

Discussion Papers of the
Max Planck Institute for
Research on Collective Goods
2017/6



**Probability Measures on
Product Spaces with Uniform
Metrics**

Martin F. Hellwig

MAX PLANCK
SOCIETY





Probability Measures on Product Spaces with Uniform Metrics

Martin F. Hellwig

May 2017

This version: July 2022

Probability Measures on Product Spaces with Uniform Metrics*

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

For a countable product of complete separable metric spaces, with a topology induced by a uniform metric, the σ -algebra generated by the open balls, which was introduced by Dudley (1966), coincides with the product σ -algebra. Any probability measure on the product space with this σ -algebra is *quasi-separable* in the sense that, for any union of open balls that has full measure, there is a countable sub-union that also has full measure. With suitably adapted definitions, the topology of weak convergence on the space of such measures is equivalent to the topology induced by the Prohorov metric. The projection mapping from such measures to sequences of measures on the first ℓ factors, $\ell = 1, 2, \dots$, is a homeomorphism if the range of this mapping is also given a uniform metric. These findings are relevant for game theory, where Chen et al. (2010, 2017) have proposed to endow the space of belief hierarchies with a uniform metric topology because this properly captures the continuity properties of strategic behaviour.

Key Words: Product spaces with uniform metrics, weak convergence of non-Borel measures, σ -algebras generated by the open balls, quasi-separable measures, Prohorov metric.

MSC Classification 60B05

JEL Classification: C02, C72

*Without implicating them, I thank Eduardo Faingold and Alia Gizatulina for helpful discussions.

1 Introduction

Let X_1, X_2, \dots be non-singleton complete separable metric spaces with metrics ρ_1, ρ_2, \dots . Suppose that the product

$$X = \prod_{k=1}^{\infty} X_k \tag{1.1}$$

has the topology induced by the uniform metric ρ^u where, for any x and \hat{x} in X ,

$$\rho^u(x, \hat{x}) = \sup_k \rho_k(\pi_k(x), \pi_k(\hat{x})) \tag{1.2}$$

and π_k is the projection from X to X_k . I use the notation X^u to indicate that X has the topology induced by ρ^u .

The space X^u is non-separable. This creates difficulties for working with the Borel σ -algebra $\mathcal{B}(X^u)$ and the space $\mathcal{M}(X^u)$ of probability measures on $(X^u, \mathcal{B}(X^u))$. To avoid these difficulties, Dudley (1966, 1967) proposed to work with the σ -algebra $\mathcal{B}_0(X^u)$ that is generated by the ρ^u -open balls and the space $\mathcal{M}_0(X^u)$ of probability measures on $(X^u, \mathcal{B}_0(X^u))$.¹ He endowed $\mathcal{M}_0(X^u)$ with the topology of weak convergence, adapted to take account of the fact that $\mathcal{B}_0(X^u)$ is not the Borel σ -algebra on X^u . Thus, a sequence $\{\mu^r\}$ of measures in $\mathcal{M}_0(X^u)$ converges weakly to a measure $\mu \in \mathcal{M}_0(X^u)$ if and only if

$$\int_{X^u} f(x) d\mu^r(x) \rightarrow \int_{X^u} f(x) d\mu(x) \tag{1.3}$$

for all f in the space $\mathcal{C}_0(X^u)$ of bounded, ρ^u -continuous, and $\mathcal{B}_0(X^u)$ -measurable real-valued functions on X .

Dudley did not actually study the topology that is induced by the concept of weak convergence on $\mathcal{M}_0(X^u)$ but merely considered the convergence behaviour of sequences of integrals of bounded continuous functions, assuming that the limit measure of such a sequence can be extended to a Borel measure for the uniform topology and is therefore concentrated on a separable set.² This analysis involves an implicit asymmetry between measures on $\mathcal{B}_0(X^u)$ that can be extended to Borel measures and measures on $\mathcal{B}(X^u)$ that cannot be extended to Borel measures. For Dudley, this asymmetry

¹See also Dudley (1978) and Pollard (1979).

²Pollard (1979) implicitly makes the same assumption. He uses the separability of the support of the limit measure in order to define a topology that is smaller than the topology of weak convergence on the space of measures on the σ -algebra generated by the open balls, but sufficiently large to provide for weak convergence in a neighbourhood of the limit measure.

did not matter because he was interested in the convergence properties of sequences of stochastic processes when the space of sample paths of the processes has the uniform topology; the limit measures in his convergence theorems satisfy his extendability condition.

In other contexts though, the asymmetry is problematic. I came across the issue in the context of games of incomplete information. In this context, a natural starting point is given by the notion of a *belief hierarchy* of an agent, beginning with the agent's first-order belief, a subjective probability distribution over states of nature, continuing with the agent's second-order belief, a subjective probability distribution over states of nature and other agents' first-order beliefs, then the agent's third-order belief, a subjective probability distribution over states of nature and other agents' first- and second-order beliefs, and so on. A belief hierarchy is an element of a product space with factors involving the spaces of beliefs of different orders.

Mertens and Zamir (1985), who introduced this formalization of belief hierarchies, assumed that, for each k , the space of beliefs of order k is appropriately topologized by the topology of weak convergence and that the space of belief hierarchies of all orders is appropriately topologized by the associated product topology. Dekel et al. (2006) and Chen et al. (2010, 2017) criticized the use of the product topology as being too coarse to capture all the continuity properties of strategic behaviour that one may be interested in.³ As an alternative, Chen et al. (2010, 2017) proposed the topology that is induced by the uniform metric; they showed that this topology is actually equivalent to the uniform strategic topology of Dekel et al. (2006).

If the spaces of different agents' belief hierarchies have the product topology, then according to Mertens and Zamir (1985), there is a homeomorphism that associates any one belief hierarchy of an agent with a probability measure over states of nature and other agents' belief hierarchies. In trying to establish a similar result for the case where belief hierarchies have the uniform topology, one finds that, with the uniform topology, the space of states of nature and other agents' belief hierarchies is non-separable and the Borel σ -algebra on this space is not well suited for the analysis. As in Dudley (1966, 1967), the difficulties are avoided if one uses the (non-Borel) σ -algebra that is generated by the ρ^u -open balls, rather than the Borel σ -algebra associated with the uniform topology.

However, in this application of the σ -algebra $\mathcal{B}_0(X^u)$ and the space

³In particular, the product topology on belief hierarchies is incompatible with the observation that beliefs of arbitrarily high orders may have a significant impact on strategic behaviour, as indicated by Rubinstein's (1989) electronic mail game.

$\mathcal{M}_0(X^u)$, an asymmetric treatment of measures in $\mathcal{M}_0(X^u)$ that can be extended to the Borel σ -algebra for X^u and measures that cannot be so extended is unwarranted. Therefore, it is useful to have more general results about the topology of weak convergence on $\mathcal{M}_0(X^u)$.

For this purpose, I introduce a Prohorov metric on $\mathcal{M}_0(X^u)$. The ρ^u -based Prohorov distance between any two measures μ and $\hat{\mu}$ in $\mathcal{M}_0(X^u)$ is defined as the greatest lower bound on the set of $\varepsilon > 0$ such that

$$\mu(B) \leq \hat{\mu}(B^\varepsilon) + \varepsilon \text{ and } \hat{\mu}(B) \leq \mu(B^\varepsilon) + \varepsilon \quad (1.4)$$

for all sets $B \in \mathcal{B}_0(X^u)$. The following lemma implies that the set B^ε in (1.4) is always well defined.

Lemma 1.1 *For any $B \in \mathcal{B}_0(X^u)$ and any $\varepsilon > 0$, the set*

$$B^\varepsilon := \bigcup_{x \in B} \{x' \in X \mid \rho^u(x', x) < \varepsilon\} \quad (1.5)$$

is also an element of $\mathcal{B}_0(X^u)$.

Thus, the ρ^u -based Prohorov distance is always well defined. The following result shows that it provides a suitable metric for the topology of weak convergence on $\mathcal{M}_0(X^u)$.

Proposition 1.2 *If the cardinal \mathfrak{c} of the continuum is not atomlessly measurable, the topology of weak convergence on $\mathcal{M}_0(X^u)$ is metrizable by the ρ^u -based Prohorov metric on $\mathcal{M}_0(X^u)$.*

For Borel measures, metrizability of the topology of weak convergence by the Prohorov metric is established in Appendix III of Billingsley (1968). The analysis there involves two steps. First, Theorem 2, p. 235, gives a condition under which any Borel measure is separable in the sense that any family of open sets to which the measure assigns probability one has a countable subfamily to which the measure also assigns probability one. Second, Theorem 5, p. 238, shows that, on a space of separable Borel measures, the topology of weak convergence is equivalent to the topology induced by the Prohorov metric.

Because $\mathcal{M}_0(X^u)$ is *not* a space of Borel measures, I cannot use Billingsley's arguments as such. However, I can use similar arguments. I replace

the notion of separability of a measure by a notion of *quasi-separability*, saying that a measure is quasi-separable if and only any family of open balls to which the measure assigns probability one has a countable subfamily to which the measure also assigns probability one.

With this change, the following two results provide suitable analogues of Theorems 2, p. 235, and 5, p. 238, in Billingsley (1968). Obviously, Proposition 1.2 is an immediate implication of these two results.

Proposition 1.3 *If the cardinal \mathfrak{c} of the continuum is not atomlessly measurable, the measures in $\mathcal{M}_0(X^u)$ are quasi-separable.*

Proposition 1.4 *If the measures in $\mathcal{M}_0(X^u)$ are quasi-separable, the topology of weak convergence on $\mathcal{M}_0(X^u)$ is metrizable by the ρ^u -based Prohorov metric.*

A proof of Proposition 1.3 is given in Section 2 below, proofs of Lemma 1.1 and Proposition 1.4 are given in Section 3. Section 4 discusses the game-theoretic application of the analysis.

In these results, the assumption that the cardinal \mathfrak{c} of the continuum is not atomlessly measurable replaces Billingsley's (1968) assumption that no discrete subset of the space has a cardinal that is atomlessly measurable.⁴ Whereas discrete sets depend on the topology, my somewhat stronger assumption does not depend on the topology of the underlying space.

By a theorem of Banach and Kuratowski (1929), the assumption that the cardinal \mathfrak{c} of the continuum is not atomlessly measurable is implied by the Continuum Hypothesis (CH). The older literature, such as Dudley (1967) or Billingsley (1968), invokes this fact to suggest that the assumption is unproblematic. Under the influence of Cohen (1966), CH has increasingly met with criticism in recent decades. It is therefore worth noting that CH is *not* necessary for the condition that \mathfrak{c} is not atomlessly measurable. Bartoszyński and Halbeisen (2003) point to the fact that, in Banach and Kuratowski (1929), the conclusion that \mathfrak{c} is not atomlessly measurable follows from the existence of what they call a *BK-matrix*. The existence of such a matrix is implied by CH but is also compatible with the negation of CH.⁵

⁴Billingsley uses the term "nonmeasurable". The more recent literature reserves the distinction "measurable - nonmeasurable" for the case of binary measures taking the values zero and one, and uses the distinction "atomlessly measurable - not atomlessly measurable" for the case considered by Billingsley.

⁵If the underlying space is the unit interval, a BK-matrix is a doubly infinite array

2 Quasi-Separability of Measures in $\mathcal{M}_0(X^u)$

This section is devoted to the proof of Proposition 1.3. I assume throughout without further mention that the cardinal \mathbf{c} of the continuum is not atomlessly measurable.

The first two lemmas concern the σ -algebra $\mathcal{B}_0(X^u)$.

Lemma 2.1 *The set X and the σ -algebra $\mathcal{B}_0(X^u)$ have the cardinal \mathbf{c} of the continuum.*

Proof. Because each the spaces X_1, X_2, X_3, \dots has more than one element, the cardinal of X is at least $2^\omega = \mathbf{c}$ where ω is the cardinal of the natural numbers and \mathbf{c} is the cardinal of the continuum. It is also no greater than \mathbf{c} because, when endowed with the product topology, X is a product of complete separable metric spaces, X itself is a complete separable metric space so the cardinal of X is either finite or ω or \mathbf{c} .⁶

To show that the cardinal of the σ -algebra $\mathcal{B}_0(X^u)$ is also equal to \mathbf{c} , I note that, by a result of Dudley (1967, p. 449), the claim is true if there is no dense set in X^u that has a cardinal less than \mathbf{c} . To see that X^u satisfies this condition, for $k = 1, 2, \dots$, let x_k^1, x_k^2 be two distinct elements of X_k and consider the set $\hat{X} = \{x = (x_1, x_2, \dots) \in X \mid x_k \in \{x_k^1, x_k^2\} \text{ for all } k\}$. Clearly, $\hat{X} \subset X$ has the cardinal \mathbf{c} , and so does any ρ^u -dense subset of X . ■

Lemma 2.2 *The σ -algebra $\mathcal{B}_0(X^u)$ is coextensive with the Borel σ -algebra $\mathcal{B}(X^p)$ for the product topology on X .*

Proof. I first prove that $\mathcal{B}_0(X^u) \subset \mathcal{B}(X^p)$. Since $\mathcal{B}_0(X^u)$ is the smallest σ -algebra that contains the ρ^u -open balls in X , it suffices to show that $\mathcal{B}(X^p)$ contains the ρ^u -open balls in X . Consider any $x = \{x_k\}_{k=1}^\infty \in X$ and any $r > 0$. The ρ^u -open r -ball around x is given as

$$B^u(x, r) = \bigcup_{n=1}^{\infty} \prod_{k=1}^{\infty} B_k(x_k, \max(r - \frac{1}{n}, 0)), \quad (2.1)$$

of sets A_k^i , $i, k \in \mathbb{N}$, such that (i) for each $i \in \mathbb{N}$, $\cup_{k \in \mathbb{N}} A_k^i = [0, 1]$, (ii) for each $i \in \mathbb{N}$, and all $k, k' \in \mathbb{N}$, $A_k^i \cap A_{k'}^i = \emptyset$ 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 show that such a matrix exists if and only if there exists a K -Lusin set of the size of the continuum.

⁶Theorem 13.1.1 in Dudley (2002).

where, for each k and $r' > 0$, $B_k(x_k, r')$ is the ρ_k -open r' -ball around x^k . Equation (2.1) can be rewritten as

$$B^u(x, r) = \bigcap_{\ell=1}^{\infty} \bigcup_{n=1}^{\infty} \left[\prod_{k=1}^{\ell} B_k(x_k, \max(r - \frac{1}{n}, 0)) \times \prod_{k=\ell+1}^{\infty} X_k \right], \quad (2.2)$$

which is a countable intersection of countable unions of cylinder sets of the form

$$C = B_1 \times \dots \times B_{\ell} \times X_{\ell+1} \times X_{\ell+2} \times \dots, \quad (2.3)$$

where, for $k = 1, \dots, \ell$, B_k is a ρ_k -open subset of X_k . Since $\mathcal{B}(X^p)$ contains the cylinder sets and is closed under countable unions and intersections, (2.2) implies that $B^u(x, r) \in \mathcal{B}(X^p)$

To prove that, conversely, $\mathcal{B}(X^p) \subset \mathcal{B}_0(X^u)$, it suffices to show that any set of the form (3.4) belongs to $\mathcal{B}_0(X^u)$. For this purpose, fix some ℓ and ρ_k -open $B_k \subset X_k$, $k = 1, \dots, \ell$, and let C be given by (3.4). For any $N > \ell$, consider the projection

$$\Pi^N(C) = B_1 \times \dots \times B_{\ell} \times X_{\ell+1} \times \dots \times X_N. \quad (2.4)$$

When endowed with the uniform metric

$$\rho^N(\{x^k\}_{k=1}^N, \{\hat{x}^k\}_{k=1}^N) := \max_{k=1}^N \rho_k(x^k, \hat{x}^k), \quad (2.5)$$

the finite product $X^1 \times \dots \times X^N$ is separable and has a countable basis $\{z^1, z^2, \dots\}$. For $m = 1, 2, \dots$ and $r > 0$, let $U^N(k, r)$ be the ρ^N -open r -ball around z^m . 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,

$$\Pi^N(C) = \bigcup_{M^*(C)} U^N(m, r_m). \quad (2.6)$$

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

$$\hat{U}^N(m, r_m) = U^N(m, r_m) \times B_{N+1}(x^{N+1}, r_m) \times B_{N+2}(x^{N+2}, r_m) \times \dots \quad (2.7)$$

and

$$\bar{U}^N(m, r_m) = \bigcup_{n=1}^{\infty} [\hat{U}^N(m, r_m - \frac{1}{n}) \times B_{N+1}(x^{N+1}, r_m - \frac{1}{n}) \times B_{N+2}(x^{N+2}, r_m - \frac{1}{n}) \times \dots]. \quad (2.8)$$

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

$$C = \bigcup_{N=\ell+1}^{\infty} \bigcap_{m=1}^{\infty} \bar{U}^N(m, r_m). \quad (2.9)$$

Since $\mathcal{B}_0(X^u)$ is closed under countable unions and intersections, it follows that $C \in \mathcal{B}_0(X^u)$. More generally, therefore, $\mathcal{B}_0(X^u) \supset \mathcal{B}(X^p)$. Upon combining the two inclusions, one obtains $\mathcal{B}_0(X^u) = \mathcal{B}(X^p)$. ■

Lemma 2.3 *For any k , the projection mapping π^k from X^u to $X_0 \times \dots \times X_k$ is continuous and open.*

Proof. Continuity is trivial. To prove openness, I note that any ρ^u -open set $U \subset X^u$ can be written in the form

$$U = \bigcup_{\ell \in \mathcal{L}} V_\ell \quad (2.10)$$

where each $V_\ell, \ell \in \mathcal{L}$, is a ρ^u -open ball around an element $x^\ell = (x_0^\ell, x_1^\ell, \dots)$ of U . Any one of these balls takes the form

$$V_\ell = \bigcup_{n=1}^{\infty} \prod_{j=0}^{\infty} B_j(x_j^\ell, r^\ell - \frac{1}{n}), \quad (2.11)$$

with projections

$$\pi^k(V_\ell) = \bigcup_{n=1}^{\infty} \prod_{j=0}^k B_j(x_j^\ell, r^\ell - \frac{1}{n}) = \prod_{j=0}^k B_j(x_j^\ell, r^\ell), \quad (2.12)$$

which are open in $X_0 \times \dots \times X_k$, $k = 0, 1, 2, \dots$. Since

$$\pi^k(U) = \bigcup_{\ell \in \mathcal{L}} \pi^k(V_\ell), \quad (2.13)$$

it follows that $\pi^k(U)$ is open if U is ρ^u -open. ■

Corollary 2.4 *For any k , the formula*

$$\Pi^k(U) = \pi^k(U) \times X_{k+1} \times X_{k+2} \times \dots \quad (2.14)$$

defines a continuous and open mapping Π^k from X^u to X^p .

For any ρ^u -open set $U \subset X^u$ and all k , one obviously has $U \subset \Pi^{k+1}(U) \subset \Pi^k(U)$. Therefore, the set

$$V(U) = \bigcap_{k=0}^{\infty} \Pi^k(U) \quad (2.15)$$

is well-defined and satisfies

$$U \subset V(U). \quad (2.16)$$

Lemma 2.5 *Let U, \hat{U} be two ρ^u -open subsets of X and suppose that, for some $\varepsilon > 0$,*

$$\rho^u(x, \hat{x}) \geq \varepsilon \quad (2.17)$$

for all $x \in U$ and \hat{x} in \hat{U} . Then $V(U) \cap V(\hat{U}) = \emptyset$.

Proof. The premise of the lemma implies that, for every $\delta > 0$ and every $x \in U$ and \hat{x} in \hat{U} , there exists k such that

$$\rho_k(\pi_k(x), \pi_k(\hat{x})) \geq \varepsilon - \delta. \quad (2.18)$$

If the lemma were false, then for some $x^* \in X$, one would have $x^* \in \Pi^k(U)$ for all k and $x^* \in \Pi^k(\hat{U})$ for all k . Since $\rho_k(\pi_k(x^*), \pi_k(x^*)) = 0$ for all k , this is incompatible with (2.18). ■

As a countable intersection of open sets in X^p , $V(U)$ belongs to the σ -algebra $\mathcal{B}(X^p)$ and therefore also to the σ -algebra $\mathcal{B}_0(X^u)$. For any ρ^u -open set $U \subset X^u$, therefore, and any measure $\mu \in \mathcal{M}_0(X^u)$, the quantity $\mu(V(U))$ is well defined. By definition,

$$\mu(V(U)) \geq \mu^*(U), \quad (2.19)$$

where μ^* , the outer measure induced by the measure μ , is defined by the formula

$$\mu^*(U) := \inf \sum_{i=1}^{\infty} \mu(B_i), \quad (2.20)$$

where the infimum is taken over all sequences $\{B_i\}_{i=1}^{\infty}$ such that $B_i \in \mathcal{B}_0(X^u)$ for all i and $U \subset \bigcup_{i=1}^{\infty} B_i$.⁷ The following result plays a similar role as Theorem 1 in Marczewski and Sikorski (1948); the proof idea is due to Banach (1930).

⁷See Dudley (2002), p. 89.

Lemma 2.6 *Let $\{U_i\}_{i \in I}$ be a collection of ρ^u -open subsets of X such that, for some $\varepsilon > 0$, for all i and j in I ,*

$$\rho^u(x^i, x^j) \geq \varepsilon \quad (2.21)$$

for all $x^i \in U_i$ and all $x^j \in U_j$. For any $\mu \in \mathcal{M}_0(X^u)$, there exists a finite or countable set $I^ \subset I$ such that*

$$\mu^*(U_i) > 0 \text{ for } i \in I^* \quad (2.22)$$

and

$$\mu^* \left(\bigcup_{i \in I \setminus I^*} U_i \right) = 0. \quad (2.23)$$

Proof. By (2.19), it suffices to prove that there exists a finite or countable set $I^{**} \subset I$ such that

$$\mu(V(U_i)) > 0 \text{ for } i \in I^{**} \quad (2.24)$$

and

$$\mu \left(V \left(\bigcup_{i \in I \setminus I^{**}} U_i \right) \right) = 0. \quad (2.25)$$

The set I^{**} is defined as the union over n of the sets

$$I_n := \left\{ i \in I \mid \mu(V(U_i)) > \frac{1}{n} \right\}.$$

I claim that each set I_n has at most n elements and, therefore, that I^{**} , as a countable union of these sets, is at most countable. For suppose that some set I_n has $n' > n$ elements, say $i_1, \dots, i_{n'}$. Then, by Lemma 2.5 and the countable additivity of the measure μ ,

$$\mu \left(\bigcup_{i=i_1}^{i_{n'}} V(U_i) \right) = \sum_{i=i_1}^{i_{n'}} \mu(V(U_i)) > n' \cdot \frac{1}{n} > 1,$$

which is impossible.

For any set $\hat{I} \subset I$, the union $\bigcup_{i \in \hat{I}} U_i$ is an open set in X^u . By the argument

given above, it follows that, for any $\hat{I} \subset I$, the set $V \left(\bigcup_{i \in \hat{I}} U_i \right)$ belongs to

the σ -algebra $\mathcal{B}(X^p)$ and therefore also to the σ -algebra $\mathcal{B}_0(X^u)$. One may therefore define a set function ν by setting

$$\nu(\hat{I}) = \mu \left(V \left(\bigcup_{i \in \hat{I}} U_i \right) \right) \quad \text{for } \hat{I} \subset I.$$

Notice that this set function is defined on all subsets of I .

By the definition of I^{**} , one obviously has

$$\nu(\{i\}) = \mu(V(U_i)) = 0 \quad \text{for } i \in I \setminus I^{**}, \quad (2.26)$$

i.e. the set function ν has no atoms in $I \setminus I^{**}$.

I also claim that ν is countably additive. Let $I_j, j = 1, 2, \dots$ be any sequence of disjoint subsets of I . Then, by Lemma 2.5 and the countable additivity of μ ,

$$\nu \left(\bigcup_{j=1}^{\infty} I_j \right) = \mu \left(V \left(\bigcup_{j=1}^{\infty} \bigcup_{i \in I_j} U_i \right) \right) = \sum_{j=1}^{\infty} \mu \left(V \left(\bigcup_{i \in I_j} U_i \right) \right) = \sum_{j=1}^{\infty} \nu(I_j),$$

which proves countable additivity of ν .

By Lemmas 2.1 and 2.5, the cardinal of $I \setminus I^{**}$ is at most \mathfrak{c} . By the assumption that the continuum is not atomlessly measurable, it follows that $\nu(I \setminus I^{**}) = 0$ and hence that

$$\mu \left(V \left(\bigcup_{i \in I \setminus I^{**}} U_i \right) \right) = \nu(I \setminus I^{**}) = 0,$$

proving (2.25). The lemma follows immediately. ■

Proposition 2.7 *For any measure $\mu \in \mathcal{M}_0(X^u)$ and any family \mathcal{G} of ρ^u -open balls covering X , there exists a countable subfamily $\mathcal{G}^* \subset \mathcal{G}$ such that $\mu(G^*) = 1$, where G^* is the union of the sets in \mathcal{G}^* .*

Proof. The proof follows along similar lines as the proof of Theorem 2, p. 235, in Billingsley (1968). Let μ and \mathcal{G} be as specified in the proposition. Because X^u is a metric space, Theorem 4.21, p. 129, in Kelley (1955), implies that \mathcal{G} has a σ -discrete refinement, i.e., there exists a family \mathcal{H} of ρ^u -open sets (not necessarily balls) covering X such that, for every $H \in \mathcal{H}$,

there exists $G(H) \in \mathcal{G}$ such that $H \subset G(H)$ and, moreover, \mathcal{H} can be written as a countable union

$$\mathcal{H} = \bigcup_{t=1}^{\infty} \mathcal{H}_t \quad (2.27)$$

where, for any t , any two sets H_{ti}, H_{tj} in \mathcal{H}_t , there exists $\varepsilon_t > 0$ such that $\rho^u(x^i, x^j) \geq \varepsilon_t$ for all $x^i \in H_{ti}$ and all $x^j \in H_{tj}$.

For any t , let I_t be the set of indices i such that $H_{ti} \in \mathcal{H}_t$. By Lemma 2.6, the set I_t has a finite or countable subset I_t^* such that

$$\mu^*(H_{ti}) > 0 \quad \text{for } i \in I_t^* \quad (2.28)$$

and

$$\mu^* \left(\bigcup_{i \in I_t \setminus I_t^*} H_{ti} \right) = 0. \quad (2.29)$$

Let \mathcal{H}_t^* be the family of sets H_{ti} , $i \in I_t^*$, and let

$$\mathcal{H}^* = \bigcup_{t=1}^{\infty} \mathcal{H}_t^*. \quad (2.30)$$

Then \mathcal{H}^* is a countable union of finite or countable sets and is itself countable.

Recalling that, for each $H \in \mathcal{H}$, there exists $G(H) \in \mathcal{G}$ such that $H \subset G(H)$, define $\mathcal{G}^* \subset \mathcal{G}$ so that $G \in \mathcal{G}^*$ if and only if $G = G(H)$ for some $H \in \mathcal{H}^*$. Since \mathcal{H}^* is countable, so is \mathcal{G}^* . Moreover, since the elements of \mathcal{G}^* belong to $\mathcal{B}_0(X^u)$, so does the countable union

$$G^* = \bigcup_{H \in \mathcal{H}^*} G(H).$$

Hence $\mu(G^*)$ is well defined and so is $\mu(X \setminus G^*) = 1 - \mu(G^*)$.

I claim that $\mu(G^*) = 1$. For suppose that $\mu(G^*) < 1$. Then also $\mu(X \setminus G^*) > 0$. Hence also $\mu^*(X \setminus G^*) > 0$, where μ^* is the outer measure defined by μ , in accordance with (2.20) above. Since $H \subset G(H)$ for all H , we also have $X \setminus G^* \subset X \setminus H^*$, where H^* is the union of the sets H_{ti} over all t and all $i \in I_t^*$. Thus, $\mu^*(X \setminus G^*) > 0$ implies $\mu^*(X \setminus H^*) > 0$. Because the refinement \mathcal{H} of the family \mathcal{G} covers X , the set $X \setminus H^*$ is a subset of the union of the sets H_{ti} over all t and all $i \in I_t \setminus I_t^*$. By the subadditivity of outer measure, it follows that

$$\mu^*(X \setminus H^*) \leq \mu^* \left(\bigcup_{t=1}^{\infty} \bigcup_{i \in I_t \setminus I_t^*} H_{ti} \right) \leq \sum_{t=1}^{\infty} \mu^* \left(\bigcup_{i \in I_t \setminus I_t^*} H_{ti} \right). \quad (2.31)$$

By (2.29) the right-hand side of (2.31) is zero. Therefore, $\mu^*(X \setminus H^*) = 0$, which is incompatible with $\mu(G^*) < 1$. This completes the proof of the proposition. ■

Proposition 1.3 follows immediately. If the cardinal \mathfrak{c} of the continuum is not atomlessly measurable, the measures in $\mathcal{M}_0(X^u)$ are quasi-separable.

3 Metrizability of the Topology of Weak Convergence on $\mathcal{M}_0(X^u)$

Metrizability of the topology of weak convergence on a space of quasi-separable non-Borel measures is proved by similar arguments as the analogous result for Borel measures, e.g., Theorem 5, p. 238, in Billingsley (1968).

As discussed in the introduction, I rely on the ρ^u -based Prohorov metric, which specifies the distance between any two measures μ and $\hat{\mu}$ in $\mathcal{M}_0(X^u)$ as the greatest lower bound on the set of $\varepsilon > 0$ such that

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

for all sets $B \in \mathcal{B}_0(X^u)$, where

$$B^\varepsilon := \bigcup_{x \in B} \{x' \in X \mid \rho^u(x', x) < \varepsilon\}. \quad (3.1)$$

Lemma 1.1 above indicates that, for any $B \in \mathcal{B}_0(X^u)$ and any $\varepsilon > 0$, the set B^ε is also an element of $\mathcal{B}_0(X^u)$ so that, for any μ and $\hat{\mu}$ in $\mathcal{M}_0(X^u)$, not only $\mu(B)$ but also $\mu(B^\varepsilon)$ is well defined.

Proof of Lemma 1.1. Consider the class \mathcal{C} of sets for which the lemma is true. From (2.1), one easily checks that, for any ρ^u -open ball $B^u(x, r)$, one has

$$(B^u(x, r))^\varepsilon = \bigcup_{x \in B^u(x, r)} \{x' \in X \mid \rho^u(x', x) < \varepsilon\} = B^u(x, r + \varepsilon) \in \mathcal{B}_0(X^u).$$

Thus, \mathcal{C} contains the ρ^u -open balls in X .

Moreover, \mathcal{C} is closed under countable unions: If $B_r, r = 1, 2, \dots$, is any countable family of sets in \mathcal{C} , a point x belongs to the ε -neighbourhood of $\bigcup_r B_r$ if and only if it belongs to B_r^ε for some r . The ε -neighbourhood of $\bigcup_r B_r$

is therefore equal to the union $\cup_r B_r^\varepsilon$. Since $B_r \in \mathcal{C}$ implies $B_r^\varepsilon \in \mathcal{B}_0(X^u)$ and $\mathcal{B}_0(X^u)$ is closed under countable unions, it follows that $\cup_r B_r^\varepsilon \in \mathcal{B}_0(X^u)$ and hence that $\cup_r B_r \in \mathcal{C}$.

Finally, \mathcal{C} is also closed under countable intersections: If $B_r, r = 1, 2, \dots$, is any countable family of sets in \mathcal{C} , a point x belongs to the ε -neighbourhood of $\cap_r B_r$ if and only if it belongs to B_r^ε for all r . The ε -neighbourhood of $\cap_r B_r$ is therefore equal to the intersection $\cap_r B_r^\varepsilon$. Since $B_r \in \mathcal{C}$ implies $B_r^\varepsilon \in \mathcal{B}_0(X^u)$ and $\mathcal{B}_0(X^u)$ is closed under countable intersections, it follows that $\cap_r B_r^\varepsilon \in \mathcal{B}_0(X^u)$ and hence that $\cap_r B_r \in \mathcal{C}$.

Since $\mathcal{B}_0(X^u)$ is the smallest σ -algebra that is closed under countable unions and countable intersections and that contains the ρ^u -open balls in X , it follows that $\mathcal{B}_0(X^u) \subset \mathcal{C}$. ■

Turning to the proof of Proposition 1.4, I note that the proof of the analogous result in Billingsley (1968) comes in two distinct steps. The first step (Theorem 3, p. 236) specifies several families of sets of measures and shows that each family is a base for the topology of weak convergence. The second step (Theorem 5, p. 238) uses this finding to establish the equivalence of the topology of weak convergence with the topology generated by the Prohorov metric.

The argument here has the same structure. For the first step of the argument, I note that the family \mathcal{F}_0 of sets taking the form

$$\left\{ \nu \in \mathcal{M}_0(X^u) \mid \left| \int_{X^u} f_i(x) d\nu(x) - \int_{X^u} f_i(x) d\mu(x) \right| < \varepsilon, i = 1, \dots, k \right\} \quad (3.2)$$

for some $\mu \in \mathcal{M}_0(X^u), \varepsilon > 0$, and f_1, \dots, f_k in $\mathcal{C}_0(X^u)$ is a base for the topology of weak convergence on $\mathcal{M}_0(X^u)$. I also consider families $\mathcal{F}_1, \mathcal{F}_2, \mathcal{F}_3$ of sets taking the forms

$$\{ \nu \in \mathcal{M}_0(X^u) \mid \nu(F_i) < \mu(F_i) + \varepsilon, i = 1, \dots, k \} \quad (3.3)$$

for some $\mu \in \mathcal{M}_0(X^u), \varepsilon > 0$, and ρ^u -closed sets F_1, \dots, F_k belonging to $\mathcal{B}_0(X^u)$ in the case of \mathcal{F}_1 ,

$$\{ \nu \in \mathcal{M}_0(X^u) \mid \nu(G_i) < \mu(G_i) + \varepsilon, i = 1, \dots, k \} \quad (3.4)$$

for some $\mu \in \mathcal{M}_0(X^u), \varepsilon > 0$, and ρ^u -open sets G_1, \dots, G_k belonging to $\mathcal{B}_0(X^u)$ in the case of \mathcal{F}_2 , and

$$\{ \nu \in \mathcal{M}_0(X^u) \mid |\nu(A_i) - \mu(A_i)| < \varepsilon, i = 1, \dots, k \} \quad (3.5)$$

for some $\mu \in \mathcal{M}_0(X^u)$, $\varepsilon > 0$, and μ -continuity sets A_1, \dots, A_k belonging to $\mathcal{B}_0(X^u)$ in the case of \mathcal{F}_3 . Each one of the families $\mathcal{F}_1, \mathcal{F}_2, \mathcal{F}_3$ is the base for a topology on $\mathcal{M}_0(X^u)$. The following result provides an analogue of Theorem 3, p. 236, in Billingsley (1968).

Proposition 3.1 *Each of the families $\mathcal{F}_1, \mathcal{F}_2, \mathcal{F}_3$ is a base for the topology of weak convergence on $\mathcal{M}_0(X^u)$.*

Proof. The proof has two parts. The first part shows that the topologies induced by the families $\mathcal{F}_1, \mathcal{F}_2, \mathcal{F}_3$ are equivalent to each other. The second part shows that the topology induced by the family \mathcal{F}_1 is equivalent to the topology induced by the family \mathcal{F}_0 . For the first part, I refer to the argument of Billingsley (1968, p. 237), which goes through without any change. Because of Lemma 1.1, the requirement that the sets F_i, G_i, A_i must all belong to $\mathcal{B}_0(X^u)$ plays no role in the argument.

For the second part, the requirement that the sets F_i in \mathcal{F}_1 must belong to $\mathcal{B}_0(X^u)$ does play a role. Therefore I give the adapted argument in detail.

I first show that any element of \mathcal{F}_1 contains an element of \mathcal{F}_0 as a subset. Any set $N \in \mathcal{F}_1$ takes the form (3.3) for some $\mu \in \mathcal{M}_0(X^u)$, $\varepsilon > 0$, and ρ^u -closed sets F_1, \dots, F_k in $\mathcal{B}_0(X^u)$. By Lemma 1.1, for any $\delta > 0$, the sets $F_i^\delta = \{x \in X^u \mid \rho^u(x, F_i) < \delta\}$, $i = 1, \dots, k$, also belong to $\mathcal{B}_0(X^u)$. Let $\delta > 0$ be such that, for $i = 1, \dots, k$, the set F_i^δ is a μ -continuity set and, moreover,

$$\mu(F_i^\delta) < \mu(F_i) + \frac{\varepsilon}{2}. \quad (3.6)$$

Next, let $\varphi : \mathbb{R}_+ \rightarrow [0, 1]$ be a continuous function such that, for any $t \in \mathbb{R}$,

$$\varphi(t) = 1 - t \text{ if } t \in [0, 1) \text{ and } \varphi(t) = 1 \text{ if } t \geq 1.$$

For any i , define a function $f_i : X^u \rightarrow [0, 1]$ by setting

$$f_i(x) = \varphi\left(\frac{1}{\delta} \cdot \rho^u(x, F_i)\right), \quad (3.7)$$

where again $\rho^u(x, F_i) = \min_{\hat{x} \in F_i} \rho^u(x, \hat{x})$. The function f_i is obviously bounded and ρ^u -continuous. Since $F_i \in \mathcal{B}_0(X^u)$, By Lemma 1.1, f_i is also measurable with respect to $\mathcal{B}_0(X^u)$.⁸ We also have $f_i(x) = 0$ for $x \in X^u \setminus F_i^\delta$ and $f_i(x) = 1$ for $x \in F_i$. For any $\nu \in \mathcal{M}_0(X^u)$, therefore,

$$\int_{X^u} f_i(x) d\nu(x) < \int_{X^u} f_i(x) d\mu(x) + \frac{\varepsilon}{2}$$

⁸ $f_i(x) > f_i(\hat{x})$ implies $\rho^u(x, F_i) < \rho^u(\hat{x}, F_i)$ so that, for some $\eta \in (\rho^u(x, F_i), \rho^u(\hat{x}, F_i))$, $x \in F_i^\eta$ and $\hat{x} \in X^u \setminus F_i^\eta$.

implies

$$\nu(F_i) \leq \int_{X^u} f_i(x) d\nu(x) < \int_{X^u} f_i(x) d\mu(x) + \frac{\varepsilon}{2} \leq \mu(F_i^\delta) + \frac{\varepsilon}{2} < \mu(F_i) + \varepsilon$$

For any $\nu \in \mathcal{M}_0(X^u)$ satisfying

$$\left| \int_{X^u} f_i(x) d\nu(x) - \int_{X^u} f_i(x) d\mu(x) \right| < \frac{\varepsilon}{2}$$

for $i = 1, \dots, k$, we therefore have

$$\nu(F_i) < \mu(F_i) + \varepsilon$$

for $i = 1, \dots, k$. Thus the set $N \in \mathcal{F}_1$ of measures $\nu \in \mathcal{M}_0(X^u)$ that satisfy (3.3) for the given $\mu \in \mathcal{M}_0(X^u)$, $\varepsilon > 0$, and F_1, \dots, F_k contains a set measures $\nu \in \mathcal{M}_0(X^u)$ that satisfy (3.2) for the same $\mu \in \mathcal{M}_0(X^u)$, $\varepsilon > 0$, and $\mu \in \mathcal{M}_0(X^u)$, $\varepsilon > 0$, and the specified functions $f_i, i = 1, \dots, k$. The latter set an element of \mathcal{F}_0 . Thus every element of \mathcal{F}_1 contains an element of \mathcal{F}_0 .

For the claim that every element of \mathcal{F}_0 also contains an element of \mathcal{F}_1 , the argument in Billingsley (1968) applies with hardly any change. For example, let $N \in \mathcal{F}_0$ be the set of measures $\nu \in \mathcal{M}_0(X^u)$ such that, for given $\mu \in \mathcal{M}_0(X^u)$, $\varepsilon > 0$, and $f \in \mathcal{C}_0(X^u)$,

$$\left| \int_{X^u} f(x) d\nu(x) - \int_{X^u} f(x) d\mu(x) \right| < \varepsilon. \quad (3.8)$$

Without loss of generality, assume that f takes values in the unit interval. Choose k' so that $\frac{1}{k'} < \varepsilon$ and, for $i = 1, \dots, k'$, let $F_i = \{x \in X^u \mid \frac{i}{k'} \leq f(x)\}$. Then, for any i , F_i is ρ^u -closed. Moreover, since f is $\mathcal{B}_0(X^u)$ -measurable, F_i is in $\mathcal{B}_0(X^u)$. By standard arguments, therefore, $\nu(F_i) < \mu(F_i) + \varepsilon$ implies

$$\int_{X^u} f(x) d\nu(x) < \frac{1}{k'} + \frac{1}{k'} \sum_{i=1}^{k'} \nu(F_i) < \frac{1}{k'} + \frac{1}{k'} \sum_{i=1}^{k'} \mu(F_i) + \varepsilon < \int_{X^u} f(x) d\mu(x) + 2\varepsilon.$$

By a parallel argument for the function $x \rightarrow 1 - f(x)$, there also exist ρ^u -closed, $\mathcal{B}_0(X^u)$ -measurable sets $F_i, i = k' + 1, \dots, 2k'$, such that, for $i = k' + 1, \dots, 2k'$, $\nu(F_i) < \mu(F_i) + \varepsilon$ implies

$$\int_{X^u} (1 - f(x)) d\nu(x) < \int_{X^u} (1 - f(x)) d\mu(x) + 2\varepsilon,$$

or, equivalently,

$$\int_{X^u} f(x) d\mu(x) < \int_{X^u} f(x) d\nu(x) + 2\varepsilon.$$

Upon combining these arguments, one finds that the set N of measures $\nu \in \mathcal{M}_0(X^u)$ that satisfy (3.8) for some given $\mu \in \mathcal{M}_0(X^u)$, $\varepsilon > 0$, and $f \in \mathcal{C}_0(X^u)$, an element of \mathcal{F}_0 , contains a subset consisting of measures that satisfy $\nu(F_i) < \mu(F_i) + \varepsilon$ for the given $\mu, \varepsilon, k = 2k'$, and ρ^u -closed, $\mathcal{B}_0(X^u)$ -measurable sets $F_i, i = 1, \dots, k$. Thus N has a subset that belongs to \mathcal{F}_1 .

By taking intersections of sets like N , with different functions f , the argument can be generalized to all sets in \mathcal{F}_0 . Thus every element of \mathcal{F}_0 contains an element of \mathcal{F}_1 as a subset. This completes the proof of the claim that the topologies induced by \mathcal{F}_0 and \mathcal{F}_1 are equivalent. ■

The following result provides an analogue of Billingsley's (1968) Theorem 5, p.238.

Proposition 3.2 *For any measure $\mu \in \mathcal{M}_0(X^u)$ that is quasi-separable, the topology induced by the Prohorov metric and the topology of weak convergence on $\mathcal{M}_0(X^u)$ are equivalent at μ .*

Proof. Let $N(\mu)$ be any \mathcal{F}_1 -neighbourhood of μ , characterized by $\varepsilon > 0$ and ρ^u -closed, $\mathcal{B}_0(X^u)$ -measurable sets $F_i, i = 1, \dots, k$, such that $\nu \in N(\mu)$ if and only if ν satisfies (3.3) for μ, ε , and F_1, \dots, F_k . Let $\delta < \varepsilon$ be such that, for $i = 1, \dots, k$, the set $F_i^\delta = \{x \in X^u | \rho^u(x, F_i) < \delta\}$ is a μ -continuity set and, moreover, $\mu(F_i^\delta) < \mu(F_i) + \frac{\varepsilon}{2}$. For any $\nu \in \mathcal{M}_0(X^u)$ such that $p(\nu, \mu) < \delta$, one has

$$\nu(F_i) < \mu(F_i^\delta) + \delta < \mu(F_i) + \varepsilon,$$

so the p -open ball with radius δ around μ is a subset of $N(\mu)$. Therefore the topology on $\mathcal{M}_0(X^u)$ that is induced by the Prohorov metric is at least as fine at μ as the topology of weak convergence.

Next, suppose that μ is quasi-separable. I will show that for any $\varepsilon > 0$, the p -open ball with radius ε around μ contains an \mathcal{F}_3 -neighbourhood $N(\mu)$ of μ . Fix a cover of X^u by ρ^u -open, μ -continuity balls with diameters less than δ , where $\delta < \frac{\varepsilon}{3}$. Appealing to quasi-separability, pass to a countable subfamily $\{B_i\}_{i=1}^\infty$ such that $\mu(\cup_i B_i) = 1$. Construct disjoint μ -continuity sets A_1, A_2, \dots by setting $A_1 = B_1$ and, for $i > 1$, $A_i = B_i \setminus \cup_{j < i} A_j$. Choose k so that

$$\mu(\cup_{i=1}^k A_i) > 1 - \delta \tag{3.9}$$

and let \mathcal{A} be the set of unions of the sets A_i over subsets of the indices $i = 1, \dots, k$. Then each $A \in \mathcal{A}$ is a μ -continuity set, and, by Proposition 3.1, there is a neighbourhood $N(\mu) \in \mathcal{F}_3$ of μ such that, for any $\nu \in N(\mu)$,

$$|\nu(A) - \mu(A)| < \delta \text{ for all } A \in \mathcal{A}. \tag{3.10}$$

I claim that $N(\mu)$ is contained in the p -open ball with radius ε around μ . To prove this claim, consider any $B \in \mathcal{B}_0(X^u)$. Let I_B be the set of indices i such that $A_i \cap B \neq \emptyset$ and let $A_B := \cup_{i \in I_B} A_i$. Then $B \subset A_B \cup (X^u \setminus \cup_{i=1}^k A_i)$. Moreover, $A_B \in \mathcal{A}$ and, because the sets A_i all have diameters less than δ , $A_B \subset B^\delta$. Using (3.9) and (3.10), one obtains

$$\mu(B) \leq \mu(A_B) + 1 - \mu(\cup_{i=1}^k A_i) < \nu(A_B) + 2\delta \leq \nu(B^\delta) + 2\delta < \nu(B^\varepsilon) + \varepsilon.$$

Similarly, taking account of the fact that (3.9) and (3.10) imply

$$\nu(\cup_{i=1}^k A_i) > 1 - 2\delta,$$

one also obtains

$$\nu(B) \leq \nu(A_B) + 1 - \nu(\cup_{i=1}^k A_i) < \mu(A_B) + 3\delta \leq \mu(B^\delta) + 3\delta < \mu(B^\varepsilon) + \varepsilon.$$

Thus, for any $\nu \in N(\mu)$, the Prohorov distance between ν and μ is less than ε . The specified set $N(\mu) \in \mathcal{F}_3$ is contained in the p -open ball with radius ε around μ . At μ , therefore the topology of weak convergence on $\mathcal{M}_0(X^u)$ is at least as fine as the topology that is induced by the Prohorov metric. Given that the topology induced by the Prohorov metric is also at least as fine at μ as the topology of weak convergence, it follows that the two topologies are equivalent at μ . ■

Proposition 1.4 is a straightforward corollary to Proposition 3.2.

4 Why Are $\mathcal{B}_0(X^u)$ and $\mathcal{M}_0(X^u)$ of Interest?

As mentioned in the introduction, I came across the σ -algebra $\mathcal{B}_0(X^u)$ and the space $\mathcal{M}_0(X^u)$ in the context of games of incomplete information. One approach to the analysis of strategic behaviour under incomplete information involves *hierarchies of beliefs of different orders*, another approach involves abstract "types" together with belief mappings assigning to each "type" a probability measure over states of nature and other agents' abstract "types". Mertens and Zamir (1985) reconciled the two approaches by showing that (i) belief hierarchies can be interpreted as "types", (ii) each belief hierarchy generates a unique probability measure over states of nature and other agents' belief hierarchies, and (iii) each abstract "type" space can be mapped into a belief-closed subspace of the space of belief hierarchies. In this formalism, belief hierarchies of an agent are elements of a product of spaces of beliefs of different orders. If this product space has the product topology, the

mapping involved in (ii) is a homeomorphism and, for each nonredundant abstract type space, the mapping in (iii) is an embedding, i.e. a homeomorphism between the abstract type space and a subspace of the space of belief hierarchies.

Dekel et al. (2006) and Chen et al. (2010, 2017) have criticized the product topology for the space of belief hierarchies because it too coarse to capture all the continuity properties of strategic behaviour that one may be interested in.⁹ They have therefore proposed finer topologies. In particular, Chen et al. (2010, 2017) have proposed the topology that is obtained from the uniform metric ρ^u in (1.2) above when the different factors of the product, i.e., the beliefs of different orders, are endowed with the Prohorov metric. Dekel et al. (2006) specify finer topologies directly in terms of the desired continuity properties of strategic behaviour, but Chen et al. (2010, 2017) show that these topologies can be reinterpreted in terms of the topology they propose. In particular, their topology is equivalent to the uniform strategic topology of Dekel et al. (2006).

What then becomes of the Mertens-Zamir reconciliation of an approach based on belief hierarchies with the Harsanyi abstract-type-space approach? In particular, what becomes of the proposition that there is a homeomorphism linking the space of belief hierarchies and the space of probability measure over states of nature and other agents' belief hierarchies? This is not just a question about the impact of having a finer topology for the domain of the mapping. It is also a question about the impact of having a finer topology for the range.

If other agents' belief hierarchies are given a uniform metric, the Borel σ -algebra on the space of states of nature and other agents's belief hierarchies is much finer than with the product topology. This creates difficulties because the mapping introduced by Mertens and Zamir (1985) only yields measures that are defined on the Borel σ -algebra for the product topology. These measures cannot generally be extended to the Borel σ -algebra for the topology associated with the uniform metric and, if they can be so extended, the extension need not be unique.

In Hellwig (2016/2022), I therefore follow Dudley and work with the σ -algebra generated by the ρ^u -open balls rather than all ρ^u -open sets. With this modification, I show that the homeomorphism and embedding theorems of Mertens and Zamir (1985) remain intact when the product topology on

⁹In particular, the product topology on belief hierarchies is incompatible with the observation that beliefs of arbitrarily high orders may have a significant impact on strategic behaviour, as indicated by Rubinstein's (1989) electronic mail game.

the spaced of belief hierarchies is replaced by the uniform topology. The argument makes essential use of the equivalence of the topology of weak convergence and the topology induced by the Prohorov metric on $\mathcal{M}_0(X^u)$.

Without going into the game-theoretic analysis, I briefly sketch the mathematical argument. For $\ell = 1, 2, \dots$, let π^ℓ be the projection from X to the finite product

$$X^\ell = \prod_{k=1}^{\ell} X_k. \quad (4.1)$$

For any $\mu \in \mathcal{M}_0(X^u)$ and any ℓ , let

$$\Pi^\ell(\mu) = \mu \circ (\pi^\ell)^{-1} \in \mathcal{M}(X^\ell), \quad (4.2)$$

and let

$$\Pi^\infty(\mu) = (\Pi^1(\mu), \Pi^2(\mu), \dots). \quad (4.3)$$

Proposition 4.1 *Suppose that each of the spaces $\mathcal{M}(X^\ell)$, $\ell = 1, 2, \dots$, is endowed with the Prohorov metric p_ℓ , and that the product space $\prod_{\ell=1}^{\infty} \mathcal{M}(X^\ell)$ is endowed with the uniform metric d^u such that, for any two sequences $\{\mu^\ell\}_{\ell=1}^{\infty}, \{\hat{\mu}^\ell\}_{\ell=1}^{\infty}$,*

$$d^u(\{\mu^\ell\}_{\ell=1}^{\infty}, \{\hat{\mu}^\ell\}_{\ell=1}^{\infty}) = \sup_{\ell} d_\ell(\mu^\ell, \hat{\mu}^\ell). \quad (4.4)$$

If $\mathcal{M}_0(X^u)$ has the topology of weak convergence, then, under Assumption 1, the mapping Π^∞ is a homeomorphism between $\mathcal{M}_0(X^u)$ and the subspace $H^u \subset \prod_{\ell=1}^{\infty} \mathcal{M}(X^\ell)$ that consists of those sequences $\{\mu^\ell\}_{\ell=1}^{\infty}$ that are mutually consistent in that $\mu^{\ell+1} \circ (\pi^\ell)^{-1} = \mu^\ell$ for all ℓ .

Proof. The mapping Π^∞ from $\mathcal{M}_0(X^u)$ to $H^u \subset \prod_{\ell=1}^{\infty} \mathcal{M}(X^\ell)$ is obviously injective and onto. By Kolmogorov's extension theorem, the inverse $(\Pi^\infty)^{-1}$ from H^u to $\mathcal{M}_0(X^u)$ is also injective and onto. Therefore it suffices to show that Π^∞ and $(\Pi^\infty)^{-1}$ are both continuous.

Continuity of Π^∞ is straightforward: For any ℓ and any set $W^\ell \in \mathcal{B}(X^\ell)$, the cylinder set

$$\hat{W}^\ell = W^\ell \times X_{\ell+1} \times X_{\ell+2} \times \dots$$

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

$$\hat{W}^{\ell\varepsilon} = (W^\ell)^\varepsilon \times X_{\ell+1} \times X_{\ell+2} \times \dots$$

that is defined by the ε -neighbourhood $(W^\ell)^\varepsilon$ of W^ℓ in X^ℓ is an ε -neighbourhood of \hat{W}^ℓ in X^u . If the Prohorov distance between two measures $\mu, \hat{\mu}$ in $\mathcal{M}_0(X^u)$ is less than ε , we have

$$\mu(\hat{W}^\ell) < \hat{\mu}(\hat{W}^{\ell\varepsilon}) + \varepsilon$$

and

$$\hat{\mu}(\hat{W}^\ell) < \mu(\hat{W}^{\ell\varepsilon}) + \varepsilon.$$

By the definition of the marginal distributions $\mu^\ell = \Pi^\ell(\mu)$ and $\hat{\mu}^\ell = \Pi^\ell(\hat{\mu})$, it follows that

$$\mu^\ell(W^\ell) < \hat{\mu}^\ell(W^{\ell\varepsilon}) + \varepsilon$$

and

$$\hat{\mu}^\ell(W^\ell) < \mu^\ell(W^{\ell\varepsilon}) + \varepsilon.$$

Because the choice of $W^\ell \in \mathcal{B}(X^\ell)$ was arbitrary, it follows that the Prohorov distance $d_\ell(\mu^\ell, \hat{\mu}^\ell)$ between $\mu^\ell = \Pi^\ell(\mu)$ and $\hat{\mu}^\ell = \Pi^\ell(\hat{\mu})$ in $\Pi^\ell(\mathcal{M}_0(X^u))$ is no greater than ε . Since ε may be taken to be arbitrarily close to the Prohorov distance $p(\mu, \hat{\mu})$ between μ and $\hat{\mu}$ in $\mathcal{M}_0(X^u)$, it follows that the Prohorov distance between $\mu^\ell = \Pi^\ell(\mu)$ and $\hat{\mu}^\ell = \Pi^\ell(\hat{\mu})$ in $\Pi^\ell(\mathcal{M}_0(X^u))$ is no greater than $p(\mu, \hat{\mu})$. Since this is true for all ℓ , it follows that

$$d^u(\Pi^\infty(\mu), \Pi^\infty(\hat{\mu})) = \sup_\ell d_\ell(\Pi^\ell(\mu), \Pi^\ell(\hat{\mu})) \leq p(\mu, \hat{\mu}).$$

Continuity of the map Π^∞ from $\mathcal{M}_0(X^u)$ to H^u follows immediately.

Next, consider the map $\beta := (\Pi^\infty)^{-1}$ from H^u to $\mathcal{M}_0(X^u)$ that is given by Kolmogorov's extension theorem. Proceeding indirectly, suppose that β is not continuous. Then there exist sequences $h^r = \{\mu^{\ell r}\}_{\ell=1}^\infty$, $r = 1, \dots, \infty$, and $h = \{\mu^\ell\}_{\ell=1}^\infty$ in H^u such that h^r converges to $h \in H^u$ but $\beta(h^r)$ does not converge to $\beta(h)$ in $\mathcal{M}_0(X^u)$. Convergence of h^r to h implies that

$$\lim_{r \rightarrow \infty} \sup_\ell d_\ell(\mu^{\ell r}, \hat{\mu}^\ell) = 0. \quad (4.5)$$

Non-convergence of $\beta(h^r)$ to $\beta(h)$ implies that, for some $\varepsilon > 0$ and some subsequence $\{h^{r'}\}$ of $\{h^r\}$,

$$p(\beta(h^{r'}), \beta(h)) \geq \varepsilon$$

for all r' . Thus, for every r' , there exists a set $B_{r'} \in \mathcal{B}_0(X^u)$ such that

$$\beta(B_{r'}|h^r) > \beta(B_{r'}^\varepsilon|h) + \varepsilon \quad (4.6)$$

or

$$\beta(B_{r'}|h) > \beta(B_{r'}^\varepsilon|h^r) + \varepsilon, \quad (4.7)$$

where $B_{r'}^\varepsilon \in \mathcal{B}_0(X^u)$ is the ε -neighbourhood of $B_{r'}$ in X^u .

For any ℓ , let $B_{r'}^\ell = \pi^\ell(B_{r'})$ be the projection of $B_{r'}$ to $X^\ell = \pi^\ell(X^u)$ and let $(B_{r'}^\ell)^\varepsilon$ be an ε -neighbourhood of $B_{r'}^\ell$ in X^ℓ . Let

$$\hat{B}_{r'}^\ell := B_{r'}^\ell \times X_{\ell+1} \times X_{\ell+2} \times \dots$$

and

$$(\hat{B}_{r'}^\ell)^\varepsilon := (B_{r'}^\ell)^\varepsilon \times X_{\ell+1} \times X_{\ell+2} \times \dots$$

be the cylinder sets in X^u that are defined by $B_{r'}^\ell$ and $(B_{r'}^\ell)^\varepsilon$. One easily verifies that the sequences $\{\hat{B}_{r'}^\ell\}$ and $\{(\hat{B}_{r'}^\ell)^\varepsilon\}$ are nonincreasing and that

$$B_{r'} = \bigcap_{\ell=1}^{\infty} \hat{B}_{r'}^\ell \quad \text{and} \quad B_{r'}^\varepsilon = \bigcap_{\ell=1}^{\infty} (\hat{B}_{r'}^\ell)^\varepsilon \quad (4.8)$$

for all r' . By elementary measure theory,¹⁰ for any r' and any $\delta > 0$, there exists an integer $L^{r'}(\delta)$ such that, for $\ell > L^{r'}(\delta)$,

$$\beta(B_{r'}^\varepsilon|h) \geq \beta((\hat{B}_{r'}^\ell)^\varepsilon|h) - \delta \quad (4.9)$$

and

$$\beta(B_{r'}^\varepsilon|h^{r'}) \geq \beta((\hat{B}_{r'}^\ell)^\varepsilon|h^{r'}) - \delta. \quad (4.10)$$

Moreover, by (4.8),

$$\beta(B_{r'}|h) \leq \beta(\hat{B}_{r'}^\ell|h) \quad (4.11)$$

and

$$\beta(B_{r'}|h^{r'}) \leq \beta(\hat{B}_{r'}^\ell|h^{r'}). \quad (4.12)$$

Set $\delta = \frac{\varepsilon}{2}$ and combine (4.9) - (4.12) with (4.6) and (4.7). Thereby one finds that, for all r' , all $\delta > 0$, and all $\ell > L^{r'}(\delta)$, either

$$\beta(\hat{B}_{r'}^\ell|h^{r'}) > \beta((\hat{B}_{r'}^\ell)^\varepsilon|h) + \frac{\varepsilon}{2} \quad (4.13)$$

or

$$\beta(\hat{B}_{r'}^\ell|h) > \beta((\hat{B}_{r'}^\ell)^\varepsilon|h^{r'}) + \frac{\varepsilon}{2}. \quad (4.14)$$

¹⁰Theorem 3.1.1, p. 86, in Dudley (2002).

Since $(\hat{B}_{r'}^\ell)^{\frac{\varepsilon}{2}} \subset (\hat{B}_{r'}^\ell)^\varepsilon$, it follows that, for all r' , all $\delta > 0$, and all $\ell > L^{r'}(\delta)$, either

$$\beta(\hat{B}_{r'}^\ell | h^{r'}) > \beta((\hat{B}_{r'}^\ell)^{\frac{\varepsilon}{2}} | h) + \frac{\varepsilon}{2} \quad (4.15)$$

or

$$\beta(\hat{B}_{r'}^\ell | h) > \beta((\hat{B}_{r'}^\ell)^{\frac{\varepsilon}{2}} | h^{r'}) + \frac{\varepsilon}{2}. \quad (4.16)$$

By the definition of β as the inverse of $\Pi^\infty = (\Pi^1, \Pi^2, \dots)$, and the cylinder nature of the sets $\hat{B}_{r'}$ and $(\hat{B}_{r'}^\ell)^{\frac{\varepsilon}{2}}$, we also have

$$\beta(\hat{B}_{r'}^\ell | h^{r'}) = \Pi^\ell(B_{r'}^\ell | \beta(h^{r'})) = \mu^{\ell r'}(B_{r'}^\ell), \quad (4.17)$$

$$\beta(\hat{B}_{r'}^\ell | h) = \Pi^\ell(B_{r'}^\ell | \beta(h)) = \mu^\ell(B_{r'}^\ell), \quad (4.18)$$

$$\beta((\hat{B}_{r'}^\ell)^{\frac{\varepsilon}{2}} | h^{r'}) = \Pi^\ell((B_{r'}^\ell)^{\frac{\varepsilon}{2}} | \beta(h^{r'})) = \mu^{\ell r'}((B_{r'}^\ell)^{\frac{\varepsilon}{2}}), \quad (4.19)$$

$$\beta((\hat{B}_{r'}^\ell)^{\frac{\varepsilon}{2}} | h) = \Pi^\ell((B_{r'}^\ell)^{\frac{\varepsilon}{2}} | \beta(h)) = \mu^\ell((B_{r'}^\ell)^{\frac{\varepsilon}{2}}). \quad (4.20)$$

Thus, (4.15) and ??) can be rewritten as

$$\mu^{\ell r'}(B_{r'}^\ell) > \mu^\ell((B_{r'}^\ell)^{\frac{\varepsilon}{2}}) + \frac{\varepsilon}{2} \quad (4.21)$$

and

$$\mu^\ell(B_{r'}^\ell) > \mu^{\ell r'}((B_{r'}^\ell)^{\frac{\varepsilon}{2}}) + \frac{\varepsilon}{2}, \quad (4.22)$$

and one of (4.21), (4.22) must hold if r' , $\delta = \frac{\varepsilon}{2}$, and $\ell > L^{r'}(\delta)$. But then, for such r' , δ , and ℓ , $d_\ell(\mu^{\ell r'}, \hat{\mu}^\ell) \geq \frac{\varepsilon}{2}$, contrary to (4.5). The assumption that $\beta = (\Pi^\infty)^{-1}$ is not continuous has thus led to a contradiction and must be false. ■

References

- [1] Banach, S. (1930), Über additive Maßfunktionen in abstrakten Mengen, *Fundamenta Mathematicae* 15, 97-101.
- [2] Banach, S., and C. Kuratowski (1929), Sur une généralisation du problème de la mesure, *Fundamenta Mathematicae* 14, 127-131.
- [3] 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.

- [4] Billingsley, P. (1968), *Convergence of Probability Measures*, Wiley, New York.
- [5] Chen, Y., A. Di Tillio, E. Faingold, and S. Xiong (2010), Uniform Topologies on Types, *Theoretical Economics* 5, 445 -478.
- [6] 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.
- [7] Cohen, P.J. (1966), *Set Theory and the Continuum Hypothesis*, W.A. Benjamin, Reading, MA.
- [8] Dekel, E., D. Fudenberg, and S. Morris (2006), Topologies on Types, *Theoretical Economics* 1, 275 - 309.
- [9] Dudley, R.M. (1966), Weak Convergence of Probability Measures on Nonseparable Metric Spaces and Empirical Measures on Euclidean Spaces, *Illinois Journal of Mathematics* 10, 109-126.
- [10] Dudley, R.M. (1967), Measures on Non-Separable Metric Spaces, *Illinois Journal of Mathematics* 11, 449-453.
- [11] Dudley, R.M. (1978), Central Limit Theorems for Empirical Measures, *Annals of Probability* 6, 899-929.
- [12] Dudley, R.M. (2002), *Real Analysis and Probability*, Cambridge Studies in Advanced Mathematics 74, Cambridge University Press, Cambridge, UK.
- [13] Harsanyi, J.C. (1967-68), Games with incomplete information played by 'Bayesian' players, I-III. *Management Science* 14, 159-182, 320-334, 486-502.
- [14] Hellwig, M.F. (2016/22), A Homeomorphism Theorem for the Universal Type Space with the Uniform Weak Topology, Preprint 17/2016, Max Planck Institute for Research on Collective Goods, Bonn, revised July 2022. http://www.coll.mpg.de/pdf_dat/2016_17online.pdf.
- [15] Kelley, J.L. (1955), *General Topology*, Van Nostrand, New York.
- [16] Marczewski, E., and R. Sikorski (1948), Measures in Non-Separable Metric Spaces, *Colloquium Mathematicum* 1, 133-139.

- [17] Mertens, J.F., and S. Zamir (1985), Formulation of Bayesian Analysis for Games with Incomplete Information, *International Journal of Game Theory* 14, 1-29.
- [18] Pollard, D. (1979), Weak Convergence on Non-Separable Metric Spaces, *Journal of the Australian Mathematical Society (Series A)* 28, 197-204.
- [19] Rubinstein, A. (1989), The Electronic Mail Game: Strategic Behavior under 'Almost Common Knowledge', *American Economic Review* 79, 385-391.