Preprints of the
Max Planck Institute for
Research on Collective Goods
Bonn 2017/6

Probability Measures on Product Spaces with Uniform Metrics

Martin F. Hellwig

MAX PLANCK SOCIETY

Preprints of the
Max Planck Institute
for Research on Collective Goods

Bonn 2017/6

Probability Measures on Product Spaces with Uniform Metrics

Martin F. Hellwig

May 2017

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

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

May 3, 2017

Abstract

For a countable product of complete separable metric spaces with a topology induced by a uniform metric, the set of Borel probability measures coincides with the set of completions of probability measures on the product σ-algebra. Whereas the product space with the uniform metric is non-separable, the support of any Bofrel measure is separable, and the topology of weak convergence on the space of Borel measures is metrizable by both the Prohorov metric and the bounded Lipschitz metric.

Key Words: Borel measures, product spaces with uniform metrics, completions of product σ-algebras, universal type space, separability of supports, metrizability of weak convergence.

JEL Classification: C02, C72

1 Introduction

Let $X_1, X_2, ...$ be complete separable metric spaces with metrics $\rho_1, \rho_2, ...$. Suppose that the product

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

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

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

$$\rho^u(x, \hat{x}) = \sup_k \rho_k(\pi_k(x), \pi_k(\hat{x})), \tag{2}$$

and π_k is the projection from X to X_k. This topology is obviously finer than the usual product topology. Indeed, as noted by Chen et al. (2010, 2012, 2016), with the uniform topology, X is not separable,[1] and the Borel σ-algebra $\mathcal{B}^u(X)$ that is induced by the uniform topology is strictly larger than the Borel σ-algebra $\mathcal{B}^p(X)$ that is induced by the product topology.[2]

Even so, the space $\mathcal{M}^u(X)$ of probability measures on $(X, \mathcal{B}^u(X))$ is not dissimilar to the space $\mathcal{M}^p(X)$ of probability measures on $(X, \mathcal{B}^p(X))$. The following proposition shows that the sets in $\mathcal{B}^u(X)$ are *universally measurable* in the completions of $\mathcal{B}^p(X)$.

Proposition 1 *Let $\mathcal{U}^p(X)$ be the class of sets that are measurable by the completion of every probability measure μ on $(X, \mathcal{B}^p(X))$. Then*

$$\mathcal{B}^p(X) \subsetneq \mathcal{B}^u(X) \subset \mathcal{U}^p(X). \tag{3}$$

Corollary 2 *For every probability measure $\mu \in \mathcal{M}^p(X)$, there exists a unique measure $\bar{\mu} \in \mathcal{M}^u(X)$ that agrees with μ on $\mathcal{B}^p(X)$. The set $\mathcal{M}^u(X)$ of probability measures on $(X,\ \mathcal{B}^u(X))$ is equal to the set of completions of probability measures $\mu \in \mathcal{M}^p(X)$.*[3]

The topology of weak convergence on $\mathcal{M}^u(X)$ is defined in terms of convergence of integrals of bounded continuous functions from X to $\mathbb{R}$ where continuity refers to the uniform topology on X. Metrizability of this topology depends on whether the elements of $\mathcal{M}^u(X)$ have separable supports. If X was an arbitrary non-separable metric space, separability of the supports

[1] Fix some $\Delta > 0$. For $k = 1, 2, ...$, let x_k^1, x_k^2 be two elements of X_k for which $\rho_k(x_k^1, x_k^2) \geq \Delta$. Let $\hat{X} \subset X$ be the set of sequences $\hat{x} = (\hat{x}_1, \hat{x}_2, ...)$ such that $\hat{x}_k \in \{x_k^1, x_k^2\}$ for all k. For any $\hat{x} \in \hat{X}$, let $B^{\Delta/2}(\hat{x})$ be the $\frac{\Delta}{2}$-sphere around $\hat{x}$. Then $\{B^{\Delta/2}(\hat{x}) | \hat{x} \in \hat{X}\}$ is an uncountable class of nonintersecting open sets.

[2] When endowed with the product topology, the set $\hat{X}$ in the preceding footnote is a complete separable metric space. As noted by Chen et al. (2016), Proposition 13.2.5, p. 496, in Dudley (2002) implies that this space contains an analytic set A that does not belong to $\mathcal{B}^p(X)$. A is however equal to the intersection of its $\frac{1}{n}$-neighbourhoods in the uniform topology and is therefore an element of $\mathcal{B}^u(X)$.

[3] Glossing over the fact that the domain of the completion of a measure $\mu \in \mathcal{M}^p(X)$ is typically larger than $\mathcal{B}^u(X)$, I also use the term completion for the extension of μ to $\mathcal{B}^u(X)$.

of measures on X would require that the cardinal of X be non-measurable, which in turn is true if the continuum hypothesis is assumed.[4] The following result shows that, in the present context, where X is a product space with a uniform metric, the supports of Borel measures on X are always separable, without any presumption about the cardinal of X, let alone the continuum hypothesis.

Proposition 3 *Every probability measure in $\mathcal{M}^u(X)$ has a separable support.*

The Prohorov metric ρ^P and the bounded Lipschitz metric ρ^{BL} on $\mathcal{M}^u(X)$ are defined so that, for any μ and $\hat{\mu}$ in $\mathcal{M}^u(X)$, $\rho^P(\mu, \hat{\mu})$ is the infimum of the set of ε such that

$$\mu(B) \leq \hat{\mu}(B^\varepsilon) + \varepsilon \text{ and } \hat{\mu}(B) \leq \mu(B^\varepsilon) + \varepsilon \tag{4}$$

for all sets $B \in \mathcal{B}^u(X)$ with ε-neighbourhoods $B^\varepsilon \in \mathcal{B}^u(X)$, and $\rho^{BL}(\mu, \hat{\mu})$ is the supremum of the expression

$$\left| \int_X f(x)\mu(dx) - \int_X f(x)\hat{\mu}(dx) \right| \tag{5}$$

over the set of real-valued functions f on X for which

$$\sup_{x \in X} |f(x)| + \sup_{\substack{x \in X \\ \hat{x} \in X}} \frac{|f(x) - f(\hat{x})|}{\rho^u(x, \hat{x})} \leq 1. \tag{6}$$

Corollary 4 *The topology of weak convergence on $\mathcal{M}^u(X)$ is metrizable by the Prohorov metric and by the bounded Lipschitz metric.*

2 Motivation: The Universal Type Space in Game Theory

The motivation for the analysis comes from game theory. In a game of incomplete information, the so-called *universal type space* has a product structure and, for some purposes, a uniform topology is more appropriate

[4] See Theorem 2, p. 235, and the subsequent discussion in Billingsley (1968).

than the product topology. The product topology is too coarse for certain desirable continuity properties of behaviour correspondences.

For simplicity, consider incomplete information in games involving two agents. In such a game, the *universal type space* of agent i is a space of sequences

$$x^i = (x^i_1, x^i_2, ...) \tag{7}$$

with the interpretation that x^i_1 is a probability measure indicating agent i's beliefs about the state of nature (agent i's first-order beliefs), x^i_2 is a probability measure indicating agent i's beliefs about the state of nature and about the other agent's first-order beliefs (agent i's second-order beliefs), x^i_3 is a probability measure indicating agent i's beliefs about the state of nature and about the other agent's first-order and second-order beliefs (agent i's third-order beliefs), and so on, with a consistency condition ensuring that the beliefs of different orders are mutually compatible. Denoting the space of possible states of nature by Θ, one may write $x^i_k \in X_k$ for $k = 1, 2, ...$, where

$$X_1 := \mathcal{M}(\Theta), X_2 := \mathcal{M}(\Theta \times X_1), X_3 := \mathcal{M}(\Theta \times X_1 \times X_2), ...,$$

If Θ is a complete separable metric space and every one of the spaces $X_1, X_2, ...$ is given the topology of weak convergence of probability measures, $X_1, X_2, ...$ are all complete separable metric spaces. The universal type space is the set of those sequences in $X = \prod_k X_k$ for which the beliefs of different orders are mutually compatible. [5]

Underlying this construction is the idea that the consistent belief hierarchies of an agent provide a complete representation of all strategically relevant aspects of an agent's information, and that strategies in any given game can be treated as functions of belief hierarchies. The continuity properties of agents' strategies depend on the topology on X.

The first generation of papers on the universal type space relied on the product topology. With this topology, in the above two-agent example, Kolmogorov's extension theorem induces a mapping from consistent belief hierarchies of agent 1 to probability distributions over the space $\Theta \times X$ of states of nature and consistent belief hierarchies of agent 2. In any strategic game with payoffs depending on the state of nature and on the participants' actions, for a given strategy of agent 2 that is anticipated by agent 1, the probability distribution over states of nature and consistent belief

[5] See Harsanyi (1967/68), Mertens and Zamir (1985), Brandenburger and Dekel (1993).

hierarchies of agent 2 that is provided by agent 1's own belief hierarchy and Kolmogorov's extension theorem provides agent 1 with a basis for assessing expected payoffs from alternative action plans. This assessment provides a basis for determining best responses or approximately best responses to the other agent's strategy. The mapping from anticipated strategies of other agents to best-response or approximately-best-reponse strategies in turn provides a basis for characterizing vectors of strategies for the different participants that satisfy whatever conditions of mutual consistence one may want to impose (equilibrium, rationalizability, approximate reationalizability).

Reliance on the product topology has however been criticized as being too coarse for certain desirable continuity properties of behaviour correspondences.[6] Specifically, the set of strictly ε-rationalizable actions can be sensitive to changes in beliefs of arbitrarily high orders so that an action that is ε-rationalizable at one type cannot be approximated by actions that are (close to) ε-rationalizable for nearby types in the product topology.[7]

Dekel et al. (2006) and Chen et al. (2010, 2012) have therefore proposed finer topologies. Chen et al. (2010, 2012) propose what they call the *uniform weak topology*, which is just the topology induced by the uniform metric (2) when each of the factor spaces $X_1 = \mathcal{M}(\Theta), X_2 = \mathcal{M}(\Theta \times X_1), ...$ has the Prohorov metric for the topology of weak convergence. They show that this topology is equivalent to the *uniform strategic topologies* of Dekel et al. (2006), which is specified directly in terms of the desired continuity properties of behaviour correspondences.

For these finer topologies on the universal type space, Corollary 2 implies that, even though $\mathcal{B}(\Theta) \times \mathcal{B}^u(X)$ is larger than $\mathcal{B}(\Theta) \times \mathcal{B}^p(X)$, the consistent belief hierarchies of agent i still contain enough information to pin down a probability measure on $\mathcal{B}(\Theta) \times \mathcal{B}^u(X)$, which can be used to assess the agent's expected payoffs from alternative action plans when he takes the

[6] Rubinstein (1989), Weinstein and Yildiz (2007), Dekel et al. (2006, 2007), Chen et al. (2010, 2012, 2016).

[7] A function from types to action plans in a given strategic game, is said to be an ε-rationalizable strategy if the specified action plan of each type is ε-rationalizable. An action plan is ε-rationalizable for a given type if there exist ε-rationalizable strategies of the other agents such that, if the other agents use these strategies, the shortfall of the agent's expected payoff from the specified action plan relative to the supremum of his expected payoff over all action plans is less than ε. More precisely, avoiding the circularity in the preceding formulation, an action plan is ε-rationalizable if it is n^{th}-order ε-rationalizable for all n, i.e. if there exist $(n-1)^{st}$-order ε-rationalizable strategies of the other agents such that, if the other agents use these strategies, the shortfall of the agent's expected payoff from the specified action plan relative to the supremum over all action plans is less than ε. See Dekel et al. (2006, 2007).

other agent's strategy as given. Corollary 4 implies that the topology of weak convergence on the space of such probability measures is metrizable. In Hellwig (2016/17), I use these results to show that, even if all agents' belief hierarchies have the uniform topology, the map from consistent belief hierarchies to probability measures over states of nature and other agents' belief hierarchies that is provided by Kolmogorov's extension theorem is a homeomorphism. Continuity or genericity properties of behaviour thus do not depend on whether one considers belief hierarchies or probability measures over states of nature and other agents' belief hierarchies.[8]

3 Proofs

Lemma 5 *If* $X = \prod_{k=1}^{\infty} X_k$ *has the uniform topology that is induced by the metric* ρ^u*, then for any* k *and* K *the projections* $\pi_k : X \to X_k$ *and* $\pi^K : X \to X_1 \times ... \times X_K$ *are open and continuous.*

Proof. As is well known, the projections are open and continuous if X has the product topology or the box topology. Since the uniform topology is finer than the product topology and coarser than the box topology, the lemma follows immediately. ■

Lemma 6 *Let* μ *be a probability measure on* $(X, \mathcal{B}^p(X))$ *and let* U *be a nonempty open set in the uniform topology on* X*. Then there exist* $\bar{V}, \hat{V}$ *in* $\mathcal{B}^p(X)$ *such that*

$$\hat{V} \subset U \subset \bar{V}, \tag{8}$$

and, moreover,

$$\mu(\hat{V}) = \mu(\bar{V}). \tag{9}$$

Proof. For $K = 1, 2, ...$, consider the sets

$$\bar{V}^K := \pi^K(U) \times X_{K+1} \times X_{K+2}, ..., \tag{10}$$

[8] In Gizatulina and Hellwig (2017), the homeomorphism theorem of Hellwig (2016/17) is used to prove that, for a residual set of compact subsets of the universal type space, it is possible to design entry fee schedules for agent i that extract an arbitrarily large fraction of the surplus that the agent expects to gain from his subsequent participation in a strategic game.

where, as before, π^K is the projection from X to $X_1 \times ... \times X_K$. By Lemma 5 each of these sets is open in the product topology on X. Their intersection

$$\bar{V} := \bigcap_{K=1}^{\infty} \bar{V}^K \tag{11}$$

is therefore an element of $\mathcal{B}^p(X)$. By the definition of the sets V^K, for $x \in U$, we also have $x \in V^K$ for all K, and therefore $x \in \bar{V}$. Thus $U \subset \bar{V}$.

If $\mu(\bar{V}) = 0$, the assertion of the lemma is trivially true with $\hat{V} = \emptyset$. Suppose therefore, that $\mu(\bar{V}) > 0$ and let $\hat{\mu}(\cdot) = \frac{\mu(\cdot)}{\mu(\bar{V})}$ be the induced conditional probability distribution on X given $\bar{V}$. For any K, let $\hat{\mu}^K := \hat{\mu} \circ (\pi^K)^{-1}$ be the associated marginal distribution on $X_1 \times ... \times X_K$ and note that $\hat{\mu}^K$ is concentrated on the set $\pi^K(\bar{V}) = \pi^K(U)$. Ulam's theorem[9] implies that, for any $\varepsilon > 0$, there exists a compact set $C^K_\varepsilon \subset \pi^K(\bar{V})$ such that

$$\hat{\mu}^K(C^K_\varepsilon) \geq 1 - \varepsilon \tag{12}$$

and therefore

$$\mu(C^K_\varepsilon \times X_{K+1} \times X_{K+2} \times ...) \geq (1 - \varepsilon)\mu(\bar{V}). \tag{13}$$

For any $x \in U$, there exists $\delta(x) > 0$ such that $B^{\delta(x)}(x) \subset U$, where

$$B^{\delta(x)}(x) := \{\hat{x} \in U || \rho^u(x, \hat{x}) < \delta(x)\} \tag{14}$$

is the open $\delta(x)$-ball around x. Moreover,

$$U = \bigcup_{x \in U} B^{\delta(x)}(x) \tag{15}$$

and therefore, for any K,

$$\pi^K(U) = \pi^K \left(\bigcup_{x \in U} B^{\delta(x)}(x) \right). \tag{16}$$

By elementary set theory, projections and unions commute, so (16) in turn yields

$$\pi^K(U) = \bigcup_{x \in U} \pi^K(B^{\delta(x)}(x)), \tag{17}$$

so the collection $\{\pi^K(B^{\delta(x)}(x)) | x \in U\}$ is a covering of $\pi^K(\bar{V}) = \pi^K(U)$ and therefore also of C^K_ε. By Lemma 5, the sets $\pi^K(B^{\delta(x)}(x)), x \in U$, are

[9] See, e.g., Theorem 7.2.5, p. 225 in Dudley (2002).

open. Because C_ε^K is compact, the open covering $\{\pi^K(B^{\delta(x)}(x))|x \in U\}$ of C_ε^K has a finite subcovering $\{\pi^K(B^{\delta(x)}(x))|x \in \{x_1^K, ..., x_{n_K}^K\} \subset U\}$. The set $\cup_{K=1}^{\infty}\{x_1^K, ..., x_{n_K}^K\}$ is countable. It can therefore be written in the form $\{y^1, y^2, ...\}$ so that, for each K and each $i \in \{1, ..., n_K\}$, there exists $\ell(K,i)$ such that $y^{\ell(K,i)} = x_i^K$. For any K, the collection $\{\pi^K(B^{\delta(y^\ell)}(y^\ell))|\ell = 1, 2, ...\}$ is an open covering of C_ε^K and the set

$$\hat{V}_\varepsilon^K := (\cup_{\ell=1}^{\infty}\pi^K(B^{\delta(y^\ell)}(y^\ell))) \times X_{K+1} \times X_{K+2} \times ... \tag{18}$$

is an open covering of the set $C_\varepsilon^K \times X_{K+1} \times X_{K+2}, ...$ so (13) implies

$$\mu(\hat{V}_\varepsilon^K) \geq (1 - \varepsilon)\mu(\bar{V}). \tag{19}$$

Now consider the set

$$\hat{V}_\varepsilon = \bigcup_{\ell=1}^{\infty} B^{\delta(y^\ell)}(y^\ell)). \tag{20}$$

By construction, $B^{\delta(y^\ell)}(y^\ell)) \subset U$ for all ℓ. Therefore, $\hat{V}_\varepsilon \subset U$. Moreover, for any ℓ,

$$B^{\delta(y^\ell)}(y^\ell)) = \bigcap_{K=1}^{\infty} [\pi^K(B^{\delta(y^\ell)}(y^\ell))) \times X_{K+1} \times X_{K+2} \times ...] \tag{21}$$

is a countable intersection of open sets in the product topology on X and is therefore an element of $\mathcal{B}^p(X)$. Because $\mathcal{B}^p(X)$ is closed under countable unions, it follows that $\hat{V}_\varepsilon$ is also an element of $\mathcal{B}^p(X)$.

Note that $\hat{V}_\varepsilon$ can also be written as

$$\hat{V}_\varepsilon = \bigcap_{K=1}^{\infty} [\pi^K \left(\bigcup_{\ell=1}^{\infty} B^{\delta(y^\ell)}(y^\ell))\right) \times X_{K+1} \times X_{K+2} \times ...]. \tag{22}$$

By the continuity of measures on monotone classes of sets,[10] it follows that

$$\mu(\hat{V}_\varepsilon) = \lim_{K\to\infty} \mu\left(\pi^K \left(\bigcup_{\ell=1}^{\infty} B^{\delta(y^\ell)}(y^\ell))\right) \times X_{K+1} \times X_{K+2} \times ...\right). \tag{23}$$

Hence, by the commutativity of unions and projections,

$$\begin{aligned}\mu(\hat{V}_\varepsilon) &= \lim_{K\to\infty} \mu\left(\bigcup_{\ell=1}^{\infty} \pi^K(B^{\delta(y^\ell)}(y^\ell))) \times X_{K+1} \times X_{K+2} \times ...\right) \\ &= \lim_{K\to\infty} \mu(\hat{V}_\varepsilon^K) \end{aligned} \tag{24}$$

[10] See, e.g. Theorem E, p. 38, in Halmos (1950),.

and, therefore, by (20), that

$$\mu(\hat{V}_\varepsilon) \geq (1-\varepsilon)\mu(\bar{V}). \tag{25}$$

To conclude the argument, note that the countable union

$$\hat{V} = \bigcup_{n=2}^{\infty} \hat{V}_{1/n} \tag{26}$$

also belongs to $\mathcal{B}^p(X)$, with a measure satisfying $\mu(\hat{V}) \geq (1-\frac{1}{n})\mu(\bar{V})$ for all n and therefore $\mu(\hat{V}) \geq \mu(\bar{V})$. Since $\hat{V}_{1/n} \subset U$ for all n, $\hat{V} \subset U$ so $U \subset \bar{V}$ implies $\hat{V} \subset \bar{V}$ and therefore $\mu(\hat{V}) \leq \mu(\bar{V})$. (9) follows immediately. ■

Proof of Proposition 1. The weak inclusion $\mathcal{B}^p(X) \subset \mathcal{B}^u(X)$ is trivial. Strictness of the inclusion follows from the example of Chen et al. (2016), see fn. 2 above. It remains to be shown that any element of $\mathcal{B}^u(X)$ belongs to the set $\mathcal{U}^p(X)$ of sets that are universally measurable, i.e. measurable by the completion of any measure $\mu \in \mathcal{M}^p(X)$. As discussed in Dudley (2002, p. 402), a set W belongs to $\mathcal{U}^p(X)$ if and only if, for every measure μ on $(X, \mathcal{B}^p(X))$, there exist elements $\bar{V}, \hat{V}$ of the product σ-algebra $\mathcal{B}^p(X)$ such that

$$\hat{V} \subset W \subset \bar{V} \tag{27}$$

and

$$\mu(\hat{V}) = \mu(\bar{V}). \tag{28}$$

Thus, by Lemma 6, any nonempty set U that is open in the uniform topology on X belongs to $\mathcal{U}^p(X)$. Moreover, if a set W belongs to $\mathcal{U}^p(X)$, then, since $\mathcal{B}^p(X)$ is closed under complements, the sets $X\backslash\bar{V}$ and $X\backslash\hat{V}$ are in $\mathcal{B}^p(X)$, and, from (27) and (28), one obtains

$$X\backslash\bar{V} \subset X\backslash W \subset X\backslash\hat{V} \tag{29}$$

and

$$\mu(X\backslash\bar{V}) = 1-\mu(\bar{V}) = 1-\mu(\hat{V}) = \mu(X\backslash\hat{V}). \tag{30}$$

Thus $\mathcal{U}^p(X)$ is also closed under complements.

Finally, suppose that W_i, $i = 1, 2, ...$, are in $\mathcal{U}^p(X)$ and, for a given measure μ on $(X, \mathcal{B}^p(X))$, let $\bar{V}_i, \hat{V}_i$, $i = 1, 2, ...$ be such that, for any i, (27) and (28) hold with $W, \bar{V}, \hat{V}$ replaced by $W_i, \bar{V}_i, \hat{V}_i$. Then obviously,

$$\cup_{i=1}^{\infty}\hat{V}_i \subset \cup_{i=1}^{\infty} W \subset \cup_{i=1}^{\infty}\bar{V}_i. \tag{31}$$

Moreover,

$$\mu(\cup_{i=1}^{\infty}\bar{V}_i) = \mu(\cup_{i=1}^{\infty}(\hat{V}_i \cup N_i)) \tag{32}$$

where, for each i, $N_i := \bar{V}_i \backslash \hat{V}_i$ and, by (28),

$$\mu(N_i) = 0. \tag{33}$$

Now (32) implies

$$\mu(\cup_{i=1}^{\infty}\bar{V}_i) \leq \mu(\cup_{i=1}^{\infty}\hat{V}_i) + \sum_{i=1}^{\infty}\mu(N_i) = \mu(\cup_{i=1}^{\infty}\hat{V}_i). \tag{34}$$

By (31), one may therefore infer that

$$\mu(\cup_{i=1}^{\infty}\hat{V}_i) = \mu(\cup_{i=1}^{\infty}\bar{V}_i) \tag{35}$$

and therefore, that $\cup_{i=1}^{\infty}W$ is also in $\mathcal{U}^p(X)$, i.e. that $\mathcal{U}^p(X)$ is closed under countable unions.

Since $\mathcal{U}^p(X)$ contains the open sets in the uniform topology and $\mathcal{U}^p(X)$ is closed under complements and under countable unions, it follows that $\mathcal{B}^u(X)$ is contained in $\mathcal{U}^p(X)$, as was to be shown. ■

Proof of Corollary 2. For any measure μ on $(X, \mathcal{B}^p(X))$, let μ^* be the outer measure that μ induces on the subsets of X, and let $\mathcal{N}(\mu)$ be the class of sets that have outer measure zero. By standard arguments, a measure $\bar{\mu}$ on the σ-algebra generated by $\mathcal{B}^p(X) \cup \mathcal{N}(\mu)$ is obtained by setting $\bar{\mu}(B) = \mu^*(B)$ for all B in this σ-algebra.[11] By Proposition 1 $\mathcal{B}^u(X)$ is contained in the σ-algebra generated by $\mathcal{B}^p(X) \cup \mathcal{N}(\mu)$, and the restriction of $\bar{\mu}$ to $\mathcal{B}^u(X)$ is a probability measure on $(X, \mathcal{B}^u(X))$ that agrees with μ on $\mathcal{B}^p(X)$. Uniqueness follows from the fact that, by the monotonicity of measures, for each set $W \in \mathcal{B}^p(X)$, the measure assigned to W must take the common value $\mu(\hat{V}) = \mu(\bar{V})$ of the sets $\hat{V}$ and $\bar{V}$ in $\mathcal{B}^p(X)$ for which (27) and (28) hold.

Since $X \in \mathcal{B}^p(X)$, for every probability measure $\bar{\mu}$ on $(X, \mathcal{B}^u(X))$, the restriction μ of $\bar{\mu}$ to $\mathcal{B}^p(X)$ is a probability measure on $(X, \mathcal{B}^p(X))$. Thus $\bar{\mu}$ is the unique extension of μ from $\mathcal{B}^p(X)$ to $\mathcal{B}^u(X)$ that is given by the first statement of the proposition. The second and third statements of the corollary follow immediately. ■

Proof of Proposition 3. Consider any measure $\mu \in \mathcal{M}^u(X)$. To prove the proposition, it suffices to show that there exists a separable set S such

[11] See, e.g., Ch. 3.3 in Dudley (2002).

that $\mu(S) = 1$. Given this set, standard arguments imply that there exists a set $S^* \subset S$ such that S^* is the smallest closed set that has μ-measure one.[12]

For any k, the separability of X_k implies that there exists a countable set $\hat{X}_k$ that is dense in X_k. The product

$$\hat{X} := \prod_{k=1}^{\infty} \hat{X}_k.$$

is obviously dense in X (in the uniform topology). For any $\delta > 0$, the open δ-balls

$$B^\delta(\hat{x}) := \{x' \in X | |\rho^u(x', \hat{x}) < \delta\} \tag{36}$$

around the points $\hat{x} \in \hat{X}$ form an (uncountable) open covering (in the uniform topology) of X.

For $K = 1, 2, ...$, consider the projection π^K from X to $X_1 \times ... \times X_K$. Proceeding as in the proof of Lemma 6, let $\mu^K = \mu \circ (\pi^K)^{-1}$ be the marginal distribution on $X_1 \times ... \times X_K$ that is induced by μ. As in th eproof of Lemma 6, Ulam's theorem implies that, for any $\varepsilon > 0$, there exists a compact set $C_\varepsilon^K \subset X_1 \times ... \times X_K$ such that

$$\mu^K(C_\varepsilon^K) \geq 1 - \varepsilon. \tag{37}$$

For any $\delta > 0$, Lemma 5 implies that the collection $\{\pi^K(B^\delta(\hat{x}))|\hat{x} \in \hat{X}\}$ is an open covering of $X_1 \times ... \times X_K$ and therefore of C_ε^K. Because C_ε^K is compact, this open covering of C_ε^K has a finite subcovering $\{\pi^K(B^\delta(\hat{x}))|\hat{x} \in \{\hat{x}_1^{\delta\varepsilon K}, ..., \hat{x}_{n_K}^{\delta\varepsilon K}\}\}$. For given δ and ε, the set $\cup_{K=1}^{\infty}\{\hat{x}_1^{\delta\varepsilon K}, ..., \hat{x}_{n_K}^{\delta\varepsilon K}\}$ is countable. It can therefore be written in the form $\{y_1^{\delta\varepsilon}, y_2^{\delta\varepsilon}, ...\}$ so that, for each k and each $i \in \{1, ..., n_K\}$, there exists $\ell^{\delta\varepsilon}(K, i)$ such that $y_{\ell^{\delta\varepsilon}(K,i)} = x_i^{\delta\varepsilon K}$. For any δ, ε, and K, the collection $\{\pi^K(B^\delta(y))|y \in \{y_1^{\delta\varepsilon}, y_2^{\delta\varepsilon}, ...\}\}$ is an open covering of C_ε^K, and the set

$$S^{\delta\varepsilon K} := (\cup_{\theta=1}^{\infty}\pi^K(B^\delta(y_\ell^{\delta\varepsilon}))) \times X_{K+1} \times X_{K+2} \times .... \tag{38}$$

is an open covering of the set $C_\varepsilon^K \times X_{K+1} \times X_{K+2} \times ....$ so (37) implies

$$\mu(S^{\delta\varepsilon K}) \geq 1 - \varepsilon. \tag{39}$$

Now consider the set

$$\begin{aligned} S^{\delta\varepsilon} \quad : \quad &= \cup_{\ell=1}^{\infty} B^\delta(y_\ell^{\delta\varepsilon}) \\ &= \bigcap_{K=1}^{\infty} [\pi^K \left(\bigcup_{\ell=1}^{\infty} B^\delta(y_\ell^{\delta\varepsilon}) \right) \times X_{K+1} \times X_{K+2} \times ....] \end{aligned} \tag{40}$$

[12] See Theorem 2.1, p. 27, in Parthasarathy (1967).

By the continuity of measures on monotone classes of sets and the commutativity of unions and projections, it follows that, for any δ and ε,

$$\begin{aligned}
\mu(S^{\delta\varepsilon}) &= \lim_{K\to\infty} \mu\left(\bigcup_{\ell=1}^{\infty}[\pi^K\left(B^\delta(y_\ell^{\delta\varepsilon})\right) \times X_{K+1} \times X_{K+2} \times ....]\right) \\
&= \lim_{K\to\infty} \mu(V^{\delta\varepsilon K})
\end{aligned}$$

and therefore, that

$$\mu(S^{\delta\varepsilon}) \geq 1-\varepsilon.. \tag{41}$$

Next, while still keeping δ fixed, for $n = 1, 2, ...$,let $\varepsilon^n = \frac{1}{n}$ and consider the set

$$\begin{aligned}
S^\delta \quad &: \quad = \cup_{n=1}^{\infty} S^{\delta\varepsilon^n} \\
&= \cup_{n=1}^{\infty} \cup_{\ell=1}^{\infty} B^\delta(y_\ell^{\delta\varepsilon^n}). \qquad (42)
\end{aligned}$$

Obviously,

$$\mu(S^\delta) \geq \mu(S^{\delta\varepsilon^n}) \tag{43}$$

for all n, so (41) implies

$$\mu(S^\delta) \geq 1 - \varepsilon^n = 1 - \frac{1}{n}$$

for all n and therefore

$$\mu(S^\delta) = 1. \tag{44}$$

Finally, for $m = 1, 2, ...,$ let $\delta^m = \frac{1}{m}$ and consider the set

$$S := \cap_{m=1}^{\infty} S^{\delta^m}. \tag{45}$$

This set satisfies

$$\begin{aligned}
\mu(S) &= 1 - \mu(\cup_{m=1}^{\infty}(X \backslash S^{\delta^m})) \\
&\geq 1 - \sum_{m=1}^{\infty}(1 - \mu(S^{\delta^m})) = 1,
\end{aligned}$$

and therefore, $\mu(S) = 1$. Moreover, since

$$S := \cap_{m=1}^{\infty}[\cup_{n=1}^{\infty} \cup_{\ell=1}^{\infty} B^\delta(y_\ell^{\delta^m\varepsilon^n})],$$

it is clear that $\{y_\ell^{\delta^m\varepsilon^n}, \ell = 1, 2, ..., m = 1, 2, ..., n = 1, 2, ...\}$ is a countable dense subset of S, i.e. that S is separable. ■

Proof of Corollary 4. The corollary follows from Theorem 11.3.3, p. 395, in Dudley upon noting that, for any measure $\mu^* \in \mathcal{M}^u(X)$ and any sequence $\{\mu^k\}$ of measures in $\mathcal{M}^u(X)$ that converges to μ^*, the countable union $S = S^* \cup [\cup_k S^k]$ of their separable supports is separable; see also Theorem 5, p. 238 in Billingsley (1968). ∎

References

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

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

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

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

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

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

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

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

[9] Gizatulina, A., and M. Hellwig (2017), The Generic Possibility of Full Surplus Extraction in Models with Large Type Spaces, Preprint 02/2017, Max Planck Institute for Research on Collective Goods, Bonn, Germany, http://www.coll.mpg.de/pdf_dat/2017_02online.pdf .

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

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

[12] Hellwig, M.F. (2016/17), 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 May 2017. http://www.coll.mpg.de/pdf_dat/2016_17online.pdf.

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

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

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

[16] Weinstein, J. and M. Yildiz (2007), A structure theorem for rationalizability with application to robust predictions of refinements. *Econometrica* 75, 365–400.