B}(\Theta \times H_j^p)$ itself has an interpretation in terms of the uniform topology, as $\mathcal{B}_0(\Theta \times H_j^u)$. This latter observation may perhaps resolve the seeming paradox involved when Dekel et al. (2006) themselves work with measures on $(\Theta \times H_j^p, \mathcal{B}(\Theta \times H_j^p))$ even as they criticize the reliance on the product topology.

The analysis of Dekel et al. (2006, pp. 283 f.) is unaffected by the restriction to the smaller σ-algebra. Their assessment of rationalizability focuses on the question whether a given action a_i of agent i can be ε-rationalized by some joint distribution ν over states of nature and "types" (belief hierarchies) and actions of agent $j \neq i$; in this assessment, any action a_j that is presumed to be taken by agent j must in turn be ε-rationalizable with probability one for agent j with the type t_j that is presumed to take this action. The realized payoff of any agent depends only on the game that is being played, the state of nature, and the vector of actions of the different agents. The *expected* payoff depends only on the agent's assessment of the joint distribution of the state of nature and the other agent's action.

A joint distribution ν over states of nature and "types" (belief hierarchies) and actions of agent $j \neq i$ is determined by the agent's type-dependent belief over the state of nature and the other player's "type" and by the agent's conjecture σ_j over the other player's strategy. This conjecture over the other player's strategy is a mapping from states of nature and types of the other player to probability distributions over actions of the other player.

In this formalism, the σ-algebra on H_j that determines the measurability properties of agent j's strategy function, as well as the σ-algebra of event in $\Theta \times H_j$ about which agent i forms beliefs, matters for the specification of the conjecture σ_j and of the domain of the agent's type-dependent belief over the state of nature and the other player's "type". However, the impact of this specification on the agent's expected payoff is integrated out and has no effect on the set of rationalizable actions. Specifically, if a joint distribution ν over states of nature and "types" (belief hierarchies) and actions of agent $j \neq i$ and conjecture σ_j are specified so as to allow for events in $\mathcal{B}(\Theta \times H_j^u)$ that do not belong to $\mathcal{B}_0(\Theta \times H_j^u) = \mathcal{B}(\Theta \times H_j^p)$, then, by taking conditional expectations of $\sigma_j(\cdot, \cdot)$ conditional on $\mathcal{B}_0(\Theta \times H_j^u)$, one obtains a new conjecture $\sigma_j^*(\cdot, \cdot)$ that is measurable with respect to $\mathcal{B}_0(\Theta \times H_j^u)$ and that yields the same joint distribution over states of nature and the other agent's actions as the original conjecture $\sigma_j(\cdot, \cdot)$. The extra richness that comes from the (larger) Borel σ-algebra $\mathcal{B}(\Theta \times H_j^u)$ is irrelevant for the analysis of rationalizability.

5 The Homeomorphism Theorem

The aim of this section is to establish an analogue of Proposition 2.2 when the domain of the Kolmogorov mapping is H_i^u and the range is the space of probability measures on $(\Theta \times H_j^u, \mathcal{B}_0(\Theta \times H_j^u))$. I denote the space of probability measures on $(\Theta \times H_j^u, \mathcal{B}_0(\Theta \times H_j^u))$ as $\mathcal{M}_0(\Theta \times H_j^u)$.

The space $\mathcal{M}_0(\Theta \times H_j^u)$ is given the topology of weak convergence of probability measures on $(\Theta \times H_j^u, \mathcal{B}_0(\Theta \times H_j^u))$: Let $\mathcal{C}_0(\Theta \times H_j^u)$ be the space of bounded, ρ^u-continuous and $\mathcal{B}(\Theta \times H_j^p)$-measurable real-valued functions on $\Theta \times H_j^u$. A sequence $\{\mu^\ell\}$ of measures in $\mathcal{M}_0(\Theta \times H_j^u)$ converges weakly to a limit μ in $\mathcal{M}_0(\Theta \times H_j^u)$ if and only if, for every $f \in \mathcal{C}_0(\Theta \times H_j^u)$, the integrals $\int_{\Theta \times H_j^u} f(\theta, \{\mu_j^k\}_{k=1}^\infty) d\mu^\ell$ converge to the limit $\int_{\Theta \times H_j^u} f(\theta, \{\mu_j^k\}_{\ell=1}^\infty) d\mu$.

The topology of weak convergence on $\mathcal{M}_0(\Theta \times H_j^u)$ should not be confused with the topology of weak convergence on $\mathcal{M}(\Theta \times H_j^p)$ considered in Section 2. To be sure, the term "topology of weak convergence" is the same, and, by Lemma 4.3, the σ-algebras $\mathcal{B}_0(\Theta \times H_j^u)$ and $\mathcal{B}(\Theta \times H_j^p)$ are coextensive, and so are the spaces $\mathcal{M}_0(\Theta \times H_j^u)$ and $\mathcal{M}(\Theta \times H_j^p)$ of measures on these σ-algebras. However, the functions f whose expectations are considered for assessing weak convergence are not the same. In the definition of the topology of weak convergence on $\mathcal{M}(\Theta \times H_j^p)$ in Section 2, these functions were required to be continuous with respect to the product topology on $\Theta \times H_j^p$; now they are required to be continuous with respect to the product topology on $\Theta \times H_j^u$. Because the uniform topology on H_j is strictly finer than the product topology, the set $\mathcal{C}_0(\Theta \times H_j^u)$ is strictly larger than the set of bounded continuous functions on $\Theta \times H_j^p$. The criterion for convergence of a sequence of measures is therefore stricter, and the induced topology on the range of the Kolmogorov mapping is strictly finer.

The proof of the homeomorphism theorem relies on the fact that, if the cardinal **c** of the continuum is not atomlessly measurable, then the topology of weak convergence on $\mathcal{M}_0(\Theta \times H_j^u)$ is metrizable by a ρ^u-based version of the Prohorov metric. The ρ^u-based Prohorov distance between any two measures μ and $\hat{\mu}$ in $\mathcal{M}_0(\Theta \times H_j^u)$ is defined as the infimum of the set

$$\{\varepsilon > 0 | \mu(B) \leq \hat{\mu}(B^\varepsilon) + \varepsilon \text{ and } \hat{\mu}(B) \leq \mu(B^\varepsilon) + \varepsilon \text{ for all } B \in \mathcal{B}_0(\Theta \times H_j^u)\}, \tag{5.1}$$

with B^ε denoting the ε-neighbourhood of B in $\Theta \times H_j^u$. In Hellwig (2017/2022) I show that, for any $\varepsilon > 0$, the ε-neighbourhood B^ε of any set $B \in \mathcal{B}_0(\Theta \times H_j^u)$ also belongs to $\mathcal{B}_0(\Theta \times H_j^u)$. For any $B \in \mathcal{B}_0(\Theta \times H_j^u)$, therefore, the terms $\hat{\mu}(B^\varepsilon)$ and $\mu(B^\varepsilon)$ in (5.1) are well defined.

Proposition 5.1 *If the cardinal* **c** *of the continuum is not atomlessly measurable, the topology of weak convergence of probability measures in* $\mathcal{M}_0(\Theta \times H_j^u)$ *is metrizable by the* ρ^u*-based Prohorov metric.*

For a proof of this proposition, the reader is referred to Hellwig (2017/2022). The argument given there is similar to the argument that Billingsley (1968, Appendix III) gave for Borel measures, but, because $\mathcal{M}_0(\Theta \times H_j^u)$ is a space of non-Borel measures, some steps must be developed from first principles.

The main result of this paper is now stated as follows.

Proposition 5.2 *Assume that the cardinal* **c** *of the continuum is not atomlessly measurable and suppose that the spaces* H_i *and* H_j *of belief hierarchies of agents* i *and* j *are endowed with the uniform topology. If the space* $\mathcal{M}_0(\Theta \times H_j^u)$ *of probability measures on* $(\Theta \times H_j^u, \mathcal{B}_0(\Theta \times H_j^u))$ *is endowed with the topology of weak convergence, then the Kolmogorov mapping defines a homeomorphism between* H_i^u *and* $\mathcal{M}_0(\Theta \times H_j^u)$.

Proof. The Kolmogorov mapping β is obviously injective and onto. Therefore it suffices to show that, if $\mathcal{M}_0(\Theta \times H_j^u)$ has the topology of weak convergence, then both β and β^{-1} are continuous. By Proposition 5.1, it suffices to show that these continuity properties hold when $\mathcal{M}_0(\Theta \times H_j^u)$ has the topology induced by the Prohorov metric.

I first show that β is continuous. Proceeding indirectly, suppose that β is not continuous. Then there exist a sequence $(\{\mu^{kr}\}_{k\geq 1})_{r=1}^{\infty}$ of belief hierarchies and a further belief hierarchy $\{\mu^{k\infty}\}_{k\geq 1}$, all in H_i, such that

$$\lim_{r\to\infty} \rho^u(\{\mu^{kr}\}_{k\geq 1}, \{\mu^{k\infty}\}_{k\geq 1}) = 0 \tag{5.2}$$

but $\beta(\{\mu^{kr}\}_{k\geq 1})$ does not converge to $\beta(\{\mu^{k\infty}\}_{k\geq 1})$. To simplify the notation, write $h^r := \{\mu^{kr}\}_{k\geq 1}$, $r = 1, 2, ...$, and $h^\infty := \{\mu^{k\infty}\}_{k\geq 1}$. Taking subsequences if necessary, suppose that, for some $\varepsilon > 0$, the Prohorov distance between $\beta(h^r)$ and $\beta(h^\infty)$ exceeds ε for all r. For each r then, there exists a set $W^r \in \mathcal{B}_0(\Theta \times H_j^u)$ such that either

$$\beta(W^r|h^r) > \beta((W^r)^\varepsilon|h^\infty) + \varepsilon \tag{5.3}$$

or

$$\beta(W^r|h^\infty) > \beta((W^r)^\varepsilon|h^r) + \varepsilon, \tag{5.4}$$

where Lemma **??** in the appendix ensures that the sets $(W^r)^\varepsilon$ also belong to $\mathcal{B}_0(\Theta \times H_j^u)$.

For any n, let π^n be the projection of H_j to the space $\mathcal{M}(X^0) \times ... \times \mathcal{M}(X^n)$. For any n and r, and let W^{rn} be the projection of W^r to $\Theta \times \pi^n(H_j)$. Further, let $(W^{rn})^\varepsilon$ be an ε-neighbourhood of W^{rn} (in $\Theta \times \pi^n(H_j)$). Let

$$\hat{W}^{rn} = W^{rn} \times \mathcal{M}(X^{n+1}) \times \mathcal{M}(X^{n+2}) \times ... \tag{5.5}$$

and

$$\hat{W}^{rn\varepsilon} = (W^{rn})^\varepsilon \times \mathcal{M}(X^{n+1}) \times \mathcal{M}(X^{n+2}) \times ... \tag{5.6}$$

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

$$W^r = \bigcap_{n=1}^{\infty} \hat{W}^{rn} \quad \text{and} \quad (W^r)^\varepsilon = \bigcap_{n=1}^{\infty} \hat{W}^{rn\varepsilon} \tag{5.7}$$

for all r. By elementary measure theory, e.g., Theorem 3.1.1, p. 86, in Dudley (2002), it follows that, for any r and any $\delta > 0$, there exists $N^r(\delta)$ such that, for any $n > N^r(\delta)$,

$$\beta((W^r)^\varepsilon|h^\infty) \geq \beta(\hat{W}^{rn\varepsilon}|h^\infty) - \delta \tag{5.8}$$

and

$$\beta((W^r)^\varepsilon|h^r) \geq \beta(\hat{W}^{rn\varepsilon}|h^r) - \delta \tag{5.9}$$

Moreover,

$$\beta(W^r|h^r) \leq \beta(\hat{W}^{rn}|h^r) \tag{5.10}$$

and

$$\beta(W^r|h^\infty) \leq \beta(\hat{W}^{rn}|h^\infty). \tag{5.11}$$

Upon combining (5.8) - (5.11) with (5.3) and (5.4), one finds that, for all r and any $\delta > 0$, there exists $N^r(\delta)$ such that, for any $n > N^r(\delta)$, either

$$\beta(\hat{W}^{rn}|h^r) \geq \beta(\hat{W}^{rn\varepsilon}|h^\infty) - \delta + \varepsilon \tag{5.12}$$

or

$$\beta(\hat{W}^{rn}|h^\infty) \geq \beta(\hat{W}^{rn\varepsilon}|h^r) - \delta + \varepsilon. \tag{5.13}$$

By the definitions of $\beta(\cdot)$, as well as $h^r := \{\mu^{kr}\}_{k\geq 1}$ and $h^\infty := \{\mu^{k\infty}\}_{k\geq 1}$, one also has

$$\beta(\hat{W}^{rn}|h^r) = \mu^{nr}(W^{rn}), \beta(\hat{W}^{rn\varepsilon}|h^r) = \mu^{nr}((W^{rn})^\varepsilon) \tag{5.14}$$

and

$$\beta(\hat{W}^{rn}|h^\infty) = \mu^{n\infty}(W^{rn}), \beta(\hat{W}^{rn\varepsilon}|h^\infty) = \mu^{n\infty}((W^{rn})^\varepsilon) \tag{5.15}$$

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

$$\mu^{n\infty}(W^{rn}) > \mu^{nr}((W^{rn})^\varepsilon) - \delta + \varepsilon \tag{5.16}$$

or

$$\mu^{nr}(W^{rn}) > \mu^{n\infty}((W^{rn})^\varepsilon) - \delta + \varepsilon. \tag{5.17}$$

If $\delta < \varepsilon$, one also has $(W^{rn})^{\varepsilon-\delta} \subset (W^{rn})^\varepsilon$ for all n and r and $(W^r)^{\varepsilon-\delta} \subset (W^r)^\varepsilon$ for all r. For all r and $\delta \in (0, \varepsilon)$, therefore, there exists $N^r(\delta)$ such that, for any $n > N^r(\delta)$, either

$$\mu^{n\infty}(W^{rn}) > \mu^{nr}((W^{rn})^{\varepsilon-\delta}) + \varepsilon - \delta \tag{5.18}$$

or

$$\mu^{nr}(W^{rn}) > \mu^{n\infty}((W^{rn})^{\varepsilon-\delta}) + \varepsilon - \delta. \tag{5.19}$$

But then, for any r, for $n > N^r(\delta)$, the Prohorov distance between the measures μ^{nr} and $\mu^{n\infty}$ is at least $\varepsilon - \delta > 0$.

It follows that, for all r,

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

contrary to (5.2). The assumption that the mapping β from H_i^u to $\mathcal{M}_0(\Theta \times H_j^u)$ is not continuous thus leads to a contradiction and must be false.

Continuity of the mapping β^{-1} from $\mathcal{M}_0(\Theta \times H_j^u)$ to H_i^u is easily obtained by observing that, by Lemma 4.3, for any n and any set $W^n \in \mathcal{B}(\Theta \times \pi^n(H_j))$, the associated cylinder set

$$\hat{W}^n = W^n \times \mathcal{M}(X^n) \times \mathcal{M}(X^{n+1}) \times ... \tag{5.20}$$

belongs to $\mathcal{B}_0(\Theta \times H_j^u)$, and, for any $\varepsilon > 0$, the cylinder set

$$\hat{W}^{n\varepsilon} = (W^n)^\varepsilon \times \mathcal{M}(X^n) \times \mathcal{M}(X^{n+1}) \times ... \tag{5.21}$$

that is defined by the ε-neighbourhood $(W^n)^\varepsilon$ of W^n in $\Theta \times \pi^n(H_j)$ is actually an ε-neighourhood of $\hat{W}^n$ in $\Theta \times H_j^u$. Hence, if the Prohorov distance between two measures μ^∞ and $\hat{\mu}^\infty$ in $\mathcal{M}_0(\Theta \times H_j^u)$ is less than ε, it must be the case that

$$\mu^\infty(\hat{W}^n) < \hat{\mu}^\infty(\hat{W}^{n\varepsilon}) + \varepsilon \tag{5.22}$$

and

$$\hat{\mu}^\infty(\hat{W}^n) < \mu^\infty(\hat{W}^{n\varepsilon}) + \varepsilon. \tag{5.23}$$

By the definition of the marginal distributions, it follows that

$$\mu^n(W^n) < \hat{\mu}^n((W^n)^\varepsilon) + \varepsilon \tag{5.24}$$

and

$$\hat{\mu}^n(W^n) < \mu^n((W^n)^\varepsilon) + \varepsilon. \tag{5.25}$$

Since the choice of $W^n \in \mathcal{B}(\Theta \times \pi^n(H_j))$ was arbitrary, it follows that the Prohorov distance between μ^n and $\hat{\mu}^n$ in $\mathcal{M}(\Theta \times \pi^n(H_j))$ is no greater than the Prohorov distance between μ^∞ and $\hat{\mu}^\infty$ in $\mathcal{M}_0(\Theta \times H_j^u)$. Since this statement holds for all n, it follows that the supremum over n of the Prohorov distances between the marginal distributions μ^n and $\hat{\mu}^n$ in $\mathcal{M}(\Theta \times \pi^n(H_j))$ is no greater than the Prohorov distance between μ^∞ and $\hat{\mu}^\infty$ in $\mathcal{M}_0(\Theta \times H_j^u)$. Continuity of the mapping β^{-1} from $\mathcal{M}_0(\Theta \times H_j^u)$ to H_i^u follows immediately. ■

Proposition 5.2 provides an analogue to Proposition 2.2 above. The homeomorphism property of the Kolmogorov mapping is unaffected if the product topology on H_i and H_j is replaced by the uniform topology. The domain H_i^u and the range $\mathcal{M}_0(\Theta \times H_j^u)$ of the Kolmogorov mapping are topologically equivalent so, again, for any statement about the continuity properties of functions and correpondences on one space, there is an equivalent statement about continuity properties of functions and correspondences on the other space.

I conclude this section with a comment on the assumption that the cardinal **c** of the continuum is not atomlessly measurable. By a theorem of Banach and Kuratowski (1929), this assumption is implied by the Continuum Hypothesis (CH).[22] The older literature, such as Dudley (1967) or Billingsley (1968), invokes this theorem to suggest that the assumption is unproblematic.[23] In recent decades though, under the influence of Cohen (1966), CH has increasingly met with criticism.

However, CH is *not* necessary for the condition that the cardinal **c** of the continuum is not atomlessly measurable. As shown by Ulam (1930), it suffices that any cardinal below **c** be weakly accessible. Bartoszynski and Halbeisen (2003) emphasize that, in Banach and Kuratowskiy (1929), the conclusion that **c** is not atomlessly measurable is obtained from an intermediate condition concerning the existence of a what they call a *BK-matrix*. Whereas the existence of a BK-matrix is implied by CH, Bartoszynski and Halbeisen (2003) show that the existence of a BK-matrix is also compatible with the negation of CH.[24]

If the cardinal **c** of the continuum is atomlessly measurable, there is no guarantee that the topology of weak convergence on $\mathcal{M}_0(\Theta \times H_j^u)$ is metrizable (or on $\mathcal{M}(\Theta \times H_j^u)$ for that matter).[25] In this case, the proof of Proposition 5.2 still shows that the Kolmogorov mapping is a homeomorphism between H_i^u and $\mathcal{M}_0(\Theta \times H_j^u)$ with the topology induced by the Prohorov metric. However, the topology of weak convergence on $\mathcal{M}_0(\Theta \times H_j^u)$ may be

[22] For an accessible presentation of the argument, see Appendix C in Dudley (2002).

[23] Thus, Dudley (1967), citing Marczewski and Sikorski (1948), writes: "One can safely assume that a finite countably additive measure on $\mathcal{B}$ is concentrated on a separable subset." Marczewski and Sikorski (1948) in turn assume CH and cite Banach and Kuratowski (1929). In a change of tone, Dudley (2002, pp. 518 f.) explains that CH is controversial, but continues "Nevertheless, it remains a useful, consistent, and rather popular assumption."

[24] If the underlying space is the unit interval, a BK-matrix is an infinite matrix of sets $A_k^i, i \in \mathbb{N}, k \in \mathbb{N}$, such that, (i) for each $i \in \mathbb{N}, \cup_{k\in\mathbb{N}} A_k^i = [0,1]$, (ii) for each $i \in \mathbb{N}$, $A_k^i \cap A_{k'}^i = \emptyset$ if $k \neq k'$, and (iii) for any sequence $\{k_i\}$ in $\mathbb{N}$, the set $\cap_{i\in\mathbb{N}}(\cup_{k\leq k_i} A_k^i)$ is at most countable. Bartoszynski and Halbeisen (2003) show that such a matrix exists if and only if there exists a K-Lusin set of the size of the continuum.

[25] In this case also, Remark 4.1 is inapplicable, so the measure defined by the Kolmogorov mapping may be extendable to $\mathcal{B}(\Theta \times H_j^u)$. Remark 4.2 however remains valid, so the extension is not generally unique, i.e., measures on $(\Theta \times H_j^u, \mathcal{B}(\Theta \times H_j^u))$ are not fully determined by belief hierarchies.

coarser than the topology induced by the Prohorov metric.[26] As a mapping from H_i^u to $\mathcal{M}_0(\Theta \times H_j^u)$ with the topology of weak convergence, therefore, the Kolmogorov mapping is continuous, but its inverse need not be.

6 An Embedding Theorem for Abstract Type Spaces

As discussed in the introduction, the universal type space is called "universal" because it has room for a complete representation of *all* strategically relevant aspects of a given specification of incomplete information. As indicated by (2.14) - (2.16), any Θ-based abstract type space with continuous belief functions can be mapped into a Θ-based space of belief hierarchies. Proposition 2.3 shows that, if the abstract type space is nonredundant and if the space of belief hierarchies is endowed with the product topology, this mapping is actually an embedding, i.e., the Θ-based abstract type space is homeomorphic to a subset of the Θ-based universal type space. In this section, I study what becomes of Proposition 2.3 when the spaces of belief hierarchies have the uniform topology.[27]

Recall that a Θ-based abstract type space model for two agents is given by a pair $T_i, i = 1, 2$, of compact metric spaces and a pair $b_i : T_i \to \mathcal{M}(\Theta \times T_j)$, $i = 1, 2$ and $j \neq i$, of continuous belief mappings with the interpretation that, if $t_i \in T_i$ is the "type" of agent i, then $b_i(t_i) \in \mathcal{M}(\Theta \times T_j)$ is the agent's probabilistic belief about the exogenous parameter $\theta \in \Theta$ and the type $t_j \in T_j$ of the other agent. In Proposition 2.3, continuity of the belief mappings was defined in terms of the given topologies on T_i and the topology of weak convergence on $\mathcal{M}(\Theta \times T_j)$, for $i = 1, 2$ and $j \neq i$. The following stronger continuity condition provides a basis for an anlogue to Proposition 2.3 when the space of belief hierarchies has the uniform topology.

Continuity in Total Variation (CTV) For $i = 1, 2$ and $j \neq i$, the belief function b_i maps T_i continuously into the space $\mathcal{M}^{TV}(\Theta \times T_j)$ of probability measures on $(\Theta \times T_j, \mathcal{B}(\Theta \times T_j))$ endowed with the topology induced by the total-variation metric ρ^{TV}, where, for any two measures $\nu, \hat{\nu}$ on $(\Theta \times T_j, \mathcal{B}(\Theta \times T_j))$,

$$\rho^{TV}(\nu, \hat{\nu}) = \sup_{B \in \mathcal{B}(\Theta \times T_j)} |\nu(B) - \hat{\nu}(B)| . \tag{6.1}$$

[26] The proof of this assertion is the same as the proof of the first statement of Theorem 5, p. 238, in Billingsley (2008).

[27] The importance of this question is discussed in Morris (2002).

Proposition 6.1 *Given the Θ-based abstract-type-space model $\{T_i, b_i\}_{i=1}^2$, with compact metric type spaces, for any i, the belief function b_i satisfies CTV, then the function $t_i \longmapsto \{\varphi_i^k(t_i)\}_{k=1}^\infty$ that is defined by (2.15) and (2.16) maps T_i continuously into H_i^u. If this function is injective, it is an embedding, i.e., a homeomorphism between T_i and a subspace of H_i^u*

Proof. For any $k > 1$, any $t_i \in T_i$ and any $B^{k-1} \in \mathcal{B}(X^{k-1})$, (2.16) implies that

$$\varphi_i^k(B^{k-1}|t_i) = b_i\left(\left(\pi_\Theta \times (\varphi_j^{k-1} \circ \pi_{T_j})\right)^{-1}(B^{k-1})|t_i\right). \qquad (6.2)$$

Thus, for any t_i and $\hat{t}_i$ in T_i,

$$\begin{aligned} \sup_{B^{k-1}\in\mathcal{B}(X^{k-1})} \left|\varphi_i^k(B^{k-1}|t_i) - \varphi_i^k(B^{k-1}|\hat{t}_i)\right| &\leq \sup_{B\in\mathcal{B}(\Theta\times T_j)} \left|b_i(B|t_i) - b_i(B|\hat{t}_i)\right| \\ &= \rho^{TV}(b_i(t_i), b_i(\hat{t}_i)). \qquad (6.3) \end{aligned}$$

Because the Prohorov distance between any two measures is no greater than the total-variation distance,[28] for any $k > 1$ and t_i and $\hat{t}_i$ in T_i, one also has

$$\rho^k(\varphi_i^k(t_i), \varphi_i^k(\hat{t}_i)) \leq \sup_{B^{k-1}\in\mathcal{B}(X^{k-1})} \left|\varphi_i^k(B^{k-1}|t_i) - \varphi_i^k(B^{k-1}|\hat{t}_i)\right| \qquad (6.4)$$

and therefore,

$$\rho^k(\varphi_i^k(t_i), \varphi_i^k(\hat{t}_i)) \leq \rho^{TV}(b_i(t_i), b_i(\hat{t}_i)). \qquad (6.5)$$

By the same argument, one also has

$$\rho^1(\varphi_i^1(t_i), \varphi_i^1(\hat{t}_i)) \leq \rho^{TV}(b_i(t_i), b_i(\hat{t}_i)) \qquad (6.6)$$

for any t_i and $\hat{t}_i$ in T_i. Hence

$$\rho^u\left(\{\varphi_i^k(t_i)\}_{k=1}^\infty . \{\varphi_i^k(\hat{t}_i)\}_{k=1}^\infty\right) \leq \rho^{TV}(b_i(t_i), b_i(\hat{t}_i)) \qquad (6.7)$$

for all t_i and $\hat{t}_i$ in T_i, and the first statement of the proposition follows. The second statement follows by standard arguments. ■

To understand the difference between this result and Proposition 2.3, it is useful to consider the following version of Rubinstein's electronic mail game. Let $I = 2$ and $\Theta = \{0, 1\}$. For $i \in I$, let $T_i = \{0, \frac{1}{2}, \frac{2}{3}, ..., 1\}$ and endow

[28] See, e.g., Gibbs and Su (2002), p. 428.

T_i with the subspace topology that is induced by the usual topology on the unit interval. Specify the belief function for agent 1 so that, for $t_1 = 0$,

$$b_1(t_1) = \delta_{(0,0)}, \tag{6.8}$$

for $t_1 = \frac{n}{n+1} > 0$,

$$b_1(t_1) = \frac{1}{2}\delta_{(1,\frac{n-1}{n})} + \frac{1}{2}\delta_{(0,\frac{n}{n+1})}, \tag{6.9}$$

and, for $t_1 = 1$,

$$b_1(t_1) = \delta_{(1,1)}, \tag{6.10}$$

where, for any $\theta \in \Theta$ and $t_2 \in T_2$, $\delta_{(\theta,t_2)}$ is the degenerate measure that assigns all probability mass to the singleton $\{(\theta, t_2)\}$. Similarly, specify the belief function for agent 2 so that, for $t_2 = 0$,

$$b_2(t_2) = \frac{1}{2}\delta_{(0,0)} + \frac{1}{2}\delta_{(1,\frac{1}{2})}, \tag{6.11}$$

for $t_2 = \frac{n}{n+1} > 0$,

$$b_2(t_2) = \frac{1}{2}\delta_{(1,\frac{n}{n+1})} + \frac{1}{2}\delta_{(0,\frac{n+1}{n+2})}, \tag{6.12}$$

and, for $t_2 = 1$,

$$b_2(t_2) = \delta_{(1,1)}, \tag{6.13}$$

where now for any $\theta \in \Theta$ and $t_1 \in T_1$, $\delta_{(\theta,t_1)}$ is the degenerate measure that assigns all probability mass to the singleton $\{(\theta, t_1)\}$.

One easily verifies that, for $i = 1, 2$, the belief function b_i is continuous if the space $\mathcal{M}(\Theta \times T_j)$ has the topology of weak convergence and that it is discontinuous at $t_i = 1$ if $\mathcal{M}(\Theta \times T_j)$ has the topology induced by the total-variation metric. The associated mapping $t_i \longmapsto \{\varphi_i^k(t_i)\}_{k=1}^{\infty}$ from T_i to H_i is continuous if H_i has the product topology and discontinuous at $t_i = 1$ if H_i has the uniform topology. Because of the discontinuity, the model fails to have the lower strategic convergence property.

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 $\theta = 1$ and if the other agent also chooses the action a_1 and otherwise the payoff $-X < 0$, where $X > Y$. For each agent i, one easily verifies that, if $\varepsilon < \frac{1}{2}(Y - X)$, then, for any $t_i \in T_i \backslash \{1\}$, a_0 is the unique interim correlated ε-rationalizable action of the agent, but for $t_i = 1$, the action a_1 is also interim rationalizable. The lower strategic convergence property of Dekel et al. (2006) fails to hold.

The failure is directly related to the discontinuity of b_i when the range of b_i has the topology induced by the total-variation metric.

The failure is also related to the discontinuity of the mapping $t_i \longmapsto \{\varphi_i^k(t_i)\}_{k=1}^{\infty}$ from T_i to H_i when H_i has the uniform topology. If the mapping $t_i \longmapsto \{\varphi_i^k(t_i)\}_{k=1}^{\infty}$ from T_i to H_i^u were continuous, then, by Theorem 1 of Chen et al. (2010), the lower strategic convergence property would hold in the belief-hierarchies formulation of the game. By Proposition 6.1, this property would also hold in the abstract-type-space formulation of the game.[29]

A remaining question is whether the topology induced by the total-variation metric is actually the coarsest topology on $\mathcal{M}(\Theta \times T_i)$ for which the mapping $t_i \longmapsto \{\varphi_i^k(t_i)\}_{k=1}^{\infty}$ from T_i to H_i^u is continuous. Could one obtain the conclusion of Proposition 6.1 for a coarser topology on the range of the belief function b_i? An answer to this question would complete the program of extending the results reviewed in Section 2 to a setting where beliefs of arbitrarily high orders can have a significant impact on strategic behaviour.

References

[1] Bachman, G., and A. Sultan (1977), Extensions of Regular Lattice Measures with Topological Applications, *Journal of Mathematical Analysis and Applications* 57, 539-559.

[2] Bachman, G., and A. Sultan (1980), On Regular Extensions of Measures, *Pacific Journal of Mathematics* 86 (2), 389-395.

[3] Banach, S., and C. Kuratowski (1929), Sur une généralisation du problème de la mesure, *Fundamenta Mathematicae* 14, 127-131.

[29] A direct argument to this effect is given by Engl (1995). For abstract type spaces, Engl shows that, if beliefs have the topology of setwise convergence, the the Nash equilibrium correspondence has the desired lower hemi-continuity property. His arguments are easily extended to the correspondence of interim ε-rationalizable outcomes. Since the topology of setwise convergence is coarser than the topology induced by the total-variation metric, the lower hemi-continuity property also holds if beliefs are endowed with the latter topology. Pintér (2005) uses the topology of setwise convergence for beliefs at different levels in a belief hierarchy, without however relating the topology to any properties of strategic behavior.

[4] Bartoszynski, T., and L. Halbeisen (2003), On a Theorem of Banach and Kuratowski and K-Lusin Sets, *Rocky Mountain Journal of Mathematics* 33 (4), 1223-1231.

[5] Berge, C. (1959), *Espaces Topologiques*, Dunod, Paris.

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

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

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

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

[10] Chen, Y., A. Di Tillio, E. Faingold, and S. Xiong (2017), Characterizing the Strategic Impact of Misspecified Beliefs, *Review of Economic Studies* 84 (4), 1424-1471.

[11] Cohen, P.J. (1966), *Set Theory and the Continuum Hypothesis*, W.A. Benjamin, Reading, MA.

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

[13] Dekel, E., D. Fudenberg, and S. Morris (2007), Interim Correlated Rationalizability, *Theoretical Economics* 2, 15-40.

[14] Dudley, R.M. (1966), Weak Convergence of Probabilities on Nonseparable Metric Spaces and Empirical Measures on Euclidean Spaces, *Illinois Journal of Mathematics* 10, 109-126.

[15] Dudley, R.M. (1967), Measures on Non-Separable Metric Spaces, *Illinois Journal of Mathematics* 11, 449-453.

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

[17] Ely, J.C., and M. Peski (2006), Hierarchies of Beliefs and Interim Rationalizability, *Theoretical Economics* 1, 19-65.

[18] Engl, G. (1995), Lower Hemi-Continuity of the Nash Equilibrium Correspondence, *Games and Economic Behavior* 9, 151-160.

[19] Fudenberg, D., and D.K. Levine (1986), Limit Games and Limit Equilibria, *Journal of Economic Theory* 38, 261-279.

[20] Gibbs A.L., and Su F.E. (2002), On Choosing and Bounding Probability Metrics, *International Statistical Review* 70, 419-435.

[21] Gizatulina, A., and M. Hellwig (2017), The Generic Possibility of Full Surplus Extraction in Models with Large Type Spaces, *Journal of Economic Theory* 170 (6), 385-416.

[22] Gul, F. (1998), A Comment on Aumann's Bayesian View, *Econometrica* 66 (4), 923-928.

[23] Harsanyi, J.C. (1967/68), Games with Incomplete Information Played by 'Bayesian' Players I - III, *Management Science* 14, 159-182, 320-334, 486-502.

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

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

[26] Heifetz, A., and D. Samet (1999), Coherent Beliefs are Not Always Types, *Journal of Mathematical Economics* 32, 475-488.

[27] Hellman, Z., and D. Samet (2012), How Common are Common Priors, *Games and Economic Behavior* 74, 517-525.

[28] Hellwig, M.F. (1996), Sequential Decisions under Uncertainty and the Maximum Theorem, *Journal of Mathetical Economics* 25 (4), 443-464.

[29] Hellwig, M.F. (2017/2022), Probability Measures on Product Spaces with a Uniform Metric, Preprint /2017, Max Planck Institute for Research on Collective Goods, Bonn, revised July 2022. http://www.coll.mpg.de/pdf_dat/2017_06online.pdf.

[30] Jordan, J.S. (1977), The Continuity of Optimal Dynamic Decision Rules, *Econometrica* 46 (6), 1365-1376.

[31] Marczewski, E., and R. Sikorski (1948), Measures in Non-Separable Metric Spaces, *Colloquium Mathematicum* 1, 133-139.

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

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

[34] Morris, S. (2002), Typical Types, mimeo, Cowles Foundation, Yale University, https://www.princeton.edu/~smorris/pdf/typicaltypes.pdf.

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

[36] Pintér, M., (2005), Types on a Purely Measurable Parameter Space, *Economic Theory* 26, 129-139.

[37] Pintér, M. (2010), The Non-Existence of a Universal Topological Type Space, *Journal of Mathematical Economics* 46, 223-229.

[38] Pollard, D. (1979), Weak Convergence on Non-Separable Metric Spaces, *Journal of the Australian Mathematical Society* (Series A) 28, 197-204.

[39] Rubinstein, A. (1989), The Electronic Mail Game: Strategic Behavior under 'Almost Common Knowledge', *American Economic Review* 79, 385-391.

[40] Ulam, S. (1930), Zur Masstheorie in der allgemeinen Mengenlehre, *Fundamenta Mathematicae* 16, 140-150.