*Bounded Reasoning and
Higher-Order Uncertainty*

Willemien Kets
Northwestern University
June 24, 2012

**NORTHWESTERN
UNIVERSITY**

Bounded Reasoning and Higher-Order Uncertainty

Willemien Kets[*]

This version: June 24, 2012

First version: November 2009

Abstract

The standard framework for analyzing games with incomplete information models players as if they have an infinite depth of reasoning. This paper generalizes the type spaces of Harsanyi (1967–1968) so that players can have a finite depth of reasoning. The innovation is that players can have a coarse perception of the higher-order beliefs of other players, thus formalizing the small-world idea of Savage (1954) in a type-space context. Unlike in other models of finite-order reasoning, players with a finite depth of reasoning can reason about higher-order events if these events are generated by events of sufficiently low order. In particular, an event F can be common belief if it is entailed by some public event. This is true even if players cannot reason about higher-order statements like "Ann believes that Bob believes that Ann believes...(58 times)...that F" in isolation. Thus, the usual equivalence between the iterative and the fixed-point account of common belief breaks down when players have a finite depth, and common belief is easier to attain than as suggested by the iterative approach.

1. Introduction

Analyzing games of incomplete information requires taking into account not only the beliefs of players, but also their *higher-order beliefs*. Consider, for example, a player who needs to decide which project to invest in. The payoff associated with each choice depends on the economic fundamentals—the state of nature—, as well as the actions of other investors. The player's optimal decision thus depends on her beliefs about the state of nature and the actions of the other players. Because the same is true for her opponents, the player's optimal action will also depend on her belief about her opponents' beliefs about the state of nature, that is, a second-order belief. And because her opponents in turn condition their action on *their* beliefs about their opponents' beliefs about the state of nature, the player's action choice will depend on her belief about her opponents' beliefs about their opponents' beliefs about nature (a third-order belief), and so on, ad infinitum (Harsanyi, 1967–1968).

Are "real" players capable of such higher-order reasoning? The answer to this question is not so clear-cut as it may seem. A statement such as "John Dean did not know that Nixon knew that Dean knew that Nixon knew that McCord had burgled O'Brien's office in the Watergate Apartments" is inherently difficult to reason about (Clark and Marshall, 1981; Kinderman et al., 1998). At the same time, other types of higher-order reasoning seem unproblematic. If two players sit across the table from each other and have eye contact with each other, then clearly each of them believes that they have eye contact, believes that the other believes that, believes that the other believes that they believe that, and so on. That is, it is common belief between the players that they have eye contact (Lewis, 1969; Chwe, 2001).[1]

Existing models do not take into account that some higher-order events are easier to reason about than others. On the one hand, the standard game-theoretic framework models players as if they have higher-order beliefs about every possible event, at all orders, i.e., as if players have an infinite depth of reasoning. On the other hand, in models in which players can have a finite depth of reasoning, such as cognitive-hierarchy models or models of level-k reasoning,[2] it is assumed that a player with a finite depth of reasoning cannot reason about *any* event at higher orders. Because beliefs at arbitrarily high order can have a significant impact on economic outcomes,[3] it is important to carefully model which higher-order events players can

---

[1]We follow the recent literature in game theory in using the terms "belief" and "common belief" rather than "knowledge" and "common knowledge." The formal distinction is that knowledge is considered to be always true, while (probability-one) belief may be true or false.

[2]See, e.g., Nagel (1995), Stahl and Wilson (1995) Ho et al. (1998), Costa-Gomes et al. (2001), Crawford and Iriberri (2007), Strzalecki (2009), and Heifetz and Kets (2011). See Crawford et al. (2012) for a survey.

[3]An action that is optimal for a player given her kth-order belief, for example, may no longer be optimal

hold beliefs about. This paper provides a framework that does just that.

We propose a class of type spaces, called *extended type spaces*, starting from the idea that a player can have a coarse perception of the state of the world. A player who has a coarse perception does not distinguish among states of the world that differ only in certain details, such as the beliefs of other players at very high orders.

More specifically, each type in an extended type space is associated with a belief (probability measure) over the states of nature and the types of other players, as in the type spaces of Harsanyi (1967–1968). Unlike in Harsanyi type spaces, however, the beliefs of different types of a given player can be defined on different σ-algebras. As a type's belief assigns a probability only to those subsets of her opponents' types that are in the type's σ-algebra, a type with a coarse σ-algebra has a coarse perception of the other players' types. And because types generate higher-order beliefs, the coarseness of a type's σ-algebra determines what features of the other players' higher-order beliefs the type can reason about.

Coarse perceptions thus model small worlds, as introduced by Savage (1954) in the context of one-person decision situations. A state in a small world describes the possible uncertainties a decision-maker faces in less detail than a state in a larger world, by neglecting certain distinctions between states. This means that "a state of the smaller world corresponds not to one state of the larger, but to a *set* of states" (Savage, 1954, p. 9, emphasis added). In the present framework, a player may ignore the distinction between types for the other player that differ only in the beliefs they generate at high order, by lumping together these types into one set in her σ-algebra.

Because a Harsanyi type space is simply an extended type space in which each type has an infinite depth of reasoning, extended type spaces generalize the Harsanyi framework. Extended type spaces can also be seen as a generalization of cognitive-hierarchy and level-k models: Kets (2012) constructs an extended type space such that a player who has a finite depth k cannot reason about any nontrivial event of order greater than k, as is also true in cognitive-hierarchy and level-k models.

The framework helps shed light on the question how players can attain common belief in an event F even if they are bounded in their reasoning. Intuitively, certain lower-order events, such as a public announcement that F is the case, immediately generate all the relevant higher-order beliefs, just like the event that players have eye contact leads them to have common belief in that event. This is true even if players cannot reason about the event that F is kth-order mutual belief for some fixed k *in isolation*, as we show. This means that we no

---

given her $(k+1)$th-order belief, for any finite k (Rubinstein, 1989; Carlsson and van Damme, 1993). Also, beliefs at arbitrarily high order may determine whether players with a common prior can have different posteriors (Aumann, 1976; Geanakoplos and Polemarchakis, 1982).

longer have the usual equivalence between common belief as a conjunction of mutual belief at all orders and common belief as an event that is induced by a public event.

The idea that 'simple' events can induce (almost) common knowledge is not new; it is central to the conceptualization of common knowledge by the philosopher David Lewis (1969) and it underlies the formalization of common knowledge and approximate common belief in Aumann (1976) and Monderer and Samet (1989), respectively. Indeed, speaking of a belief hierarchy such as the one that describes a player's higher-order beliefs about the event that she and another player have eye contact, Lewis writes: "this is a chain of implications, [it does not represent] steps in anyone's actual reasoning. Therefore, there is nothing improper about its infinite length" (p. 53). Our contribution here is to point out that this idea applies beyond the context of common belief and, more fundamentally, to formalize it in the context of players with bounded reasoning abilities, and to use it to show that the usual equivalence between different notions of common belief can break down.

The remainder of this paper is organized as follows. The next section illustrates the main results with some simple examples. Section 3 formally introduces the notion of an extended type space, and Section 4 considers the higher-order beliefs that types can have. Section 5 considers the different notions of common belief that have appeared in the literature, and studies their relations. Section 6 presents a direct characterization of the depth of reasoning of types. Section 7 discusses the related literature, and Section 8 concludes. Most proofs are relegated to the appendices.

2. Examples

2.1. Extended type spaces

We present some examples to introduce the framework and to illustrate the main results. We consider a setting in which two players, Ann (a) and Bob (b), are uncertain about the state of nature $\theta \in \Theta$. The uncertainty faced by the players, including their uncertainty about the beliefs of the other player, is represented by an *(extended) type space*. As in the type spaces of Harsanyi (1967–1968), each player $i = a, b$ is endowed with a type space T_i, and each type $t_i \in T_i$ is associated with a belief (probability measure) $\beta_i(t_i)$ about the state of nature and the other player's type. Unlike in a Harsanyi type space, the beliefs of types can be defined on different σ-algebras. Specifically, Ann's type set is endowed with a collection Π_a of σ-algebras, and the belief $\beta_b(t_b)$ for a type t_b for Bob about Ann's type is defined on a σ-algebra $\Sigma_b(t_b)$ from that collection, and similarly with the player labels interchanged. We will see that the σ-algebra on which a type's belief is defined reflects the extent to which the

type "thinks through" the beliefs of the other player. The σ-algebras have to satisfy some conditions that ensure that types induce well-defined belief hierarchies; see Section 2.3.

2.1.1. Infinite depth

We start with a simple example to illustrate how types generate higher-order beliefs. The state of nature θ can be either high (h) or low (ℓ), and each player $i = a, b$ has two types, labeled t_i^1, t_i^2. Each type t_a for Ann is endowed with the power set, that is, $\Sigma_a(t_a)$ is the collection of all subsets of Bob's type set, and likewise for the types for Bob.

$\beta_a(t_a^1)$	h	ℓ
t_b^1	$\frac{1}{2}$	0
t_b^2	$\frac{1}{2}$	0

$\beta_a(t_a^2)$	h	ℓ
t_b^1	0	0
t_b^2	0	1

$\beta_b(t_b^1)$	h	ℓ
t_a^1	$\frac{1}{3}$	0
t_a^2	$\frac{2}{3}$	0

$\beta_b(t_b^2)$	h	ℓ
t_a^1	0	$\frac{1}{4}$
t_a^2	$\frac{3}{4}$	0

Figure 1: An extended type space where every type has an infinite depth of reasoning, with the beliefs for types for Ann on the left, and those for Bob on the right; we write x for the singleton $\{x\}$.

Because Bob's type sets is finite, it suffices to specify the belief $\beta_a(t_a)$ for a type t_a for Ann on the (unique) partition of Bob's type set that generates the σ-algebra $\Sigma_a(t_a)$. In this case, specifying the belief for type t_a on the pairs (θ, t_b) for every state of nature θ and every type t_b for Bob specifies its belief on the full σ-algebra. The beliefs for the types for Bob can be described similarly; see Figure 1.

The types and their beliefs specify players' higher-order beliefs. For example, type t_a^1 for Ann believes (with probability 1) that the state of nature is h and assigns probability $\frac{1}{2}$ to the event that Bob believes that, as it assigns probability $\frac{1}{2}$ to type t_b^1, which believes that $\theta = h$. Because each type for Bob similarly has a belief about Ann's belief about θ, t_a^1 also induces a belief about Bob's beliefs about Ann's belief about θ, and so on.

In this case, the reasoning does not stop at some fixed order, and each type has an infinite depth of reasoning. Indeed, the type space in Figure 1 is just a Harsanyi type space. In general, a Harsanyi type space is an extended type space in which the beliefs of types are defined on σ-algebras that are sufficiently fine to induce beliefs at all orders. Extended type spaces thus generalize the Harsanyi framework.

2.1.2. Finite depth

What happens if types are endowed with a coarser σ-algebra? Consider the extended type space in Figure 2. The collection Π_a of σ-algebras on Ann's type set consists of the trivial σ-algebra $\mathcal{F}_a^0$ and the σ-algebra $\mathcal{F}_a^*$ that is generated by the partition $\{\{t_a^1, t_a^2\}, \{t_a^3, t_a^4\}\}$;

the collection Π_b is similarly defined. Each type t_b for Bob is endowed with the σ-algebra $\Sigma_b(t_b) = \mathcal{F}_a^*$, and likewise for the types for Ann.

$\beta_a(t_a^1)$	h	ℓ
$\{t_b^1, t_b^2\}$	1	0
$\{t_b^3, t_b^4\}$	0	0

$\beta_a(t_a^2)$	h	ℓ
$\{t_b^1, t_b^2\}$	0	0
$\{t_b^3, t_b^4\}$	1	0

$\beta_b(t_b^1)$	h	ℓ
$\{t_a^1, t_a^2\}$	1	0
$\{t_a^3, t_a^4\}$	0	0

$\beta_b(t_b^2)$	h	ℓ
$\{t_a^1, t_a^2\}$	0	0
$\{t_a^3, t_a^4\}$	1	0

$\beta_a(t_a^3)$	h	ℓ
$\{t_b^1, t_b^2\}$	0	1
$\{t_b^3, t_b^4\}$	0	0

$\beta_a(t_a^4)$	h	ℓ
$\{t_b^1, t_b^2\}$	0	0
$\{t_b^3, t_b^4\}$	0	1

$\beta_b(t_b^3)$	h	ℓ
$\{t_a^1, t_a^2\}$	0	1
$\{t_a^3, t_a^4\}$	0	0

$\beta_b(t_b^4)$	h	ℓ
$\{t_a^1, t_a^2\}$	0	0
$\{t_a^3, t_a^4\}$	0	1

Figure 2: An extended type space with coarse σ-algebras.

Each type t_a for Ann has a first-order belief, that is, it assigns a probability to each event that involves the state of nature. Type t_a^1, for example, believes that the state of nature is h. Each type also induces a second-order belief, that is, a belief about any belief Bob may have about θ. Type t_a^1, for example, assigns probability 1 to the event that Bob has type t_b^1 or t_b^2 (i.e., to $\{t_b^1, t_b^2\}$), and thus to the event that Bob believes that the state of nature is h. However, type t_a^1 does not have a well-articulated belief about Bob's beliefs about her belief about the state of nature, since t_b^1 and t_b^2 differ in their beliefs about Ann's belief about nature: type t_b^1 believes that Ann believes that $\theta = h$, and t_b^2 believes that Ann believes that $\theta = \ell$. This means that t_a^1 has a finite depth of reasoning.

By endowing different types for a player with different σ-algebras, we can model situations in which players are uncertain about the depth of reasoning of their opponent; Example 1 in Section 6 provides an instance of this. Before discussing the depth of reasoning of types in more detail, we turn to the question what higher-order beliefs players can have if they are bounded in their reasoning.

2.2. Higher-order beliefs

Some forms of higher-order reasoning seem unproblematic, even if players have a finite depth of reasoning. Can we understand that in the context of this framework? And can that help us understand how common belief can be attained?

2.2.1. Eye contact

Consider the type space in Figure 3. The set of states of nature is $\Theta = \{eh, e\ell, n\}$. If the state of nature θ is eh or $e\ell$, then Ann and Bob have eye contact with each other. If the state is eh, then the value of a particular good is high; if the state is $e\ell$, then the value is low. If

$\theta = n$, then Ann and Bob do not have eye contact (and the value of the good is immaterial). The collection Π_i of σ-algebras on the type set for player $i = a, b$ contains the trivial σ-algebra $\mathcal{F}_i^0$ and the σ-algebra $\mathcal{F}_i$ that is generated by the partition $\{\{t_i^1, t_i^2\}, \{t_i^3, t_i^4\}, \{t_i^5\}\}$; every type t_a for Ann is endowed with the σ-algebra $\Sigma_a(t_a) = \mathcal{F}_b$, and likewise for the types for Bob.

$\beta_a(t_a^1)$	eh	el	n
$\{t_b^1, t_b^2\}$	1	0	0
$\{t_b^3, t_b^4\}$	0	0	0
$\{t_b^5\}$	0	0	0

$\beta_a(t_a^2)$	eh	el	n
$\{t_b^1, t_b^2\}$	0	0	0
$\{t_b^3, t_b^4\}$	1	0	0
$\{t_b^5\}$	0	0	0

$\beta_a(t_a^3)$	eh	el	n
$\{t_b^1, t_b^2\}$	0	1	0
$\{t_b^3, t_b^4\}$	0	0	0
$\{t_b^5\}$	0	0	0

$\beta_a(t_a^4)$	eh	el	n
$\{t_b^1, t_b^2\}$	0	0	0
$\{t_b^3, t_b^4\}$	0	1	0
$\{t_b^5\}$	0	0	0

Figure 3: The beliefs for Ann's types $t_a^1, t_a^2, t_a^3, t_a^4$. Type t_a^5 assigns probability 1 to $(\theta, t_b) = (n, t_b^5)$. The beliefs for the types for Bob can be obtained by exchanging the indices a and b.

The types for Ann have beliefs about Bob's beliefs about the value of the good, but not about Bob's beliefs about Ann's belief about the value. Type t_a^1, for example, believes that the value of the good is high (because t_a^1 assigns probability 1 to $\theta = eh$) and believes that Bob believes that (because t_a^1 assigns probability 1 to the event that Bob has type t_b^1 or t_b^2, both of which believe that $\theta = eh$). The type, however, cannot reason about the event that Bob believes that Ann believes that the value is high, because the type's σ-algebra does not separate the types that differ in their beliefs about Ann's beliefs about the value of the good, such as types t_b^1 and t_b^2. Thus, t_a^1 does not have a well-articulated third-order belief about certain events, and the same, in fact, holds for the other types.

Players can nevertheless have nontrivial beliefs at higher orders. For example, type t_a^1 for Ann believes that there is eye contact (i.e., t_a^1 assigns probability 1 to $\theta \in \{eh, el\}$), believes that Bob believes that (t_a^1 assigns probability 1 to types for Bob that assign probability 1 to $\theta \in \{eh, el\}$), and so on, and similarly for type t_b^1 for Bob. The key is that the high-order events that both players believe that there is eye contact, believe that the other believes that, and so on, are equivalent to the low-order event $E_e = \Theta \times \{t_a^1, t_a^2, t_a^3, t_a^4\} \times \{t_b^1, t_b^2, t_b^3, t_b^4\}$ that the players believe that they have eye contact. Because players can reason about this low-order event, they can form beliefs about the high-order events as well.

By contrast, events concerning players' higher-order beliefs about the value of the good cannot be expressed in terms of an event of sufficiently low order. Intuitively, that Ann and

Bob believe that the value is high does not imply that they believe that the other believes that. More formally, the higher-order event $E_{hh} = \Theta \times \{t_a^1\} \times \{t_b^1\}$ (both players believe that the value is high and that the other believes it is high) is strictly contained in the lower-order event $E_h = \Theta \times \{t_a^1, t_a^2\} \times \{t_b^1, t_b^2\}$ (both players believe that the value is high), and no type can reason about the higher-order event.

This holds generally: a player that has a finite depth of reasoning can form beliefs about a higher-order event if and only if the event corresponds to an event of sufficiently low order (Proposition 4.4).

2.2.2. Common belief

The event E_e that Ann and Bob believe they have eye contact is a so-called *public event*: whenever it occurs, both players believe it. This event induces *common belief* in the event that they have eye contact: both players believe they have eye contact, believe that the other believes that, and so on. Indeed, in Harsanyi type spaces, public events generate common belief in the events that they entail, and in fact generate *kth-order mutual belief* in such events, for any k. That is, any sentence of the form "both players believe that... (k times) that the event is true" holds. What is the relationship between these notions when players have a finite depth of reasoning?

Consider the type space in Figure 4. The set of states of nature is $\Theta = \{h, m, \ell\}$, and the collection Π_i of σ-algebras on the type set for player $i = a, b$ contains the trivial σ-algebra and the σ-algebra generated by the partition $\{\{t_i^1, t_i^2\}, \{t_i^3, t_i^4\}, \{t_i^5, t_i^6\}, \{t_i^7, t_i^8\}\}$.

The event $E := \Theta \times \{t_a^1, t_a^2, t_a^3, t_a^4\} \times \{t_b^1, t_b^2, t_b^3, t_b^4\}$ is a public event: whenever the players have a type that is consistent with E (e.g., $t_a = t_a^1$ and $t_b = t_b^2$), both of them assign probability 1 to E. Now consider players' higher-order beliefs in the event F that the state of nature θ is h or ℓ, that is, $F := \{h, m\} \times T_a \times T_b$. Then, whenever the players have a type that is consistent with E, both of them believe F; and believe that both believe F; and believe that both believe that both believe F; and... That is, the public event E—perhaps E is a public announcement that F is the case—can induce common belief in F.

However, there is no type that believes the sentence $s_2 =$ "both players believe that both players believe that F is the case (and players may or may not believe that both players believe that both players believe that F holds)." The problem is that no player can reason about the event $E_{s_2} = \Theta \times \{t_a^1, t_a^2, t_a^3, t_a^4, t_a^5\} \times \{t_b^1, t_b^2, t_b^3, t_b^4, t_b^5\}$ that corresponds to the sentence s_2: it does not correspond to a suitable low-order event. This implies that F is not third-order mutual belief.

Nonetheless, there are types that believe the sentence $s_2' =$ "both players believe that both players believe that F *and* E are the case": the sentence s_2' corresponds to the event E that

8

$\beta_a(t_a^1)$	h	m	ℓ
$\{t_b^1,t_b^2\}$	1	0	0
$\{t_b^3,t_b^4\}$	0	0	0
$\{t_b^5,t_b^6\}$	0	0	0
$\{t_b^7,t_b^8\}$	0	0	0

$\beta_a(t_a^2)$	h	m	ℓ
$\{t_b^1,t_b^2\}$	0	0	0
$\{t_b^3,t_b^4\}$	1	0	0
$\{t_b^5,t_b^6\}$	0	0	0
$\{t_b^7,t_b^8\}$	0	0	0

$\beta_a(t_a^3)$	h	m	ℓ
$\{t_b^1,t_b^2\}$	0	1	0
$\{t_b^3,t_b^4\}$	0	0	0
$\{t_b^5,t_b^6\}$	0	0	0
$\{t_b^7,t_b^8\}$	0	0	0

$\beta_a(t_a^4)$	h	m	ℓ
$\{t_b^1,t_b^2\}$	0	0	0
$\{t_b^3,t_b^4\}$	0	1	0
$\{t_b^5,t_b^6\}$	0	0	0
$\{t_b^7,t_b^8\}$	0	0	0

$\beta_a(t_a^5)$	h	m	ℓ
$\{t_b^1,t_b^2\}$	0	0	0
$\{t_b^3,t_b^4\}$	0	0	0
$\{t_b^5,t_b^6\}$	$\frac{1}{2}$	$\frac{1}{2}$	0
$\{t_b^7,t_b^8\}$	0	0	0

$\beta_a(t_a^6)$	h	m	ℓ
$\{t_b^1,t_b^2\}$	0	0	0
$\{t_b^3,t_b^4\}$	0	0	0
$\{t_b^5,t_b^6\}$	0	0	0
$\{t_b^7,t_b^8\}$	$\frac{1}{2}$	$\frac{1}{2}$	0

$\beta_a(t_a^7)$	h	m	ℓ
$\{t_b^1,t_b^2\}$	0	0	0
$\{t_b^3,t_b^4\}$	0	0	0
$\{t_b^5,t_b^6\}$	0	0	$\frac{1}{2}$
$\{t_b^7,t_b^8\}$	0	0	$\frac{1}{2}$

$\beta_a(t_a^8)$	h	m	ℓ
$\{t_b^1,t_b^2\}$	0	0	0
$\{t_b^3,t_b^4\}$	0	0	0
$\{t_b^5,t_b^6\}$	0	0	0
$\{t_b^7,t_b^8\}$	0	0	1

Figure 4: The beliefs for the types for Ann in an extended type space in which there is common belief that θ is h or m, even if there is no third-order mutual belief in that event.

there is a public announcement of F, and that event is of sufficiently low order. Intuitively, players cannot reason about the other player's higher-order belief about θ *in isolation*, while they are able to reason about such beliefs *in the context* of the public event. This makes sense: when there is no public announcement of F, a player, say Ann, may still believe that F is the case (as do the types t_a^5 and t_a^6), and even believe that the other player believes F (as does t_a^5), but it is not transparent to her whether or not the other player believes that she believes F, and considering the different possibilities is too hard for her. If there is a public announcement, on the other hand, then the higher-order beliefs are fixed by the low-order public event, and are completely transparent to her.

Thus, when players have a finite depth of reasoning, common belief in an event F does not necessarily imply that there is mutual belief in F at each order (Proposition 5.3). In other words, it is not the case that attaining common belief is at least as hard as attaining kth-order mutual belief for all k. This is not the case in the Harsanyi context: in a Harsanyi type space, an event is common belief if and only if it is mutual belief at every order.

2.3. Depth of reasoning

We now consider a type's depth of reasoning in more detail. A type for Ann has a *depth (of reasoning)* of at least k if the type induces a well-articulated kth-order belief, i.e., it can assign a probability to all events involving Bob's $(k-1)$th-order beliefs. How can we ensure that each type has a well-defined depth of reasoning? And how can we characterize a type's depth?

As we have seen, a type with a coarse σ-algebra can have a finite depth of reasoning. But, because the depth of reasoning of a type is inherently a property of the belief hierarchy that it induces, it is not clear a priori how the type's σ-algebra determines its depth of reasoning.

Some intuition can be gleaned from the type space in Figure 2. Consider an arbitrary type t_b for Bob, say t_b^1, with σ-algebra $\Sigma_b(t_b^1) = \mathcal{F}_a^*$. As we have seen, t_b^1 can form a second-order belief, owing to the fact that its σ-algebra $\mathcal{F}_a^*$ lumps together the types for Ann that put probability 1 on h (viz., types t_a^1 and t_a^2) and separates them from those types that put probability 1 on ℓ (viz., types t_a^3 and t_a^4); and more generally, contains all the subsets

$$\{t_a : \ \beta_a(t_a) \text{ assigns probability at least } p \text{ to } E\}, \tag{2.1}$$

for any probability p and any event E of the form $E = E' \times T_b$, where $E' \subseteq \{h, \ell\}$, or, equivalently, any event E in the product σ-algebra $\mathcal{F}_{\{h,\ell\}} \times \{T_b, \emptyset\}$, where $\mathcal{F}_{\{h,\ell\}}$ is the usual σ-algebra on the set of states of nature, and $\{T_b, \emptyset\}$ is the trivial σ-algebra on Bob's type set. Hence, the σ-algebra $\mathcal{F}_a^*$ separates the types for Ann according to their first-order belief, and the types for Bob have a second-order belief.

This, of course, holds generally: a type for Bob has a kth-order belief if its σ-algebra separates the types for Ann according to their $(k-1)$th-order beliefs, and similarly for Ann's types. Thus, the challenge is to identify conditions that ensure that a type's σ-algebra separates the types for the other player according to their beliefs up to some order, and to classify the σ-algebras according to the order up to which they separate the types.

In principle, we can check whether a σ-algebra separates the types according to their higher-order beliefs by writing out the belief hierarchies that the types induce. But the belief hierarchy induced by a type depends on the σ-algebras on the type sets, and may be ill-defined if the σ-algebras do not separate the types according to their higher-order beliefs. We thus want to find a condition *on the type space* that guarantees that the σ-algebras separate the types according to their higher-order beliefs.

We can again gain some intuition from the type space in Figure 2. As we noted, a type for Ann generates a third-order belief if its σ-algebra separates the types for Bob according to their second-order belief. How can we identify the types for Bob that have the same second-order belief? Recall that the σ-algebra $\mathcal{F}_a^*$ lumps together exactly those types for Ann that have the same first-order belief. Therefore, the types for Bob that have the same second-order belief are precisely those types that have the same belief about nature and whose beliefs about

Ann's type coincide on the σ-algebra $\mathcal{F}_a^*$. That is, a σ-algebra separates the types for Bob according to their second-order belief if it contains the subsets of the form

$$\big\{ t_b : \ \beta_b(t_b) \text{ assigns probability at least } p \text{ to } E \big\},$$

for any probability p and event E that belongs to the product σ-algebra $\mathcal{F}_{\{h,\ell\}} \times \mathcal{F}_a^*$. If a σ-algebra $\Sigma_a(t_a)$ associated with a type t_a for Bob contains these sets, then we say that $\Sigma_a(t_a)$ *dominates* the σ-algebra $\mathcal{F}_a^*$.

This can of course be generalized: a σ-algebra $\widetilde{\mathcal{F}}_b$ separates the types for Bob according to their kth-order beliefs if $\widetilde{\mathcal{F}}_b$ dominates a σ-algebra $\widetilde{\mathcal{F}}_a$ on Ann's type set that separates her types according to their $(k-1)$th-order belief, which is the case if $\widetilde{\mathcal{F}}_a$ dominates a σ-algebra $\widetilde{\mathcal{F}}_b'$ on Bob's type set that separates the types according to their $(k-2)$th-order beliefs, and so on.

2.3.3. No extraneous sets

However, even if a type's σ-algebra dominates the appropriate σ-algebras, its depth of reasoning may not be well-defined. Consider the structure in Figure 5. The collection Π_a of σ-algebras on Ann's type set contains the trivial σ-algebra and the σ-algebra generated by the partition $\{\{t_a^1, t_a^2\}, \{t_a^3\}, \{t_a^4\}\}$; the collection Π_b of σ-algebras on Bob's type set is defined similarly.

$\beta_a(t_a^1)$	h	ℓ
$\{t_b^1, t_b^2\}$	$\frac{1}{2}$	0
$\{t_b^3\}$	$\frac{1}{2}$	0
$\{t_b^4\}$	0	0

$\beta_a(t_a^2)$	h	ℓ
$\{t_b^1, t_b^2\}$	0	0
$\{t_b^3\}$	0	0
$\{t_b^4\}$	1	0

$\beta_b(t_b^1)$	h	ℓ
$\{t_a^1, t_a^2\}$	$\frac{1}{2}$	0
$\{t_a^3\}$	$\frac{1}{2}$	0
$\{t_a^4\}$	0	0

$\beta_b(t_b^2)$	h	ℓ
$\{t_a^1, t_a^2\}$	0	0
$\{t_a^3\}$	0	0
$\{t_a^4\}$	1	0

$\beta_a(t_a^3)$	h	ℓ
$\{t_b^1, t_b^2\}$	0	$\frac{1}{2}$
$\{t_b^3\}$	0	$\frac{1}{2}$
$\{t_b^4\}$	0	0

$\beta_a(t_a^4)$	h	ℓ
$\{t_b^1, t_b^2\}$	0	0
$\{t_b^3\}$	0	0
$\{t_b^4\}$	0	1

$\beta_b(t_b^3)$	h	ℓ
$\{t_a^1, t_a^2\}$	0	$\frac{1}{2}$
$\{t_a^3\}$	0	$\frac{1}{2}$
$\{t_a^4\}$	0	0

$\beta_b(t_b^4)$	h	ℓ
$\{t_a^1, t_a^2\}$	0	0
$\{t_a^3\}$	0	0
$\{t_a^4\}$	0	1

Figure 5: The σ-algebra associated with the types for Ann is not the coarsest σ-algebra that dominates the trivial σ-algebra.

The σ-algebra $\Sigma_b(t_b^1)$ associated with type t_b^1 for Bob dominates the trivial σ-algebra on Bob's type set. As we have seen in (2.1), this means that the type has a belief about Ann's belief about nature. In addition, the type can assign a probability to some events involving Ann's beliefs about nature and about Bob's beliefs about nature, but not to others. For

11

example, t_b^1 can form a belief about the event that Ann assigns equal probability to the event $E_{\ell h}$ that $\theta = \ell$ and that Bob believes that $\theta = h$, and to the event $E_{\ell\ell}$ that $\theta = \ell$ and that Bob believes that $\theta = \ell$. The reason is that t_b^1 can assign a probability to the subset of types (the singleton $\{t_a^3\}$) that have this second-order belief. On the other hand, t_b^1 cannot form a belief about another event that requires the same level of sophistication: it cannot reason about the event that Ann assigns equal probability to the event E_{hh} that $\theta = h$ and that Bob believes that $\theta = h$ and to the event $E_{h\ell}$ that $\theta = h$ and that Bob believes that $\theta = \ell$, because it can only assign a probability to the set $\{t_a^1, t_a^2\}$, not to the individual types; and t_a^1 and t_a^2 have different beliefs about E_{hh} and $E_{h\ell}$.

Unlike in the type spaces in Figures 3 and 4, the event $\{t_a^3\}$ that Ann assigns equal probability to the events $E_{\ell h}$ and $E_{\ell\ell}$ is not generated by some lower-order event, so it is not the case that Bob can reason about this event because it is entailed by a suitable lower-order event. Thus, the "depth" of t_b^1 depends on the event it reasons about, and if we allow that, there is nothing we can rule out when it comes to a player's reasoning about others.

The problem is that the σ-algebra $\Sigma_b(t_b^1)$ contains a set, viz. $\{t_a^3\}$, that is of higher order than the other events that the type can reason about. The coarsest σ-algebra that dominates the trivial σ-algebra does not have this problem: it contains precisely the subsets of types for Ann that differ in their first-order beliefs, but no 'extraneous' sets like $\{t_a^3\}$.

So, a type has a well-defined depth of reasoning if its σ-algebra is the coarsest σ-algebra that dominates the coarsest σ-algebra. . . that dominates the trivial σ-algebra. This is the case if every (nontrivial) σ-algebra in the collection Π_b of σ-algebras on Bob's type set is the coarsest σ-algebra that dominates some σ-algebra in Π_a (and likewise with the labels a and b interchanged).

2.3.4. More than two players

For the case of two players, this condition guarantees that each type generates a belief hierarchy of a well-defined depth. When there are three or more players, the situation is a bit more involved. Suppose there are three players, Ann, Bob, and Carol. Each player $i = a, b, c$ has four types, labeled as $t_i^1, t_i^2, t_i^3, t_i^4$. The σ-algebras in Π_i are the trivial σ-algebra $\mathcal{F}_i^0$ and the σ-algebra $\mathcal{F}_i^1$ generated by the partition $\{\{t_i^1, t_i^2\}, \{t_i^3, t_i^4\}\}$. We now endow each type for Ann with a product σ-algebra, describing the beliefs it can have about the types for Bob and Carol, and similarly for the types for Bob and Carol. For each player i, the types t_i^1 and t_i^2 believe that the state of nature is h; the other types believe that the state of nature is ℓ. Suppose that type t_a^1 has the product σ-algebra $\Sigma_a(t_a^1) = \mathcal{F}_b^1 \times \mathcal{F}_a^0$, and that it assigns probability 1 to the event $\{h\} \times \{t_b^1, t_b^2\} \times \{t_c^1, t_c^2, t_c^3, t_c^4\}$. Then, the type believes that Bob believes that the state of nature is h (as both t_b^1 and t_b^2 assign probability 1 to that event), but it does not have a

well-articulated belief about Carol's belief about nature. That is, Ann can reason about Bob's first-order beliefs, but not about Carol's, meaning that her "depth" of reasoning depends on the identity of the player she reasons about.

A type t_a for Ann has the same ability to reason about Bob's beliefs as about Carol's if there is some k such that its σ-algebra $\Sigma_a(t_a)$ separates the types for Bob and Carol according to their kth-order belief. So, to bring the extent to which Ann can reason about Bob's higher-order beliefs in line with the degree to which she can reason about Carol's belief, we have to impose a condition not on the σ-algebras on individual type sets, as we did earlier, but on the combinations of σ-algebras, that is, on the product σ-algebras. We can then extend our earlier analysis in a straightforward way.

More specifically, we organize the σ-algebras on players' type sets into so-called *profiles* $\boldsymbol{F} := \mathcal{F}_a \times \mathcal{F}_b \times \mathcal{F}_c$. The set of profiles, denoted Π, then represents the possible combinations of σ-algebras that types can have. We can extend the notion of dominance to profiles, and the natural analogue of our condition for the two-player case is the following:[4]

> *For every (nontrivial) profile $\boldsymbol{F} \in \Pi$, there is a profile $\boldsymbol{F}' = \mathcal{F}'_a \times \mathcal{F}'_b \times \mathcal{F}'_c \in \Pi$ such that $\mathcal{F}_a$ is the coarsest σ-algebra that contains all the sets of the form*
>
> $$\left\{ t_a : \ \beta_a(t_a) \ \text{assigns probability at least } p \text{ to } E \right\},$$
>
> *where p is a probability and E belongs to $\mathcal{F}_{\{h,\ell\}} \times \mathcal{F}'_b \times \mathcal{F}'_c$, and similarly for $\mathcal{F}_b$ and $\mathcal{F}_c$.*

This condition is sufficient to ensure that each type generates a belief hierarchy of a well-defined depth (Theorem 4.2). In fact, it is stronger than necessary; in the formal treatment, we therefore work with a somewhat weaker condition (Assumption 1) which turns out to be a weakening of the standard condition on Harsanyi type spaces that the belief maps be measurable.

2.3.5. *Characterizing a type's depth*

We can exploit the recursive nature of our condition to characterize a type's depth of reasoning. Every type has a σ-algebra that is consistent with some profile, and that profile dominates a profile which in turn dominates another profile, and so on. Thus, we can classify

---

[4]This condition, which puts restrictions on the possible profiles of σ-algebras, is somewhat stronger for the case of two players than the condition on individual type sets we identified in Section 2.3.3. However, for every extended type space for two players that satisfies the weaker condition, there exists an extended type space that satisfies the stronger variant and in which types generate the same belief hierarchies. In that sense, there is no loss of generality in adopting the stronger version also for two-player games.

the profiles according to the profiles they dominate. We can start with the profiles that do not dominate any other profile; those are the profiles that do not separate the types according to their beliefs, and thus correspond to depth 1. The profiles that dominate profiles that do not dominate any profiles themselves (but no other profiles) separate the types according to their beliefs about θ, and thus correspond to depth 2, and so on. This gives us a simple recursive procedure to determine a type's depth (Proposition 6.3).[5] .

Indeed, in the type space in Figure 2, the profile $\mathcal{F}_a^0 \times \mathcal{F}_b^0$ of trivial σ-algebras does not dominate another profile, and thus correspond to depth 1. The profile $\mathcal{F}_a^* \times \mathcal{F}_b^*$ dominates $\mathcal{F}_a^0 \times \mathcal{F}_b^0$, but no other profile. The types for Ann, which are endowed with the σ-algebra $\Sigma_a(t_a) = \mathcal{F}_b^*$, therefore have depth 2, and similarly for the types for Bob. There are no types that have depth 1; the trivial σ-algebras $\mathcal{F}_a^0, \mathcal{F}_b^0$ are used only to the specify the beliefs of types that have a higher depth.

We can thus determine a type's depth without explicitly writing out its belief hierarchy. All together, extended type spaces provide an implicit description of players' finite and infinite hierarchies of beliefs, by specifying types, beliefs about types, and a collection of σ-algebras, just like the Harsanyi type spaces implicitly model players' infinite belief hierarchies, by specifying types and beliefs about types.

3. Extended type spaces

We now begin the formal treatment. Section 3.1 defines the class of extended type spaces, and Section 3.2 demonstrates that each Harsanyi type space can be seen as an extended type space.

3.1. Definition

There is a set of players N; to avoid trivialities, we assume that the number of players is at least 2. Players are uncertain about the *state of nature*, which is drawn from a nonempty set Θ. A state of nature $\theta \in \Theta$ could, for instance, specify the payoff functions of the players, or their actions. The set Θ of states of nature is endowed with some σ-algebra $\mathcal{F}_\Theta$. Given a family of sets $\{Y_j\}_{j \in N}$ and a player $i \in N$, we denote by Y and Y_{-i} the product spaces $\bigtimes_{j \in N} Y_j$ and $\bigtimes_{j \neq i} Y_j$, respectively, and we write $y \in Y$ and $y_{-i} \in Y_{-i}$ for their typical elements. Given a family of measurable spaces $(Y_j, \mathcal{F}_j)_{j \in N}$, we write $\boldsymbol{F}$ and $\boldsymbol{F}_{-i}$ for the product σ-algebras

---

[5]There are some subtleties, though. First, in some type spaces, this procedure only gives a lower bound on a type's depth. Second, profiles can also be part of a cycle or infinite chain of profiles that dominate each other, so that this "bottom-up" procedure does not work for these profiles; see Section 6.

$\bigtimes_{j \in N} \mathcal{F}_j$ and $\bigtimes_{j \neq i} \mathcal{F}_j$ on Y and Y_{-i}, respectively.

A (Θ-based) *(extended) type space* is a tuple

$$\mathcal{T} = ((T_i, \Sigma_i, \beta_i)_{i \in N}, \Pi),$$

that satisfies Assumption 1 below. The set T_i is a nonempty set of *types* for player $i \in N$, and Π is a set of *profiles*, where a profile is a product σ-algebra $\boldsymbol{F}$ on the set T of type profiles. Given the set Π of profiles and a player $i \in N$, we can define

$$\Pi_{-i} := \{\boldsymbol{F}_{-i} : \text{ there is } \mathcal{F}_i \text{ such that } \mathcal{F}_i \times \boldsymbol{F}_{-i} \in \Pi\}$$

to be the set of σ-algebras on T_{-i} that are induced by one of the profiles in Π. The function Σ_i maps each type $t_i \in T_i$ into a (product) σ-algebra $\Sigma_i(t_i)$ in Π_{-i}. The function β_i maps each type t_i into a probability measure $\beta_i(t_i)$ on the product σ-algebra $\mathcal{F}_\Theta \times \Sigma_i(t_i)$ on $\Theta \times T_{-i}$. The function β_i is the *belief map* for player i, and the probability measure $\beta_i(t_i)$ is the *belief* of $t_i \in T_i$ about the state of nature and the types of the other players.

To state Assumption 1, we need some more notation. Fix product σ-algebras $\boldsymbol{F} = \bigtimes_{j \in N} \mathcal{F}_j$ and $\boldsymbol{F}' = \bigtimes_{j \in N} \mathcal{F}'_j$ on the set T of type profiles, and let $i \in N$. Then, the σ-algebra $\mathcal{F}_i$ *i-dominates* the product σ-algebra $\boldsymbol{F}'$ if for each event $E \in \mathcal{F}_\Theta \times \boldsymbol{F}'_{-i}$, and $p \in [0, 1]$,

$$\{t_i \in T_i : E \in \mathcal{F}_\Theta \times \Sigma_i(t_i), \beta_i(t_i)(E) \geq p\} \in \mathcal{F}_i.$$

If $\mathcal{F}_i$ *i-dominates* $\boldsymbol{F}'$ for each player $i \in N$, then we say that $\boldsymbol{F}$ *dominates* $\boldsymbol{F}'$, and we write $\boldsymbol{F} \succ \boldsymbol{F}'$. Similarly, if for each player i, $\mathcal{F}_i$ is the coarsest σ-algebra that *i-dominates* $\boldsymbol{F}'$, then we write $\boldsymbol{F} \succ^* \boldsymbol{F}'$.[6] If the product σ-algebra $\boldsymbol{F}$ dominates itself, that is, $\boldsymbol{F} \succ \boldsymbol{F}$, then we say that $\boldsymbol{F}$ is *self-dominating*.

We can now state the assumption:

Assumption 1. For every profile $\boldsymbol{F} = \bigtimes_{j \in N} \mathcal{F}_j \in \Pi$ such that $\mathcal{F}_i \neq \{T_i, \emptyset\}$ for some $i \in N$, one of the following holds:

(a) $\boldsymbol{F}$ is self-dominating; or

(b) there is a profile $\boldsymbol{F}' \in \Pi$ such that for each player $j \in N$, the σ-algebra $\mathcal{F}_j$ is the coarsest σ-algebra that *j-dominates* $\boldsymbol{F}'$, that is, $\boldsymbol{F} \succ^* \boldsymbol{F}'$.

Remark. Assumption 1 is weaker than the condition discussed in Section 2.3.4: part (a) of Assumption 1 says that if a profile $\boldsymbol{F}$ is self-dominating, then $\mathcal{F}_j$ need not be the *coarsest*

---

[6]The coarsest σ-algebra that *i-dominates* $\boldsymbol{F}'$ is well-defined: it is the σ-algebra that is the intersection of all σ-algebras on T_i (such as the power set) that *i-dominate* $\boldsymbol{F}'$.

σ-algebra that j-dominates some profile $\boldsymbol{F}'$. As we will see, types whose σ-algebra is derived from a self-dominating profile have an infinite depth of reasoning (Lemma A.3(e)), so that 'extraneous' sets in a type's σ-algebra (cf. Section 2.3.3) do not affect what higher-order events a type can reason about.[7][8] $\triangleleft$

3.2. Harsanyi type spaces

As we show now, extended type spaces generalize Harsanyi type spaces, in the sense that every Harsanyi type space can be viewed as an extended type space. A (Θ-based) *Harsanyi type space* is a tuple $(T_i^{\mathcal{H}}, \mathcal{F}_i^{\mathcal{H}}, \beta_i^{\mathcal{H}})_{i \in N}$, where for each player i, $T_i^{\mathcal{H}}$ is a nonempty set of types, and $\mathcal{F}_i^{\mathcal{H}}$ is a σ-algebra on T_i. The function $\beta_i^{\mathcal{H}}$ maps each type $t_i \in T_i^{\mathcal{H}}$ into a probability measure $\beta_i^{\mathcal{H}}(t_i)$ on the product σ-algebra $\mathcal{F}_\Theta \times \boldsymbol{F}_{-i}^{\mathcal{H}}$. The set of probability measures on $\mathcal{F}_\Theta \times \boldsymbol{F}_{-i}^{\mathcal{H}}$ is endowed with the σ-algebra that is generated by the sets

$$\left\{ \mu : \mu(E) \geq p \right\} : \qquad\qquad E \in \mathcal{F}_\Theta \times \boldsymbol{F}_{-i}^{\mathcal{H}}, p \in [0,1].$$

That is, the set of probability measures on $\mathcal{F}_\Theta \times \boldsymbol{F}_{-i}^{\mathcal{H}}$ is endowed with the coarsest σ-algebra that contains the sets of probability measures that assign probability at least p to the event E, for any $E \in \mathcal{F}_\Theta \times \boldsymbol{F}_{-i}^{\mathcal{H}}$ and $p \in [0,1]$. The belief maps $\beta_i^{\mathcal{H}}$ are assumed to be measurable. (This specifications covers most of the alternative definitions in the literature, such as those that require that type sets be separable metrizable or Polish, and assume that the belief maps are Borel measurable or continuous.)

Given a Harsanyi type space $(T_i^{\mathcal{H}}, \mathcal{F}_i^{\mathcal{H}}, \beta_i^{\mathcal{H}})_{i \in N}$, we can define the tuple $((T_i^*, \Sigma_i^*, \beta_i^*)_{i \in N}, \Pi^*)$, where for each player $i \in N$, $T_i^* := T_i^{\mathcal{H}}$, the collection Π^* of profiles is the singleton $\{\boldsymbol{F}^{\mathcal{H}}\}$, $\Sigma_i^*(t_i) := \boldsymbol{F}_{-i}^{\mathcal{H}}$ for each $t_i \in T_i^*$, and $\beta_i^* := \beta_i^{\mathcal{H}}$. This structure is an extended type space, as the next result shows.

Proposition 3.1. *The tuple $((T_i^*, \Sigma_i^*, \beta_i^*)_{i \in N}, \Pi^*)$ satisfies Assumption 1. Therefore, any Harsanyi type space can be seen as an extended type space.*

In fact, as can be seen from the proof of Proposition 3.1, the requirement that the belief maps be measurable is equivalent to the condition that the product σ-algebra $\boldsymbol{F}^{\mathcal{H}}$ is self-dominating. Thus, Assumption 1 relaxes the condition that the belief maps be measurable, by allowing for profiles that do not dominate themselves (as in the type space in Figure 2, for example).

---

[7]However, such sets can affect strategic behavior, at least in a Harsanyi type space, depending on the solution concept (Ely and Peski, 2006; Dekel et al., 2007).

[8]While every Harsanyi type space satisfies Assumption 1 (Proposition 3.1), not every Harsanyi type space satisfies the stronger condition (Friedenberg and Meier, 2012).

4. Higher-order beliefs

We are interested in the question of what higher-order beliefs types can have. Section 4.1 demonstrates how each type can be mapped into a belief hierarchy, where a belief hierarchy describes a type's beliefs at all orders. Section 4.2 shows that each type has a well-defined depth of reasoning, and characterizes the events that a type with a finite depth can reason about.

4.1. Belief hierarchies

We first construct the space of belief hierarchies. In contrast to constructions of the space of belief hierarchies for the Harsanyi case (e.g., Mertens and Zamir, 1985; Brandenburger and Dekel, 1993), we need to take into account the fact that types can have a finite depth of reasoning. To accommodate that, the set of kth-order beliefs for a player i is endowed with a *collection* of σ-algebras, as opposed to a single one, with different σ-algebras capturing different perceptions for player $j \neq i$ of player i's kth-order beliefs.

To construct the space of belief hierarchies, we need some more notation. Given a set X and a (nonempty) collection $\mathscr{S}$ of σ-algebras on X, let $\Delta(X, \mathscr{S})$ be the set of probability measures that are defined on one of the σ-algebras in $\mathscr{S}$. If μ is a probability measure in $\Delta(X, \mathscr{S})$, then the σ-algebra on which μ is defined is denoted by $\Sigma(\mu) \in \mathscr{S}$. The set $\Delta(X, \mathscr{S})$ of probability measures is endowed with the σ-algebra $\mathcal{F}_{\Delta(X,\mathscr{S})}$ generated by the sets

$$\big\{\mu \in \Delta(X, \mathscr{S}) : E \in \Sigma(\mu), \mu(E) \geq p\big\} : \qquad E \in \mathcal{F}, \mathcal{F} \in \mathscr{S}, p \in [0, 1].$$

This σ-algebra naturally separates beliefs (probability measures) according to the probability they assign to events; this makes it possible to talk about "beliefs about beliefs," and so on (Heifetz and Samet, 1998). Finally, given a collection $\{f_j : X_j \to Y_j\}_{j \in N}$ of functions and a player $i \in N$, we write f_{-i} and $f_{-i}(x_{-i})$ for $(f_j)_{j \neq i}$ and $(f_j(x_j))_{j \neq i}$, respectively.

We are now ready to construct the space of belief hierarchies. We construct a sequence $B_i^0, B_i^1, \ldots$ of spaces for each player i that describe the higher-order beliefs for that player. Formally, for each player $i \in N$, fix an arbitrary "seed" μ_i^0. Then define

$$B_i^0 := \{\mu_i^0\},$$

and let $\mathcal{F}_{B_i^0} := \{B_i^0, \emptyset\}$ be the trivial σ-algebra on B_i^0.

For $k > 0$, suppose that for each $i \in N$ and $n \leq k - 1$, the set B_i^n has been defined, and that $\mathcal{F}_{B_i^n}$ is a σ-algebra on B_i^n. Then, for each player i, define

$$\mathscr{S}_i^k := \big\{\mathcal{F}_\Theta \times \boldsymbol{A}_{-i}^{k-1} : \ \mathcal{A}_j^{k-1} \text{ is a sub-}\sigma \text{ algebra of } \mathcal{F}_{B_j^{k-1}}, j \neq i\big\},$$

and

$$B_i^k := B_i^{k-1} \times \Delta(\Theta \times B_{-i}^{k-1}, \mathscr{S}_i^k).$$

Let $\mathcal{F}_{B_i^k}$ be the product σ-algebra $\mathcal{F}_{B_i^{k-1}} \times \mathcal{F}_{\Delta(\Theta \times B_{-i}^{k-1}, \mathscr{S}_i^k)}$.

This defines a sequence of spaces $B_i^0, B_i^1, \ldots$ for each player i. The space B_i^k is the set of kth-order belief hierarchies $(\mu_i^0, \mu_i^1, \ldots, \mu_i^k)$, with μ_i^k a kth-order belief, i.e., a belief about the state of nature and the $(k-1)$th-order beliefs of the other players.

We can map each type into a belief hierarchy. Fix an extended type space $\mathcal{T} = ((T_i, \Sigma_i, \beta_i)_{i \in N}, \Pi)$ and a player $i \in N$. Define the function $h_i^{\mathcal{T},0}$ from T_i into B_i^0 in the obvious way: $h_i^{\mathcal{T},0}(t_i) := \mu_i^0$ for $t_i \in T_i$. For $k > 0$, define the function $h_i^{\mathcal{T},k}$ from T_i into B_i^k by:

$$h_i^{\mathcal{T},k}(t_i) = \left(h_i^{\mathcal{T},k-1}(t_i), \beta_i(t_i) \circ \left(\mathrm{Id}_\Theta, h_{-i}^{\mathcal{T},k-1} \right)^{-1} \right),$$

where $\beta_i(t_i) \circ \left(\mathrm{Id}_\Theta, h_{-i}^{\mathcal{T},k-1} \right)^{-1}$ is the probability measure induced by t_i over the set of states of nature and the $(k-1)$th-order beliefs of the other players. That is, for each $E \subseteq \Theta \times B_{-i}^{k-1}$,

$$\beta_i(t_i) \circ \left(\mathrm{Id}_\Theta, h_{-i}^{\mathcal{T},k-1} \right)^{-1}(E) = \beta_i(t_i)\left(\left\{ (\theta, t_{-i}) : (\theta, h_{-i}^{\mathcal{T},k-1}(t_{-i})) \in E \right\} \right)$$

is the probability that type t_i assigns to the event that the state of nature and the $(k-1)$th-order belief hierarchies for the other players lie in E, whenever this probability is well-defined. Thus, $h_i^{\mathcal{T},k}(t_i)$ is the kth-order belief hierarchy induced by t_i, and $\beta_i(t_i) \circ \left(\mathrm{Id}_\Theta, h_{-i}^{\mathcal{T},k-1} \right)^{-1}$ is its kth-order belief.

A (full) *belief hierarchy* of a type t_i is simply a sequence of probability measures that specifies the kth-order beliefs induced by t_i for every k. Formally, define the function $h_i^{\mathcal{T}}$ from T_i into the product space $\bigtimes_{k=1}^{\infty} \Delta(\Theta \times B_{-i}^{k-1}, \mathscr{S}_i^k)$ by

$$h_i^{\mathcal{T}}(t_i) := \left(\beta_i(t_i) \circ \left(\mathrm{Id}_\Theta, h_{-i}^{\mathcal{T},0} \right)^{-1}, \beta_i(t_i) \circ \left(\mathrm{Id}_\Theta, h_{-i}^{\mathcal{T},1} \right)^{-1}, \ldots \right).$$

Then, $h_i^{\mathcal{T}}(t_i)$ is the *belief hierarchy* generated by type t_i in type space $\mathcal{T}$. The next result says that the functions $h_i^{\mathcal{T},k}$ and h_i are well-defined.

Proposition 4.1. *Let* $\mathcal{T} = ((T_i, \Sigma_i, \beta_i)_{i \in N}, \Pi)$ *be an extended type space. Then, for each player* $i \in N$ *and type* $t_i \in T_i$, *the following hold:*

(a) *Type* t_i *induces a* kth-*order belief hierarchy for every* k, *that is,* $h_i^{\mathcal{T},k}(t_i)$ *is an element of the set* B_i^k *for every* k.

(b) *Type* t_i *induces a full belief hierarchy, that is,* $h_i^{\mathcal{T}}(t_i)$ *is an element of the set* $\bigtimes_{k=1}^{\infty} \Delta(\Theta \times B_{-i}^{k-1}, \mathscr{S}_i^k)$.

We can now define H_i to be the set of belief hierarchies generated by some type for player i in some extended type space $\mathcal{T}$, so that H_i is a subset of $\times_{k=1}^{\infty} \Delta(\Theta \times B_{-i}^{k-1}, \mathscr{S}_i^k)$. By Proposition 4.1(b), the set H_i of belief hierarchies is nonempty.

The present construction extends the construction of Heifetz and Samet (1998) for the Harsanyi case to the case where types can have a finite (as well as infinite) depth of reasoning.[9] The next section defines the depth of reasoning of types, and shows that each type has a well-defined depth.

4.2. Depth of reasoning

Intuitively, a type has an infinite depth of reasoning if it can assign a probability to every kth-order event for every k, where a kth-order event concerns the state of nature and the $(k-1)$th-order belief of the other players. A type has a finite depth of reasoning k if it induces a well-defined belief about every mth-order event for $m \leq k$, but there exist higher-order events to which it cannot assign a probability.

Formally, fix an extended type space $\mathcal{T} = ((T_i, \Sigma_i, \beta_i)_{i \in N}, \Pi)$, and for each player $i \in N$ and $k = 0, 1, \ldots$, let $B_i^{\mathcal{T},k}$ be the image of the kth-order hierarchy map $h_i^{\mathcal{T},k}$ in B_i^k, that is, $B_i^{\mathcal{T},k} := h_i^{\mathcal{T},k}(T_i)$. By Proposition 4.1(a), the set $B_i^{\mathcal{T},k}$ is a (nonempty) subset of B_i^k, and we endow $B_i^{\mathcal{T},k}$ with the relative σ-algebra $\mathcal{F}_{B_i^{\mathcal{T},k}}$ induced by the σ-algebra $\mathcal{F}_{B_i^k}$.

Then, a *kth-order event* for a player $i \in N$ is an event that involves the state of nature and the $(k-1)$th-order beliefs of the other players, that is, a kth-order event for i is an event in $\mathcal{F}_\Theta \times \times_{j \neq i} \mathcal{F}_{B_j^{\mathcal{T},k-1}}$. A type t_i *can reason at order k* if for every kth-order event $E \in \mathcal{F}_\Theta \times \times_{j \neq i} \mathcal{F}_{B_j^{\mathcal{T},k-1}}$, the probability

$$\beta_i(t_i) \circ \left(\mathrm{Id}_\Theta, h_{-i}^{\mathcal{T},k-1}\right)^{-1}(E) = \beta_i(t_i)\left(\left\{(\theta, t_{-i}) : (\theta, h_{-i}^{\mathcal{T},k-1}(t_{-i})) \in E\right\}\right)$$

is well-defined, or, equivalently, if for every $E' \in \times_{j \neq i} \mathcal{F}_{B_j^{\mathcal{T},k-1}}$, we have that

$$\left\{t_{-i} \in T_{-i} : h_{-i}^{\mathcal{T},k-1}(t_{-i}) \in E'\right\} \in \Sigma_i(t_i).$$

If a type can reason at order k, then it can reason at every order $\ell < k$, due to the recursive structure of belief hierarchies. Conversely, a type t_i cannot reason at order $k+1$ if there is an event $E' \in \times_{j \neq i} \mathcal{F}_{B_j^{\mathcal{T},k}}$ such that

$$\left\{t_{-i} \in T_{-i} : h_{-i}^{\mathcal{T},k}(t_{-i}) \in E'\right\} \notin \Sigma_i(t_i).$$

Clearly, a type that cannot reason at order $k+1$ cannot reason at any order $\ell > k+1$.

We can now define the depth of reasoning of types.

---

[9]Indeed, the space constructed by Heifetz and Samet (1998) is equivalent to the subset $H_i^{\mathcal{H}} \subsetneq H_i$ of belief hierarchies that are induced by extended type spaces that are derived from Harsanyi type spaces.

Definition. Fix a type $t_i \in T_i$ for player i. Then:

(a) type t_i has *finite depth (of reasoning)* $k = 1, 2, \ldots$ if t_i can reason at order k, but not at order $k + 1$;

(b) type t_i has an *infinite depth (of reasoning)* if t_i can reason at every order.

We write $d_i(t_i)$ for the depth of reasoning of a type t_i, where $d_i(t_i) = \infty$ if t_i has an infinite depth of reasoning, and $d_i(t_i) = k$ if t_i has finite depth k. The next result states that each type has a well-defined depth:[10]

Theorem 4.2. *Let* $\mathcal{T} = ((T_i, \Sigma_i, \beta_i)_{i \in N}, \Pi)$ *be an extended type space and fix a player* $i \in N$ *and a type* $t_i \in T_i$ *for* i*. Then, there is a unique* k *such that* $d_i(t_i) = k$*, where* $k = 1, 2, \ldots$ *or* $k = \infty$*.*

It is straightforward to show that types from Harsanyi type spaces have an infinite depth of reasoning:

Proposition 4.3. *Let* $\mathcal{T} = ((T_i, \Sigma_i, \beta_i)_{i \in N}, \Pi)$ *be an extended type space derived from a Harsanyi type space, and fix a player* $i \in N$ *and type* $t_i \in T_i$ *for* i*. Then,* t_i *has an infinite depth of reasoning.*

The next result shows that types with a finite depth can reason about a higher-order event if and only if the event correspond to an event at lower order.

Proposition 4.4. *Fix a player* i *and a type* $t_i \in T_i$*. Suppose* $d_i(t_i) = k < \infty$*. Then, for each event* $B \subseteq H_{-i}$ *that describes the belief hierarchies for the players* $j \neq i$*, we have*

$$\left\{ t_{-i} \in T_{-i} : h_{-i}^{\mathcal{T}}(t_{-i}) \in B \right\} \in \Sigma_i(t_i)$$

if and only if there is an event $B_{k-1} \in \bigtimes_{j \neq i} \mathcal{F}_{B_j^{\mathcal{T},k-1}}$ *such that*

$$\left\{ t_{-i} \in T_{-i} : h_{-i}^{\mathcal{T}}(t_{-i}) \in B \right\} = \left\{ t_{-i} \in T_{-i} : h_{-i}^{\mathcal{T},k-1}(t_{-i}) \in B_{k-1} \right\}.$$

In the next section, we use these results to study the conditions under which common belief and kth-order mutual belief can be attained. Section 6 provides a characterization of the depth of reasoning of types.

---

[10]It is easy to verify that if different types (possibly from different type spaces) generate the same belief hierarchy $(\mu^1, \mu^2, \ldots)$, then they have the same depth of reasoning.

5. Common belief

We apply the framework to understand how common belief can be attained when players have a finite depth. The next section defines a belief operator for our setting. We then consider the various accounts of common belief that have appeared in the literature, and show that common belief can obtain even if mutual belief cannot be obtained at all orders.

5.1. Belief operators

We extend the standard belief operators to our setting (cf. Brandenburger and Dekel, 1987; Monderer and Samet, 1989). Fix an extended type space $\mathcal{T} = ((T_i, \Sigma_i, \beta_i)_{i \in N}, \Pi)$. A *state of the world* $(\theta, t) \in \Theta \times T$ specifies a state of nature θ and a type t_i for each player $i \in N$. Fix an event $E \subseteq \Theta \times T$, and for player $i \in N$ and type $t_i \in T_i$, define

$$E_{t_i} := \{(\theta, t_{-i}) : (\theta, t) \in E\}.$$

Then, define

$$\mathcal{B}_i(E) := \{(\theta, t) \in \Theta \times T : E_{t_i} \in \mathcal{F}_\Theta \times \Sigma_i(t_i), \beta_i(t_i)(E_{t_i}) = 1\}$$

to be the event that player i *believes* E (with probability 1). Let

$$\mathcal{B}(E) := \bigcap_{i \in N} \mathcal{B}_i(E)$$

be the event that all players believe the event E, that is, $\mathcal{B}(E)$ is the event that E is *mutual belief*. The set of states at which E is *kth-order mutual belief* is $\mathcal{B}^k(E)$, where $\mathcal{B}^1(E) := \mathcal{B}(E)$, and $\mathcal{B}^k := \mathcal{B} \circ \mathcal{B}^{k-1}$.

The belief operator $\mathcal{B}_i$ coincides with the standard one for the Harsanyi case, and indeed it satisfies many of the usual properties, as we show in Appendix B. However, the belief operator is not monotonic, unlike the standard operator: if a player believes an event E, she need not believe an event E' that is implied by E. Intuitively, a player can only believe events that she can reason about (i.e., events that belong to her σ-algebra), and that she can reason about E does not imply that she can reason about E'.

5.2. Common belief: different accounts

We investigate the relationship between different notions of common belief when players can have a finite depth of reasoning. Game-theoretic applications often focus on the iterative account: an event $F \subseteq \Theta \times T$ is common belief if all players believe F, believe that all players

believe F, and so on. That is, F is *common belief in the iterative sense* at a state of the world $(\theta, t) \in \Theta \times T$ if

$$(\theta, t) \in \bigcap_{k=1}^{\infty} \mathcal{B}^k(F).$$

Denote the set of states of the world in which F is common belief in the iterative sense by $C^{\text{IT}}(F)$.

While this definition of common belief is close to the informal description, it seems unlikely that players check whether each of the infinitely many events $\mathcal{B}^k(F)$ obtains, even if they have an infinite depth of reasoning. An alternate approach is the fixed-point approach. The idea is that if an event F is common belief, then every player believes F and believes that F is common belief; and conversely, if every player believes F and believes that F is common belief, then F is common belief (Aumann, 1976; Lewis, 1969; Halpern and Moses, 1990). Formally, the event F is *common belief in the fixed-point sense* at a state of the world $(\theta, t) \in \Theta \times T$ if there is an event $X \subseteq \Theta \times T$ such that $(\theta, t) \in X$, and

$$X \subseteq \mathcal{B}(X \cap F), \text{ and}$$
$$X \supseteq \mathcal{B}(X \cap F).$$

Equivalently, F is common belief in the fixed-point sense at (θ, t) if (θ, t) belongs to a fixed point of the mapping f_F defined by

$$f_F(A) = \mathcal{B}(A \cap F),$$

where $A \subseteq \Theta \times T$. Denote the set of states of the world in which F is common belief in the fixed-point sense by $C^{\text{FP}}(F)$.

The fixed-point account does not refer to infinitely many statements regarding players' beliefs, but it is silent on how common belief can be attained in practice. This motivates the following definition (Monderer and Samet, 1989; Clark and Marshall, 1981). An event is *common belief due to a public event* at (θ, t) if there exists an event $E \subseteq \Theta \times T$ such that $(\theta, t) \in E$ and for each player $i \in N$, we have

$$E \subseteq \mathcal{B}_i(E), \text{ and} \tag{5.1}$$
$$E \subseteq \mathcal{B}_i(F).$$

An event E that satisfies (5.1) for each $i \in N$ is called a *public* or *self-evident* event: whenever E occurs, every player believes E. Denote the set of states of the world in which F is common belief due to a public event by $C^{\text{PUB}}(F)$.

We are particularly interested in the case of common belief about the state of nature, that is, common belief in events that do not make reference to players' higher-order beliefs. Say

that $F \subseteq \Theta \times T$ is a *primitive event* if $F \in \mathcal{F}_\Theta \times \bigtimes_{j \in N}\{T_j, \emptyset\}$ (The restriction to primitive events is not essential; what matters is that the relevant types can reason about F.)

The fixed-point notion and the notion of common belief due to a public event are known to be equivalent for Harsanyi type spaces; the next result shows that this holds even if players have a finite depth of reasoning:

Proposition 5.1. *Let F be a primitive event. Then, $C^{\mathrm{FP}}(F) = C^{\mathrm{PUB}}(F)$.*

Since the two notions are equivalent, we write C for $C^{\mathrm{FP}} = C^{\mathrm{PUB}}$ in the remainder of this section. In a Harsanyi context, the iterative account is equivalent to the other two accounts, as is well-known.

Proposition 5.2. *Suppose $\mathcal{T}$ is derived from a Harsanyi type space with the set of profiles given by $\Pi = \{\boldsymbol{F}^{\mathcal{H}}\}$. Then, for any primitive event F, we have that $C^{\mathrm{IT}}(F) = C(F)$.*

However, the iterative account need not be equivalent to the other two in general type spaces:

Proposition 5.3. *There exists a type space $\mathcal{T}$ and a primitive event F in $\mathcal{T}$ such that $C(F)$ is nonempty, but $C^{\mathrm{IT}}(F) = \emptyset$. That is, a public event can generate common belief, but common belief in the iterative sense does not obtain.*

Proof. Take $\mathcal{T}$ to be the type space in Figure 4 in Section 2.2.2. As in Section 2.2.2, let $F := \{h, m\} \times T_a \times T_b$, and $E := \Theta \times \{t_a^1, t_a^2, t_a^3, t_a^4\} \times \{t_b^1, t_b^2, t_b^3, t_b^4\}$. Then, for each player $i = a, b$, we have

$$E \subseteq \mathcal{B}_i(E);$$
$$E \subseteq \mathcal{B}_i(F).$$

Hence, $C(F) = \Theta \times \{t_a^1, t_a^2, t_a^3, t_a^4\} \times \{t_b^1, t_b^2, t_b^3, t_b^4\}$.

However, for every $i = a, b$ and type $t_i \in T_i$, we have that either

$$[\mathcal{B}^2(F)]_{t_i} = \Theta \times \{t_i^1, t_i^2, t_i^3, t_i^4, t_i^5\} \notin \mathcal{F}_\Theta \times \Sigma_i(t_i)$$

or $[\mathcal{B}^2(F)]_{t_i} = \emptyset$. Hence, the set $\mathcal{B}_i(\mathcal{B}^2(F))$ is empty for every $i = a, b$, and so is $C^{\mathrm{IT}}(F) = \bigcap_{k=1}^\infty \mathcal{B}^k(F)$. $\qquad\square$

Technically, the reason that the proof of Proposition 5.2 does not go through when players have a finite depth of reasoning is that the belief operator is not monotonic. Indeed, in the type space in the proof of Proposition 5.3, we have that

$$\mathcal{B}_i(\Theta \times \{t_a^1, t_a^2, t_a^3, t_a^4, t_a^5\} \times \{t_b^1, t_b^2, t_b^3, t_b^4, t_b^5\}) = \emptyset,$$

while
$$\mathcal{B}_i(\Theta \times \{t_a^1, t_a^2, t_a^3, t_a^4\} \times \{t_b^1, t_b^2, t_b^3, t_b^4\}) = \Theta \times \{t_a^1, t_a^2, t_a^3, t_a^4\} \times \{t_b^1, t_b^2, t_b^3, t_b^4\}.$$

In general, it can be shown that $C^{\text{IT}}(F)$ is a (possibly empty) fixed point of f_F. Hence, if common belief in an iterative sense obtains, then so does common belief in the fixed-point sense (and thus common belief due to a public event). The converse does not hold, however: Proposition 5.3 demonstrates that common belief can be attained even if common belief in the iterative sense does not obtain.[11]

6.   Characterizing a type's depth

We provide a characterization of the depth of reasoning of types in terms of the properties of the type space alone. The value of such a characterization is that we can determine a type's depth of reasoning without writing out its belief hierarchy. We first classify the profiles according to the profiles that they dominate. We then use this classification to characterize a type's depth.

Formally, fix an extended type space $\mathcal{T} = ((T_i, \Sigma_i, \beta_i)_{i \in N}, \Pi)$. Define

$$\mathcal{O}^1 := \left\{ \boldsymbol{F} \in \Pi : \text{ there is no } \boldsymbol{F}' \in \Pi \text{ such that } \boldsymbol{F} \succ \boldsymbol{F}' \right\}$$

to be the set of profiles that do not dominate any profile. For $k = 2, 3, \ldots$, let

$$\mathcal{O}^k := \Big\{ \boldsymbol{F} \in \Pi : \text{ there is } \boldsymbol{F}' \in \mathcal{O}^{k-1} \text{ such that } \boldsymbol{F} \succ \boldsymbol{F}', \text{ and}$$
$$\text{there is no } \boldsymbol{F}'' \in \Pi \setminus \bigcup_{m \le k-1} \mathcal{O}^m \text{ such that } \boldsymbol{F} \succ \boldsymbol{F}'' \Big\}.$$

In words, a profile $\boldsymbol{F}$ belongs to the collection $\mathcal{O}^k$ if it dominates some profile in $\mathcal{O}^{k-1}$, and, moreover, any profile that is dominated by $\boldsymbol{F}$ belongs to a class $\mathcal{O}^m$ for $m \le k - 1$. Finally, let

$$\mathcal{O}^\infty := \Big\{ \boldsymbol{F} \in \Pi : \text{ for each } k = 1, 2, \ldots, \text{ there is } \boldsymbol{F}' \in \Pi \setminus \bigcup_{m \le k-1} \mathcal{O}^m \text{ such that } \boldsymbol{F} \succ \boldsymbol{F}' \Big\}$$

be the collection of profiles $\boldsymbol{F} \in \Pi$ such that for any k, there is a profile that is dominated by $\boldsymbol{F}$ that does not belong to any of the classes $\mathcal{O}^m$, $m \le k - 1$, where $\mathcal{O}^0 := \emptyset$.

The classes $\mathcal{O}^\infty$, $\mathcal{O}^1, \mathcal{O}^2, \ldots$ thus classify the profiles in Π according to the profiles that they dominate. Each profile belongs to precisely one of these subsets, as the next result demonstrates.

---

[11]One could, of course, try to restore equivalence by defining a belief operator that is monotonic, such as one that is based on inner measures. This involves, however, relaxing the (σ-)additivity condition that probability measures satisfy; we therefore do not pursue this direction further.

Lemma 6.1. *The collection* $\{\mathcal{O}^\infty, \mathcal{O}^1, \mathcal{O}^2, \ldots\}$ *is a partition of* Π.

We say that a profile $\boldsymbol{F} \in \Pi$ has *order* k if $\boldsymbol{F} \in \mathcal{O}^k$, where $k = 1, 2, \ldots$ or $k = \infty$. By Lemma 6.1, the order of each profile is well-defined.

The *rank* $r_i(t_i)$ of a type $t_i \in T_i$ is the maximum order of the profiles that are consistent with its σ-algebra:

$$r_i(t_i) := \max\{k = \infty, 1, 2, \ldots : \mathcal{F}_i \times \Sigma_i(t_i) \in \mathcal{O}^k \text{ for some } \mathcal{F}_i\}.$$

The rank of a type provides a lower bound to its depth.

Proposition 6.2. *Fix a player* $i \in N$ *and a type* $t_i \in T_i$. *Then, the depth of* t_i *is at least its rank, that is,* $d_i(t_i) \geq r_i(t_i)$. *In particular, if* $r_i(t_i) = \infty$, *then* $d_i(t_i) = \infty$.

However, the bound need not be tight, as the following two examples illustrate.

Example 1. There are two players, Ann (a) and Bob (b) and the state of nature θ is either h or ℓ. The type space is given in Figure 6. The set of profiles is $\Pi = \{\mathcal{F}_a^0 \times \mathcal{F}_b^0, \mathcal{F}_a^1 \times \mathcal{F}_b^0, \mathcal{F}_a^2 \times \mathcal{F}_b^2\}$, where $\mathcal{F}_i^0$ and $\mathcal{F}_i^2$, $i = a, b$, are the trivial σ-algebra and the power set, respectively, on the type set of player i, and $\mathcal{F}_a^1$ is the σ-algebra on Ann's type set that is generated by the partition $\{\{t_a^1, t_a^2\}, \{t_a^3, t_a^4\}\}$. It can be checked that $\mathcal{F}_a^1 \times \mathcal{F}_b^0 \succ^* \mathcal{F}_a^0 \times \mathcal{F}_b^0$, and that $\mathcal{F}_a^2 \times \mathcal{F}_b^2$ is self-dominating, so that Assumption 1 is satisfied.
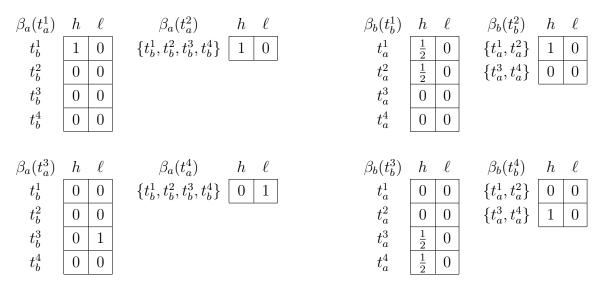
$\beta_a(t_a^1)$	h	ℓ
t_b^1	1	0
t_b^2	0	0
t_b^3	0	0
t_b^4	0	0

$\beta_a(t_a^2)$	h	ℓ
$\{t_b^1, t_b^2, t_b^3, t_b^4\}$	1	0

$\beta_b(t_b^1)$	h	ℓ
t_a^1	$\frac{1}{2}$	0
t_a^2	$\frac{1}{2}$	0
t_a^3	0	0
t_a^4	0	0

$\beta_b(t_b^2)$	h	ℓ
$\{t_a^1, t_a^2\}$	1	0
$\{t_a^3, t_a^4\}$	0	0

$\beta_a(t_a^3)$	h	ℓ
t_b^1	0	0
t_b^2	0	0
t_b^3	0	1
t_b^4	0	0

$\beta_a(t_a^4)$	h	ℓ
$\{t_b^1, t_b^2, t_b^3, t_b^4\}$	0	1

$\beta_b(t_b^3)$	h	ℓ
t_a^1	0	0
t_a^2	0	0
t_a^3	$\frac{1}{2}$	0
t_a^4	$\frac{1}{2}$	0

$\beta_b(t_b^4)$	h	ℓ
$\{t_a^1, t_a^2\}$	0	0
$\{t_a^3, t_a^4\}$	1	0

Figure 6: The type space for Example 1.

The rank of a type can be lower than its depth. Type t_b^2, for example, is endowed with the σ-algebra $\mathcal{F}_a^1$, and the order of the corresponding profile $\mathcal{F}_a^1 \times \mathcal{F}_b^0$ is $r_b(t_b^2) = 2$. However, the depth of t_b^2 is $d_b(t_b^2) = 3$. The same, of course, applies to type t_b^4. ◁

Thus, a type with a finite depth of reasoning can be assigned a rank that is strictly lower than its depth. The next example demonstrates that a type with an infinite depth of reasoning can be assigned a finite rank.

Example 2. Consider the type space in Figure 7. The collection Π of profiles consists of $\widetilde{\mathcal{F}}_a^0 \times \widetilde{\mathcal{F}}_b^0$ and $\widetilde{\mathcal{F}}_a^1 \times \widetilde{\mathcal{F}}_b^0$, where $\widetilde{\mathcal{F}}_i^0$, $i = a, b$, is the trivial σ-algebra on the type set for player i, and $\widetilde{\mathcal{F}}_a^1$ is the power set on Ann's type set. It is straightforward to verify that $\widetilde{\mathcal{F}}_a^1 \times \widetilde{\mathcal{F}}_b^0 \succ^* \widetilde{\mathcal{F}}_a^0 \times \widetilde{\mathcal{F}}_b^0$. Again, some types have a rank that is strictly lower than their depth. Types q_b^1 and q_b^2 are assigned rank 2, but have an infinite depth. ◁

$\beta_a(q_a^1)$	h	ℓ
$\{q_b^1, q_b^2\}$	1	0

$\beta_a(q_a^2)$	h	ℓ
$\{q_b^1, q_b^2\}$	0	1

$\beta_b(q_b^1)$	h	ℓ
q_a^1	1	0
q_a^2	0	0

$\beta_b(q_b^2)$	h	ℓ
q_a^1	0	0
q_a^2	1	0

Figure 7: The type space for Example 2.

In both examples, the beliefs of Ann's types at some higher order are completely determined by their beliefs at some lower order. In Example 1, Ann's beliefs about θ fix her belief about θ and Bob's belief about θ (because all types for Bob have the same belief about θ), and in Example 2, Ann's belief about θ completely determine her belief at all orders (because Ann has no beliefs beyond the first order). This implies that the pertinent σ-algebras on Ann's type set ($\mathcal{F}_a^1$ in Example 1, $\widetilde{\mathcal{F}}_a^1$ in Example 2) can form a profile with σ-algebras on Bob's type set that separate Bob's types at the relevant lower order, but not at the higher order ($\mathcal{F}_b^0$ and $\widetilde{\mathcal{F}}_b^0$, respectively), so that the "order" of the σ-algebras on Ann's type set is greater than the order of the profiles as a whole ($\mathcal{F}_a^1 \times \mathcal{F}_b^0$ and $\widetilde{\mathcal{F}}_a^1 \times \widetilde{\mathcal{F}}_b^0$, respectively), and a type for Bob with such a σ-algebra is assigned a rank that is lower than its depth.

This is ruled out if all σ-algebras $\mathcal{F}_j$ that make up a profile $\boldsymbol{F}$ have the same "order," as required by the following condition:[12,13]

Assumption 2. Fix a profile $\boldsymbol{F} \in \Pi$. Then one of the following holds:

---

[12]One might think that one could alternatively assign an order to the product σ-algebras $\boldsymbol{F}_{-i} \in \Pi_{-i}$, rather than to the profiles $\boldsymbol{F} \in \Pi$. However, this approach does not seem to be straightforward if there are more than two players. More importantly, even in the case of two players, a type's 'rank' defined in this way can be lower than its depth, at least without additional conditions. In Example 2, for example, the 'order' of the σ-algebra $\widetilde{\mathcal{F}}_a^1$ associated with q_b^1 and q_b^2 would be 3 (because $\widetilde{\mathcal{F}}_a^1$ a-dominates the σ-algebra $\widetilde{\mathcal{F}}_b^0$ which b-dominates $\widetilde{\mathcal{F}}_a^0$, and the latter does not a-dominate any σ-algebras), but the depth of q_b^1 and q_b^2 is infinite.

[13]While this condition rules out that the beliefs of types at higher order are completely determined by their beliefs at lower order (as in Examples 1 and 2), it does *not* rule out that the beliefs of a type *about a particular event* are fixed by its beliefs at lower order. In particular, the type spaces in Section 2.2 satisfy Assumption 2.

(1) For each $j \in N$, the σ-algebra $\mathcal{F}_j$ contains the sets

$$\left\{ t_j \in T_j : h_j^{\mathcal{T},m}(t_j) \in B_m \right\}$$

for all $m = 0, 1, \ldots$ and $B_m \in \mathcal{F}_{B_j^{\mathcal{T},m}}$; or

(2) There exists $k = 1, 2, \ldots$ such that for each $j \in N$, the σ-algebra $\mathcal{F}_j$ contains the sets

$$\left\{ t_j \in T_j : h_j^{\mathcal{T},m}(t_j) \in B_m \right\}$$

for $m = 0, 1, \ldots, k-1$ and $B_m \in \mathcal{F}_{B_j^{\mathcal{T},m}}$; and

$$\left\{ t_j \in T_j : h_j^{\mathcal{T},k}(t_j) \in B_k \right\} \notin \mathcal{F}_j$$

for some $B_k \in \mathcal{F}_{B_j^{\mathcal{T},k}}$.

That is, if a σ-algebra $\mathcal{F}_j$ in a profile $\boldsymbol{F}$ does not contain every $(k+1)$th-order event, then neither does the σ-algebra $\mathcal{F}_i$, for any player $i \neq j$.[14]

Under this additional assumption, a type's rank fully characterizes its depth.

Proposition 6.3. *Suppose $\mathcal{T}$ satisfies Assumption 2, and fix a player $i \in N$ and a type $t_i \in T_i$. Then, the depth of t_i is equal to its rank, that is, $d_i(t_i) = r_i(t_i)$.*

While Assumption 2 is intuitive, it may be hard to check whether a particular type space satisfies it, given that it is a condition on the belief hierarchies. We thus present a condition that is formulated in terms of the properties of the type space alone, and show that it is equivalent to Assumption 2.

Assumption 3. Fix a profile $\boldsymbol{F} \in \Pi$. If $\mathcal{F}_i$ i-dominates $\boldsymbol{F}$ for some $i \in N$, then there exist profiles $\boldsymbol{F}^1, \boldsymbol{F}^2, \ldots \in \Pi$ (not necessarily distinct) such that

$$\boldsymbol{F} \succ \boldsymbol{F}^1 \succ \boldsymbol{F}^2 \succ \cdots .$$

That is, $\boldsymbol{F}$ is self-dominating, or is part of a cycle or infinite chain of profiles that dominate each other.

Proposition 6.4. *Let $\mathcal{T} = ((T_i, \Sigma_i, \beta_i)_{i \in N}, \Pi)$ be an extended type space. Then $\mathcal{T}$ satisfies Assumption 2 if and only if it satisfies Assumption 3.*

---

[14]Indeed, Examples 1 and 2 do not satisfy Assumption 2. In Example 1, the σ-algebra $\mathcal{F}_a^1$ contains all third-order events, but the σ-algebra $\mathcal{F}_b^0$ (with which $\mathcal{F}_a^1$ forms a profile) does not. In Example 2, the σ-algebra $\widetilde{\mathcal{F}}_a^1$ contains all higher-order events, but the σ-algebra $\widetilde{\mathcal{F}}_b^0$ (that forms the profile $\widetilde{\mathcal{F}}_a^1 \times \widetilde{\mathcal{F}}_b^0$ with $\widetilde{\mathcal{F}}_a^1$) does not contain any nontrivial higher-order events.

Together with Proposition 6.3, Proposition 6.4 implies that the depth of each type is given by its rank provided that the type space satisfies Assumption 3. Even if a type space does not satisfy Assumption 3 (or, equivalently, Assumption 2), the classification of profiles of orders can still be used to obtain a complete characterization of each type's depth, using an approach that is notationally more involved; see the online appendix.

7.   Related literature

7.1.   Bounded reasoning

Besides cognitive hierarchy and level-k models and the present framework, there are other ways of modeling that players are somehow bounded in their reasoning about the higher-order beliefs of other players. One approach is to model a player who does not have a belief about a kth-order event as having ambiguous beliefs. Ahn (2007) defines type spaces with ambiguous beliefs, but does not apply his framework to model players' depth of reasoning. Indeed, there seems to be a conceptual distinction between small worlds and a finite depth of reasoning on the one hand, and ambiguous beliefs about higher-order beliefs on the other.

Another possible approach is to model a player with a bounded depth of reasoning as a player who has incomplete preferences over states that differ only in the beliefs of the other player at high orders. Di Tillio (2008) considers a class of type structures that allow for incomplete preferences, but does not consider type spaces in which the incompleteness of a type's preferences reflect its depth of reasoning.

Finally, one could view a player with a finite depth of reasoning as a player who is unaware that she can reason about events at higher orders. Fagin and Halpern (1988) propose a logic that can capture very general forms of unawareness; and their formalism can in principle be used to model that players have a finite depth of reasoning. However, the inability of players to reason about other players' higher-order beliefs seems to be more a matter of limited computational powers than of unawareness; see, e.g., Davis and Pratt (1995) and Kinderman et al. (1998) for some supporting evidence.

Some papers study environments in which players can be unaware of some aspects of the state of nature and have a finite depth of reasoning. Pintér and Udvari (2011) generalize the extended type spaces defined here to allow for unawareness as well as bounded reasoning. However, because Pintér and Udvari do not impose a condition like Assumption 1, the depth of reasoning of a type need not be defined in their framework. Heinsalu (2011) considers a class of type spaces in which players may be unaware of certain states of nature and can be bounded in their reasoning about others in the sense of level-k or cognitive-hierarchy models

(cf. Strzalecki, 2009; Heifetz and Kets, 2011). The present formalism has the advantage that it separates the issue of bounded reasoning from the conceptually distinct issue of unawareness.[15]

7.2. Iterative versus fixed-point notions of common belief

The relation between the iterative and fixed-point account of common knowledge or common belief has been the focus of a number of papers in logic, linguistics, and game theory (e.g., Lewis, 1969; Aumann, 1976; Clark and Marshall, 1981; Milgrom, 1981; Barwise, 1988; Monderer and Samet, 1989; Halpern and Moses, 1990). The different notions of common knowledge/common belief coincide for Harsanyi type spaces and related belief structures, but the equivalence can break down in more general structures. The nature of the breakdown observed in these papers here is strikingly different from the one demonstrated here: in the structures studied thus far, common belief in the fixed-point sense implies the conjunction of mutual belief at all orders, while the opposite is true for the example in Section 2.2 (Halpern and Moses, 1990; Lismont and Mongin, 1995; Heifetz, 1999).

The failure of the equivalence result in the present case is due to the fact that the belief operator is not monotonic. Nonmonotonic belief operators have been used in epistemic game theory to characterize various solution concepts (Battigalli and Siniscalchi, 2002; Brandenburger et al., 2008, 2012). These papers use the iterative notion of common belief.

7.3. Measurable structures on type sets

One insight of the present paper is that, by choosing the measurable sets on which a type's belief is defined, we can get types that can reason about only finitely many orders of beliefs. Indeed, the technical contribution of this paper is to formulate conditions on the type space that guarantee that the σ-algebra of a type with a finite depth k lumps together precisely the types that induce belief hierarchies that coincide up to order $k-1$. The idea that a type's σ-algebra can determine its depth of reasoning fits in with a broader literature that studies how the measurable structure associated with types in Harsanyi type spaces can implicitly impose restrictions on reasoning, i.e., on belief hierarchies (e.g., Brandenburger and Keisler, 2006; Friedenberg and Meier, 2012); see Friedenberg and Keisler (2011) for a detailed discussion and further references.

---

[15]Also, Jehiel and Koessler (2008) characterize equilibrium behavior in a game when players do not condition their behavior on their higher-order beliefs; and Ciancaruso and Germano (2011) define an equilibrium notion for quotient type spaces based on Harsanyi type spaces.

8. Concluding remarks

This paper introduces a novel framework for analyzing situations in which players can have a finite depth of reasoning, building on the notion of a small world of Savage (1954). The present approach makes it possible to model situations in which players with a finite depth of reasoning are able to reason about 'simple' higher-order events, where an event is simple if it is expressible in terms of an event of sufficiently low order.

It is an open question whether, and if so, how, the strategic behavior of players with a finite depth of reasoning depends on what higher-order events they can form a belief about, that is, whether the behavior of a player with a fixed, finite depth k depends on the context she is in.

To investigate this, it is important to understand what contexts we should include in our model given certain assumptions on players' reasoning process and given the strategic situation that players face. Experiments can of course be of great help here. In addition, a syntactic model, that is, a framework that explicitly specifies what statements (about the economic fundamentals and about beliefs) players can reason about, may provide complementary insights. Assumptions on players' reasoning process put restrictions on the syntax, which in turn translate into restrictions on the class of type spaces and thus on the context; see, e.g., Samet (1990) for an early example of this approach. Together, these lines of research can perhaps ultimately lead to a better understanding of the strategic behavior of players with a finite depth of reasoning.

Appendix A   Proofs for Sections 3 and 4

A.1   Preliminary results

We start with some auxiliary results. The first result says that taking inverse images preserves σ-algebras:

Lemma A.1. *Let $f : X \to Y$ be a function from X into Y, and let $\mathcal{E}$ be a nonempty collection of subsets of Y. Then,*

$$\sigma\big(\{f^{-1}(E) : E \in \mathcal{E}\}\big) = \big\{f^{-1}(E) : E \in \sigma(\mathcal{E})\big\},$$

where $\sigma(\mathcal{E})$ is the σ-algebra generated by $\mathcal{E}$.

The proof is standard, and thus omitted.

We next show that the belief maps are measurable if and only if the relevant product σ-algebra is self-dominating. To state the result, we need some more notation. Fix an extended

type space $\mathcal{T} = ((T_i, \Sigma_i, \beta_i)_{i \in N}, \Pi)$, and for each player $i \in N$, define

$$\mathscr{S}_i := \{\mathcal{F}_\Theta \times \boldsymbol{F}_{-i} : \boldsymbol{F}_{-i} \in \Pi_{-i}\}$$

to be the collection of product σ-algebras on $\Theta \times T_{-i}$ on which the belief of a type for i can be defined. Recall that the set of probability measures on a σ-algebra in $\mathscr{S}_i$ is denoted by $\Delta(\Theta \times T_{-i}, \mathscr{S}_i)$, and that $\Delta(\Theta \times T_{-i}, \mathscr{S}_i)$ is endowed with the σ-algebra $\mathcal{F}_{\Delta(\Theta \times T_{-i}, \mathscr{S}_i)}$ that is generated by the sets

$$\big\{\mu \in \Delta(\Theta \times T_{-i}, \mathscr{S}_i\}) : E \in \Sigma(\mu), \mu(E) \geq p\big\} : \qquad E \in \mathcal{F}, \mathcal{F} \in \mathscr{S}_i, p \in [0,1].$$

We can also define the set of σ-algebras on T_j, $j \in N$, that are consistent with some profile:

$$\Pi_j := \{\mathcal{F}_j : \text{ there is } \boldsymbol{F}_{-j} \text{ such that } \mathcal{F}_j \times \boldsymbol{F}_{-j} \in \Pi\}.$$

Fix a player $i \in N$. Clearly, if for each player $j \neq i$, there is a σ-algebra $\mathcal{F}_j^* \in \Pi_j$ such that $\mathcal{F}_j \subseteq \mathcal{F}_j^*$ for all $\mathcal{F}_j \in \Pi_j$, then $\mathcal{F}_{\Delta(\Theta \times T_{-i}, \mathscr{S}_i)}$ is generated by the sets

$$\big\{\mu \in \Delta(\Theta \times T_{-i}, \mathscr{S}_i) : E \in \Sigma(\mu), \mu(E) \geq p\big\} : \qquad E \in \mathcal{F}_\Theta \times \boldsymbol{F}_{-i}^*, p \in [0,1].$$

We can now state the result:

Lemma A.2. *Let $((T_i, \Sigma_i, \beta_i)_{i \in N}, \Pi)$ be an extended type space, and suppose that for each player $j \in N$, there is $\mathcal{F}_j^* \in \Pi_j$ such that $\mathcal{F}_j \subseteq \mathcal{F}_j^*$ for all $\mathcal{F}_j \in \Pi_j$. Then, the following are equivalent:*

(a) for each player $i \in N$, the belief map β_i is measurable with respect to the σ-algebras $\mathcal{F}_i^*$ and $\mathcal{F}_{\Delta(\Theta \times T_{-i}, \mathscr{S}_i)}$;

(b) the profile $\boldsymbol{F}^*$ is self-dominating.

Proof. Fix $i \in N$. It follows from Lemma A.1 that the belief map β_i is measurable with respect to $\mathcal{F}_i^*$ and $\mathcal{F}_{\Delta(\Theta \times T_{-i}, \mathscr{S}_i)}$ if and only if for every $E \in \mathcal{F}_\Theta \times \boldsymbol{F}_{-i}^*$ and $p \in [0,1]$,

$$\big\{t_i \in T_i : E \in \mathcal{F}_\Theta \times \Sigma_i(t_i), \beta_i(t_i)(E) \geq p\big\} \in \mathcal{F}_i^*.$$

Hence, the functions $\beta_i, i \in N$, are measurable if and only if $\boldsymbol{F}^* \succ \boldsymbol{F}^*$, that is, $\boldsymbol{F}^*$ is self-dominating. $\qquad \square$

To state the next few results, we need some more notation. Given a type space $\mathcal{T} = ((T_i, \Sigma_i, \beta_i)_{i \in N}, \Pi)$, a player $i \in N$ and $k = 0, 1, \ldots$, denote the image $h_i^{\mathcal{T},k}(T_i)$ of T_i in B_i^k by $B_i^{\mathcal{T},k}$, whenever the function $h_i^{\mathcal{T},k}$ is well-defined. In that case, we endow $B_i^{\mathcal{T},k}$ with the relative

σ-algebra induced by $\mathcal{F}_{B_i^k}$. Define $\sigma(h_i^{\mathcal{T},k})$ to be the σ-algebra generated by the function $h_i^{\mathcal{T},k}$, that is,

$$\sigma\big(h_i^{\mathcal{T},k}\big) := \big\{\{t_i \in T_i : h_i^{\mathcal{T},k}(t_i) \in B\} : B \in \mathcal{F}_{B_i^{\mathcal{T},k}}\big\}.$$

It is easy to see that if $h_i^{\mathcal{T},k}$ and $h_i^{\mathcal{T},k+1}$ are well-defined, then

$$\sigma\big(h_i^{\mathcal{T},k}\big) \subseteq \sigma\big(h_i^{\mathcal{T},k+1}\big),$$

a result we use without explicit mention. Also, note that $\sigma(h_i^{\mathcal{T},0}) = \{T_i, \emptyset\}$, an observation we will use frequently. Finally, given a product space $Z = X \times Y$, we define proj_X^Z to be the projection function from Z into X.

The next result plays a central role in many of our results. Part (a) makes it possible to use an inductive proof, even though we do not have an a priori ordering on the profiles in Π. Part (b) shows that each type induces a well-defined belief hierarchy, and part (d) is critical for the characterization of profiles in Corollary A.4 below. Parts (c) and (e) are auxiliary results.

Lemma A.3. *For each* $k = 1, 2, \ldots$*, the following hold:*

(a) *For each profile* $\boldsymbol{F} \in \Pi$*, we have* $\mathcal{F}_i \subseteq \sigma(h_i^{\mathcal{T},k-1})$ *for all* $i \in N$*, or* $\mathcal{F}_i \supseteq \sigma(h_i^{\mathcal{T},k-1})$ *for all* $i \in N$*.*

(b) *For each player* $i \in N$ *and type* $t_i \in T_i$*, we have* $h_i^{\mathcal{T},k}(t_i) \in B_i^k$*.*

(c) *We have* $\bigtimes_{j \in N} \sigma(h_j^{\mathcal{T},k}) \succ^* \bigtimes_{j \in N} \sigma(h_j^{\mathcal{T},k-1})$*.*

(d) *For each profile* $\boldsymbol{F} \in \Pi$ *such that* $\mathcal{F}_i \supseteq \sigma(h_i^{\mathcal{T},k-1})$ *and* $\mathcal{F}_i \subsetneq \sigma(h_i^{\mathcal{T},k})$ *for some* $i \in N$*, there is* $m \leq k-1$ *such that* $\mathcal{F}_j = \sigma(h_j^{\mathcal{T},m})$ *for all* $j \in N$*, and* $\sigma(h_n^{\mathcal{T},m}) \subsetneq \sigma(h_n^{\mathcal{T},m+1})$ *for some player* $n \in N$*.*

(e) *For each profile* $\boldsymbol{F} \in \Pi$ *that is self-dominating and for each player* $j \in N$*, we have that* $\mathcal{F}_j \supseteq \sigma(h_j^{\mathcal{T},k})$*.*

Proof. The proof is by induction. Clearly, for each profile $\boldsymbol{F}$ and a player $i \in N$, we have that $\mathcal{F}_i \supseteq \sigma(h_i^{\mathcal{T},0})$. Also, for each player $i \in N$ and type $t_i \in T_i$, we have $h_i^{\mathcal{T},0}(t_i) \in B_i^0$. We next want to show that $h_i^{\mathcal{T},1}(t_i) \in B_i^1$. This holds if and only if $\beta_i(t_i) \circ (\mathrm{Id}_\Theta, h_{-i}^{\mathcal{T},0})^{-1}$ is a probability measure in $\Delta(\Theta \times B_{-i}^0, \mathscr{S}_i^1)$. But this is immediate: if we write $\Sigma_i(t_i) = \widetilde{\boldsymbol{F}}_{-i}$, then clearly $\widetilde{\mathcal{F}}_j \supseteq \sigma(h_j^{\mathcal{T},0})$ for all $j \neq i$, and the result follows.

The next step is to prove that $\bigtimes_{j \in N} \sigma(h_j^{\mathcal{T},1}) \succ^* \bigtimes_{j \in N} \sigma(h_j^{\mathcal{T},0})$. To see this, note that for each player $i \in N$,

$$\sigma\big(h_i^{\mathcal{T},1}\big) = \Big\{\big\{t_i \in T_i : \beta_i(t_i) \circ \big(\mathrm{Id}_\Theta, h_{-i}^{\mathcal{T},0}\big)^{-1} \in B\big\} : B \in \mathcal{F}_{\Delta(\Theta \times B_{-i}^0, \mathscr{S}_i^1)}\Big\},$$

so that by Lemma A.1, the σ-algebra $\sigma(h_i^{\mathcal{T},1})$ is generated by the sets

$$\left\{t_i \in T_i : E \in \mathcal{F}_\Theta \times \Sigma_i(t_i), \beta_i(t_i)(E) \geq p\right\} : \qquad E \in \mathcal{F}_\Theta \times \bigtimes_{j \neq i} \sigma(h_j^{\mathcal{T},0}), p \in [0,1],$$

and the result follows. We can apply this immediately to show that if a profile $\boldsymbol{F} \in \Pi$ is self-dominating, then $\mathcal{F}_i \supseteq \sigma(h_i^{\mathcal{T},1})$ for all $i \in N$. To show this, fix a profile $\boldsymbol{F} \in \Pi$ that is self-dominating and a player $i \in N$. Because $\mathcal{F}_j \supseteq \sigma(h_j^{\mathcal{T},0})$ for all $j \in N$, it follows from the fact that $\boldsymbol{F} \succ \boldsymbol{F}$ that $\mathcal{F}_i \supseteq \sigma(h_i^{\mathcal{T},1})$.

Finally, fix a profile $\boldsymbol{F} \in \Pi$ and suppose that there is a player $i \in N$ such that $\mathcal{F}_i \supseteq \sigma(h_i^{\mathcal{T},0})$ and $\mathcal{F}_i \subsetneq \sigma(h_i^{\mathcal{T},1})$. We claim that $\mathcal{F}_j = \sigma(h_j^{\mathcal{T},0})$ for all $j \in N$. Suppose by contradiction that there is a player $n \in N$ such that $\mathcal{F}_n \neq \sigma(h_n^{\mathcal{T},0})$. Then, by Assumption 1, there is a profile $\boldsymbol{F}' \in \Pi$ such that $\boldsymbol{F} \succ \boldsymbol{F}'$ (where possibly $\boldsymbol{F}' = \boldsymbol{F}$). Since $\mathcal{F}_j' \supseteq \sigma(h_j^{\mathcal{T},0})$ for each player $j \in N$, this implies that $\boldsymbol{F} \succ \bigtimes_{j \in N} \sigma(h_j^{\mathcal{T},0})$. But because for each player $j \in N$, the σ-algebra $\sigma(h_j^{\mathcal{T},1})$ is the coarsest σ-algebra that j-dominates $\bigtimes_{j' \in N} \sigma(h_{j'}^{\mathcal{T},0})$, we have $\mathcal{F}_i \supseteq \sigma(h_i^{\mathcal{T},1})$, a contradiction.

For $k > 1$, suppose that for all $\ell \leq k - 1$, we have established the following:

(a') For each $\boldsymbol{F} \in \Pi$, we have $\mathcal{F}_i \subseteq \sigma(h_i^{\mathcal{T},\ell-1})$ for all $i \in N$, or $\mathcal{F}_i \supseteq \sigma(h_i^{\mathcal{T},\ell-1})$ for all $i \in N$.

(b') For each $i \in N$ and $t_i \in T_i$, we have $h_i^{\mathcal{T},\ell}(t_i) \in B_i^\ell$.

(c') We have $\bigtimes_{j \in N} \sigma(h_j^{\mathcal{T},\ell}) \succ^* \bigtimes_{j \in N} \sigma(h_j^{\mathcal{T},\ell-1})$.

(d') For each $\boldsymbol{F} \in \Pi$ such that $\mathcal{F}_i \supseteq \sigma(h_i^{\mathcal{T},\ell-1})$ and $\mathcal{F}_i \subsetneq \sigma(h_i^{\mathcal{T},\ell})$ for some $i \in N$, there is $m \leq \ell - 1$ such that $\mathcal{F}_j = \sigma(h_j^{\mathcal{T},m})$ for all $j \in N$, and $\sigma(h_n^{\mathcal{T},m}) \subsetneq \sigma(h_n^{\mathcal{T},m+1})$ for some player $n \in N$.

(e') For each profile $\boldsymbol{F} \in \Pi$ that is self-dominating and for each player $j \in N$, we have that $\mathcal{F}_j \supseteq \sigma(h_j^{\mathcal{T},\ell})$.

The proof then follows from the following claims.

Claim 1. For each $\boldsymbol{F} \in \Pi$, we have $\mathcal{F}_i \subseteq \sigma(h_i^{\mathcal{T},k-1})$ for all $i \in N$, or $\mathcal{F}_i \supseteq \sigma(h_i^{\mathcal{T},k-1})$ for all $i \in N$.

Proof of Claim 1. Let $\boldsymbol{F} \in \Pi$. We want to show that either $\mathcal{F}_i \subseteq \sigma(h_i^{\mathcal{T},k-1})$ for each player $i \in N$, or $\mathcal{F}_i \supseteq \sigma(h_i^{\mathcal{T},k-1})$ for each player $i \in N$. If $\boldsymbol{F}$ is self-dominating, then the result follows from the induction hypothesis (e'). So suppose $\boldsymbol{F}$ is not self-dominating. If $\mathcal{F}_j = \{T_j, \emptyset\}$ for each player $j \in N$, then clearly $\mathcal{F}_j \subseteq \sigma(h_j^{\mathcal{T},k-1})$ for each player $j \in N$. Hence, suppose there is a player $n \in N$ such that $\mathcal{F}_n \neq \{T_n, \emptyset\}$. Then, by Assumption 1, there is a profile $\boldsymbol{F}' \in \Pi$ such that $\boldsymbol{F} \succ^* \boldsymbol{F}'$. We claim that $\boldsymbol{F}'$ is not self-dominating. For suppose not. Then $\boldsymbol{F} \succ^* \boldsymbol{F}'$

implies that $\mathcal{F}_j' \supseteq \mathcal{F}_j$ for each player $j \in N$. But then the fact that $\boldsymbol{F}$ dominates $\boldsymbol{F}'$ implies that $\boldsymbol{F}$ dominates itself, a contradiction.

It will be useful to define $\mathcal{F}_j^0 := \mathcal{F}_j$ and $\mathcal{F}_j^1 := \mathcal{F}_j'$ for $j \in N$. Then, by Assumption 1, there exist profiles $\boldsymbol{F}^2, \boldsymbol{F}^3, \ldots \in \Pi$ such that $\boldsymbol{F}^\ell$, $\ell \geq 0$, is not self-dominating, and one of the following holds:

(i) the profiles form a cycle or infinite chain, that is,

$$\boldsymbol{F}^0 \succ^* \boldsymbol{F}^1 \succ^* \boldsymbol{F}^2 \succ^* \cdots ;$$

(ii) the profiles form a finite chain, that is, there is $m < \infty$ such that

$$\boldsymbol{F}^0 \succ^* \boldsymbol{F}^1 \succ^* \boldsymbol{F}^2 \succ^* \cdots \succ^* \boldsymbol{F}^m = \bigtimes_{j \in N} \{T_j, \emptyset\}.$$

(Note that there is no profile $\boldsymbol{F}^{m+1} \in \Pi$ such that $\boldsymbol{F}^m \succ^* \boldsymbol{F}^{m+1}$, given that $\boldsymbol{F}^m = \bigtimes_{j \in N} \{T_j, \emptyset\}$ is not self-dominating.)

First consider (i). Because $\mathcal{F}_j^1 \supseteq \sigma(h_j^{\mathcal{T},0})$ for $j \in N$, it follows from the induction hypothesis (c') that $\mathcal{F}_j^0 \supseteq \sigma(h_j^{\mathcal{T},1})$ for each $j \in N$. By a similar argument, $\mathcal{F}_j^1 \supseteq \sigma(h_j^{\mathcal{T},1})$ for all $j \in N$. Repeating this argument, we find that $\mathcal{F}_j \supseteq \sigma(h_j^{\mathcal{T},k-1})$ for all $j \in N$.

Next consider (ii). If $m \leq k - 1$, then it follows from the induction hypothesis (c') that $\mathcal{F}_j = \sigma(h_j^{\mathcal{T},m}) \subseteq \sigma(h_j^{\mathcal{T},k-1})$ for all $j \in N$. If $m > k - 1$, then, by the induction hypothesis (c'), we have

$$\boldsymbol{F}^0 \succ^* \boldsymbol{F}^1 \succ^* \boldsymbol{F}^2 \succ^* \cdots \succ^* \boldsymbol{F}^{m-(k-1)} \succ^* \bigtimes_{j \in N} \sigma(h_j^{\mathcal{T},k-2}).$$

By the induction hypothesis (c'), we have $\mathcal{F}_j^{m-(k-1)} = \sigma(h_j^{\mathcal{T},k-1})$ for each $j \in N$; and the fact that $\boldsymbol{F}^{m-(k-1)-1} \succ^* \boldsymbol{F}^{m-(k-1)}$ implies that $\boldsymbol{F}^{m-(k-1)-1} \succ^* \bigtimes_{j \in N} \sigma(h_j^{\mathcal{T},k-2})$. Hence, $\mathcal{F}_j^{m-(k-1)-1} \supseteq \sigma(h_j^{\mathcal{T},k-1})$ for each $j \in N$. Repeating this argument gives that for each $j \in N$, $\mathcal{F}_j \supseteq \sigma(h_j^{\mathcal{T},k-1})$. $\qquad \square$

Claim 2. For each $i \in N$ and $t_i \in T_i$, we have $h_i^{\mathcal{T},k}(t_i) \in B_i^k$.

Proof of Claim 2. Fix a player i and a type $t_i \in T_i$. By the induction hypothesis (b'), we have $h_i^{\mathcal{T},k}(t_i) \in B_i^k$ if and only if $\beta_i(t_i) \cdot (\mathrm{Id}_\Theta, h_{-i}^{\mathcal{T},k-1})^{-1}$ is an element of $\Delta(\Theta \times B_{-i}^{k-1}, \mathscr{S}_i^k)$, i.e., for each $j \neq i$, there is a sub-σ algebra $\mathcal{A}_j^{k-1}$ of $\mathcal{F}_{B_j^{k-1}}$ such that for each $E \in \boldsymbol{A}_{-i}^{k-1}$, $(h_{-i}^{\mathcal{T},k-1})^{-1}(E) \in \Sigma_i(t_i)$.

Let $\boldsymbol{F} \in \Pi$ be a profile consistent with $\Sigma_i(t_i)$, that is, $\Sigma_i(t_i) = \boldsymbol{F}_{-i}$. By Claim 1, we have $\mathcal{F}_j \subseteq \sigma(h_j^{\mathcal{T},k-1})$ for all $j \in N$, or $\mathcal{F}_j \supseteq \sigma(h_j^{\mathcal{T},k-1})$ for all $j \in N$.

First suppose that $\mathcal{F}_j \subseteq \sigma(h_j^{\mathcal{T},k-1})$ for all $j \neq i$ and $\mathcal{F}_n \subsetneq \sigma(h_n^{\mathcal{T},k-1})$ for some $n \neq i$. By the induction hypotheses (a') and (d'), there is $m' < k - 1$ such that $\mathcal{F}_j = \sigma(h_j^{\mathcal{T},m'})$ for all $j \neq i$;

34

take m to be the maximum $m' < k - 1$ for which this holds. Using that for each $j \neq i$, the σ-algebra $\mathcal{F}_{B_j^{\mathcal{T},m}}$ is the relative σ-algebra on $B_j^{\mathcal{T},m} = h_j^{\mathcal{T},m}(T_j)$ induced by $\mathcal{F}_{B_j^m}$, we have

$$
\begin{aligned}
\sigma\big(h_j^{\mathcal{T},m}\big) &= \Big\{ \{t_j \in T_j : h_j^{\mathcal{T},m}(t_j) \in B\} : B \in \mathcal{F}_{B_j^{\mathcal{T},m}} \Big\} \\
&= \Big\{ \{t_j \in T_j : h_j^{\mathcal{T},m}(t_j) \in B \cap B_j^{\mathcal{T},m}\} : B \in \mathcal{F}_{B_j^m} \Big\} \\
&= \Big\{ \{t_j \in T_j : h_j^{\mathcal{T},m}(t_j) \in B\} : B \in \mathcal{F}_{B_j^m} \Big\}.
\end{aligned}
$$

Consequently, $\beta_i(t_i) \circ (\mathrm{Id}_\Theta, h_{-i}^{\mathcal{T},k-1})^{-1}$ is a probability measure on the product σ-algebra $\mathcal{F}_\Theta \times \big\{ \big(\mathrm{proj}_{B_{-i}^m}^{B_{-i}^{k-1}}\big)^{-1}(B) : B \in \bigtimes_{j \neq i} \mathcal{F}_{B_j^m} \big\}$. Since the projection function is measurable, this is a sub-σ algebra of $\mathcal{F}_\Theta \times \bigtimes_{j \neq i} \mathcal{F}_{B_j^{k-1}}$, and it follows that $h_i^{\mathcal{T},k}(t_i) \in B_i^k$.

Finally, if $\mathcal{F}_j \supseteq \sigma(h_j^{\mathcal{T},k-1})$ for all $j \neq i$, then $\beta_i(t_i) \circ (\mathrm{Id}_\Theta, h_{-i}^{\mathcal{T},k-1})^{-1}$ is a probability measure on $\mathcal{F}_\Theta \times \bigtimes_{j \neq i} \mathcal{F}_{B_j^{k-1}}$, and the result is immediate. $\qquad\square$

By Claim 2, the function $h_i^{\mathcal{T},k}$ is well-defined for $i \in N$, so that the σ-algebra $\sigma(h_i^{\mathcal{T},k})$ is well-defined.

Claim 3. We have $\bigtimes_{j \in N} \sigma(h_j^{\mathcal{T},k}) \succsim^* \bigtimes_{j \in N} \sigma(h_j^{\mathcal{T},k-1})$.

Proof of Claim 3. If we prove that for each player $i \in N$, the σ-algebra $\sigma(h_i^{\mathcal{T},k})$ is generated by the sets

$$
\Big\{ t_i \in T_i : \big(\mathrm{Id}_\Theta, h_{-i}^{\mathcal{T},k-1}\big)^{-1}(E) \in \mathcal{F}_\Theta \times \Sigma_i(t_i), \beta_i(t_i) \circ \big(\mathrm{Id}_\Theta, h_{-i}^{\mathcal{T},k-1}\big)^{-1}(E) \geq p \Big\} :
$$
$$
E \in \mathcal{F}_\Theta \times \bigtimes_{j \neq i} \mathcal{F}_{B_j^{\mathcal{T},k-1}}, p \in [0,1], \quad \text{(A.1)}
$$

then the result is immediate: If $\sigma(h_i^{\mathcal{T},k})$ is generated by the sets in (A.1), then it follows directly from the definition of $\sigma(h_j^{\mathcal{T},k-1})$, $j \neq i$, that $\sigma(h_i^{\mathcal{T},k})$ is the coarsest σ-algebra that includes the sets

$$
\{t_i \in T_i : E \in \mathcal{F}_\Theta \times \Sigma_i(t_i), \beta_i(t_i)(E) \geq p\}
$$

for $E \in \mathcal{F}_\Theta \times \bigtimes_{j \neq i} \sigma(h_j^{\mathcal{T},k-1})$ and $p \in [0,1]$, or, equivalently, that $\bigtimes_{j \in N} \sigma(h_j^{\mathcal{T},k}) \succsim^* \bigtimes_{j \in N} \sigma(h_j^{\mathcal{T},k-1})$.

Hence, it remains to show that $\sigma(h_i^{\mathcal{T},k})$ is generated by the sets in (A.1). We can write

$$
\sigma(h_i^{\mathcal{T},k}) = \Big\{ \{t_i \in T_i : h_i^{\mathcal{T},k}(t_i) \in B\} : B \in \mathcal{F}_{B_i^{\mathcal{T},k-1}} \times \mathcal{F}_{\Delta(\Theta \times B_{-i}^{k-1}, \mathscr{S}_i^k)} \Big\}.
$$

Because every σ-algebra in $\mathscr{S}_i^k$ is a sub-σ algebra of $\mathcal{F}_\Theta \times \bigtimes_{j \neq i} \mathcal{F}_{B_j^{k-1}}$, the σ-algebra $\mathcal{F}_{\Delta(\Theta \times B_{-i}^{k-1}, \mathscr{S}_i^k)}$ is generated by the sets

$$
\big\{ \mu_i^k \in \Delta(\Theta \times B_{-i}^{k-1}, \mathscr{S}_i^k) : E \in \Sigma(\mu_i^k), \mu_i^k(E) \geq p \big\} : \quad E \in \mathcal{F}_\Theta \times \bigtimes_{j \neq i} \mathcal{F}_{B_j^{k-1}}, p \in [0,1].
$$

It then follows from these observations and Lemma A.1 that the σ-algebra $\sigma(h_i^{\mathcal{T},k})$ is generated by the sets in $\sigma(h_i^{\mathcal{T},k-1})$ and the sets

$$\left\{ t_i \in T_i : \left(\mathrm{Id}_\Theta, h_{-i}^{\mathcal{T},k-1} \right)^{-1}(E) \in \mathcal{F}_\Theta \times \Sigma_i(t_i), \beta_i(t_i) \circ \left(\mathrm{Id}_\Theta, h_{-i}^{\mathcal{T},k-1} \right)^{-1}(E) \geq p \right\} :$$
$$E \in \mathcal{F}_\Theta \times \underset{j\neq i}{\bigtimes} \mathcal{F}_{B_j^{k-1}}, p \in [0,1]. \quad \text{(A.2)}$$

The proof is complete if we show that the σ-algebra $\sigma(h_i^{\mathcal{T},k-1})$ is generated by a subset of the family of sets in (A.2), given that the collection of sets in (A.2) is just the generating family of sets in (A.1) (by a similar transformation as in the proof of Claim 2).

By the induction hypothesis (c'), the σ-algebra $\sigma(h_i^{\mathcal{T},k-1})$ is generated by the sets

$$\left\{ t_i \in T_i : \left(\mathrm{Id}_\Theta, h_{-i}^{\mathcal{T},k-2} \right)^{-1}(E) \in \mathcal{F}_\Theta \times \Sigma_i(t_i), \beta_i(t_i) \circ \left(\mathrm{Id}_\Theta, h_{-i}^{\mathcal{T},k-2} \right)^{-1}(E) \geq p \right\} :$$
$$E \in \mathcal{F}_\Theta \times \underset{j\neq i}{\bigtimes} \mathcal{F}_{B_j^{\mathcal{T},k-2}}, p \in [0,1].$$

Fix $E \in \mathcal{F}_\Theta \times \bigtimes_{j\neq i} \mathcal{F}_{B_j^{\mathcal{T},k-2}}$ and $p \in [0,1]$, and define $E' := \left(\mathrm{proj}_{\Theta \times B_{-i}^{\mathcal{T},k-2}}^{\Theta \times B_{-i}^{\mathcal{T},k-1}} \right)^{-1}(E)$. Then $E' \in \mathcal{F}_\Theta \times \bigtimes_{j\neq i} \mathcal{F}_{B_j^{\mathcal{T},k-1}}$, and

$$\left\{ t_i \in T_i : \left(\mathrm{Id}_\Theta, h_{-i}^{\mathcal{T},k-2} \right)^{-1}(E) \in \mathcal{F}_\Theta \times \Sigma_i(t_i), \beta_i(t_i) \circ \left(\mathrm{Id}_\Theta, h_{-i}^{\mathcal{T},k-2} \right)^{-1}(E) \geq p \right\} =$$
$$\left\{ t_i \in T_i : \left(\mathrm{Id}_\Theta, h_{-i}^{\mathcal{T},k-1} \right)^{-1}(E') \in \mathcal{F}_\Theta \times \Sigma_i(t_i), \beta_i(t_i) \circ \left(\mathrm{Id}_\Theta, h_{-i}^{\mathcal{T},k-1} \right)^{-1}(E') \geq p \right\}.$$

Hence, the σ-algebra $\sigma(h_i^{\mathcal{T},k-1})$ is generated by a subset of the sets in (A.2). $\qquad \square$

Claim 4. For each profile $\boldsymbol{F} \in \Pi$ that is self-dominating, we have $\mathcal{F}_j \supseteq \sigma(h_j^{\mathcal{T},k})$ for each player $j \in N$.

Proof of Claim 4. Immediate from the induction hypothesis (e') and Claim 3. $\qquad \square$

Claim 5. For each $\boldsymbol{F} \in \Pi$ such that $\mathcal{F}_i \supseteq \sigma(h_i^{\mathcal{T},k-1})$ and $\mathcal{F}_i \subsetneq \sigma(h_i^{\mathcal{T},k})$ for some $i \in N$, there is $m \leq k-1$ such that $\mathcal{F}_j = \sigma(h_j^{\mathcal{T},m})$ for all $j \in N$, and $\sigma(h_n^{\mathcal{T},m}) \subsetneq \sigma(h_n^{\mathcal{T},m+1})$ for some player $n \in N$.

Proof of Claim 5. By Claim 1, either $\mathcal{F}_j \subseteq \sigma(h_j^{\mathcal{T},k-1})$ for all $j \in N$, or $\mathcal{F}_j \supseteq \sigma(h_j^{\mathcal{T},k-1})$ for all $j \in N$. If $\mathcal{F}_j = \{T_j, \emptyset\}$ for $j \in N$, then the result follows immediately. So suppose there is some player $n \in N$ such that $\mathcal{F}_n \neq \{T_n, \emptyset\}$. By Claim 4, the profile $\boldsymbol{F}$ does not dominate itself. Hence, by Assumption 1, there is a profile $\boldsymbol{F}' \in \Pi$ such that $\boldsymbol{F} \succ \boldsymbol{F}'$. Define $\mathcal{F}_j^0 := \mathcal{F}_j$ and $\mathcal{F}_j^1 := \mathcal{F}_j'$ for $j \in N$. Then, by a similar argument as in the proof of Claim 1, it follows that there exist $m < \infty$ and profiles $\boldsymbol{F}^2, \boldsymbol{F}^3, \ldots, \boldsymbol{F}^m \in \Pi$ such that $\boldsymbol{F}^\ell$, $\ell = 0, \ldots, m$, does not dominate itself, and

$$\boldsymbol{F}^0 \succ^* \boldsymbol{F}^1 \succ^* \boldsymbol{F}^2 \succ^* \cdots \succ^* \boldsymbol{F}^m = \underset{j\in N}{\bigtimes} \{T_j, \emptyset\}.$$

Since $\mathcal{F}_i \subsetneq \sigma(h_i^{\mathcal{T},k})$, it follows from (c') and Claim 3 that the maximum m for which this holds is at most $k-1$, and $\mathcal{F}_j = \mathcal{F}_j^0 = \sigma(h_j^{\mathcal{T},m})$ for all j. $\qquad\square$

This completes the induction. $\qquad\square$

Lemma A.3 has a useful corollary:

Corollary A.4. *Fix a profile $\boldsymbol{F} \in \Pi$. Then one of the following holds:*

(a) *for each player $j \in N$, $\mathcal{F}_j \supseteq \sigma(h_j^{\mathcal{T},m})$ for all m; or*

(b) *there is $k = 0, 1, \ldots$ such that $\mathcal{F}_j = \sigma(h_j^{\mathcal{T},k})$ for each player $j \in N$, and $\mathcal{F}_i \subsetneq \sigma(h_i^{\mathcal{T},k+1})$ for some player $i \in N$.*

Proof. By Lemma A.3(a), for every $k = 0, 1, \ldots$, one of the following is the case:

- for each player j, $\mathcal{F}_j \subseteq \sigma(h_j^{\mathcal{T},k})$; or

- for each player j, $\mathcal{F}_j \supseteq \sigma(h_j^{\mathcal{T},k})$.

If for each $j \in N$, $\mathcal{F}_j \supseteq \sigma(h_j^{\mathcal{T},k})$ for all k, then we have (a). So suppose there is $k = 0, 1, \ldots$ and a player $i \in N$ such that $\mathcal{F}_i \subsetneq \sigma(h_i^{\mathcal{T},k+1})$. Then, by Lemma A.3, there is $m \leq k$ such that $\mathcal{F}_j = \sigma(h_j^{\mathcal{T},m})$ for each $j \in N$, and $\mathcal{F}_n \subsetneq \sigma(h_n^{\mathcal{T},m+1})$ for some player $n \in N$. $\qquad\square$

A.2 Proof of Proposition 3.1

It follows from Lemma A.2 that the structure $((T_i^*, \Sigma_i^*, \beta_i^*)_{i \in N}, \Pi^*)$ satisfies Assumption 1 (with $\mathscr{S}_i := \{\mathcal{F}_\Theta \times \boldsymbol{F}_{-i}^{\mathcal{H}}\}$). $\qquad\square$

A.3 Proof of Proposition 4.1

This follows directly from Lemma A.3(b). $\qquad\square$

A.4 Proof of Theorem 4.2

Write $\boldsymbol{F}_{-i} = \bigtimes_{j \neq i} \mathcal{F}_j$ for $\Sigma_i(t_i)$. By Corollary A.4, one of the following is the case:

(a) for all m, $\mathcal{F}_j \supseteq \sigma(h_j^{\mathcal{T},m})$ for all $j \neq i$; or

(b) there is $k = 0, 1, \ldots$ such that for each $j \neq i$, $\mathcal{F}_j = \sigma(h_j^{\mathcal{T},k-1})$, and $\mathcal{F}_n \subsetneq \sigma(h_n^{\mathcal{T},k})$ for some $n \neq i$.

The proof now follows directly from the definitions. Fix k. If for each $j \neq i$, $\mathcal{F}_j \supseteq \sigma(h_j^{\mathcal{T},k})$, then it follows from the definition of a product σ-algebra and Lemma A.1 that for any $B \in \bigtimes_{j \neq i} \mathcal{F}_{B_j^{\mathcal{T},k}}$,

$$\left\{ t_{-i} \in T_{-i} : h_{-i}^{\mathcal{T},k}(t_{-i}) \in B \right\} \in \Sigma_i(t_i).$$

If $\mathcal{F}_n \subsetneq \sigma(h_n^{\mathcal{T},k})$ for some $n \neq i$, on the other hand, then there is $B \in \bigtimes_{j \neq i} \mathcal{F}_{B_j^{\mathcal{T},k}}$ such that

$$\left\{ t_{-i} \in T_{-i} : h_{-i}^{\mathcal{T},k}(t_{-i}) \in B \right\} \notin \Sigma_i(t_i).$$

Consequently, if for all $j \neq i$, $\mathcal{F}_j \supseteq \sigma(h_j^{\mathcal{T},m})$ for all m, then $d_i(t_i) = \infty$. If there is $k = 1, 2, \ldots$ such that $\mathcal{F}_j = \sigma(h_j^{\mathcal{T},k-1})$ for all $j \neq i$, and $\mathcal{F}_n \subsetneq \sigma(h_n^{\mathcal{T},k})$ for some $n \neq i$, then $d_i(t_i) = k$. $\qquad\square$

A.5   Proof of Proposition 4.3

The proof follows directly from Lemmas A.2 and Lemma A.3(e). $\qquad\square$

A.6   Proof of Proposition 4.4

Because the depth $d_i(t_i)$ of type t_i is $k < \infty$, it follows from Corollary A.4 that

$$\Sigma_i(t_i) = \bigtimes_{j \neq i} \sigma(h_j^{\mathcal{T},k-1}) = \left\{ \left\{ t_{-i} \in T_{-i} : h_{-i}^{\mathcal{T},k-1}(t_{-i}) \in B_{k-1} \right\} : B_{k-1} \in \bigtimes_{j \neq i} \mathcal{F}_{B_j^{\mathcal{T},k-1}} \right\}. \qquad \text{(A.3)}$$

(Again, use the definition of a product σ-algebra and Lemma A.1.) For any $B \subseteq H_{-i}$, if there is $B_{k-1} \in \bigtimes_{j \neq i} \mathcal{F}_{B_j^{\mathcal{T},k-1}}$ such that

$$\left\{ t_{-i} \in T_{-i} : h_{-i}^{\mathcal{T}}(t_{-i}) \in B \right\} = \left\{ t_{-i} \in T_{-i} : h_{-i}^{\mathcal{T},k-1}(t_{-i}) \in B_{k-1} \right\},$$

then, clearly,

$$\left\{ t_{-i} \in T_{-i} : h_{-i}^{\mathcal{T}}(t_{-i}) \in B \right\} \in \Sigma_i(t_i).$$

The converse is also immediate from A.3. $\qquad\square$

Appendix B   Properties of the belief operator

This appendix considers the properties of the belief operator. The belief operator coincides with the standard one in Harsanyi type spaces. The next few results show that it satisfies the usual introspection properties even if players have a finite depth of reasoning.

Lemma B.1. (Positive introspection (1)) *For each $G \subseteq \Theta \times T$ and $i \in N$, we have that $\mathcal{B}_i(G) \subseteq \mathcal{B}_i(\mathcal{B}_i(G))$.*

Proof. Fix $i \in N$, and suppose $(\theta, t) \in \mathcal{B}_i(G)$. Then, $[\mathcal{B}_i(G)]_{t_i} = \Theta \times T_{-i}$, so $[\mathcal{B}_i(G)]_{t_i} \in \Sigma_i(t_i)$, and $\beta_i(t_i)([\mathcal{B}_i(G)]_{t_i}) = 1$. Hence, $(\theta, t) \in \mathcal{B}_i(\mathcal{B}_i(G))$. $\qquad\square$

Lemma B.2. (Positive introspection (2)) *For each $G \subseteq \Theta \times T$ and $i \in N$, we have that $\mathcal{B}_i(\mathcal{B}_i(G)) \subseteq \mathcal{B}_i(G)$.*

Proof. Fix $i \in N$, and let $(\theta, t) \in \mathcal{B}_i(\mathcal{B}_i(G))$. Then, $[\mathcal{B}_i(G)]_{t_i} \neq \emptyset$, that is, there exist θ', t'_{-i} such that $(\theta', t_i, t'_{-i}) \in \mathcal{B}_i(G)$. Since $\mathcal{B}_i(G)$ is of the form $\Theta \times Q_i \times T_{-i}$, with $Q_i \subseteq T_i$, we can set $\theta' = \theta$ and $t'_{-i} = t_{-i}$, and the result follows. $\qquad\square$

Given an event $G \subseteq \Theta \times T$ and a player $i \in N$, define

$$\mathcal{NB}_i(G) := \left\{ (\theta, t) : G_{t_i} \notin \mathcal{F}_\Theta \times \Sigma_i(t_i), \text{ or } \beta_i(t_i)(G_{t_i}) < 1 \right\}$$

to be the event that i cannot reason about G, or assigns less than full probability to G.

Lemma B.3. (Negative introspection) *For each $G \subseteq \Theta \times T$ and $i \in N$, we have that $\mathcal{NB}_i(G) \subseteq \mathcal{B}_i(\mathcal{NB}_i(G))$.*

Proof. Immediate $\qquad\square$

However, the belief operator fails some of the other standard properties of belief operators, at least partly. Importantly, the belief operator is not monotonic when players can have a finite depth, in the sense that the usual monotonicity property holds only for events that belong to the σ-algebras of the pertinent types.

Lemma B.4. (Partial monotonicity) *Let $G, G' \subseteq \Theta \times T$ and $i \in N$. If $G \subseteq G'$ and $G'_{t_i} \in \mathcal{F}_\Theta \times \Sigma_i(t_i)$ for all $t_i \in T_i$ such that $(\theta, t_i, t_{-i}) \in \mathcal{B}_i(G)$ for some $(\theta, t_{-i}) \in \Theta \times T_{-i}$, then $\mathcal{B}_i(G) \subseteq \mathcal{B}_i(G')$.*

Proof. Let $t_i \in T_i$ be such that $(\theta, t_i, t_{-i}) \in \mathcal{B}_i(G)$ for some $(\theta, t_{-i}) \in \Theta \times T_{-i}$. By assumption, $G'_{t_i} \in \mathcal{F}_\Theta \times \Sigma_i(t_i)$, so that $\beta_i(t_i)(G'_{t_i})$ is well-defined. It then follows immediately that $\beta_i(t_i)(G'_{t_i}) = 1$. $\qquad\square$

Clearly, if the type space is derived from a Harsanyi type space (so that every type t_i for player i has the same σ-algebra $\Sigma_i(t_i) = \boldsymbol{F}^{\mathcal{H}}_{-i}$), then the belief operator is monotonic for events in the standard σ-algebra $\mathcal{F}_\Theta \times \boldsymbol{F}^{\mathcal{H}}$.

Finally, we note that while the belief operator fails one the directions of the standard conjunction property when it is not monotonic, it does satisfy the other:

Lemma B.5. (Partial conjunction) *Let $E_1, E_2, \ldots \subseteq \Theta \times T$ and fix $i \in N$. Then,* $\bigcap_{k=1}^{\infty} \mathcal{B}_i(E_k) \subseteq \mathcal{B}_i\left(\bigcap_{k=1}^{\infty} E_k\right)$.

Proof. Suppose that $(\theta, t) \in \bigcap_{k=1}^{\infty} \mathcal{B}_i(E_k)$. Then, $[E_k]_{t_i} \in \mathcal{F}_\Theta \times \Sigma_i(t_i)$ for all k, so that $[\bigcap_{k=1}^{\infty} E_k]_{t_i} \in \mathcal{F}_\Theta \times \Sigma_i(t_i)$. Moreover,

$$
\beta_i(t_i)\left(\Big[\bigcap_{k=1}^{\infty} E_k\Big]_{t_i}\right) \;=\; 1 - \beta_i(t_i)\left(\bigcup_{k=1}^{\infty} \big\{(\theta', t'_{-i}) : (\theta', t_i, t'_{-i}) \notin E_k\big\}\right)
$$

$$
\geq\; 1 - \sum_{k=1}^{\infty} \beta_i(t_i)\big(\big\{(\theta', t'_{-i}) : (\theta', t_i, t'_{-i}) \notin E_k\big\}\big) = 1.
$$

Hence, $(\theta, t) \in \mathcal{B}_i\left(\bigcap_{k=1}^{\infty} E_k\right)$. $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad$ □

Again, if the type space is derived from a Harsanyi type space, then the reverse inclusion also holds for the case that $E_1, E_2, \ldots \in \mathcal{F}_\Theta \times \boldsymbol{F}^{\mathcal{H}}$, as the belief operator is monotonic in that case.

Appendix C Proofs for Sections 5 and 6

C.1 Proof of Proposition 5.1

Suppose $(\theta, t) \in C^{\text{FP}}(F)$. That is, there is $X \subseteq \Theta \times T$ such that $(\theta, t) \in X$, and $X = \mathcal{B}(F \cap X)$. Fix $i \in N$. Then, $[F \cap X]_{t_i} \in \mathcal{F}_\Theta \times \Sigma_i(t_i)$, and $\beta_i(t_i)([F \cap X]_{t_i}) = 1$. As F is a primitive event, this implies that $F_{t_i} \in \mathcal{F}_\Theta \times \Sigma_i(t_i)$ and $X_{t_i} \in \mathcal{F}_\Theta \times \Sigma_i(t_i)$. By Lemma B.4, $\beta_i(t_i)(F_{t_i}) = 1$ and $\beta_i(t_i)(X_{t_i}) = 1$. Hence, for all $i \in N$, $X \subseteq \mathcal{B}_i(F)$ and $X \subseteq \mathcal{B}_i(X)$, and $(\theta, t) \in C^{\text{PUB}}(F)$.

To prove the converse, suppose that $(\theta, t) \in C^{\text{PUB}}(F)$. That is, there is $E \subseteq \Theta \times T$ such that $(\theta, t) \in E$, and for all $i \in N$, $E \subseteq \mathcal{B}_i(E)$ and $E \subseteq \mathcal{B}_i(F)$. Define

$$
E' := \{\theta' \in \Theta : \text{ there is } t' \in T \text{ s.t. } (\theta', t') \in E\} \times
$$

$$
\underset{i \in N}{\times} \{t'_i \in T_i : \text{ there is } (\theta', t'_{-i}) \in \Theta \times T_{-i} \text{ s.t. } (\theta', t') \in E\}
$$

to be the product event corresponding to E; note that $(\theta, t) \in E'$. It is easy to verify that for each $i \in N$, we have that $E' \subseteq \mathcal{B}_i(F)$. Similarly, $E' \subseteq \mathcal{B}_i(E)$. To show that $E' \subseteq \mathcal{B}_i(E')$, note that for each $i \in N$ and $t''_i \in T_i$, we have that $E'_{t''_i} \in \mathcal{F}_\Theta \times \Sigma_i(t''_i)$ whenever $E_{t''_i} \in \mathcal{F}_\Theta \times \Sigma_i(t''_i)$. Hence, by Lemma B.4, $E' \subseteq \mathcal{B}_i(E) \subseteq \mathcal{B}_i(E')$ for each $i \in N$.

We show that E' is a fixed point of f_F. Fix $(\theta', t') \in E'$ and $i \in N$. Then, $E'_{t'_i}, F_{t'_i} \in \mathcal{F}_\Theta \times \Sigma_i(t'_i)$, so that $[E' \cap F]_{t'_i} \in \mathcal{F}_\Theta \times \Sigma_i(t'_i)$. Moreover, by Lemma B.5, we have that $\beta_i(t'_i)([E' \cap F]_{t'_i}) = 1$. Hence, $(\theta', t') \in \mathcal{B}(E' \cap F)$, and it follows that $E' \subseteq \mathcal{B}(E' \cap F)$.

To prove that $\mathcal{B}(E' \cap F) \subseteq E'$, fix $(\theta', t') \in \mathcal{B}(E' \cap F)$ and $i \in N$. Then, $[E' \cap F]_{t_i} \neq \emptyset$, so that there exists $(\theta'', t''_{-i}) \in \Theta \times T_{-i}$ such that $(\theta'', t'_i, t''_{-i}) \in E'$. Since E' is a product event, we can take $\theta'' = \theta'$ and $t''_{-i} = t'_{-i}$. Hence, $\mathcal{B}(E' \cap F) \subseteq E'$. $\qquad\square$

C.2 Proof of Proposition 5.2

We first show that the operator $\mathcal{B}_i$, $i \in N$, is downward continuous in the Harsanyi case.

Lemma C.1. *Suppose $\mathcal{T}$ is derived from a Harsanyi type space with the set of profiles given by $\Pi = \{\boldsymbol{F}^{\mathcal{H}}\}$. Suppose $A_1, A_2, \ldots$ is a decreasing sequence in $\mathcal{F}_\Theta \times \boldsymbol{F}^{\mathcal{H}}$, and fix $i \in N$. Then,*

$$\mathcal{B}_i\Big(\bigcap_{k=1}^\infty A_k\Big) = \bigcap_{k=1}^\infty \mathcal{B}_i(A_k).$$

That is, $\mathcal{B}_i$ is downward continuous.

Proof. It follows from Lemma B.4 that $\mathcal{B}_i\big(\bigcap_{k=1}^\infty A_k\big) \subseteq \bigcap_{k=1}^\infty \mathcal{B}_i(A_k)$. To show the reverse inclusion, note that for any $t_i \in T_i$, we have that $[\bigcap_{k=1}^\infty A_k]_{t_i} \in \mathcal{F}_\Theta \times \Sigma_i(t_i)$. Let $(\theta, t) \in \bigcap_{k=1}^\infty \mathcal{B}_i(A_k)$. Then,

$$\beta_i(t_i)\Big(\Big[\bigcap_{k=1}^\infty A_k\Big]_{t_i}\Big) = \lim_{k \to \infty} \beta_i(t_i)([A_k]_{t_i}) = 1,$$

and it follows that $(\theta, t) \in \mathcal{B}_i\big(\bigcap_{k=1}^\infty A_k\big)$. $\qquad\square$

Note that if $\mathcal{B}_i$ is downward continuous, then so is $f_F(\cdot) = \mathcal{B}_i(F \cap \cdot)$.

We are now ready to prove Proposition 5.2. It suffices to show that $C^{\mathrm{IT}}(F)$ is the greatest fixed point of f_F. For if f_F has a greatest fixed point $gfp(f_F)$, then

$$gfp(f_F) = \bigcup \{X : X = f_F(X)\} = C(F).$$

We use the following lemma, which is due to Halpern and Moses (1990).

Lemma C.2. *If f is a downward continuous mapping from the subsets of some space Y to the subsets of Y, then the greatest fixed point of f exists and is given by*

$$gfp(f) = \bigcap_{k=0}^\infty f^k(Y),$$

where $f^0(Y) = Y$, and $f^k(Y) = f(f^{k-1}(Y))$ for $k > 0$.

Hence, by Lemma C.2, the greatest fixed point of f_F is

$$
\begin{aligned}
gfp(f_F) &= (\Theta \times T) \cap \mathcal{B}(F) \cap \mathcal{B}(F \cap \mathcal{B}(F)) \cap \cdots \\
&= \mathcal{B}(F) \cap \mathcal{B}^2(F) \cap \mathcal{B}^3(F) \cap \cdots \\
&= C^{\mathrm{IT}}(F),
\end{aligned}
$$

where we have used the conjunction property, i.e., $\mathcal{B}(V \cap Z) = \mathcal{B}(V) \cap \mathcal{B}(Z)$ for $V, Z \in \mathcal{F}_\Theta \times \boldsymbol{F}^{\mathcal{H}}$.

$\square$

C.3   Proof of Lemma 6.1

We start with some preliminary results. The first result shows that profiles that are self-dominating or are part of a cycle or infinite chain have an infinite order.

Lemma C.3. *Suppose there exist profiles $\boldsymbol{F}^0, \boldsymbol{F}^1, \ldots \in \Pi$ (not necessarily distinct) such that*

$$\boldsymbol{F}^0 \succ \boldsymbol{F}^1 \succ \boldsymbol{F}^2 \succ \cdots ,$$

that is, the profiles are self-dominating, form a cycle, or form an infinite chain. Then, $\boldsymbol{F}^\ell \in \mathcal{O}^\infty$ for all $\ell \geq 0$.

Proof. We prove the result for $\ell = 0$; the proof for $\ell > 0$ is similar. Because $\boldsymbol{F}^1 \succ \boldsymbol{F}^2$, we have $\boldsymbol{F}^1 \notin \mathcal{O}^1$. By a similar argument, $\boldsymbol{F}^2 \notin \mathcal{O}^1$, so that $\boldsymbol{F}^1 \notin \mathcal{O}^2$. If we repeat this argument, we have that $\boldsymbol{F}^1 \in \Pi \setminus \bigcup_{m \leq k-1} \mathcal{O}^m$ for every $k = 1, 2, \ldots$, and the result follows. $\square$

The next result helps us to classify the profiles in Π.

Lemma C.4. *Let $\boldsymbol{F} \in \Pi$. Then:*

(a) *If for each $i \in N$, $\mathcal{F}_i \supseteq \sigma(h_i^{\mathcal{T},m})$ for all m, then $\boldsymbol{F} \in \mathcal{O}^\infty$.*

(b) *If there is $k = 0, 1, \ldots$ such that $\mathcal{F}_j = \sigma(h_j^{\mathcal{T},k})$ for all $j \in N$, and $\mathcal{F}_i \subsetneq \sigma(h_i^{\mathcal{T},k+1})$ for some $i \in N$, then $\boldsymbol{F} \in \mathcal{O}^{k+1}$.*

Proof. (a) First suppose that for each $i \in N$, we have $\mathcal{F}_i = \sigma(h_i^{\mathcal{T},0}) = \sigma(h_i^{\mathcal{T},1}) = \ldots$. Then, by Lemma A.3(c), the profile $\boldsymbol{F}$ is self-dominating. The result then follows from Lemma C.3. So suppose there is $\ell \in N$ such that $\mathcal{F}_\ell \neq \sigma(h_\ell^{\mathcal{T},0})$.

Again, if $\boldsymbol{F}$ dominates itself, then it follows from Lemma C.3 that $\boldsymbol{F} \in \mathcal{O}^\infty$. Hence, it remains to consider the case that $\boldsymbol{F}$ is not self-dominating. By Assumption 1, there is $\boldsymbol{F}' \in \Pi$ such that $\boldsymbol{F} \succ^* \boldsymbol{F}'$. By a similar argument as in the proof of Lemma A.3 (Claim 1), the profile $\boldsymbol{F}'$ is not self-dominating. Define $\mathcal{F}_j^0 := \mathcal{F}_j$ and $\mathcal{F}_j^1 := \mathcal{F}_j'$ for $j \in N$. Then there exist profiles $\boldsymbol{F}^2, \boldsymbol{F}^3, \ldots \in \Pi$ such that $\boldsymbol{F}^x$ is not self-dominating for any $x \geq 0$, and one of the following holds:

(i) the profiles form a cycle or infinite chain, that is,

$$\boldsymbol{F}^0 \succ^* \boldsymbol{F}^1 \succ^* \boldsymbol{F}^2 \succ^* \cdots ;$$

(ii) the profiles form a finite chain, that is, there is $m < \infty$ such that

$$\boldsymbol{F}^0 \succ^* \boldsymbol{F}^1 \succ^* \boldsymbol{F}^2 \succ^* \cdots \succ^* \boldsymbol{F}^m = \underset{j \in N}{\bigtimes} \{T_j, \emptyset\}.$$

If (i) is the case, then the result follows from Lemma C.3. We claim that (ii) does not hold. For suppose not. Then, by Lemma A.3(c), we have $\mathcal{F}_j = \mathcal{F}_j^0 = \sigma(h_j^{T,m})$ for $j \in N$, so that for each player $j \in N$ and $\ell \geq m$, $\sigma(h_j^{T,\ell}) = \sigma(h_j^{T,m})$. Using Lemma A.3(c) again, we see that $\boldsymbol{F}$ is self-dominating, a contradiction.

(b) The proof is by induction. Fix $\boldsymbol{F} \in \Pi$, and suppose that $\mathcal{F}_j = \sigma(h_j^{T,0})$ for all $j \in N$, and $\mathcal{F}_i \subsetneq \sigma(h_i^{T,1})$ for some i. Then $\boldsymbol{F} \not\succ \bigtimes_{j \in N}\{T_j, \emptyset\}$ (by Lemma A.3(c)), and it follows that there is no $\boldsymbol{F}' \in \Pi$ such that $\boldsymbol{F} \succ \boldsymbol{F}'$. Hence, $\boldsymbol{F} \in \mathcal{O}^1$.

For $k > 0$, suppose that for all $\ell \leq k - 1$ and $\boldsymbol{F} \in \Pi$, we have

$$\boldsymbol{F} \in \mathcal{O}^{\ell+1} \tag{C.1}$$

whenever $\mathcal{F}_j = \sigma(h_j^{T,\ell})$ for all $j \in N$, and $\mathcal{F}_i \subsetneq \sigma(h_i^{T,\ell+1})$ for some i.

Fix $\boldsymbol{F} \in \Pi$, and suppose that $\mathcal{F}_j = \sigma(h_j^{T,k})$ for all j, and $\mathcal{F}_i \subsetneq \sigma(h_i^{T,k+1})$ for some i. By Lemma A.3(e), the profile $\boldsymbol{F}$ is not self-dominating. We claim that there is a player $n \in N$ such that $\mathcal{F}_n \neq \{T_n, \emptyset\}$. For suppose not. Then, for all $j \in N$, we have $\mathcal{F}_j = \sigma(h_j^{T,0}) = \cdots = \sigma(h_j^{T,k})$. By Lemma A.3(c), the profile $\boldsymbol{F}$ is self-dominating (given that $k > 0$), a contradiction.

Hence, by Assumption 1, there is $\boldsymbol{F}' \in \Pi$ such that $\boldsymbol{F} \succ^* \boldsymbol{F}'$. By Lemma A.3(e), $\boldsymbol{F}'$ does not dominate itself. Again, as in the proof of part (a), we define $\mathcal{F}_j^0 := \mathcal{F}_j$ and $\mathcal{F}_j^1 := \mathcal{F}_j'$ for $j \in N$, and there exist profiles $\boldsymbol{F}^2, \boldsymbol{F}^3, \ldots \in \Pi$ such that $\boldsymbol{F}^x$ is not self-dominating for $x \geq 0$, and the profiles form a cycle or infinite chain (case (i) above), or a finite chain (case (ii) above).

We claim that the profiles do not form a cycle or infinite chain. Suppose not. Then, because $\mathcal{F}_j^1 \supseteq \sigma(h_j^{T,0})$ for $j \in N$, it follows from Lemma A.3(c) that $\mathcal{F}_j^0 \supseteq \sigma(h_j^{T,1})$ for $j \in N$. Similarly, $\mathcal{F}_j^1 \supseteq \sigma(h_j^{T,1})$ for $j \in N$. Applying Lemma A.3(c) repeatedly gives $\mathcal{F}_i = \mathcal{F}_i^0 \supseteq \sigma(h_i^{T,k+1})$, a contradiction.

Consequently, there is $m < \infty$ such that

$$\boldsymbol{F}^0 \succ^* \boldsymbol{F}^1 \succ^* \boldsymbol{F}^2 \succ^* \cdots \succ^* \boldsymbol{F}^m = \underset{j \in N}{\bigtimes} \{T_j, \emptyset\}.$$

We use the following claim:

Claim. For each $\ell = 0, 1, \ldots, m$, the following hold:

- $\mathcal{F}_j^{m-\ell} = \sigma(h_j^{T,\ell})$ for $j \in N$; and

- there is $i_\ell \in N$ such that $\mathcal{F}_{i_\ell}^{m-\ell} \subsetneq \sigma(h_{i_\ell}^{\mathcal{T},\ell+1})$.

Proof of Claim. The proof is by induction. By assumption, $\mathcal{F}_j^m = \sigma(h_j^{\mathcal{T},0})$ for all $j \in N$. We need to show that there is i_0 such that $\mathcal{F}_{i_0} \subsetneq \sigma(h_{i_0}^{\mathcal{T},1})$. Suppose not. Then $\boldsymbol{F}^m$ is self-dominating (by Lemma A.3(c)), which leads to a contradiction. For $\ell > 0$, suppose the result holds for $\ell - 1$, and consider $\boldsymbol{F}^{m-\ell}$. By Lemma A.3(c), we have $\mathcal{F}_j^{m-\ell} = \sigma(h_j^{\mathcal{T},\ell})$ for $j \in N$. Again, a simple proof by contradiction can be used to show that there is $i_\ell \in N$ such that $\mathcal{F}_{i_\ell}^{m-\ell} \subsetneq \sigma(h_{i_\ell}^{\mathcal{T},\ell+1})$. $\square$

It follows that $m = k$, and, by the induction hypothesis (C.1), we have $\boldsymbol{F}^1 \in \mathcal{O}^k$. Since $\boldsymbol{F} \succ \boldsymbol{F}^1$ by assumption, it remains to show that there is no $\widetilde{\boldsymbol{F}} \in \Pi \setminus \bigcup_{\ell \leq k} \mathcal{O}^\ell$ such that $\boldsymbol{F} \succ \widetilde{\boldsymbol{F}}$. Suppose by contradiction that there exists such a profile $\widetilde{\boldsymbol{F}}$. Then, by (C.1), there is no $\ell = 0, 1, \ldots, k-1$ such that $\widetilde{\mathcal{F}}_j = \sigma(h_j^{\mathcal{T},\ell})$ for $j \in N$, and $\widetilde{\mathcal{F}}_n \subsetneq \sigma(h_n^{\mathcal{T},\ell+1})$ for some player $n \in N$. By Corollary A.4, therefore, $\widetilde{\mathcal{F}}_j \supseteq \sigma(h_j^{\mathcal{T},k})$ for $j \in N$. But then it follows from Lemma A.3(c) that $\mathcal{F}_i \supseteq \sigma(h_i^{\mathcal{T},k+1})$, a contradiction. $\square$

We can now prove Lemma 6.1. A straightforward argument by contradiction establishes that $\mathcal{O}^\ell \cap \mathcal{O}^m = \emptyset$ whenever $\ell \neq m$. By definition, $\mathcal{O}^\infty \cup \bigcup_{m=1}^\infty \mathcal{O}^m \subseteq \Pi$. Hence, it remains to show the reverse inclusion, that is, for each $\boldsymbol{F} \in \Pi$, there is $k \in \mathbb{N} \cup \{\infty\}$ such that $\boldsymbol{F} \in \mathcal{O}^k$. This follows directly from Corollary A.4 and Lemma C.4. $\square$

C.4 Proof of Proposition 6.2

We start with an observation that follows directly from Corollary A.4 and Lemma C.4.

Corollary C.5. *For each $k = 1, 2, \ldots$, there is a nonempty subset N_k of N such that*

$$\mathcal{O}^k = \left\{ \bigtimes_{j \in N} \sigma(h_j^{\mathcal{T},k-1}) \in \Pi : \sigma(h_i^{\mathcal{T},k-1}) \subsetneq \sigma(h_i^{\mathcal{T},k}) \text{ iff } i \in N_k \right\}.$$

Moreover, we have that

$$\mathcal{O}^\infty = \left\{ \boldsymbol{F} \in \Pi : \mathcal{F}_j \supseteq \sigma(h_j^{\mathcal{T},j}) \text{ for all } m \right\}.$$

The next preliminary result implies that a type cannot be assigned an infinite rank, unless its σ-algebra is consistent with a profile of infinite order.

Lemma C.6. *Let $i \in N$ and $\boldsymbol{F}_{-i} \in \Pi_{-i}$. If there is no $\mathcal{F}_i \in \Pi_i$ such that $\mathcal{F}_i \times \boldsymbol{F}_{-i} \in \mathcal{O}^\infty$, then*

$$\max\{\ell = \infty, 1, 2, \ldots : \mathcal{F}_i \times \boldsymbol{F}_{-i} \in \mathcal{O}^\ell \text{ for some } \mathcal{F}_i\} < \infty$$

Proof. Suppose there is no $\mathcal{F}_i \in \Pi_i$ such that $\mathcal{F}_i \times \boldsymbol{F}_{-i} \in \mathcal{O}^\infty$. Suppose by contradiction that $\max\{\ell = \infty, 1, 2, \dots : \mathcal{F}_i \times \boldsymbol{F}_{-i} \in \mathcal{O}^\ell \text{ for some } \mathcal{F}_i\} = \infty$. Then, by Assumption 1 and Corollary C.5, there exist $k_1 < k_2 < \infty$ and $\mathcal{F}_i^1, \mathcal{F}_i^2 \in \Pi_i$ such that:

- $\{T_i, \emptyset\} \subsetneq \mathcal{F}_i^1 \subsetneq \mathcal{F}_i^2$; and

- $\mathcal{F}_i^\ell$ is the coarsest σ-algebra that dominates $\{T_i, \emptyset\} \times \boldsymbol{F}_{-i}$, $\ell = 1, 2$. (The choice of σ-algebra on T_i is immaterial here.)

But this is of course a contradiction. $\qquad\square$

We can now prove Proposition 6.2. Fix $i \in N$ and $t_i \in T_i$. It will be convenient to write $\boldsymbol{F}_{-i} = \bigtimes_{j \neq i} \mathcal{F}_j$ for $\Sigma_i(t_i)$. Recall that $d_i(t_i) = \infty$ if and only if for each $j \neq i$, $\mathcal{F}_j \supseteq \sigma(h_j^{\mathcal{T},m})$ for all m. Also, $d_i(t_i) = k < \infty$ if and only if for each $j \neq i$, $\mathcal{F}_j = \sigma(h_j^{\mathcal{T},k-1})$, and $\mathcal{F}_\ell \subsetneq \sigma(h_j^{\mathcal{T},k})$ for some $\ell \neq i$.

First suppose that for all $j \neq i$, $\mathcal{F}_j \supseteq \sigma(h_j^{\mathcal{T},m})$ for all m. If there is $\mathcal{F}_i \in \Pi_i$ such that $\mathcal{F}_i \supseteq \sigma(h_i^{\mathcal{T},m})$ for all m and $\mathcal{F}_i \times \boldsymbol{F}_{-i} \in \Pi$, then it follows from Corollary C.5 that $r_i(t_i) = \infty = d_i(t_i)$. Otherwise, there exists $k < \infty$ such that $\mathcal{F}_j = \sigma(h_j^{\mathcal{T},k-1}) = \sigma(h_j^{\mathcal{T},k}) = \dots$ for all $j \neq i$ (by Corollary A.4); and by Lemmas 6.1 and C.6, there is $\mathcal{F}_i^m \in \Pi_i$ such that $\mathcal{F}_i^m \times \boldsymbol{F}_{-i} \in \mathcal{O}^m$, and there is no $\mathcal{F}_i^\ell \in \Pi_i$ such that $\mathcal{F}_i^\ell \times \boldsymbol{F}_{-i} \in \mathcal{O}^\ell$ for $\ell > m$. In that case, $r_i(t_i) = m < \infty = d_i(t_i)$.

Next suppose that there is $k < \infty$ such that for each $j \neq i$, $\mathcal{F}_j = \sigma(h_j^{\mathcal{T},k-1})$ and $\mathcal{F}_\ell \subsetneq \sigma(h_\ell^{\mathcal{T},k})$ for some $\ell \neq i$. If there is $\mathcal{F}_i^k \in \Pi_i$ such that $\mathcal{F}_i^k \times \boldsymbol{F}_{-i} \in \mathcal{O}^k$, then $r_i(t_i) = k = d_i(t_i)$ (by Corollary C.5). Otherwise, by Lemma 6.1 and Corollary C.5, there is $\mathcal{F}_i \in \Pi_i$ such that $\mathcal{F}_i \times \boldsymbol{F}_{-i} \in \mathcal{O}^m$ for some $m \leq k - 1$. We claim that $m = k - 1$. To see this, note that because $\sigma(h_\ell^{\mathcal{T},k-1}) \subsetneq \sigma(h_\ell^{\mathcal{T},k})$ by assumption, we have that $\sigma(h_i^{\mathcal{T},k-2}) \subsetneq \sigma(h_i^{\mathcal{T},k-1})$ by Lemma A.3 (given that $\sigma(h_j^{\mathcal{T},k-2}) = \sigma(h_j^{\mathcal{T},k-1})$ for $j \neq i$ by assumption). But then, by a similar argument, there is $j \neq i$ such that $\sigma(h_j^{\mathcal{T},k-3}) \subsetneq \sigma(h_j^{\mathcal{T},k-2})$. Hence, $m \geq k - 1$, and thus $m = k - 1$. It follows that $r_i(t_i) = k - 1 < k = d_i(t_i)$. $\qquad\square$

C.5 Proof of Proposition 6.3

Fix $i \in N$ and $t_i \in T_i$. Again, it will be convenient to write $\boldsymbol{F}_{-i} = \bigtimes_{j \neq i} \mathcal{F}_j$ for $\Sigma_i(t_i)$. First suppose that for each $j \neq i$, $\mathcal{F}_j \supseteq \sigma(h_j^{\mathcal{T},m})$ for all m. By Assumption 2, there is $\mathcal{F}_i \in \Pi_i$ such that $\mathcal{F}_i \supseteq \sigma(h_i^{\mathcal{T},m})$ for all m, and $\mathcal{F}_i \times \boldsymbol{F}_{-i} \in \Pi$. By Corollary C.5, we then have that $r_i(t_i) = \infty = d_i(t_i)$.

Next suppose there is $k = 1, 2, \dots$ such that for each $j \neq i$, $\mathcal{F}_j = \sigma(h_j^{\mathcal{T},k-1}) \subsetneq \sigma(h_j^{\mathcal{T},k})$. Then, by Assumption 2, $\sigma(h_i^{\mathcal{T},k-1}) \times \boldsymbol{F}_{-i} \in \Pi$, and it follows from Corollary C.5 that $r_i(t_i) =$

$k = d_i(t_i).$ $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad$ $\square$

C.6 Proof of Proposition 6.4

We first show that if a type space satisfies Assumption 3, then it satisfies Assumption 2. Suppose $\mathcal{T}$ satisfies Assumption 3, and suppose by contradiction that it violates Assumption 2. Then, by Corollary A.4, there exist a profile $\boldsymbol{F} \in \Pi$ and $k < \infty$ such that $\mathcal{F}_j = \sigma(h_j^{\mathcal{T},k-1})$ for all $j \in N$, $\sigma(h_i^{\mathcal{T},k-1}) = \sigma(h_i^{\mathcal{T},k})$ for some $i \in N$, and $\sigma(h_\ell^{\mathcal{T},k-1}) \subsetneq \sigma(h_\ell^{\mathcal{T},k})$ for some $\ell \in N$. Then, by a similar argument as in the proof of Lemma C.4, $\mathcal{F}$ does not dominate itself, is not part of a cycle or an infinite chain. But $\mathcal{F}_i$ i-dominates $\boldsymbol{F}$, contradicting that $\mathcal{T}$ satisfies Assumption 3.

Next, suppose that $\mathcal{T}$ satisfies Assumption 2. Suppose there exist a profile $\boldsymbol{F} \in \Pi$ and a player $i \in N$ such that $\mathcal{F}_i$ i-dominates $\boldsymbol{F}$. Assume by contradiction that $\boldsymbol{F}$ is not self-dominating, and is not part of a cycle or infinite chain.

It follows from Corollary A.4 that there is $k < \infty$ such that $\mathcal{F}_j = \sigma(h_j^{\mathcal{T},k-1})$ for all $j \in N$ and $\mathcal{F}_\ell \subsetneq \sigma(h_\ell^{\mathcal{T},k})$ for some $\ell \in N$. Hence, by Assumption 2, we have that $\mathcal{F}_j \subsetneq \sigma(h_j^{\mathcal{T},k})$ for all $j \in N$. Because $\mathcal{F}_i$ i-dominates $\boldsymbol{F}$, we have that $k > 1$, $\mathcal{F}_j = \sigma(h_j^{\mathcal{T},k-2})$ for all $j \neq i$, and $\mathcal{F}_i \supsetneq \sigma(h_i^{\mathcal{T},k-2})$ (by Lemma A.3(c)).

By Corollary C.5, we have that $\boldsymbol{F} \in \mathcal{O}^k$, and, by definition, $\mathcal{O}^{k-1} \neq \emptyset$. By Corollary C.5, $\mathcal{O}^{k-1} = \{\bigtimes_{j \in N} \sigma(h_j^{\mathcal{T},k-2})\}$. But then, because $\sigma(h_j^{\mathcal{T},k-2}) = \sigma(h_j^{\mathcal{T},k-1})$ for $j \neq i$ and $\mathcal{F}_i \supsetneq \sigma(h_i^{\mathcal{T},k-2})$, $\mathcal{T}$ does not satisfy Assumption 2, a contradiction. $\qquad\square$

References

Ahn, D. S. (2007). Hierarchies of ambiguous beliefs. *Journal of Economic Theory 136*, 286–301.

Aumann, R. J. (1976). Agreeing to disagree. *Annals of Statistics 4*(6), 1236–1239.

Barwise, J. (1988). Three views of common knowledge. In *TARK '88: Proceedings of the 2nd Conference on Theoretical Aspects of Reasoning about Knowledge*, San Francisco, CA, pp. 365–379. Morgan Kaufmann Publishers Inc.

Battigalli, P. and M. Siniscalchi (2002). Strong belief and forward-induction reasoning. *Journal of Economic Theory 106*, 356–391.

Brandenburger, A. and E. Dekel (1987). Common knowledge with probability 1. *Journal of Mathematical Economics 16*, 237–245.

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.

Brandenburger, A., A. Friedenberg, and H. J. Keisler (2012). Fixed points in epistemic game theory. In *Proceedings of the 2008 Clifford Lectures*, AMS Proceedings of Symposia in Applied Mathematics. Forthcoming.

Brandenburger, A. and H. J. Keisler (2006). An impossibility theorem on beliefs in games. *Studia Logica 84*, 211–240.

Carlsson, H. and E. van Damme (1993). Global games and equilibrium selection. *Econometrica 61*, 989–1018.

Chwe, M. S.-Y. (2001). *Rational Ritual: Culture, Coordination, and Common Knowledge*. Princeton, NJ: Princeton University Press.

Ciancaruso, D. and F. Germano (2011). Quotient spaces of boundedly rational types. Working paper.

Clark, H. H. and C. Marshall (1981). Definite reference and mutual knowledge. In A. K. Joshi, B. L. Webber, and I. A. Sag (Eds.), *Elements of Discourse Understanding*, Chapter 1. Cambridge, UK: Cambridge University Press.

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., M. A. Costa-Gomes, and N. Iriberri (2012). Structural models of nonequilibrium strategic thinking: Theory, evidence, and applications. *Journal of Economic Literature*. Forthcoming.

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.

Davis, H. L. and C. Pratt (1995). The development of children's theory of mind: The working memory explanation. *Australian Journal of Psychology 47*, 25–31.

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.

Ely, J. and M. Peski (2006). Hierarchies of beliefs and interim rationalizability. *Theoretical Economics 1*, 19–65.

Fagin, R. and J. Y. Halpern (1988). Belief, awareness and limited reasoning. *Artificial Intelligence 34*, 39–76.

Friedenberg, A. and H. J. Keisler (2011). Iterated dominance revisited. Working paper, Arizona State University.

Friedenberg, A. and M. Meier (2012). On the relationship between hierarchy and type morphisms. *Economic Theory 46*, 377–399.

Geanakoplos, J. D. and H. M. Polemarchakis (1982). We can't disagree forever. *Journal of Economic Theory 28*, 192–200.

Halpern, J. Y. and Y. Moses (1990). Knowledge and common knowledge in a distributed environment. *Journal of the ACM 37*, 549–587.

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. (1999). Iterative and fixed point common belief. *Journal of Philosophical Logic 28*, 61–79.

Heifetz, A. and W. Kets (2011). All types naive and canny. Working paper.

Heifetz, A. and D. Samet (1998). Topology-free typology of beliefs. *Journal of Economic Theory 82*, 324–341.

Heinsalu, S. (2011). Universal type spaces with unawareness. Working paper, Yale University.

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.

Jehiel, P. and F. Koessler (2008). Revisiting games of incomplete information with analogy-based expectations. *Games and Economic Behavior 62*, 533–557.

Kets, W. (2012). Bounded reasoning: A canonical type space. Working paper.

Kinderman, P., R. Dunbar, and R. P. Bentall (1998). Theory-of-Mind deficits and casual attributions. *British Journal of Psychology 89*, 191–204.

Lewis, D. (1969). *Convention: A Philosophical Study.* Cambridge, MA: Harvard University Press.

Lismont, L. and P. Mongin (1995). Belief closure: A semantics of common knowledge for modal propositional logic. *Mathematical Social Sciences 30*, 127–153.

Mertens, J. F. and S. Zamir (1985). Formulation of Bayesian analysis for games with incomplete information. *International Journal of Game Theory 14*, 1–29.

Milgrom, P. (1981). An axiomatic characterization of common knowledge. *Econometrica 49*, 219–222.

Monderer, D. and D. Samet (1989). Approximating common knowledge with common beliefs. *Games and Economic Behavior 1*, 170–190.

Nagel, R. (1995). Unraveling in guessing games: An experimental study. *American Economic Review 85*, 1313–1326.

Pintér, M. and Z. Udvari (2011). Generalized type spaces. Working paper 34107, MPRA.

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

Samet, D. (1990). Ignoring ignorance and agreeing to disagree. *Journal of Economic 52*, 190–207.

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.

Strzalecki, T. (2009). Depth of reasoning and higher order beliefs. Working paper, Harvard University.