Common assumption of cautious rationality and iterated admissibility*

Emiliano Catonini† Nicodemo De Vito‡

June 2014

Abstract. Iterated admissibility, i.e. the iterated deletion of weakly dominated strategies, is an important and extensively applied solution concept for complete information games. To understand when it is the appropriate one, conditions under which players want to avoid strategies that are weakly dominated in some reduced game along the procedure (although possibly not in the final set!) must be provided. It is intuitive that these conditions have to incorporate some cautious attitude of the players. Yet, to what extent players are cautious and assume that opponents are must be carefully defined in order to provide a correct motivation for iterated admissibility. Brandenburger, Friedenberg and Keisler (ECMA, 2008) define a notion of rationality, including a full-support requirement for lexicographic beliefs, which delivers iterated admissibility when players adopt it, assume (to a defined extent) that opponents adopt it, and so on, up to some finite level. This notion of rationality cannot be commonly assumed in their sense by players unless heavy exogenous restrictions to beliefs apply. Here, we provide meaningful but weaker notions of cautiousness and assumption such that cautious rationality can be commonly assumed by players and iterated admissibility is captured.

Keywords: iterated admissibility, weak dominance, lexicographic probability systems, assumption, rationality, cautiousness

1 Introduction

Iterated admissibility, coinciding with the iterated deletion of weakly dominated strategies, is among the most succesful solution concepts for complete information games. First, it does not rely on any exogenous equilibrium motivation: Players can perform it from scratch through nothing else than their strategic reasoning. Second, it reflects an intuitively

*Working paper. We want to thank Adam Brandenburger, Amanda Friedenberg and Jerome Keisler for inspiring this work. A special thank to Pierpaolo Battigalli for precious discussions and suggestions about our work. Thanks also to Edward Green, Byung Soo Lee and Burkhard Schipper for their comments.

†ICEF, Higher School of Economics, Moscow, emiliano.catonini@gmail.com

‡Bocconi University, nicodemo.devito@unibocconi.it

reasonable behavior: A strategy is not chosen when there is another one that, when it makes a difference, can only do better.[1] Yet, if we exclude some opponents' strategies with the same motivation, how our strategies perform against them should in principle not matter, and as Samuelson [13] pointed out a tension emerges. Stahl [14] already noticed that lexicographic beliefs [5], i.e lists of conjectures in a priority order, can solve this tension. Having different conjectures at uncomparable levels of likelihood allows to disregard some scenarios in the first place, yet using them as "tie-breakers" if her primary conjecture leaves the player undecided about what to do. Therefore, the focus then shifted to identify the principles that shape the right lexicographic beliefs, which inform the strategic reasoning of players who choose iteratively admissible strategies.

Brandenburger, Friedenberg and Keisler [6] (henceforth BFK), to whom this work is much indebted, define notions of *rationality* and *assumption* that, opportunely combined, deliver the iteratively admissible strategies (in a finite game). First, they incorporate in rationality a full support requirement, which we will call "*open-mindedness*": Players put every possible state of world (including opponents' types, i.e. opponents' beliefs) in the joint support of their lexicographic probability system[2] (henceforth LPS). Second, they set that an event is assumed by a LPS if every part[3] of the event is deemed as infinitely more likely than every part of the complementary event. Then, the Authors show an impossibility result: In a rich enough type structure (complete and continuous), players are unable to commonly assume rationality, i.e. the corresponding event is empty. The impossibility ceases to hold for poorer type structures, but this means imposing exogenous restrictions to the hierarchies of beliefs, which could find no justification in the context at hand. Now, suppose that players are willing to assume that everyone is rational; assume that everyone is rational and assumes that everyone else is rational; and so on. But if common assumption of rationality is impossible, at some point players must start having doubts. Why should they? This puzzling result has inspired other papers. Keisler and Lee [10] show the existence of complete type structures with discontinuous belief maps where the impossibility ceases to hold. Heifetz, Meier and Schipper [9] take a more radical way out by changing the solution concept.[4]

The aim of this paper instead is to characterize epistemically precisely iterated admissibility, while obtaining a non-empty "cautious rationality and common assumption of cautious rationality" event in a rich type structure, through different but meaningful notions of caution and assumption, and independently of the topology on the type structure.[5] Consequentialist players do not bother to put in the support of their LPS every possible belief of the opponents; nor they focus on every part of an event when they want to assume

[1] Moreover, in dynamic games without relevant ties among payoffs, iterated admissibility operationalizes extensive form rationalizability ([12] and [4]). Yet, the analysis of the extensive form solution concept is required to understand the epistemic motivations: see [4].

[2] Their notion of lexicographic probability system differs from the one in Blume, Brandenburger and Dekel [5], in that it requires mutual singularity.

[3] Intersection between any open set and the event.

[4] Also Asheim and Dufwemberg [1] defined a solution concept (fully admissible sets) that captures a form of cautiousness and full belief in rationality and that does not refine, nor is refined, by iterated admissibility. Barelli and Galanis [2] instead, characterize iterated admissibility with tie-breaking sets instead of lexicographic beliefs.

[5] A topology will be needed only to construct a Borel sigma-algebra of events.

it. Yet, if cautious, they care about considering all opponents' strategies as possible to some extent; coherently, when they assume an event, they consider each relevant component of the event (characterized by a particular payoff implication) as infinitely more likely than each component of the complementary event. *Cautiousness* is invariant to the type structure topology and finds an explicit correspondence in hierarchies: full support of first-order beliefs. In our view, instead, the interpretation of open-mindedness is more problematic: paradoxically it can be obtained even putting probability zero on the most relevant event, the open-mindedness of the opponents (see the Discussion).[6] Our notion of assumption is invariant to the type structure topology too and still has a meaningful preference-based representation (provided in the Appendix).

As a type structure, we adopt the mutually singular canonical one developed by Catonini and De Vito [7] (henceforth CDV). The canonical type structure allows to represent all meaningful hierarchies of lexicographic beliefs about strategies, so that no exogenous restriction is super-imposed and the states of interest will be entirely identified by the conceptually relevant events. Such type structure is complete, continuous and mutually singular, and it allows to compare results and identify the differences with BFK. Moreover, mutual singularity requires each measure of the LPS to represent the player's hypothesis conditional on the fact that the preceding ones do not realize, consistently with the preference-based representation of assumption we provide.

2 Cautiousness, assumption and iterated admissibility

2.1 Preliminaries

For every Polish space X, let $\mathcal{M}(X)$ denote the set of Borel probability measures on it. Let $\mathcal{N}_k(X) := (\mathcal{M}(X))^k$ and $\mathcal{N}(X) := \cup_{k\in\mathbb{N}}\mathcal{N}_k(X)$ denote the sets of lenght k and any-lenght lexicographic beliefs. Moreover, let $\mathcal{L}_k(X) \subseteq \mathcal{N}_k(X)$ and $\mathcal{L}(X) = \cup_{k\in\mathbb{N}}\mathcal{L}_k(X)$ denote the sets of lenght k and any-lenght LPS's, i.e. the lexicographic beliefs $\lambda = (\lambda^1, ..., \lambda^k)$ satisfying *mutual singularity*: There exist Borel sets $E_1, ..., E_k$ in X such that for every $p \leq k$, $\lambda^p(E_p) = 1$ and $\lambda^p(E_q) = 0$ for $q \neq p$. Finally, let $\mathcal{N}^+(X) \subseteq \mathcal{N}(X)$ and $\mathcal{L}^+(X) \subseteq \mathcal{L}(X)$ denote the sets of *full-support* lexicographic beliefs and LPS's, i.e. such that $\cup_{k\in\mathbb{N}}\text{Supp}\lambda^k = X$

For any $\lambda = (\lambda^1, ..., \lambda^k) \in \mathcal{N}_k(X\times Y)$, we denote by $\overline{\text{marg}}_Y\lambda := (\text{marg}_Y\lambda^1, ..., \text{marg}_Y\lambda^k)$ the marginal lexicographic belief over the subspace Y.

Consider the strategic form of a finite complete information game $\langle I, (S_i, u_i)_{i\in I}\rangle$, where I is the finite set of players and, for every $i \in I$, S_i is the finite set of strategies and $u_i : S \to \mathbb{R}$ is the payoff function. For all profiles of sets $(X_i)_{i\in I}$ we will denote by X_{-i} the cartesian product over all players but i, and by X the cartesian product over all players. Define the expected payoff function $\pi_i : \mathcal{M}(S_i) \times \mathcal{M}(S_{-i}) \to \mathbb{R}$ in the usual way. A pure strategy s_i or a pure opponents' subprofile of strategies s_{-i} as argument of π_i will indicate the probability distribution putting probability 1 on it.

[6] The intuitive interpretation of open-mindedness is rescued in BFK by the definition of assumption, according to which players to put positive probability on every part of the rationality event if they assume it. The use of open-mindedness with our notion of assumption, instead, would remain misleading.

The tuple $\mathcal{T} = \langle S_i, T_i, \beta_i \rangle_{i \in I}$ is a *lexicographic type structure* [6] if, for every $i \in I$, T_i is a Polish type space and $\beta_i : T_i \to \mathcal{L}(T_{-i} \times S_{-i})$ is the measurable belief map that associates each type with a LPS over opponents' strategies and types. A lexicographic type structure is *complete* if each β_i is onto; it is *continuous* if each β_i is continuous. Each cartesian product of sets is endowed with the product topology and a Borel σ-algebra of *events*. Given any two type structures $\mathcal{T} = \langle S_i, T_i, \beta_i \rangle_{i \in I}$ and $\mathcal{T}' = \langle S_i, T_i', \beta_i' \rangle_{i \in I}$, a *type morphism* between them, if it exists, is measurable map $\tau = (\tau_i : T_i \to T_i')_{i \in I}$ such that for every $i \in I$, $t_i \in T_i$ and measurable $E \subseteq S_{-i} \times T_{-i}$, $\beta_i(t_i)(E) = \beta_i'(\tau_i(t_i))(\overline{\tau}_i(E))$, where $\overline{\tau}_i : (s_i, t_i) \mapsto (s_i, \tau_i(t_i))$.

2.2 Iterated admissibility

Iterated admissibility is a reduction procedure of the set of strategy profiles that relies on the admissibility criterion.

Definition 1 *Fix a cartesian set $X_i \times X_{-i} \subseteq S_i \times S_{-i}$. A strategy $s_i \in X_i$ is* ***admissible with respect*** *to $X_i \times X_{-i}$ if there exists $\mu_i \in \mathcal{M}(S_{-i})$ such that* $\mathrm{Supp}\mu_i = X_{-i}$ *and $\pi_i(s_i, \mu_i) \geq \pi_i(s_i', \mu_i)$ for every $s_i' \in X_i$.*

The *iteration of admissibility* delivers a chain of cartesian subsets of strategy profiles $S^0 = S, S^1, S^2, ...$ such that for every $i \in I$, $s_i \in S_i$ and $n \in \mathbb{N}$, $s_i \in S_i^n$ if and only if $s_i \in S_i^{n-1}$ and it is admissible with respect to $S_i^{n-1} \times S_{-i}^{n-1}$.

In a finite game, for every $n \in \mathbb{N}$, S^n is non-empty and there exists $M \in \mathbb{N}$ such that $S^n = S^M$ for all $n \geq M$. A standard result due to Pearce [12] allows to claim that a strategy $s_i \in X_i$ is admissible with respect to $X_i \times X_{-i}$ if and only it is not weakly dominated over $X_i \times X_{-i}$. Thus, iterated admissibility coincides with the iterated deletion of weakly dominated strategies.

As already argued, looking at mere, fully mixed conjectures may wrongly justify the choice of an iteratively inadmissible strategy: For a player i there may be strategies that are not weakly dominated over S^M and yet do not belong to S_i^M. The reason is that a player who performs iterated admissibility wants to avoid also strategies that are weakly dominated over some previous set of the chain. Thus, she considers every opponents' subprofile in that set still possible to some extent. Yet, the ones that do not survive the following step must not be considered nearly as likely as the ones that do, otherwise we would run the opposite risk of rescuing strategies that are weakly dominated over S^M. Hence, we need lists of conjectures at uncomparable levels of likelihood. These lists are nothing else than lexicographic beliefs over opponents' strategy subprofiles, which we call *lexicographic conjectures*. One *cannot* impose them to be mutually singular, otherwise some iteratively admissible strategy that are never the only best reply to conjectures over S_{-i}^M may find no justification (see the original example of [3] in BFK).

With respect to lexicographic conjectures, we take the standard definition of lexicographic best reply. For any two vectors $x, y \in \mathbb{R}^k$, we write $x \geq^L y$ if either $x_p = y_p$ for every $p \leq k$ or there exists $q \leq k$ such that $x_q > y_q$ and $x_p = y_p$ for every $p < q$.

Definition 2 *Consider a player $i \in I$ and a lexicographic conjecture $\lambda \in \mathcal{N}(S_{-i})$. A strategy $s_i \in S_i$ is a* ***lexicographic best reply*** *to $\lambda_i = (\lambda_i^1, ..., \lambda_i^k)$ if for every $s_i' \neq s_i$, $(\pi_i(s_i, \lambda_i^p))_{p=1}^k \geq^L (\pi_i(s_i', \lambda_i^p))_{p=1}^k$.*

That is, a strategy is optimal against a lexicographic conjecture if it defeats each other strategy before (along the list) the opposite occurs (or if they are equivalent against all conjectures of the list). This optimality notion is the starting point of our construction.

2.3 Rationality and cautiousness

Fix any lexicographic type structure $\mathcal{T} = \langle S_i, T_i, \beta_i \rangle_{i \in I}$. We are primarily interested in strategy-type pairs where each player plays a lexicographic best reply to her lexicographic conjecture.

Definition 3 *A strategy-type pair $(s_i, t_i) \in S_i \times T_i$ is* ***rational*** *if s_i is a lexicographic best reply to $\overline{\mathrm{marg}}_{S_{-i}} \beta_i(t_i) \in \mathcal{N}(S_{-i})$. We denote by R_i the set of all rational $(s_i, t_i) \in S_i \times T_i$.*

Cautiousness instead is a property of the type alone. A cautious type deems all opponents' strategies as possible to some extent.

Definition 4 *A type $t_i \in T_i$ is* ***cautious*** *if $\overline{\mathrm{marg}}_{S_{-i}} \beta_i(t_i) \in \mathcal{N}^+(S_{-i})$. We denote by C_i the set of all $(s_i, t_i) \in S_i \times T_i$ where t_i is cautious.*

In terms of induced hierarchies, cautiousness corresponds to having a full-support first-order belief. Moreover, cautiousness is invariant to the type structure topology and to type morphisms in the following strong sense: A cautious type in the original type structure is mapped into a cautious type in the destination structure by a type morphism.

In BFK, a strategy-type pair (s_i, t_i) is rational only if $\beta_i(t_i)$ has full support (other than s_i being a lexicographic best reply to $\overline{\mathrm{marg}}_{S_{-i}} \beta_i(t_i)$). We call this different full-support property "open-mindedness" to distinguish it from cautiousness.

Definition 5 *A type $t_i \in T_i$ is* ***open-minded*** *if $\beta_i(t_i) \in \mathcal{L}^+(T_{-i} \times S_{-i})$.*

In terms of hierarchies, open-mindedness guarantees full-support of first-order beliefs too; full support of higher-order beliefs depends on the type structure.[7] Instead, open-mindedness does no have the same invariance properties of cautiousness with respect to topology and type morphisms.

Overall, open-mindedness is clearly stronger than cautiousness: full-support of the entire LPS implies full support of the marginal LPS, but the vice versa is not true. So, an open-minded type is cautious but a cautious type needs not be open-minded. Yet, cautious rationality is still a sufficient condition for admissibility. Therefore we can provide the following weakening of the analogous result in BFK.

[7] We conjecture that also in a complete but not continuous type structure, open-mindedness does not imply full support of all orders of belief in the induced hierarchy. The type structure of [10] could be a case in point. See CDV for details.

Proposition 1 *Fix a type structure* $\mathcal{T} = \langle S_i, T_i, \beta_i \rangle_{i \in I}$. *If strategy-type pair* $(s_i, t_i) \in S_i \times T_i$ *is cautiously rational, then* s_i *is admissible.*

Proof: If $(s_i, t_i) \in S_i \times T_i$ is cautiously rational, then s_i is a lexicographic best reply to $\overline{\mathrm{marg}}_{S_{-i}} \beta_i(t_i) \in \mathcal{N}^+(S_{-i})$, where $\beta_i(t_i) = (\lambda_i^1, ..., \lambda_i^k) \in \mathcal{L}(S_{-i} \times T_{-i})$. By Proposition 1 in [5], to every $\overline{\mathrm{marg}}_{S_{-i}}(\lambda_i^1, ..., \lambda_i^k) \in \mathcal{N}^+(S_{-i})$ there corresponds a probability measure $\nu_i \in \mathcal{M}(S_{-i})$, with $\mathrm{Supp}\nu_i = S_{-i}$, such that $\pi_i(s_i, \nu_i) \geq \pi_i(s_i', \nu_i)$ for every $s_i' \in S_i$. ■

We denote by $R_i^1 := R_i \cap C_i$ the set of **cautiously rational** strategy-type pairs; Cautious rationality, R^1, will be the first building block of our epistemic characterization of iterated admissibility, capturing the first step.

2.4 Assumption

To capture the further iterations of admissibility, we need to define the events that identify and motivate the right lexicographic conjectures over opponents' behavior. These events will be based on the concept of *assumption*. When a player assumes an event E, she should not completely rule out the possibility that the complement of E occurs. Yet, she must consider E *infinitely more likely* than its complement. This translates first into E having probability 1 up to some measure of the LPS and 0 thereafter. But this is not enough. In BFK it is required that every part of the event (intersection between an open set and the event) is given positive probability by some measure in the LPS. Here we only require that every "relevant part" (intersection between a strategy-based cylinder and the event) is given positive probability by some measure in the LPS. This implies that a player, when assuming an event, gives a positive probability to every payoff-relevant component of the event and deems it infinitely more likely than the complementary event.

Definition 6 *A LPS* $\lambda_i = (\lambda_i^1, ..., \lambda_i^k) \in \mathcal{L}(S_{-i} \times T_{-i})$ ***assumes*** *an event* $E \subseteq S_{-i} \times T_{-i}$ *at level* $l \leq k$ *if:*

1. $\lambda_i^q(E) = 1$ *for every* $q \leq l$*;*
2. $\lambda_i^q(E) = 0$ *for every* $q > l$*;*
3. *for every cylinder* $Y = \{s_{-i}\} \times T_{-i}$*, if* $Y \cap E \neq \emptyset$*, then* $\lambda_i^p(Y \cap E) > 0$ *for some* $p \leq k$*.*

On the other hand, λ_i **assumes** E ***à la*** **BFK** if 1 and 2 hold and 3 holds for every open $Y \subseteq S_{-i} \times T_{-i}$. In BFK, this is not exactly the definition of assumption but a characterization, which they then use as a working definition. We chose to take directly a definition of this kind for our notion of assumption because, like in BFK, it is the one that can be easily given a preference-based representation, which we expose in the Appendix. The gist of it is the following: If i assumes the event E, she prefers to bet on the realization of any relevant part in E (i.e. a part of E that can be distinguished by its observable implication) rather than on the complement of E.

It is straightforward to notice that our notion of assumption is weaker than the one in BFK: If a LPS assumes an event E in BFK, it also assumes E here because cylinders are open sets. Clearly the vice versa is not true. Yet, it is interesting to notice that if λ_i is open-minded and E is open, the two notions coincide: Every part of E is an intersection between two open sets, so it is open, so by full support is given positive probability by some measure of the LPS like assumption a la BFK requires.

Proposition 2 *If $\lambda_i \in \mathcal{L}(S_{-i} \times T_{-i})$ assumes E à la BFK, then λ_i assumes E. Conversely, suppose that λ_i has full-support and E is open. Then, if λ_i assumes E, λ_i assumes E à la BFK.*

Our notion of assumption has a very convenient property: invariance with respect to the topology of the type structure (keeping the same sigma-algebra). As Keisler and Lee [10] show, instead, there exist different topologies on the type structure that generate the same Borel sigma-algebra where one can find types whose LPS displays common BFK-assumption of rationality, which is impossible with the topology that makes the belief maps continuous. Interestingly, this is also connected to the equivalence with our notion of assumption, as shown later in the Discussion.

We finally define the corresponding operator A_i, which takes an event $E \subseteq S_{-i} \times T_{-i}$ and yields the set of strategy-type pairs that assume E. The operator is non-monotonic: If $F \subseteq E$, it needs not be the case that $A_i(F) \subseteq A_i(E)$. Otherwise, assumption would not be suitable to characterize iterated admissibility: as already observed, in a reduced game along the procedure, previously eliminated strategies may be not weakly dominated anymore.

2.5 Common assumption of cautious rationality

In this section, we adopt the *mutually singular canonical type structure* $\mathcal{T}_u^* = \langle S_i, \Lambda_i, g_i \rangle_{i \in I}$ developed by CDV, because it features some fundamental properties for our purposes. First, it is mutually singular, complete and continuous, making results comparable with BFK. Second and most importanlty, as CDV prove, it is *terminal* with respect to every structure $\mathcal{T} = \langle S_i, T_i, \beta_i \rangle_{i \in I}$ where each type induces a mutually singular hierarchy.[8] Focusing on a two-players game for notational simplicity, the hierarchy induced by a type t_i is the sequence $h_i(t_i) = (b_i^1(t_i), b_i^2(t_i), ...)$ where $b_i^1 = \overline{\mathrm{marg}}_{S_{-i}} \beta_i(t_i)$, $b_i^2 \in \mathcal{N}(S_{-i} \times \mathcal{N}(S_i))$ is the level-by-level pushforward of $\beta_i(t_i)$ on $S_{-i} \times \mathcal{M}(S_i)$ through $f : (s_{-i}, t_{-i}) \mapsto (s_{-i}, \beta_{-i}(t_{-i}))$, and so on. For our purposes we will not need to induce beliefs beyond the second order. The hierarchy is mutually singular if for some $m \in \mathbb{N}$, $b_i^m(t_i)$ is a LPS. Terminality means that there exists a type morphism between every such $\mathcal{T}$ and $\mathcal{T}_u^*$. In words, $\mathcal{T}$ can be embedded in $\mathcal{T}_u^*$ as a belief-closed subspace.

[8] As Lee [11] noticed, mutual singularity can be just "cosmetic" when putting probability on different types at different levels does not mean putting probability on different opponents' hierarchies, because of possible redundancies. For this reason, mutual singularity of the type structure is not sufficient for the existence of a type morphism in the canonical one. Mutual singularity of the hierarchy induced by each type, instead, is sufficient for common belief in mutual singularity (and coherence) of the hierarchies. Such hierarchies are exactly those captured by the ones captured by the mutually singular canonical type structure (see CDV).

In the canonical type structure, we can define the *cautious rationality and m–th order assumption of cautious rationality*, as well as the *cautious rationality and common assumption of rationality* events[9] inductively as follows:

$$\begin{aligned} R_i^{m+1} &: = R_i^m \cap (A_i(R_{-i}^m)),\ \forall m \geq 1; \\ R^\infty &: = \cap_{m\in\mathbb{N}} R^m. \end{aligned}$$

The behavioral implications of the first events correspond step-by-step to the iteratively admissible strategy profiles. The last event is non-empty too and its behavioral implications coincide with the final set of the iterated admissibility procedure. These facts are summarized in the following characterization theorem.

Theorem 1 *In the canonical mutually singular type structure, for every* $n \geq 1$, $S^n = \mathrm{Proj}_S R^n$. *Moreover,* $S^M = \mathrm{Proj}_S R^\infty$.

The strategy of the proof is the following. We construct a finite mutually singulare type structure with both uncautious and cautious types. To the cautious types we attach LPS's that justify playing the admissible strategies and will display assumption of cautious rationality up to different orders, or common assumption of cautious rationality. We obtain the latter by letting a subset of cautious types that can rationally play the iteratively admissible strategies "believe in each other" in their primary hypothesis. Then, we embed this structure into the mutually singular canonical one, exploiting its terminality property. Cautious rationality is preserved by the type morphism and types are constructed in such a way to assume the orders of cautious rationality they have been intended for in the canonical type structure.

Proof: For every $n \leq M+1$, $i \in I$ and $s_i \in S_i^n$, take a $\mu_{s_i}^n \in \mathcal{M}(S_{-i}^{n-1})$ such that $\mathrm{Supp}\mu_{s_i}^n = S_{-i}^{n-1}$ and for every $s_i' \in S_i$, $\pi_i(s_i, \mu_{s_i}^n) \geq \pi_i(s_i', \mu_{s_i}^n)$.

Now we build the finite type structure that we will embed in the canonical one. For every $i \in I$ and $0 \leq n \leq M+1$, create a set of types $T_i^n = \left\{t_{s_i}^n\right\}_{s_i \in S_i^n}$. Let $T_i := \cup_{n\leq M+1} T_i^n$. For every $i \in I$, define a belief map $\beta_i : T_i \to \mathcal{L}(S_{-i} \times T_{-i})$ by the following inductive procedure.

- For every $s_i \in S_i$, take a $\nu_i^{s_i} \in \mathcal{M}(S_{-i} \times T_{-i})$ such that $\mathrm{Suppmarg}_{S_{-i}} \nu_i^{s_i} \neq S_{-i}$. Let $\beta_i(t_{s_i}^0) := (\nu_i^{s_i})$.

- For every $s_i \in S_i^1$, take a $\eta_i^{s_i} \in \mathcal{M}(S_{-i} \times T_{-i}^0)$ such that $\mathrm{marg}_{S_{-i}} \eta_i^{s_i} = \mu_{s_i}^1$. Let $\beta_i(t_{s_i}^1) := (\eta_i^{s_i})$.

- For every $1 < n \leq M$ and $s_i \in S_i^n$, take *the* $\mu_i \in \mathcal{M}(S_{-i} \times T_{-i}^{n-1})$ such that for every $s_{-i} \in S_{-i}^{n-1}$, $\mu_i((s_{-i}, t_{s_{-i}}^{n-1})) = \mu_{s_i}^n(s_{-i})$. Let $\beta_i(t_{s_i}^n) := (\mu_i, \beta_i(t_{s_i}^{n-1}))$.

- For every $s_i \in S_i^{M+1}$, take *the* $\mu_i \in \mathcal{M}(S_{-i} \times T_{-i}^{M+1})$ such that for every $s_{-i} \in S_{-i}^{M+1}$, $\mu_i((s_{-i}, t_{s_{-i}}^{M+1})) = \mu_{s_i}^{M+1}(s_{-i})$. Let $\beta_i(t_{s_i}^{M+1}) := (\mu_i, \beta_i(t_{s_i}^{M-1}))$.

[9] We prove in the Appendix that all the sets we use are events.

Notice that for every $i \in I$ and $t_i \in T_i$, t_i is cautious by marginal full support of the last measure in $\beta_i(t_i)$, and $\beta_i(t_i)$ is mutually singular by disjoint supports. Moreover, $b_i^2(t_i) = (\lambda^1, ..., \lambda^k)$ is mutually singular: notice that for every $p < k$, λ^p puts positive probability only on opponents' first-order beliefs of lenght $k-p$; morever, λ^k and λ^{k-1} put positive probability only respectively on non full-support and full-support opponents' first-order beliefs. Hence, the structure can be embedded in the mutually singular canonical type structure $\mathcal{T}_u^* = \langle S_i, \Lambda_i, g_i \rangle_{i \in I}$. Keep the same labelling for types after the embedding. Then, with a slight abuse of notation, we can write that for every $i \in I$ and $t_i \in T_i$, $\beta_i(t_i) = g_i(t_i)$.

First, we prove by induction that for every $i \in I$, $n \leq M$ and $s_i \in S_i^n$, $(s_i, t_{s_i}^n) \in R_i^n \backslash R_i^{n+1}$, which implies $S_i^n \subseteq \text{Proj}_{S_i} R_i^n$.

- **Inductive hypothesis ($n - 1 \leq M$):** For every $i \in I$, $m \leq n - 1$ and $s_i \in S_i^m$, $(s_i, t_{s_i}^m) \in R_i^m \backslash R_i^{m+1}$.
- **Basis step ($n = 1$).** For every $i \in I$ and $s_i \in S_i$, $t_{s_i}^0 \notin R_i^1$ because $t_{s_i}^0$ is not cautious. Hence, for every $s_i \in S_i^1$, $g_i(t_{s_i}^1)$ puts probability 0 on R_{-i}^1, but $g_i(t_{s_i}^1)$ is cautious and has s_i as a lexicographic best reply, so $(s_i, t_{s_i}^1) \in R_i^1 \backslash R_i^2$.
- **Inductive step ($n \leq M$).** Take any $i \in I$ and $s_i \in S_i^n$. By the inductive hypothesis, the measures in $g_i(t_{s_i}^n) = (\lambda_i^1, ..., \lambda_i^n)$ have supports in, respectively, $R_{-i}^{n-1} \backslash R_{-i}^n, ..., (S_{-i} \times \Lambda_i) \backslash R_{-i}^1$. Moreover, $\text{Suppmarg}_{S_{-i}} \lambda_i^m = S_{-i}^{n-m}$ for every $m < n$. Therefore, $g_i(t_{s_i}^n)$ assumes R_{-i}^m at level $n - m$, so $(s_i, t_{s_i}^n) \in R_i^n$. Yet, since λ_i^1 puts probability 0 on R_{-i}^n, $(s_i, t_{s_i}^n) \notin R_i^{n+1}$.

In a similar way,[10] for every $i \in I$, $s_i \in S_i^{M+1}$ and $m \leq M$, $g_i(t_{s_i}^{M+1}) = (\lambda_i^1, ..., \lambda_i^M)$ assumes R_{-i}^m (at level $M - m$), so that $(s_i, t_i^{M+1}) \in R_i^{M+1}$. But then, $\text{Supp}\lambda_i^1 \subseteq R_{-i}^{M+1}$ and $\text{Suppmarg}_{S_{-i}} \lambda_i^1 = S_{-i}^{M+1}$, so $g_i(t_{s_i}^{M+1})$ assumes also R_{-i}^{M+1}. Hence, $(s_i, t_{s_i}^{M+1}) \in R_i^{M+2}$. By induction, it is immediate to show that then for every $n \in \mathbb{N}$, $g_i(t_{s_i}^{M+1})$ assumes R_{-i}^n. Hence, $(s_i, t_{s_i}^{M+1}) \in R_i^\infty$ and $S_i^M = S_i^{M+1} \subseteq \text{Proj}_{S_i} R_i^\infty$.

For the opposite inclusion, take as inductive hypothesis that $S^n \supseteq \text{Proj}_S R^n$, that together with the opposite inclusion already proved implies $S^n = \text{Proj}_S R^n$.

Take any $(s_i, t_i) \in R_i^n \subseteq R_i^{n-1}$. Notice that s_i is a lexicographic best reply to the lexicographic conjecture $\overline{\text{marg}}_{S_{-i}} g_i(t_i) = (\lambda_i^1, ..., \lambda_i^k)$, where, if $n > 1$, $g_i(t_i)$ assumes R^{n-1} at some level $l \leq k$; if $n = 1$, set $l = k$. By the inductive hypothesis, it follows that $s_i \in S_i^{n-1}$, for every $i \in I$. Hence it is enough to show the existence of a probability measure $\nu_i \in \mathcal{M}(S_{-i})$ such that $\text{Supp}\nu_i = S_{-i}^{n-1}$ and $\pi_i(s_i, \nu_i) \geq \pi_i(s_i', \nu_i)$ for every $s_i' \in S_i^{n-1}$. If s_i is indifferent to all s_i' against the first l conjectures, s_i is a best reply to any convex combination $\nu_i := \alpha_1 \lambda_i^1 + ... + \alpha_l \lambda_i^l$, and since *by the third requirement of assumption*

[10] Notice that, since pairs of the kind $(s_{-i}, t_{s_{-i}}^{M+1})$ appear in the first measure of $g_i(t_{s_i}^{M+1})$, one can show that $t_{s_i}^{M+1}$ assumes R_{-i}^n only provided that $(s_{-i}, t_{-i}^{M+1}) \in R_{-i}^n$, for every $s_{-i} \in S_{-i}^{M+1}$. The inductive procedure can start because for every $j \in I$, $(s_j, t_{s_j}^{M+1}) \in R_j^1$.

$\text{Supp}\nu_i = S_{-i}^{n-1}$, $s_i \in S_i^n$. Else, let d be the minimum *positive* value $\pi_i(s_i, \lambda_i^q) - \pi_i(s_i', \lambda_i^q)$ over $s_i' \in S_i$ and $q \leq l$. Then, s_i is a best reply to any convex combination $\nu_i := \alpha_1\lambda_i^1 + ... + \alpha_l\lambda_i^l$ such that for every $p < l$, $\alpha_{p+1} \cdot l \cdot (\max_{s\in S} \pi_i(s) - \min_{s\in S} \pi_i(s)) \leq \alpha_p \cdot d$. Since $\text{Supp}\nu_i = S_{-i}^{n-1}$, then $s_i \in S_i^n$.

Moreover, since $S^M \supseteq \text{Proj}_S R^M \supseteq \text{Proj}_S R^\infty$, it holds $S^M \supseteq \text{Proj}_S R^\infty$. ■

3 Discussion

Two natural questions may arise at this point. First: What drives the non-emptiness of the common assumption of cautious rationality event with respect to BFK's emptiness result? The new notion of assumption, the switch from open-mindedness to cautiousness, or both? Second: Does the non-emptiness hold in every complete and continuous type structures like BFK's emptiness?

We have a sharp answer to the first question: the new notion of assumption alone. The key is that our notion of assumption allows to assume at the same level of an LPS nested events whenever they have the same behavioral implications.[11] One may think that this is impossible with open-mindedness in place of cautiousness: the intuitive understanding of full support leads to think that, for instance, the event[12] "my opponent is open-minded and rational but does not assume I am open-minded and rational" must be assigned positive probability. In that case, it would not be possible to assume open-minded rationality and assumption of open-minded rationality at the same level. But that is not the case. Even the whole event "my opponent is open-minded" can be assigned probability zero by a full support measure over the opponent's type structure. The crucial point is the following: The event that the opponent is open-minded is "thin", i.e. the event that she is not open-minded is dense. As we prove in the Appendix, the set of non-full support measures on a space is dense in the set of all measures and the statement can be extended to LPS's; then, being the belief map a homeomorphism, the denseness is transfered to non-openminded types.[13] Hence, for a LPS to have full support, it is sufficient that one measure puts positive probability on each point of the countable dense subset of opponents' non-open-minded types.[14]

A constructive way to show the non-emptiness of *common assumption of open-minded rationality* is then the following. The construction in the previous section must be repli-

[11] In another perspective, a high enough order of assumption of rationality can imply all the further ones, up to the common assumption of rationality.

[12] All the following sets are proved to be events (i.e. measurable) in BFK for a complete and continuous type structure which is metrized like the canonical we adopt (see the Appendix). Hence, they are events also in our context.

[13] Thus, the set of open-minded types has empty interior in the mutually singular canonical type structure. It is shown in CDV that such set is also dense - specifically, open-mided types form a residual subset of the canonical type structure. However, there is a sense in which the set of open-mided types is "thin". As we show in our companion paper, if the BFK indifference condition holds (cf. BFK, Definition 10.1), then rational and open-minded strategy-type pairs form a non-dense set (with empty interior) in the canonical type space.

[14] A Borel subset of a Polish space is separable.

cated with two variations. First, the whole canonical structure must be dragged into the picture and the new types are additional to it. Second, instead of defining the set of uncautious types, the last measure of the LPS attached to each new type must be concentrated on the countable dense subset of opponents' non open-minded types. In the so obtained type structure, the new types still induce mutually singular hierarchies: the last measure of the attached LPS puts probability one on types from the canonical with non-full-support first order beliefs, while all other measures put probability one on other new cautious types. Therefore, there is a type morphism that maps the types from the canonical into their original ones and the new ones into their "twins".

Yet, although technically feasible, we find the joint use of open-mindedness, which represents caution by focusing on every part of the state space, and our notion of assumption, which does *not* focus on every part of the assumed event, somewhat conceptually contradictory.[15]

Keisler and Lee [10] obtain that different orders of rationality can be assumed (*à la* BFK) at the same level by choosing a particular topology that makes the belief maps discontinuous. Interestingly, since the m-th order assumption of rationality events are open in their type structure and all considered LPS's are of full support, by Proposition 2 in their context the BFK notion of assumption coincides with ours. Both for open-mindedness and assumption *à la* BFK (conditional on a particular topology), the non-emptiness results seem to be granted by a mathematical translation of the concepts that departs from their intuitive understanding, making them more similar to cautiousness and our version of assumption.

The answer to the second question is no. Every complete and continuous lexicographic type structure is not compact, because the set of LPS's is not compact, so continuity and ontoness require also the type spaces not to be compact. As a consequence, the typical finite intersection property arguments used in other epistemic characterizations to show non-emptiness of the "common belief" sets would not work here. From another viewpoint, the closed construction of types that "believe in each other", granting non-emptiness of common assumption of cautious rationality in the canonical structure, may not be found in a complete and continuous structure. This is the reason why, differently from other characterization results, our analysis is not carried on in a generic continuous and complete structure, but in the canonical one.

4 Appendix

4.1 Preference-based representation of assumption

Fix a lexicographic type structure $\mathcal{T} = \langle S_i, T_i, \beta_i \rangle_{i \in I}$. To shorten notation, it will be convenient to set $\Omega := T_{-i} \times S_{-i}$ and to drop i's subscript from LPS's λ_i on Ω. Fix a LPS $\lambda = (\lambda^1, ..., \lambda^k) \in \mathcal{L}(\Omega)$. Let $\mathcal{A}$ be the set of all measurable functions from Ω to $[0, 1]$. For

[15] Yang [15] shows the non-emptiness of common assumption of open-minded rationality (in our sense) in a non-mutually singular canonical structure. Dekel, Friedenberg and Siniscalchi [8] show that dropping mutual singularity overall the results of BFK can be replicated in the new setting.

every $x, y \in \mathcal{A}$, define $\succsim^{\lambda}$ on $\mathcal{A}$ as follows:

$$x \succsim^{\lambda} y \Longleftrightarrow \left(\int_{\Omega} x(\omega) d\lambda^{m}(\omega)\right)_{m=1}^{k} \geq^{L} \left(\int_{\Omega} y(\omega) d\lambda^{m}(\omega)\right)_{m=1}^{k}.$$

Given a Borel set $E \subseteq \Omega$ and acts $x, z \in \mathcal{A}$, let $(x_E, z_{\Omega \backslash E}) \in \mathcal{A}$ be the act defined as $(x_E, z_{\Omega \backslash E})(\omega) = x(\omega)$ if $\omega \in E$ and $(x_E, z_{\Omega \backslash E})(\omega) = z(\omega)$ if $\omega \in \Omega \backslash E$. Moreover, let $\succsim_E^{\lambda}$ denote the conditional preference given E, that is, $x \succsim_E^{\lambda} y$ if and only if $(x_E, z_{\Omega \backslash E}) \succsim^{\lambda} (y_E, z_{\Omega \backslash E})$ for some (implying all) $z \in \mathcal{A}$.

Definition 7 *A non-empty Borel set $E \subseteq S_{-i} \times T_{-i}$ is assumed under $\succsim^{\lambda}$ if:*

1. *(**relevance**) for every $s_{-i} \in S_{-i}$ and every cylinder $U = \{s_{-i}\} \times T_{-i}$, if $E \cap U \neq \emptyset$ then there exist two acts $x, y \in \mathcal{A}$ such that $x \succ_{E \cap U}^{\lambda} y$;*

2. *(**strict determination**) for all $x, y \in \mathcal{A}$, $x \succ_E^{\lambda} y$ implies $x \succ^{\lambda} y$.*

Proposition 3 *An Event $E \subseteq \Omega$ is assumed by λ if and only if it is assumed under $\succsim^{\lambda}$.*

Proof: If λ assumes E at level l, then for every $x, y, z \in \mathcal{A}$ and $q \leq l$,

$$\int_{\Omega} x(\omega) d\lambda^{q}(\omega) = \int_{\Omega} (x_E, z_{\Omega \backslash E})(\omega) d\lambda^{q}(\omega); \ \int_{\Omega} y(\omega) d\lambda^{q}(\omega) = \int_{\Omega} (y_E, z_{\Omega \backslash E})(\omega) d\lambda^{q}(\omega). \tag{A}$$

Moreover, for every $q > l$,

$$\int_{\Omega} (x_E, z_{\Omega \backslash E})(\omega) d\lambda^{q}(\omega) = \int_{\Omega} (y_E, z_{\Omega \backslash E})(\omega) d\lambda^{q}(\omega).$$

Thus, if $x \succ_E^{\lambda} y$, there exists $p \leq l$ such that for every $q < p$

$$\begin{aligned} \int_{\Omega} (x_E, z_{\Omega \backslash E})(\omega) d\lambda^{p}(\omega) \quad &> \quad \int_{\Omega} (y_E, z_{\Omega \backslash E})(\omega) d\lambda^{p}(\omega); \\ \int_{\Omega} (x_E, z_{\Omega \backslash E})(\omega) d\lambda^{q}(\omega) \quad &= \quad \int_{\Omega} (y_E, z_{\Omega \backslash E})(\omega) d\lambda^{q}(\omega). \end{aligned}$$

Then by A, $x \succ^{\lambda} y$ (strict determination).

Moreover, for every $s_{-i} \in S_{-i}$, if $E \cap (\{s_{-i}\} \times T_{-i}) \neq \emptyset$, then there exists $q \leq l$ such that $\lambda^{q}(E \cap U) > 0$. So, for every $x \in \mathcal{A}$ taking a strictly positive constant value on $E \cap U$ and the null act $y \in \mathcal{A}$, $x \succ_{E \cap U}^{\lambda} y$ (relevance).

If λ does not assume E, one of the following holds.

1. There exists $\overline{s}_{-i} \in \text{Proj}_{S_{-i}} E$ such that $\lambda^{m}(E \cap (\{\overline{s}_{-i}\} \times T_{-i})) = 0$ for every $m \leq k$.[16] This implies that $x \sim_{E \cap (\{\overline{s}_{-i}\} \times T_{-i})}^{\lambda} y$ for all $x, y \in \mathcal{A}$, violating relevance.

[16] A LPS $\lambda = (\lambda^{1}, ..., \lambda^{k})$ such that $\lambda^{p}(E) = 0$ for every $p \leq k$ falls into this case.

2. There exist $p \leq k$ and $q \geq p$ such that $\lambda^p(E) \neq 1$ and $\lambda^q(E) \neq 0$. Take events $E_{1,...,}E_k$ that verify the definition of mutual singularity for λ. Take $x, y \in \mathcal{A}$ such that

$$\begin{array}{rcll} x(\omega) & = & 0,\ y(\omega) = 1, & \text{if } \omega \in E_p \backslash E \\ x(\omega) & = & \lambda^p(E_p \backslash E)/2,\ y(\omega) = 0, & \text{if } \omega \in E_q \bigcap E \\ x(\omega) & = & y(\omega) = 0, & \text{else.} \end{array}$$

Thus $x \succ_E^\lambda y$ but $y \succ^\lambda x$, violating strict determination. ■

4.2 Proofs of measurability of relevant sets

Endow $\mathcal{M}(\Omega)$ with the weak-* topology. Then, for every event $E \subseteq \Omega$, the set of probability measures μ such that $\mu(E) = p$ for $p \in Q \cap [0,1]$ is Borel in $\mathcal{M}(\Omega)$. The weak-* topology is induced by the Prokhorov metric. Call $p(\mu, \mu')$ the Prohorov distance between $\mu, \mu' \in \mathcal{M}(\Omega)$. For each $k \in \mathbb{N}$, endow $\mathcal{N}_k(\Omega)$ with the product topology, $\mathcal{N}(\Omega)$ with the disjoint union topology and $\mathcal{L}_k(\Omega)$ and $\mathcal{L}(\Omega)$ with the relative topologies. Such topologies are induced by the following metric: if $\lambda_i, \lambda_i' \in \mathcal{N}(\Omega)$ have different lenghts, their distance $\pi(\lambda, \lambda')$ is 1; if $\lambda_i, \lambda_i' \in \mathcal{N}_k(\Omega)$ for some k, $\pi(\lambda, \lambda') = \max_{q \leq k} p(\lambda_i^q, \lambda_i'^q)$. This topology is the one employed by BFK for their type structures and it is the one arising naturally in the canonical structure of CDV from the construction. Therefore, the following results will apply to the canonical structure.

Lemma 1 *Fix $s_i \in S_i$. The set $\mathcal{B}^{s_i}$ of all $\lambda \in \mathcal{L}(\Omega)$ such that s_i is a lexicographic best reply to $\overline{\text{marg}}_{S_{-i}}\lambda$ is Borel in $\mathcal{L}(\Omega)$.*

Proof: Fix $s_i' \neq s_i$. Define

$$\begin{array}{rcl} O_{s_i,s_i'}^{\mathcal{W}} & : & = \left\{ \mu \in \mathcal{M}(\Omega) \left| \pi_i(s_i, \text{marg}_{S_{-i}}\mu) \geq \pi_i(s_i', \text{marg}_{S_{-i}}\mu) \right. \right\}, \\ O_{s_i,s_i'}^{\mathcal{S}} & : & = \left\{ \mu \in \mathcal{M}(\Omega) \left| \pi_i(s_i, \text{marg}_{S_{-i}}\mu) > \pi_i(s_i', \text{marg}_{S_{-i}}\mu) \right. \right\}. \end{array}$$

Take a $\mu \in O_{s_i,s_i'}^{\mathcal{S}}$. There exists $\varepsilon > 0$ such that for every $\mu' \in \mathcal{M}(\Omega)$ with $p(\mu, \mu') < \varepsilon$, $\pi_i(s_i,\text{marg}_{S_{-i}}\mu') - \pi_i(s_i',\text{marg}_{S_{-i}}\mu') > 0$. So $O_{s_i,s_i'}^{\mathcal{S}}$ is open. Since $O_{s_i,s_i'}^{\mathcal{W}}$ is the complement of $O_{s_i',s_i}^{\mathcal{S}}$, it is closed.

The set $U_k^{s_i} = \mathcal{B}^{s_i} \cap \mathcal{L}_k(\Omega)$ can be expressed as

$$U_k^{s_i} = \bigcap_{s_i' \neq s_i} \left(O_{s_i,s_i'}^{\mathcal{S}} \times \mathcal{L}_{k-1}(\Omega) \cup \left(O_{s_i,s_i'}^{\mathcal{W}} \times O_{s_i,s_i'}^{\mathcal{S}} \times \mathcal{L}_{k-2}(\Omega) \right) \cup ... \cup \left(O_{s_i,s_i'}^{\mathcal{W}} \times O_{s_i,s_i'}^{\mathcal{W}} \times ... \times O_{s_i,s_i'}^{\mathcal{W}} \right) \right),$$

so it is Borel in $\mathcal{L}(\Omega)$. Finally, $\mathcal{B}^{s_i} = \cup_{k \in \mathbb{N}} U_k^{s_i}$, so it is Borel too. ■

Lemma 2 *The set $\mathcal{C}$ of all $\lambda \in \mathcal{L}(\Omega)$ such that $\overline{\text{marg}}_{S_{-i}}\lambda \in \mathcal{N}^+(S_{-i})$ is Borel.*

Proof: For every $s_{-i} \in S_{-i}$, the set $\mathcal{C}_k^{s_{-i}} := \mathcal{L}_k(\Omega) \setminus \{\lambda \in \mathcal{L}_k(\Omega) \,|\lambda^m (s_{-i} \times T_{-i}) = 0, \forall m \leq k\}$ is Borel by the product topology. Hence, $\mathcal{C} = \cup_{k \in \mathbb{N}} \cap_{s_{-i} \in S_{-i}} \mathcal{C}_k^{s_{-i}}$ is Borel. ■

Lemma 3 *Fix an event $E \subseteq \Omega$. The set $\mathcal{A}^E$ of all $\lambda \in \mathcal{L}(\Omega)$ that assume E is Borel.*

Proof: Fix $k \in \mathbb{N}$ and $l \leq k$. By the product topology, the sets

$$\begin{aligned} \mathcal{A}_{k,l}^1 &= \{\lambda \in \mathcal{L}_k(\Omega) \,|\lambda^m (E) = 1, \forall m \leq l\}, \\ \mathcal{A}_{k,l}^2 &= \{\lambda \in \mathcal{L}_k(\Omega) \,|\lambda^m (E) = 0, \forall m > l\}, \end{aligned}$$

are Borel in $\mathcal{L}_k(\Omega)$. Thus, $\mathcal{A}_{k,l}^1 \cap \mathcal{A}_{k,l}^2$ is the Borel set of all $\lambda \in \mathcal{L}_k(\Omega)$ satisfying conditions 1 and 2 of assumption for l.

Moreover, for every $s_{-i} \in \mathrm{Proj}_{S_{-i}}(E)$, the set

$$\mathcal{A}_{k,l}^{s_{-i}} = \{\lambda \in \mathcal{L}_k(\Omega) \,|\lambda^m (\{s_{-i}\} \times T_{-i}) = 0, \forall m \leq l\}$$

is Borel in $\mathcal{L}_k(\Omega)$. Notice that

$$\mathcal{A}_{k,l}^3 = \cap_{s_{-i} \in \mathrm{Proj}_{S_{-i}}(E)} \left(\mathcal{L}_k(\Omega) \diagdown \mathcal{A}_{k,l}^{s_{-i}}\right)$$

is the Borel set of all $\lambda \in \mathcal{L}_k(\Omega)$ satisfying condition 3 of assumption for l. Hence $\mathcal{A}_{k,l} = \mathcal{A}_{k,l}^1 \cap \mathcal{A}_{k,l}^2 \cap \mathcal{A}_{k,l}^3$ is the Borel set of all $\lambda \in \mathcal{L}_k(\Omega)$ that assume E at level l. Finally, $\mathcal{A}^E = \cup_{k \in \mathbb{N}} \cup_{l \leq k} \mathcal{A}_{k,l}$, so it is Borel too. ■

We have the following corollary:

Corollary 1 *C_i, R_i and $A_i(E)$ are Borel subsets of $S_i \times T_i$.*

Proof: $R_i = \cup_{s_i \in S_i} (\{s_i\} \times \beta_i^{-1}(\mathcal{B}^{s_i}))$, $C_i = S_i \times \beta_i^{-1}(\mathcal{C})$ and $A_i(E) = S_i \times \beta_i^{-1}(\mathcal{A}^E)$ are Borel in $S_i \times T_i$ by respectively Lemma 1, 2 and 3, plus the measurability of β_i. ■

So finally, for our characterization result we can claim the following

Lemma 4 *For each $n \geq 1$, R_i^n is Borel in $S_i \times T_i$.*

Proof: By Corollary 1, for every $i \in I$, the sets $R_i^1 = C_i \cap R_i$ and $A_i\left(R_{-i}^1\right)$ are Borel in $S_i \times T_i$. So, an easy induction argument shows that, for all $n \geq 1$, the sets $R_i^{n+1} = R_i^1 \cap \left(\cap_{m \leq n} A_i\left(R_{-i}^m\right)\right)$ are Borel in $S_i \times T_i$. ■

4.3 Denseness of non full-support LPS's

This is a mathematical result of more general interest, so we abstract from Ω to any infinite metric space (X, d).

Claim 1 *Fix $\varepsilon > 0$. Take a $\mu \in \mathcal{M}(X)$ such that $0 \leq \mu(O) < \varepsilon$ for some open $O \subset X$. There exists a $\nu \in \mathcal{M}(X)$ such that:*

1. $\nu(O) = 0$*;*

2. for every $Y \in \mathcal{B}(X)$*, if* $\mu(Y) = 0$*, then* $\nu(Y) = 0$*;*

3. $p(\mu, \nu) < \varepsilon$*.*

Proof. Take the measure $\nu \in \mathcal{M}(X)$ such that for every $Y \in \mathcal{B}(X)$,

$$\nu(Y) = (\mu(Y) - \mu(Y \cap O))/(1 - \mu(O)),$$

i.e. put probability 0 on O and rescale elsewhere. The measure is well defined and clearly satisfies 1 and 2, so it only remains to show 3. For any $Y \subset X$, let $Y^\gamma := \{x \in X : \exists y \in Y, d(x, y) < \gamma\}$ be the γ-neighbourhood of Y. The Prokhorov distance between $\mu, \nu \in \mathcal{M}(X)$ is defined as:

$$p(\mu, \nu) := \inf \{\gamma > 0 : \forall Y \in \mathcal{B}(X), \mu(Y) \leq \nu(Y^\gamma) + \gamma \wedge \nu(Y) \leq \mu(Y^\gamma) + \gamma\}.$$

For every $Y \in \mathcal{B}(X)$, $\mu(Y) < v(Y^\varepsilon) + \varepsilon$, since

$$\begin{gathered} \mu(Y) - v(Y^\varepsilon) < \mu(Y) - v(Y) = \\ = (\mu(O) - v(O)) + (\mu(Y \backslash O) - v(Y \backslash O)) - (\mu(O \backslash Y) - v(O \backslash Y)) = \\ = \mu(O) + (\mu(Y \backslash O) - (\mu(Y \backslash O)/(1 - \mu(O)))) + (-\mu(O \backslash Y)) \end{gathered}$$

and the first term is smaller than ε, while the second and the third are negative.

For every $Y \in \mathcal{B}(X)$, $\nu(Y) < \mu(Y^\varepsilon) + \varepsilon$, otherwise there would be $Y \in \mathcal{B}(X)$ such that

$$\nu(X) = \nu(Y) + \nu(X \backslash Y) \geq \mu(Y) + \varepsilon + \nu(X \backslash Y) > \mu(Y) + \varepsilon + \mu(X \backslash Y) - \varepsilon = 1,$$

where the strong inequality comes from $\mu(Y) - v(Y) < \varepsilon$, as discussed above.

So the two conditions are verified also for some $\eta < \varepsilon$, implying that $\varepsilon > p(\mu, \nu)$. ■

Claim 2 *Fix* $\varepsilon > 0$*. Take a* $\lambda = (\lambda^1, ..., \lambda^k) \in \mathcal{L}^+(X)$*. There exists an open* $O \subset X$ *such that for every* $q \leq k$*,* $0 \leq \lambda^q(O) < \varepsilon$*.*

Proof. Let e be the smallest integer strictly bigger than $1/\varepsilon$. Take a set $Y_1 = \{y^1, ..., y^{ke}\}$ of $k \cdot e$ distinct points in X and call $\delta := \min_{w \leq ke, q \neq w} d(y^w, y^q)$. Take a set of open balls $O_1, ..., O_{ke}$ with radius $\delta/2$ centered in $y^1, ..., y^{ke}$. For every $q \leq k$, call c_q the cardinality of the set $W_q := \{w \leq ke : \lambda^q(O_w) \geq \varepsilon\}$. All open balls are pairwise disjoint, so it must hold $\sum_{w \in W_q} \lambda^q(O_w) \leq 1$. Then, $c_q < e$. So, there must exist $b \leq ke$ such that for every $q \leq k$, $b \notin W_q$. ■

Theorem 2 *Fix* $\varepsilon > 0$*. Take a* $\lambda = (\lambda^1, ..., \lambda^k) \in \mathcal{L}^+(X)$*. There exists* $\lambda' \in \mathcal{L}(X) \backslash \mathcal{L}^+(X)$ *such that* $\pi(\lambda, \lambda') < \varepsilon$*.*

Proof. By Claim 2, there exists an open $O \subset X$ such that $0 \leq \lambda^q(O) < \varepsilon$ for every $q \leq k$. Then, by Claim 1, for every $q \leq k$, we can take a $\nu^q \in \mathcal{M}(X)$ such that (1) $\nu^q(O) = 0$, (2) for every $Y \in \mathcal{B}(X)$, if $\lambda^q(Y) = 0$, then $\nu^q(Y) = 0$, and (3) $p(\lambda^q, \nu^q) < \varepsilon$. By mutual singularity, there exist sets $E_1, ...E_k$ in X such that for every $p \leq k$, $\lambda^p(E_p) = 1$ and $\lambda^p(E_q) = 0$ for $q \neq p$. Thus, by 2, for every $p \leq k$, $\nu^p(X \backslash E_p) = 0$ and $\nu^p(E_q) = 0$ for $q \neq p$. Hence, $\lambda' = (\nu^1, ..., \nu^k) \in \mathcal{L}(X)$. Moreover, by 3 $\pi(\lambda, \lambda') = \max_{q \leq k} p(\lambda^q, \nu^q) < \varepsilon$. Finally, $\lambda' \in \mathcal{L}^+(X)$ because by 1, for every $q \leq k$, $\nu^q(O) = 0$. ■

References

[1] Asheim, G., Dufwenberg, M., "Admissibility and common belief", *Games and Economic Behavior*, **42**, 2003, 208-234.

[2] Barelli, P., Galanis, S., "Admissibility and event-rationality", *Games and Economic Behavior*, **77**, 2013, 21-40.

[3] Battigalli, P., *Restrizioni Razionali su Sistemi di Probabilità Soggettive e Soluzioni di Giochi ad Informazione Completa*, EGEA, Milan, 1993.

[4] Battigalli, P., and M. Siniscalchi, "Strong Belief and Forward Induction Reasoning," *Journal of Economic Theory*, **106**, 2002, 356-391.

[5] Blume, L., Brandenburger, A., Dekel, E., "Lexicographic probabilities and equilibrium refinements", *Econometrica*, **59**, 1991, 81-98.

[6] Brandenburger, A., Friedenberg, A., Keisler, J., "Admissibility in games", *Econometrica*, **76**, 2008, 307-352.

[7] Catonini, E., De Vito, N., "Hierarchies of lexicographic beliefs", working paper, 2014.

[8] Dekel, E., Friedenberg, A., Siniscalchi, S., "Lexicographic beliefs and assumption", working paper, 2014.

[9] Heifetz, A., Meier, M., Schipper, B., "Comprehensive rationalizability", working paper, 2013.

[10] Keisler, J., Lee, B., "Common assumption of rationality", working paper, 2013.

[11] Lee, B., "Conditional Beliefs and Higher-Order Preferences", working paper, 2013.

[12] Pearce, D., "Rational Strategic Behavior and the Problem of Perfection," *Econometrica*, **52**, 1984, 1029-1050.

[13] Samuelson, L., "Dominated Strategies and Common Knowledge", Games and Economic Behavior, **4**, 1992, 284-313.

[14] Stahl, D., "Lexicographic rationalizability and iterated admissibility", *Economic Letters*, **47**, 1995, 155-159.

[15] Yang, C., "Weak assumption and iterative admissibility", working paper, 2014.