Appendix

Mathematical Preliminary

The following results are standard and have been adopted from various sources across Hammond (2004), Patrick Billingsley (2012), Patrick. Billingsley (1999), Cohn (2013), and McDonald and Weiss (2013)

*Weak and Weak*Topology*: For a topological vector space (T,$\mathcal{T}$), the weak topology is defined to be the weakest/coarsest topology under which each element of its continuous dual X* remains continuous on X. The dual X* is just the space of linear functions $f : X \mapsto \mathbb{R}$. In simpler terms if there is a set X and there is a function f which maps X to some topological space $(Y, \mathcal{Y})$, then the weak topology on X is the coarsest topology generated by sets $\{f^{-1}(G) : G \in \mathcal{Y}\}$, i.e. it ensures the function f is continuous.

The Weak*Topology is defined as the coarsest topology under which every element $x \in X$ corresponds to a continuous map on X^*.

Homeomorphism: Let $(T_1, \mathcal{T}_1)$ and $(T_2, \mathcal{T}_2)$ be topological spaces. A bijection $f : T_1 \mapsto T_2$ is a homeomorphism if both f and f^{-1} are continuous. Homeomorphisms preserve topological properties and makes the two sets topologically equivalent or indistinguishable. Thus it is a structure preserving operation.

Polish spaces and Measure Theory:

A few key properties related to Polish spaces are stated here. A Polish space is a complete separable metric space equipped with it's Borel σ-algebra, the smallest σ-algebra generated by sets which are open in the metric topology. Each closed subspace and each open subspace of a Polish space are Polish. The disjoint union of a finite or infinite sequence of Polish spaces is Polish. The product of a finite or infinite sequence of Polish spaces are Polish. Given a Polish space X, with it's Borel σ algebra($\mathcal{B}$), the set of all $\Delta(X, \mathcal{B})$ probability measures on the measurable space $(X, \mathcal{B})$ is Polish when equipped with the topology of weak convergence of probability measures. This topology corresponds to the weak* topology and thus implies, a sequence of measures $(\pi_n)_{n=0}^{\infty} \in \Delta(X, \mathcal{B})$ converges to a limit $\pi \in \Delta(X, \mathcal{B})$ if and only if for every bounded continuous function $f : X \mapsto \mathbb{R}$, $\int_X f(x)\pi_n$ converges in $\mathbb{R}$ to $\int_X f(x)\pi$. The metric which induces this topology is the Prohorov metric.

Polish spaces are what are often referred to as standard measurable spaces. An alternative definition provided for standard measurable spaces is— a measurable space $(X, \mathcal{B})$ is standard if it is Borel isomorphic to a compact metric space($(K, \mathcal{K})$. Borel Isomorphism corresponds to a bijection $f : X \mapsto K$, such that f and f^{-1} are both measurable. An example is $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ is Borel Isomorphic to $([0,1], \mathcal{B}([0,1]))$. More generally every separable metric space is homeomorphic to a subset of the Hilbert cube, and homeomorphisms on topological spaces are analogous to Borel isomorphisms on Borel spaces, separable metric spaces are standard.

Extension to the General Measurable Case

This section provides a rough sketch of how to proceed if the assumption that S is Polish is dropped to S being a general measurable space, the rest of the structure is retained. The player identities are in [0,1], with the corresponding completed Borel sigma algebra $\mathcal{B}_0([0,1])$ and S has some sigma algebra $\mathcal{S}$. As in Heifetz and Samet (1998) let $\Delta(X)$ be equipped with the sigma algebra Σ_Δ of the form $\beta^p(E) = \{\mu|\mu(E) \geq p\}$ for any event $E \in X$ and $0 \leq p \leq 1$.

Starting with a general measurable space, the Kolmogorov extension theorem cannot be applied to to form a unique belief over the infinite product space. The Ionescu tulcea theorem is used in general measurable spaces.

Definition 8.1 (Markov kernel)**.** *Cosma Shalizi (2007)*

Let $(X, \mathcal{A})$ and $(Y, \mathcal{B})$ be measurable spaces. A Markov kernel with source $(X, \mathcal{A})$ and target $(Y, \mathcal{B})$ is a map

$$\kappa : \mathcal{B} \times X \to [0,1]$$

with the following properties:

1. *For every (fixed) $B \in \mathcal{B}$, the map $x \mapsto \kappa(B, x)$ is $\mathcal{A}$-measurable.*
2. *For every (fixed) $x \in X$, the map $B \mapsto \kappa(B, x)$ is a probability measure on $(Y, \mathcal{B})$.*

In other words, it associates to each point $x \in X$ a probability measure $\kappa(dy|x) : B \mapsto \kappa(B, x)$ on $(Y, \mathcal{B})$ such that, for every measurable set $B \in \mathcal{B}$, the map $x \mapsto \kappa(B, x)$ is measurable with respect to the σ-algebra $\mathcal{A}$.

Markov kernels are essentially transition probabilities and they provide the mechanism for how to move from the space X to space Y. A Markov kernel is thus a generalisation of the notion of a conditional probability.Now the Ionescu-Tuclea Theorem from Tulcea (1963) is stated.

Theorem 8.1 (Ionescu-Tulcea theorem)**.** *Tulcea (1963)*

Suppose that $(\Omega_0, \mathcal{A}_0, P_0)$ is a probability space and $(\Omega_i, \mathcal{A}_i)$ for $i \in \mathbb{N}$ is a sequence of measurable spaces. For each $i \in \mathbb{N}$, let

$$\kappa_i : (\Omega^{i-1}, \mathcal{A}^{i-1}) \to (\Omega_i, \mathcal{A}_i)$$

be the Markov kernel derived from $(\Omega^{i-1}, \mathcal{A}^{i-1})$ and $(\Omega_i, \mathcal{A}_i)$, where

$$\Omega^i := \prod_{k=0}^{i} \Omega_k \text{ and } \mathcal{A}^i := \bigotimes_{k=0}^{i} \mathcal{A}_k.$$

Then there exists a sequence of probability measures

$$P_i := P_0 \otimes \bigotimes_{k=1}^{i} \kappa_k$$

defined on the product space for the sequence $(\Omega^i, \mathcal{A}^i)$, $i \in \mathbb{N}$, *and there exists a uniquely defined probability measure* P *on* $\left(\prod_{k=0}^{\infty} \Omega_k, \bigotimes_{k=0}^{\infty} \mathcal{A}_k\right)$, *so that*

$$P_i(A) = P\left(A \times \prod_{k=i+1}^{\infty} \Omega_k\right)$$

is satisfied for each $A \in \mathcal{A}^i$ *and* $i \in \mathbb{N}$. *(The measure* P *has conditional probabilities equal to the stochastic kernels.)*

The Ionescu-Tulcea theorem states that given a base space, and a probability measure over the base space, and a sequence of Markov kernels which provide transitions from one measurable space to the next, a unique probability measure on the infinite product space exists. In our setting, the base space is $(S \times [0,1])$ and per the Ionescu-Tulcea theorem any probability measure is allowed to be picked over the base space. From Hammond (2017) the uniform marginal condition (3.1) on $\Delta(S \times [0,1])$ provides the existence of a Markov kernel. For simplicity it is easier to pick a joint probability measure over $(S \times [0,1])$ which is equal to the product measure of $(S \times [0,1])$, i.e. a measure which reflects independence between beliefs about states and player identities.The order of hierarchies of beliefs are the sequence of measurable spaces in the Ionescu-Tulcea theorem. The transition probabilities in the setup are defined by the belief hierarchies of the players. Each player's belief of order k about the lower-order types is a transition probability which maps the belief of order $k-1$ to a probability distribution on the k^{th} order type space(Markov kernel).

Starting from the base space each step creates a new measure via the Markov kernel and one can construct the Markov kernel such that this measure is in line with the coherency conditions, then by the Ionescu-Tulcea theorem we have a unique probability measure on the infinite product space, which only implies the bijection result of (1.1). To get a result equivalent to the homeomorphsim for measurable spaces, as in Heifetz and Samet (1998), type morphisms, which preserve the types and states structure is required. Thus the Ionescu-Tulcea theorem is a strong tool when working with general measurable spaces to build a unique probability measure on the infinite order hierarchy. Note, however, as in Heifetz and Samet (1999) if such a type space constructed using Markov kernels(which follow the coherency conditions) is not the universal type space or there exists some Markov kernel which generates a universal type space, which is the proper subset of the space of all coherent hierarchies is to be proved in future work and beyond the scope of this paper.