

# Bounded Reasoning and Higher-Order Uncertainty

Willemien Kets\*

JOB MARKET PAPER

November 15, 2010

## Abstract

Harsanyi type structures, the device traditionally used to model players' beliefs in games, generate infinite hierarchies of beliefs. Can the standard framework nevertheless be used to model situations in which players potentially have a finite depth of reasoning? This paper extends the Harsanyi framework to allow for higher-order uncertainty about players' depth of reasoning. The basic principle is that players with a finite depth of reasoning cannot distinguish states that differ only in players' beliefs at high orders. I apply the new framework to the electronic mail game of Rubinstein (1989). Coordination on the Pareto-efficient action is possible when there is higher-order uncertainty about players' depth of reasoning, unlike in the standard case, provided that one player thinks it is sufficiently likely that the other player has a finite (though potentially very high) depth of reasoning. Finally, I construct a type space that allows for bounded reasoning that contains the universal type space (which generates all infinite belief hierarchies) as a subspace, showing that the present framework fully generalizes the Harsanyi formalism.

---

\*U.C. Irvine and Santa Fe Institute. E-mail: willemien.kets@gmail.com. Phone: +1-505-204 8012. This paper supersedes Kets (2009). I am grateful to Adam Brandenburger, Yossi Feinberg, and Matthew Jackson for their guidance and support. This paper owes much to the input of Adam Brandenburger, Alfredo Di Tillio, Amanda Friedenberg, and Aviad Heifetz. I thank Songzi Du, Matt Elliott, Yossi Feinberg, Alex Frankel, Ben Golub, Joe Halpern, Yair Livne, Martin Meier, Rosemarie Nagel, Andrés Perea, Tomasz Szadzik, Burkhard Schipper, Elias Tsakas, and Takuro Yamashita for stimulating discussions and helpful comments and suggestions. Much of this research was carried out during visits to Stanford University and NYU Stern School of Business, and I thank these institutions for their hospitality.

# 1 Introduction

In many economic contexts, an agent’s payoff depends on his own action, the actions of others, and perhaps some unknown state of nature. A rational player chooses a best response to whatever belief she may have about the actions of the other players, that is, her first-order beliefs. But, if she expects that the others’ choice depend on their beliefs about their opponents’ actions (as in the case when her opponents’ are rational), then she also has to think about her opponents’ first-order beliefs—and so on, to arbitrarily high orders (cf. Harsanyi, 1967–1968).

Such higher-order beliefs can affect economic conclusions. Higher-order uncertainty about the seller’s information, for example, can lead to a breakdown of the Coase conjecture (Feinberg and Skrzypacz, 2005), and whether a game has multiple equilibria may depend on players’ higher-order beliefs (Carlsson and van Damme, 1993). Also, full rent extraction in auctions is possible under certain forms on higher-order uncertainty but not others (Neeman, 2004).<sup>1</sup>

How to model these beliefs? Strategic models traditionally use Harsanyi type structures. Each type in a Harsanyi type structure is associated with a belief about the state of nature and the other players’ types. Each type thus generates an infinite hierarchy of beliefs: Since each type has a belief about the state of nature, i.e., a *first-order belief*, a type’s belief over the other players’ types induces a belief both about the state of nature and about the other players’ beliefs about the state of nature, that is, a *second-order belief*. Following this logic, we see that each type induces a *k*th-order belief for any *k*, so that each type can be viewed as having an infinite depth of reasoning.

But while the logic of the Harsanyi framework requires that a player’s reasoning about other players’ beliefs continues indefinitely, limited cognitive resources generally force it to stop at some finite order. And, even if a player can reason about others’ beliefs at (sufficiently) high orders, she may be uncertain whether other players can do so, whether other players believe that their opponents can do so, and so on. That is, players’ depth of reasoning can be bounded, and there is higher-order uncertainty about players’ reasoning abilities.

How does such higher-order uncertainty affect rational behavior? As every Harsanyi type generates an infinite hierarchy of belief, this question can only be addressed in a richer framework.<sup>2</sup> This paper therefore develops a framework that explicitly allows for the possibility that players have a bounded depth of reasoning. The basic idea is that a player who does

---

<sup>1</sup>Results like these have stimulated a large literature that studies the robustness of predictions to perturbations of higher-order beliefs (e.g., Monderer and Samet, 1989; Kajii and Morris, 1997; Friedenberg and Meier, 2007; Weinstein and Yildiz, 2007). Higher-order beliefs also play a central role in the epistemic characterization of solution concepts; see, e.g., Aumann and Brandenburger (1995).

<sup>2</sup>Also see the discussion in Section 6.2.

not reason about other players' beliefs at very high orders cannot distinguish states of the world that differ only in others' beliefs at these orders. This approach, which extends the notion of “small worlds” of Savage (1954) to strategic settings, naturally leads to a class of type structures, called *extended type structures*, that contains the Harsanyi type structures.

More specifically, each type in an extended type structure is associated with a belief (probability measure) about the state of nature and the types of other players, as in the Harsanyi framework, but the beliefs of different types can now be defined on different  $\sigma$ -algebras.<sup>3</sup> A  $\sigma$ -algebra is the collection of events a type can form beliefs about, so if a type's belief about the other players' types is defined on a coarse  $\sigma$ -algebra, then the type cannot make very fine distinctions when it comes to others' beliefs.

As an illustration, suppose that there are two players, Ann and Bob. Suppose the belief of a type  $t^a$  for Ann over Bob's types is defined on the (trivial)  $\sigma$ -algebra that does not distinguish any types for Bob (i.e., the type's  $\sigma$ -algebra is  $\{T^b, \emptyset\}$ , where  $T^b$  is the set of types for Bob). In that case, the only aspects of Bob's beliefs the type  $t^a$  can form beliefs about are those that are common to all types for Bob. In particular, if there is uncertainty about Bob's belief about the state of nature, then  $t^a$  cannot form beliefs about any nontrivial first-order belief for Bob, so that the type does not have second-order beliefs. On the other hand, if the  $\sigma$ -algebra for  $t^a$  distinguishes the types for Bob that have different beliefs over the state of nature, then  $t^a$  can reason about Bob's first-order beliefs, so that it has a well-articulated second-order belief. Generally, a type's depth of reasoning equals  $k < \infty$  if the type has a well-articulated  $k$ th-order belief, but it cannot distinguish types for the other player that differ only in their beliefs at higher orders.

While this extends the Harsanyi framework in a natural way, it does not yet allow us to analyze the effects of players' limited depth of reasoning in the same straightforward way as the implications of players' higher-order beliefs can be studied in the Harsanyi framework. The Harsanyi framework provides a description of players' higher-order beliefs that is entirely *implicit*—i.e., does not make reference to players' higher-order beliefs—whereas the depth of reasoning of a type is inherently a property of the belief hierarchy it induces. This suggests that the framework can retain the tractability of the Harsanyi formalism only if the depth of types can be characterized without making reference to the belief hierarchies.

The methodological contribution of the paper is that it provides such a characterization. Theorem 4.3 shows that the depth of reasoning for a type can be determined using a simple iterative procedure. Roughly, a type has a high depth of reasoning if the type can form beliefs

---

<sup>3</sup>A  $\sigma$ -algebra on a set  $X$  is a collection of subsets of  $X$  that contains  $X$  itself and is closed under complements and countable unions. If a probability measure  $\mu$  is defined on a particular  $\sigma$ -algebra  $\Sigma$  on  $X$ , then  $\mu$  can only assign a probability to subsets of  $X$  that belong to  $\Sigma$ .

about the beliefs of other players defined on  $\sigma$ -algebras that correspond to a high depth of reasoning. The depth of types is thus defined recursively, starting with the types associated with the coarsest  $\sigma$ -algebras. This means that the depth of a type can be determined from the type structure alone, without explicitly describing the belief hierarchies.

Just like the Harsanyi framework makes it possible to model players' infinite belief hierarchies implicitly, by specifying types and beliefs about types, the current framework therefore provides an implicit description of players' finite and infinite hierarchies of beliefs, including higher-order uncertainty about others' depth of reasoning, by specifying types, beliefs about types, and a collection of  $\sigma$ -algebras on each set of types. Theorem 5.2 makes precise the sense in which the current formalism generalizes the Harsanyi framework. It shows that there is a structure that allows for bounded reasoning which embeds every Harsanyi type structure. In particular, it contains the universal type space of Mertens and Zamir (1985) which generates all infinite hierarchies of beliefs. This establishes that the extended type structures fully generalize the Harsanyi framework.

I apply the new framework to the electronic mail game of Rubinstein (1989). In the electronic mail game, two players exchange messages about their beliefs about the payoffs in a coordination game for a finite number of rounds, so that each player knows the payoffs, knows that each player knows the payoffs, and so on, for a potentially large but finite number of iterations. Rubinstein shows the counterintuitive result that the Pareto-efficient action is not rationalizable when that action is risk-dominated, regardless of the number of messages sent.

However, this result depends critically on the assumption that players are able to make complicated inferences the higher-order beliefs of the other player. I show that if a type assigns sufficiently high probability to the event that the other player has a finite depth of reasoning, then there is a finite number of messages  $M$  such that players can rationally choose the Pareto-efficient action whenever at least  $M$  messages have been sent, even if both players in fact have an infinite depth of reasoning.<sup>4</sup> The intuition is precisely that if a player believes it is sufficiently likely that the other player does not make arbitrarily complex deductions, then he may believe that the other player is willing to choose the Pareto-efficient action, so that it is rational for him to choose the Pareto-efficient action as well, even if he has an infinite depth of reasoning.

The framework developed here is complementary to models of bounded reasoning that have been used to explain strategic behavior in a wide range of experiments, such as cognitive-hierarchy models and models of level- $k$  reasoning.<sup>5</sup> In these models, each player is endowed

---

<sup>4</sup>Stinchcombe (1988), Jehiel and Koessler (2008) and Strzalecki (2009) establish similar results; see the discussion below.

<sup>5</sup>See, e.g., Nagel (1995), Stahl and Wilson (1995) Ho et al. (1998), Costa-Gomes et al. (2001),

with a set of “cognitive types.” A cognitive type specifies the player’s beliefs about other players’ cognitive types, where each player believes that other players have a lower cognitive type than he does. Each cognitive type plays a best response to its beliefs about the actions of other players, except the so-called level-0 type, which have no beliefs.

An important difference between level- $k$  models and the present approach is that there is no clear separation between types and decision rules in level- $k$  models. These models therefore do not provide a framework to describe players’ belief hierarchies per se, unlike extended type structures, but rather offer a behavioral model based on intuitive assumptions on players’ reasoning processes. This makes it difficult to separate the implications of higher-order uncertainty from the effects of different behavioral assumptions in these models.<sup>6</sup>

Level- $k$  models thus provide a framework that offers an intuitive explanation for experimental behavior in a large class of games, while the present formalism makes it possible to isolate the implications of higher-order uncertainty about others’ depth of reasoning within a framework that is otherwise completely standard, without committing to a specific behavioral model.

The next section provides an informal discussion of the framework and applies it to the electronic mail game; the formal treatment begins in Section 3. The related literature is discussed in Section 6, and Section 7 concludes.

## 2 Heuristic treatment

### 2.1 Unbounded reasoning

Higher-order uncertainty can affect what actions a player can rationally choose, even if that is not always intuitive. As an illustration, consider the following example, based on the electronic mail game of Rubinstein (1989).<sup>7</sup> Ann (denoted  $a$ ) and Bob ( $b$ ) play a coordination game with two actions,  $A$  (“attack”) and  $N$  (“not attack”), with the payoffs to each action given in Figure 2.1(a). There is uncertainty about payoffs. The payoff  $\theta$  of coordinating on  $A$  is either smaller or greater than the payoff to the other action, taking the values  $\theta_1 = -2$  or  $\theta_2 = 1$ .

Players have a common prior which gives equal probability to both realizations of  $\theta$ . Bob

---

Camerer et al. (2004), and Crawford and Iriberri (2007). Strzalecki (2009), Rogers, Palfrey, and Camerer (2009), and Heifetz and Kets (2010) provide generalizations.

<sup>6</sup>In particular, level- $k$  models require the modeler to specify the behavior of the level-0 types; this is not necessary in the present framework, as no assumptions are made a priori about the relation between a player’s depth and his actions or his beliefs about others’ actions.

<sup>7</sup>I use the parametrization of Strzalecki (2009).

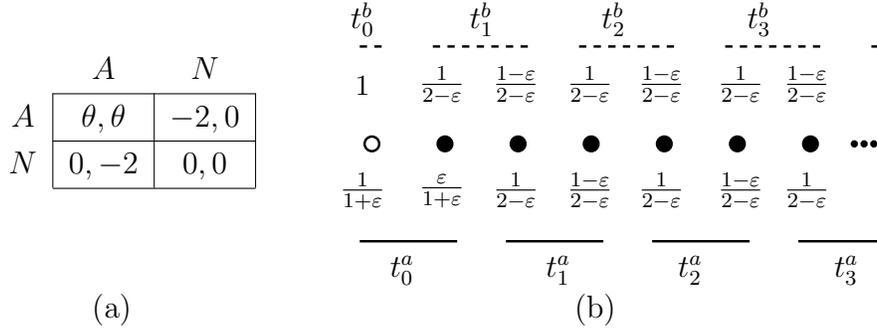


Figure 2.1: (a) The payoff matrix, where  $\theta$  is either  $\theta_1 = -2$  or  $\theta_2 = 1$ ; (b) The Harsanyi type structure. The solid and dashed lines represent the types for Ann and Bob, respectively, where  $t_\ell^i$  is the type for player  $i$  that has sent  $\ell$  messages. The white dot ( $\circ$ ) is the state of the world at which  $\theta = \theta_1$ , and the numbers are the posterior beliefs of each type. For example, types  $t_0^a$  and  $t_0^b$  assign probability  $\frac{1}{1+\varepsilon}$  and 1 to  $(\theta_1, t_0^a, t_0^b)$ , respectively.

knows the parameter, and if the parameter is  $\theta = \theta_2$ , he automatically sends a message to Ann to inform her. Ann then responds automatically to confirm the receipt of the message, and so on. Each message is lost with some small but positive probability  $\varepsilon$ . Each player knows the number of messages he has sent, but does not know whether the other player received his or her last message

This defines a Harsanyi type structure, denoted by  $\langle T_H^a, T_H^b, \beta_H^a, \beta_H^b \rangle$ , where  $T_H^a = \{t_0^a, t_1^a, \dots\}$  is the set of *types* for Ann, and type  $t_\ell^a$  has sent  $\ell$  messages; see Figure 2.1(b) for an illustration. The *belief map*  $\beta_H^a$  is a measurable function that assigns to each type for Ann a belief about the parameter  $\theta$  and Bob's type, so that  $\beta_H^a(t_\ell^a)$  is the *belief* of a type  $t_\ell^a$  for Ann over the payoff parameter and Bob's types, and similarly for Bob.

Each type generates an infinite hierarchy of beliefs: Type  $t_0^b$  for Bob, for instance, believes that  $\theta = \theta_1$ , which is the type's *first-order belief*. Moreover, type  $t_0^b$  believes that  $\theta = \theta_1$  and that Ann assigns probability  $1/(1 + \varepsilon)$  to  $\theta_1$ , which defines the type's *second-order belief*, and so on. Measurability of the belief maps implies that these hierarchies are well defined.

Suppose both players are rational, believe that the other is rational, and so on. What actions can they choose? Clearly, if Bob believes (with probability 1) that  $\theta = \theta_1 < 0$  (i.e., his type is  $t_0^b$ ), then the rational choice for him is to play  $N$ . But what if Bob has sent two messages (i.e., his type is  $t_2^b$ ), so that he believes that  $\theta = \theta_2$  and believes that Ann believes this, so that coordinating on  $A$  is Pareto-efficient? Is it rational for him to choose  $A$ ?

The answer is no. If Bob has sent (precisely) two messages, then he thinks it is likely that Ann thinks it is likely that he thinks it is likely that Ann thinks it is likely that  $\theta = \theta_1$ . In that case, he thinks it is likely that Ann thinks it is likely that he thinks it is likely that Ann chooses  $N$  (which is the rational choice if she thinks it is likely that  $\theta = \theta_1$ ), and therefore,

he thinks it is likely that Ann thinks it is likely that he chooses  $N$ . Since  $N$  is the rational choice for Ann if she thinks it is likely that Bob chooses  $N$ , type  $t_2^b$  for Bob thinks it is likely that Ann plays  $N$ , in which case the unique rational choice for him is to play  $N$ . Indeed, the same argument applies no matter how many messages have been sent:

**Theorem (Rubinstein, 1989)** *The unique rationalizable action for each player is  $N$ , regardless of the number of messages that have been exchanged.*

Rubinstein calls this result “paradoxical,” and suggests that the problem lies in the fact that players have well-articulated beliefs at all orders. Can the Harsanyi framework be extended to allow reasoning to stop at some finite order? And can this lead to more intuitive predictions?

## 2.2 Small worlds

To answer these questions, we need to think about what it means if a player does not form a belief about other players’ higher-order beliefs. In the context of single-person decision theory, Savage (1954) suggested that a decision-maker who does not reason about all aspects of a situation reasons within a small world, where “a state of [a small] world corresponds not to one state of [a larger world], but to a *set of states*” (Savage, 1954, p. 9, *emph. added*). In that case, the decision-maker has a coarse perception of the choice situation, as he does not distinguish states of the world that differ only in details he does not reason about.

In the current strategic context, the interest is in the perception of higher-order beliefs. A coarse perception of a player’s higher-order beliefs corresponds to a coarse perception of his type, where a coarse perception can be modeled using  $\sigma$ -algebras, i.e., the collection of events a player can form beliefs about. If a type for Ann, for example, does not reason about Bob’s beliefs at some order, then that type’s belief is defined on a  $\sigma$ -algebra that does not separate types for Bob that induce belief hierarchies which differ only at high orders.

## 2.3 Extended type structures

The above discussion suggests that higher-order uncertainty about players’ depth of reasoning can be modeled with a type structure in which the beliefs of types of a given player are possibly defined on different  $\sigma$ -algebras. As before, suppose that there are two players, Ann ( $a$ ) and Bob ( $b$ ), and let  $\Theta$  be the set of states of nature.

Then, an *extended type structure* is a structure

$$\langle T^a, T^b, \mathcal{X}^a, \mathcal{X}^b, (\Sigma_k^a)_{k \in \mathbb{N} \cup \{\infty\}}, (\Sigma_k^b)_{k \in \mathbb{N} \cup \{\infty\}} \rangle, \quad (2.1)$$

where  $T^a$  and  $T^b$  are the sets of types for Ann and Bob, which can be taken to be countable here, and  $\Sigma_k^a$  and  $\Sigma_k^b$  are  $\sigma$ -algebras on  $T^a$  and  $T^b$ , respectively, for each  $k \in \mathbb{N} \cup \{\infty\}$ . Assume that  $\Sigma_1^a$  is the trivial  $\sigma$ -algebra on the type set  $T^a$ , and that  $\Sigma_\infty^a$  is the  $\sigma$ -algebra containing all subsets of  $T^a$ , that is, its power set. Similarly,  $\Sigma_1^b$  and  $\Sigma_\infty^b$  are the trivial  $\sigma$ -algebra on  $T^b$  and its power set. The set of states of nature  $\Theta$  can be taken to be countable for the purposes of this section, and is endowed with some  $\sigma$ -algebra  $\Sigma_\Theta$ .

The function  $\chi^a$  maps each of Ann's types  $t^a \in T^a$  to a probability space  $\chi^a(t^a) = (\Theta \times T^b, \Sigma^a(t^a), \beta^a(t^a))$ , where  $\Sigma^a(t^a) = \Sigma_\Theta \otimes \Sigma_k^b$  is a  $\sigma$ -algebra on the product space  $\Theta \times T^b$  for some  $k \in \mathbb{N} \cup \{\infty\}$ .<sup>8</sup> The function  $\chi^b$  is defined analogously. If  $\Sigma^a(t^a) = \Sigma_\Theta \otimes \Sigma_k^b$  for some type  $t^a$  for Ann, then  $\Sigma_k^b$  represents her *perception* of Bob's type. The perception of a type determines its depth of reasoning, as discussed below, that is, the order up to which the type can form beliefs. This means that a countable collection of  $\sigma$ -algebras on each type set suffices.

Finally, note that there need not be any type for Ann with a particular perception  $\Sigma_k^b$ , and similarly for Bob. This means that the assumption that the collection of  $\sigma$ -algebras on a type set includes the trivial  $\sigma$ -algebra and the power set is in fact without loss of generality. The role of such "virtual" perceptions is discussed below.

Extended type structures satisfy a number of conditions (beyond those listed here) that ensure that the  $\sigma$ -algebras on the types represent possible perceptions of higher-order beliefs. If that is the case, then the depth of reasoning of a type can be determined directly from the relations between the  $\sigma$ -algebras, as discussed next.

## 2.4 Depth of reasoning

How does a type's perception determine its depth of reasoning? It will be instructive to focus on a subclass of structures where the intuition is especially clear.

The notion of dominance will be useful in defining this class of structures. Say that a  $\sigma$ -algebra  $\Sigma_k^i$  on the type set  $T^i$  of player  $i$  *dominates* the  $\sigma$ -algebra  $\Sigma_m^j$  on the type set  $T^j$  of the other player, denoted  $\Sigma_k^i \succ \Sigma_m^j$ , if for every event  $E$  in  $\Sigma_\Theta \otimes \Sigma_m^j$  and  $p \in [0, 1]$ , it holds that

$$\{t^i \in T^i : E \in \Sigma^i(t^i), \beta^i(t^i)(E) \geq p\} \in \Sigma_k^i.$$

That is, if the perception  $\Sigma_k^a$  dominates  $\Sigma_m^b$ , for example, then Bob can form beliefs about the beliefs for Ann about his type that can be expressed in terms of the events in  $\Sigma_m^b$  whenever his perception of Ann's type is at least as fine as  $\Sigma_k^a$ .

---

<sup>8</sup>A probability space is a triple  $(X, \Sigma, \mu)$ , where  $X$  is some set,  $\Sigma$  is a  $\sigma$ -algebra on  $X$ , and  $\mu$  is a probability measure defined on  $\Sigma$ .

We can now define the class of extended type structures we focus on here. Consider an extended type structure such that the perception  $\Sigma_{k+1}^a$  is the coarsest perception that dominates  $\Sigma_k^b$ , for every  $k \in \mathbb{N}$ . Assume that there is no  $k < \infty$  such that  $\Sigma_\ell^a = \Sigma_k^a$  for all  $\ell \geq k$  (the formal treatment allows for the general case). Likewise for the  $\sigma$ -algebras  $(\Sigma_k^b)_k$  on the type set  $T^b$ . Note that all perceptions are well defined, since the perceptions  $\Sigma_1^a$  and  $\Sigma_1^b$  are fixed to be the trivial  $\sigma$ -algebras.

What is the depth of reasoning of a type in such a structure? To answer that question, we of course need to define the depth of reasoning of a type. At this point it is not possible to give a precise definition, but the idea is intuitive. Informally, a type has an infinite *depth (of reasoning)* if the type induces a  $k$ th-order belief for every  $k$ ; a type has a finite depth  $k$  if induces a  $k$ th-order belief, but no belief at higher orders.

While the depth of a type is defined in terms of the belief hierarchy it induces, it can be characterized in terms of the coarseness of its perception of others' types. This is done using the *rank* of perceptions. Start with the trivial  $\sigma$ -algebras. If a perception  $\Sigma_m^b$  (of Bob's types) Ann does not dominate the trivial  $\sigma$ -algebra on  $T^a$ , then its rank equals 1; otherwise it is at least 2. Likewise with  $a$  and  $b$  interchanged. In the next step, the perception with rank at least 2 can be bounded further by checking whether it dominates a perception that dominates the trivial  $\sigma$ -algebra (i.e., a perception of rank at least 2), and so on. The rank of the power sets  $\Sigma_\infty^a$  and  $\Sigma_\infty^b$  is defined to be infinite. Given the way the procedure is set up, it is straightforward to show that two perceptions (on a given type set) have the same rank if and only if they are identical (Lemma 3.1).

Since the interest is in the depth of types, rather than the properties of perceptions per se, one might ask why this procedure orders perceptions, as opposed to directly ordering the types. The reason is that all perceptions are important for determining the depth of types, as discussed below, including the “virtual” perceptions, that is, perceptions such that there is no type with that perception. Applying a similar procedure to the types, therefore, does not always yield the types' depth of reasoning.

There is a direct relation between the depth of a type and the rank of its perception (Proposition 3.2 and Theorem 4.3):

*If the rank of the perception of a type equals  $k$ , then the type has depth  $k$ .*

The ordering of the perceptions in terms of their ranks thus translates into an ordering of the types in terms of their depth. The value of this result is that it characterizes the depth of a type—that is, a property of the belief hierarchy it induces—in terms of a property of its perception, i.e., a  $\sigma$ -algebra. The rank of a type's perception, in turn, is determined by the

rank of the perceptions it dominates, so that a perception has a high rank if it dominates perceptions of a high rank. Starting from the coarsest  $\sigma$ -algebras  $\Sigma_1^a$  and  $\Sigma_1^b$  on  $T^a$  and  $T^b$ , respectively, it is therefore possible to derive the depth of any type using a simple iterative procedure based on the dominance relations. The virtual perceptions thus serve to give the other perceptions the appropriate rank.

The intuition behind the result is the following. Recall that the perceptions  $\Sigma_1^a$  and  $\Sigma_1^b$  are the trivial  $\sigma$ -algebras on  $T^a$  and  $T^b$ , respectively. Assume for simplicity that  $\Sigma_1^a$  does not dominate  $\Sigma_1^b$ , and  $\Sigma_1^b$  does not dominate  $\Sigma_1^a$  (the formal treatment considers the general case). By construction, the  $\sigma$ -algebra  $\Sigma_2^a$  of Ann's types contains all subsets of types for Ann that can be characterized in terms of their belief over the set of states of nature  $\Theta$  (as  $\Sigma_2^a \succ \Sigma_1^b$ ). A type for Bob with perception at least as fine as  $\Sigma_2^a$  can thus form beliefs about Ann's first-order belief, so that its depth is at least 2. Likewise for  $\Sigma_2^b$ . Going further, the  $\sigma$ -algebra  $\Sigma_3^a$  contains all subsets of types for Ann that can be characterized in terms of their beliefs over the set of states of nature and Bob's beliefs about the state of nature (since  $\Sigma_3^a$  dominates  $\Sigma_2^b$ , and  $\Sigma_2^b$  dominates  $\Sigma_3^a$ ). A type for Bob with perception at least as fine as  $\Sigma_3^a$  can therefore reason about Ann's second-order belief, so that the type has depth at least 3. Again, likewise for  $\Sigma_3^b$ . A perception of the type set  $T^a$  thus groups together types for Ann that coincide in their belief up to some fixed order, and similarly for a perception of  $T^b$ . This holds for the perceptions of actual types as well as for virtual perceptions.

In this subclass of extended type structures, there is a direct relation between the perceptions of types and the perceptions of higher-order beliefs, as sketched above (though the intuition carries over to the general case). The result that relates the depth of a type and the rank of its perception, however, uses only the following properties:

**Filtrations:** For each  $k \in \mathbb{N}$ , it holds that  $\Sigma_{k+1}^a \supseteq \Sigma_k^a$  and  $\Sigma_{k+1}^b \supseteq \Sigma_k^b$ . Moreover,  $\Sigma_k^a \subseteq \Sigma_\infty^a$  for all  $k$ .

**Strict refinements and dominance:** If  $\Sigma_{k+1}^a$  is a strict refinement of  $\Sigma_k^a$ , then there is a  $\sigma$ -algebra  $\Sigma_m^b$  that is undominated by  $\Sigma_k^a$  such that  $\Sigma_{k+1}^a$  dominates  $\Sigma_m^b$ . Likewise with  $a$  and  $b$  interchanged.

**Ordering:** Suppose  $\Sigma_k^a \subsetneq \Sigma_\infty^a$ . If  $\Sigma_m^b$  dominates  $\Sigma_k^a$ , then  $\Sigma_k^a$  does not dominate  $\Sigma_m^b$ . Moreover, if  $\Sigma_m^b$  dominates every perception  $\Sigma_k^a$  that is a proper subset of the power set  $\Sigma_\infty^a$ , then  $\Sigma_k^a$  dominates  $\Sigma_\infty^b$ . Similarly with  $a$  and  $b$  interchanged.

**Mutual dominance:** The perception  $\Sigma_\infty^a$  dominates  $\Sigma_\infty^b$  and vice versa.

Every extended type structure satisfies these properties (by definition; see Section 3). Theorem 4.3 shows that these properties are sufficient to determine the depth of types also for

the general case. Importantly, the properties refer only to the  $\sigma$ -algebras on the type sets; they do not refer to the hierarchies of beliefs induced by the types. Proposition 5.1 and Theorem 5.2 characterize the relation between Harsanyi type structures and extended type structures, showing that the present framework fully generalizes the Harsanyi formalism.

Why are the properties listed above natural in the current context? The first property says that the  $\sigma$ -algebras on a given type set form a *filtration*: each  $\sigma$ -algebra is either coarser or finer than another one. This is an intuitive property, as the  $\sigma$ -algebras on a type set represent perceptions of higher-order beliefs in this framework. Since there is a complete order on the perceptions of higher-order beliefs (that is, the depth of belief hierarchies), it is reasonable if there is a complete order on the  $\sigma$ -algebras as well.<sup>9</sup> The second property, in conjunction with the first, says that a finer perception dominates a strictly larger set of perceptions than a coarser one. These two properties are the essential ingredients for showing that there is a one-to-one relation between the  $\sigma$ -algebras on a type set and the rank (Lemma 3.1).

The third property says that a type for Ann with a coarse perception  $\Sigma_m^b$  of Bob's type can only form beliefs about Bob's belief about Ann's belief that are expressible in a perception that is strictly coarser than  $\Sigma_m^b$  (as  $\Sigma_m^b \succ \Sigma_k^a$  implies that  $\Sigma_k^a$  does not dominate  $\Sigma_m^b$ ); and likewise for the types for Bob with a coarse perception. A type with a coarse perception therefore has a coarse perception of Bob's higher-order beliefs, as we would like.

The fourth property states that a type for Ann with perception  $\Sigma_\infty^b$  can form beliefs about Bob's beliefs over her types expressible in  $\Sigma_\infty^a$  (as  $\Sigma_\infty^b \succ \Sigma_\infty^a$ ), and therefore about his beliefs about her beliefs about his types expressible in  $\Sigma_\infty^b$  (as  $\Sigma_\infty^a \succ \Sigma_\infty^a$ ), and so on; likewise for types for Bob with perception  $\Sigma_\infty^a$ . This means that types with a fine perception of types have a fine perception of higher-order beliefs. Again, this is precisely what we want.

While these properties are natural when perceptions of types represent perceptions of higher-order beliefs, it is an open question whether these properties are necessary to determine the depth of reasoning of types. However, it seems that any structure that models higher-order uncertainty in this way should have the intuitive feature that  $\sigma$ -algebras group together types that coincide in their beliefs up to some order, as in the case considered here.

Now that we can determine the depth of reasoning of types, we can revisit the electronic mail game, to study the potential implications of higher-order uncertainty of others' depth of reasoning.

---

<sup>9</sup>This is not a logical necessity, however, as there can be redundant types. If there are redundant types in a structure, then this property says that for any pair of redundant types, there is some minimum rank  $r \leq \infty$  such that all perceptions of rank at least  $r$  distinguish the types.

## 2.5 The electronic mail game revisited

Recall that coordinating on the action  $A$  is not rationalizable in the electronic mail game if both players have an infinite depth of reasoning, believe that the other has an infinite depth of reasoning, and so on. How does that change when there is higher-order uncertainty about players' depth of reasoning?

### 2.5.1 Bounded reasoning

We will consider an extended type structure that is a slight modification of the Harsanyi type structure discussed in Section 2.1. In particular, one player has an infinite depth of reasoning in every state of the world, while the other player may have both a finite or infinite depth of reasoning. This means that there is higher-order uncertainty: even if both players have an infinite depth of reasoning, one of them may think it is possible that the other has a finite depth, or may believe that the other thinks this is possible, and so on.

The extended type structure is defined as follows. The type sets for Ann and Bob are

$$T_*^a := \{t_{\ell,k}^a : \ell = 0, 1, \dots, k \in \mathbb{N} \cup \{\infty\}\}$$

and

$$T_*^b := \{t_{\ell,\infty}^b : \ell \in \mathbb{N}\},$$

respectively. Type  $t_{\ell,\infty}^b$  is the type for Bob that has sent  $\ell$  messages; similarly,  $t_{m,k}^a$  is a type for Ann that has sent  $m$  messages, where  $k \in \mathbb{N} \cup \{\infty\}$ . The belief maps for Ann and Bob are denoted by  $\beta_*^a$  and  $\beta_*^b$ , respectively, and the  $\sigma$ -algebras on  $T_*^a$  and  $T_*^b$  are  $(\Sigma_{k,*}^a)_{k \in \mathbb{N} \cup \{\infty\}}$  and  $(\Sigma_{k,*}^b)_{k \in \mathbb{N} \cup \{\infty\}}$ .

Each type for Bob can distinguish all types for Ann. That is, the (marginal of) belief  $\beta_*^b(t_{\ell,\infty}^b)$  of each type  $t_{\ell,\infty}^b$  over Ann's types is defined on the power set  $\Sigma_{\infty,*}^a$  of  $T_*^a$ . The belief of Bob about the payoff parameter and number of messages sent by Ann (given the number of messages he has sent) is the same as under the original Harsanyi type structure. For example, type  $t_2^b$  for Bob believes that  $\theta = \theta_1$  (with probability 1) and assigns probability  $1/(2-\varepsilon)$  to the event that Ann has sent precisely one message. Generally, for every  $\ell$  and event  $E \subseteq \Theta \times T_*^a$ ,

$$\beta_*^b(t_{\ell,\infty}^b)(E) = \beta_H^b(t_\ell^b)(\{(\theta, t_m^a) \in \Theta \times T_H^a : (\theta, t_{m,k}^a) \in E \text{ for all } k\}).$$

Conditional on the event that Ann has sent  $m$  messages, where  $m = \ell - 1$  or  $m = \ell$ , type  $t_\ell^b$  assigns probability  $p_k$  to type  $t_{m,k}^a$ , where  $p_k$  does not depend on  $\ell$ . It will be convenient to assume that  $p_k > 0$  for all  $k$ . The specification of the other  $\sigma$ -algebras  $\Sigma_{1,*}^a, \Sigma_{2,*}^a, \dots$  on  $T_*^a$  is immaterial for now.

The types  $t_{\ell,k}^a$  for Ann that have sent the same number of messages  $\ell$  differ in the coarseness of their perception of Bob's types. The belief of type  $t_{\ell,k}^a$  on  $T_*^b$  is defined on the  $\sigma$ -algebra  $\Sigma_{k,*}^b$ , where

$$\begin{aligned}\Sigma_{1,*}^b &:= \{T_*^b, \emptyset\}, \\ \Sigma_{2,*}^b &:= \sigma\left(\{\{t_{0,\infty}^b\}, \{t_{1,\infty}^b, t_{2,\infty}^b, \dots\}\}\right), \\ \Sigma_{3,*}^b &:= \sigma\left(\{\{t_{0,\infty}^b\}, \{t_{1,\infty}^b\}, \{t_{2,\infty}^b, t_{3,\infty}^b, \dots\}\}\right), \\ &\dots\dots\dots \\ \Sigma_{k,*}^b &:= \sigma\left(\{\{t_{0,\infty}^b\}, \dots, \{t_{k-2,\infty}^b\}, \{t_{k-1,\infty}^b, t_{k,\infty}^b, \dots\}\}\right), \\ &\dots\dots\dots\end{aligned}$$

with  $\sigma(\mathcal{E})$  the  $\sigma$ -algebra generated by the collection of subsets  $\mathcal{E}$ .<sup>10</sup> As always, the  $\sigma$ -algebra  $\Sigma_{\infty,*}^b$  is the power set of  $T_*^b$ .

The beliefs for the types of Ann about the communication protocol are again the same as before. However, a type for Ann with a bounded depth of reasoning does not realize that the messages may affect Bob's beliefs at high orders. Formally, the belief of type  $t_{\ell,k}^a$  is the same as the belief of the Harsanyi type  $t_\ell^a \in T_H^a$  (under the appropriate transformation of Bob's type set) when beliefs are restricted to the  $\sigma$ -algebra  $\Sigma_{k,*}^b$ . That is, for every event  $E \in \Sigma_\Theta \otimes \Sigma_{k,*}^b$ ,

$$\beta_*^a(t_{\ell,k}^a)(E) = \beta_H^a(t_\ell^a)(\{(\theta, t_m^b) \in \Theta \times T_H^b : (\theta, t_{m,\infty}^b) \in E\}).$$

This means that the belief of type  $t_{\ell,k}^a$  coincides with that of  $t_\ell^a$  on a coarser  $\sigma$ -algebra  $\sigma$ -algebra (under the bijection  $t_\ell^b \mapsto t_{\ell,\infty}^b$ ). In particular, the belief of  $t_{\ell,\infty}^a$  coincides with that of  $t_\ell^a$  under this transformation of Bob's types.

It is important to note that the fact that type' beliefs in the extended type structure can be defined in terms of the beliefs of the types in the original Harsanyi structure does not mean that the higher-order beliefs of the types are close. In particular the type  $t_{\ell,\infty}^b$  believes it is possible that Ann has a finite depth of reasoning, unlike the type  $t_\ell^b$  in the Harsanyi type structure. This in turn means that the higher-order beliefs of the types of Ann in the extended type structure are different from those of her types in the Harsanyi type structure, and so on. Such differences in higher-order beliefs can lead to different predictions, as shown below. However, this requires specifying the depth of each type first.

Clearly, each type for Bob has an infinite depth, as it can form beliefs over the individual types of Ann. The depth of the types for Ann can be determined in an intuitive way by noting

---

<sup>10</sup>A  $\sigma$ -algebra on  $X$  that is generated by some collection  $\mathcal{E}$  of subsets of  $X$  is the smallest  $\sigma$ -algebra on  $X$  that contains the subsets in  $\mathcal{E}$ .

that the  $\sigma$ -algebras  $\Sigma_{2,*}^b, \Sigma_{3,*}^b, \dots$  on  $T_*^b$  make finer and finer distinctions between the types for Bob on the basis of their higher-order beliefs. For example, types  $t_{1,\infty}^b, t_{2,\infty}^b, \dots$  for Bob all have the same first-order belief (each type  $t_{\ell,\infty}^b$  assigns probability 1 to  $\theta_2$  for  $\ell \geq 1$ ), while  $t_{0,\infty}^b$  has a deviant first-order belief (the type assigns probability 1 to  $\theta_1$ ). This means that a type  $t_{\ell,k}^a$  for Ann whose  $\sigma$ -algebra  $\Sigma_{k,*}^b$  is at least as fine as  $\Sigma_{2,*}^b$  can form a belief about Bob's belief about  $\Theta$ , so that its depth is at least 2.

Similarly, type  $t_{1,\infty}^b$  has a different belief about Ann's belief about the payoff parameter than  $t_{2,\infty}^b, t_{3,\infty}^b, \dots$  ( $t_{1,\infty}^b$  thinks it is possible that Ann thinks it is possible that the payoff parameter is  $\theta_1$ , while  $t_{2,\infty}^b, t_{3,\infty}^b, \dots$  all assign zero probability to that event), so that type  $t_{1,\infty}^b$  has a different second-order belief than the types  $t_{2,\infty}^b, t_{3,\infty}^b, \dots$ . A type  $t_{\ell,k}^a$  for Ann with  $\sigma$ -algebra  $\Sigma_{3,*}^b$  or finer can thus reason about Bob's second-order belief, so that its depth is at least 3. The  $\sigma$ -algebras  $\Sigma_{1,*}^a, \Sigma_{2,*}^a, \Sigma_{3,*}^a, \dots$  (not specified here) similarly make finer and finer distinctions between the types for Ann on the basis of their higher-order beliefs (taking into account that the depth of certain types is bounded).

Determining the depth of Ann's types in this way requires one to carefully consider the higher-order beliefs of the types for Bob. However, the same result can be obtained by considering the dominance relations among  $\sigma$ -algebras. The iterative procedure outlined in the previous section gives a partitioning of the  $\sigma$ -algebras  $\Sigma_{\infty,*}^b$  and  $\Sigma_{1,*}^b, \Sigma_{2,*}^b, \dots$  (which represent Ann's perception of Bob's types) into classes of rank  $r^a \in \{1, 2, 4, 6, \dots\} \cup \{\infty\}$ . Similarly, the  $\sigma$ -algebras for Bob  $\Sigma_{\infty,*}^a$  and  $\Sigma_{1,*}^a, \Sigma_{2,*}^a, \dots$  can be partitioned into classes of rank  $r^b \in \{1, 2, 3, 5, 7, \dots\} \cup \{\infty\}$ , as illustrated in Figure 2.2. Note that the  $\sigma$ -algebras  $\Sigma_{1,*}^a, \Sigma_{2,*}^a, \dots$  on  $T_*^a$  are all virtual perceptions: they do not represent the perception of actual types for Bob, but only serve to fix the rank of the perceptions of Ann's types, and thus the depth of these types.

Theorem 4.3 then says that a type  $t_{\ell,k}^a$  for Ann has depth  $r(k)$ , where  $r(k)$  is the rank of the type's perception, given by  $r(1) = 1, r(2) = 2, r(3) = 4, r(4) = 6$ , and so on. The reason is that it is "common belief" (in an informal sense) that Bob knows the payoff parameter. To see this, suppose that a type for Ann has depth 3. By definition, the type can reason about Bob's second-order beliefs (i.e., Bob's beliefs about Ann's beliefs about  $\theta$ ), but cannot form beliefs about Bob's third-order belief, that is, his beliefs about Ann's beliefs about his beliefs about  $\theta$ . But, since Bob knows  $\theta$  in every state of the world, any two types for Bob that have the same beliefs about Ann's beliefs about  $\theta$  have the same beliefs about Ann's beliefs about Bob's beliefs about  $\theta$ .<sup>11</sup> This implies that if Ann has a belief about Bob's second-order belief, then she has a belief about Bob's third-order belief, and similarly for higher-orders.

---

<sup>11</sup>This does rely on the assumption that the probability  $p_k$  that Bob assigns to Ann's perception  $\Sigma_{k,*}^b$  does not depend on his type.

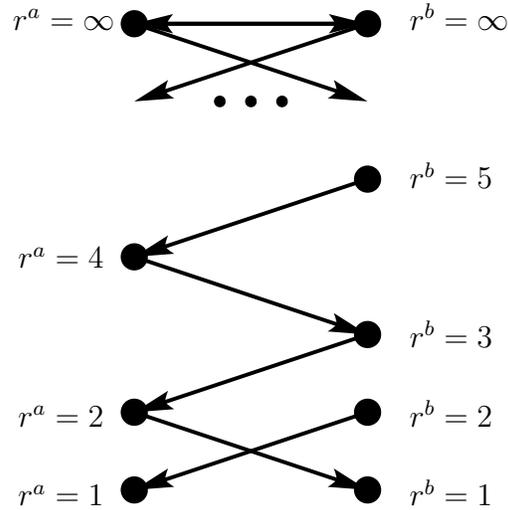


Figure 2.2: An ordering of the  $\sigma$ -algebras on the basis of their ranks, where a dot represents a class of  $\sigma$ -algebras, with the  $\sigma$ -algebras representing Ann’s perception on the left, and those for Bob on the right. An arrow from a class of  $\sigma$ -algebras to another class means that the  $\sigma$ -algebras in the former dominate those in the latter. Arrows representing dominance relations implied by other dominance relations have been omitted.

Before turning to the strategic analysis it is worth remarking that while the  $\sigma$ -algebras were “hand-picked” in this example so that differences in perceptions of types could be directly traced back to differences in perceptions of higher-order beliefs, this need not be done in general: the depth of types can be characterized from the dominance relations, regardless whether the  $\sigma$ -algebras have a direct interpretation in terms of higher-order beliefs.<sup>12</sup>

### 2.5.2 Coordinated attack

What actions can players choose, if they are rational, believe that the other is rational (to the extent they can reason about others’ rationality), believe that the other believes that his opponent is rational whenever the other player can reason about his opponent’s rationality (to the extent they can reason about others’ beliefs about their opponents’ rationality), and so on?

In the standard case, the rationalizable actions for a player can be found by iteratively eliminating actions that are not a best response to any “reasonable” conjecture over the actions of others, where a “reasonable” conjecture assigns positive probability only to actions of the

<sup>12</sup>Of course, the  $\sigma$ -algebras do represent perceptions of higher-order beliefs—that is at the core of the formalism. The point is that this relation is a result, not an input.

opponent that have not yet been eliminated.<sup>13</sup>

Here, Ann may have a bounded depth of reasoning, in which case she will be unable to verify whether her conjectures are reasonable beyond a certain point. A type for Ann of depth 1, for instance, has a conjecture about Bob's action, but that conjecture may assign positive probability to actions that are irrational for Bob, as reasoning about Bob's rationality requires the type to reason about Bob's beliefs, something which it is unable to do. More generally, a type of finite depth  $k$  does not eliminate any actions on the basis of considerations that require reasoning beyond order  $k$ .

What does this imply for behavior? A type  $t_{0,k}^a$  for Ann that has not sent any messages will still choose not to attack, as the action  $N$  is strictly dominant given her beliefs about  $\theta$ . However, the action  $A$  ("attack") is ruled out for a type  $t_{1,k}^a$  that has sent one message only if its depth  $r(k)$  is at least 3, since eliminating  $A$  requires the type to reason about Bob's beliefs about her beliefs about  $\theta$ : for that, she has to realize that Bob may think it is likely that she thinks it is likely that the state is  $\theta_1$ . Similarly, if Ann has received 2 messages, she will eliminate  $A$  only if her depth is at least 5, and so on.

This has implications for the actions that Bob can rationalize as well. A type  $t_{\ell,\infty}^b$  for Bob that has sent  $\ell > 0$  messages assigns a positive probability to the event that Ann has not eliminated the action  $A$ , and this probability increases with  $\ell$ . If  $\ell$  is sufficiently large, then this can make  $A$  rationalizable for him, provided the probability that Ann has an infinite depth of reasoning is not too high. In turn, this can make  $A$  rationalizable for types for Ann with a high, or even infinite depth of reasoning. This gives the following result:

*Suppose the probability  $p_\infty$  that a type for Ann has an infinite depth of reasoning is strictly less than  $\frac{1}{3}$ . Then there is a finite number of messages  $M$  such that  $A$  is rationalizable for a player whenever he has sent at least  $M$  messages.*

This is the outcome one would expect (cf. Rubinstein, 1989, p. 389–399), and it is driven entirely by the assumption that while players are rational, expect others to be rational, and so on, they are not always able to make complex inferences about others' beliefs, or may be uncertain whether others can do so, and so on.<sup>14</sup>

<sup>13</sup>See Dekel et al. (2007) and Battigalli and Siniscalchi (2003).

<sup>14</sup>The result does not hinge on the specific assumptions made here. For example, the result still holds if the probability that the probability  $p_k$  that Ann has perception  $\Sigma_{k,*}^b$  is zero for some  $k$ , as long as  $p_\infty < \frac{1}{3}$ . In particular, one could assume that Ann can reason up to at least order  $K$ , so that  $p_k = 0$  whenever  $k \leq r^{-1}(K)$ , and so on. One could also consider a larger extended type structure, in which there are types that have an infinite depth of reasoning, believe that the other player has an infinite depth of reasoning, ..., but think it is possible that the other player ... thinks it is possible that the other player has a finite depth of reasoning.

The next section begins the formal treatment by defining the class of extended type structure. Section 4 shows how the depth of reasoning of a type can be characterized, and Section 5 establishes the precise relation between Harsanyi type structures and extended type structures.

### 3 Extended type structures

#### 3.1 Definition

Suppose for simplicity that there are two players, Ann and Bob, indexed by  $i = a, b$ , and denote by  $j$  the player other than  $i$ , i.e.,  $j \neq i$ . The results can be extended to any finite number of players. Players are uncertain about the *state of nature*, which is drawn from a set  $\Theta$ ; the elements of  $\Theta$  could for instance be a player's payoff function or action. The set of states of nature  $\Theta$  is assumed to be Polish and to contain at least two elements.<sup>15</sup> The Borel  $\sigma$ -algebra on a space  $X$  is denoted  $\mathcal{B}(X)$ , and the set of Borel probability measures on  $X$  is  $\mathcal{M}(X)$ .

Consider a structure of the form

$$\langle T^a, T^b, \chi^a, \chi^b, (\Sigma_k^a)_{k \in \mathbb{N} \cup \{\infty\}}, (\Sigma_k^b)_{k \in \mathbb{N} \cup \{\infty\}} \rangle,$$

where  $T^a, T^b$  are the sets of *types* for Ann and Bob, respectively, assumed to be nonempty and Polish, and  $\Sigma_k^a$  and  $\Sigma_k^b$  are  $\sigma$ -algebras on  $T^a$  and  $T^b$ , where  $\Sigma_1^a = \{T^a, \emptyset\}$  and  $\Sigma_\infty^a = \mathcal{B}(T^a)$  is the Borel  $\sigma$ -algebra on  $T^a$ ; similarly,  $\Sigma_1^b = \{T^b, \emptyset\}$  and  $\Sigma_\infty^b = \mathcal{B}(T^b)$ .

For each player  $i = a, b$ , the function  $\chi^i$  maps every type  $t^i \in T^i$  to a probability space  $(\Theta \times T^j, \mathcal{B}(\Theta) \otimes \Sigma_m^j, \beta^i(t^i))$ , where  $\beta^i(t^i)$  is a probability measure defined on the  $\sigma$ -algebra  $\mathcal{B}(\Theta) \otimes \Sigma_m^j$  on  $\Theta \times T^j$ . If  $\chi^i(t^i)$  is the probability space  $(\Theta \times T^j, \mathcal{B}(\Theta) \otimes \Sigma_k^j, \beta^i(t^i))$ , let  $\Sigma^i(t^i) := \mathcal{B}(\Theta) \otimes \Sigma_k^j$  be the  $\sigma$ -algebra associated with  $t^i$ . The function  $\beta^i$  is the *belief map* for player  $i$ , and  $\beta^i(t^i)$  is the *belief* of  $t^i \in T^i$ .

The interpretation is that a  $\sigma$ -algebra  $\Sigma_k^a$  on Ann's type set  $T^a$  reflects the coarseness of perception of Ann's higher-order beliefs for a type  $t^b$  for Bob such that (the marginal of) the belief  $\beta^b(t^b)$  is defined on that  $\sigma$ -algebra. Notice that there need not be any types for player  $i$  with a given  $\sigma$ -algebra  $\Sigma_k^j$  on  $T^j$ . Members of  $\Theta \times T^a \times T^b$  are called *states (of the world)*.

Extended type structures satisfy a number of conditions: A structure  $\langle T^a, T^b, \chi^a, \chi^b, (\Sigma_k^a)_k, (\Sigma_k^b)_k \rangle$  as specified above is an extended type structure if there exists  $\mathcal{Z} \in \mathbb{N} \cup \{\infty\}$  such that for each player  $i = a, b$ , Conditions 3.1–3.4 below hold.

Condition 3.1 concerns the  $\sigma$ -algebras associated with the types of a given player.

---

<sup>15</sup>A topological space is *Polish* if it is separable and completely metrizable. Examples of Polish spaces include finite sets, the set of natural numbers  $\mathbb{N}$  (under the discrete metric) or any Euclidean space  $\mathbb{R}^m$ .

**Condition 3.1 [Filtrations]** For every  $k \in \mathbb{N}$  and  $\ell \leq \mathcal{Z}$  such that  $\ell > k$ , it holds that  $\Sigma_\ell^i \supseteq \Sigma_k^i$ . Moreover,  $\Sigma_\ell^i = \Sigma_{\mathcal{Z}}^i$  for all  $\ell \geq \mathcal{Z}$ . If  $\mathcal{B}(T^i) \neq \{T^i, \emptyset\}$ , then  $\Sigma_{k+1}^j = \Sigma_k^j$  for some  $\Sigma_k^j \subsetneq \Sigma_{\mathcal{Z}}^j$  implies that  $\Sigma_{k+1}^i \supsetneq \Sigma_k^i$ .<sup>16</sup>  $\triangleleft$

The condition states that the  $\sigma$ -algebras associated with the types of a player can be completely ordered in terms of set inclusion. That is, the collection of  $\sigma$ -algebras for a player form a *filtration*. Moreover, the  $\sigma$ -algebra  $\Sigma_{\mathcal{Z}}^i$ , which coincides with the Borel  $\sigma$ -algebra on  $T^i$ , is at least as fine as any other  $\sigma$ -algebra in the filtration.

While Condition 3.1 puts conditions on the relations between the  $\sigma$ -algebras on the type set of a single player, Conditions 3.2–3.4 impose restrictions on the relations between  $\sigma$ -algebras on different type sets. To state these conditions, some more notation is needed. Say that the  $\sigma$ -algebra  $\Sigma_k^i$  on  $T^i$  *dominates* the  $\sigma$ -algebra  $\Sigma_\ell^j$  on  $T^j$ , denoted  $\Sigma_k^i \succ \Sigma_\ell^j$ , if the following holds:

$$\forall E \in \mathcal{B}(\Theta) \otimes \Sigma_\ell^j, p \in [0, 1] : \quad \{t^i \in T^i : E \in \Sigma^i(t^i), \beta^i(t^i)(E) \geq p\} \in \Sigma_k^i.$$

That is, if the  $\sigma$ -algebra  $\Sigma_k^i$  dominates  $\Sigma_\ell^j$ , then a type  $t^j$  of player  $j$  with  $\sigma$ -algebra  $\Sigma_k^i$  on  $T^i$  can distinguish types for  $i$  that differ in their beliefs on  $\mathcal{B}(\Theta) \otimes \Sigma_\ell^j$ . In that case, type  $t^j$  has a belief about player  $i$ 's beliefs on  $\mathcal{B}(\Theta) \otimes \Sigma_\ell^j$ . Clearly, if  $\Sigma_k^i \succ \Sigma_\ell^j$ , then  $\Sigma_k^i \succ \Sigma_m^j$  for every  $m \leq \ell$  (use Condition 3.1).

The next condition states that if a  $\sigma$ -algebra on  $T^i$  is a strict refinement of another one, than the former dominates a  $\sigma$ -algebra on  $T^j$  that the latter does not dominate:

**Condition 3.2 [Strict refinements and dominance]** If  $\Sigma_k^i \supsetneq \Sigma_{k-1}^i$ , then there exists a  $\sigma$ -algebra  $\Sigma_\ell^j$  on  $T^j$  such that  $\Sigma_k^i \succ \Sigma_\ell^j$  and  $\Sigma_{k-1}^i \not\succeq \Sigma_\ell^j$ .  $\triangleleft$

To state the next condition, write  $\Sigma_{T^j}^i(t^i)$ , where  $t^i \in T^i$ , for the  $\sigma$ -algebra on which the marginal on  $T^j$  of the belief  $\beta^i(t^i)$  for type  $t^i$  is defined, that is, if  $\Sigma^i(t^i) = \mathcal{B}(\Theta) \otimes \Sigma_k^j$ , then  $\Sigma_{T^j}^i(t^i) = \Sigma_k^j$ .

**Condition 3.3 [Ordering]** Suppose there is a player  $q = a, b$  and a type  $t \in T^q$  such that  $\Sigma_{T^r}^q(t) \subsetneq \Sigma_{\mathcal{Z}}^r$ , where  $r \neq q$ . Then, for every  $\Sigma_\ell^j$  such that  $\Sigma_\ell^j \subsetneq \Sigma_{\mathcal{Z}}^j$ , it holds that  $\Sigma_\ell^j \not\succeq \Sigma_k^i$  whenever  $\Sigma_k^i \succ \Sigma_\ell^j$ . Moreover, if  $\Sigma_\ell^j \succ \Sigma_k^i$  for all  $\Sigma_k^i \subsetneq \Sigma_{\mathcal{Z}}^i$ , then  $\Sigma_\ell^j \succ \Sigma_{\mathcal{Z}}^i$ .  $\triangleleft$

**Condition 3.4 [Mutual dominance]** It holds that  $\Sigma_{\mathcal{Z}}^a \succ \Sigma_{\mathcal{Z}}^b$  and  $\Sigma_{\mathcal{Z}}^b \succ \Sigma_{\mathcal{Z}}^a$ .  $\triangleleft$

It follows from Condition 3.1 that the  $\sigma$ -algebra  $\Sigma_{\mathcal{Z}}^i$  dominates any  $\sigma$ -algebra  $\Sigma_\ell^j$  on  $T^j$ . Moreover, by Condition 3.3,  $\Sigma_\ell^j \not\succeq \Sigma_{\mathcal{Z}}^i$  whenever  $\Sigma_\ell^j$  is a proper subset of  $\Sigma_{\mathcal{Z}}^j$  (if there is some type

---

<sup>16</sup>Since  $\mathcal{B}(T^i)$  contains the singletons, the Borel  $\sigma$ -algebra  $\mathcal{B}(T^i)$  equals  $\{T^i, \emptyset\}$  if and only if player  $i$  only has one type.

$t$  with a coarse perception  $\Sigma_{T^r}^q(t) \subsetneq \Sigma_{\mathcal{Z}}^r$ . These conditions imply that the  $\sigma$ -algebras of the different players can naturally be ordered in terms of the induced depth of reasoning. The next subsection elaborates on this.

To complete the definition of an extended type structure, it remains to specify a  $\sigma$ -algebra on the set of possible beliefs. Let  $\mathcal{M}_k(\Theta \times T^j)$  be the collection of probability measures on the  $\sigma$ -algebra  $\mathcal{B}(\Theta) \otimes \Sigma_k^j$  on  $\Theta \times T^j$ , where  $k \leq \mathcal{Z}$ , so that, in particular, the set  $\mathcal{M}_{\mathcal{Z}}(\Theta \times T^j)$  is the collection of Borel probability measures  $\mathcal{M}(\Theta \times T^j)$  on  $\Theta \times T^j$ . Let

$$\mathcal{M}^+(\Theta \times T^j) := \mathcal{M}(\Theta \times T^j) \cup \bigcup_{k < \mathcal{Z}} \mathcal{M}_k(\Theta \times T^j),$$

be the union of these sets; note that  $\mathcal{M}^+(\Theta \times T^j)$  is the codomain  $\beta^i(T^i)$  of the belief map  $\beta^i$ . The set of probability measures  $\mathcal{M}^+(\Theta \times T^j)$  is endowed with the  $\sigma$ -algebra  $\mathcal{F}^i$  generated by sets of the form

$$\{\nu^i \in \mathcal{M}^+(\Theta \times T^j) : E \in \Sigma(\nu^i), \nu^i(E) \geq p\} : \quad E \in \mathcal{B}(\Theta) \otimes \Sigma_{\mathcal{Z}}^j, p \in [0, 1], \quad (3.1)$$

where  $\Sigma(\mu)$  for a probability measure  $\mu$  denotes the  $\sigma$ -algebra on which  $\mu$  is defined. The (relative)  $\sigma$ -algebra on  $\mathcal{M}(\Theta \times T^j)$  generated by sets of the form (3.1) coincides with the Borel  $\sigma$ -algebra on  $\mathcal{M}(\Theta \times T^j)$  associated with the weak topology.

### 3.2 Dominance and rank

The dominance relation  $\succ$  plays an important role in determining the depth of reasoning of a type. To see the intuition, recall that a  $\sigma$ -algebra  $\Sigma_k^a$  on Ann's type set  $T^a$  represents the perception of Ann's higher-order beliefs for a type for Bob whose belief is defined on that  $\sigma$ -algebra. If  $\Sigma_k^a$  dominates some  $\sigma$ -algebra  $\Sigma_\ell^b$  on Bob's types, then a type for Bob whose beliefs are defined on  $\Sigma_k^a$  can reason about the event that Ann assigns some probability to any of Bob's higher-order beliefs expressible in  $\Sigma_\ell^b$ . If  $\Sigma_\ell^b$  does not dominate  $\Sigma_k^a$ , then a type for Ann whose belief is defined on  $\Sigma_\ell^b$  cannot reason about the event that Bob has some belief about her higher-order beliefs that are expressible only in the  $\sigma$ -algebra  $\Sigma_k^a$ , suggesting that a type for Bob with  $\sigma$ -algebra  $\Sigma_k^a$  has a (strictly) greater depth of reasoning than a type for Ann with  $\sigma$ -algebra  $\Sigma_\ell^b$ .

Section 4 will make these intuitions precise. As a step in that direction, it will be useful to partition the  $\sigma$ -algebras for each players into classes, with the ordering of  $\sigma$ -algebras into classes being determined by the dominance relations.

For each player  $i = a, b$ , let  $\mathcal{S}^i := \{\Sigma_k^i : k \in \mathbb{N} \cup \{\infty\}\}$  be the set of  $\sigma$ -algebras on  $T^i$ , and define

$$\Lambda_{\geq 1}^i := \mathcal{S}^j,$$

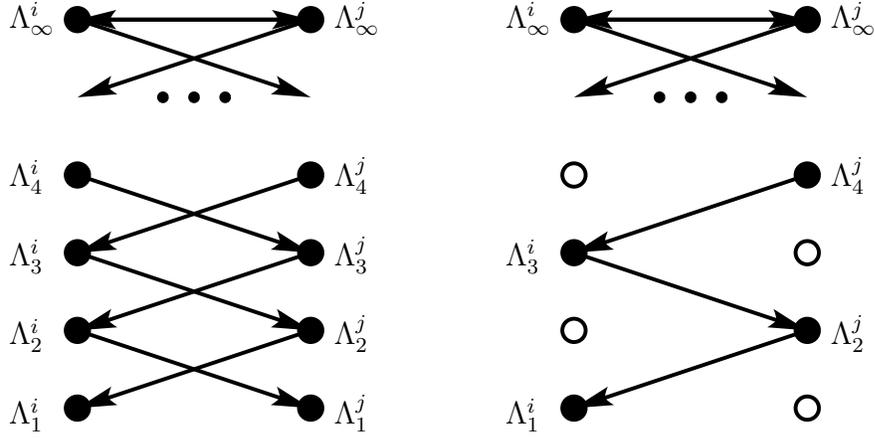


Figure 3.1: Two possible configurations for the dominance classes and the dominance relations between them. Each black dot ( $\bullet$ ) represents a nonempty dominance class, and each white dot ( $\circ$ ) is an empty dominance class. An arrow from a dominance class  $\Lambda_k^i$  to a dominance class  $\Lambda_\ell^j$  means that the  $\sigma$ -algebras in  $\Lambda_k^i$  dominate those in  $\Lambda_\ell^j$ . For clarity, arrows representing dominance relations implied by other dominance relations have been omitted.

where  $j \neq i$ . For  $k \in \mathbb{N}$ , define

$$\Lambda_{\geq k+1}^i := \{\mathcal{S}^j \in \Lambda_{\geq k}^i : \text{there is } \mathcal{S}^i \in \Lambda_{\geq k}^j \text{ s.t. } \mathcal{S}^j \succ \mathcal{S}^i\},$$

and

$$\Lambda_k^i := \Lambda_{\geq k}^i \setminus \Lambda_{\geq k+1}^i.$$

That is,  $\Lambda_{\geq k+1}^i$  is the collection of  $\sigma$ -algebras on  $T^j$  that dominate a  $\sigma$ -algebra in  $\Lambda_{\geq k}^j$ , and  $\Lambda_k^i$  is the subset of  $\sigma$ -algebras in  $\Lambda_{\geq k}^i$  that do not dominate a  $\sigma$ -algebra in  $\Lambda_{\geq k}^j$ . Also define

$$\Lambda_\infty^i := \{\Sigma_k^j \in \mathcal{S}^j : \Sigma_k^j \succ \Sigma_\ell^j \text{ for all } \Sigma_\ell^j \subsetneq \Sigma_k^j\}$$

By Conditions 3.1 and 3.4, it holds that  $\Sigma_Z^j \in \Lambda_\infty^i$ . In general, a class  $\Lambda_k^i$  can be empty for some  $k \in \mathbb{N}$ . Clearly,  $\Lambda_k^i \cap \Lambda_\ell^i = \emptyset$  whenever  $k \neq \ell$  for  $k, \ell \in \mathbb{N}$ . Moreover, the collection of classes  $\Lambda_k^i$ ,  $k \in \mathbb{N} \cup \{\infty\}$  forms a partition of the collection of  $\sigma$ -algebras  $\mathcal{S}^j$  on  $T^j$ .

This gives the following preliminary result:

**Lemma 3.1** *Let  $i = a, b$ ,  $j \neq i$ . If  $\mathcal{S}^j, \mathcal{S}_*^j \in \Lambda_k^i$  for some  $k \in \mathbb{N} \cup \{\infty\}$ , then  $\mathcal{S}^j = \mathcal{S}_*^j$ .*

This result says that the classes  $\Lambda_k^i$  are equivalence classes: Any two  $\sigma$ -algebras in one such class coincide. Say that a  $\sigma$ -algebra  $\mathcal{S}^j$  in  $\Lambda_k^i$  has *rank*  $k$ .

The next result shows how the ranks of the equivalence classes for different players are related via the dominance relation:

**Proposition 3.2** *Suppose  $\mathcal{S}^a \in \mathcal{S}^a, \mathcal{S}^b \in \mathcal{S}^b$  are such that  $\mathcal{S}^a \neq \Sigma_{\mathcal{Z}}^a, \mathcal{S}^b \neq \Sigma_{\mathcal{Z}}^b$ . Then, for  $i = a, b$  and  $j \neq i$ , the  $\sigma$ -algebra  $\mathcal{S}^i$  dominates  $\mathcal{S}^j$  if and only if the rank of  $\mathcal{S}^i$  belongs is strictly greater than the rank of  $\mathcal{S}^j$ .*

That is, a  $\sigma$  algebra on  $T^i$  from an equivalence class  $\Lambda_k^i$  dominates every  $\sigma$ -algebras in  $\Lambda_\ell^i$  whenever  $k > \ell$ . Moreover, the  $\sigma$ -algebras in  $\Lambda_\infty^a$  dominate the  $\sigma$ -algebras in  $\Lambda_\infty^b$  and vice versa (Condition 3.4); see Figure 3.1 for an illustration. I will refer to the equivalence classes  $\Lambda_k^i, k \in \mathbb{N}$ , as the *dominance classes* for player  $i$  in the following.

The argument above then suggests that types for Bob with a  $\sigma$ -algebra of rank  $k$  have a greater depth of reasoning than types for Ann with a  $\sigma$ -algebra with rank  $\ell$  if and only if  $k > \ell$ . This is formalized in the next section.

## 4 Depth of reasoning

This section relates the coarseness of the  $\sigma$ -algebra associated with a type with the depth of the belief hierarchy it generates. Section 4.1 constructs a space of finite and infinite hierarchies of beliefs, and defines the depth of each belief hierarchy. Section 4.2 then maps each type into a hierarchy in this space, and establishes the relation between the depth of a type and the rank of its  $\sigma$ -algebra.

### 4.1 Belief hierarchies

This section constructs a space of belief hierarchies. Start with some preliminary notation. The product  $Y = \times_\ell Y_\ell$  of a collection of spaces  $Y_\ell$  is endowed with the product topology, and the projection mapping from  $U \subseteq Y$  to  $Y_\ell$  is denoted  $\pi_{Y_\ell}^U$ . The marginal of a probability measure  $\mu$  on  $U$  on  $Y^\ell$  is  $\text{marg}_{Y^\ell} \mu := \mu \circ (\pi_{Y^\ell}^U)^{-1}$ . Recall that if  $f$  is a function from a space  $X$  to a measurable space  $(Y, \Sigma)$ , then the  $\sigma$ -algebra on  $X$  generated by the function  $f$ , denoted  $\sigma(f)$ , is the smallest  $\sigma$ -algebra on  $X$  that makes  $f$  measurable.

As before, assume that the set of states of nature  $\Theta$  is Polish and that it contains at least two elements. To construct a space of hierarchies for each player  $i$ , I define two sequences of spaces for each player (cf. Heifetz, 1993). The first sequence,  $(\Omega_k^i)_{k \in \mathbb{N}}$ , represents player  $i$ 's *k*th-order uncertainty domain for each  $k$ . The second,  $(B_k^i)_{k \in \mathbb{N}}$  consists of  $i$ 's coherent beliefs on these spaces. Formally, for  $i = a, b$  and  $j \neq i$ , define

$$\Omega_1^i := \Theta, \quad B_1^i := \mathcal{M}^+(\Omega_1^i),$$

where  $\mathcal{M}^+(\Omega_1^i) = \mathcal{M}(\Omega_1^i)$  is the set of Borel probability measures on  $\Omega_1^i$ , endowed with the weak topology. For  $k = 1, 2, \dots$ , assume that  $\Omega_1^i, \dots, \Omega_k^i$ , and  $B_1^i, \dots, B_k^i$  have been defined

for each player  $i = a, b$ , and let

$$\Omega_{k+1}^i := \Theta \times B_k^j.$$

That is, the  $(k+1)$ th-order uncertainty domain for player  $i$  consists of her basic uncertainty domain  $\Theta$  and the coherent beliefs of  $j$  at order  $k$ .

Turning to a player's beliefs about the  $(k+1)$ th-order uncertainty domain, let

$$B_{k+1}^i := \left\{ (\mu_1^i, \dots, \mu_k^i, \mu_{k+1}^i) \in B_k^i \times \mathcal{M}^+(\Omega_{k+1}^i) : \text{marg}_{\Omega_k^i} \mu_{k+1}^i = \mu_k^i \right\}, \quad (4.1)$$

where

$$\mathcal{M}^+(\Omega_{k+1}^i) := \mathcal{M}(\Omega_{k+1}^i) \cup \left( \bigcup_{\ell=1}^k \mathcal{M}_\ell(\Omega_{k+1}^i) \right),$$

with  $\mathcal{M}_\ell(\Omega_{k+1}^i)$  the collection of probability measures on  $\Omega_{k+1}^i$  with the  $\sigma$ -algebra

$$\sigma \left( \pi_{\Omega_\ell^i}^{\Omega_{k+1}^i} \right) = \left\{ \left( \pi_{\Omega_\ell^i}^{\Omega_{k+1}^i} \right)^{-1} (B) : B \in \mathcal{B}(\Omega_\ell^i) \right\}$$

generated by the function  $\pi_{\Omega_\ell^i}^{\Omega_{k+1}^i}$  that projects  $\Omega_{k+1}^i$  into  $\Omega_\ell^i$ . That is, the events in the  $\sigma$ -algebra  $\sigma(\pi_{\Omega_\ell^i}^{\Omega_{k+1}^i})$  are the extensions of the Borel sets of  $\Omega_\ell^i$  to the higher-order space  $\Omega_{k+1}^i$ . The interpretation is that the  $\sigma$ -algebra is not refined further beyond the Borel  $\sigma$ -algebra on  $\Omega_\ell^i$ , in the sense that it cannot distinguish elements of  $\Omega_{k+1}^i$  that coincide up to  $\Omega_\ell^i$ .<sup>17</sup>

The condition on the marginal of  $\mu_{k+1}^i$  in the definition (4.1) of  $B_{k+1}^i$  is a coherency condition. The recursive construction of the spaces makes that the probability measures  $\mu_{k+1}^i$  and  $\mu_k^i$  both specify a belief about the space of uncertainty  $\Omega_k^i$ , and these need to be coherent: the “new belief”  $\mu_{k+1}^i$  cannot contradict beliefs at lower orders. The coherency condition is a standard ingredient of a construction of a space of belief hierarchies, but has an extra bite in the current context. Unlike in other constructions, beliefs can be defined on different  $\sigma$ -algebras, so that the  $\sigma$ -algebras in a given hierarchy also need to be consistent across different orders.

It follows from this coherency condition and the selection of possible beliefs  $\mathcal{M}^+(\Omega_k^a)$  that if a player does not refine her  $\sigma$ -algebra at a certain order, she will not refine it at higher orders. This means there are two cases. In the first case, player  $i$  has a sequence of beliefs  $(\mu_1^i, \dots, \mu_k^i) \in B_k^i$  such that each  $\mu_\ell^i$  is a Borel probability measure on  $\Omega_\ell^i$ . In that case, her belief  $\mu_{k+1}^i$  defined on  $\Omega_{k+1}^i$  is either a Borel probability measure or the unique probability measure defined on the  $\sigma$ -algebra  $\sigma(\pi_{\Omega_\ell^i}^{\Omega_{k+1}^i})$  that is consistent with  $(\mu_1^a, \dots, \mu_k^a)$ . That is, either

---

<sup>17</sup>Indeed, Lemma D.4 in the appendix shows that the  $\sigma$ -algebra generated by  $\pi_{\Omega_\ell^i}^{\Omega_{k+1}^i}$  is strictly finer than the  $\sigma$ -algebra generated by  $\pi_{\Omega_m^i}^{\Omega_{k+1}^i}$  when  $\ell > m$ , and that each of those is coarser than the Borel  $\sigma$ -algebra  $\mathcal{B}(\Omega_{k+1}^i)$ .

player  $i$  refines her  $\sigma$ -algebra to the Borel  $\sigma$ -algebra on  $\Omega_{k+1}^i$ , or her  $\sigma$ -algebra is not refined any further. The second case to consider is that in which the belief  $\mu_k^i$  on her  $k$ th-order uncertainty domain  $\Omega_k^i$  is given by a probability measure with  $\sigma$ -algebra  $\sigma(\pi_{\Omega_\ell^i}^{\Omega_k^i})$  for some  $\ell \leq k-1$ . In that case,  $\mu_{k+1}^i$  is the probability measure on  $\Omega_{k+1}^i$  with  $\sigma$ -algebra  $\sigma(\pi_{\Omega_\ell^i}^{\Omega_{k+1}^i})$  that agrees with her beliefs at lower orders.

The space  $\mathcal{M}^+(\Omega_{k+1}^i)$  will be endowed with the topology induced by the metric  $\rho_{k+1}^i$ , defined by:

$$\rho_{k+1}^i(\mu, \mu') := \begin{cases} \tilde{\rho}_{k+1}^i(\mu, \mu') & \text{if } \mu, \mu' \in \mathcal{M}(\Omega_{k+1}^i); \\ \tilde{\rho}_\ell^i(\mu_\ell, \mu'_\ell) & \text{if } \mu, \mu' \in \mathcal{M}_\ell(\Omega_{k+1}^i); \\ 1 & \text{otherwise;} \end{cases}$$

for  $\mu, \mu' \in \mathcal{M}^+(\Omega_{k+1}^i)$ , where  $\tilde{\rho}_m^i$  is the Prohorov metric on the set of Borel probability measures  $\mathcal{M}(\Omega_m^i)$  on  $\Omega_m^i$ , and  $\nu_m = \text{marg}_{\Omega_m^i} \nu$  is the marginal of a probability measure  $\nu \in \mathcal{M}_m(\Omega_{k+1}^i)$  on  $\Omega_m^i$ . This topology is natural in the sense that it coincides with the weak topology on each subset of probability measures with a given  $\sigma$ -algebra.<sup>18</sup>

The sequence  $(B_k^i)_{k \in \mathbb{N}}$  endowed with the projection operators  $(\pi_{\Omega_\ell^i}^{\Omega_k^i})_{\ell \leq k}$  forms an inverse limit sequence. Let  $\widehat{T}^i$  be the inverse limit space, i.e., the set of belief hierarchies  $(\mu_1^i, \mu_2^i, \dots)$  such that any finite segment  $(\mu_1^i, \dots, \mu_k^i)$  belongs to  $B_k^i$ . The inverse limit space of belief hierarchies can be empty even if each of the spaces  $B_k^i$  is nonempty. The conditions on the set of states of nature  $\Theta$  and the choice of topology on each of the higher-order uncertainty domains ensure however that this is not the case:

**Proposition 4.1** *The set  $\widehat{T}^i$  is nonempty and Polish.*

It follows from the above argument that the set  $\widehat{T}^i$  of belief hierarchies for player  $i$  can be partitioned into different classes. Let  $\widehat{T}_k^i$ , where  $k \in \mathbb{N}$ , be the collection of  $(\mu_1^i, \mu_2^i, \dots)$  of belief hierarchies such that  $\mu_\ell^i \in \mathcal{M}(\Omega_\ell^i)$  for  $\ell \leq k$  and  $\mu_\ell^i \in \mathcal{M}_k(\Omega_\ell^i)$  for  $\ell > k$ . Say that  $(\mu_1^i, \mu_2^i, \dots)$  has *depth*  $k$ . Also, let  $\widehat{T}_\infty^i$  be the class of belief hierarchies  $(\mu_1^i, \mu_2^i, \dots)$  such that  $\mu_k^i$  is a Borel probability measure on  $\Omega_k^i$  for all  $k$ . The depth of a belief hierarchy  $(\mu_1^i, \mu_2^i, \dots) \in \widehat{T}_\infty^i$  is said to be infinite. That is, the depth of a finite belief hierarchy is  $k$  if it distinguishes the individual elements of its uncertainty domain up to order  $k$ . An infinite belief hierarchy distinguishes the individual elements in every uncertainty domain.

<sup>18</sup>More precisely, the metric space  $(\mathcal{M}_\ell(\Omega_{k+1}^i), \rho_{k+1}^i)$  is isometric to the metric space  $(\mathcal{M}(\Omega_\ell^i), \tilde{\rho}_\ell^i)$ . This means that the topology on  $\mathcal{M}^+(\Omega_{k+1}^i)$  is essentially the sum topology of the spaces  $\mathcal{M}(\Omega_1^i), \dots, \mathcal{M}(\Omega_{k+1}^i)$  of Borel probability measures endowed with the weak topology. Lemma D.5 in the appendix shows that this topology indeed induces a natural Borel  $\sigma$ -algebra on the collection of probability measures  $\mathcal{M}^+(\Omega_{k+1}^i)$ , with the measurable sets given by the collection of probability measures that assign probability at least  $p$  to events in  $\Omega_{k+1}^i$ , over all  $p \in [0, 1]$ .

The next section uses the definition of the depth of belief hierarchies to characterize the depth of reasoning of types in extended type structures.

## 4.2 Types and their depth

This section characterizes the relation between a type's  $\sigma$ -algebra and its depth of reasoning, as given by the depth of the hierarchy it induces, using the rank of the  $\sigma$ -algebras. To obtain this result, I inductively define a collection of functions, which map each type into a hierarchy of beliefs. Proposition 4.2 below shows that these functions are well defined. Let  $\Theta$  be the set of states of nature, as before, and consider a  $\Theta$ -based extended type structure with players indexed  $i = a, b$ , and type sets and belief maps of the players given by  $T^a, T^b$ , and  $\beta^a, \beta^b$ , respectively. Let  $i = a, b$ , and  $j \neq i$ , as before. Recall that  $\Lambda_k^i$  is a dominance class of rank  $k$ , containing  $\sigma$ -algebras on  $T^j$ ; the collection of  $\sigma$ -algebras with rank at least  $k$  is  $\Lambda_{\geq k}^i$ .

Define the function  $\gamma_1^i$  from  $\Theta \times T^j$  to  $\Omega_1^i$  by

$$\forall (\theta, t^j) \in \Theta \times T^j : \quad \gamma_1^i(\theta, t^j) := \theta.$$

Clearly,  $\gamma_1^i(\Theta \times T^j) \subseteq \Theta = \Omega_1^i$ . Also, for every  $k \leq \mathcal{Z}$ ,

$$\forall B \in \mathcal{B}(\Omega_1^i) : \quad (\gamma_1^i)^{-1}(B) \in \mathcal{B}(\Theta) \otimes \Sigma_k^j$$

for each  $\Sigma_k^j \in \Lambda_{\geq 1}^j = \mathcal{S}^j$ .

For  $m > 1$ , suppose the function  $\gamma_{m-1}^i$  from  $\Theta \times T^j$  to  $\Omega_{m-1}^i$  has been defined for each player  $i$ , so that in particular,

$$\gamma_{m-1}^i(\Theta \times T^j) \subseteq \Omega_{m-1}^i.$$

Also, suppose that

$$\forall B \in \mathcal{B}(\Omega_{m-1}^i) : \quad (\gamma_{m-1}^i)^{-1}(B) \in \mathcal{B}(\Theta) \otimes \Sigma_k^j \tag{4.2}$$

for every  $\Sigma_k^j \in \Lambda_{\geq m-1}^j$ .<sup>19</sup> Define the function  $\gamma_m^i : \Theta \times T^j \rightarrow \Omega_m^i$  as follows. For any state of nature  $\theta \in \Theta$  and type  $t^j \in T^j$  for player  $j$  such that  $\Sigma_{T^i}^j(t^j) \in \Lambda_{\geq m-1}^j$ , define

$$\gamma_m^i(\theta, t^j) := (\gamma_{m-1}^i(\theta, t^j), \beta^j(t^j) \circ (\gamma_{m-1}^j)^{-1});$$

alternatively, if  $\Sigma_{T^i}^j(t^j) \in \Lambda_\ell^j$  for  $\ell < m - 1$ , let

$$\gamma_m^i(\theta, t^j) := (\gamma_{m-1}^i(\theta, t^j), \mu_{\ell, m-1}^j(t^j)),$$

---

<sup>19</sup>Note that  $\Lambda_{\geq m-1}^j$  is nonempty, as it contains the Borel  $\sigma$ -algebra  $\mathcal{B}(T^j)$ .

where  $\mu_{\ell, m-1}^j(t^j)$  is the unique probability measure in  $\mathcal{M}_\ell(\Omega_{m-1}^j)$  such that

$$\text{marg}_{\Omega_\ell^j} \mu_{\ell, m-1}^j(t^j) = \beta^j(t^j) \circ (\gamma_\ell^j)^{-1}.$$

The functions  $\gamma_m^i$  are used to map each type for player  $i$  into a belief hierarchy in the space  $\widehat{T}^i$  of beliefs hierarchies constructed in Section 4.1. Define the function  $h^i$  from  $T^i$  to  $\widehat{T}^i$  as follows. Let  $t^i \in T^i$  be a type for player  $i$ . If  $\Sigma_{T^j}^i(t^i) \in \Lambda_m^i$  with  $m < \infty$ , define

$$h^i(t^i) := (\beta^i(t^i) \circ (\gamma_1^i)^{-1}, \dots, \beta^i(t^i) \circ (\gamma_m^i)^{-1}, \mu_{m, m+1}^i(t^i), \mu_{m, m+2}^i(t^i), \dots),$$

where the probability measure  $\mu_{m, \ell}^i(t^i) \in \mathcal{M}_m(\Omega_\ell^i)$ ,  $\ell > m$ , is defined as above. If  $\Sigma_{T^j}^i(t^i) = \mathcal{B}(T^j)$ , then define

$$h^i(t^i) := (\beta^i(t^i) \circ (\gamma_1^i)^{-1}, \beta^i(t^i) \circ (\gamma_2^i)^{-1}, \dots).$$

The next result states that this gives well-defined hierarchies of beliefs:

**Proposition 4.2** *For each player  $i$ , the functions  $\gamma_m^i : \Theta \times T^j \rightarrow \Omega_m^i$ ,  $m \in \mathbb{N}$ , and the function  $h^i : T^i \rightarrow \widehat{T}^i$  are well defined.*

Theorem 4.3 is the main result of this section. The result says that a type  $t^i \in T^i$  with a  $\sigma$ -algebra in the dominance class  $\Lambda_k^i$  induces a belief hierarchy of depth  $k$ :

**Theorem 4.3** *Let  $t^i \in T^i$  be a type for player  $i$  in an extended type structure. If  $\Sigma_{T^j}^i(t^i)$  has rank  $k \leq \infty$ , then  $t^i$  induces a belief hierarchy  $h^i(t^i)$  of depth  $k$ , i.e.,  $h^i(t^i) \in \widehat{T}_k^i$ .*

The result implies that a type for player  $i$  with a  $\sigma$ -algebra in the dominance class  $\Lambda_k^i$  can reason about every event in her  $\ell$ th-order uncertainty domain  $\Omega_\ell^i$  for any  $\ell \leq k$ . Moreover, it follows from the proof of Proposition 4.2 that there are events in the Borel  $\sigma$ -algebra on the  $(k+1)$ th domain of uncertainty  $\Omega_{k+1}^i$  that the type cannot reason about.<sup>20</sup> Say that a type  $t^i$  that induces a belief hierarchy  $h^i(t^i)$  of depth  $k$  has a *depth of reasoning* equal to  $k$ .

The next section shows that standard Harsanyi type structures can be characterized by common certainty that each player has an infinite depth of reasoning.

## 5 Harsanyi type structures

This section relates Harsanyi type structures to extended type structures. Section 5.1 shows that a Harsanyi type structure can be seen as a special class of extended type structures. Section 4.1 shows that there is an extended type structure that contains all Harsanyi type structures.

---

<sup>20</sup>Under an assumption that beliefs are sufficiently rich, a stronger result holds: a type of depth  $k$  cannot assign a probability to *any* nontrivial event concerning the other player's beliefs at order  $k$  or higher.

## 5.1 Common belief in infinite depth

This section shows that Harsanyi type structures are characterized by common belief in the event that each player has an infinite depth of reasoning. To state this result, adopt the standard belief operator to allow for the possibility that a player may not have beliefs at all orders. Fix a  $\Theta$ -based extended type structure  $\mathcal{T}$ , with type sets  $T^a, T^b$  for Ann and Bob, respectively, and let  $E \in \mathcal{B}(\Theta) \otimes \mathcal{B}(T^a) \otimes \mathcal{B}(T^b)$  be an event. Given a type  $t^a \in T^a$  for Ann, let

$$E_{t^a} := \{(\theta, t^b) \in \Theta \times T^b : (\theta, t^a, t^b) \in E\}$$

where  $\beta^a$  is the belief map of Ann in the extended type structure  $\mathcal{T}$ .<sup>21</sup> The event that Ann believes an event  $E \in \mathcal{B}(\Theta) \otimes \mathcal{B}(T^a) \otimes \mathcal{B}(T^b)$  (with probability  $p = 1$ ) is then:

$$B^a(E) := \{(\theta, t^a, t^b) \in \Theta \times T^a \times T^b : E_{t^a} \in \Sigma^a(t^a), \beta^a(t^a)(E_{t^a}) = 1\};$$

the event  $B^b(E)$  that Bob believes an event  $E \in \mathcal{B}(\Theta) \otimes \mathcal{B}(T^a) \otimes \mathcal{B}(T^b)$  is defined similarly. Appendix C.1 shows that the belief operator satisfies the usual properties. Define  $B(E) := B^a(E) \cap B^b(E)$  to be the event that both Ann and Bob believe  $E$ ;  $B(E)$  is the collection of states of the world where  $E$  is *mutual belief*.

Common belief in an event obtains if both players believe an event, believe that both believe it, and so on. Since the set of states where an event is mutual belief is itself an event, common belief in an event can be defined by iterating the mutual belief operator  $B(\cdot)$ : Say that an event  $E \in \mathcal{B}(\Theta) \otimes \mathcal{B}(T^a) \otimes \mathcal{B}(T^b)$  is *common belief* at state  $(\theta, t^a, t^b) \in \Theta \times T^a \times T^b$  if

$$(\theta, t^a, t^b) \in \bigcap_{\ell \in \mathbb{N}} [B]^\ell(E),$$

where  $[B]^1(E) := B(E)$ , and  $[B]^{\ell+1} = B \circ [B]^\ell$ .

Recall that a Harsanyi type structure on the set of states of nature  $\Theta$  is a structure  $\langle T_H^a, T_H^b, \beta_H^a, \beta_H^b \rangle$ , where  $T_H^a$  and  $T_H^b$  are Polish spaces, endowed with their Borel  $\sigma$ -algebras, and the belief map  $\beta_H^a$  is a measurable function from  $T_H^a$  into the set of Borel probability measures  $\mathcal{M}(\Theta \times T_H^b)$  on  $\Theta$  and Bob's type; the belief map  $\beta_H^b$  is defined similarly. Given a  $\Theta$ -based extended type structure with type sets  $T^a, T^b$  for Ann and Bob, respectively, let  $T_\infty^a, T_\infty^b$  be the set of types for Ann and Bob which induce a hierarchy of infinite depth. This gives the following result:

**Proposition 5.1** *Let  $\mathcal{T}_H := \langle T_H^a, T_H^b, \beta_H^a, \beta_H^b \rangle$  be a Harsanyi type structure. Then:*

---

<sup>21</sup>Implicit in the current formalism is thus that a player knows his own type. One could alternatively assume that a player cannot reason about his own higher-order beliefs above a certain order; see Kets (2009).

(a) There is an extended type structure  $\mathcal{T}$  which is equivalent to the Harsanyi type structure  $\mathcal{T}_H$ , in the sense that there is a bijection  $b^i$  for each player  $i$  from  $T_H^i$  to the type set  $T^i$  in the extended type structure  $\mathcal{T}$  such that  $\beta^i(b^i(t_H^i)) = \beta_H^i(t^i)$ , where  $\beta^i$  is the belief map of player  $i$  in the extended type structure.

(b) At each state in the extended type structure  $\mathcal{T}$  corresponding to the Harsanyi type structure  $\mathcal{T}_H$ , there is common belief in the event that players have an infinite depth of reasoning:

$$\Theta \times T^a \times T^b = \bigcap_{\ell \in \mathbb{N}} [B]^\ell (\Theta \times T_\infty^a \times T_\infty^b).$$

(c) Conversely, if for some extended type structure  $\mathcal{T}$  there is common belief in the event that players have an infinite depth of reasoning at every state of the world, then there is a Harsanyi type structure  $\mathcal{T}_H$  such that  $\mathcal{T}$  is equivalent to  $\mathcal{T}_H$  in the sense defined in (a).

On the other hand, there are clearly extended type structures that are not Harsanyi type structures (the extended type structure in Section 2 is an example).

## 5.2 Universality

This section shows that there is an extended type structure such that any extended type structure derived from a Harsanyi type structure can be embedded into this structure in a way that preserves higher-order beliefs.

Formally, fix the extended type structures  $\mathcal{T}$  and  $\mathcal{Q}$ , with type sets  $T^a, T^b$  and  $Q^a, Q^b$ , respectively. The belief maps in  $\mathcal{T}$  and  $\mathcal{Q}$  for player  $i = a, b$  are denoted by  $\beta^i$  and  $\lambda^i$ , respectively. Let  $T_\infty^i$  be the set of types for player  $i$  that induce an infinite belief hierarchy, i.e., for each type  $t^i \in T_\infty^i$ , the associated  $\sigma$ -algebra on  $T^j$  is  $\Sigma_{\mathcal{Z}}^j$ .

Following Mertens and Zamir (1985), let  $\varphi^a : T^a \rightarrow Q^a$  and  $\varphi^b : T^b \rightarrow Q^b$  be Borel measurable functions. The induced function  $\varphi := (\varphi^a, \varphi^b)$  is a *type morphism* for types of infinite depth if for each player  $i$ , and each Borel set  $E \in \mathcal{B}(\Theta) \otimes \mathcal{B}(T^j)$ ,

$$\lambda^i(h^i(t^i))(E) = \beta^i(t^i)((\text{Id}_\Theta, h^j)^{-1}(E))$$

for each type  $t^i \in T_\infty^i$ , where  $j \neq i$  is the player other than  $i$ , as before, and  $\text{Id}_\Theta$  is the identity function on  $\Theta$ . A type morphism preserves the beliefs structure of spaces as given by the belief maps.

Appendix D shows that the spaces  $\widehat{T}^a, \widehat{T}^b$  of finite and infinite belief hierarchies constructed in Section 4 can be used to define an extended type structure  $\widehat{\mathcal{T}}$ , called the *canonical structure*. The next result states that each extended type structure derived from a Harsanyi type

structure can be embedded into the canonical structure via a unique type morphism, so that the canonical structure  $\widehat{\mathcal{T}}$  is *universal* with respect to this class of extended type structures:

**Theorem 5.2** *Let  $\mathcal{T}$  be an extended type structure derived from a Harsanyi type structure as in Proposition 5.1. There is a unique type morphism from  $\mathcal{T}$  to the canonical structure  $\widehat{\mathcal{T}}$ .*

In particular, the universal type space of infinite hierarchies defined by Mertens and Zamir (1985), Brandenburger and Dekel (1993), and others can be embedded in the canonical structure in this way.<sup>22</sup>

## 6 Related literature

### 6.1 Higher-order beliefs and limited sophistication

A number of recent papers have developed type structures that generalize or provide an alternative to the standard Harsanyi framework. Strzalecki (2009) provides a general type space for cognitive hierarchy models and models of level- $k$  reasoning. An important difference with the present framework is that the formalism of Strzalecki is not a generalization of a Harsanyi type space; rather, Strzalecki follows the literature on level- $k$  reasoning in assuming that a player of a given cognitive type believes that other players have a lower cognitive type.

Heifetz and Kets (2010) develop a framework that allows for uncertainty both about the state of nature and the depth of reasoning of other players, and construct a universal type structure which contains the universal type structure of infinite belief hierarchies.<sup>23</sup> A difference with the current approach is that Heifetz and Kets assume that types with a limited depth of reasoning can conceive only of types that are less sophisticated than they are, as opposed to assuming that such types have a coarse perception of other players' higher-order beliefs, as is done here. They characterize the set of types in this universal structure whose behavior is fixed by the behavior of nonstrategic types in the context of a level- $k$  model.

Other related papers include Di Tillio (2008) and Ahn (2007).<sup>24</sup> Di Tillio (2008) relaxes the assumption that players are subjective expected utility maximizers, and constructs a space

---

<sup>22</sup>To be precise, this holds if the set of states of nature satisfies the appropriate topological conditions. The topological assumptions made here are the same as those of Brandenburger and Dekel (1993), and are somewhat stronger than those used by Heifetz (1993). Heifetz and Samet (1998b) do not require any topological conditions.

<sup>23</sup>Nielsen (2010) similarly constructs a space of belief hierarchies, though the space of hierarchies he constructs does not contain all infinite belief hierarchies, in contrast with the space constructed by Heifetz and Kets (2010).

<sup>24</sup>Also see Fagin, Halpern, and Vardi (1991).

of preference hierarchies to model players' higher-order preferences. He allows for incomplete preferences, which could be interpreted as a player not having beliefs about certain events, though he does not investigate this issue further. Ahn (2007) constructs a universal space of ambiguous beliefs, where a player has ambiguous beliefs about an event if his beliefs are represented by a set of priors. A possible interpretation of the case in which a player's belief over a collection of events is given by all possible priors is that a player does not have a belief about these events; see Halpern (2003, Thm. 2.3.3) for a result along these lines. Ahn does not explore this interpretation or the implications of bounded reasoning.

## 6.2 Information structures of a finite depth

While it seems unrealistic to assume that players have an infinite depth of reasoning when their beliefs are nontrivial at every order, there are Harsanyi type structures such that the higher-order beliefs of all types are completely determined by their beliefs up to some finite order, as in the often-studied cases that type sets are finite, or that beliefs are determined by a common prior and players' private information about the state of nature. Say that a Harsanyi type structure has a finite depth in this case.<sup>25</sup> When a Harsanyi type structure has a finite depth, the assumption that players have an infinite depth of reasoning does not seem to be so problematic.

An important difference between Harsanyi type structures of a finite depth and extended type structures is that the order of beliefs that completely determines the belief of a type in a Harsanyi type structure is common belief, whereas uncertainty about other players' depth of reasoning is a central element of the current approach.

This conceptual difference can lead to fundamentally different predictions. For example, while repeated communication will lead players to agree on their posteriors when higher-order beliefs are described by a finite Harsanyi type structure with a common prior (Geanakoplos and Polemarchakis, 1982), this need not be the case when there is higher-order uncertainty about the depth of reasoning of other players (depending on assumptions of how players update their beliefs when they have a limited depth of reasoning). Intuitively, the result of Geanakoplos and Polemarchakis (1982) relies on players' ability to distinguish states on the basis of the higher-order beliefs at those states. When there is a positive probability that a player does not have beliefs at the relevant orders, his announcements can become uninformative, and no agreement will be reached, even if both players in fact have an infinite depth of reasoning.

---

<sup>25</sup>Heifetz and Samet (1998a) and Morris, Postlewaite, and Shin (1995) explicitly study the rank or depth of state spaces. Also see Qin and Yang (2010).

### 6.3 Frames, analogies and unawareness

The current framework models players' finite depth of reasoning by assuming that players may not be able to distinguish states of the world that only differ in their beliefs at high orders. The literature on analogy-based expectation equilibria and learning with analogy classes also assumes that players may fail to distinguish different states of the world, motivated by the assumption that players use simplified representations to learn about their strategic environment (e.g., Jehiel, 2005; Mengel, 2007; Jehiel and Koessler, 2008). The solution concept used in that literature reflects the learning process of players, not their inability to reason about others' beliefs.

Another related literature studies the implications of relaxing the assumption that players have a common language (Crawford and Haller, 1990; Bacharach, 1993; Bacharach and Stahl, 2000; Blume, 2000; Blume and Gneezy, 2010). In these models, players have languages of different coarseness to describe the state of nature. By contrast, this paper assumes that players have the same language to describe the state of nature, but allows for heterogeneity in the perception of higher-order beliefs.

The inability of players to distinguish certain states on the basis of others' higher-order beliefs can also be interpreted in terms of unawareness. This is a different form of unawareness than the one commonly studied in the literature, which assumes that players may be unaware of certain aspects of the state of nature; see Geanakoplos (1989), Feinberg (2004), Halpern and Rego (2006), Heifetz et al. (2006), Board and Chung (2006), and Li (2009), among others. An interesting open question is what the interactions are between the different forms of unawareness, and to what strategic effects this gives rise.

### 6.4 Coordination in the electronic mail game

Section 2 showed that the Pareto-efficient action can be rationalized in the electronic mail game when there is higher-order uncertainty about players' depth of reasoning. Other papers have shown similar results. The result of Jehiel and Koessler (2008) is closest to the one presented here in the sense that they also assume that players may fail to distinguish certain states. In their model, players follow an equilibrium strategy, and ignore all information that is not payoff relevant. The possibility of coordinated attack therefore does not depend on how many messages have been sent. Allowing for approximate equilibria, Stinchcombe (1988) similarly shows that players can attack under the standard Harsanyi framework when players ignore certain payoff irrelevant information. Finally, Strzalecki (2009) shows that coordination on the Pareto-efficient action is possible in the electronic mail game using a model of level- $k$  reasoning, under the condition that very sophisticated types do not put too much weight

on other very sophisticated types. Such a condition is not needed here; in particular, in the extended type structure in Section 2, one player always believes that the other is fully sophisticated—she just cannot always infer the implications.

## 7 Concluding remarks

Aumann (1992) proposed to use epistemic models “as a very general kind of framework for studying limited rationality; a unified framework for considering rationality in environments that may include irrationality. . .” This paper takes a further step in this direction by developing an epistemic framework that explicitly allows for the possibility that players have a bounded depth of reasoning. The framework is fully consistent with the Harsanyi formalism, so that the effects of allowing for higher-order uncertainty about players’ depth of reasoning can be directly investigated.

The preliminary strategic analysis in Section 2 suggests that intuitive predictions can be obtained by allowing for higher-order uncertainty about others’ depth of reasoning. While the result presented there could also be obtained under some form of higher-order uncertainty of players’ rationality, it is not clear whether this holds more generally under related epistemic conditions, as when rational players are assumed to avoid weakly dominated strategies (cf. Brandenburger, Friedenberg, and Keisler, 2008).

More generally, a fundamental difference between the Harsanyi framework and the present one is that the current framework allows for higher-order uncertainty about players’ depth of reasoning (also see the discussion in Section 6.2). Whether allowing for such higher-order uncertainty leads to different conclusions as to what actions are rational under general assumptions on players’ beliefs is an important one, and the contribution of the present paper is that it provides a framework in which this question can be formalized.

## Appendix A Proofs for Section 3

### A.1 Proof of Lemma 3.1

Suppose  $k = \infty$ . Then, by Conditions 3.1 and 3.4, it holds that  $\Sigma_{\mathcal{Z}}^j \in \Lambda_{\infty}^i$ . Conversely, if  $\mathcal{S}^j \succ \mathcal{S}^i$  for all  $\mathcal{S}^i \subsetneq \Sigma_{\mathcal{Z}}^i$  (so that  $\mathcal{S}^j \in \Lambda_{\infty}^i$ ), then  $\mathcal{S}^j = \Sigma_{\mathcal{Z}}^j$  by Conditions 3.3 and 3.4.

So suppose  $k \in \mathbb{N}$ . Suppose  $\mathcal{S}^j \neq \mathcal{S}_*^j$ . By Condition 3.1, it is without loss of generality to assume  $\mathcal{S}_*^j \supsetneq \mathcal{S}^j$ . Suppose  $\mathcal{S}^j, \mathcal{S}_*^j \in \Lambda_1^i$ . Then, by the definition of  $\Lambda_1^i$ , there is no  $\mathcal{S}^i$  such that  $\mathcal{S}_*^j \succ \mathcal{S}^i$ . But this contradicts Condition 3.2.

For  $k > 1$  and  $q = a, b, r \neq q$ , suppose that  $\mathcal{S}^q, \hat{\mathcal{S}}^q \in \Lambda_{k-1}^r$  implies that  $\mathcal{S}^q = \hat{\mathcal{S}}^q$  and suppose  $\mathcal{S}^j, \mathcal{S}_*^j \in \Lambda_k^i$ . Using the definition of  $\Lambda_k^i$  and the induction hypothesis,  $\mathcal{S}^j \succ \mathcal{S}^i$  and  $\mathcal{S}_*^j \succ \mathcal{S}^i$  for all  $\mathcal{S}^i \in \mathcal{S}^i \setminus \Lambda_{\geq k}^j$ . But this contradicts Condition 3.2.  $\square$

### A.2 Proof of Proposition 3.2

[If] Suppose  $\mathcal{S}^i \in \Lambda_k^j$  and  $\mathcal{S}^j \in \Lambda_{\ell}^i$  such that  $\mathcal{S}^i \neq \Sigma_{\mathcal{Z}}^i, \mathcal{S}^j \neq \Sigma_{\mathcal{Z}}^j$ , and  $k > \ell$ . We need to show that  $\mathcal{S}^i \succ \mathcal{S}^j$ . Using the definition of  $\Lambda_k^j$  and Lemma 3.1, it follows that  $\mathcal{S}^i \succ \Sigma^j$  for all  $\Sigma^j \in \Lambda_{k-1}^i$ . By Condition 3.1, this implies that  $\mathcal{S}^i \succ \Sigma^j$  for all  $\Sigma^j \in \mathcal{S}^j \setminus \Lambda_{\geq k}^i$ , so  $\mathcal{S}^i \succ \mathcal{S}^j$ .

[Only if] Suppose  $\mathcal{S}^i \succ \mathcal{S}^j$  for  $\mathcal{S}^i \neq \Sigma_{\mathcal{Z}}^i, \mathcal{S}^j \neq \Sigma_{\mathcal{Z}}^j$ , where  $\mathcal{S}^i \in \Lambda_k^j$  for some  $k$ . We need to show that  $\mathcal{S}^j \in \Lambda_{\ell}^i$  for  $\ell < k$ . But it follows directly from the definition of  $\Lambda_k^j$  that  $\mathcal{S}^j \in \mathcal{S}^j \setminus \Lambda_{\geq k}^i$ .  $\square$

## Appendix B Proofs for Section 4

### B.1 Proof of Proposition 4.1

The proof of Proposition 4.1 relies on a number of lemmas. Throughout, we use that the space  $\mathcal{M}(X)$  of Borel probability measures on a Polish space  $X$  endowed with the weak topology is again a Polish space (Aliprantis and Border, 2005, Thm. 15.15).

**Lemma B.1** *For  $i = a, b$  and  $k \in \mathbb{N}$ ,  $\Omega_k^i$  and  $B_k^i$  are nonempty Polish spaces.*

**Proof.** The proof is by induction. Clearly,  $\Omega_1^i$  and  $B_1^i$  are nonempty and Polish for  $i = a, b$ . Suppose that  $\Omega_1^i, \dots, \Omega_k^i$  and  $B_1^i, \dots, B_k^i$  are nonempty and Polish. It is immediate that  $\Omega_{k+1}^i$ , the product of nonempty Polish spaces, is nonempty and Polish. It remains to show that  $B_{k+1}^i$  is nonempty and Polish. The proof is complete if the following two claims hold:

**Claim 1**  $\mathcal{M}^+(\Omega_{k+1}^i)$  is nonempty and Polish;

**Claim 2**  $B_{k+1}^i$  is a nonempty, closed subset of  $B_k^i \times \mathcal{M}^+(\Omega_{k+1}^i)$ .

To see that this suffices to prove the lemma, note that by Claim 1,  $B_k^i \times \mathcal{M}^+(\Omega_{k+1}^i)$  is Polish; since a closed subset of a Polish space is Polish, the proof follows.

It remains to establish the two claims. To prove the first claim, note that  $\mathcal{M}^+(\Omega_{k+1}^i)$  is clearly nonempty, as it contains the set of Borel probability measures  $\mathcal{M}(\Omega_{k+1}^i)$  on the nonempty Polish space  $\Omega_{k+1}^i$ . To show that  $\mathcal{M}^+(\Omega_{k+1}^i)$  is Polish, it is sufficient to show that  $\mathcal{M}_\ell(\Omega_{k+1}^i)$  is Polish for all  $\ell \leq k$ , since the sum of a (countable) sequence of Polish spaces is Polish (Kechris, 1995, Prop. 3.3). But this follows directly from the fact that  $\mathcal{M}_\ell(\Omega_{k+1}^i)$  and  $\mathcal{M}(\Omega_\ell^i)$  are isometric.

To prove the second claim, let  $\hat{t}^i = (\mu_1^i, \mu_2^i, \dots, \mu_{k+1}^i) \in B_k^i \times \mathcal{M}^+(\Omega_{k+1}^i)$  and suppose there is a sequence  $(\hat{t}_n^i)_{n \in \mathbb{N}}$  in  $B_k^i$ , where  $\hat{t}_n^i = (\mu_{1,n}^i, \mu_{2,n}^i, \dots, \mu_{k+1,n}^i)$ , such that  $\hat{t}_n^i \rightarrow \hat{t}^i$ . It is sufficient to show that  $\hat{t}^i \in B_k^i$ . Since  $B_k^i \times \mathcal{M}^+(\Omega_{k+1}^i)$  is endowed with the product topology,

$$\text{marg}_{\Omega_k^i} \mu_{k+1,n}^i \rightarrow \text{marg}_{\Omega_k^i} \mu_{k+1}^i,$$

and

$$\mu_{k,n}^i \rightarrow \mu_k^i.$$

Because  $\hat{t}_n^i \in B_{k+1}^i$  for all  $n$ , it follows that

$$\text{marg}_{\Omega_k^i} \mu_{k+1}^i = \mu_k^i$$

so that  $\hat{t}^i \in B_{k+1}^i$ . □

**Lemma B.2 (Heifetz, 1993, Thm. 6)** *For any  $(\mu_1^i, \dots, \mu_k^i) \in B_k^i$  with  $\mu_\ell^i \in \mathcal{M}(\Omega_\ell^i)$  for all  $\ell \leq k$ , there exists  $\mu_{k+1}^i \in \mathcal{M}(\Omega_{k+1}^i)$  such that  $(\mu_1^i, \dots, \mu_k^i, \mu_{k+1}^i) \in B_{k+1}^i$ .*

**Proof.** Let  $i = a, b$ , and fix  $(\mu_1^i, \dots, \mu_k^i) \in B_k^i$  such that  $\mu_\ell^i \in \mathcal{M}(\Omega_\ell^i)$  for all  $\ell \leq k$ . It suffices to show that there exists a continuous mapping  $f_k^i : \mathcal{M}(\Omega_k^i) \rightarrow \mathcal{M}(\Omega_{k+1}^i)$  such that  $(\mu_1^i, \dots, \mu_k^i, f_k^i(\mu_k^i)) \in B_{k+1}^i$ . To show this, we will construct a continuous mapping  $F_k^i : \Omega_k^i \rightarrow \Omega_{k+1}^i$  such that the composition  $\pi_{\Omega_k^i}^{\Omega_{k+1}^i} \circ F_k^i$  is the identity function on  $\Omega_k^i$ . Suppose we have defined such a continuous mapping  $F_k^i$ . Let  $f_k^i$  be the inverse image, i.e.,

$$f_k^i(\mu_k^i) := \mu_k^i \circ (F_k^i)^{-1}$$

for  $\mu_k^i \in \mathcal{M}(\Omega_k^i)$ . By standard results,  $f_k^i$  is continuous (Aliprantis and Border, 2005, Thm. 15.14). By assumption,  $\pi_{\Omega_k^i}^{\Omega_{k+1}^i} \circ F_k^i$  is the identity function on  $\Omega_k^i$ , so that

$$f_k^i(\mu_k^i) \circ (\pi_{\Omega_k^i}^{\Omega_{k+1}^i})^{-1} = \mu_k^i \circ (F_k^i)^{-1} \circ (\pi_{\Omega_k^i}^{\Omega_{k+1}^i})^{-1} = \mu_k^i.$$

Hence, it remains to construct  $F_k^i$  for  $i = a, b$  and  $k \in \mathbb{N}$ . Fix an arbitrary  $\theta_0^i \in \Theta$ , and define  $F_1^i : \Omega_1^i \rightarrow \Omega_2^i$  by

$$F_1^i(\theta^i) = (\theta^i, \delta_{\theta_0^i})$$

for  $\theta^i \in \Theta$ , where  $\delta_x$  is the delta-function on  $x$ . Then clearly  $F_1^i$  is continuous, and  $\pi_{\Omega_1^i}^{\Omega_2^i} \circ F_1^i$  is the identity function on  $\Omega_1^i$ . Suppose that  $F_k^i : \Omega_k^i \rightarrow \Omega_{k+1}^i$  has been defined for  $i = a, b$  so that  $F_k^i$  is continuous and  $\pi_{\Omega_k^i}^{\Omega_{k+1}^i} \circ F_k^i$  is the identity function. Then  $f_k^i$  is defined for both players as above, and is continuous; in fact, it is the Borel probability measure on  $\Omega_{k+1}^i$  under which  $i$  assigns probability 1 to  $j$  assigning probability 1 to  $\dots$  ( $k$  times) that  $\theta_0^i \in \Theta$  has been realized. Define  $F_{k+1}^i : \Omega_{k+1}^i \rightarrow \Omega_{k+2}^i$  by

$$F_{k+1}^i(\theta^i, \mu_1^j, \dots, \mu_k^j) = F_{k+1}^i(\theta^i, \mu_1^j, \dots, \mu_k^j, f^j(\mu_k^j)),$$

where, as always,  $j \neq i$ . It follows that  $F_{k+1}^i$  is continuous, and that  $\pi_{\Omega_{k+1}^i}^{\Omega_{k+2}^i} \circ F_{k+1}^i$  is the identity function.  $\square$

**Lemma B.3** *For  $i = a, b$  and  $k \in \mathbb{N}$ , the projection  $\pi_{B_{k-1}^i}^{B_k^i}$  is a surjection.*

**Proof.** Fix  $(\mu_1^i, \dots, \mu_{k-1}^i) \in B_{k-1}^i$ . It suffices to show that there exists  $\mu_k^i \in \mathcal{M}^+(\Omega_k^i)$  such that  $(\mu_1^i, \dots, \mu_{k-1}^i, \mu_k^i) \in B_k^i$ . If  $\mu_m^i$  is a Borel probability measure for all  $m$ , i.e.,  $\mu_m^i \in \mathcal{M}(\Omega_m^i)$  for  $m \leq k-1$ , then this follows from Lemma B.2. Hence, suppose there exists  $\ell < k-1$  such that  $\mu_m^i \in \mathcal{M}(\Omega_m^i)$  for  $m \leq \ell$  and  $\mu_m^i \in \mathcal{M}_\ell(\Omega_m^i)$  otherwise. In that case, there exists a unique probability measure  $\mu_k^i \in \mathcal{M}_\ell(\Omega_k^i)$  that coincides with  $\mu_{k-1}^i$  on  $\mathcal{B}(\Omega_\ell^i)$ .  $\square$

The proof of Proposition 4.1 is now straightforward:

**Proof of Proposition 4.1.** If the spaces in an inverse limit sequence are nonempty and the projections between them are surjective, then the inverse limit space is itself nonempty (e.g., Hocking and Young, 1988, Lemma 2.84). It therefore follows from Lemmas B.1 and B.3 that the inverse limit  $\widehat{T}^i$  is nonempty, where  $\widehat{T}^i \subset \times_{k \in \mathbb{N}} \mathcal{M}^+(\Omega_k^i)$  is endowed with the relative product topology. Since an inverse limit space is a closed subset of a product of Polish spaces (Hocking and Young, 1988), the inverse limit space  $\widehat{T}^i$  is Polish.  $\square$

## B.2 Proof of Proposition 4.2

The proof is by induction. As noted in the main text, the claim holds for  $m = 1$ . Let  $m > 1$  and suppose that the claim is true for  $m-1$ . The proof that the claim holds for  $m$  follows from the following steps:

**Step 1:** For every  $(\theta, t^j) \in \Theta \times T^j$ , it holds that  $\gamma_m^i(\theta, t^j) \in \Omega_m^i$ .

**Proof of Step 1.** The first thing to show is that the last component of  $\gamma_m^i(\theta, t^j)$  is an element of  $\mathcal{M}^+(\Omega_{m-1}^j)$  for all  $(\theta, t^j) \in \Theta \times T^j$ . Let  $t^j \in T^j$  be a type for player  $j$ . First suppose that  $\Sigma_{T^i}^j(t^j) \in \Lambda_k^j$  for  $k < m-1$ . In that case, the last component of  $\gamma_m^i(\theta, t^j)$  is  $\mu_{k,m-1}^j(t^j) \in \mathcal{M}_k(\Omega_{m-1}^j)$  for any  $\theta \in \Theta$ , so that  $\mu_{k,m-1}^j(t^j) \in \mathcal{M}^+(\Omega_{m-1}^j)$ . So suppose that  $\Sigma_{T^i}^j(t^j) \in \Lambda_{\geq m-1}^j$ . It follows from the induction hypothesis (4.2) that  $(\gamma_{m-1}^j)^{-1}(B) \in \mathcal{B}(\Theta) \otimes \Sigma_\ell^i$  for every  $B \in \mathcal{B}(\Omega_{m-1}^j)$ , so that for each  $B \in \mathcal{B}(\Omega_{m-1}^j)$ , the probability  $\beta^j(t^j) \circ (\gamma_{m-1}^j)^{-1}(B)$  is well defined. It follows that  $\beta^j(t^j) \circ (\gamma_{m-1}^j)^{-1} \in \mathcal{M}(\Omega_{m-1}^j)$  for every  $t^j \in T_k^j$ .

To show that beliefs are coherent (in the sense defined in Section 4.1), it will again be convenient to distinguish two cases. Let  $t^j \in T^j$  and let  $m > 2$ .<sup>26</sup> First suppose that  $\Sigma_{T^i}^j(t^j) \in \Lambda_k^j$  for  $k < m-1$ . For  $k = m-2$ , it holds by construction that

$$\text{marg}_{\Omega_{m-2}^j} \mu_{k,m-1}^j(t^j) = \beta^j(t^j) \circ (\gamma_{m-2}^j)^{-1},$$

so that beliefs are coherent. So suppose  $k < m-2$ . We need to show that

$$\text{marg}_{\Omega_{m-2}^j} \mu_{k,m-1}^j(t^j) = \mu_{k,m-2}^j(t^j),$$

where  $\mu_{k,m-2}^j(t^j)$  is defined as above. To show this, note that  $\text{marg}_{\Omega_{m-2}^j} \mu_{k,m-1}^j(t^j)$  is a probability measure on the  $\sigma$ -algebra  $\sigma(\pi_{\Omega_k^j}^{\Omega_{m-2}^j})$  on  $\Omega_{m-2}^j$ . Then, for every  $E \in \sigma(\pi_{\Omega_k^j}^{\Omega_{m-2}^j})$ ,

$$\begin{aligned} \text{marg}_{\Omega_{m-2}^j} \mu_{k,m-1}^j(t^j)(E) &= \left( \text{marg}_{\Omega_{k,m-2}^j} \beta^j(t^j) \circ (\gamma_k^j)^{-1} \circ (\pi_{\Omega_k^j}^{\Omega_{m-1}^j}) \right)(E) \\ &= \left( \beta^j(t^j) \circ (\gamma_k^j)^{-1} \circ (\pi_{\Omega_k^j}^{\Omega_{m-2}^j})^{-1} \right) \left( (\pi_{\Omega_k^j}^{\Omega_{m-2}^j})(E) \right) \\ &= \mu_{k,m-2}^j(t^j)(E), \end{aligned}$$

Next suppose that  $\Sigma_{T^i}^j(t^j) \in \Lambda_{\geq m-1}^j$ , and let  $E \in \mathcal{B}(\Omega_{m-2}^i)$ . Since the projection function  $\pi_{\Omega_{m-2}^i}^{\Omega_{m-1}^i}$  is continuous, it is Borel measurable, so that

$$\left( \pi_{\Omega_{m-2}^i}^{\Omega_{m-1}^i} \right)^{-1}(E) \in \mathcal{B}(\Omega_{m-1}^i).$$

Also, note that by the induction hypothesis,  $\gamma_{m-1}^j(\theta, t^j) \in \Omega_{m-1}^i$  for every  $(\theta, t^j) \in \Theta \times T^j$ . Then,

$$\begin{aligned} (\gamma_{m-2}^j)^{-1}(E) &= \{(\theta, t^j) \in \Theta \times T^j : \gamma_{m-2}^j(\theta, t^j) \in E\} \\ &= \left\{ (\theta, t^j) \in \Theta \times T^j : \gamma_{m-1}^j(\theta, t^j) \in \left( \pi_{\Omega_{m-2}^i}^{\Omega_{m-1}^i} \right)^{-1}(E) \right\} \\ &= \left( \gamma_{m-1}^j \right)^{-1} \left( \left( \pi_{\Omega_{m-2}^i}^{\Omega_{m-1}^i} \right)^{-1}(E) \right), \end{aligned}$$

---

<sup>26</sup>The requirement that beliefs are coherent for  $m \leq 2$  is trivially satisfied.

so that

$$\begin{aligned}\beta^j(t^j)((\gamma_{m-2}^j)^{-1}(E)) &= \beta^j(t^j) \circ (\gamma_{m-1}^j)^{-1} \circ (\pi_{\Omega_{m-2}^i}^{\Omega_{m-1}^i})^{-1}(E) \\ &= \text{marg}_{\Omega_{m-2}^i} \beta^j(t^j) \circ (\gamma_{m-1}^j)^{-1}(E),\end{aligned}$$

so that beliefs are coherent also in this case.  $\square$

For the next step, some new notation is needed. Let  $k \leq \mathcal{Z}$  and  $m > 1$ . For  $i = a, b$ , define  $\underline{\gamma}_{m-1}^i : \mathcal{M}^+(\Theta \times T^j) \rightarrow \mathcal{M}^+(\Omega_{m-1}^i)$  by:<sup>27</sup>

$$\forall \mu \in \mathcal{M}^+(\Theta \times T^j) : \quad \underline{\gamma}_{m-1}^i(\mu) := \mu \circ (\gamma_{m-1}^i)^{-1}.$$

Also, let  $\mathcal{S}(\mathcal{M}^+(\Theta \times T^j); k)$  be the sub- $\sigma$  algebra (of the Borel  $\sigma$ -algebra on  $\mathcal{M}^+(\Theta \times T^j)$ ) generated by sets of the form

$$\{\mu \in \mathcal{M}^+(\Theta \times T^j) : E \in \Sigma(\mu), \mu(E) \geq p\} : \quad E \in \mathcal{B}(\Theta) \otimes \mathcal{S}_k^j, p \in [0, 1],$$

where  $\mathcal{S}_k^j \in \bigcup_{\ell \leq k} \Lambda_\ell^i$ , and where  $\Sigma(\mu)$  is the  $\sigma$ -algebra on which the probability measure  $\mu$  is defined.<sup>28</sup> By Lemma D.5, it holds that the Borel  $\sigma$ -algebra on  $\mathcal{M}^+(\Omega_m^i)$  is generated by sets of the form

$$\{\mu \in \mathcal{M}^+(\Omega_m^i) : E \in \Sigma(\mu), \mu(E) \geq p\} : \quad E \in \mathcal{B}(\Omega_m^i), p \in [0, 1].$$

**Step 2:** Suppose that the induction hypothesis (4.2) holds. Then,

$$(\underline{\gamma}_{m-1}^j)^{-1}(M) \in \mathcal{S}(\mathcal{M}^+(\Theta \times T^i); m-1)$$

for every  $M \in \mathcal{B}(\mathcal{M}^+(\Omega_{m-1}^j))$ .

**Proof of Step 2.** To show that for all  $M \in \mathcal{B}(\mathcal{M}^+(\Omega_{m-1}^j))$ ,

$$(\underline{\gamma}_{m-1}^j)^{-1} \in \mathcal{S}(\mathcal{M}^+(\Theta \times T^i); m-1),$$

note that it follows from Corollary 4.24 of Aliprantis and Border (2005) that this holds if and only if

$$\{\mu \in \mathcal{M}^+(\Theta \times T^i) : \underline{\gamma}_{m-1}^j(\mu) \in \{\nu \in \mathcal{M}^+(\Omega_{m-1}^j) : E \in \Sigma(\nu), \nu(E) \geq p\}\}$$

---

<sup>27</sup>That is, the definition of  $\underline{\gamma}_{m-1}^i$  is very similar to the definition of the image measure induced by the function  $\gamma_m^i$ . The difference is that the domain of an image measure (as commonly defined) consists of probability measures defined on a single  $\sigma$ -algebra, and similarly for the range space.

<sup>28</sup>Note that  $\bigcup_{\ell \leq k} \Lambda_\ell^i$  is nonempty, as  $\Sigma_1^j \in \bigcup_{\ell \leq k} \Lambda_\ell^i$ .

is an element of  $\mathcal{S}(\mathcal{M}^+(\Theta \times T^i); m-1)$  for all  $E \in \mathcal{B}(\Omega_{m-1}^j)$  and  $p \in [0, 1]$ . Equivalently, the claim holds if and only if

$$\{\mu \in \mathcal{M}^+(\Theta \times T^i) : E \in \Sigma(\mu \circ (\gamma_{m-1}^j)^{-1}), \mu \circ (\gamma_{m-1}^j)^{-1}(E) \geq p\}$$

is in  $\mathcal{S}(\mathcal{M}^+(\Theta \times T^i); m-1)$  for every  $E \in \mathcal{B}(\Omega_{m-1}^j)$  and  $p \in [0, 1]$ . To prove this, recall that  $\mathcal{S}(\mathcal{M}^+(\Theta \times T^i); m-1)$  is generated by sets of the form

$$\{\mu \in \mathcal{M}^+(\Theta \times T^i) : B \in \Sigma(\mu), \mu(B) \geq p\} : \quad B \in \mathcal{B}(\Theta) \times \mathcal{S}_{m-1}^i, p \in [0, 1],$$

where  $\mathcal{S}_{m-1}^i \in \Lambda_{m-1}^j$ . The result therefore follows if for every  $E \in \mathcal{B}(\Omega_{m-1}^j)$ , there exists  $B \in \mathcal{B}(\Theta) \otimes \mathcal{S}_{m-1}^i$  such that  $B = (\gamma_{m-1}^j)^{-1}(E)$ . But this is immediate from (4.2).  $\square$

**Step 3:** For every  $M \in \mathcal{S}(\mathcal{M}^+(\Theta \times T^i); m-1)$  and  $\Sigma_k^j \in \Lambda_{\geq m}^i$ , it holds that  $(\beta^j)^{-1}(M) \in \Sigma_k^j$ .

**Proof of Step 3.** Note that  $\Lambda_{\geq m}^i$  is nonempty, as it contains  $\Sigma_{\mathcal{Z}}^j$ . Let  $\Sigma_k^j \in \Lambda_{\geq m}^i$  and let  $\mathcal{S}_{m-1}^i \in \Lambda_{m-1}^j$ . By Corollary 4.24 of Aliprantis and Border (2005), it is sufficient to show that for every  $E \in \mathcal{B}(\Theta) \otimes \mathcal{S}_{m-1}^i$  and  $p \in [0, 1]$ ,

$$\{t^j \in T^j : E \in \Sigma^j(t^j), \beta^j(t^j)(E) \geq p\} \in \Sigma_k^j.$$

But this follows from Proposition 3.2.  $\square$

**Step 4:** Suppose the induction hypothesis (4.2) holds. Then, for every  $M \in \mathcal{B}(\mathcal{M}^+(\Omega_{m-1}^j))$ , it holds that  $(\gamma_{m-1}^j \circ \beta^j)^{-1}(M) \in \Sigma_k^j$  for every  $\Sigma_k^j \in \Lambda_{\geq m}^i$ .

**Proof of Step 4.** This follows directly from Steps 2 and 3.  $\square$ .

Suppose  $\Lambda_{m-1}^i \neq \emptyset$  and let  $\Sigma_k^j \in \Lambda_{\geq m}^i$ . Using (4.2), it follows from Steps 2–4 that

$$\forall B \in \mathcal{B}(\Omega_m^i) : \quad (\gamma_m^i)^{-1}(B) \in \mathcal{B}(\Theta) \otimes \Sigma_k^j.$$

Together with Step 1, this implies that  $\gamma_\ell^i$  is well defined for every  $\ell \in \mathbb{N}$ . It follows immediately that the function  $h^i$  is well defined.  $\square$

### B.3 Proof of Theorem 4.3

Let  $t^i \in T^i$  be a type for player  $i$ , and suppose  $\Sigma_{T^j}^i(t^i) \in \Lambda_k^i$ ,  $k < \infty$ . It follows from Proposition 4.2 that

$$(\beta^i(t^i) \circ (\gamma_1^i)^{-1}, \dots, \beta^i(t^i) \circ (\gamma_k^i)^{-1}) \in B_k^i.$$

Moreover,  $\beta^i(t^i) \circ (\gamma_\ell^i)^{-1} \in \mathcal{M}(\Omega_\ell^i)$  for  $\ell \leq k$ . It follows from the definition of  $\mu_{k,m}^i(t^i)$ ,  $m > k$ , that  $\mu_{k,m}^i(t^i) \in \mathcal{M}_k(\Omega_m^i)$ , and

$$(\beta^i(t^i) \circ (\gamma_1^i)^{-1}, \dots, \beta^i(t^i) \circ (\gamma_{k-1}^i)^{-1}, \mu_{k,k+1}^i(t^i), \dots, \mu_{k,m}^i(t^i)) \in B_m^i.$$

Hence,  $h^i(t^i) \in \widehat{T}_k^i$ , so that  $h^i(t^i)$  is a finite hierarchy of depth  $k$ .

Next suppose  $\Sigma_{T^j}^i(t^i) = \mathcal{B}(T^j)$ . By Proposition 4.2, it holds that

$$(\beta^i(t^i) \circ (\gamma_1^i)^{-1}, \dots, \beta^i(t^i) \circ (\gamma_{k-1}^i)^{-1}) \in B_k^i$$

for all  $k$ . Also,  $\beta^i(t^i) \circ (\gamma_\ell^i)^{-1} \in \mathcal{M}(\Omega_\ell^i)$  for all  $\ell \in \mathbb{N}$ . Conclude that  $h^i(t^i) \in \widehat{T}_\infty^i$ , so that  $h^i(t^i)$  has infinite depth.  $\square$

## Appendix C Proofs for Section 5

### C.1 The belief operator

To discuss the properties of the belief operator, it will be useful to define the complement of the event that a player believes some event  $E$ . As before, let  $\mathcal{T}$  be an extended type structure, with type sets  $T^a, T^b$  for Ann and Bob, respectively, and let  $E \in \mathcal{B}(\Theta) \otimes \mathcal{B}(T^a) \otimes \mathcal{B}(T^b)$  be an event. Recall the definition of  $E_{t^a}$ , where  $t^a$  is an arbitrary type for Ann. The event that Ann does not believe  $E$ , denoted  $NB^a(E)$ , is then

$$NB^a(E) := \{(\theta, t^a, t^b) \in \Theta \times T^a \times T^b : E_{t^a} \notin \Sigma^a(t^a) \text{ or } \beta^a(t^a)(E_{t^a}) < 1\},$$

with  $NB^b(E)$  defined analogously. That is, if a type for Ann does not believe  $E$ , then either she assigns probability less than 1 to  $E$ , or her perception is too coarse to reason about  $E$ , in the sense that  $E$  is not an event for her type.<sup>29</sup>

Allowing for the fact that players may not be able to reason about all higher-order events, the belief operator satisfies the usual properties, as can be readily verified:

**Necessitation:**  $B^a(\Theta \times T^a \times T^b) = \Theta \times T^a \times T^b$ ;

**Monotonicity:** For any  $E \in \mathcal{B}(\Theta) \otimes \mathcal{B}(T^a) \otimes \Sigma_k^b$  and  $F \in \mathcal{B}(\Theta) \otimes \mathcal{B}(T^a) \otimes \Sigma_\ell^b$ , where  $k \leq \mathcal{Z}, \ell \leq k$ , such that  $E \subseteq F$ , it holds that  $B^a(E) \subseteq B^a(F)$ ;

**Conjunction:** For any  $E, F \in \mathcal{B}(\Theta) \otimes \mathcal{B}(T^a) \otimes \mathcal{B}(T^b)$ ,

$$B^a(E) \cap B^a(F) \subseteq B^a(E \cap F);$$

if, in addition,  $E, F \in \mathcal{B}(\Theta) \otimes \mathcal{B}(T^a) \otimes \Sigma_k^b$ ,  $k \leq \mathcal{Z}$ , implies  $E \cap F \in \mathcal{B}(\Theta) \otimes \mathcal{B}(T^a) \otimes \Sigma_k^b$ , then

$$B^a(E) \cap B^a(F) \supseteq B^a(E \cap F);$$

---

<sup>29</sup>This definition is similar to the one employed by Heifetz et al. (2006) in the context of unawareness. However, their definition satisfies different properties, which is not surprising given the differences between unawareness and the issues studied here; also see Halpern and Moses (1990).

**Positive introspection:** For any  $E \in \mathcal{B}(\Theta) \otimes \mathcal{B}(T^a) \otimes \mathcal{B}(T^b)$ ,

$$B^a(E) \subseteq B^a(B^a(E));$$

**Negative introspection:** For any  $E \in \mathcal{B}(\Theta) \otimes \mathcal{B}(T^a) \otimes \mathcal{B}(T^b)$ ,

$$NB^a(E) \subseteq B^a(NB^a(E));$$

with similar conditions holding for Bob's belief operator  $B^b$ . Necessitation is a weak consistency condition. Monotonicity says that if a player believes an event  $E$ , then she also believes any event  $F$  implied by  $E$ . Conjunction means that a player believes both  $E$  and  $F$  if and only if she believes  $E$  and she believes  $F$ . Positive introspection is an important property that says that if a player believes an event  $E$ , then she believes that she believes it. Negative introspection, on the other hand, states that a player knows what she does not believe: If Ann does not believe  $E$ , then she believes that she does not believe  $E$ .<sup>30</sup>

The properties listed above are similar to the ones satisfied by the usual belief operator, defined for the case where players have an infinite depth of reasoning. The only difference is that in the current setting, monotonicity and conjunction only hold for certain classes of events. Specifically, monotonicity is only guaranteed to hold if the implied event  $F$  is also an event in the  $\sigma$ -algebra in which  $E$  is an event. Similarly, conjunction holds if  $E \cap F$  is an event for a type whenever  $E$  and  $F$  are events for that type; otherwise, it need not hold. In the standard setting, where the beliefs of all types of a player are defined on the same  $\sigma$ -algebra, monotonicity and conjunction hold for all events in the  $\sigma$ -algebra associated with the types of the player.

## C.2 Proof of Proposition 5.1

(a) Let  $T^i = T_H^i$  for  $i = a, b$ , and choose  $\chi^i$  such that for each  $t^i \in T^i$ ,  $\Sigma^i(t^i) = \mathcal{B}(\Theta) \otimes \mathcal{B}(T^j)$  and  $\beta^i(t^i) = \beta_H^i(t_H^i)$ , where  $t_H^i \in T_H^i$  is the type in the Harsanyi structure corresponding to  $t^i$ .

Suppose each player has only one type in the Harsanyi type structure. In that case,  $\Sigma_1^i = \{T^i, \emptyset\} = \mathcal{B}(T^i)$  for each player  $i$ , as the Borel  $\sigma$ -algebra contains the singletons. It is easy to see that Condition 3.1 is satisfied with  $\mathcal{Z} = 1$ ; the other conditions are trivially satisfied. If some player has more than one type. In that case, all conditions are satisfied when  $\mathcal{Z} = 2$ .

(b) Fix an extended type structure  $\mathcal{T}$  (with type sets  $T^a, T^b$ ) that is equivalent to a Harsanyi type structure  $\mathcal{T}_H$  in the sense defined in part (a). Then, by Condition 3.4, it holds that

---

<sup>30</sup>That the belief operator satisfies positive and negative introspection is a direct implication from the assumption that a player knows her own type.

$T_\infty^a = T^a$  and  $T_\infty^b = T^b$ , so that trivially  $\Theta \times T_\infty^b \in \Sigma^a(t^a)$  for each type  $t^a \in T^a$  for Ann, and similarly for Bob. The result is then immediate.

(c) By definition,  $\Theta \times T^a \times T^b \supseteq \Theta \times T_\infty^a \times T_\infty^b$ . Using the definition of common belief and that  $\Theta \times T^a \times T^b = \cap_\ell B^\ell(\Theta \times T_\infty^a \times T_\infty^b)$ , it follows that

$$\Theta \times T^a \times T^b = \cap_\ell B^\ell(\Theta \times T_\infty^a \times T_\infty^b) \subseteq B(\Theta \times T_\infty^a \times T_\infty^b) = \subseteq \Theta \times T_\infty^a \times T_\infty^b,$$

so that  $T_\infty^a = T^a$ , and  $T_\infty^b = T^b$ .<sup>31</sup> A Harsanyi type structure  $\mathcal{T}_H = \langle T_H^a, T_H^b, \beta_H^a, \beta_H^b \rangle$  such that the extended type structure  $\mathcal{T}$  is equivalent to  $\mathcal{T}_H$  can now be constructed by setting  $T_H^i := T^i$  and  $\beta_H^i(t_H^i) := \beta^i(t^i)$  for each  $t^i \in T^i$ , where  $t_H^i$  is the type in  $T_H^i$  corresponding to  $t^i$ . It can easily be checked that this is a Harsanyi type structure.  $\square$

### C.3 Proof of Theorem 5.2

Let  $\mathcal{T}$  be an extended type structure with type sets  $T^a, T^b$ , and recall that the function  $h^i : T^i \rightarrow \widehat{T}^i$  defined in Section 4 maps each type in  $T^i$  to a belief hierarchy in  $\widehat{T}^i$ . The function  $h^i$  is the *hierarchy map* from  $T^i$  into  $\widehat{T}^i$ . Denote the subset of types for player  $i$  with infinite depth by  $T_\infty^i \subseteq T^i$ .

The following preliminary result shows that type morphisms preserve infinite hierarchies of beliefs:

**Lemma C.1** *Let  $\varphi = (\varphi^a, \varphi^b)$  be a type morphism from an extended type structure  $\mathcal{T}$  to an extended type structure  $\mathcal{Q}$ . Then, for each player  $i$ ,*

$$h_Q^i(\varphi^i(t^i)) = h_T^i(t^i)$$

for each type  $t^i \in T_\infty^i$ , where  $h_T^i$  and  $h_Q^i$  are the hierarchy maps from  $T^i$  and  $Q^i$  into  $\widehat{T}^i$ , respectively.

**Proof.** Denote the belief maps of player  $i$  in  $\mathcal{T}$  and  $\mathcal{Q}$  by  $\beta_T^i$  and  $\beta_Q^i$ , respectively. It is sufficient to show that  $\gamma_{k,Q}^i \circ (\text{Id}_\Theta, \varphi^j) = \gamma_{k,T}^i$ , where  $\gamma_{k,Q}^i$  and  $\gamma_{k,T}^i$  are the mappings defined in Section 4 with domains  $\Theta \times Q^j$  and  $\Theta \times T^j$ , respectively, and  $\text{Id}_\Theta$  is the identity function on  $\Theta$ . The proof is by induction. For every  $(\theta, t^j) \in \Theta \times T_\infty^j$ ,  $\gamma_{1,Q}^i(\theta, \varphi^j(t^j)) = \theta = \gamma_{1,T}^i(\theta, t^j)$ . For  $k > 1$ , suppose the claim is true for  $k - 1$ . Then, using the induction hypothesis and that

---

<sup>31</sup>Indeed, it can readily be verified that  $\Theta \times T_\infty^j \in \Sigma^j(t^j)$  for any  $t^j \in T_\infty^j$ .

$\varphi$  is a type morphism, it holds for each  $(\theta, t^j) \in \Theta \times T_\infty^j$  that

$$\begin{aligned}
\gamma_{k,Q}^i(\theta, \varphi^j(t^j)) &= (\gamma_{k-1,Q}^i(\theta, \varphi^j(t^j)), \beta_Q^j(\varphi^j(t^j)) \circ (\gamma_{k-1,Q}^j)^{-1}) \\
&= (\gamma_{k-1,T}^i(\theta, t^j), \beta_T^j(t^j) \circ (\text{Id}_\Theta, \varphi^i)^{-1} \circ (\gamma_{k-1,Q}^j)^{-1}) \\
&= (\gamma_{k-1,T}^i(\theta, t^j), \beta_T^j(t^j) \circ (\gamma_{k-1,Q}^j \circ (\text{Id}_\Theta, \varphi^i))^{-1}) \\
&= (\gamma_{k-1,T}^i(\theta, t^j), \beta_T^j(t^j) \circ (\gamma_{k-1,T}^j)^{-1}).
\end{aligned}$$

It follows that

$$\begin{aligned}
h_Q^i(\varphi^i(t^i)) &= (\beta_Q^i(\varphi^i(t^i)) \circ (\gamma_{1,Q}^i)^{-1}, \beta_Q^i(\varphi^i(t^i)) \circ (\gamma_{2,Q}^i)^{-1}, \dots) \\
&= (\beta_T^i(t^i) \circ (\gamma_{1,T}^i)^{-1}, \beta_T^i(t^i) \circ (\gamma_{2,T}^i)^{-1}, \dots) \\
&= h_T^i(t^i),
\end{aligned}$$

for each  $t^i \in T_\infty^i$  □

We are now ready to prove Theorem 5.2. Let  $\Theta$  be the set of states of nature, and let  $\mathcal{T}_H$  be a Harsanyi type structure on  $\Theta$ , and let  $\mathcal{T}$  be the extended type structure that is equivalent to the Harsanyi structure, as defined in Proposition 5.1, with type sets  $T^a, T^b$  and belief maps  $\beta^a, \beta^b$  for Ann and Bob, respectively. Denote the  $\sigma$ -algebras in the filtration on  $T^i$  by  $\Sigma_k^i$ ,  $k \leq \infty$ , as before, and recall that there exists  $\mathcal{Z} \leq \infty$  such that  $\Sigma_{\mathcal{Z}}^i = \mathcal{B}(T^i)$ . It follows from Proposition 5.1 that each type in  $T^i$  has infinite depth, i.e., for each  $t^i \in T^i$ , the associated  $\sigma$ -algebra is  $\Sigma^i(t^i) = \mathcal{B}(\Theta) \otimes \Sigma_{\mathcal{Z}}^j$ , so that  $T_\infty^i = T^i$ . By Proposition D.3, there is a canonical structure  $\widehat{\mathcal{T}}$ , where the type set  $\widehat{T}^i$  of player  $i$  is given by the space of finite and infinite belief hierarchies on  $\Theta$ , and the belief map  $\psi^i$  is canonical, as defined in Proposition D.1.

To show that each extended type structure derived from a Harsanyi type structure can be embedded in the canonical structure, define  $\varphi^i := h^i$  for each player  $i$ , where  $h^i : T^i \rightarrow \widehat{T}^i$  is the hierarchy map for player  $i$ . Note that, since  $T^i = T_\infty^i$ , the function  $\varphi^i$  is indeed defined only for types of infinite depth. It follows from the proof of Proposition 4.2 that  $\varphi^i$  is Borel measurable, i.e., for each  $B \in \mathcal{B}(\widehat{T}^i)$ ,

$$(\varphi^i)^{-1}(B) \in \mathcal{B}(T^i).$$

Let  $E \in \mathcal{B}(\Theta) \otimes \mathcal{B}(\widehat{T}^j)$  be a Borel set. Using that the belief map  $\psi^i$  is canonical,

$$\begin{aligned}
\psi^i(\varphi^i(t^i))(E) &= \psi^i(\beta^i(t^i) \circ (\gamma_1^i)^{-1}, \beta^i(t^i) \circ (\gamma_2^i)^{-1}, \dots) \\
&= \beta^i(t^i)((\text{Id}_\Theta, \varphi^j)^{-1}(E)),
\end{aligned}$$

so that  $\varphi$  is a type morphism. Since type morphisms preserve belief hierarchies (Lemma C.1),  $\varphi$  is the unique type morphism into the spaces of belief hierarchies  $\widehat{T}^a, \widehat{T}^b$ . □

## Appendix D The canonical structure

Section 4 constructs finite and infinite belief hierarchies to characterize the depth of types in extended type structures (Theorem 4.3). Here I demonstrate that the resulting spaces of belief hierarchies  $\widehat{T}^a, \widehat{T}^b$  define an extended type structure (Proposition D.3), called the *canonical structure*. This result is used in the proof of Theorem 5.2 to show that any extended type structure derived from a Harsanyi type structure can be embedded into this structure.

Recall the definitions from Section 4.1. It will also be useful to define  $\Omega^i := \Theta \times \widehat{T}^j$  for each player  $i$  and  $j \neq i$ , so that  $\Omega^i$  is the full domain of uncertainty for player  $i$ . It follows directly from 4.1 that  $\Omega^i$  is nonempty and Polish. Also define  $B_0^i := \{\widehat{T}^i, \emptyset\}$  to be the trivial  $\sigma$ -algebra on  $\widehat{T}^i$ .

The next result states that a (finite or infinite) hierarchy in  $\widehat{T}^i$  fully describes a player's uncertainty: Any hierarchy for player  $i$  uniquely determines a belief on  $\Omega^i = \Theta \times \widehat{T}^j$ . Moreover, the mapping from a hierarchy for a player  $i$  to a belief over  $\Theta \times \widehat{T}^j$  is *canonical* in the sense that the belief about  $\Theta \times \widehat{T}^j$  agrees with the beliefs on the spaces  $\Omega_1^i, \Omega_2^i, \dots$  specified by the hierarchy that induces that belief:

**Proposition D.1** (a) *If  $(\mu_1^i, \mu_2^i, \dots) \in \widehat{T}_\infty^i$ , then there exists a unique probability measure  $\mu^i$  on  $(\Theta \times \widehat{T}^j, \mathcal{B}(\Theta) \otimes \mathcal{B}(\widehat{T}^j))$  such that*

$$\text{marg}_{\Omega_k^i} \mu^i = \mu_k^i$$

for all  $k \in \mathbb{N}$ .

(b) *If  $(\mu_1^i, \mu_2^i, \dots) \in \widehat{T}_\ell^i$  for  $\ell \in \mathbb{N}$ , then there exists a unique probability measure  $\mu^i$  on  $(\Theta \times \widehat{T}^j, \mathcal{B}(\Theta) \otimes \sigma(\pi_{B_{\ell-1}^j}^{\widehat{T}^j}))$  such that*

$$\text{marg}_{\Omega_k^i} \mu^i = \mu_k^i.$$

for all  $k \in \mathbb{N}$ .

**Proof.** (a) First note that it follows from Proposition 4.1 that  $\Omega^i = \Theta \times \widehat{T}^j$  is a nonempty Polish space. By a version of the Kolmogorov extension theorem (Parthasarathy, 1978, Prop. 27.4), there exists a unique Borel probability measure  $\mu^i$  on  $\Theta \times \widehat{T}^j$  for any  $(\mu_1^i, \mu_2^i, \dots) \in \widehat{T}_\infty^i$  such that  $\pi_{\Omega_k^i}^{\Omega^i}(\mu^i) = \mu_k^i$  for all  $k$ .

(b) Fix  $k \in \mathbb{N}$  and consider a hierarchy  $\hat{t}^i = (\mu_1^i, \dots, \mu_k^i, \mu_{k+1}^i, \dots) \in \widehat{T}_k^i$  of depth  $k$ . Then the unique probability measure  $\mu^i$  on the measurable space  $(\Theta \times \widehat{T}^j, \mathcal{B}(\Theta) \otimes \sigma(\pi_{\Theta \times B_{k-1}^j}^{\Theta \times \widehat{T}^j}))$  defined by

$$\mu^i(A) = \mu_k^i \left( \pi_{\Theta \times B_{k-1}^j}^{\Theta \times \widehat{T}^j} (A) \right)$$

for all  $A \in \{(\pi_{\Theta \times B_{k-1}^j}^{\Theta \times \widehat{T}^j})^{-1}(B) : B \in \mathcal{B}(\Theta \times B_{k-1}^j)\}$  satisfies the requirements.  $\square$

That is, each belief hierarchy can be associated with a belief about  $\Theta$  and the belief hierarchy of the other player. It follows from Lemma D.4 that the beliefs associated with hierarchies of greater depth are defined on a strictly finer  $\sigma$ -algebra than hierarchies of smaller depth. In particular, the beliefs of a hierarchy whose depth is infinite are defined on the Borel  $\sigma$ -algebra, which is finer than any of the other  $\sigma$ -algebras.

This result can be used to define an extended type structure, where the set of types for each player  $i$  is the set of finite and infinite belief hierarchies  $\widehat{T}^i$ . To state this result, a topology on this set of probability measures needs to be specified. Recall that  $\mathcal{M}(\Theta \times \widehat{T}^j)$  is the set of Borel probability measures on  $\Theta \times \widehat{T}^j$  and that  $\mathcal{M}_\ell(\Theta \times \widehat{T}^j)$  is the set of probability measures on  $(\Theta \times \widehat{T}^j, \mathcal{B}(\Theta) \otimes \sigma(\pi_{B_{\ell-1}^j}^{\widehat{T}^j}))$ , for  $\ell \in \mathbb{N}$ . Let

$$\mathcal{M}^+(\Theta \times \widehat{T}^j) := \mathcal{M}(\Theta \times \widehat{T}^j) \cup \bigcup_{\ell \in \mathbb{N}} \mathcal{M}_\ell(\Theta \times \widehat{T}^j)$$

be the collection of these probability measures. Define the following function on  $\mathcal{M}^+(\Theta \times \widehat{T}^j) \times \mathcal{M}^+(\Theta \times \widehat{T}^j)$ : For  $\mu, \mu' \in \mathcal{M}^+(\Theta \times \widehat{T}^j)$ ,

$$\rho_{\widehat{T}^j}^i(\mu, \mu') = \begin{cases} \tilde{\rho}_{\widehat{T}^j}^i(\mu, \mu') & \text{if } \mu, \mu' \in \mathcal{M}(\Theta \times \widehat{T}^j); \\ \tilde{\rho}_k^i(\mu_k, \mu'_k) & \text{if } \mu, \mu' \in \mathcal{M}_k(\Theta \times \widehat{T}^j); \\ 1 & \text{otherwise;} \end{cases}$$

where  $\tilde{\rho}_{\widehat{T}^j}^i$  is the Prohorov metric on  $\mathcal{M}(\Theta \times \widehat{T}^j)$ ,  $\nu_k := \text{marg}_{\Theta \times B_{k-1}^j} \nu$ , and  $\tilde{\rho}_k^i$  is the Prohorov metric on  $\Theta \times B_{k-1}^j = \Omega_k^i$ , as before. It is easy to verify that  $\rho_{\widehat{T}^j}^i$  is a metric on  $\mathcal{M}^+(\Theta \times \widehat{T}^j)$ .

By Proposition D.1 above, each belief hierarchy  $(\mu_1^i, \mu_2^i, \dots)$  of player  $i$  can be associated with a unique belief  $\mu^i$  over the basic space of uncertainty  $\Theta$  and the other player's belief hierarchies  $\widehat{T}^j$  in such a way that  $i$ 's belief over his  $k$ th-order space of uncertainty (as specified by the marginal of  $\mu^i$  on  $\Omega_k^i$ ) coincides with  $\mu_k^i$ . The inverse mapping assigns to each belief  $\mu^i$  over  $\Omega^i = \Theta \times \widehat{T}^j$  a belief hierarchy  $(\text{marg}_{\Omega_k^i} \mu^i)_{k \in \mathbb{N}}$ . It turns out that both mappings are continuous bijections, so that this defines a homeomorphism, as the next result demonstrates:

**Proposition D.2** *Let  $i, j = a, b$ ,  $j \neq i$ .*

- (a) *There is a homeomorphism  $\psi_\infty^i : \widehat{T}_\infty^i \rightarrow \mathcal{M}(\Theta \times \widehat{T}^j)$ .*
- (b) *For each  $k \in \mathbb{N}$ , there is a homeomorphism  $\psi_k^i : \widehat{T}_k^i \rightarrow \mathcal{M}_k(\Theta \times \widehat{T}^j)$ , where  $\mathcal{M}_k(\Theta \times \widehat{T}^j)$  is the set of probability measures on  $\Theta \times \widehat{T}^j$  with  $\sigma$ -algebra  $\mathcal{B}(\Theta) \otimes \sigma(\pi_{B_{k-1}^j}^{\widehat{T}^j})$ .*
- (c) *There is a homeomorphism  $\psi^i : \widehat{T}^i \rightarrow \mathcal{M}^+(\Theta \times \widehat{T}^j)$ .*

(d) For any  $k \in \mathbb{N}$ ,  $\sigma(\pi_{B_{k-1}^j}^{\widehat{T}^j})$  is a strict subset of  $\mathcal{B}(\widehat{T}^j)$ . Moreover,  $\sigma(\pi_{B_{k-1}^j}^{\widehat{T}^j})$  is a strict subset of  $\sigma(\pi_{B_{k'-1}^j}^{\widehat{T}^j})$  whenever  $k < k'$ .

**Proof.** (a) By Proposition D.1, each hierarchy  $(\mu_1^i, \mu_2^i, \dots) \in \widehat{T}_\infty^i$  can be associated with a unique Borel probability measure  $\mu^i$  on  $\Theta \times \widehat{T}^j$ . Denote the function that maps  $\widehat{T}^i$  to  $\mathcal{M}(\Theta \times \widehat{T}^j)$  in this way by  $\psi_\infty^i$ . Conversely, let  $r_\infty^i : \mathcal{M}(\Theta \times \widehat{T}^j) \rightarrow \widehat{T}^i$  be the mapping that assigns to each  $\mu^i \in \mathcal{M}(\Theta \times \widehat{T}^j)$  the hierarchy  $(\text{marg}_{\Omega_k^i} \mu^i)_{k \in \mathbb{N}}$ . The function  $r_\infty^i$  is the inverse of  $\psi_\infty^i$ ; it remains to show that  $\psi_\infty^i$  and  $r_\infty^i$  are continuous. The function  $\psi_\infty^i$  is continuous if and only if  $\hat{t}_n \rightarrow \hat{t}^i$  in  $\widehat{T}_\infty^i$  implies  $\psi_\infty^i(\hat{t}_n) \rightarrow \psi_\infty^i(\hat{t}^i)$  in  $\mathcal{M}(\Theta \times \widehat{T}^j)$ . But the cylinders form a convergence determining class in  $\Omega^i = \Theta \times \widehat{T}^j$  (Billingsley, 1999, Thm. 2.4), and the value of  $\psi_\infty^i(\hat{t}^i)$  for  $\hat{t}^i = (\mu_1^i, \mu_2^i, \dots)$  on the cylinders is given by the  $\mu_k^i$ 's. Finally, it follows directly from the continuity of the marginal operator that  $r_k^i$  is continuous.

(b) By Proposition D.1(b), each finite hierarchy  $\hat{t}^i = (\mu_1^i, \mu_2^i, \dots) \in \widehat{T}_\ell^i$  of depth  $\ell$  can be associated with a unique probability measure  $\mu^i$  on the measurable space  $(\Theta \times \widehat{T}^j, \mathcal{B}(\Theta) \otimes \sigma(\pi_{B_{\ell-1}^j}^{\widehat{T}^j}))$  such that its marginal on  $\Omega_k^i$  coincides with  $\mu_k^i$  for all  $k$ . Denote the mapping that associates each  $\hat{t}^i \in \widehat{T}_\ell^i$  with such a probability measure by  $\psi_\ell^i$ . Conversely, every  $\mu^i \in \mathcal{M}_\ell(\Theta \times \widehat{T}^j)$  can be mapped into  $\widehat{T}_\ell^i$  by taking the appropriate marginal. This mapping clearly defines the inverse of  $\psi_\ell^i$ , and is continuous. The continuity of  $\psi_\ell^i$  is immediate since  $\psi_\ell^i(\hat{t}^i)$  is completely determined by the  $\ell$ th element of  $\hat{t}^i$ .

(c) Immediate from (a) and (b).

(d) To establish the first result, it will be useful to first prove the following:

$$\mathcal{B}(\Omega^i) = \sigma \left( \left\{ \left( \pi_{\Omega_k^i}^{\Omega^i} \right)^{-1}(B) : B \in \mathcal{B}(\Omega_k^i), k \in \mathbb{N} \right\} \right).$$

To show this, let  $\mathcal{G}(\Omega_k^i)$  be the collection of open sets in  $\Omega_k^i$ , where  $k \in \mathbb{N}$ , and recall that  $\sigma(\mathcal{E})$  for a collection of subsets of a set  $X$  denotes the  $\sigma$ -algebra on  $X$  generated by  $\mathcal{E}$ . Then, as the open sets generate the Borel  $\sigma$ -algebra, and  $\Omega^i$  is endowed with the product topology,

$$\mathcal{B}(\Omega^i) = \sigma \left( \left\{ \left( \pi_{\Omega_k^i}^{\Omega^i} \right)^{-1}(B) : B \in \mathcal{G}(\Omega_k^i), k \in \mathbb{N} \right\} \right).$$

Hence, it is sufficient to show that

$$\sigma \left( \left\{ \left( \pi_{\Omega_k^i}^{\Omega^i} \right)^{-1}(B) : B \in \mathcal{B}(\Omega_k^i), k \in \mathbb{N} \right\} \right) = \sigma \left( \left\{ \left( \pi_{\Omega_k^i}^{\Omega^i} \right)^{-1}(B) : B \in \mathcal{G}(\Omega_k^i), k \in \mathbb{N} \right\} \right).$$

Since the Borel sets include the open sets, it is immediate that

$$\sigma \left( \left\{ \left( \pi_{\Omega_k^i}^{\Omega^i} \right)^{-1}(B) : B \in \mathcal{B}(\Omega_k^i), k \in \mathbb{N} \right\} \right) \supseteq \sigma \left( \left\{ \left( \pi_{\Omega_k^i}^{\Omega^i} \right)^{-1}(B) : B \in \mathcal{G}(\Omega_k^i), k \in \mathbb{N} \right\} \right).$$

It remains to show the reverse inclusion. The reverse inclusion holds if

$$\left\{ \left( \pi_{\Omega_k^i}^{\Omega^i} \right)^{-1}(B) : B \in \mathcal{B}(\Omega_k^i), k \in \mathbb{N} \right\} \subseteq \sigma \left( \left\{ \left( \pi_{\Omega_k^i}^{\Omega^i} \right)^{-1}(B) : B \in \mathcal{G}(\Omega_k^i), k \in \mathbb{N} \right\} \right).$$

By Lemma 4.23 of Aliprantis and Border (2005), this holds if

$$\left\{ \left( \pi_{\Omega_k^i}^{\Omega^i} \right)^{-1}(B) : B \in \mathcal{B}(\Omega_k^i), k \in \mathbb{N} \right\} \subseteq \left\{ \left( \pi_{\Omega_k^i}^{\Omega^i} \right)^{-1}(B) : B \in \sigma(\mathcal{G}(\{\Omega_k^i : k \in \mathbb{N}\})) \right\}.$$

Using that the spaces  $\Omega_k^i$  are disjoint, where  $k \in \mathbb{N}$ , the above follows if  $\mathcal{B}(\Omega_k^i) \subseteq \sigma(\mathcal{G}(\Omega_k^i))$  for all  $k$ . But this is immediate, since the open sets generate the Borel  $\sigma$ -algebra. Conclude that  $\sigma(\pi_{\Omega_k^i}^{\Omega^i}) \subseteq \mathcal{B}(\Omega^i)$  for all  $k$ . To show that the inclusion is strict, it needs to be shown that there exists  $E \in \mathcal{B}(\Omega^i)$  such that  $E \notin \sigma(\pi_{\Omega_k^i}^{\Omega^i})$ . Taking  $E$  to be a singleton of  $\Omega^i$ , a similar argument as in the proof of Lemma B.2 can be applied, and the result follows.

Turning to the second claim, note that  $\Omega_k^i = \Theta \times B_k^j$  is Polish for all  $k$ , so that  $\mathcal{B}(\Omega_k^i) = \mathcal{B}(\Theta) \otimes \mathcal{B}(B_k^j)$ . By Lemma D.4,  $\mathcal{B}(B_k^j) \subsetneq \mathcal{B}(B_{k'}^j)$  whenever  $k < k'$ , and the result follows.  $\square$

Proposition D.2 can be used to define an extended type structure, where the types for each player are the belief hierarchies:

**Proposition D.3** *For  $i = a, b$  and  $k \in \mathbb{N}$ , let  $\mathcal{S}_k^i := \sigma(\pi_{B_{k-1}^i}^{\hat{T}^i})$  and define  $\mathcal{S}_1^i := \{\hat{T}^i, \emptyset\}$ , and  $\mathcal{S}_\infty^i := \mathcal{B}(\hat{T}^i)$ . Then, the structure*

$$\langle \hat{T}^a, \hat{T}^b, \chi^a, \chi^b, (\mathcal{S}_k^a)_{k \in \mathbb{N} \cup \{\infty\}}, (\mathcal{S}_k^b)_{k \in \mathbb{N} \cup \{\infty\}} \rangle,$$

with  $\chi^i$  a mapping from the type set  $\hat{T}^i$  to a collection of probability spaces defined by

$$\forall \hat{t}^i \in \hat{T}_k^i : \quad \chi^i(\hat{t}^i) := (\Theta \times \hat{T}^j, \mathcal{B}(\Theta) \otimes \mathcal{S}_k^j, \psi^i(\hat{t}^i))$$

is an extended type structure.

The proof of this result is relegated to the next section, as it requires some additional results. The extended type structure identified in Proposition D.3 will be referred to as the *canonical structure*, since the belief maps  $\psi^a, \psi^b$  are canonical (Proposition D.1). Additionally, since the belief map  $\psi^i$  is a surjection for each player  $i$ , there is a type with every possible belief about  $\Theta \times T^j$ , so that this extended type structure is *complete* (Brandenburger, 2003).

## D.1 Proof of Proposition D.3

### D.1.1 Preliminary results

The following results will be useful.

**Lemma D.4** *Let  $k \in \mathbb{N}, k > 1$ , and let  $\ell, m \leq k$  be such that  $m > \ell$ . Then*

$$\sigma(\pi_{\Omega_{k+1}^i}^{\Omega_\ell^i}) \subsetneq \sigma(\pi_{\Omega_{k+1}^i}^{\Omega_m^i}).$$

Moreover,

$$\sigma(\pi_{\Omega_{k+1}^i}^{\Omega_\ell^i}) \subsetneq \mathcal{B}(\Omega_{k+1}^i)$$

for all  $\ell$ .

**Proof.** To prove the first claim, suppose  $\ell < m$  for  $\ell, m \leq k$ , and suppose  $E_{\ell, k+1} \in \sigma(\pi_{\Omega_\ell^i}^{\Omega_{k+1}^i})$ . That is, there is  $E_\ell \in \mathcal{B}(\Omega_\ell^i)$  such that  $E_{\ell, k+1} = (\pi_{\Omega_\ell^i}^{\Omega_{k+1}^i})^{-1}(E_\ell)$ . To show that  $E_{\ell, k+1} \in \sigma(\pi_{\Omega_m^i}^{\Omega_{k+1}^i})$ , it suffices to show that there exists  $E_m \in \mathcal{B}(\Omega_m^i)$  such that  $E_{\ell, k+1} = (\pi_{\Omega_m^i}^{\Omega_{k+1}^i})^{-1}(E_m)$ . By construction,

$$E_{\ell, k+1} = \left(\pi_{\Omega_m^i}^{\Omega_{k+1}^i}\right)^{-1} \circ \left(\pi_{\Omega_\ell^i}^{\Omega_m^i}\right)^{-1}(E_\ell).$$

Since the projection  $\pi_{\Omega_\ell^i}^{\Omega_m^i}$  is continuous, it is Borel measurable, so that  $(\pi_{\Omega_\ell^i}^{\Omega_m^i})^{-1}(E_\ell) \in \mathcal{B}(\Omega_m^i)$ . This establishes that  $\sigma(\pi_{\Omega_\ell^i}^{\Omega_{k+1}^i}) \subseteq \sigma(\pi_{\Omega_m^i}^{\Omega_{k+1}^i})$ .

To show that the inclusion is strict, it suffices to show that there exists  $E_{m, k+1} \in \sigma(\pi_{\Omega_m^i}^{\Omega_{k+1}^i})$  such that  $E_{m, k+1} \notin \sigma(\pi_{\Omega_\ell^i}^{\Omega_{k+1}^i})$ . By Lemma B.1 below,  $\Omega_c^i$  is nonempty for every  $c \in \mathbb{N}$ . Define

$$E_m := \{(\theta^i, \mu_1^j, \dots, \mu_m^j)\}$$

for some  $(\theta^i, \mu_1^j, \dots, \mu_m^j) \in \Omega_m^i$  such that  $\mu_c^j$  is a Borel probability measure on  $\Omega_c^j$  for all  $c \leq m$ , i.e.,  $\mu_c^j \in \mathcal{M}(\Omega_c^j)$ . Note that such an element of  $\Omega_m^i$  always exists; for instance, we can take  $(\theta, \mu_1^j, \dots, \mu_m^j)$  such that  $\mu_1^j$  assigns probability 1 to some  $\theta' \in \Theta$ ,  $\mu_2^j$  assigns probability 1 to the event that  $i$  assigns probability 1 to  $\theta'$ , and so on. Again by Lemma B.1,  $\Omega_m^i$  is Polish, so that  $E_m$  is indeed a Borel set of  $\Omega_m^i$ . Hence,

$$\left(\pi_{\Omega_m^i}^{\Omega_{k+1}^i}\right)^{-1}(E_m) \in \sigma\left(\pi_{\Omega_\ell^i}^{\Omega_{k+1}^i}\right).$$

If it can be shown that there is no  $E_\ell \in \mathcal{B}(\Omega_\ell^i)$  such that

$$\left(\pi_{\Omega_\ell^i}^{\Omega_{k+1}^i}\right)^{-1}(E_\ell) = \left(\pi_{\Omega_m^i}^{\Omega_{k+1}^i}\right)^{-1}(E_m), \quad (\text{D.1})$$

then the proof of the claim is complete. Suppose by contradiction that there exists  $E_\ell$  such that (D.1) holds. Then,

$$\left(\pi_{\Omega_\ell^i}^{\Omega_m^i}\right)^{-1}(E_\ell) = E_m,$$

or, equivalently,

$$\left\{ (u^i, \nu_1^j, \dots, \nu_m^j) \in \Omega_m^i : \pi_{\Omega_\ell^i}^{\Omega_m^i}(u^i, \nu_1^j, \dots, \nu_m^j) \in E_\ell \right\} = \{(\theta^i, \mu_1^j, \dots, \mu_m^j)\}.$$

This can only hold if  $E_\ell = \{(\theta^i, \mu_1^j, \dots, \mu_\ell^j)\}$ , so that

$$\left\{ (u^i, \nu_1^j, \dots, \nu_m^j) \in \Omega_m^i : \pi_{\Omega_\ell^i}^{\Omega_m^i}(u^i, \nu_1^j, \dots, \nu_m^j) \in E_\ell \right\} = \left\{ (\theta^i, \nu_1^j, \dots, \nu_m^j) \in \Omega_m^i : \nu_c^j = \mu_c^j \text{ for } c \leq \ell \right\}.$$

Consequently, it suffices to show that this set is not a singleton. Since  $\Theta$  contains at least two elements, it follows from the proof of Lemma B.2 below that it is possible to construct two sequences of continuous functions, denoted  $g_\ell^i, g_{\ell+1}^i, \dots, g_{m-1}^i$  and  $f_\ell^i, f_{\ell+1}^i, \dots, f_{m-1}^i$ , respectively, for each player  $i$ , where, for all  $c$ ,  $g_c^i$  and  $f_c^i$  are functions from  $\mathcal{M}(\Omega_c^i)$  to  $\mathcal{M}(\Omega_{c+1}^i)$ ,  $g_c^i \neq f_c^i$ , and

$$\begin{aligned} (\theta^i, \mu_1^j, \dots, \mu_\ell^j, g_\ell^j(\mu_\ell^j), g_{\ell+1}^j(g_\ell^j(\mu_\ell^j)), \dots, g_{m-1}^j(g_{m-2}^j(\dots(g_\ell^j(\mu_\ell^j)))))) &\in \Omega_m^i, \\ (\theta^i, \mu_1^j, \dots, \mu_\ell^j, f_\ell^j(\mu_\ell^j), f_{\ell+1}^j(f_\ell^j(\mu_\ell^j)), \dots, f_{m-1}^j(f_{m-2}^j(\dots(f_\ell^j(\mu_\ell^j)))))) &\in \Omega_m^i, \end{aligned}$$

i.e., the set of extensions of  $(\theta^i, \mu_1^j, \dots, \mu_\ell^j)$  to  $\Omega_m^i$  is not a singleton.

Now turn to the second claim. Again, it follows directly from the fact that  $\pi_{\Omega_\ell^i}^{\Omega_{k+1}^i}$  is Borel measurable that  $\sigma(\pi_{\Omega_\ell^i}^{\Omega_{k+1}^i}) \subseteq \mathcal{B}(\Omega_{k+1}^i)$ . The proof that the inclusion is strict is identical to the proof above, and thus omitted.  $\square$

**Lemma D.5** *The Borel  $\sigma$ -algebra on  $\mathcal{M}^+(\Omega_{k+1}^i)$  induced by the topology generated by  $\rho_{k+1}^i$  coincides with the  $\sigma$ -algebra generated by sets of the form*

$$\{\mu \in \mathcal{M}(\Omega_{k+1}^i) : \mu(E) \geq p\} : \quad E \in \mathcal{B}(\Omega_{k+1}^i), p \in [0, 1],$$

and

$$\{\mu \in \mathcal{M}_\ell(\Omega_{k+1}^i) : \mu(E) \geq p\} : \quad E \in \sigma(\pi_{\Omega_\ell^i}^{\Omega_{k+1}^i}), p \in [0, 1],$$

for  $\ell = 1, \dots, k$ .

**Proof.** First consider the subspace  $\mathcal{M}(\Omega_{k+1}^i)$ , endowed with the topology induced by the Prohorov metric  $\tilde{\rho}_{k+1}^i$ . The Prohorov metric metrizes the weak topology. It follows from the proof of Lemma 14.16 of Aliprantis and Border (2005) that the Borel  $\sigma$ -algebra induced by the weak topology coincides with the  $\sigma$ -algebra generated by the maps  $\mu \mapsto \mu(E)$  for  $E \in \mathcal{B}(\Omega_{k+1}^i)$ . But this  $\sigma$ -algebra is identical to the  $\sigma$ -algebra generated by sets of the form

$$\{\mu \in \mathcal{M}(\Omega_{k+1}^i) : \mu(E) \geq p\},$$

for  $E \in \mathcal{B}(\Omega_{k+1}^i)$  and  $p \in [0, 1]$  (e.g. Kechris, 1995, p. 67).

Next consider a subspace  $\mathcal{M}_\ell(\Omega_{k+1}^i)$ , where  $\ell = 1, \dots, k$ . By the same argument as above, the  $\sigma$ -algebra on  $\mathcal{M}(\Omega_\ell^i)$  induced by the weak topology (metrized by the Prohorov metric  $\tilde{\rho}_\ell^i$ ), denoted  $\mathcal{Q}_\ell$ , coincides with the  $\sigma$ -algebra on  $\mathcal{M}(\Omega_\ell^i)$  generated by sets of the form

$$\{\mu \in \mathcal{M}(\Omega_\ell^i) : \mu(E) \geq p\},$$

for  $E \in \mathcal{B}(\Omega_\ell^i)$  and  $p \in [0, 1]$ , which will be denoted  $\mathcal{Q}'_\ell$ . Define a function  $f : \mathcal{M}_\ell(\Omega_{k+1}^i) \rightarrow \mathcal{M}(\Omega_\ell^i)$  by:

$$\forall \mu \in \mathcal{M}_\ell(\Omega_{k+1}^i) : \quad f(\mu) = \text{marg}_{\Omega_\ell^i} \mu.$$

For a function  $g : X \rightarrow Y$  and a collection of subsets  $\mathcal{E}$  of  $Y$ , define

$$g^{-1}(\mathcal{E}) := \{g^{-1}(E) : E \in \mathcal{E}\}.$$

It is immediate that the fact that  $\mathcal{Q}_\ell$  and  $\mathcal{Q}'_\ell$  coincide implies that

$$f^{-1}(\mathcal{Q}_\ell) = f^{-1}(\mathcal{Q}'_\ell).$$

Clearly, the proof is complete if the following claims hold:

**Claim 1.**  $f^{-1}(\mathcal{Q}'_\ell)$  is the  $\sigma$ -algebra on  $\mathcal{M}_\ell(\Omega_{k+1}^i)$  generated by sets of the form

$$\{\mu \in \mathcal{M}_\ell(\Omega_{k+1}^i) : \mu(E) \geq p\} : \quad E \in \sigma(\pi_{\Omega_\ell^i}^{\Omega_{k+1}^i}), p \in [0, 1].$$

**Proof of Claim 1.** Using Lemma 4.23 of Aliprantis and Border (2005), it holds that

$$\begin{aligned} f^{-1}(\mathcal{Q}'_\ell) &= f^{-1}\left(\sigma(\{\{\mu \in \mathcal{M}(\Omega_\ell^i) : \mu(E) \geq p\} : E \in \mathcal{B}(\Omega_\ell^i), p \in [0, 1]\})\right) \\ &= \sigma\left(\{f^{-1}(\{\mu \in \mathcal{M}(\Omega_\ell^i) : \mu(E) \geq p\}) : E \in \mathcal{B}(\Omega_\ell^i), p \in [0, 1]\}\right) \\ &= \sigma\left(\{\{\mu \in \mathcal{M}_\ell(\Omega_{k+1}^i) : \mu(E) \geq p\} : E \in \sigma(\pi_{\Omega_\ell^i}^{\Omega_{k+1}^i}), p \in [0, 1]\}\right), \end{aligned}$$

which proves the claim. □

**Claim 2.**  $f^{-1}(\mathcal{Q}_\ell)$  is the  $\sigma$ -algebra on  $\mathcal{M}_\ell(\Omega_{k+1}^i)$  induced by the relative topology generated by  $\rho_{k+1}^i$ .

**Proof of Claim 2.** We first introduce some notation. Given a metric space  $(X, d)$  and  $r > 0$ , let

$$B_r(x) := \{z \in Z : d(x, z) < r\}$$

be the open ball centered at  $x$  of radius  $r$ . A subset  $A$  of  $X$  is  $d$ -open if for each  $a \in A$ , there exists  $r > 0$  (dependent on  $a$ ) such that  $B_r(a) \subseteq A$ . Then, using that the open sets generate the Borel  $\sigma$ -algebra, it follows from Lemma 4.23 of Aliprantis and Border (2005) that

$$\begin{aligned} f^{-1}(\mathcal{Q}_\ell) &= f^{-1}(\sigma(\{A \subseteq \mathcal{M}(\Omega_\ell^i) : A \text{ is } \tilde{\rho}_\ell^i\text{-open}\})) \\ &= \sigma(f^{-1}(\{A \subseteq \mathcal{M}(\Omega_\ell^i) : A \text{ is } \tilde{\rho}_\ell^i\text{-open}\})). \end{aligned}$$

The following lemmas will be helpful:

**Lemma D.6** *Let  $A \subseteq \mathcal{M}(\Omega_\ell^i)$ . Then  $A$  is  $\tilde{\rho}_\ell^i$ -open if and only if  $f^{-1}(A)$  is  $\rho_{k+1}^i$ -open.*

**Proof.** Suppose  $A \subseteq \mathcal{M}(\Omega_\ell^i)$  is  $\tilde{\rho}_\ell^i$ -open. Let  $b \in f^{-1}(A)$ , that is,  $\text{marg}_{\Omega_\ell^i} b \in A$ . Then there exists  $r > 0$  such that

$$\{\mu_\ell \in \mathcal{M}(\Omega_\ell^i) : \tilde{\rho}_\ell^i(\text{marg}_{\Omega_\ell^i} b, \mu_\ell) < r\}.$$

Using the definition of  $\rho_{k+1}^i$ , it is sufficient to show that

$$\{\mu_{k+1} \in \mathcal{M}_\ell(\Omega_{k+1}^i) : \tilde{\rho}_\ell^i(\text{marg}_{\Omega_\ell^i} b, \text{marg}_{\Omega_{k+1}^i} \mu_{k+1}) < r\} \subseteq f^{-1}(A) = \{\mu_{k+1} \in \mathcal{M}_\ell(\Omega_{k+1}^i) : \text{marg}_{\Omega_\ell^i} \mu_{k+1} \in A\}.$$

To show this, suppose that  $\mu_{k+1} \in \mathcal{M}_\ell(\Omega_{k+1}^i)$  is such that  $\tilde{\rho}_\ell^i(\text{marg}_{\Omega_\ell^i} b, \text{marg}_{\Omega_{k+1}^i} \mu_{k+1}) < r$ , and define  $\mu_\ell := \text{marg}_{\Omega_\ell^i} \mu_{k+1}$ . Then, by assumption,  $\tilde{\rho}_\ell^i(\text{marg}_{\Omega_\ell^i} b, \mu_\ell) < r$ , and  $\mu_\ell = \text{marg}_{\Omega_\ell^i} \mu_{k+1} \in A$ . It follows that  $\mu_{k+1} \in f^{-1}(A)$ .

To prove the other direction, fix  $a \in A$ , and let  $b \in \mathcal{M}_\ell(\Omega_{k+1}^i)$  be such that  $\text{marg}_{\Omega_\ell^i} b = a$ . By assumption, there exists  $r > 0$  such that

$$\{\mu_{k+1} \in \mathcal{M}_\ell(\Omega_{k+1}^i) : \tilde{\rho}_\ell^i(\mu_{k+1}, a) < r\} \subseteq \{\mu_{k+1} \in \mathcal{M}_\ell(\Omega_{k+1}^i) : \text{marg}_{\Omega_\ell^i} \mu_{k+1} \in A\}. \quad (\text{D.2})$$

It is sufficient to show that

$$\{\mu_\ell \in \mathcal{M}(\Omega_\ell^i) : \tilde{\rho}_\ell^i(\mu_\ell, a) < r\} \subseteq A.$$

Suppose  $\mu_\ell \in \mathcal{M}(\Omega_\ell^i)$  is such that  $\tilde{\rho}_\ell^i(\mu_\ell, a) < r$ , and let  $\mu_{k+1} \in \mathcal{M}_\ell(\Omega_{k+1}^i)$  be such that  $\mu_\ell = \text{marg}_{\Omega_\ell^i} \mu_{k+1}$ , so that  $\tilde{\rho}_\ell^i(\text{marg}_{\Omega_\ell^i} \mu_{k+1}, a) < r$ . By (D.2),  $\text{marg}_{\Omega_\ell^i} \mu_{k+1} \in A$ , so  $\mu_\ell \in A$ .  $\square$

**Lemma D.7** *The collection of sets  $\{f^{-1}(A) \subseteq \mathcal{M}_\ell(\Omega_{k+1}^i) : A \text{ is } \tilde{\rho}_\ell^i\text{-open}, A \in \mathcal{M}(\Omega_\ell^i)\}$  is equal to the collection  $\{B \subseteq \mathcal{M}_\ell(\Omega_{k+1}^i) : B \text{ is } \tilde{\rho}_\ell^i\text{-open}\}$ .*

**Proof.** Clearly,

$$\{f^{-1}(A) \subseteq \mathcal{M}_\ell(\Omega_{k+1}^i) : A \text{ is } \tilde{\rho}_\ell^i\text{-open}, A \in \mathcal{M}(\Omega_\ell^i)\} \subseteq \{B \subseteq \mathcal{M}_\ell(\Omega_{k+1}^i) : B \text{ is } \tilde{\rho}_\ell^i\text{-open}\},$$

so it remains to show the other inclusion. Let  $B \subseteq \mathcal{M}_\ell(\Omega_{k+1}^i)$  such that  $B$  is  $\tilde{\rho}_\ell^i$ -open. We need to show that there is  $A \subseteq \mathcal{M}(\Omega_\ell^i)$  such that  $B = f^{-1}(A)$ . But  $f$  is a bijection, so we can take  $f(A) = B$ . Since  $B$  is open and  $f$  is continuous,  $f^{-1}(B) = A$  is open.  $\square$

The proof of Lemma D.5 now follows directly.  $\square$

**Lemma D.8** *The Borel  $\sigma$ -algebra on  $\mathcal{M}^+(\Theta \times \widehat{T}^j)$  induced by the topology generated by the metric  $\rho_{\widehat{T}^j}^i$  is generated by sets of the form*

$$\{\mu \in \mathcal{M}^+(\Theta \times \widehat{T}^j) : E \in \Sigma(\mu), \mu(E) \geq p\} : \quad E \in \mathcal{B}(\Theta) \otimes \mathcal{B}(\widehat{T}^j), p \in [0, 1],$$

where  $\Sigma(\mu)$  denotes the  $\sigma$ -algebra on which the probability measure  $\mu$  is defined.

**Proof.** The set of types  $\widehat{T}^i$  for player  $i$  is endowed with the relative topology of the product topology on  $\times_{k \in \mathbb{N}} \mathcal{M}^+(\Omega_k^i)$ . Using Lemma D.5, observe that the Borel  $\sigma$ -algebra  $\mathcal{B}(\widehat{T}^i)$  is generated by sets of the form

$$\{(\mu_1^i, \mu_2^i, \dots) \in \widehat{T}^i : E \in \Sigma(\mu_k^i), \mu_k^i(E) \geq p\} : \quad k \in \mathbb{N}, E \in \mathcal{B}(\Omega_k^i), p \in [0, 1],$$

where  $\Sigma(\mu)$  is the  $\sigma$ -algebra associated with the probability measure  $\mu$ . Using that  $\psi^i$  is canonical, it follows that  $\mathcal{B}(T^i)$  is generated by sets of the form

$$\{\hat{t} \in \widehat{T}^i : E \in \Sigma^i(\hat{t}^i), \psi^i(\hat{t}^i)(E) \geq p\} : \quad E \in \mathcal{B}(\Theta) \otimes \mathcal{B}(\widehat{T}^j), p \in [0, 1].$$

To show that the Borel  $\sigma$ -algebra on  $\mathcal{M}^+(\Theta \times T^j)$  is generated by sets of the form

$$\{\mu \in \mathcal{M}^+(\Theta \times \widehat{T}^j) : \mu(E) \geq p\} : \quad E \in \mathcal{B}(\Theta) \otimes \mathcal{B}(\widehat{T}^j), p \in [0, 1],$$

it is useful to start with some preliminary observations. Consider a homeomorphism  $f$  from a topological space  $X$  to a topological space  $Y$ . Since a homeomorphism and its inverse are both continuous, the functions  $f$  and  $f^{-1}$  are Borel measurable, so that

$$\mathcal{B}(Y) = \{f(B) : B \in \mathcal{B}(X)\}.$$

By Proposition D.2,  $\psi^i$  is a homeomorphism. Using Lemma D.9 and that  $\psi^i$  is surjective, it follows that

$$\begin{aligned} \mathcal{B}(\Theta \times \widehat{T}^j) &= \{\psi^i(B) : B \in \mathcal{B}(\widehat{T}^i)\} \\ &= \sigma\left(\{\psi^i(\{\hat{t}^i \in \widehat{T}^i : E \in \Sigma^i(\hat{t}^i), \psi^i(\hat{t}^i)(E) \geq p\}) : \right. \\ &\quad \left. E \in \mathcal{B}(\Theta) \times \mathcal{B}(\widehat{T}^j), p \in [0, 1]\}\right) \\ &= \sigma\left(\{\{\mu \in \mathcal{M}^+(\Theta \times \widehat{T}^j) : E \in \Sigma(\mu), \mu(E) \geq p\} : \right. \\ &\quad \left. E \in \mathcal{B}(\Theta) \times \mathcal{B}(\widehat{T}^j), p \in [0, 1]\}\right), \end{aligned}$$

which is what needed to be shown. □

**Lemma D.9** *The  $\sigma$ -algebra  $\sigma(\pi_{B_{k-1}^i}^{\widehat{T}^i})$  on  $\widehat{T}^i$  is generated by sets of the form*

$$\{\hat{t}^i \in \widehat{T}^i : E \in \Sigma^i(\hat{t}^i), \psi^i(\hat{t}^i)(E) \geq p\} : \quad E \in \mathcal{B}(\Theta) \otimes \mathcal{S}_{k-1}^j, p \in [0, 1].$$

**Proof.** The cases  $k = 2$  and  $k > 2$  are treated separately for notational reasons, but the structure of the proofs is similar. First suppose  $k = 2$ . Recall that, since  $\Theta$  is Polish, the Borel  $\sigma$ -algebra on  $B_1^i = \mathcal{M}(\Theta)$  is generated by sets of the form

$$\{\mu \in \mathcal{M}(\Theta) : \mu(E) \geq p\} : \quad E \in \mathcal{B}(\Theta), p \in [0, 1].$$

Then, using that the structure is canonical,

$$\begin{aligned} \sigma(\pi_{\mathcal{M}(\Theta)}^{\widehat{T}^i}) &= \{(\pi_{\mathcal{M}(\Theta)}^{\widehat{T}^i})^{-1}(B) : B \in \mathcal{B}(\Theta)\} \\ &= \sigma\left(\{(\pi_{\mathcal{M}(\Theta)}^{\widehat{T}^i})^{-1}(\{\mu \in \mathcal{M}(\Theta) : \mu(E) \geq p\}) : E \in \mathcal{B}(\Theta), p \in [0, 1]\}\right) \\ &= \sigma\left(\{\{\hat{t}^i \in \widehat{T}^i : \text{marg}_{\Theta} \psi^i(\hat{t}^i)(E) \geq p\} : E \in \mathcal{B}(\Theta), p \in [0, 1]\}\right). \end{aligned}$$

It follows that  $\sigma(\pi_{\mathcal{M}(\Theta)}^{\widehat{T}^i})$  is generated by sets of the form

$$\{\{\hat{t}^i \in \widehat{T}^i : E \in \Sigma^i(\hat{t}^i), \psi^i(\hat{t}^i)(E) \geq p\} : E \in \mathcal{B}(\Theta) \otimes \{\widehat{T}^j, \emptyset\}, p \in [0, 1].$$

Let  $k > 2$ . Using that the belief hierarchies are coherent, it holds that

$$\sigma(\pi_{\mathcal{M}(\Theta)}^{\widehat{T}^i}) = \{(\pi_{\mathcal{M}^+(\Omega_k^i)}^{\widehat{T}^i})^{-1}(B) : B \in \mathcal{B}(\mathcal{M}^+(\Omega_k^i))\}.$$

By Lemma D.5, and using that  $\sigma(\pi_{\Omega_\ell^i}^{\Omega_k^i}) \subseteq \mathcal{B}(\Omega_k^i)$  for  $\ell = 1, \dots, k-1$ , it follows that the Borel  $\sigma$ -algebra on  $\mathcal{M}^+(\Omega_k^i)$  is generated by sets of the form

$$\{\mu \in \mathcal{M}^+(\Omega_k^i) : E \in \Sigma(\mu), \mu(E) \geq p\} : \quad E \in \mathcal{B}(\Omega_k^i), p \in [0, 1],$$

where  $\Sigma(\mu)$  for a probability measure  $\mu$  denotes the  $\sigma$ -algebra on which that measure is defined.

Again using that the structure is canonical and that  $\Omega_k^i = \Theta \times B_{k-1}^j$ , it holds that

$$\begin{aligned} \sigma(\pi_{\mathcal{M}(\Theta)}^{\widehat{T}^i}) &= \{(\pi_{\mathcal{M}^+(\Omega_k^i)}^{\widehat{T}^i})^{-1}(B) : B \in \mathcal{B}(\mathcal{M}^+(\Omega_k^i))\} \\ &= \sigma\left(\left\{(\pi_{\mathcal{M}^+(\Omega_k^i)}^{\widehat{T}^i})^{-1}(\{\mu \in \mathcal{M}^+(\Omega_k^i) : E \in \Sigma(\mu), \mu(E) \geq p\}) \geq p : \right. \right. \\ &\quad \left. \left. E \in \mathcal{B}(\Omega_k^i), p \in [0, 1]\right\}\right) \\ &= \sigma\left(\left\{\{\hat{t}^i \in \widehat{T}^i : E \in \Sigma(\text{marg}_{\Theta \times B_{k-1}^j} \psi^i(\hat{t}^i)), \psi^i(\hat{t}^i)(E) \geq p\} : \right. \right. \\ &\quad \left. \left. E \in \mathcal{B}(\Theta) \otimes \mathcal{B}(\Theta \otimes B_{k-1}^j)\right\}\right) \\ &= \sigma\left(\left\{\{\hat{t}^i \in \widehat{T}^i : E \in \Sigma^i(\hat{t}^i), \psi^i(\hat{t}^i)(E) \geq p\} : E \in \mathcal{B}(\Theta) \otimes \sigma(\pi_{B_{k-1}^j}^{\widehat{T}^j})\right\}\right). \end{aligned}$$

□

### D.1.2 Proof

Consider the structure  $\langle \widehat{T}^a, \widehat{T}^b, \chi^a, \chi^b, (\mathcal{S}_k^a)_{k \in \mathbb{N} \cup \{\infty\}}, (\mathcal{S}_k^b)_{k \in \mathbb{N} \cup \{\infty\}} \rangle$  defined in the proposition, with  $\mathcal{S}_k^i := \sigma(\pi_{B_{k-1}^i}^{\widehat{T}^i})$ ,  $k \geq 2$ , and  $\mathcal{S}_1^i = \{\widehat{T}^i, \emptyset\}$ , and  $\mathcal{S}_\infty^i := \mathcal{B}(\widehat{T}^i)$ . For ease of reference, denote the structure by  $\mathcal{T}$ . To show that the structure  $\mathcal{T}$  is an extended type structure, it needs to be verified that there exists  $\mathcal{Z} \in \mathbb{N} \cup \{\infty\}$  such that Conditions 3.1–3.4 are satisfied for both players.

It follows from Lemma D.4 that **Condition 3.1** is satisfied if and only if  $\mathcal{Z} = \infty$ . To show that **Conditions 3.2** and **Condition 3.3** hold, it is sufficient to show that

$$\sigma(\pi_{B_{k-1}^j}^{\widehat{T}^j}) \succ \sigma(\pi_{B_{\ell-1}^i}^{\widehat{T}^i}) \iff k > \ell. \quad (\text{D.3})$$

To show this, fix  $k < \infty$ , and suppose  $\ell < k$ , so that  $k \geq 2$ . Let  $E \in \mathcal{B}(\Theta) \otimes \mathcal{S}_\ell^j = \mathcal{B}(\Theta) \otimes \sigma(\pi_{B_{\ell-1}^j}^{\widehat{T}^j})$  and  $p \in [0, 1]$ . It then follows directly from Lemma D.9 that

$$\{\hat{t}^i \in \widehat{T}^i : E \in \Sigma^i(\hat{t}^i), \psi^i(\hat{t}^i)(E) \geq p\} \in \sigma(\pi_{B_{k-1}^i}^{\widehat{T}^i}).$$

Conversely, suppose that for every  $E \in \mathcal{B}(\Theta) \otimes \mathcal{S}_\ell^j$  and  $p \in [0, 1]$ ,

$$\{\hat{t}^i \in \widehat{T}^i : E \in \Sigma^i(\hat{t}^i), \psi^i(\hat{t}^i)(E) \geq p\} \in \mathcal{S}_k^i.$$

We need to show that  $k > \ell$ . Suppose by contradiction that  $k \leq \ell$ . It will be convenient to distinguish two cases. First consider the case  $\ell = 1$ , so that  $k = 1$ . Let  $E = (\theta, \widehat{T}^j)$  for some  $\theta \in \Theta$ , and fix  $p \in [0, 1]$ . By assumption,

$$\{\hat{t}^i \in \widehat{T}^i : E \in \Sigma^i(\hat{t}^i), \psi^i(\hat{t}^i)(E) \geq p\} \in \{\widehat{T}^i, \emptyset\}.$$

Note that  $E \in \Sigma^i(\hat{t}^i)$  for every type  $\hat{t}^i \in \widehat{T}^i$ . Since  $\Theta$  contains at least two elements, and  $B_1^i = \mathcal{M}(\Theta)$ , there exist  $\hat{t}^i, u^i \in \widehat{T}^i$  such that  $\psi^i(\hat{t}^i)(E) \geq p > \psi^i(u^i)(E)$ , a contradiction.

Next suppose  $\ell \geq 2$ . If  $k = 1$ , it is easy to see that a contradiction results using a similar argument as above, so suppose  $k \geq 2$  (where  $k \leq \ell$ ). Recall that the Borel  $\sigma$ -algebra on a Polish space contains the singletons. Fix  $\tilde{\mu}_{\ell-1} := (\tilde{\mu}_1^j, \dots, \tilde{\mu}_{\ell-1}^j) \in B_{\ell-1}^j$ , and let  $E = (\theta, \{(\mu_1^j, \dots, \mu_{\ell-1}^j, \mu_\ell^j, \dots) \in \widehat{T}^j : (\mu_1^j, \dots, \mu_{\ell-1}^j) = \tilde{\mu}\})$  for some  $\theta \in \Theta$ , and let  $p \in (0, 1)$ . Note that if a contradiction is obtained for  $k = \ell$ , then a contradiction will result for any  $k < \ell$ . So suppose  $k = \ell$ . Let  $u^i, z^i \in \widehat{T}_\infty^i$  (so that  $E \in \Sigma^i(u^i), \Sigma^i(z^i)$ ) such that  $\pi_{B_{\ell-1}^i}^{\widehat{T}^i}(u^i) = \pi_{B_{\ell-1}^i}^{\widehat{T}^i}(z^i)$  and

$$\text{marg}_{\Theta \times B_\ell^j} \psi^i(u^i)(E) = 1, \quad \text{marg}_{\Theta \times B_\ell^j} \psi^i(z^i)(E) = 0.$$

Then, for any  $B \in \sigma(\pi_{B_{\ell-1}^i}^{\widehat{T}^i})$ , either  $u^i, z^i \in B$  or it is the case that  $u^i \notin B$  and  $z^i \notin B$ , that is,  $\sigma(\pi_{B_{\ell-1}^i}^{\widehat{T}^i})$  does not separate  $u^i$  and  $z^i$ . But

$$u^i \in \{\hat{t}^i \in \widehat{T}^i : E \in \Sigma^i(\hat{t}^i), \psi^i(\hat{t}^i)(E) \geq p\},$$

and

$$z^i \notin \{\hat{t}^i \in \hat{T}^i : E \in \Sigma^i(\hat{t}^i), \psi^i(\hat{t}^i)(E) \geq p\},$$

a contradiction.

To see that **Condition 3.4** is satisfied, note that  $\psi^i$  is continuous, so that it is Borel measurable. It follows that for all  $M \in \mathcal{B}(\mathcal{M}^+(\Theta \times \hat{T}^j))$ ,

$$\{\hat{t}^i \in \hat{T}^i : \psi^i(\hat{t}^i) \in M\} \in \mathcal{B}(\hat{T}^i).$$

By Lemma D.8, the Borel  $\sigma$ -algebra on  $\mathcal{M}^+(\Theta \times \hat{T}^j)$  is generated by sets of the form

$$\{\mu \in \mathcal{M}^+(\Theta \times \hat{T}^j) : E \in \Sigma(\mu), \mu(E) \geq p\} : \quad E \in \mathcal{B}(\Theta) \otimes \mathcal{B}(\hat{T}^j), p \in [0, 1],$$

where  $\Sigma(\mu)$  denotes the  $\sigma$ -algebra associated with a probability measure  $\mu$ . Using Corollary 4.24 of Aliprantis and Border (2005) and the fact that  $\psi$  is Borel measurable, it follows that for every  $E \in \mathcal{B}(\Theta) \otimes \mathcal{B}(\hat{T}^j)$  and  $p \in [0, 1]$ ,

$$\{\hat{t}^i \in \hat{T}^i : E \in \Sigma^i(\hat{t}^i), \psi^i(\hat{t}^i)(E) \geq p\} \in \mathcal{B}(\hat{T}^i).$$

Moreover, it follows from (D.3) that there is no  $k \in \mathbb{N}$  such that  $\mathcal{S}_k^i \succ \mathcal{S}_\infty^j$ . □

## References

- Ahn, D. S. (2007). Hierarchies of ambiguous beliefs. *Journal of Economic Theory* 136, 286–301.
- Aliprantis, C. D. and K. C. Border (2005). *Infinite Dimensional Analysis: A Hitchhiker's Guide* (3rd ed.). Berlin: Springer.
- Aumann, R. (1992). Irrationality in game theory. In P. Dasgupta, D. Gale, O. Hart, and E. Maskin (Eds.), *Economic Analysis of Markets and Games: Essays in Honor of Frank Hahn*. Cambridge: MIT Press.
- Aumann, R. and A. Brandenburger (1995). Epistemic conditions for Nash equilibrium. *Econometrica* 63, 1161–1180.
- Bacharach, M. (1993). Variable universe games. In K. Binmore, A. Kirman, and P. Tani (Eds.), *Frontiers of Game Theory*. Cambridge, MA: MIT Press.
- Bacharach, M. and D. O. Stahl (2000). Variable-frame level- $n$  theory. *Games and Economic Behavior* 32, 220–246.

- Battigalli, P. and M. Siniscalchi (2003). Rationalization and incomplete information. *B. E. Journal of Theoretical Economics* 3, article 3.
- Billingsley, P. (1999). *Convergence of Probability Measures* (2nd ed.). Wiley Series in Probability and Statistics. New York: Wiley.
- Blume, A. (2000). Coordination and learning with a partial language. *Journal of Economic Theory* 95, 1–36.
- Blume, A. and U. Gneezy (2010). Cognitive forward induction and coordination without common knowledge: An experimental study. *Games and Economic Behavior* 68, 488–511.
- Board, O. and K.-S. Chung (2006). Object-based unawareness. Working paper, University of Pittsburgh.
- Brandenburger, A. (2003). On the existence of a ‘complete’ possibility structure. In M. Basili, N. Dimitri, and I. Gilboa (Eds.), *Cognitive processes and economic behavior*, pp. 30–34. London: Routledge.
- Brandenburger, A. and E. Dekel (1993). Hierarchies of beliefs and common knowledge. *Journal of Economic Theory* 59, 189–198.
- Brandenburger, A., A. Friedenberg, and H. J. Keisler (2008). Admissibility in games. *Econometrica* 76, 307–352.
- Camerer, C. F., T.-H. Ho, and J.-K. Chong (2004). A cognitive hierarchy model of games. *Quarterly Journal of Economics* 119(3), 861–898.
- Carlsson, H. and E. van Damme (1993). Global games and equilibrium selection. *Econometrica* 61, 989–1018.
- Costa-Gomes, M., V. P. Crawford, and B. Broseta (2001). Cognition and behavior in normal-form games: An experimental study. *Econometrica* 69, 1193–1235.
- Crawford, V. P. and H. Haller (1990). Learning how to cooperate: Optimal play in repeated coordination games. *Econometrica* 58, 581–596.
- Crawford, V. P. and N. Iriberri (2007). Level- $k$  auctions: Can a nonequilibrium model of strategic thinking explain the winners curse and overbidding in private-value auctions? *Econometrica* 75, 1721–1770.

- Dekel, E., D. Fudenberg, and S. Morris (2007). Interim correlated rationalizability. *Theoretical Economics* 2, 15–40.
- Di Tillio, A. (2008). Subjective expected utility in games. *Theoretical Economics* 3, 287–323.
- Fagin, R., J. Y. Halpern, and M. Y. Vardi (1991). A model-theoretic analysis of knowledge. *Journal of the ACM* 38, 382–428.
- Feinberg, Y. (2004). Subjective reasoning: Games with unawareness. Working paper, Stanford GSB.
- Feinberg, Y. and A. Skrzypacz (2005). Uncertainty about uncertainty and delay in bargaining. *Econometrica* 73, 69–91.
- Friedenberg, A. and M. Meier (2007). The context of the game. Working paper.
- Geanakoplos, J. (1989). Game theory without partitions and applications to speculation and consensus. Working paper, Cowles Foundation for Research in Economics.
- Geanakoplos, J. D. and H. M. Polemarchakis (1982). We can’t disagree forever. *Journal of Economic Theory* 28, 192–200.
- Halpern, J. Y. (2003). *Reasoning about Uncertainty*. Cambridge, MA: MIT Press.
- Halpern, J. Y. and Y. Moses (1990). Knowledge and common knowledge in a distributed environment. *Journal of the ACM* 37, 549–587.
- Halpern, J. Y. and L. Rego (2006). Extensive games with possibly unaware players. In *Proceedings of the Fifth International Joint Conference on Autonomous Agents and Multiagent Systems*, pp. 744–751.
- Harsanyi, J. C. (1967–1968). Games on incomplete information played by Bayesian players. Parts I–III. *Management Science* 14, 159–182, 320–334, 486–502.
- Heifetz, A. (1993). The Bayesian formulation of incomplete information—The non-compact case. *International Journal of Game Theory* 21, 329–338.
- Heifetz, A. and W. Kets (2010). A grain of finite-level reasoning. Working paper.
- Heifetz, A., M. Meier, and B. Schipper (2006). Interactive unawareness. *Journal of Economic Theory* 130, 78–94.

- Heifetz, A. and D. Samet (1998a). Knowledge spaces with arbitrarily high rank. *Games and Economic Behavior* 22, 260–273.
- Heifetz, A. and D. Samet (1998b). Topology-free typology of beliefs. *Journal of Economic Theory* 82, 324–341.
- Ho, T.-H., C. Camerer, and K. Weigelt (1998). Iterated dominance and iterated best response in experimental “ $p$ -beauty contests”. *American Economic Review* 88, 947–969.
- Hocking, J. G. and G. S. Young (1988). *Topology*. Mineola, NY: Dover Publications.
- Jehiel, P. (2005). Analogy-based expectation equilibrium. *Journal of Economic Theory* 123, 81–104.
- Jehiel, P. and F. Koessler (2008). Revisiting games of incomplete information with analogy-based expectations. *Games and Economic Behavior* 62, 533–557.
- Kajii, A. and S. Morris (1997). The robustness of equilibria to incomplete information. *Econometrica* 65, 1283–1309.
- Kechris, A. S. (1995). *Classical Descriptive Set Theory*. Graduate Texts in Mathematics. Berlin: Springer-Verlag.
- Kets, W. (2009). Do you think about what I think you think? Finite belief hierarchies in games. Working paper.
- Li, J. (2009). Information structures with unawareness. *Journal of Economic Theory* 144, 977–993.
- Mengel, F. (2007). Learning across games. Working paper, Maastricht University.
- Mertens, J. F. and S. Zamir (1985). Formulation of Bayesian analysis for games with incomplete information. *International Journal of Game Theory* 14, 1–29.
- Monderer, D. and D. Samet (1989). Approximating common knowledge with common beliefs. *Games and Economic Behavior* 1, 170–190.
- Morris, S., A. Postlewaite, and H. Shin (1995). Depth of knowledge and the effect of higher order uncertainty. *Economic Theory* 6, 453–467.
- Nagel, R. (1995). Unraveling in guessing games: An experimental study. *American Economic Review* 85, 1313–1326.

- Neeman, Z. (2004). The relevance of private information in mechanism design. *Journal of Economic Theory* 117, 55–77.
- Nielsen, C. R. (2010). Stepwise thinking in strategic games with incomplete information. Working paper, University of Copenhagen.
- Parthasarathy, K. R. (1978). *Introduction to probability and measure*. New York: Springer-Verlag.
- Qin, C. and C. Yang (2010). An explicit approach to modeling finite-order type spaces and applications. Working paper, U.C. Santa Barbara.
- Rogers, B. W., T. R. Palfrey, and C. F. Camerer (2009). Heterogeneous quantal response equilibrium and cognitive hierarchies. *Journal of Economic Theory* 144, 1440–1467.
- Rubinstein, A. (1989). The electronic mail game: Strategic behavior under “almost common knowledge”. *American Economic Review* 79, 385–391.
- Savage, L. J. (1954). *The Foundations of Statistics*. John Wiley & Sons.
- Stahl, D. O. and P. W. Wilson (1995). On players’ models of other players: Theory and experimental evidence. *Games and Economic Behavior* 10, 218–254.
- Stinchcombe, M. (1988). Approximate common knowledge. Working paper, U.C. San Diego.
- Strzalecki, T. (2009). Depth of reasoning and higher order beliefs. Working paper, Harvard University.
- Weinstein, J. and M. Yildiz (2007). A structure theorem for rationalizability with application to robust predictions of refinements. *Econometrica* 75, 365–400.