

# Hierarchies of belief and interim rationalizability

JEFFREY C. ELY

Department of Economics, Northwestern University

MARCIN PEŃSKI

Department of Economics, University of Chicago

In games with incomplete information, conventional hierarchies of belief are incomplete as descriptions of the players' information for the purposes of determining a player's behavior. We show by example that this is true for a variety of solution concepts. We then investigate what is essential about a player's information to identify behavior. We specialize to two player games and the solution concept of interim rationalizability. We construct the universal type space for rationalizability and characterize the types in terms of their beliefs. Infinite hierarchies of beliefs *over conditional beliefs*, which we call  $\Delta$ -hierarchies, are what turn out to matter. We show that any two types in any two type spaces have the same rationalizable sets in all games if and only if they have the same  $\Delta$ -hierarchies.

KEYWORDS. Interim rationalizability, belief hierarchies.

JEL CLASSIFICATION. C72, D82

## 1. INTRODUCTION

Games with incomplete information are indispensable tools in modern economic analysis. John Harsanyi, in a series of papers (Harsanyi 1967–68), introduced the Bayesian game framework which is now the nearly universally adopted model of strategic behavior under incomplete information. Harsanyi himself observed that prior to his innovation, the basic theory of incomplete information was almost non-existent despite the obvious wealth of potential applications. The problem seemed to arise at the most fundamental level: how to formalize a player's beliefs and higher-order beliefs in a manageable way.

“It seems to me that the basic reason why the theory of games with incomplete information has made so little progress so far lies in the fact that these games give rise, or at least appear to give rise, to an infinite regress in reciprocal expectation on the part of the players .... In such a game player 1's

---

Jeffrey C. Ely: [jeffely@northwestern.edu](mailto:jeffely@northwestern.edu).

Marcin Peński: [mpeski@uchicago.edu](mailto:mpeski@uchicago.edu).

Financial support from NSF grant #9985462 and the Alfred P. Sloan foundation is gratefully acknowledged. We thank Bart Lipman (in his role as colleague), Aviad Heifetz, Phil Reny, Bill Sandholm, Asher Wolinsky, Itai Sher and Marciano Siniscalchi for valuable comments. Also, the suggestions of Bart Lipman (in his role as editor) and two anonymous referees helped to significantly improve the exposition of this paper.

Copyright © 2006 Jeffrey C. Ely and Marcin Peński. Licensed under the [Creative Commons Attribution-NonCommercial License 2.5](https://creativecommons.org/licenses/by-nc/2.5/). Available at <http://econtheory.org>.

strategy choice will depend on what he expects (or believes) to be player 2's payoff function  $U_2$ , as the latter will be an important determinant of player 2's behavior in the game. ... But his strategy choice will also depend on what he expects to be player 2's first-order expectation about his own payoff function  $U_1$ . Indeed player 1's strategy choice will also depend on what he expects to be player 2's second-order expectation—that is, on what player 1 thinks that player 2 thinks that player 1 thinks about player 2's payoff function  $U_2$ . ... and so on *ad infinitum*."

To completely describe an incomplete information environment, one must specify the players' *infinite hierarchies of belief*, and this appeared intractable. Harsanyi's solution was based on the compact model of information that is now called a *type space*. Suppose the players are uncertain about which events in some set  $\Omega$  of *states of the world* hold. Typically  $\Omega$  will represent the possible payoff functions in the game. In a type space over  $\Omega$ , all strategically relevant aspects of a player's information about  $\Omega$  are encapsulated in a single variable, referred to as the player's type. Each player has a set of possible types  $T_i$ , and for each type  $t_i$  there is specified a belief  $\mu_i(t_i)$  about the underlying payoffs and the types of the other players in the game. This structure is quite simple formally, and yet within this simple model are embedded the complex hierarchies of belief that seemed to produce infinite regress. Here is a simple example to illustrate.

### 1.1 Example

In this example, there are two players and two possible payoff-relevant states of the world,  $\Omega = \{-1, +1\}$ . Each player  $i$  has two possible types,  $T_i = \{-1, +1\}$ , and the state of the world together with the players' types are drawn from a common prior<sup>1</sup> distribution  $\mu \in \Delta(T_1 \times T_2 \times \Omega)$  given by

$$\mu(t_i, t_{-i}, \omega) = \begin{cases} \frac{1}{4} & \text{if } \omega = t_i \cdot t_{-i} \\ 0 & \text{otherwise.} \end{cases}$$

The belief  $\mu_i(t_i) \in \Delta(\Omega \times T_{-i})$  is then derived by updating  $\mu$  in a Bayesian fashion, conditional on the realized type  $t_i$ .

In this type space, each player assigns equal probability to each payoff-relevant event. These are the first-order beliefs. Since player  $i$  holds this first-order belief regardless of whether  $i$  is type is  $-1$  or  $+1$ , and since player  $-i$  assigns probability 1 to player  $i$  having one of these two types, it follows that each player is certain of the others' first-order beliefs. These are the second-order beliefs. The same reasoning implies that each player is certain of the other's second-order beliefs and so on. Indeed, in this simple type space it is always *common knowledge* that the two states are equally likely. We can see that Harsanyi's Bayesian game model is a parsimonious way to formalize the complex hierarchies of beliefs that previously seemed to be intractable. It has also

<sup>1</sup>The common prior model was introduced by Harsanyi as a special case of his model. To sharpen our point this example has a common prior (as this is typical in economic applications) but in our analysis we consider type spaces of the most general form.

proved straightforward to apply using versions of traditional complete-information solution concepts.

### 1.2 Foundations of the Harsanyi framework

In the Harsanyi framework, type spaces are a convenient modeling device used to describe the players' beliefs and higher-order beliefs. One potential concern with the use of a type space is the following. If hierarchies of belief are what really matter, then we must be assured that any hierarchies we might wish to model can be captured in some type space. This concern has been resolved by [Mertens and Zamir \(1985\)](#) and [Brandenburger and Dekel \(1993\)](#) who showed that when the set of states of the world  $\Omega$  has some minimal structure, then any internally consistent ("coherent") hierarchy can be modeled using a type space. In fact, there exists a single *universal* type space  $U(\Omega)$  that simultaneously captures them all: for every coherent hierarchy there is a type in  $U(\Omega)$  with that hierarchy.<sup>2</sup> Thus, the Harsanyi framework is sufficiently general to model any incomplete information scenario.

Another potential concern has not received the same attention. The type space we used in our example is but one of many that would capture those specific hierarchies of belief. Indeed, any specification of the players' hierarchies can be equally well generated by many different type spaces. If hierarchies are what matter, and if type spaces are simply a convenient device used to model them, then we should be assured that the *outcomes* we predict for a given hierarchy should not depend on the particular type space used to model it. However, the type space can matter for outcomes, as can be seen in the following game.

Consider the following two player game of incomplete information. There are two states of the world  $\Omega = \{-1, +1\}$ . Each player  $i$  has three actions  $A_i = \{a_i, b_i, c_i\}$  and a payoff  $u_i$  that depends on the actions chosen by each player and the state of the world. The payoffs are summarized in [Figure 1](#).

|               |          |          |        |               |          |          |        |
|---------------|----------|----------|--------|---------------|----------|----------|--------|
|               | $a_2$    | $b_2$    | $c_2$  |               | $a_2$    | $b_2$    | $c_2$  |
| $a_1$         | 1, 1     | -10, -10 | -10, 0 | $a_1$         | -10, -10 | 1, 1     | -10, 0 |
| $b_1$         | -10, -10 | 1, 1     | -10, 0 | $b_1$         | 1, 1     | -10, -10 | -10, 0 |
| $c_1$         | 0, -10   | 0, -10   | 0, 0   | $c_1$         | 0, -10   | 0, -10   | 0, 0   |
| $\omega = +1$ |          |          |        | $\omega = -1$ |          |          |        |

FIGURE 1. A game with incomplete information.

Focusing only on the subsets  $\{a_i, b_i\}$ , we have a common interest game in which the players wish to choose the same action in the positive state and the opposite action in the negative state. Failing to coordinate is costly, and the action  $c_i$  is a "safe" alternative when, conditional on the state,  $i$  is uncertain of his opponent's action.

<sup>2</sup>On the other hand, when the set of states lacks the topological structure assumed by these authors, [Heifetz and Samet \(1999\)](#) showed that the type space framework may not be sufficiently general to model all coherent hierarchies. See also [Meier \(2005\)](#)

Within the Harsanyi framework, once we have described the set of states of the world, the actions in the game, and the payoffs, we complete the description by specifying the players' information, i.e. their hierarchies of belief about payoffs, and then finding a suitable type space to model them. Let us suppose that it is common knowledge among the players that the two states of the world are equally likely. The type space we have already introduced is one way to capture this assumption. We can now apply standard solution concepts to the resulting Bayesian game. It is a Bayesian Nash equilibrium for the players to achieve perfect coordination where types  $t_i = +1$  play  $a_i$  and types  $t_i = -1$  play  $b_i$ . Symmetrically, there is another equilibrium where  $t_i = +1$  play  $b_i$  and  $t_i = -1$  play  $a_i$ . Obviously it is also an equilibrium for both to play  $c_i$  independent of type. It follows that all actions are interim rationalizable for every type.

In the type space under consideration, each player has the same hierarchy of beliefs regardless of his type. It appears that there is a spurious duplication of types. So instead consider the simpler type space in which each player has exactly one type and this type knows the other player's type and assigns equal probability to the two states of the world. Formally,  $T_i^* = \{*\}$  and there is a common prior  $\mu^*$  given by  $\mu^*(*, *, +1) = \mu^*(*, *, -1)$ . This type space generates exactly the same hierarchies of belief as in our first example: common knowledge that the states are equally likely. However, when the game in our example is played over this type space, the unique Bayesian Nash equilibrium, the unique correlated equilibrium, indeed the unique rationalizable outcome, is for both players to play  $c_i$ .<sup>3</sup>

We cannot be assured that our predictions are invariant to the choice of the type space. Indeed, specifying the hierarchies was not enough to complete the description of the environment as it is *not* only hierarchies that matter for (correlated) equilibrium and rationalizability. While the additional types in the original type space are duplicates in terms of their hierarchies, they are not redundant because they generate a payoff-relevant means of correlating behavior *with the state of the world*.

This observation has a significance for the philosophical debate (see [Aumann 1987](#), [Brandenburger and Dekel 1993](#), [Gul 1998](#), and [Aumann 1998](#)) about whether or not the information structure in a game is common knowledge. The universal type space has been interpreted as precisely that information structure that can be assumed without loss of generality to be common knowledge. For example, [Brandenburger and Dekel \(1993\)](#) suggest that the universal type space realizes Aumann's hypothesis of a completely specified "state space". This is certainly true if, as in [Brandenburger and Dekel \(1993\)](#), one considers the information structure purely as a model of beliefs (about beliefs) about uncertain events. But if what is important is the range of possible behaviors

<sup>3</sup>It deserves emphasis that the issue we are pointing to here is distinct from the familiar one that adding redundant types to an information structure creates the possibility that the players can correlate their action choices and thus increases the set of equilibrium outcomes. That observation is equivalent to the statement that the set of correlated equilibria of a game is larger than the set of Nash equilibria of a game. To see that something different is happening in our example, note that the sets of correlated equilibria in the two games are distinct. Adding redundant types in order to generate correlation in play can never affect the set of correlated equilibria (see [Brandenburger and Dekel 1987](#)). Indeed, it can never affect the set of rationalizable outcomes as it does here.

in a game and not just beliefs, then our example shows that there is a loss of generality in assuming that the universal type space is commonly known. In particular, this assumption would imply that whenever the players commonly know that each state in the example is equally likely, they must play action  $c$ .<sup>4</sup>

One of the themes of the debate concerns the appropriateness of different ways of modeling interactive beliefs. Gul (1998) claims that hierarchies of beliefs are more appropriate if one takes Savage's (1954) "personalistic" view. Aumann (1998) replies that the hierarchy and type space models are equivalent and could be used interchangeably depending on the convenience in a particular application.<sup>5</sup> Our observations echo the claims made in the recent literature (see Battigalli and Siniscalchi (2003, Section 6) and Bergemann and Morris (2005, Section 2.5)) that, in terms of the behavior that can be modeled, Harsanyi type spaces are more general. Our objective is to understand how the hierarchies model must be augmented in order to re-establish this equivalence.

### 1.3 $\Delta$ -hierarchies

As argued above, from the point of view of rationalizability, conventional hierarchies of belief are incomplete as descriptions of a player's information. The goal of this paper is to identify a characteristic of a player's information that is sufficient and necessary for determining what is rationalizable in any game. Ideally, we would like this characteristic to be expressible in a language common to the players and the modeler. In particular, just as with conventional hierarchies of belief it should be described in a way that does not refer to any particular type space, the symbols of which are not assumed to have any meaning to the players.<sup>6</sup>

For the solution concept of rationalizability, we are able to identify a characteristic of information, which we call the  $\Delta$ -hierarchy, that is both necessary and sufficient for predicting rationalizable behavior in two player games.  $\Delta$ -hierarchies are infinite hierarchies of beliefs *about conditional beliefs*. Perhaps the easiest way to understand  $\Delta$ -hierarchies is to see how they can be extracted from a type space.

If player  $i$  could learn his opponent's information,  $i$  would obtain some conditional

<sup>4</sup>In Section 7 we present an example that makes an even stronger point in this regard. There we describe a game with an action that is not rationalizable for any type when it is assumed that the universal type space is common knowledge. Nevertheless, this action could be played as a part of Bayesian Nash equilibrium in a very simple, completely standard type space where there is common knowledge of rationality and common knowledge that the players' beliefs are coherent.

<sup>5</sup>Aumann (1998) refers to partition model, which for our purposes is equivalent to Harsanyi's type space.

<sup>6</sup>It is tempting to suggest that the solution is to describe types by their hierarchies of belief over *all* events in  $\Omega \times T$ , not just those that are payoff-relevant. Indeed, the type  $*$  and the types  $\{-1, +1\}$  can be distinguished by their beliefs over states *and* types of the opponent. However, this approach fails for two reasons. First, the payoff-irrelevant events are type-space specific, so it would be impossible to describe these beliefs in a type-space independent way. Second, distinguishing types by their beliefs over all events only pushes the problem to the other extreme: types that are truly equivalent would not be treated so. For example, the two types  $-1$  and  $+1$  have the same rationalizable behavior in every game (this is a consequence of our main result) and yet they have distinct first-order beliefs over the types of the opponent and  $\Omega$ . Even worse, *any* type space can be duplicated by re-labeling the types or expanded by adding additional irrelevant types generating a new set of strategically equivalent types whose hierarchies over  $\Omega \times T$  are nevertheless distinct.

belief about the state of the world. So, before knowing the opponent's information,  $i$  has a prior belief over the many different conditional beliefs he *could* obtain, were he to learn it. The first-order belief is this probability distribution over possible conditional beliefs. Within a type space, the first-order belief is derived for a type  $t_i$  as follows. First, determine for each type  $t_{-i}$  of the opponent what would be the conditional belief  $\beta(t_i, t_{-i}) \in \Delta\Omega$  of  $t_i$  if the opponent's type were known. Then, the probability of any set  $Y \subset \Delta\Omega$  of possible conditional beliefs is the probability  $t_i$  assigns to the set of types  $t_{-i}$  for which  $\beta(t_i, t_{-i})$  belongs to  $Y$ . Once we have derived first-order beliefs of this form for every type, we can in the usual way derive the second-order beliefs: the probability any type  $t_i$  assigns to the events consisting of first-order beliefs of the opponent and the conditional beliefs of  $t_i$ . Higher-order beliefs are defined analogously.

Let us see how the types in our type spaces are distinguished by their  $\Delta$ -hierarchies. In the first type space, conditional on learning the opponent's type, each type would learn the state with certainty. Because the two types of the opponent have equal probability, the first-order belief of each type is an equal mixture over Dirac measures. With probability  $\frac{1}{2}$  the conditional belief assigns probability 1 to state  $\omega = +1$  and with probability  $\frac{1}{2}$  the conditional belief assigns probability 1 to state  $\omega = -1$ . Since all types have this first-order belief, the second-order and higher-order beliefs are again degenerate: there is common-knowledge of these first-order beliefs. In the second type space, each type's first-order belief is instead a Dirac measure on an equal mixture. Here, each type already knows the opponent's type, so there is nothing new to learn. If  $i$  were to learn the opponent's type, then with probability 1 the conditional belief would assign equal probability to the two states. Hence, the  $\Delta$ -hierarchy of type  $*$  is equal to common knowledge of a Dirac measure on an equal mixture over both states. On the other hand, not all types have distinct  $\Delta$ -hierarchies. For example the derivation above shows that the  $\Delta$ -hierarchies of types  $+1$  and  $-1$  are the same.

#### 1.4 Overview of results

To recap, a  $\Delta$ -hierarchy consists of a probability distribution over conditional beliefs (the first-order belief) in  $\Delta\Omega$ , a joint probability distribution over the opponent's first-order belief and own conditional beliefs (the second-order belief), etc. Just as with conventional hierarchies of belief, while they are implicitly captured within a type space, their explicit description uses only the natural language of probabilities and conditional probabilities. In particular the description does not refer to any particular type space. Furthermore, by an extension of the results of [Mertens and Zamir \(1985\)](#) and [Brandenburger and Dekel \(1987\)](#) (our [Theorem 1](#)), all  $\Delta$ -hierarchies can be collected into a single, "universal" type space  $U(\Delta\Omega)$  over  $\Delta\Omega$ .

The main result of this paper, [Theorem 2](#), shows that  $U(\Delta\Omega)$  is, in a sense, the universal type space for rationalizability. Precisely, we divide the statement into three parts:

1. *Sufficiency*: Two types with the same  $\Delta$ -hierarchy have the same rationalizable behavior in every game.
2. *Necessity*: For any two types with different  $\Delta$ -hierarchies there is a game in which

both types have different rationalizable sets.

3. *Non-Redundancy*: For each  $\Delta$ -hierarchy there exists a type space that can be used to represent it.

To show non-redundancy, we construct a single type space over  $\Omega$ , called  $L(\Omega)$  (Section 4.1) where the set of types of player  $i$  is equal to  $L_i(\Omega) = U_i(\Delta\Omega) \times \Delta\Omega$ . Here each type in  $L_i(\Omega)$  has a label  $(u_i^\Delta, \tau_i)$ , where  $u_i^\Delta \in U_i(\Delta\Omega)$  is a  $\Delta$ -hierarchy and  $\tau \in \Delta\Omega$  is a useful “dummy” variable. We show that every  $\Delta$ -hierarchy is represented in  $L_i(\Omega)$ . Indeed, the  $\Delta$ -hierarchy of type  $(u_i^\Delta, \tau_i)$  is equal to  $u_i^\Delta$ .

In order to prove sufficiency, we study type mappings: mappings that associate the types in two separate type spaces. We show in Section 4.2 that a type mapping preserves rationalizable behavior if it preserves *conditional* beliefs.<sup>7</sup> Consider the conditional belief about  $\Omega$  of type  $t_i$  in the source type space conditional on a given type  $t_{-i}$  of the opponent. In a type mapping that preserves conditional beliefs, these should be equal to the belief of the image of  $t_i$  in the target space, conditional on the image of  $t_{-i}$ . We show sufficiency by focusing on the type mapping that sends types in  $T$  to types in  $L(\Omega)$  via  $t_i \rightarrow (u_i^\Delta(t_i), \tau^*)$ , where  $u_i^\Delta(t_i)$  is the  $\Delta$ -hierarchy of  $t_i$  and  $\tau^*$  is some fixed dummy. We show that this mapping preserves conditional beliefs, and therefore preserves rationalizable behavior. Since two types with the same  $\Delta$ -hierarchy are mapped to the same type under this mapping, it follows that their rationalizable behavior is the same as that of this (common) image type.

Finally, to show necessity, we pursue a parallel construction (Section 6.1). Any type in any type space can be interpreted as a *rule* that associates each game form with the set of actions that are rationalizable for that type. We construct a space of all rules  $\mathcal{R}$  by collecting every rule associated with any type in any type space. We show that  $\mathcal{R}$  is naturally seen as a proper type space with  $\Delta\Omega$  as the space of basic uncertainty. A key step is to characterize the beliefs of type-rules, which is the main subject of Section 6.2.

We show that  $\mathcal{R}$  and  $U(\Delta\Omega)$  are essentially the same type space. The mapping that associates each type in  $\mathcal{R}$  with its hierarchy of beliefs is an isomorphism of type spaces,  $\iota : U(\Delta\Omega) \rightarrow \mathcal{R}$  and  $\iota^{-1} : \mathcal{R} \rightarrow U(\Delta\Omega)$  (Theorem 4). The mapping has the property that for any type in any type space, the rule of this type can be obtained by first computing its  $\Delta$ -hierarchy and then applying the mapping  $\iota$ . Since by definition two distinct rules in  $\mathcal{R}$  are distinguished by some game, it follows that two distinct  $\Delta$ -hierarchies can be distinguished by some game.

Along the way, we present some new results on the structure of the rationalizability correspondence. These are presented in Section 5. Finally, Section 7 comments and contains further examples.

<sup>7</sup>Mertens and Zamir (1985) show that *belief-preserving* mappings preserve conventional hierarchies of beliefs. Under these mappings the conditional belief of the image type, conditional on type of the opponent  $t_{-i}$ , must be equal to the conditional belief of the source type, conditional on *all* source types of the opponent that are mapped into  $t_{-i}$ . Belief-preserving type mappings do not preserve rationalizable behavior. See Section 3.1.

### 1.5 Related literature

An example similar to ours was discovered independently by Dekel et al. (2005a).<sup>8</sup> They introduce a new version of rationalizability in which players can conjecture correlations between the opponent's action and the state beyond those correlations that are explicitly modeled in the type space. Conventional hierarchies are sufficient to identify the sets that are rationalizable under this alternative definition. In the context of our example, all actions satisfy their definition regardless of the type space. Indeed, there is no type in any type space for which action  $c_i$  is the only rationalizable action under the alternative concept. Liu (2005) develops a framework for explicitly describing the correlations that distinguish any type space from its reduced-form counterpart in which each hierarchy of belief is represented by at most one type.

## 2. BACKGROUND AND PRELIMINARY RESULTS

### 2.1 Notation

We refer to the identity mapping on a set  $Y$  by  $\text{id}_Y : Y \rightarrow Y$ . For any mappings  $f : A \rightarrow B$ ,  $g : A' \rightarrow B'$ , we define a product mapping  $f \times g : A \times B \rightarrow A' \times B'$  by  $(f \times g)(a, b) = (f(a), g(b))$ .

If  $A$  is a measurable space, then  $\Delta A$  is the set of all probability measures on  $A$  (relative to its given  $\sigma$ -field). If  $A$  and  $B$  are measurable spaces, then the space  $A \times B$  is endowed with its product  $\sigma$ -field. If  $A$  is a topological space, we treat it as a measurable space with its Borel  $\sigma$ -field, denoted  $\mathcal{B}_A$ , and the space of Borel probability measures on  $A$  is denoted  $\Delta A$ . If  $A$  is a Polish space, then  $\Delta A$  endowed with the weak\* topology is also Polish. For any  $a \in A$  let  $\delta(a) \in \Delta A$  be the Dirac measure concentrated on a point  $a$ .

For any measure  $\mu \in \Delta A$  and integrable function  $f : A \rightarrow \mathbf{R}$  we use  $\mu[f]$  to denote the expectation of  $f$  with respect to  $\mu$ . For measures on product spaces,  $\mu \in \Delta(A \times B)$ , denote by  $\text{marg}_A \mu \in \Delta(A)$  the marginal on  $A$  and by  $C_A \mu(\cdot) : B \rightarrow \Delta A$  a version of conditional probability over  $A$  given  $b \in B$  which exists whenever  $A$  is Polish. Our results do not depend on the choice of version. Similarly, for any measurable subset  $B' \subseteq B$ , we adopt the notation  $C_A \mu(B') \in \Delta A$  to signify the conditional probability measure over  $A$  given  $B'$ .

When  $A$  is a metric space then we denote by  $\mathcal{K}A$  the space of all non-empty closed subsets of  $A$  with the Hausdorff metric. If  $A$  is Polish, then so is  $\mathcal{K}A$ .

Given two measurable spaces,  $A, B$  and a measurable mapping  $\phi : A \rightarrow B$  we can in a natural way define a mapping that transports probability measures  $\Delta \phi : \Delta A \rightarrow \Delta B$ , such that for any measure  $\mu \in \Delta A$  and any measurable subset  $B' \subseteq B$ , we have  $\Delta \phi(\mu)(B') = \mu(\phi^{-1}(B'))$ .

<sup>8</sup>See also Bergemann and Morris (2005, Section 2.5) and Battigalli and Siniscalchi (2003, Section 6) for related discussions.

## 2.2 Games

We consider games with two players and take as given a Polish space of basic uncertainty  $\Omega$ . A *game form* (or simply game) over  $\Omega$  is a tuple  $G = (u_i, A_i)_{i=1,2}$ , where for each  $i$ ,  $A_i$  is a Polish space of actions and  $u_i : A_i \times A_{-i} \times \Omega \rightarrow \mathbf{R}$  is a bounded, measurable payoff function.

For some purposes it is useful to consider *product games*: take any two games  $G^1, G^2$ ,  $G^k = (A_i^k, u_i^k)$ . We construct a product game  $G = G_1 \times G_2 = (A_i, u_i)$ , where the action sets in  $G$  are the products of the actions sets from the original games,  $A_i = A_i^1 \times A_i^2$ , and payoffs are given by

$$u_i((a_i^1, a_i^2), (a_{-i}^1, a_{-i}^2), \omega) = u_i^1(a_i^1, a_{-i}^1, \omega) + u_i^2(a_i^2, a_{-i}^2, \omega).$$

A game  $G = (u_i, A_i)_i$  is *compact* if  $u_i$  are continuous and  $A_i$  are compact. A game  $G = (u_i, A_i)_i$  is *finite* if  $u_i$  are continuous and  $A_i$  are finite. We let  $\mathcal{G}$  denote the class of all compact games and  $\mathcal{G}^F$  the smaller class of all finite games. Unless we specifically state a restriction to compact or finite games, our results apply to all games. Note that if  $G^1, G^2$  are compact or finite, then so is  $G^1 \times G^2$ .

## 2.3 Type spaces

A *type space* over  $\Omega$ ,  $T = (T_i, \mu_i)_{i=1,2}$ , is a pair of measurable spaces  $T_i$  and two mappings  $\mu_i : T_i \rightarrow \Delta(\Omega \times T_{-i})$ . We say that a type space has *weakly measurable beliefs* if for any measurable function  $f : \Omega \times T_{-i} \rightarrow \mathbf{R}$ , the sets

$$\{t_i : \mu(t_i)[f] < 0\}$$

are measurable. We say that a type space has *strongly measurable beliefs* if there exist jointly measurable functions  $\beta_i : T_i \times T_{-i} \rightarrow \Delta\Omega$ , such that

$$\beta_i(t_i, t_{-i}) = C_{\Omega} \mu(t_i)(t_{-i}).$$

Let  $\mathcal{T}_w(\Omega)$  be the collection of all type spaces over  $\Omega$  with weakly measurable beliefs and  $\mathcal{T}_s(\Omega)$  be the collection of all type spaces over  $\Omega$  with strongly measurable beliefs.<sup>9</sup>

A type space over  $\Omega$  is an implicit way of modeling a collection of hierarchies of belief over  $\Omega$ . As we discussed in the introduction, under most commonly used solution concepts, what matters for behavior are hierarchies of belief about *conditional beliefs*. We can model these in an analogous way using type spaces over the space of basic uncertainty  $\Delta\Omega$ . The beliefs of a type  $t_i$  in such a type space  $T$  are probabilities over  $\Delta(\Delta\Omega \times T_{-i})$ . We interpret these as joint probabilities over the types of the opponent and conditional beliefs about  $\Omega$ . We will consider the class of all type spaces over  $\Delta\Omega$  with weakly measurable beliefs, denoted  $\mathcal{T}_w(\Delta\Omega)$ .

<sup>9</sup>Obviously any type space with strongly measurable beliefs has also weakly measurable beliefs,  $\mathcal{T}_s(\Omega) \subseteq \mathcal{T}_w(\Omega)$ . The connection in the other way is not clear. For any type space  $T \in \mathcal{T}_w(\Omega)$ , standard theorems guarantee existence of conditional beliefs  $\beta(t_i, t_{-i})$  that are measurable in  $t_{-i}$  for given  $t_i$ . We do not know, in general, whether we can choose conditional beliefs that are jointly measurable. We conjecture that this is possible for Polish type spaces.

There is a natural way in which any type space  $T = (T_i, \mu_i)$  over  $\Omega$  can be transformed into a type space  $T^\Delta = (T_i^\Delta, \mu_i^\Delta) \in \mathcal{T}_w(\Delta\Omega)$ . Let  $T_i^\Delta = T_i$  and for any  $t_i \in T_i^\Delta$  define  $\mu_i^\Delta(t_i) \in \Delta(\Delta\Omega \times T_{-i})$  to be the unique probability measure satisfying the following two conditions:

1. beliefs about opponent types are unchanged,

$$\text{marg}_{T_{-i}^\Delta} \mu_i^\Delta(t_i) = \text{marg}_{T_{-i}} \mu_i(t_i),$$

2. conditional beliefs of  $t_i \in T_i^\Delta$  about  $\Delta\Omega$  given type  $t_{-i}$  are a point mass on the conditional belief  $t_i$  in  $T_i$  given type  $t_{-i}$ :

$$C_{\Delta\Omega} \mu_i^\Delta(t_i)(t_{-i}) = \delta(C_\Omega \mu_i(t_i)(t_{-i})).$$

(Recall that  $\delta(\cdot)$  denotes Dirac delta measure.)

The logic behind this construction is the following. The translated beliefs  $\mu_i^\Delta(t_i)$  capture exactly the joint probability over the opponent's types and the resulting conditional beliefs as embodied in  $\mu_i(t_i)$ . In [Appendix C.1](#) we show that this defines a weakly measurable belief mapping.

LEMMA 1. *Suppose that  $T \in \mathcal{T}_s(\Omega)$ . Then  $T^\Delta \in \mathcal{T}_w(\Delta\Omega)$ .*

#### 2.4 Type mappings and the universal type space

Take two type spaces  $T = (T_i, \mu_i)$ ,  $T' = (T'_i, \mu'_i)$  over the same space of basic uncertainty  $X$ . A *type mapping* between  $T$  and  $T'$ , denoted  $\phi : T \rightarrow T'$ , is a pair of measurable mappings  $\phi_i : T_i \rightarrow T'_i$  transporting types from one space to the other.

One can think of a class of type mappings as capturing some notion of equivalency between types in different type spaces. Two types belong to the same equivalence class under that notion if they are associated under some type mapping in the class. Roughly speaking, *universal* type spaces are those that contain exactly one representative for each equivalence class. The literature has previously worked with the class of *belief-preserving* type mappings.

DEFINITION 1. A type mapping  $\phi$  between type spaces  $T, T' \in \mathcal{T}_w(X)$  preserves beliefs if for any measurable subset  $S' \subset X \times T'_{-i}$ ,

$$\mu'(\phi(t_i))(S') = \mu(t_i)(\hat{\phi}^{-1}(S')), \quad (1)$$

where  $\hat{\phi} = \text{id} \times \phi$ .

For an example of a belief-preserving type mapping, recall the two type spaces presented in the introduction. There is exactly one type mapping between the larger and the smaller type space. This mapping sends both types  $t_i = +1, -1$  into one type  $t'_i = *$ .

This type mapping also preserves beliefs. Indeed, type  $*$  believes that the two states are equally likely conditional on the opponent having type  $*$  and this is the same belief held by any type  $t_i$  conditional on the opponent having a type in  $\phi^{-1}(*) = \{+1, -1\}$ .

For the class of belief-preserving type mappings, the existence of a universal type space has been shown by [Mertens and Zamir \(1985\)](#), [Mertens et al. \(1994\)](#), [Brandenburger and Dekel \(1993\)](#), and [Battigalli and Siniscalchi \(1999\)](#) within various formal settings. We present a version of this result below.

**THEOREM 1.** *Let  $X$  be a Polish space and  $\mathcal{T}_w(X)$  the class of all type spaces over  $X$  with weakly measurable beliefs. There exists a universal type space  $U(X) \in \mathcal{T}_w(X)$  such that for any type space  $T \in \mathcal{T}_w(X)$ , there is a unique beliefs-preserving type mapping  $u^T : T_i \rightarrow U(X)$ .<sup>10</sup> Moreover,  $U_i(X)$  is a Polish space and the belief mapping  $\mu_i^{U(X)} : U_i(X) \rightarrow \Delta(X \times U_{-i}(X))$  is a homeomorphism.*

The theorem is a slight generalization of the aforementioned results in that it covers the case of purely measurable type spaces that we consider. (Previous studies used topological type spaces with continuous belief mappings.) The proof is an easy adaptation of the proof due to [Mertens et al. \(1994\)](#) and can be found in [Appendix A](#).

### 3. INTERIM RATIONALIZABILITY

Throughout the paper, we assume that payoffs in a game  $G \in \mathcal{G}$  depend only on the actions of both players and the state of the world. Sometimes it is convenient to use notation that indirectly makes payoff a function of types of the opponent. Given a payoff function  $u_i : A \times \Omega \rightarrow \mathbf{R}$ , we derive a new payoff function  $\pi_i : A \times T \rightarrow \mathbf{R}$ , defined directly in terms of the types as follows:

$$\pi_i(a, t) = \beta_i(t_i, t_{-i})[u_i(a_i, a_{-i}, \cdot)].$$

This payoff is calculated by “integrating out” the residual uncertainty over  $\Omega$  conditional on a realized type profile  $t$ .

An assessment is a pair of subsets  $\alpha = (\alpha_1, \alpha_2)$  where  $\alpha_i \subset T_i \times A_i$ . Alternatively an assessment can be defined by the pair of correspondences  $\alpha_i : T_i \rightrightarrows A_i$ , with  $\alpha_i(t_i) := \{a_i : (t_i, a_i) \in \alpha_i\}$ . The image  $\alpha_i(t_i)$  is interpreted as the set of actions that player  $i$  of type  $t_i$  could conceivably play.

Fix a type space  $T \in \mathcal{T}_s(\Omega)$  and a game  $G = (A_i, u_i)$ . A behavioral strategy for player  $i$  is a measurable function  $\sigma_i : T_i \rightarrow \Delta A_i$ . The expected payoff to type  $t_i$  of player  $i$  from

<sup>10</sup>In the language of category theory, the universal type space is a *terminal* object within the category of type spaces  $\mathcal{T}_w(X)$  connected with belief-preserving type mappings as morphisms.

choosing action  $a_i$  when the opponent's strategy is  $\sigma_{-i}$  is given by<sup>11</sup>

$$\begin{aligned} U_i(a_i, \sigma_{-i} \mid t_i) &= \mu_i^T(t_i)[\sigma_{-i}(t_{-i})[u_i(a_i, \cdot, \cdot)]] \\ &= \int_{\Omega \times T_{-i}} \int_{A_{-i}} u_i(a_i, \cdot, \omega) d\sigma_{-i}(t_{-i}) d\mu_i^T(t_i). \end{aligned}$$

The strategy  $\sigma_i$  is a selection from the assessment  $\alpha$  if for each  $i$ ,  $\sigma_i(t_i) \in \Delta \alpha_i(t_i)$  for all  $t_i \in T_i$ . Let  $\Sigma_i(\alpha)$  be the set of all strategies for  $i$  that are selections from  $\alpha$ .

A *conjecture* for player  $i$  of type  $t_i$  is a probability measure  $\sigma_{-i}^\Delta \in \Delta(T_{-i} \times A_{-i})$  such that its marginal on types of the opponent is equal to the marginal beliefs of type  $t_i$ ,

$$\text{marg}_{T_{-i}} \sigma_{-i}^\Delta = \text{marg}_{T_{-i}} \mu_i^T(t_i).$$

We say that the conjecture is consistent with the assessment  $\alpha$  if  $\sigma_{-i}^\Delta(\alpha_{-i}) = 1$ . (When  $\alpha_{-i}$  is not a measurable set, then we require that  $\sigma_{-i}^\Delta(Y) = 1$  for some measurable subset  $Y \subset \alpha_{-i}$ . When this is the case, for ease of exposition we simply write  $\sigma_{-i}^\Delta(\alpha_{-i}) = 1$ .) We denote the set of all conjectures of type  $t_i$  consistent with  $\alpha$  by  $\Sigma^\Delta(\alpha \mid t_i)$ .

Behavioral strategies and conjectures are equivalent representations of strategic uncertainty. For any behavioral strategy  $\sigma_{-i} \in \Sigma_{-i}(\alpha)$  and for any  $t_i$ , there is a conjecture  $\sigma_{-i}^\Delta \in \Sigma^\Delta(\alpha \mid t_i)$  such that

$$U_i(a_i, \sigma_{-i} \mid t_i) = \sigma_{-i}^\Delta[\pi_i(a_i, \cdot, t_i, \cdot)] := U_i(a_i, \sigma_{-i}^\Delta \mid t_i)$$

for every  $a_i \in A_i$ . Conversely, if  $\sigma_{-i}^\Delta$  is a conjecture for  $t_i$ , then there is a behavioral strategy  $\sigma_{-i}$  satisfying the same equalities.<sup>12</sup> We work with behavioral strategies or conjectures, whichever is most convenient.

An action  $a_i$  is an interim best-response for  $t_i$  against a conjecture  $\sigma_{-i}^\Delta$  if  $U_i(a_i, \sigma_{-i}^\Delta \mid t_i) \geq U_i(a'_i, \sigma_{-i}^\Delta \mid t_i)$  for all  $a'_i \in A_{-i}$ . Let  $B(\alpha \mid t_i)$  denote the set of all interim best-responses for  $t_i$  to any conjecture  $\sigma_{-i}^\Delta \in \Sigma^\Delta(\alpha \mid t_i)$ .

An assessment  $\alpha$  has the *best-response property* if every action attributed to player  $i$  is an interim best-reply to some conjecture concentrated on  $\alpha$ , i.e.,

$$\alpha_i \subset \{(t_i, a_i) : a_i \in B(\alpha \mid t_i)\}.$$

If the above is satisfied with equality, then we say that  $\alpha$  has the fixed-point property.

<sup>11</sup>This payoff function is defined with respect to a specific type space. In order to minimize notation, here and with similar constructs defined later, we omit explicit reference to this dependence whenever the context is clear.

<sup>12</sup>The conjecture associated with a given behavioral strategy is unique, but for any conjecture there will be many equivalent behavioral strategies which differ on sets of measure zero. Also note that if  $\alpha$  is an assessment that is empty for some type  $t_{-i}$ , and  $\sigma_{-i}^\Delta \in \Sigma^\Delta(\alpha \mid t_i)$ , then an equivalent behavioral strategy  $\sigma_{-i}$  will satisfy  $\sigma_{-i}(t_{-i}) \in \alpha_{-i}(t_{-i})$  for  $\mu(t_i)$ -almost every  $t_{-i}$ . But due to the emptiness, the equivalent behavioral strategies are not, strictly speaking, selections from  $\alpha$ .

PROPOSITION 1. *There exists a maximal (in the sense of set inclusion) assessment with the fixed-point property.*

PROOF. It is easy to verify that the union of assessments with the best-response property has the best-response property. Let  $R$  be the union of all assessments with the best-response property. Obviously  $R$  is the maximal set with the best-response property. We claim that  $R$  has the fixed-point property, in which case it will be the maximal such set. If  $R$  does not have the fixed-point property then there exists a type  $t_i$  and action  $a_i$  such that  $a_i$  is an interim best-reply to some conjecture consistent with  $R$ . But then we can add the pair  $(t_i, a_i)$  to  $R_i$  and obtain a larger assessment with the best-response property, a contradiction.  $\square$

DEFINITION 2. Given a type space  $T$  and a game  $G$ , the interim rationalizable correspondence is the maximal assessment with the fixed-point property, denoted  $R^{G,T}$ . We say that  $a_i$  is interim rationalizable for type  $t_i$  if  $a_i \in R_i^{G,T}(t_i)$ .

Note that in general the set  $R_i^{G,T}(t_i)$  may be empty. This can happen in games with discontinuities where best-replies need not exist. In Section 5 we give conditions under which the rationalizable sets are non-empty. Those results are used only in the parts of Section 6 where explicitly noted.

### 3.1 Type mappings that preserve rationalizability

We conclude this section with a result that is central to our analysis. It gives conditions under which two types, possibly from different type spaces, are equivalent in terms of rationalizable behavior. In Section 2.4, we discussed how equivalence relations among types can be captured by classes of type mappings. Equivalence in terms of belief hierarchies was captured by the class of belief-preserving type mappings. We now introduce a new class of type mappings that capture *behavioral* equivalence under rationalizability.

Let  $\phi$  be a type mapping between two type spaces  $T = (T_i, \mu_i)$ ,  $T' = (T'_i, \mu'_i)$ ,  $T, T' \in \mathcal{T}_s(X)$ , defined over the same space of basic uncertainty  $X$ . Suppose further that for every  $t_i$  there is a measurable mapping  $\phi_{t_i} : T_{-i} \rightarrow T'_{-i}$ , such that, for any measurable subset  $S' \subseteq T'_{-i}$ ,

$$\text{marg}_{T'_{-i}} \mu'(\phi(t_i))(S') = \text{marg}_{T_{-i}} \mu(t_i)(\phi_{t_i}^{-1}(S')). \quad (2)$$

The mappings  $\phi_{t_i}$  are referred to as the *dual* mappings. We can interpret the dual mapping  $\phi_{t_i}$  as describing how player  $i$  type  $t_i$  “thinks” that types of the opponent are transported.

We say that the dual mappings are *consistent* if for every  $t_i \in T_i$  and every  $t_{-i} \in T_{-i}$ ,

$$\mu'(\phi_{t_i}(t_{-i})) = \mu'(\phi(t_{-i})). \quad (3)$$

In other words, under the dual mapping  $\phi_{t_i}$ , the beliefs of the opponent are mapped in a way that is consistent with the underlying type mapping  $\phi$ .

We will consider type mappings with consistent duals. First note that if  $\phi$  is a belief-preserving type mapping, then the mappings  $\phi_i$  are themselves consistent duals. Indeed, (3) is satisfied by taking  $\phi_{t_i} = \phi_i$  for all  $t_i$ . But belief-preserving type mappings do not in general preserve rationalizable sets. Indeed, in [Section 2.4](#) we noted that the type mapping between type spaces from the introduction preserves beliefs. However this type mapping does not preserve rationalizable sets. Indeed, as we pointed out in the introduction, rationalizable behavior is different for types  $t_i = +1, -1$  and for type  $\phi(t_i) = t'_i = *$ .

Our discussion in the introduction points to the main problem: belief-preserving type mappings do not in general preserve the structure of conditional beliefs about payoff-relevant events. This motivates the following definition.

**DEFINITION 3.** Let  $T, T' \in \mathcal{T}_s(X)$ . A type mapping  $\phi : T \rightarrow T'$  preserves conditional beliefs iff there exist consistent dual mappings  $\phi_{t_i}$  such that for any type  $t_i \in T_i$ , for  $\mu_i(t_i)$ -almost any type  $t_{-i} \in T_{-i}$ ,

$$C_X \mu(t_i)(t_{-i}) = C_X \mu'(\phi(t_i))(\phi_{t_i}(t_{-i})). \quad (4)$$

In other words, the beliefs of type  $\phi(t_i)$  about  $X$  conditional on the opponent being of type  $\phi_{t_i}(t_{-i})$  are the same as the beliefs of  $t_i$  conditional on  $t_{-i}$ . Our result is that such type mappings preserve rationalizable sets.

**LEMMA 2.** Suppose that for type spaces  $T, T' \in \mathcal{T}_s(\Omega)$  there is a type mapping  $\phi : T \rightarrow T'$  that preserves conditional beliefs. Then it preserves rationalizable sets, i.e. for every game  $G$ , for all types  $t_i \in T_i$ ,

$$R_i^{G,T}(t_i) = R_i^{G,T'}(\phi(t_i)).$$

**PROOF.** Fix a game  $G$ . The proof naturally divides into two parts. First, we show that for any  $t_i \in T_i$ ,

$$R_i^{G,T}(t_i) \subset R_i^{G,T'}(\phi(t_i)) \quad (5)$$

and then we show the opposite direction

$$R_i^{G,T'}(\phi(t_i)) \subset R_i^{G,T}(t_i). \quad (6)$$

To show (5), consider the following assessment for type space  $T'$ :

$$\alpha'_i = \hat{\phi}_i(R_i^{G,T}) \bigcup \bigcup_{t'_{-i} \in T'_{-i}} \hat{\phi}_{t_{-i}}(R_i^{G,T}),$$

where  $\hat{\phi}_i = \phi_i \times \text{id}_A$ . We will show that  $\alpha'_i$  has the best-response property. This directly implies (5).

Let us write  $S'_i = \phi(T_i) \cup \bigcup_{t_{-i}} \phi_{t_{-i}}(T_i)$ . Note that for any  $t'_i \in S'_i$ ,

$$\alpha'_i(t'_i) = \bigcup_{t_i \in \phi^{-1}(t'_i)} R_i^{G,T}(t_i) \bigcup \bigcup_{t_{-i}} \bigcup_{t_i \in \phi_{t_{-i}}^{-1}(t'_i)} R_i^{G,T}(t_i) \quad (7)$$

and for any  $t'_i \notin S'_i$ ,  $\alpha'_i(t'_i) = \emptyset$ .

Pick  $t_i \in T_i$ . Let  $t'_i = \phi(t_i)$ ,  $a_i \in R_i^{G,T}(t_i)$ , and  $\sigma_{-i}^\Delta \in \Sigma^\Delta(R_{-i}^{G,T} | t_i)$ , such that  $a_i$  is a best response for  $t_i$  against a conjecture  $\sigma_{-i}^\Delta$ . We construct a conjecture  $\hat{\sigma}_{-i}^\Delta$  for  $t'_i$  as follows:

$$\hat{\sigma}_{-i}^\Delta = \sigma_{-i}^\Delta \circ \hat{\phi}_{t_i}^{-1}. \quad (8)$$

To show that this is indeed a conjecture for  $t'_i$ , we verify

$$\begin{aligned} \text{marg}_{T'_{-i}} \hat{\sigma}_{-i}^\Delta &= \text{marg}_{T'_{-i}} [\sigma_{-i}^\Delta \circ \hat{\phi}_{t_i}^{-1}] \\ &= \left[ \text{marg}_{T_i} \sigma_{-i}^\Delta \right] \circ \phi_{t_i}^{-1} \\ &= \mu_i^T(t_i) \circ \phi_{t_i}^{-1} \\ &= \mu_i^{T'}(t'_i). \end{aligned}$$

The third equality holds because  $\sigma_{-i}^\Delta$  is a conjecture for  $t_i$  and the fourth because  $\phi_{t_i}$  is a consistent dual.

We claim (8) implies that for any  $z_i \in A_i$ ,

$$U_i(z_i, \hat{\sigma}_{-i}^\Delta | \hat{t}_i) = U_i(z_i, \sigma_{-i}^\Delta | t_i).$$

To show this, we first use the fact that  $\phi$  preserves conditional beliefs to establish that the type-dependent payoff function  $\pi_i$  is preserved under  $\phi$ . For any action profile  $a$ , and type profile  $\hat{t} \in T$ ,

$$\begin{aligned} \pi_i^T(a, \hat{t}) &= (C_\Omega \mu_i^T(\hat{t}_i)(\hat{t}_{-i})) [u_i(a, \cdot)] \\ &= (C_\Omega \mu_i^{T'}(\phi(\hat{t}_i))(\phi_{t_i}(\hat{t}_{-i}))) [u_i(a, \cdot)] \\ &= \pi_i^{T'}(a, \phi(\hat{t}_i), \phi_{t_i}(\hat{t}_{-i})). \end{aligned}$$

Next, it follows that for any  $\hat{t}_i \in T_i$ ,

$$\begin{aligned} U_i(z_i, \sigma_{-i}^\Delta | \hat{t}_i) &= \sigma_{-i}^\Delta[\pi_i^T(z_i, a_{-i}, \hat{t}_i, t_{-i})] \\ &= \sigma_{-i}^\Delta[\pi_i^{T'}(z_i, a_{-i}, \phi(\hat{t}_i), \phi_{t_i}(t_{-i}))] \quad (\text{integrating over } a_{-i} \text{ and } t_{-i}) \\ &= \hat{\sigma}_{-i}^\Delta[\pi_i^{T'}(z_i, a_{-i}, \phi(\hat{t}_i), t'_{-i})] \quad (\text{integrating over } a_{-i} \text{ and } t'_{-i}) \\ &= U_i(z_i, \hat{\sigma}_{-i}^\Delta | \phi(\hat{t}_i)). \end{aligned}$$

The third equality holds because by the construction of  $\sigma_{-i}^\Delta$  (see (8)), for any measurable subset  $C \subset \mathbf{R}$ ,

$$\hat{\sigma}_{-i}^\Delta(\{(t'_{-i}, a_{-i}) : \pi_i^{T'}(z_i, a_{-i}, t'_i, t'_{-i}) \in C\}) = \sigma_{-i}^\Delta(\{(t_{-i}, a_{-i}) : \pi_i^{T'}(z_i, a_{-i}, t'_i, \phi_{t_i}(t_{-i})) \in C\}).$$

This establishes our claim.

Next, note that  $\hat{\sigma}_{-i}^\Delta$  is consistent with  $\alpha'_{-i}$ , i.e.  $\hat{\sigma}_{-i}^\Delta \in \Sigma^\Delta(\alpha'_{-i} | t'_i)$ . In particular, because  $\phi_{t_i}(R_{-i}^{G,T}) \subset \alpha'_{-i}$ , we have  $\sigma_{-i}^\Delta(\alpha'_{-i}) \geq \sigma_{-i}^\Delta(R_{-i}^{G,T}) = 1$ . We have therefore shown that  $a_i \in B(\alpha' | t'_i)$ . Since  $a_i$  was arbitrary, we conclude  $R_i^{G,T}(t_i) \subset B(\alpha' | t'_i)$ .

Next consider  $t_i'' = \phi_{t_{-i}}(t_i)$  for some  $t_{-i}$ . Because  $\phi_{t_{-i}}$  is a consistent dual,  $\mu_i^{T'}(t_i') = \mu_i^{T'}(t_i'')$ . It follows that

$$R_i^{G,T}(t_i) \subset B(\alpha' | t_i') = B(\alpha' | t_i'').$$

It now follows from (7) that  $\alpha'$  has the best-response property.

We turn to (6). Construct an assessment for  $T$  as follows:

$$\alpha_i(t_i) = \bigcup_{t_{-i}} R_i^{G,T'}(\phi_{t_{-i}}(t_i)).$$

Pick  $t_i \in T_i$ . For any  $t_{-i}$ , the two types  $\phi_{t_{-i}}(t_i)$  and  $\phi(t_i)$  have the same beliefs in  $T_{-i}$  because  $\phi_{t_{-i}}$  is a consistent dual. Therefore  $R_i^{G,T'}(\phi_{t_{-i}}(t_i)) = R_i^{G,T'}(\phi(t_i))$ . Thus,  $\alpha_i(t_i) = R_i^{G,T'}(\phi(t_i))$ . We will show that  $\alpha$  has the best-response property, which implies (6).

Let  $a_i \in R_i^{G,T'}(\phi(t_i))$  so that there is a conjecture  $\hat{\sigma}_{-i}^\Delta \in \Sigma^\Delta(R_{-i}^{G,T'} | \phi(t_i))$  such that  $a_i$  is a best response of  $t_i$  against  $\hat{\sigma}_{-i}^\Delta$ . The following conditions define the corresponding conjecture  $\sigma_{-i}^\Delta$  for  $t_i$ :

$$\text{marg}_{T_i} \sigma_{-i}^\Delta = \text{marg}_{T_i} \mu_i(t_i) \quad (9)$$

$$C_{A-i} \sigma_{-i}^\Delta(t_{-i}) = C_{A-i} \hat{\sigma}_{-i}^\Delta(\phi_{t_i}(t_{-i})). \quad (10)$$

We claim that this construction yields (8). Indeed, the marginals are equal:

$$\begin{aligned} \text{marg}_{T_{-i}'} \hat{\sigma}_{-i}^\Delta &= \text{marg}_{T_{-i}'} \mu_i^{T'}(\phi_i(t_i)) \\ &= \left[ \text{marg}_{T_{-i}} \mu_i^T(t_i) \right] \circ \phi_{t_i}^{-1} && \text{by (2)} \\ &= \left[ \text{marg}_{T_{-i}} \sigma_{-i}^\Delta \right] \circ \phi_{t_i}^{-1} && \text{by (9)} \\ &= \text{marg}_{T_{-i}'} [\sigma_{-i}^\Delta \circ \hat{\phi}_{t_i}^{-1}], \end{aligned}$$

and by (10), for  $\hat{t}_{-i} \in \phi_{t_i}(T_i)$  the conditionals are equal:

$$C_{A-i} \hat{\sigma}_{-i}^\Delta(\hat{t}_{-i}) = C_{A-i} [\sigma_{-i}^\Delta \circ \hat{\phi}_{t_i}^{-1}](\hat{t}_{-i}).$$

Since  $\phi_{t_i}(T_i)$  has  $\mu^{T'}(\phi_i(t_i))$ -probability 1, this completes the argument that (8) is satisfied. Thus, the claim from the first half of the proof applies and we can conclude  $a_i$  is a best response against  $\sigma_{-i}^\Delta$ . It remains to show that  $\sigma_{-i}^\Delta \in \Sigma^\Delta(\alpha | t_i)$ . This follows because  $\alpha_{-i} = \hat{\phi}_i^{-1}(R_{-i}^{G,T'})$ , so that

$$\sigma_{-i}^\Delta(\alpha_{-i}) = [\sigma_{-i}^\Delta \circ \hat{\phi}_i^{-1}](R_{-i}^{G,T'}) = \hat{\sigma}_{-i}^\Delta(R_{-i}^{G,T'}) = 1$$

where we have again used (8). □

## 4. THE MAIN RESULT

In this section we present our main result:  $\Delta$ -hierarchies are necessary and sufficient for identifying rationalizable behavior in all games. In the introduction we defined  $\Delta$ -hierarchies and demonstrated their construction. Here we present an equivalent definition that will be more convenient for the formal results.

Recall that any type space  $T \in \mathcal{T}_s(\Omega)$  is naturally transformed into a unique type space  $T^\Delta \in \mathcal{T}_w(\Delta\Omega)$  (Section 2.3). By Theorem 1, the space  $T_i^\Delta$  is mapped via the belief-preserving  $u_i^{T^\Delta}$  into the universal space  $U(\Delta\Omega)$  of all hierarchies of belief over  $\Delta\Omega$ . The  $\Delta$ -hierarchy of a type is its image in  $U(\Delta\Omega)$  under this mapping.

DEFINITION 4. Let  $T \in \mathcal{T}_s(\Omega)$ . The  $\Delta$ -hierarchy of a type  $t_i \in T_i$  is the image  $u_i^{T^\Delta}(t_i) \in U(\Delta\Omega)$ .

THEOREM 2.  *$\Delta$ -hierarchies are necessary and sufficient for identifying rationalizable behavior. In particular for any  $T, \hat{T} \in \mathcal{T}_s(\Omega)$ , any  $t_i \in T_i, \hat{t}_i \in \hat{T}_i$ ,*

1. (Sufficiency) *if  $u_i^{T^\Delta}(t_i) = u_i^{\hat{T}^\Delta}(\hat{t}_i)$  then  $R_i^{G,T}(t_i) = R_i^{G,\hat{T}}(\hat{t}_i)$  for all games  $G$*
2. (Necessity) *if  $u_i^{T^\Delta}(t_i) \neq u_i^{\hat{T}^\Delta}(\hat{t}_i)$  then there exists a finite game  $G$  such that  $R_i^{G,T}(t_i) \neq R_i^{G,\hat{T}}(\hat{t}_i)$ .*

Moreover,  $U(\Delta\Omega)$  does not contain any redundancies:

3. (Non-Redundancy) *for any  $u_i \in U_i(\Delta\Omega)$ , there is a type and a type space  $t_i \in T_i \in \mathcal{T}_s(\Omega)$  such that  $u_i = u_i^{T^\Delta}(t_i)$ .*

In the remainder of this section we prove the first and last claims in the statement. The proof of necessity requires some additional results on the structure of the rationalizable correspondence. These are presented in subsequent sections.

## 4.1 Non-redundancy

We begin with the last claim, that  $U(\Delta\Omega)$  contains no redundancies. To that end, we construct a type space  $L(\Omega)$  that itself includes all possible  $\Delta$ -hierarchies, and such that the mapping  $u^{L(\Omega)\Delta} : L(\Omega) \rightarrow U(\Delta\Omega)$  is onto. Let the space of types be

$$L_i(\Omega) = U_i(\Delta\Omega) \times \Delta\Omega.$$

The label of each type  $(u_i, \tau_i) \in L_i(\Omega)$  has two coordinates. The first corresponds to some  $\Delta$ -hierarchy,  $u_i \in U_i(\Delta\Omega)$ . The second,  $\tau_i$ , is a convenient “dummy” variable whose role will be seen shortly. For every type we define the beliefs

$$\mu_i^{L(\Omega)}(u_i, \tau_i) \in \Delta(\Omega \times L_{-i}(\Omega))$$

as the unique measure satisfying the following two conditions:

1. marginal beliefs over  $L_{-i}(\Omega)$  are obtained from  $U(\Delta\Omega)$ ,

$$\text{marg}_{L_{-i}(\Omega)}(u_i, \tau_i) = \mu_i^{U(\Delta\Omega)}(u_i)$$

2. conditional on the opponent's type  $(u_{-i}, \tau_{-i}) \in L_{-i}(\Omega) = U_{-i}(\Delta\Omega) \times \Delta\Omega$ , the conditional beliefs about  $\Omega$  are equal to  $\tau_{-i}$ ,

$$C_\Omega \mu_i^{L(\Omega)}(u_i, \tau_i)(u_{-i}, \tau_{-i}) = \tau_{-i}.$$

These conditional probabilities are measurable (in fact continuous) so that these beliefs properly define a type space with strongly measurable beliefs,  $L(\Omega) \in \mathcal{T}_s(\Omega)$ . Note also that the beliefs of type  $(u_i, \tau_i) \in L_i(\Omega)$  depend only on the first coordinate  $u_i$ .

For every  $u_i \in U_i(\Delta\Omega)$  there is a type in  $L(\Omega)$  with that label. It suffices to verify that the actual  $\Delta$ -hierarchy of any type in  $L(\Omega)$  in fact coincides with its label. Indeed, it is straightforward to check that  $u_i^{L(\Omega)\Delta}(u_i, \tau_i) = u_i$  for every  $(u_i, \tau_i) \in L_i(\Omega)$ .

#### 4.2 Sufficiency

To prove sufficiency, we begin by constructing for any  $T \in \mathcal{T}_s(\Omega)$  a type mapping into  $L(\Omega)$ . First, fix an arbitrary  $\tau \in \Delta\Omega$  and let  $\text{in}_i : U_i(\Delta\Omega) \rightarrow L_i(\Omega)$  be the inclusion mapping  $\text{in}_i^T(u_i) = (u_i, \tau)$  (none of the results below depends on the choice of  $\tau$ ). We may then define a type mapping  $l_i^T : T \rightarrow L(\Omega)$  as the following composition:

$$l^T = \text{in}_i^T \circ u_i^{T\Delta}. \quad (11)$$

Recall that the mapping  $u_i^{T\Delta}$  is the unique belief-preserving type mapping from  $T^\Delta$  to  $U(\Delta\Omega)$ . We show in the following lemma that, as a consequence, the mapping  $l^T$  preserves conditional beliefs.

**LEMMA 3.** *For any type space  $T \in \mathcal{T}_s(\Omega)$ , the type mapping  $l^T : T \rightarrow L(\Omega)$  preserves conditional beliefs.*

**PROOF.** Let  $T = (T_i, \mu_i^T)$ . We define the dual mapping for type  $t_i \in T_i$  by

$$\phi_{t_i}(t_{-i}) = \left( u_{-i}^{T\Delta}(t_{-i}), C_\Omega \mu_i^T(t_i)(t_{-i}) \right).$$

This is the product of measurable mappings, hence measurable. (The first coordinate is measurable because of **Theorem 1** and the second because  $T \in \mathcal{T}_s(\Omega)$ .) We verify that  $\phi_{t_i}^T$  satisfies equation (2) and is therefore a valid dual mapping. Take any measurable set of types  $S'$  in the space  $L(\Omega)$ , i.e.  $S' \subseteq L_{-i}(\Omega) = U_{-i}(\Delta\Omega) \times \Delta\Omega$ . We have a sequence of

equalities:

$$\begin{aligned}
\text{marg}_{L-i(\Omega)} \mu_i^{L(\Omega)}(l^T(t_i))(S') &= \text{marg}_{L-i(\Omega)} \mu_i^{L(\Omega)}(u_i^{T^\Delta}(t_i), \tau)(S') \\
&= \mu_i^{U(\Delta\Omega)}(u_i^{T^\Delta}(t_i))(S') \\
&= \mu_i^{T^\Delta}(t_i) \left( \left\{ (t_{-i}, \tau_{-i}) : (u_{-i}^{T^\Delta}(t_{-i}), \tau_{-i}) \in S' \right\} \right) \\
&= \text{marg}_{T_{-i}} \mu_i(t_i) \left( \left\{ t_{-i} : (u_{-i}^{T^\Delta}(t_{-i}), C_\Omega \mu_i^T(t_i)(t_{-i})) \in S' \right\} \right) \\
&= \text{marg}_{T_{-i}} \mu_i(t_i) (\phi_{t_i}^{-1}(S')).
\end{aligned}$$

The third equality follows from the fact that  $u^{T^\Delta} : T^\Delta \rightarrow U(\Delta\Omega)$  preserves beliefs and the fourth from the definition of beliefs on the space  $T^\Delta$ .

We check that  $\phi$  is consistent: note that for any  $u_i \in U_i(\Delta\Omega)$ , any  $\tau_i, \tau'_i \in \Delta\Omega$ ,  $\mu_i^{L(\Omega)}(u_i, \tau_i) = \mu_i^{L(\Omega)}(u_i, \tau'_i)$  so that for any  $t_i \in T_i$ ,  $t_{-i} \in T_{-i}$ ,

$$\begin{aligned}
\mu_{-i}^{L(\Omega)}(l^T(t_{-i})) &= \mu_{-i}^{L(\Omega)}(u_{-i}^{T^\Delta}(t_{-i}), \tau) \\
&= \mu_{-i}^{L(\Omega)}(u_{-i}^{T^\Delta}(t_{-i}), C_\Omega \mu_i^T(t_i)(t_{-i})) = \mu_{-i}^{L(\Omega)}(\phi_{t_i}(t_{-i})).
\end{aligned}$$

Verification that  $l$  preserves conditional beliefs becomes straightforward:

$$\begin{aligned}
C_\Omega \mu_i^{L(\Omega)}(l^T(t_i))(\phi_{t_i}(t_{-i})) &= C_\Omega \mu_i^{L(\Omega)}(l^T(t_i))(u_{-i}^{T^\Delta}(t_{-i}), C_\Omega \mu_i^T(t_i)(t_{-i})) \\
&= C_\Omega \mu_i(t_i)(t_{-i}).
\end{aligned} \quad \square$$

We can thus apply **Lemma 2** and conclude that the type mapping  $l^T$  preserves rationalizable sets. In particular, for any game  $G$ ,

$$R^{G,T}(t_i) = R^{G,L(\Omega)}(l^T(t_i))$$

and

$$R^{G,\hat{T}}(\hat{t}_i) = R^{G,L(\Omega)}(l^{\hat{T}}(\hat{t}_i)).$$

The sufficiency part of **Theorem 2** now follows immediately because  $u_i^{T^\Delta}(t_i) = u_i^{\hat{T}^\Delta}(\hat{t}_i)$  implies  $l^T(t_i) = l^{\hat{T}}(\hat{t}_i)$  so that the same set appears on the right-hand sides of the above equations.

## 5. STRUCTURE OF THE RATIONALIZABLE CORRESPONDENCE

In this section we analyze the correspondence  $R^{G,T}$  mapping types in a fixed type space  $T$  to their rationalizable sets in a fixed game  $G$ . When  $G$  is a compact and continuous, this correspondence is non-empty, closed-valued, and measurable in a strong

sense.<sup>13,14</sup> The measurability of the rationalizable correspondence is used extensively in the proof of necessity.

**PROPOSITION 2.** *For any game  $G \in \mathcal{G}$ , for each type space  $T \in \mathcal{T}_s(\Omega)$ , for each type  $t_i \in T_i$ , the set  $R_i^{G,T}(t_i)$  of interim rationalizable actions is non-empty and closed. Thus, we can view  $R_i^{G,T}$  as a function from  $T_i$  to  $\mathcal{K}A_i$ . This function is measurable: for every  $\mathcal{B} \in \mathcal{B}_{\mathcal{K}A_i}$ , the set*

$$\{t_i \in T_i : R_i^{G,T}(t_i) \in \mathcal{B}\}$$

*is a measurable subset of  $T_i$ .*

The proof of **Proposition 2** proceeds in two steps. First, we show that the rationalizable correspondence is non-empty valued and closed if the type space is a Polish space and satisfies the additional property that the mapping from types to beliefs is continuous in a strong sense.

Precisely, suppose that  $T$  is a type space and each  $T_i$  is a Polish space. Say that  $T$  is *continuous* if for each  $i$ , the mapping  $\mu_i^T : T_i \rightarrow \Delta(\Omega \times T_{-i})$  is continuous. If, in addition, there is a version of the conditional belief mapping  $\beta_i$  that is continuous then we say that  $T$  is  $\Delta$ -*continuous*. Note that  $\Delta$ -continuity is a stronger property than continuity alone. When the type space is  $\Delta$ -continuous and the game is compact, the rationalizable correspondence is upper hemicontinuous with non-empty values. The proof is in **Appendix E**.

**PROPOSITION 3.** *Suppose that  $T$  is  $\Delta$ -continuous and  $G$  is a compact game. Then  $R_i^{G,T}$  is upper hemicontinuous with non-empty values.*

It now follows from standard results that  $R_i^{G,T}$  is measurable in this case. The second step in the proof of **Proposition 2** is to show that any type space in  $\mathcal{T}_s(\Omega)$  can be measurably mapped into a  $\Delta$ -continuous type space in a way that preserves rationalizable sets. In particular, the space  $L(\Omega)$  is  $\Delta$ -continuous.

**PROOF OF PROPOSITION 2.** First we assume that  $T$  is  $\Delta$ -continuous. In this case, we can make use of the following fact. Let  $Y$  be a topological space,  $Z$  be a Polish space and  $\varphi : Y \rightarrow Z$  a correspondence with non-empty compact values. If  $\varphi$  is upper hemicontinuous, then the function  $\hat{\varphi} : Y \rightarrow \mathcal{K}Z$  defined by  $\hat{\varphi}(y) = \varphi(y)$  is Borel-measurable.<sup>15</sup> By **Proposition 3**,  $R_i^{G,T}$  is upper hemicontinuous with non-empty values. The former implies in particular that it has closed and hence compact values. This establishes the Proposition in the case of  $\Delta$ -continuous  $T$ .

<sup>13</sup>It is somewhat surprising that in general the correspondence need not be closed (i.e. upper hemicontinuous) even for compact games with continuous type spaces and  $\Omega$  compact, indeed even for finite games when  $\Omega$  is finite. This is shown by example in **Section 7**.

<sup>14</sup>For compact games  $G$ , it can also be shown that our fixed-point definition of rationalizability is equivalent to the outcome of iterative elimination of never-best replies. Since none of our results use the latter characterization, we have focused on the fixed-point characterization that applies to any game, even those with discontinuities. For a proof of the iterative characterization, see our working paper, **Ely and Pęski (2004, Proposition 6)**.

<sup>15</sup>See **Aliprantis and Border (1994, Corollary 14.70)**.

Notice that  $L(\Omega)$  is  $\Delta$ -continuous: the beliefs conditional on a given type of the opponent  $(\tau_{-i}, u_{-i})$  are equal to  $\tau_{-i}$  regardless of the type of player  $i$ . Hence, the rationalizable correspondence on  $L(\Omega)$  is upper hemicontinuous (hence closed) and measurable.

Now let  $T$  be any type space in  $\mathcal{T}_s(\Omega)$ . By Lemmas 2 and 3 there is a measurable mapping  $l^T : T \rightarrow L(\Omega)$  that preserves rationalizable sets. Thus, for any  $t_i \in T_i$ ,

$$R_i^{G,T}(t_i) = R_i^{G,L(\Omega)}(l^T(t_i)).$$

In other words, the correspondence  $R_i^{G,T}$  is the composition of this measurable mapping and the closed and measurable correspondence  $R_i^{G,L(\Omega)}$ . It follows that  $R_i^{G,T}$  is non-empty, closed-valued, and when viewed as a function, measurable.  $\square$

## 6. NECESSITY

This section deals with the proof of the necessity part in Theorem 2. To do this, we introduce a new type space,  $\mathcal{R}(\Delta\Omega)$ , which can be viewed as the universal type space for rationalizability. Here we briefly outline the construction.

For every type  $t_i$  in any type space, we can define a *rule of rationalizable behavior* (or simply a *rule*),  $\rho(t_i)$ , which is a mapping associating any compact game  $G \in \mathcal{G}$  with a subset of rationalizable actions of type  $t_i$  in this game  $G$ . By definition, for any two types with different rules, there is a game  $G$  such that these two types have different rationalizable sets in  $G$ . We may collect all rules of player  $i$  that are derived from some type in some type space into the set  $\mathcal{R}_i$ .

Now, we can restate the result that we are looking for: we want to show that two types with different  $\Delta$ -hierarchies,  $u^{T^\Delta}(t_i) \neq u^{T'^\Delta}(t'_i)$ , necessarily lead to two different rules of behavior,  $\rho(t_i) \neq \rho(t'_i)$ . In fact, we show below something stronger. There is a natural bijective mapping,  $\iota$ , between rules in  $\mathcal{R}$  and hierarchies in  $U(\Delta\Omega)$ , with the property that for any type  $t_i \in T_i$ , its rule is equal to the value of the mapping computed at this type's  $\Delta$ -hierarchy,

$$\rho(t_i) = \iota \left( u^{T^\Delta}(t_i) \right).$$

The proof proceeds as follows. The crucial step is to observe that  $\mathcal{R}$  can be seen as a type space over  $\Delta\Omega$ . For this we need to define beliefs of a rule  $r_i \in \mathcal{R}_i$  about  $\Delta\Omega$  and rules of the opponent  $\mathcal{R}_{-i}$ . Here is how it can be done. Take any type  $t_i \in T_i$  that leads to the rule  $r_i$  and find the beliefs of  $t_i$  about  $\Delta\Omega$  and types of the opponent  $T_{-i}$  (these are the beliefs of  $t_i$  in type space  $T^\Delta$  as in Section 2.3). Next, we identify the types of the opponent with their rules and transform in this way the beliefs of  $t_i$  into beliefs about  $\Delta\Omega$  and the rules of the opponent.

The first main result in this section (Theorem 3) is that any two types whose beliefs are distinct when viewed in this way must also give rise to distinct rationalizable rules. The proof involves constructing a finite game in which the two types have distinct rationalizable sets. This allows us to complete the description of the type space  $\mathcal{R}(\Delta\Omega)$  by defining the belief mapping. The beliefs of a type  $r_i$  are derived as above from the beliefs of any type  $t_i$  with such a rule. The theorem implies (see Proposition 4) that the beliefs will not depend on which such  $t_i$  we choose (there are of course many).

Having constructed the type space  $\mathcal{R}(\Delta\Omega)$  we present a result that justifies its interpretation as the universal type space for rationalizability. We show how to embed  $\mathcal{R}(\Delta\Omega)$  in a type space over  $\Omega$ ,  $\mathcal{R}(\Omega)$ . This is analogous to our previous derivation of the type space  $L(\Omega) \in \mathcal{T}_s(\Omega)$  from  $U(\Delta\Omega)$ . So, types in  $\mathcal{R}(\Omega)$  are labeled by two parts: a rule  $r_i$  and a “dummy” variable  $\tau_i \in \Delta\Omega$ . We show that this labeling is internally consistent—types  $\mathcal{R}(\Omega)$  with a label  $r_i$  (and whatever dummy) actually have  $r_i$  as their rule,

$$r_i = \rho(r_i, \tau_i).$$

Thus, all possible rationalizable rules are represented by a type.

Finally, we collect together all the mappings defined in this and the previous sections. Each rule  $r_i$  can be mapped into a type  $(r_i, \tau^*) \in \mathcal{R}(\Omega)$  (for some fixed dummy  $\tau^*$ ). Since  $(r_i, \tau^*)$  is a type in a type space over  $\Omega$ , we can compute its  $\Delta$ -hierarchy (using methods from [Section 4](#)). This defines a mapping from  $\mathcal{R}(\Delta\Omega)$  to  $U(\Delta\Omega)$ . In fact, this is the same mapping as the Mertens-Zamir belief preserving mapping guaranteed by [Theorem 1](#). Going the other way, we can map any  $\Delta$ -hierarchy  $u_i$  into type  $(u_i, \tau^*) \in L(\Omega)$ . Since  $(u_i, \tau^*)$  is a type in a type space over  $\Omega$ , we may compute its rationalizable rule  $r_i$ . This defines a mapping from  $U(\Delta\Omega)$  to  $\mathcal{R}(\Delta\Omega)$ . We show that the former mapping is an inverse of the latter ([Theorem 4](#)). This finishes the proof of the necessity part of [Theorem 2](#).

As a by-product (and a method) of the proof, we have showed that the space of all rules  $\mathcal{R}$ , which is defined in a highly abstract way, can be naturally reinterpreted as a type space. This fact is surprising—there is no a priori reason why a universal space for a solution concept could be characterized in such a way. We do not know whether this fact is an accident or consequence of some deeper relation.

### 6.1 The space of rationalizable rules

Here we define rationalizable rules. The set of all such rules will be used to define a type space over  $\Delta\Omega$ . We begin with the following sets<sup>16</sup> for each  $i$ :

$$\mathcal{S}_i = \prod_{G \in \mathcal{G}} \mathcal{K}A_i^G.$$

Any element  $r_i$  of  $\mathcal{S}_i$  can be viewed as a rule that assigns a (closed) subset of  $A_i^G$  to each  $G \in \mathcal{G}$ —recall that  $\mathcal{K}A_i^G$  is a compact Polish space with the Hausdorff metric. The value of a rule  $r_i$  on a particular game  $G \in \mathcal{G}$  is denoted by  $r_i(G)$ . We equip  $\mathcal{S}_i$  with the associated product topology and Borel  $\sigma$ -algebra.

A rule  $r_i$  is *rationalizable* if there exists a type space  $T$  and a type  $t_i \in T_i$  such that  $R_i^{G,T}(t_i) = r_i(G)$  for every  $G \in \mathcal{G}$ . We use the notation  $\rho_i^T : T_i \rightarrow \mathcal{S}_i$  for the mapping

<sup>16</sup>We are implicitly treating the class of compact, continuous games  $\mathcal{G}$  as a set in the following definition and elsewhere. To see that this is valid, recall that by Uryshon’s Metrization Theorem (see [Aliprantis and Border 1994](#)), any Polish action space can be embedded in the Hilbert cube  $\mathbf{H} = [0, 1]^{\mathbb{N}}$ . Thus any compact game can be equivalently viewed as a subset of the set  $\mathcal{K}\mathbf{H} \times \mathcal{K}\mathbf{H} \times F(\mathbf{H} \times \mathbf{H} \times \Omega)$  where  $F(X)$  is the set of real-valued functions on  $X$ .

that associates types in  $T$  with their corresponding rules. Let  $\mathcal{R}_i$  be the subset of  $\mathcal{S}_i$  consisting of all rationalizable rules, i.e.

$$\mathcal{R}_i = \{r_i \in \mathcal{S}_i : r_i = \rho_i^T(t_i) \text{ for some } T \in \mathcal{T}_s(\Omega) \text{ and some } t_i \in T_i\}.$$

The space  $\mathcal{R}_i$  inherits the topology and  $\sigma$ -algebra from  $\mathcal{S}_i$ . The derived  $\sigma$ -algebra is denoted  $\mathcal{B}_{\mathcal{R}_i}$ . We pause here to record the following important fact used repeatedly later (the proof can be found in [Appendix C.2](#)).

LEMMA 4. *For any type space  $T \in \mathcal{T}_s(\Omega)$ , the mapping  $\rho_i^T : T_i \rightarrow \mathcal{R}_i$  is measurable.*

Our goal is to treat  $\mathcal{R}$  as a type space over  $\Delta\Omega$  by assigning beliefs to each rule. To do this, first find some type space  $T$  and type  $t_i \in T_i$ , such that  $\rho_i^T(t_i) = r_i$ . Now consider the corresponding type space  $T^\Delta$  over  $\Delta\Omega$ . The beliefs  $\mu_i^\Delta(t_i)$  of  $t_i$  in  $T^\Delta$  assign probabilities to the types of the opponents and  $\Delta\Omega$ . The mapping  $\rho_{-i}$  induces a mapping that transforms these into beliefs over the opponent's rules and  $\Delta\Omega$ , as follows:

$$\mu_i^\Delta(t_i) \rightarrow \Delta(\text{id}_{\Delta\Omega} \times \rho_{-i}^T)(\mu_i^\Delta(t_i)) \in \Delta(\Delta\Omega \times \mathcal{R})$$

We would like to take these to be the beliefs of the rule  $r_i = \rho_i(t_i)$ . One potential complication arises from the fact that many different types (in many different type spaces) can induce the same rule. The beliefs of  $r_i$  would be well-defined only if all types that generate the rule  $r_i$  also generate the same beliefs. The result proved in the next section implies that this is indeed the case.

## 6.2 Rationalizable beliefs in a given game

In this section we prove a fundamental intermediate result in the proof of necessity. We fix a game  $G \in \mathcal{G}$ , a type space  $T \in \mathcal{T}_s(\Omega)$ , and  $T^\Delta$ , its corresponding type space over  $\Delta\Omega$ . Let us identify the types of player  $-i$  with their sets of rationalizable actions in  $G$ , i.e. via the mapping  $t_{-i} \rightarrow R_{-i}^{G,T}(t_{-i})$ . Now for any  $t_i$  the beliefs  $\mu_i^\Delta(t_i)$  naturally induce beliefs over  $\Delta\Omega$  and the possible rationalizable sets of the opponent:

$$\varpi^{G,T}(t_i) := \Delta(\text{id}_{\Delta\Omega} \times R_{-i}^{G,T})(\mu_i^\Delta(t_i)) \in \Delta(\Delta\Omega \times \mathcal{K}A_{-i}^G).$$

This way of viewing the beliefs of  $t_i$  strips away all type-space specific details except those that are intrinsically relevant for rationalizability in the particular game  $G$ . We call  $\varpi^{G,T}(t_i)$  the *rationalizable* beliefs for  $G$ . Note the relationship between  $\varpi^{G,T}(t_i)$  and the beliefs over rules presented in the previous subsection. Since  $R_{-i}^{G,T}(t_{-i}) = \rho_{-i}(t_{-i})[G]$ , they are the marginals of the latter corresponding to just the specific game  $G$ .

Consider two types  $t_i, t'_i$  whose beliefs agree when viewed in this way, i.e.  $\varpi^{G,T}(t_i) = \varpi^{G,T'}(t'_i)$ . It is not hard to see that these types must have the same rationalizable sets in  $G$ .<sup>17</sup> We are interested in a converse. However, if  $\varpi^{G,T}(t_i) \neq \varpi^{G,T'}(t'_i)$ , it does not follow

<sup>17</sup>The rationalizable actions of a type are just the set of best-replies to all conjectures about rationalizable actions of the opponent and states of the world. The condition implies that the two types have the same set of possible conjectures, hence the same set of best-replies.

that  $t_i$  and  $t'_i$  necessarily have distinct rationalizable sets in  $G$ .<sup>18</sup> Nevertheless we can show that if two types have differing beliefs relative to  $G$ , then there is another game  $G'$  in which their rationalizable sets differ. Moreover  $G'$  is a finite game even if  $G$  is not.

**THEOREM 3.** *Let  $G$  be any compact game. Let  $T, T' \in \mathcal{T}_s(\Omega)$  and consider any two types,  $t_i \in T_i$  and  $t'_i \in T'_i$ . If  $t_i$  and  $t'_i$  differ in terms of their rationalizable belief in game  $G$ , i.e.*

$$\varpi^{G,T}(t_i) \neq \varpi^{G,T'}(t'_i),$$

*then there is a finite game  $G'$  in which  $t_i$  and  $t'_i$  have distinct rationalizable sets, i.e.*

$$R_i^{G',T}(t_i) \neq R_i^{G',T'}(t'_i).$$

For the proof of **Theorem 3**, we need a technical lemma whose proof is in **Appendix B**. The lemma establishes that any two distinct beliefs can be “separated” by a certain class of function. The function will be used to construct payoffs in the game mentioned in the statement of **Theorem 3**.

**LEMMA 5.** *For any game  $G = (A_i, u_i)$  and  $\mu, \mu' \in \Delta(\Delta\Omega \times \mathcal{K}A_{-i})$ , if  $\mu \neq \mu'$  then there are natural numbers  $N_1, N_2$  and a continuous bounded function  $\psi : \{1, \dots, N_1\} \times A_{-i}^{N_2} \times \Omega \rightarrow [0, \infty)$  such that for  $f : \Delta\Omega \times \mathcal{K}A_{-i} \rightarrow \mathbf{R}$  defined by*

$$f(\tau, K) = \max_{k=1, \dots, N_1} \sup_{a_1, \dots, a_{N_2} \in K} \tau[\psi(k, a_1, \dots, a_{N_2}, \omega)],$$

*we have*

$$\mu[f] \neq \mu'[f].$$

**PROOF OF THEOREM 3.** Let  $N_1, N_2, \psi$ , and  $f$  be as given by **Lemma 5**. Suppose w.l.o.g. that in particular

$$\varpi^{G^*,T}(t_i)[f] < \varpi^{G^*,T'}(t'_i)[f].$$

Find  $\lambda > 0$ , such that

$$\varpi^{G^*,T}(t_i)[\lambda f - 1] < 0 < \varpi^{G^*,T'}(t'_i)[\lambda f - 1].$$

We will prove the theorem by constructing a game  $G = (A_j, u_j)$  such that  $R^{G,T}(t_i) \neq R^{G,T'}(t'_i)$ .

First, find a game  $G_0 = (A_j^0, u_j^0)$  such that  $A_{-i}^0 = \{1, \dots, N_1\}$  and all actions of player  $-i$  are rationalizable for all types of player  $-i$ , i.e. for any  $t_{-i} \in T_{-i}$  and any  $t'_{-i} \in T'_{-i}$ ,

<sup>18</sup>A simple counterexample, but by no means the only one, is a game in which player  $i$  is indifferent among all of his actions in any state  $\omega \in \Omega$ , but player  $-i$  is not. Different types of  $-i$  may have distinct rationalizable sets so that different types of player  $i$  will have distinct beliefs about the rationalizable sets of  $-i$ . Despite this, the rationalizable actions will be the same for all types of  $i$ .

$R_{-i}^{G_0, T}(t_{-i}) = R_{-i}^{G_0, T'}(t'_{-i}) = A_{-i}^0$ . (Such a game always exists.) Denote  $Z = \{0, 1\}$  and define sets of actions in game  $G$  as

$$\begin{aligned} A_i &= A_i^0 \times (A_i^*)^{N_2} \times Z, \\ A_{-i} &= A_{-i}^0 \times (A_{-i}^*)^{N_2}. \end{aligned}$$

Payoffs of player  $-i$  are given by

$$\begin{aligned} u_{-i} \left( (a_{-i}^0, a_{-i,1}^*, \dots, a_{-i,N_2}^*), (a_i^0, a_{i,1}^*, \dots, a_{i,N_2}^*), z \right), \omega \\ = u_{-i}^0(a_{-i}^0, a_i^0, \omega) + \sum_{k=1}^{N_2} u_{-i}^*(a_{-i,k}^*, a_{i,k}^*, \omega) \end{aligned}$$

(in particular they do not depend on  $z$ ) and payoffs of player  $i$  are given by

$$\begin{aligned} u_i \left( (a_i^0, a_{i,1}^*, \dots, a_{i,N_2}^*), (a_{-i}^0, a_{-i,1}^*, \dots, a_{-i,N_2}^*), \omega \right) \\ = u_i^0(a_i^0, a_{-i}^0, \omega) + \sum_{k=1}^{N_2} u_i^*(a_{i,k}^*, a_{-i,k}^*, \omega) + z \left[ \lambda \psi(a_{-i}^0, a_{-i,1}, \dots, a_{-i,N_2}, \omega) - 1 \right]. \end{aligned}$$

We show that the rationalizable sets for types  $t_i$  and  $t'_i$  are different in  $G$ . First observe that due to the product structure of the game  $G$ , for any type space  $S$ , for any type  $s_{-i} \in S_{-i}$ ,

$$R_{-i}^{G,S}(s_{-i}) = A_{-i}^0 \times [R_{-i}^{G^*,S}(s_{-i})]^{N_2}$$

and for any  $s_i \in S_i$ ,

$$R_i^{G,S}(s_i) = R_i^{G_0}(s_i) \times [R_i^{G^*,S}(s_i)]^{N_2} \times Z_i(s_i),$$

where  $Z_i(s_i) \subseteq Z$ .

In the type space  $T$ , consider the (pure) behavioral strategy of player  $-i$  that for type  $s_{-i}$  selects  $a_{-i}^0 \in \{1, \dots, N_1\}$  and  $(a_{-i,1}^*, \dots, a_{-i,N_2}^*)$  from  $R_{-i}^{G^*,S}(s_{-i})$  to maximize the expression  $\beta_i(t_i, s_{-i})[\psi(a_{-i}^0, a_{-i,1}^*, \dots, a_{-i,N_2}^*, \omega)]$ . By the measurable maximum theorem, this defines a measurable selection from  $R_{-i}^{G,T}$ .<sup>19</sup> Call this strategy  $\sigma_{-i}$ . We can define the analogous strategy  $\sigma'_{-i}$  for type space  $T'$  where type  $t'_i$  replaces  $t_i$  in the definition.

We calculate the payoff to type  $t'_i$  of player  $i$  from playing  $z = 1$  against  $\sigma'_{-i}$ :

$$\begin{aligned} \mu'_i(t'_i) \left[ \lambda \cdot \max_{k=1, \dots, N_1} \sup_{a_1, \dots, a_{N_2} \in K} \rho_i^{T'}(t'_i)(s'_{-i}) [\psi(k, a_1, \dots, a_{N_2}, \omega)] - 1 \right] \\ = \varpi^{G^*, T'}(t'_i) [\lambda f - 1] > 0. \end{aligned}$$

<sup>19</sup>See Aliprantis and Border (1994, Theorem 14.91). We need to check only that the mapping  $\beta_i(t_i, \cdot) [\psi(a_{-i}^0, a_{-i,1}^*, \dots, a_{-i,N_2}^*, \omega)]$  is measurable in  $s_{-i}$  and  $R_{-i}^{G^*,S}(s_{-i})$  is a measurable correspondence.

Thus,  $1 \in Z_i(t'_i)$ . On the other hand, the strategy  $\sigma_{-i}$  clearly maximizes, among all rationalizable strategies for player  $-i$  in type space  $T$ , the payoff that type  $t_i$  could receive from playing  $z = 1$  and

$$\begin{aligned} \mu_i(t_i) \left[ \lambda \cdot \max_{k=1, \dots, N_1} \sup_{a_1, \dots, a_{N_2} \in K} \rho_i^T(t_i)(s_{-i}) [\psi(k, a_1, \dots, a_{N_2}, \omega)] - 1 \right] \\ = \varpi^{G^*, T}(t_i) [\lambda f - 1] < 0. \end{aligned}$$

Hence  $1 \notin Z(t_i)$ .  $\square$

### 6.3 The universal type space for rationalizability

**Theorem 3** allows us to complete the description of  $\mathcal{R}(\Delta\Omega)$  by defining beliefs. Let  $r_i$  be any rule and  $t_i$  an arbitrary type such that  $\rho_i(t_i) = r_i$ . We define

$$\mu_i^{\mathcal{R}(\Delta\Omega)}(r_i) := \Delta(\text{id}_{\Delta\Omega} \times \rho_{-i}^T)(\mu_i^\Delta(t_i)).$$

The question raised in **Section 6.1** is whether these beliefs are well-defined, in particular whether they depend on the choice of  $t_i$ . We can interpret **Theorem 3** as stating that all single-game *marginals* of these beliefs are independent of the choice of  $t_i$ . The following proposition extends the conclusion to the entire measure by an appeal to product games and an application of Kolmogorov's extension theorem.

**PROPOSITION 4.** *For any two type spaces  $T, T' \in \mathcal{T}_s(\Omega)$  if two types  $t_i \in T_i, t'_i \in T'_i$  have the same rationalizable sets in every finite game*

$$R_i^{G, T}(t_i) = R_i^{G, T'}(t'_i) \quad \text{for all } G \in \mathcal{G}^F$$

*then they also generate the same beliefs about the conditional beliefs and the opponent's rules,*

$$\Delta(\text{id}_{\Delta\Omega} \times \rho_{-i}^T)(\mu_i^\Delta(t_i)) = \Delta(\text{id}_{\Delta\Omega} \times \rho_{-i}^{T'})(\mu_i^\Delta(t'_i)).$$

**PROOF.** By the Kolmogorov extension theorem, it is enough to show that for any finite number of games  $G_1, \dots, G_k \in \mathcal{G}$  we have

$$\text{marg}_{\Delta\Omega \times \mathcal{H} A_{-i}^{G_1} \times \dots \times \mathcal{H} A_{-i}^{G_k}} \Delta(\text{id}_{\Delta\Omega} \times \rho_{-i}^T)(\mu_i^\Delta(t_i)) = \text{marg}_{\Delta\Omega \times \mathcal{H} A_{-i}^{G_1} \times \dots \times \mathcal{H} A_{-i}^{G_k}} \Delta(\text{id}_{\Delta\Omega} \times \rho_{-i}^T)(\mu_i^\Delta(t'_i)).$$

The last equality is equivalent to

$$\Delta(\text{id}_{\Delta\Omega} \times R_{-i}^{G_1, T} \times \dots \times R_{-i}^{G_k, T})(\mu_i^\Delta(t_i)) = \Delta(\text{id}_{\Delta\Omega} \times R_{-i}^{G_1, T} \times \dots \times R_{-i}^{G_k, T})(\mu_i^\Delta(t'_i)).$$

To show this we consider product games (defined in **Section 2**). Observe that the set of rationalizable actions in the product game  $G = G_1 \times \dots \times G_k$  is just the product of the sets rationalizable actions in the component games  $G_1, \dots, G_k$ . This means that

$$\begin{aligned} \Delta(\text{id}_{\Delta\Omega} \times R_{-i}^{G_1, T} \times \dots \times R_{-i}^{G_k, T})(\mu_i^\Delta(t_i)) &= \varpi^{G, T}(t_i) \\ \Delta(\text{id}_{\Delta\Omega} \times R_{-i}^{G_1, T} \times \dots \times R_{-i}^{G_k, T})(\mu_i^\Delta(t'_i)) &= \varpi^{G, T}(t'_i). \end{aligned}$$

By [Theorem 3](#), if  $t_i$  and  $t'_i$  have the same rationalizable sets in all finite games, then the same belief appears on the right hand sides. This concludes the proof.  $\square$

Having defined the beliefs, we must now show that the belief mapping is weakly measurable, so that  $\mathcal{R}(\Delta\Omega)$  is a well-defined type space. The proof can be found in [Appendix C.2](#).

**PROPOSITION 5.** *The belief mapping  $\mu_i^{\mathcal{R}} : \mathcal{R}_i \rightarrow \Delta(\Delta\Omega \times \mathcal{R}_{-i})$  is weakly measurable. Thus  $\mathcal{R}(\Delta\Omega) \in \mathcal{T}_w(\Delta\Omega)$ .*

Next, we will embed  $\mathcal{R}(\Delta\Omega)$  within a type space over the basic space of uncertainty  $\Omega$  to verify that our construction is internally consistent, i.e. the rationalizable behavior of a type  $r_i$  is indeed  $r_i$ . The construction is directly analogous to our previous construction of  $L(\Omega)$  from  $U(\Delta\Omega)$ .

Precisely, we define a type space  $\mathcal{R}(\Omega)$  over  $\Omega$  in the following way. Let  $\mathcal{R}_i \times \Delta\Omega$  be the space of types for player  $i$ . For every type  $(r_i, \tau_i) \in \mathcal{R}_i \times \Delta\Omega$  we define the beliefs

$$\mu_i^{\mathcal{R}(\Omega)}(r_i, \tau_i) \in \Delta(\Omega \times (\mathcal{R}_{-i} \times \Delta\Omega))$$

as the unique measure satisfying the following two conditions:

1. marginal beliefs over the opponent's types are obtained from  $\mathcal{R}(\Delta\Omega)$ ,

$$\text{marg}_{\Delta\Omega \times \mathcal{R}_{-i}} \mu_i^{\mathcal{R}(\Omega)}(r_i, \tau_i) = \mu_i^{\mathcal{R}}(r_i)$$

2. conditional on the opponent's type  $(r_{-i}, \tau_{-i})$ , the beliefs about  $\Omega$  are equal to  $\tau_{-i}$ ,

$$C_{\Omega} \mu_i^{\mathcal{R}(\Omega)}(r_i, \tau_i)(r_{-i}, \tau_{-i}) = \tau_{-i}.$$

These conditional probabilities are measurable (in fact continuous) so that these beliefs properly define a type space with strongly measurable beliefs,  $\mathcal{R}(\Omega) \in \mathcal{T}_s(\Omega)$ .<sup>20</sup>

Finally, to justify referring to  $\mathcal{R}(\Delta\Omega)$  as the universal space for rationalizability (in addition to more instrumental purposes later on), we now show that the construction is internally consistent. In particular, each type in  $\mathcal{R}(\Omega)$  is “labeled” by a rationalizable rule. We will demonstrate that the actual rule of a type coincides with its label.

Fix an arbitrary  $\tau \in \Delta\Omega$ . For any type space  $T \in \mathcal{T}_s(\Omega)$ , consider the type-mapping  $\phi^T : T \rightarrow \mathcal{R} \times \Delta\Omega$  defined by

$$\phi_i^T(t_i) = (\rho_i^T(t_i), \tau).$$

The measurability of  $\phi^T$  follows from [Lemma 4](#).

<sup>20</sup>The construction of the type space  $\mathcal{R}(\Omega)$  over  $\Omega$  is somewhat arbitrary. In fact, there are many ways to construct a valid model (among type spaces over  $\Omega$ ) for  $\mathcal{R}(\Delta)$ . (By contrast, the type space  $\mathcal{R}(\Delta)$  is unique up to isomorphisms, as we prove later.) We have chosen the construction in the text as it is the most natural and convenient for expositional purposes.

PROPOSITION 6. *For any type space  $T \in \mathcal{T}_s(\Omega)$ , the type mapping  $\phi^T : T \rightarrow \mathcal{R}(\Omega)$  preserves conditional beliefs. Hence, it preserves rationalizable sets. Moreover, for any rule  $r_i \in \mathcal{R}_i$  and any  $\tau_i \in \Delta\Omega$ ,*

$$r_i = \rho_i^{\mathcal{R}(\Omega)}(r_i, \tau_i).$$

PROOF. We define the dual mapping as follows: for each type  $t_i \in T_i$  let

$$\phi_{t_i}(t_{-i}) = \left( \rho_{-i}^T(t_{-i}), C_{\Omega}\mu_i(t_i)(t_{-i}) \right).$$

The measurability follows from Lemma 4. Following exactly the line of argument used in the proof of Lemma 3, we may check that these are consistent duals and that  $\phi$  is a conditional-belief preserving type mapping.

Take now any rationalizable rule  $r_i \in \mathcal{R}_i$  and an arbitrary  $\hat{\tau} \in \Delta\Omega$ . There is a type in a type space  $t_i \in T_i$  such that  $r_i = \rho_i^T(t_i)$ . Since the type mapping  $\phi^T$  preserves rationalizable sets, we have

$$r_i = \rho_i^T(t_i) = \rho_i^{\mathcal{R}(\Omega)}(r_i, \tau) = \rho_i^{\mathcal{R}(\Omega)}(r_i, \hat{\tau}).$$

The last equality comes from the fact that beliefs of types in  $\mathcal{R}(\Omega)$ , hence also their rationalizable sets, depend only on the  $r_i$ -coordinate.  $\square$

COROLLARY 1. *The following statements are equivalent.*

- (i)  $r_i = r'_i$ .
- (ii)  $r_i$  and  $r'_i$  coincide for all finite games.
- (iii)  $\mu_i^{\mathcal{R}(\Delta\Omega)}(r_i) = \mu_i^{\mathcal{R}(\Delta\Omega)}(r'_i)$ .

PROOF. The implication (i)  $\implies$  (ii) is trivial. The implication (ii)  $\implies$  (iii) is a restatement of Proposition 4. To show (iii)  $\implies$  (i), consider the types  $(r_i, \tau)$  and  $(r'_i, \tau)$  in  $\mathcal{R}(\Omega)$  for arbitrary  $\tau \in \Delta\Omega$ . By (iii) and the construction of  $\mathcal{R}(\Omega)$ , they have the same beliefs. Thus, they have the same rationalizable sets in every game, i.e. the same rule. By Proposition 6 the rule of  $(r_i, \tau)$  is  $r_i$  and the rule of  $(r'_i, \tau)$  is  $r'_i$ . Thus  $r_i = r'_i$ .  $\square$

#### 6.4 $\mathcal{R}(\Delta\Omega)$ is isomorphic to $U(\Delta\Omega)$

Finally, we show the central result of the paper: the two spaces  $U(\Delta\Omega)$  and  $\mathcal{R}(\Delta\Omega)$ , although described in different ways, are equivalent in terms of their relevant structure. In particular we show in this subsection that they are isomorphic under a belief-preserving type mapping. The following lemma plays an important role.

Define the projection  $\text{proj}_T : T \times \Delta\Omega \rightarrow T$  by  $\text{proj}_T(t_i, \tau_i) = t_i$ , and for arbitrary  $\tau \in \Delta\Omega$  the inclusion  $\text{in}^{T, \tau} : T \rightarrow T \times \Delta\Omega$  by  $\text{in}^{T, \tau}(t_i) = (t_i, \tau)$ , so that

$$\text{proj}_T \circ \text{in}^{T, \tau} = \text{id}_T.$$

The next lemma states a handy property that connects the category of type spaces over  $\Omega$  with type spaces over  $\Delta\Omega$ . (It is proved in Appendix D.)

LEMMA 6. Let  $S, T \in \mathcal{T}_w(\Delta\Omega)$ .

(i) For any belief preserving type mapping  $\phi^\Delta : S \rightarrow T$ , the composite mapping

$$\phi = \text{in}^{T,\tau} \circ \phi^\Delta \circ \text{proj}_S$$

is a conditional beliefs preserving type mapping  $\phi : S \times \Delta\Omega \rightarrow T \times \Delta\Omega$ .

(ii) Suppose different types in  $T$  have different beliefs,  $t_i \neq t'_i \implies \mu_i(t_i) \neq \mu_i(t'_i)$ . Then for any conditional beliefs preserving type mapping  $\phi : S \times \Delta\Omega \rightarrow T \times \Delta\Omega$ , the composite mapping

$$\phi^\Delta = \text{proj}_T \circ \phi \circ \text{in}^{S,\tau}$$

is a belief preserving type mapping  $\phi^\Delta : S \rightarrow T$ .<sup>21</sup>

The property assumed in the second part of the lemma is satisfied by both spaces  $U(\Delta\Omega)$  and  $\mathcal{R}(\Delta\Omega)$ . In the former case, this is the last part of the statement of **Theorem 1**. In the latter case, this is **Corollary 1**.

THEOREM 4. There are unique beliefs-preserving type mappings  $\iota : U(\Delta\Omega) \rightarrow \mathcal{R}(\Delta\Omega)$  and  $\iota^{-1} : \mathcal{R}(\Delta\Omega) \rightarrow U(\Delta\Omega)$ . Either mapping is the inverse of the other:  $\iota^{-1} \circ \iota = \text{id}_{U(\Delta\Omega)}$  and  $\iota \circ \iota^{-1} = \text{id}_{\mathcal{R}(\Delta\Omega)}$ .

PROOF. Existence and uniqueness of the type mapping  $\iota^{-1}$  is assured by **Theorem 1**. In order to show existence of a beliefs-preserving type mapping from  $U(\Delta\Omega)$  to  $\mathcal{R}(\Delta\Omega)$ , note first that by the first half of **Proposition 6**, there is a conditional beliefs-preserving type mapping  $\phi^{L(\Omega),\tau} : L(\Omega) \rightarrow \mathcal{R}(\Omega)$ . The second part of **Lemma 6** then guarantees existence of a beliefs-preserving mapping  $(\phi^{L(\Omega),\tau})^\Delta : U(\Delta\Omega) \rightarrow \mathcal{R}(\Delta\Omega)$ .

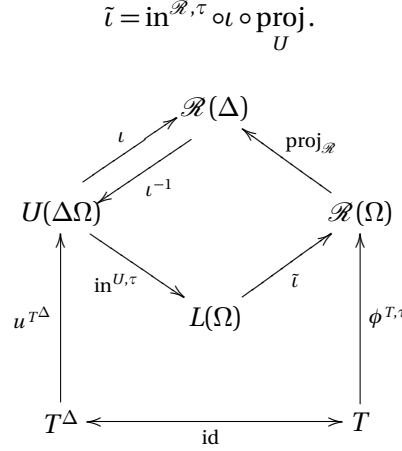
Suppose now that we have two different belief-preserving mappings  $\iota_1, \iota_2 : U(\Delta\Omega) \rightarrow \mathcal{R}(\Delta\Omega)$ . There is  $u_i \in U_i(\Delta\Omega)$  such that  $\iota_1(u_i) \neq \iota_2(u_i)$ . By the first part of **Lemma 6** there are then two type mappings  $\phi_1, \phi_2 : L(\Omega) \rightarrow \mathcal{R}(\Omega)$  that preserve conditional beliefs and satisfy  $\phi_1(u_i, \tau) = (i_1(u_i), \tau) \neq (i_2(u_i), \tau) = \phi_2(u_i, \tau)$ . By **Lemma 2**,  $\phi_1(u_i, \tau)$  and  $\phi_2(u_i, \tau)$  must have the same rationalizable rules. But this is a contradiction because **Proposition 6** shows that the rationalizable rule for  $\phi_1(u_i, \tau)$  is  $\iota_1(u_i)$  while the rationalizable rule for  $\phi_2(u_i, \tau)$  is  $\iota_2(u_i)$ .

The equality  $\iota^{-1} \circ \iota = \text{id}_{U(\Delta\Omega)}$  comes from the uniqueness of belief-preserving type mapping from  $U(\Delta\Omega)$  to itself. The second equality  $\iota \circ \iota^{-1} = \text{id}_{\mathcal{R}(\Delta\Omega)}$  is a consequence of the fact that  $\iota \circ \iota^{-1}$  would generate a conditional belief preserving mapping from  $\mathcal{R}(\Omega)$  to itself (the first part of **Lemma 6** guarantees that). Such a mapping has to preserve rules, so it has to preserve  $r$ -coordinates of types in  $\mathcal{R}(\Omega)$ . This implies that  $\iota \circ \iota^{-1} = \text{id}_{\mathcal{R}(\Delta\Omega)}$ .  $\square$

<sup>21</sup>Within the framework of category theory, the lemma defines functors between the category of type spaces over  $\Omega$  (with morphisms corresponding to conditional-belief-preserving type mappings) and the category of type spaces over  $\Delta\Omega$  (with morphisms corresponding to belief-preserving type mappings).

## 6.5 Proof of necessity

The structure of the proof can be seen via the following diagram which collects various mappings defined in the paper, where the type mapping  $\tilde{\iota} : L(\Omega) \rightarrow \mathcal{R}(\Omega)$  is defined as follows



Suppose  $u_i^{T^\Delta}(t_i) \neq u_i^{T^\Delta}(t_{-i})$ . By [Corollary 1](#), in order to show that the two types have distinct rationalizable sets in some finite game, it suffices to show that they have distinct rationalizable rules. Consider the route in the diagram  $T \rightarrow T^\Delta \rightarrow U(\Delta\Omega) \rightarrow L(\Omega)$ , which by [Lemma 3](#) preserves rationalizable sets. The images of  $t_i$  and  $t'_i$  under this mapping are elements of  $L(\Omega)$  with distinct first coordinates (namely  $u_i^{T^\Delta}(t_i)$  and  $u_i^{T^\Delta}(t'_i)$ ). We show that they have distinct rationalizable sets which implies the same for  $t_i$  and  $t'_i$ , concluding the proof.

Because  $\iota$  is one-to-one (by [Theorem 4](#)), if two types in  $L(\Omega)$  have distinct first coordinates, then their images under  $\tilde{\iota}$  also have distinct first coordinates. And by [Proposition 6](#), if two types in  $\mathcal{R}(\Omega)$  have distinct first coordinates, then they have distinct rationalizable rules. In particular there is some game in which they have distinct rationalizable sets. Now by [Theorem 4](#) and [Lemma 6](#) the mapping  $\tilde{\iota}$  preserves conditional beliefs. By [Lemma 2](#) it preserves rationalizable sets. We conclude that two types in  $L(\Omega)$  with distinct first coordinates have distinct rationalizable rules.  $\square$

## 7. COMMENTS AND OTHER EXAMPLES

## 7.1 Example

The conventional universal type space  $U(\Omega)$  is not rich enough from the point of view of solution concepts such as Bayesian equilibrium or Rationalizability. We have previously shown this by demonstrating that there are types whose rationalizable rules are not represented by any type in  $U(\Omega)$ . Here we present an example that makes the point even stronger: there is an action that is not rationalizable for any type in  $U(\Omega)$ , yet as we show below, it is easy to construct simple, perfectly standard type spaces in which the action is rationalizable. Consider the two-player game with two states of the world with payoff matrix given in [Figure 2](#).

|        | $a_2$  | $b_2$  | $b'_2$ | $a'_2$ |
|--------|--------|--------|--------|--------|
| $a_1$  | 1, 1   | 1, -9  | -1, -9 | -1, -1 |
| $b_1$  | -9, 1  | 0, 0   | -9, -9 | -9, -1 |
| $b'_1$ | -9, -1 | -9, -9 | 0, 0   | -9, 1  |
| $a'_1$ | -1, -1 | -1, -9 | 1, -9  | 1, 1   |

$\omega = +1$

|        | $a_2$  | $b_2$  | $b'_2$ | $a'_2$ |
|--------|--------|--------|--------|--------|
| $a_1$  | 1, 1   | 1, -9  | -1, -9 | -1, -1 |
| $b_1$  | -9, 1  | -9, -9 | 0, 0   | -9, -1 |
| $b'_1$ | -9, -1 | 0, 0   | -9, -9 | -9, 1  |
| $a'_1$ | -1, -1 | -1, -9 | 1, -9  | 1, 1   |

$\omega = -1$

FIGURE 2. Actions  $b_i$  and  $b'_i$  are not rationalizable in  $U(\Omega)$ .

We will show that neither  $b_i$  nor  $b'_i$  are rationalizable for any type in  $U(\Omega)$ . Note first that an equal mixture between  $a_i$  and  $a'_i$  guarantees a payoff of 0. Thus,  $b_i$  and  $b'_i$  are best-replies only if player  $i$  is certain that the opponent plays an action in  $\{b_{-i}, b'_{-i}\}$ , and the action is correlated with the state. Now if  $i$  assigns greater than  $\frac{1}{2}$  probability to state  $+1$ , then it is easily verified that action  $a_i$  achieves strictly higher payoff than  $b_i$ , and action  $a'_i$  achieves strictly higher payoff than  $b'_i$ , regardless of the opponent's strategy. Likewise, if the probability of state  $+1$  is less than  $\frac{1}{2}$ , then  $a_i$  must do better than  $b'_i$  and  $a'_i$  better than  $b_i$ . Thus, actions  $b_i$  or  $b'_i$  can be rationalizable only for types who assign the two states equal probability and who assign probability 1 to opponent's types for whom  $b_{-i}$  or  $b'_{-i}$  are rationalizable. Now the game is symmetric, so the same analysis applies to player  $-i$  with the players' roles reversed. Putting these two conclusions together, actions  $b_i$  and  $b'_i$  are rationalizable only for types of player  $i$  who assign equal probability to the two states and probability 1 to the event that player  $-i$  has the same beliefs and assigns probability 1 to the event that  $b_i$  and  $b'_i$  are rationalizable for  $i$ . By induction,  $b_i$  and  $b'_i$  are rationalizable only for those types of player  $i$  for whom it is common-knowledge that the two states are equally likely. Let  $v_i$  be the type in  $U_i(\Omega)$  with this hierarchy of beliefs, and  $v_{-i}$  the analogous type for player  $-i$ . Note that in  $U(\Omega)$ , type  $v_i$  assigns probability 1 to  $v_{-i}$  and equal probability to the two states. But then no matter what mixed action is played by  $v_{-i}$ , it is never correlated with the state. Thus  $b_i$  and  $b'_i$  can never be best-replies, hence never be rationalizable for type  $v_i$ .

Nevertheless, both  $b_i$  and  $b'_i$  are rationalizable for all types in the type space from the introduction. Indeed, any pure strategy profile in which the two types of each player play different actions in  $\{b_i, b'_i\}$  is a Bayesian Nash equilibrium.

## 7.2 More than two players

In games with more than two players, there are many natural ways to extend the definition of rationalizability. This is well-understood already in the context of complete-

information where *correlated* rationalizability allows players to conjecture that the opponents' play is correlated while *independent* rationalizability does not. When information is incomplete, there is an even greater variety of plausible extensions due to the greater variety of ways in which conjectures can be correlated. First, a player may suppose that the opponents achieve correlation by conditioning their play on the outcome of some randomization device that is uncorrelated with the state of the world. A different set of conjectures would result if a player were to entertain the possibility that the opponents communicated their private information to one another prior to choosing their actions. Combining these two possibilities would lead to a larger set of conjectures. Of course the assumption that the opponents cannot correlate at all would lead to the smallest set of conjectures. Each of these four possibilities would be extensions of interim rationalizability in the sense that they would reduce in two player games to the definition we have employed. Our method of analysis can be adapted to these versions of rationalizability. Our conjecture is that the necessary and sufficient conditions will be different for each of these different versions. The task of characterizing these conditions is left for future work.

Finally, one version of rationalizability that is not an extension of interim rationalizability is the concept of *correlated* interim rationalizability studied by Dekel et al. (2005a). In that concept a player may conjecture that the opponents condition on a randomization device that is correlated in arbitrary ways with the unknown state of the world. This removes the distinction between the two type spaces we presented in the introduction. Dekel et al. (2005a) show more generally that correlated interim rationalizability depends only on conventional hierarchies of belief.

### 7.3 Upper hemicontinuity

The literature has had some interest in finding the “right” topology on the universal type space to capture similarity of types with respect to their strategic behavior.<sup>22</sup> One requirement of such a topology is that the rationalizable correspondence should be upper hemicontinuous. Our results shed some light on this requirement. We have shown (Proposition 3) that a sufficient condition for upper hemicontinuity is that the topology be fine enough so that conditional beliefs are continuous. Here we present an example where this strong form of continuity fails and the correspondence fails to be upper hemicontinuous.

Suppose that  $\Omega = \{-1, 1\}$  and consider the game from the introduction and the following type space:  $T_i = [-1, 1]$  and beliefs defined by

$$\mu_i(t_i)(t_{-i}, \omega) = \begin{cases} \frac{1}{2} & \text{if } t_{-i} \in \{t_i, -t_i\} \text{ and } \text{sign}[t_i \cdot \omega] = \text{sign}[t_{-i}] \\ 0 & \text{otherwise.} \end{cases}$$

It is easy to verify that these beliefs are generated by a common prior and are continuous as a function of  $t_i$ . However conditional beliefs over  $\Omega$  exhibit a discontinuity at

<sup>22</sup>See Monderer and Samet (1997), Kajii and Morris (1998), Morris (2002), Weinstein and Yildiz (2003), Weinstein and Yildiz (2004), and Dekel et al. (2005b).

$t_i = t_{-i} = 0$ . Indeed, if  $t_i = t_{-i} \neq 0$ , then  $t_i$  assigns probability 1 to state  $\omega = 1$  conditional on  $t_{-i}$ , but if  $t_i = t_{-i} = 0$ , then both states have equal conditional probability.

For  $t_i = t_{-i} \neq 0$ , the set  $\{-t_i, +t_i\} \times \{-t_{-i}, +t_{-i}\}$  is a belief-closed subspace that is isomorphic to the first type space from the introduction. Thus, all actions are rationalizable for these types. However, the belief-closed subspace  $\{0\} \times \{0\}$  is isomorphic to the second type space and hence action  $c_i$  is the unique rationalizable action for types  $t_i = t_{-i} = 0$ . Thus, for this finite game with a finite set of states of the world and continuous belief-mapping, the rationalizable correspondence is not upper hemicontinuous.

#### 7.4 Universal type space for the measurable case

Following the literature, we say that a type space  $U$  over a space of basic uncertainty  $X$  is *universal* among type spaces with property  $Y$  if for every such type space there is a unique mapping into  $U$  that preserves beliefs. Mertens et al. (1994) showed that there exists a universal type space for all continuous type spaces, assuming  $X$  is a Polish space. On the other hand, Heifetz and Samet (1999) showed that there is no universal type space for measurable (not necessarily continuous) type spaces when  $X$  is assumed only to be measurable. Our Theorem 1, whose proof is only a slight adaptation of the one in Mertens et al. (1994), is a positive result for an in-between case. It shows the existence of a universal measurable type space under the assumption that  $X$  is Polish. This may be comforting because while it is questionable to assume some particular structure on a type space (as it is nothing more than an artificial modeling construct), there may be good reason to assume structure on the physical world  $X$ .

#### 7.5 Type-space independent description of $U(\Delta\Omega)$ .

We sought in this paper a characterization of interactive beliefs that is necessary and sufficient to identify rationalizable behavior. We wanted this characterization to be describable in natural language terms that are independent of any modeling device such as a type space. Here we observe that just as with conventional hierarchies of belief, a player's  $\Delta$ -hierarchy can be elicited by a sequence of questions concerning beliefs and higher-order beliefs. One way to pose these is to begin with the following basic question: "Given any set of actions for your opponent, what is the set of distributions over actions and states of the world that you could face?" The answer to this question reveals the player's first-order  $\Delta$ -belief. After receiving an answer to this question, the remainder of the hierarchy is elicited analogously as with conventional hierarchies: "Given any set of actions for your opponent, what would you consider possible within the set of joint probability distributions over the state, your opponent's action, and your opponent's answer to the previous question?"

#### 7.6 Other solution concepts

The universal type space  $U(\Omega)$  is the smallest (i.e. most parsimonious) type space that is complete in terms of conventional hierarchies of belief. We showed however that in order to analyze games using the concept of rationalizability, we need the larger space

$U(\Delta\Omega)$ . Can we conclude that  $U(\Delta\Omega)$  is the right universal type space? Certainly not: depending on the application the modeler will need a different property of players' information, and this will correspond to potentially different type spaces that are "universal" with respect to this property. For example, we may carry out a similar investigation for other solution concepts such as Bayesian Nash equilibrium (BNE). It is easy to construct examples to show that even  $U(\Delta\Omega)$  is not rich enough for BNE: two types can have the same  $\Delta$ -hierarchy and yet play differently in BNE.<sup>23</sup> This raises a few questions. First, is there a primitive characteristic of information that is sufficient to characterize BNE behavior? Second, is there a single universal type space that is sufficient to characterize behavior in all "standard" solution concepts?

The first question is a topic for further research. For the second question we have a partial answer. Suppose there exists a type space  $P(\Omega)$  with the property that *any* type space can be mapped by an *injective*, belief-preserving type mapping into  $P(\Omega)$ . Then  $P(\Omega)$  would be the type space we seek (although it may not be the most parsimonious). Effectively,  $P(\Omega)$  would include a complete replica of every type space.

In general such a type space will not exist. However, if we restrict attention to the category of *Polish* type spaces, then existence can be established.

CLAIM. *There is a Polish type space over  $\Omega$ ,  $P(\Omega) \in \mathcal{T}_w(\Omega)$ , such that for any Polish type space  $T \in \mathcal{T}_w(\Omega)$ , there is an injective, belief-preserving type mapping  $\pi^T : T \rightarrow P(\Omega)$ .*

Details are available from the authors.

## APPENDIX

### A. SKETCH OF PROOF OF THEOREM 1

The proof is a minor adaptation of Mertens et al. (1994). We mention here only the changes necessary to accommodate type spaces with weakly measurable (i.e. not necessarily continuous) belief mappings. The remainder is identical.

We first recall the definition of a projective limit. Suppose that we have a sequence of Polish spaces  $\{P_n\}_{n \geq 0}$  together with a sequence of continuous mappings  $h_n : P_{n+1} \rightarrow P_n$ . Then the projective limit is a subset  $P \subseteq P_0 \times P_1 \times \dots$  of all points  $(p_0, p_1, \dots)$  such that  $h_n(p_{n+1}) = p_n$  for all  $n$ . Such a set is a Polish space in the product topology. The induced projection mappings  $H_n : P \rightarrow P_n$  are continuous.

Mertens et al. (1994) construct the universal type space  $U(X)$  for Polish  $X$  as the projective limit of hierarchies of beliefs. Hierarchies of beliefs of order  $n$  are defined inductively as follows

$$\begin{aligned} U_i^0 &= \{*\} \\ U_i^n &= \Delta(X \times U_{-i}^{n-1}). \end{aligned}$$

<sup>23</sup>The simplest example is a game of complete information, i.e. when  $\Omega$  is a singleton. Then any type in any type space over  $\Omega$  has the same (degenerate)  $\Delta$ -hierarchy. But the set of BNE outcomes expands as the type space expands due to added possibilities for correlation *in behavior*.

The mappings  $h_i^n : U_i^{n+1} \rightarrow U_i^n$  are defined inductively by

$$\begin{aligned} h_i^0(u_i) &= * \text{ for all } u_i \in U_i^1 \\ h_i^n &= \Delta(\text{id}_X \times h_{-i}^{n-1}). \end{aligned}$$

We take  $U_i(X)$  to be the projective limit of the system  $(U_i^n, h_i^n)$ . The induced projection mappings  $H_i^n : U_i(X) \rightarrow U_i^n$  are continuous and satisfy

$$H_i^n = h_i^n \circ H_i^{n+1}.$$

Mertens et al. (1994) show that one can construct a continuous belief mapping  $\mu_i^* : U_i(X) \rightarrow \Delta(X \times U_{-i}(X))$  with the property that for any  $u_i \in U_i(X)$ ,

$$H_i^n(u_i) = \Delta(\text{id}_X \times H_{-i}^{n-1})(\mu_i^*(u_i)). \quad (12)$$

This property implies that the mapping  $\mu_i^*$  is a homeomorphism between  $U_i(X)$  and  $\Delta(X \times U_{-i}(X))$ . (Observe that for any  $u_i \in U_i(X)$ , there is exactly one belief  $\mu_i^*$  such that (12) is satisfied: by the definition of the projective limit, two different measures on  $\Delta(X \times U_{-i}(X))$  have to have different images under  $\Delta(\text{id}_X \times H_{-i}^{n-1})$  for at least one  $n$ . Similarly, for any two  $u_i \neq u'_i$ ,  $H_i^n(u_i) \neq H_i^n(u'_i)$  for at least one  $n$  and  $\mu_i^*(u_i) \neq \mu_i^*(u'_i)$ . The inverse mapping

$$\mu_i \rightarrow \{\Delta(\text{id}_X \times H_{-i}^0)(\mu_i), \dots, \Delta(\text{id}_X \times H_{-i}^{n-1})(\mu_i), \dots\}$$

is naturally continuous.)

Mertens et al. (1994) show that if the type space  $T$  is continuous (the belief mapping is continuous) then there exists a sequence of continuous  $u_i^{T,n} : T_i \rightarrow U_i^n$ ,  $h_i^n \circ u_i^{T,n+1} = u_i^{T,n}$ , that extend to the continuous mapping  $u_i^T : T_i \rightarrow U_i$ , which is the unique mapping that preserves beliefs. The difference in our case is that we are not able to guarantee the continuity of the mappings  $u_i^{T,n}$ . However, we show that weak measurability of the belief mapping assures that the maps  $u_i^{T,n}$  defined exactly in the same way as in Mertens et al. (1994) are measurable. Moreover, they generate pointwise converging measurable mappings  $u_i^{T,n,u_i^*} : T_i \rightarrow U_i$ , for some  $u_i^* \in U_i$ , which converge to the measurable mapping  $u_i^T : T_i \rightarrow U_i$ . For the same reasons, this mapping preserves beliefs.

Precisely, we use the following lemma.

LEMMA 7. *Let  $T \in \mathcal{T}_w(X)$  and suppose that there is a Polish space  $B$  and measurable mapping  $\phi_{-i} : T_{-i} \rightarrow B$ . Then the mapping  $\Phi_i : T_i \rightarrow \Delta(X \times B)$  defined by*

$$\Phi_i(t_i) = \Delta(\text{id}_X \times \phi_{-i})\mu_i(t_i)$$

*is measurable.*

PROOF. We need to check whether for any measurable function  $f : X \times B \rightarrow \mathbf{R}$  the sets  $\{t_i : \Phi_i(t_i)[f] < 0\}$  are measurable. But

$$\{t_i : \Phi_i(t_i)[f] < 0\} = \{t_i : \mu_i(t_i)[f(x, \phi_{-i}(t_{-i}))] < 0\}$$

and the last set is measurable by the definition of weak measurability of  $\mu_i(t_i)$ .  $\square$

Choose now an arbitrary  $u_i^* \in U_i$  and construct mappings  $u_i^{T,n} : T_i \rightarrow U_i$  with  $u_i^{T,n}(t_i) = u_i^*$  and later inductively

$$u_i^{T,n+1} = \Delta(\text{id}_X \times u_{-i}^n) \mu_i(t_i).$$

By the lemma, each of these mappings is measurable. Moreover, they converge pointwise to the mapping  $u_i^T$ , which is also measurable (as a pointwise limit of measurable mappings). Exactly as in [Mertens et al. \(1994\)](#), we can verify that it is the unique belief-preserving mapping.

#### B. PROOF OF [LEMMA 5](#)

We begin with some notation. For a fixed game  $G = (A_j, u_j) \in \mathcal{G}$ , define the following set of continuous bounded functions  $f : \Delta\Omega \times \mathcal{K}A_{-i} \rightarrow [0, \infty)$ :

$$\mathcal{F}^G = \left\{ \begin{array}{l} f(\tau, K) = \max_{k=1, \dots, N_1} \sup_{a_1, \dots, a_{N_2} \in K} \tau[\psi(k, a_1, \dots, a_{N_2}, \omega)] : \\ \text{for some natural } N_1, N_2 \text{ and} \\ \text{continuous bounded function } \psi : \{1, \dots, N_1\} \times A_{-i}^{N_2} \times \Omega \rightarrow [0, \infty) \end{array} \right\}.$$

On the first coordinate  $\tau \in \Delta\Omega$ , the functions  $f \in \mathcal{F}^G$  are “piecewise linear” and convex. On the second coordinate  $K \in \mathcal{K}A_{-i}$ , they are set-increasing: for any two sets  $K \subseteq K'$ ,  $K, K' \in \mathcal{K}A_{-i}$ ,  $f(\tau, K) \leq f(\tau, K')$ . Next, define the set of differences of functions from  $\mathcal{F}^G$

$$\mathcal{L}^G = \{f - g : f, g \in \mathcal{F}^G\} \subseteq C(\Delta\Omega \times \mathcal{K}A_{-i}).$$

The following stronger version of [Lemma 5](#) will be used also in the proof of [Proposition 5](#).

**LEMMA 5'.** *For any game  $G$  the collection of sets  $\{\mu : \mu[f] < 0\} \subseteq \Delta(\Delta\Omega \times \mathcal{K}A_{-i})$  for  $f \in \mathcal{L}^G$  generates the weak\* topology on  $\Delta(\Delta\Omega \times \mathcal{K}A_{-i})$ . In particular, for any  $\mu, \mu' \in \Delta(\Delta\Omega \times \mathcal{K}A_{-i})$ ,  $\mu \neq \mu'$  there is a function  $f^G \in \mathcal{F}^G$ , such that*

$$\mu[f^G(\tau, A)] \neq \mu'[f^G(\tau, A)].$$

The following notation and preliminary results will be used in the proof. Let  $\mathbf{H}$  denote the Hilbert cube  $[0, 1]^{\mathbb{N}}$ . Since  $\Omega$  is Polish, there is a countable sequence of functions  $h_k^* : \Omega \rightarrow [0, 1]$  that define a compatible metric on  $\Delta\Omega$ ,

$$d_{\Delta\Omega}(\tau, \tau') = \sum_{k=1}^{\infty} \frac{1}{2^k} \left| \tau[h_k^*] - \tau'[h_k^*] \right|,$$

and a mapping  $H : \Delta\Omega \rightarrow \mathbf{H}$  with

$$H(\tau) = (\tau[h_1^*], \tau[h_2^*], \dots).$$

The mapping  $H$  embeds  $\Delta\Omega$  (with the weak\* topology) in the Hilbert cube (with the product topology). Suppose we have a family  $\mathcal{F}$  of continuous functions  $f : \mathbf{H} \times \mathcal{K}A_{-i} \rightarrow \mathbf{R}$  such that the collection of sets

$$\{\mu : \mu[f(h, A)] < 0\} \subseteq \Delta(\mathbf{H} \times \mathcal{K}A_{-i})$$

is a subbase for the weak\* topology on  $\Delta(\mathbf{H} \times \mathcal{H}A_{-i})$  (in which case we say that  $\mathcal{F}$  generates the topology). Then, because  $H$  is an embedding, it will follow that the corresponding family  $\mathcal{F}' \subset C(\Delta\Omega \times \mathcal{H}A_{-i})$  with

$$\mathcal{F}' = \{f' : f'(\tau, K) = f(H(\tau), K) \text{ for some } f \in \mathcal{F}\}$$

generates the topology on  $\Delta(\Delta\Omega \times \mathcal{H}A_{-i})$ . The strategy of proof is to find such an  $\mathcal{F}$  so that the corresponding  $\mathcal{F}'$  is included in  $\mathcal{L}^G$ .

For each natural number  $n$ , define the following set of continuous functions  $f : [0, 1]^n \times \mathcal{H}A_{-i} \rightarrow [0, \infty)$ :

$$\mathcal{F}_n = \left\{ \begin{array}{l} f(z_1, \dots, z_n, A) = \max_{k=1, \dots, N_1} \sup_{a_1, \dots, a_{N_2} \in A} \eta^k(a_1, \dots, a_{N_2}) \cdot z : \\ \text{for some natural numbers } N_1, N_2 \text{ and} \\ \text{some continuous bounded functions } \eta^1, \dots, \eta^{N_1} : A^{N_2} \rightarrow [0, 1]^n \end{array} \right\}$$

where  $\cdot$  is a scalar product of two vectors in  $\mathbf{R}^n$ . Next, define the set of differences of functions from  $\mathcal{F}_n$

$$\mathcal{L}_n = \{f - g : f, g \in \mathcal{F}_n\} \subseteq C([0, 1]^n \times \mathcal{H}A_{-i}).$$

We have a lemma.

LEMMA 8. *The set  $\mathcal{L}_n$  is uniformly dense in the set  $C([0, 1]^n \times \mathcal{H}A_{-i})$ .*

PROOF. This is a standard argument applying the lattice version of the Stone-Weierstrass theorem (see [Aliprantis and Border \(1994, Theorem 7.45\)](#)). We need to verify that  $\mathcal{L}_n$ :

- is closed under scalar multiplication: If  $(f - g) \in \mathcal{L}_n$ , then for any  $\lambda \in \mathbf{R}$ ,  $\lambda(f - g) \in \mathcal{L}_n$  as well;
- contains a constant function:  $\mathbf{1} \in \mathcal{L}_n$ ;
- is closed under finite sums: first note that for any  $f, g \in \mathcal{F}_n$ ,  $z = (z_1, \dots, z_m) \in [0, 1]^n$  and  $K \in \mathcal{H}A_{-i}$

$$\begin{aligned} f(z, K) + g(z, K) &= \max_{k=1, \dots, N_1^f} \sup_{a_1, \dots, a_{N_2^f} \in K} \eta^k(a_1, \dots, a_{N_2^f}) \cdot z \\ &\quad + \max_{l=1, \dots, N_1^g} \sup_{a_1, \dots, a_{N_2^g} \in K} \eta^l(a_1, \dots, a_{N_2^g}) \cdot z \\ &= \max_{\substack{k=1, \dots, N_f \\ l=1, \dots, N_g}} \sup_{\substack{a_1, \dots, a_{N_2^f} \in K \\ a_1, \dots, a_{N_2^g} \in K}} \left( \eta^k(a_1, \dots, a_{N_2^f}) + \eta^l(a_1, \dots, a_{N_2^g}) \right) \cdot z \end{aligned}$$

so that  $f + g \in \mathcal{F}_n$ . But this implies that for any  $(f - g), (f' - g') \in \mathcal{L}_n$  we also have  $(f + f') - (g + g') \in \mathcal{L}_n$ ;

- is closed with respect to taking the maximum of two functions: for any  $f, g \in \mathcal{F}_n$ ,  $z \in [0, 1]^n$ ,  $K \in \mathcal{K}A_{-i}$

$$\begin{aligned} \max\{f(z, A), g(z, A)\} &= \max \left\{ \begin{array}{l} \max_{k=1, \dots, N_1^f} \sup_{a_1, \dots, a_{N_2^f} \in K} \eta^k(a_1, \dots, a_{N_2^f}) \cdot z, \\ \max_{l=1, \dots, N_1^g} \sup_{a_1, \dots, a_{N_2^g} \in K} \nu^l(a_1, \dots, a_{N_2^g}) \cdot z \end{array} \right\} \\ &= \max_{k=1, \dots, N_1^f + N_1^g} \sup_{a_1, \dots, a_{N_2^f + N_2^g} \in K} \varphi^k(a_1, \dots, a_{N_2^f + N_2^g}) \cdot z \end{aligned}$$

where

$$\begin{aligned} \varphi^k(a_1, \dots, a_{N_2^f + N_2^g}) &= \eta^k(a_1, \dots, a_{N_2^f}) \text{ for } k \leq N_1^f \text{ and} \\ \varphi^k(a_1, \dots, a_{N_2^f + N_2^g}) &= \nu^{k - N_1^f}(a_{N_2^f + 1}, \dots, a_{N_2^f + N_2^g}) \text{ for } N_1^f < k \leq N_1^f + N_1^g. \end{aligned}$$

Then  $h = \max(f, g) \in \mathcal{F}_n$ . Together with the fact that

$$\max\{f - g, f' - g'\} = \max\{f + g', f' + g\} - (g + g')$$

and the previous point, it implies that  $\max\{f - g, f' - g'\} \in \mathcal{L}_n$  for any  $f - g, f' - g' \in \mathcal{L}_n$ ;

- separates points: for any  $z, z' \in [0, 1]^n$ ,  $z \neq z'$ , there is vector  $\eta \in \mathbf{R}^n$ , such that  $\eta \cdot z \neq \eta \cdot z'$ . Similarly, for any  $K, K' \in \mathcal{K}A_{-i}$ ,  $K \neq K'$ , there is a continuous function  $s : A_{-i} \rightarrow [0, 1]$ , such that

$$f(A) = \sup_{a \in K} s(a) = 1 > 0 = \sup_{a \in K'} s(a) = f(A'). \quad \square$$

Finally we can prove **Lemma 5'**. Any  $f \in C([0, 1]^n \times \mathcal{K}A_{-i})$  can be viewed as an element  $f' \in C(\mathbf{H} \times \mathcal{K}A_{-i})$  by writing  $f'(h, K) = f(h_1, \dots, h_n, K)$ . By the Stone-Weierstrass theorem (algebraic version, see **Aliprantis and Border (1994, Theorem 7.46)**) the subset  $\cup_n C([0, 1]^n \times \mathcal{K}A_{-i})$  is uniformly dense in  $C(\mathbf{H} \times \mathcal{K}A_{-i})$ . By **Lemma 8**, the family  $\mathcal{L}_n$  is uniformly dense in  $C([0, 1]^n \times \mathcal{K}A_{-i})$ . Thus  $\cup_n \mathcal{L}_n$  is uniformly dense in  $\cup_n C([0, 1]^n \times \mathcal{K}A_{-i})$  and hence in  $C(\mathbf{H} \times \mathcal{K}A_{-i})$ . We conclude that the family  $\cup_n \mathcal{L}_n$  generates the topology on  $\Delta(\mathbf{H} \times \mathcal{K}A_{-i})$  (see **Aliprantis and Border (1994, Theorem 12.2)**).

The proof is now completed by showing that each  $f \in \mathcal{L}_n$  corresponds to a function  $f'$  belonging to  $\mathcal{L}^G$  by the formula  $f'(\tau, K) = f(H(\tau), K)$ . By the linear structure of  $\mathcal{L}^G$  it suffices to show that for each  $g \in \mathcal{F}_n$ , the composition  $g \circ H : \Delta\Omega \times \mathcal{K}A_{-i}^G \rightarrow \mathbf{R}$  belongs to  $\mathcal{F}^G$ . This is verified by noting that

$$\begin{aligned} (g \circ H)(\tau, K) &= \max_{k=1, \dots, N_1} \sup_{a_1, \dots, a_{N_2} \in K} \eta^k(a_1, \dots, a_{N_2}) \cdot (\tau[h_1^*(\omega)], \dots, \tau[h_n^*(\omega)]) \\ &= \max_{k=1, \dots, N_1} \sup_{a_1, \dots, a_{N_2} \in A} \tau \left[ \sum_{m=1}^n h_m^*(\omega) \eta_m^k(a_1, \dots, a_{N_2}) \right]. \end{aligned}$$

Since  $\sum_{m=1}^n h_m^*(\omega) \eta_m^k(a_1, \dots, a_{N_2})$  is a bounded continuous function mapping  $\{1, \dots, N_1\} \times (A_{-i}^G)^{N_2} \times \Omega$  into  $\mathbf{R}$ , we conclude that  $g \circ H$  is an element of  $\mathcal{F}^G$ .  $\square$

## C. RESULTS ON MEASURABILITY

## C.1 Proof of Lemma 1

We need the following result.

LEMMA 9. *Let  $A$  and  $B$  be measurable spaces and  $g : A \times B \rightarrow [0, 1]$  a jointly measurable map. If  $m : A \rightarrow \Delta(B)$  is measurable, then the map  $L^g : A \rightarrow \mathbf{R}$  defined by  $L^g(a) = m(a)[g(a, \cdot)]$  is measurable.*

PROOF. There exists a sequence of simple functions  $g_n : A \times B \rightarrow \mathbf{R}$  such that  $g_n \uparrow g$  and by the definition of the Lebesgue integral, for any probability measure  $\nu \in \Delta(A \times B)$ ,

$$\nu[g_n] \rightarrow \nu[g].$$

In particular, for any given  $a \in A$ , if  $\nu$  is the measure whose marginal on  $B$  is  $m(a)$  and whose marginal on  $A$  is  $\delta_a$ ,

$$L^{g_n}(a) = \nu[g_n] \rightarrow \nu[g] = L^g(a).$$

Thus, if we can show that  $L^f$  is measurable for all simple functions  $f$ , then we will have shown that  $L^g$  is measurable as the pointwise limit of measurable mappings.

First consider  $f = 1_{\alpha \times \beta}$  for  $\alpha \subset A$  and  $\beta \subset B$  (measurable). We have  $L^f(a) = 1_\alpha(a) \cdot m(a)(\beta)$  which is measurable since  $m$  was assumed to be measurable. Thus,  $L^f$  is measurable for all  $f$  that are indicators of product sets. Now for any finite  $k$ , let  $\alpha_1, \dots, \alpha_k$  and  $\beta_1, \dots, \beta_k$  be measurable subsets of  $A$  and  $B$  respectively and note that for  $f = \prod_{l=1}^k 1_{\alpha_l \times \beta_l}$ ,

$$L^f(a) = \prod_l 1_{\alpha_l} \cdot m(a)(\cap_l \beta_l)$$

is measurable. Thus if  $f = 1_{\cap_l (\alpha_l \times \beta_l)} = \prod_l 1_{\alpha_l \times \beta_l}$ , then  $L^f$  is measurable, and if

$$f = 1_{\cup_l (\alpha_l \times \beta_l)} = \sum_l 1_{\alpha_l \times \beta_l} - \sum_{S \subset \{1, \dots, k\}} (|S| - 1) \prod_{l \in S} 1_{\alpha_l \times \beta_l}$$

then  $L^f$  is measurable as a linear combination of measurable functions. Note also that  $L^{1-E} = L^{1-1_E} = 1 - L^{1_E}$ . Thus  $L^f$  is measurable for all indicator functions  $f$  of sets in the algebra generated by the product sets.

Now consider any sequence  $X_n$  with  $X_n \subset A \times B$ ,  $X_n \subset X_{n+1}$  for all  $n$  and  $\cup_n X_n = X$ . The corresponding sequence  $L^{1_{X_n}}$  is an increasing sequence of maps converging pointwise to  $L^{1_X}$ . Thus if  $L^{1_{X_n}}$  are measurable for all  $n$ , so is  $L^{1_X}$ . It follows that the collection of sets  $X$  such that  $L^{1_X}$  is measurable is a monotone class. Since it includes the algebra generated by the product sets, by the monotone class lemma it includes the corresponding  $\sigma$ -algebra, i.e. the product sigma-algebra on  $A \times B$ . Finally, since any simple function  $f : A \times B \rightarrow \mathbf{R}$  has the form

$$f(a, b) = \sum_{j=1}^k c_j 1_{X_j}(a, b)$$

for some coefficients  $c_j$  and measurable sets  $X_j \subset A \times B$ , any such  $L^f$  is measurable as a linear combination of measurable functions.  $\square$

PROOF OF LEMMA 1. We must show that for any measurable  $f : \Delta\Omega \times T_{-i} \rightarrow \mathbf{R}$ , the mapping

$$t_i \rightarrow \mu_i^\Delta(t_i)[f]$$

is measurable. Define  $g(t_i, t_{-i}) = f(\beta_i(t_i, t_{-i}), t_{-i})$ . Note that  $g$  is jointly measurable and

$$\mu_i^\Delta(t_i)[f] = \mu_i(t_i)[g(t_i, \cdot)].$$

Now apply Lemma 9.  $\square$

### C.2 Proof of Lemma 4 and Proposition 5

PROOF OF LEMMA 4. By the monotone class theorem and the choice of topology on  $\mathcal{R}_i$ , we need to check that for any finite number of games  $G^1, \dots, G^k \in \mathcal{G}$  and open set  $KA \subseteq \mathcal{K}A_i^{G^1} \times \dots \times \mathcal{K}A_i^{G^k}$ , the set

$$\left(\rho_i^T\right)^{-1}(KA \times \prod_{G \in \mathcal{G}, G \neq G_i \text{ for } i=1, \dots, k} \mathcal{K}A_i^G) = \left(R_i^{G^1, T} \times \dots \times R_i^{G^k, T}\right)^{-1}(KA)$$

is measurable. Consider the product game  $G = G^1 \times \dots \times G^k = (A_i^G, u_i^G)$ . Observe that  $KA$  is an open subset  $KA \subseteq \mathcal{K}A_i^G$  and  $R_i^{G^1, T} \times \dots \times R_i^{G^k, T} = R_i^{G, T}$ . Now by Proposition 2, the set  $\left(R_i^{G, T}\right)^{-1}(KA)$  is measurable in  $T_i$ .  $\square$

For the proof of Proposition 5 we need a lemma identifying some measurable subsets of rationalizable rules.

LEMMA 10. For any game  $G = (A_i, u_i) \in \mathcal{G}$ , for any closed subset  $A' \subseteq A_i$ , the subset of rationalizable rules  $\{r_i \in \mathcal{R}_i : r_i(G) \subseteq A'\}$  is closed in  $\mathcal{R}_i$ .

PROOF. For closed  $A'$ , the set  $\mathcal{K}A' = \{K \in \mathcal{K}A_i : K \subseteq A'\}$  is closed in  $\mathcal{K}A_i$  (see Aliprantis and Border (1994, Theorem 3.63)). Thus, by the definition of the product topology on  $\mathcal{R}_i$ ,

$$\{r_i \in \mathcal{R}_i : r_i(G) \subseteq A'\} = \{r_i \in \mathcal{R}_i : r_i(G) \in \mathcal{K}A'\}$$

is closed.  $\square$

PROOF OF PROPOSITION 5. Let  $\mathcal{D}$  be the collection of all subsets  $E$  of  $\Delta\Omega \times \mathcal{R}_{-i}$  such that the mapping

$$r_i \rightarrow \mu_i^{\mathcal{R}}(r_i)[1_E]$$

is Borel measurable. We prove the proposition by showing that  $\mathcal{D}$  includes all measurable sets.

Let  $\mathcal{P}^*(\mathcal{G})$  be the collection of all finite subsets of  $\mathcal{G}$ , and define

$$\mathcal{C} = \bigcup_{\Gamma \in \mathcal{P}^*(\mathcal{G})} \left\{ V = V^0 \times \prod_{G \in \Gamma} V^G : \begin{array}{l} V^0 \subset \Delta\Omega, V^G \subset \mathcal{H}A_{-i}^G \text{ are measurable} \\ \text{and } V^G = \mathcal{H}A_{-i}^G \text{ for all } G \notin \Gamma \end{array} \right\}.$$

Note that  $\mathcal{C}$  is an algebra (closed under taking complements, finite intersections and unions) and generates the product topology and hence the  $\sigma$ -algebra on  $\Delta\Omega \times \mathcal{R}_{-i}$ . We first show that  $\mathcal{C} \subset \mathcal{D}$ .

Consider any element  $V \in \mathcal{C}$  for which  $\Gamma = \{G\}$  is a singleton. If we can show that  $\{r_i : \text{marg}_{\Delta\Omega \times \mathcal{H}A_{-i}^G} \mu_i^{\mathcal{R}}(r_i)(V^0 \times V^G) \in I\}$  is a measurable set of rules for every interval  $I \subset [0, 1]$ , it will follow that  $V \in \mathcal{D}$ . Since  $\{\mu \in \Delta(\Delta\Omega \times \mathcal{H}A_{-i}^G) : \mu(V^0 \times V^G) \in I\}$  is a measurable set, it suffices to show that  $\text{marg}_{\Delta\Omega \times \mathcal{H}A_{-i}^G} \mu_i^{\mathcal{R}} : \mathcal{R}_i \rightarrow \Delta(\Delta\Omega \times \mathcal{H}A_{-i}^G)$  is a measurable mapping.

By **Lemma 5'** there is a base for the Borel  $\sigma$ -algebra on  $\Delta\Omega \times \mathcal{H}A_{-i}^G$  consisting of sets of the form

$$W^f = \{\mu : \mu[f] < 0\}$$

for all functions  $f \in \mathcal{L}^G$ . In the course of the proof of **Theorem 3**, we showed that for any  $f^G \in \mathcal{F}^G$  and  $s \in \mathbf{R}$ , there is a game  $\tilde{G}$  and a closed subset of actions  $\tilde{A} \in A_{-i}$  such that for any type space  $T \in \mathcal{T}_s(\Omega)$  the following two sets of types  $t_i \in T_i$  are equal

$$\{t_i : R_i^{\tilde{G}, T}(t_i) \subseteq \tilde{A}\} = \{t_i : \varpi^{G, T}(t_i)[f^G] < s\}.$$

This implies that the following two sets of rationalizable rules are equal

$$\{r_i \in \mathcal{R}_i : r_i(\tilde{G}) \subseteq \tilde{A}\} = \left\{ r_i \in \mathcal{R}_i : \text{marg}_{\Delta\Omega \times \mathcal{H}A_{-i}^G} \mu_i^{\mathcal{R}}(r_i)[f^G] < s \right\}.$$

By **Lemma 10**, the first set is measurable. Now since  $f = f_1^G - f_2^G$  for some  $f_1^G, f_2^G \in \mathcal{F}^G$ , we have

$$W^f = \bigcup_{\substack{x_1, x_2 \in \mathbb{Q}, m=1,2 \\ x_1 + x_2 < 0}} \bigcap \{ \mu : \mu[f_m^G] < x_m \},$$

where  $\mathbb{Q}$  is the set of rational numbers. Then

$$\left[ \text{marg}_{\Delta\Omega \times \mathcal{H}A_{-i}^G} \mu_i^{\mathcal{R}} \right]^{-1} (W^f) = \bigcup_{\substack{x_1, x_2 \in \mathbb{Q}, m=1,2 \\ x_1 + x_2 < 0}} \bigcap \left\{ r_i \in \mathcal{R}_i : \text{marg}_{\Delta\Omega \times \mathcal{H}A_{-i}^G} \mu_i^{\mathcal{R}}(r_i)[f_m^G] < x_m \right\}$$

is measurable as countable union of finite intersections of measurable sets. Therefore the inverse image of every set in a base for the sigma-algebra is measurable and this implies that  $\text{marg}_{\Delta\Omega \times \mathcal{H}A_{-i}^G} \mu_i^{\mathcal{R}}$  is measurable. (See **Aliprantis and Border (1994, Lemma 8.16)**.) Now consider an element  $V \in \mathcal{C}$  for which  $\Gamma$  is an arbitrary finite set. Consider the

product game  $\tilde{G} = \prod_{\Gamma} G$ , where the product set  $V^{\tilde{G}} = \prod_{G \in \Gamma} V^G$  is a measurable subset of  $\mathcal{H}A_{-i}^{\tilde{G}}$ . By the product structure, for any rationalizable rule  $r_{-i}(\tilde{G}) = \prod_{G \in \Gamma} r_{-i}(G)$ . Thus if we define  $V' = V^0 \times V^{\tilde{G}} \times \prod_{G \notin \Gamma} \mathcal{H}A_{-i}^{G'}$ , we have  $1_{V'} = 1_V$ , and we have already shown that  $V'$  belongs to  $D$ .

We have shown  $\mathcal{C} \subset \mathcal{D}$ . Now consider any sequence of measurable subsets  $E_n \subset \Delta\Omega \times \mathcal{R}_{-i}$  such that  $E_n \subset E_{n+1}$ ,  $E \in \mathcal{D}$ , and let  $E = \cup E_n$ . The sequence of indicator functions  $1_{E_n}$  increases pointwise to  $1_E$ . By countable additivity,  $\mu_i^{\mathcal{R}}(r_i)(E) = \lim \mu_i^{\mathcal{R}}(r_i)(E_n)$ , and hence for any open interval  $I$ ,

$$\{r_i : \mu_i^{\mathcal{R}}(r_i)[1_E] \in I\} = \bigcup_n \bigcap_{m > n} \{r_i : \mu_i^{\mathcal{R}}(r_i)[1_{E_n}] \in I\}$$

which is measurable. Thus  $E \in \mathcal{D}$  and  $\mathcal{D}$  is a monotone class that includes the algebra  $\mathcal{C}$ . By the monotone class lemma,  $\mathcal{D}$  includes all Borel sets.  $\square$

#### D. PROOF OF LEMMA 6

PROOF. We begin with a simple observation: For any belief-preserving type mapping  $\psi : S \rightarrow T$ ,  $S, T \in \mathcal{T}_w(X)$ , for any type  $s_i \in S_i$ , and for any measurable subsets  $B, C \subseteq T_{-i} \times X$  we have

$$\begin{aligned} \mu^T(\psi(s_i))(B) &= \mu^S(s_i)\left([\psi_{-i} \times \text{id}_X]^{-1}(B)\right), \\ \mu^T(\psi(s_i))(B \mid C) &= \mu^S(s_i)\left([\psi_{-i} \times \text{id}_X]^{-1}(B) \mid [\psi_{-i} \times \text{id}_X]^{-1}(C)\right). \end{aligned}$$

The first is just the definition of a belief-preserving type mapping, while the second is a consequence of the first.

*Part 1.* Suppose that  $\phi^\Delta : S \rightarrow T$  preserves beliefs. For any  $(s_i, \tau_i) \in S_i \times \Delta\Omega$ , define  $\phi_{(s_i, \tau_i)} : S_{-i} \times \Delta\Omega \rightarrow T_{-i} \times \Delta\Omega$  with

$$\phi_{(s_i, \tau_i)}(s_{-i}, \tau_{-i}) = (\phi^\Delta(s_{-i}), \tau_{-i}).$$

We check that (2) holds: for any subset  $B \subseteq T_{-i} \times \Delta\Omega$ ,

$$\begin{aligned} \text{marg}_{T_{-i} \times \Delta\Omega} \mu^{T \times \Delta\Omega}(\phi(s_i, \tau_i))(B') &= \text{marg}_{T_{-i} \times \Delta\Omega} \mu^{T \times \Delta\Omega}(\phi^\Delta(s_i), \tau)(B) \\ &= \mu^T(\phi^\Delta(s_i))((t_i, \tau_{-i}) \in B) \\ &= \mu^{S_i}(s_i)\left((\phi^\Delta \times \text{id}_X)^{-1}(B)\right) \\ &= \text{marg}_{S_{-i} \times \Delta\Omega} \mu^{S \times \Delta\Omega}(s_i, \tau_i)\left(\phi_{(s_i, \tau_i)}^{-1}(B)\right) \end{aligned}$$

where the third equality comes from the observation above and the fact that  $\phi^\Delta$  preserves beliefs. We verify immediately that  $\phi$  satisfies (3) and (4).

*Part 2.* If  $\phi$  preserves conditional beliefs, then there is for any  $(s_i, \tau_i) \in S \times \Delta\Omega$  a measurable mapping  $\phi_{(s_i, \tau_i)} : S_{-i} \times \Delta\Omega \rightarrow T_{-i} \times \Delta\Omega$  such that equations (2), (3), (4) hold.

Equation (4) and the definition of beliefs in the type space  $T \times \Delta\Omega$  imply that  $\text{proj}_{\Delta\Omega} \phi_{(s_i, \tau_i)}(s_{-i}, \tau_{-i}) = \tau_{-i}$ . The definition of beliefs in the space  $T \times \Delta\Omega$  together with the assumption that different types in  $T$  have different beliefs imply that  $\text{proj}_T \circ \phi_{(s_i, \tau_i)} = \text{proj}_T \circ \phi$ .

These two facts taken together lead us to the conclusion that the mapping  $\phi_{(s_i, \tau_i)}$  does not depend on  $(s_i, \tau_i)$ :

$$\begin{aligned} \phi_{(s_i, \tau_i)}(s_{-i}, \tau_{-i}) &= \left( \text{proj}_T \phi(s_{-i}, \tau_{-i}), \tau_{-i} \right) \\ &= \left( \left( \phi^\Delta \circ \text{proj}_S \right) \times \text{proj}_{\Delta\Omega} \right) (s_{-i}, \tau_{-i}) =: \hat{\phi}_{-i}(s_{-i}, \tau_{-i}). \end{aligned}$$

It is immediate to see that the type mapping  $\hat{\phi} : S \times \Delta\Omega \rightarrow T \times \Delta\Omega$  preserves conditional beliefs with a dual equal to itself,  $\hat{\phi}_i = \hat{\phi}_{(s_i, \tau_i)}$  for any  $(s_i, \tau_i)$ . A straightforward argument shows that  $\hat{\phi}$  preserves beliefs.

Then, for any  $s_i \in S_i$ , for any  $t_{-i} \in T_{-i}$ , any measurable  $B \subseteq \Delta\Omega$ ,

$$\begin{aligned} C_{\Delta\Omega} \mu_i^S(s_i) \left( (\phi_i^\Delta)^{-1}(t_{-i}) \right) (B) &= \mu_i^S(s_i) \left( (\phi_i^\Delta)^{-1}(t_{-i}) \times B \mid (\phi_i^\Delta)^{-1}(t_{-i}) \times \Delta\Omega \right) \\ &= \mu_i^{S \times \Delta\Omega}(s_i, \tau) \left\{ \hat{\phi}^{-1}(\{t_{-i}\} \times B) \times \Omega \mid \hat{\phi}^{-1}(\{t_{-i}\} \times \Delta\Omega) \times \Omega \right\} \\ &= \mu_i^{T \times \Delta\Omega}(\hat{\phi}(s_i, \tau)) \left\{ \{t_{-i}\} \times B \times \Omega \mid \{t_{-i}\} \times \Delta\Omega \times \Omega \right\} \\ &= \mu_i^T(\phi^\Delta(s_{-i})) (\{t_{-i}\} \times B \mid \{t_{-i}\} \times \Delta\Omega) \\ &= C_{\Delta\Omega} \mu_i^T(\phi_i^\Delta(s_i)) (t_{-i})(B), \end{aligned}$$

and  $\phi^\Delta$  preserves beliefs (the third equality comes from the second observation above).  $\square$

#### E. UPPER HEMICONTINUITY OF THE RATIONALIZABLE CORRESPONDENCE

This appendix contains the proof of **Proposition 3**. The proof builds on two lemmas.

**LEMMA 11.** *If  $T$  is  $\Delta$ -continuous and  $G$  is a compact, continuous game, then  $U_i(a_i, \sigma_{-i}^\Delta \mid t_i)$  is jointly continuous in  $(a_i, \sigma_{-i}^\Delta, t_i)$ .*

**PROOF.** First we show that  $\pi_i$  is jointly continuous. Pick  $M > \sup |u_i(a, \omega)|$  (recall that we assume that  $u_i$  is bounded for this class of games.) Let  $(a^k, t^k) \rightarrow (a^\infty, t^\infty) \in A \times T$ . The set  $\{t^k\}_{k=1}^\infty$  is compact, and so by  $\Delta$ -continuity, the corresponding family of measures  $\{\beta_i(t^k)\} \subset \Delta\Omega$  is also compact. Because  $\Omega$  is a Polish space, the family is tight, i.e. for every  $\varepsilon > 0$ , there exists a compact  $K^\varepsilon \subset \Omega$  such that  $\beta_i(t^k)(K^\varepsilon) > 1 - \varepsilon$  for all  $k \in \{1, \dots, \infty\}$ . We have

$$\left| \pi_i(a^k, t^k) - \pi_i(a^\infty, t^\infty) \right| \leq \left| \int_{K^\varepsilon} u_i(a^k, \cdot) d\beta_i(t^k) - \int_{K^\varepsilon} u_i(a^\infty, \cdot) d\beta_i(t^\infty) \right| + 2\varepsilon M.$$

Since  $K^\varepsilon$  is compact and  $u_i$  is jointly continuous,

$$\sup_{\omega \in K^\varepsilon} |u_i(a^k, \omega) - u_i(a^\infty, \omega)| \rightarrow 0$$

i.e., the sequence of functions  $u_i(a^k, \cdot) : K^\varepsilon \rightarrow \mathbf{R}$  converges uniformly to  $u_i(a^\infty, \cdot)$ . It follows that

$$\left| \int_{K^\varepsilon} u_i(a^k, \omega) d\beta_i(t^k) - \int_{K^\varepsilon} u_i(a^\infty, \omega) d\beta_i(t^\infty) \right| \rightarrow 0$$

and so

$$\begin{aligned} \limsup_k |\pi_i(a^k, t^k) - \pi_i(a^\infty, t^\infty)| \\ \leq \limsup_k \left| \int_{K^\varepsilon} u_i(a^k, \cdot) d\beta_i(t^k) - \int_{K^\varepsilon} u_i(a^\infty, \cdot) d\beta_i(t^\infty) \right| + 2\varepsilon M \\ = 2\varepsilon M \end{aligned}$$

and since  $\varepsilon$  was arbitrary, we have shown  $\pi_i(a^k, t^k) \rightarrow \pi_i(a^\infty, t^\infty)$ .

Having shown that  $\pi_i$  is jointly continuous, we can apply exactly the same argument to show that

$$U_i(a_i, \sigma_i^\Delta | t_i) = \int_{A_{-i} \times T_{-i}} \pi_i(a_i, \cdot, t_i, \cdot) d\sigma_i^\Delta(a_{-i}, t_{-i})$$

is jointly continuous as well.  $\square$

**LEMMA 12.** *Suppose  $T$  is a continuous type space and let  $S_i \subset T_i$  be a compact subset of types and  $B \subset A_{-i} \times T_{-i}$  is a closed assessment. Then the correspondence*

$$\Sigma^\Delta(B | \cdot) : S_i \rightrightarrows \Delta(T_{-i} \times A_{-i})$$

*has compact graph.*

**PROOF.** The proof uses the following result (see [Aliprantis and Border \(1994, Theorem 12.20\)](#)): If  $X$  is a Polish space, then a family  $\mathcal{F} \subset \Delta(X)$  has compact closure if and only if  $\mathcal{F}$  is tight, i.e. for every  $\varepsilon > 0$  there is a compact  $K \subset X$  such that  $\nu(K) > 1 - \varepsilon$  for all  $\nu \in \mathcal{F}$ .

Let  $(t_i^k, \sigma_{-i}^{\Delta, k})$  be a sequence from the graph. Since  $S_i$  is assumed compact,  $t_i^k$  has a subsequence converging to some  $t_i \in S_i$ . Now pick  $\varepsilon > 0$ . Since  $S_i$  is compact, by the continuity of  $T$  so is  $\mu_i^T(S_i) = \{\mu_i^T(t_i) : t_i \in S_i\}$  and by the continuity of marginals, so is  $\text{marg}_{T_{-i}} \mu_i^T(S_i)$ . By the above result, there is a compact  $K \subset T_{-i}$  such that  $\text{marg}_{T_{-i}} \mu_i^T(t_i)(K) > 1 - \varepsilon$  for all  $t_i \in S_i$ . Thus for any  $t_i \in S_i$  and  $\sigma_{-i}^\Delta \in \Sigma^\Delta(B | t_i)$ , we have  $\sigma_{-i}^{\Delta, t_i}(K \times A_{-i}) = \text{marg}_{T_{-i}} \mu_i^T(t_i)(K) > 1 - \varepsilon$ . Since  $K \times A_{-i}$  is compact, this shows that the family

$$\bigcup_{t_i \in S_i} \Sigma^\Delta(B | t_i)$$

is tight and therefore has compact closure.

Because  $\sigma_{-i}^{\Delta,k}$  is a sequence from this set, it has a convergent subsequence  $\sigma_{-i}^{\Delta,l} \rightarrow \sigma_{-i}^{\Delta}$ . The proof is concluded by showing that  $\sigma_{-i}^{\Delta} \in \Sigma^{\Delta}(B | t_i)$ .

1. By definition,  $\text{marg}_{T_{-i}} \sigma_{-i}^{\Delta,l} = \text{marg}_{T_{-i}} \mu_i^T(t_i^l)$ . Because the mapping  $T$  is continuous,  $\text{marg}_{T_{-i}} \mu_i^T(t_i^l) \rightarrow \text{marg}_{T_{-i}} \mu_i^T(t_i)$  and since  $\text{marg}_{T_{-i}} \sigma_{-i}^{\Delta,l} \rightarrow \text{marg}_{T_{-i}} \sigma_{-i}^{\Delta}$ , we have  $\text{marg}_{T_{-i}} \sigma_{-i}^{\Delta} = \text{marg}_{T_{-i}} \mu_i^T(t_i)$ .
2. The set of measures assigning probability 1 to a given closed set is closed. Since  $\sigma_{-i}^{\Delta,l}(B) = 1$  for all  $l$ , we have  $\sigma_{-i}^{\Delta}(B) = 1$  as well.  $\square$

For the next result, we need to define a procedure of iterative elimination of not-best responses. For any game  $G$  and any type space  $T$  define inductively assessments:

$$R_{0,i}^{G,T} = T_i \times A_i \text{ and for any } n > 0 \text{ let } R_{n,i}^{G,T}(t_i) = B(R_{n-1}^{G,T} | t_i).$$

**PROOF OF PROPOSITION 3.** We start by showing inductively that each  $R_{n,i}^{G,T}$  is a closed correspondence. It is obviously true for  $n = 0$ . Suppose now it is true for some arbitrary  $n$ , and let  $(t_i^k, a_i^k) \rightarrow (t_i, a_i)$  with  $(t_i^k, a_i^k) \in R_{n,i}^{G,T}$ . Then for each  $k$  there is a conjecture  $\sigma_{-i}^{\Delta,k} \in \Sigma^{\Delta}(R_{n-1,-i}^{G,T} | t_i^k)$  such that  $a_i^k$  is a best response of  $t_i^k$  against  $\sigma_{-i}^{\Delta,k}$ . By Lemma 12, there is a subsequence  $\sigma_{-i}^{\Delta,l}$  converging to  $\sigma_{-i}^{\Delta} \in \Sigma^{\Delta}(R_{n-1,-i}^{G,T} | t_i)$ .

Finally, by Lemma 11,

$$\begin{aligned} U_i(a_i^l, \sigma_{-i}^{\Delta,l} | t_i^l) &\rightarrow U_i(a_i, \sigma_{-i}^{\Delta} | t_i) \\ U_i(z_i, \sigma_{-i}^{\Delta,l} | t_i^l) &\rightarrow U_i(z_i, \sigma_{-i}^{\Delta} | t_i^l). \end{aligned}$$

Thus,  $a_i^l$  is a best response of  $t_i^l$  against  $\sigma_{-i}^{\Delta,l}$  for all  $l$  implies that  $a_i$  is a best response of  $t_i$  against  $\sigma_{-i}^{\Delta}$ . We have shown that  $(t_i, a_i) \in R_{n,i}^{G,T}$  and hence that the latter is closed.

The first step implies that the correspondence  $\mathcal{R}_i^{G,T}$  is closed as an intersection of closed sets. Now, we move to show that  $R_i^{G,T} = \mathcal{R}_i^{G,T}$ . Because  $R_i^{G,T}$  has the fixed-point property, we have  $R_i^{G,T} \subset R_{k,i}^{G,T}$  for every  $k$ , hence  $R_i^{G,T}$  is contained in  $\mathcal{R}_i^{G,T}$ . To show equality, therefore, it suffices to show that  $\mathcal{R}_i^{G,T}$  also has the fixed-point property and is therefore a subset of  $R_i^{G,T}$ . We need to show

$$\mathcal{R}_i^{G,T} = \{(t_i, a_i) : a_i \in B(\mathcal{R}_i^{G,T} | t_i)\}.$$

Suppose  $a_i$  is a best response of  $t_i$  against  $\sigma_{-i}^{\Delta} \in \Sigma^{\Delta}(R_k^{G,T} | t_i)$ . Then  $\sigma_{-i}^{\Delta} \in \Sigma^{\Delta}(R_k^{G,T} | t_i)$  for every  $k$  and hence  $(t_i, a_i) \in R_{k,i}^{G,T}$  for every  $k$ . This shows that  $(t_i, a_i) \in \mathcal{R}_i^{G,T}$ .

Suppose  $(t_i, a_i) \in \mathcal{R}_i^{G,T}$ , i.e.  $(t_i, a_i) \in R_{k,i}^{G,T}$  for every  $k$ . Then for each  $k$  there is a  $\sigma_{-i}^{\Delta,k} \in \Sigma^{\Delta}(R_k^{G,T} | t_i)$  such that  $a_i$  is a best response of  $t_i$  against  $\sigma_{-i}^{\Delta,k}$ . Since  $\mathcal{R}_i^{G,T}$  is closed, we can use Lemma 12 to extract a convergent subsequence  $\sigma_{-i}^{\Delta,l} \rightarrow \sigma_{-i}^{\Delta}$ . Repetition of the continuity argument above shows that  $a_i$  is a best response of  $t_i$  against  $\sigma_{-i}^{\Delta}$ . In order to conclude that  $(t_i, a_i)$  is a best response to some conjecture from  $\Sigma^{\Delta}(\mathcal{R}_i^{G,T} |$

$t_i$ ), it is enough to check that  $\sigma_{-i}^\Delta \in \Sigma^\Delta(\mathcal{R}^{G,T} \mid t_i)$ . Notice however that this is immediate consequence of two facts:

$$\begin{aligned} \text{marg } \sigma_{-i}^\Delta &= \lim_{T_i} \text{marg } \sigma_{-i}^{\Delta,l} = \text{marg } \mu_i(t_i) \\ \sigma_{-i}^\Delta(\mathcal{R}_{-i}^{G,T}) &= \lim_{n \rightarrow \infty} \sigma_{-i}^\Delta(R_{n,i}^{G,T}) = \lim_{n \rightarrow \infty} \lim_{l \rightarrow \infty} \sigma_{-i}^{\Delta,l}(R_{n,i}^{G,T}) = 1. \end{aligned}$$

The last equality follows from the definition of  $\sigma_{-i}^{\Delta,l} \in \Sigma^\Delta(R_{k(l)}^{G,T} \mid t_i)$  for some  $k(l)$ : for  $k \leq k(l)$  we have  $\sigma_{-i}^{\Delta,l}(R_{k,i}^{G,T}) = \sigma_{-i}^{\Delta,l}(R_{k(l),i}^{G,T}) = 1$ . This ends the proof.  $\square$

#### REFERENCES

- Aliprantis, Charalambos D. and Kim C. Border (1994), *Infinite Dimension Analysis: A Hitchhiker's Guide*. Springer-Verlag, New York. [38, 40, 43, 55, 56, 58, 59, 62]
- Aumann, Robert J. (1987), “Correlated equilibrium as an expression of Bayesian rationality.” *Econometrica*, 55, 1–18. [22]
- Aumann, Robert J. (1998), “Common priors: a reply to Gul.” *Econometrica*, 66, 929–938. [22, 23]
- Battigalli, Pierpaolo and Marciano Siniscalchi (1999), “Hierarchies of conditional beliefs and interactive epistemology in dynamic games.” *Journal of Economic Theory*, 88, 188–230. [29]
- Battigalli, Pierpaolo and Marciano Siniscalchi (2003), “Rationalization with incomplete information.” *Advances in Theoretical Economics*, 3. <http://www.bepress.com/bejte/advances/vol3/iss1/art3>. [23, 26]
- Bergemann, Dirk and Stephen Morris (2005), “Robust mechanism design.” *Econometrica*, 73, 1771–1813. [23, 26]
- Brandenburger, Adam and Eddie Dekel (1987), “Rationalizability and correlated equilibria.” *Econometrica*, 55, 1391–1402. [22, 24]
- Brandenburger, Adam and Eddie Dekel (1993), “Hierarchies of beliefs and common knowledge.” *Journal of Economic Theory*, 59, 189–198. [21, 22, 29]
- Dekel, Eddie, Drew Fudenberg, and Stephen Morris (2005a), “Interim rationalizability.” Discussion paper 2064, Harvard Institute of Economic Research. [26, 50]
- Dekel, Eddie, Drew Fudenberg, and Stephen Morris (2005b), “Topologies on types.” Working Paper 2093, Harvard Institute of Economic Research. [50]
- Ely, Jeffrey C. and Marcin Pęski (2004), “Hierarchies of belief and interim rationalizability.” Discussion Paper 1388, Northwestern University CMS-EMS. [38]

Gul, Faruk (1998), "A comment on Aumann's Bayesian view." *Econometrica*, 66, 923–928. [22, 23]

Harsanyi, John C. (1967–68), "Games with incomplete information played by 'Bayesian' players, Parts I, II, III." *Management Science*, 14, 159–182, 320–334, 486–502. [19]

Heifetz, Aviad and Dov Samet (1999), "Coherent beliefs are not always types." *Journal of Mathematical Economics*, 32, 475–488. [21, 51]

Kajii, Atsushi and Stephen Morris (1998), "Payoff continuity in incomplete information games." *Journal of Economic Theory*, 82, 267–276. [50]

Liu, Qingmin (2005), "Representation of belief hierarchies in games with incomplete information." Unpublished paper, Graduate School of Business, Stanford University. [26]

Meier, Martin (2005), "On the non-existence of universal information structures." *Journal of Economic Theory*, 122, 132–139. [21]

Mertens, Jean-François, Silvain Sorin, and Shmuel Zamir (1994), "Repeated games." Discussion Papers 9420, 9421, 9422, CORE, Louvain-la-Neuve. [29, 51, 52, 53, 54]

Mertens, Jean-François and Shmuel Zamir (1985), "Formulation of Bayesian analysis for games with incomplete information." *International Journal of Game Theory*, 14, 1–29. [21, 24, 25, 29]

Monderer, Dov and Dov Samet (1997), "Proximity of information in games with incomplete information." *Mathematics of Operations Research*, 21, 707–725. [50]

Morris, Stephen (2002), "Typical types." Unpublished paper, Department of Economics, Yale University. [50]

Savage, Leonard J. (1954), *The Foundations of Statistics*. John Wiley and Sons, New York. [23]

Weinstein, David and Muhamet Yildiz (2003), "Impact of higher-order uncertainty." Working Paper 03-14, Department of Economics, MIT. [50]

Weinstein, David and Muhamet Yildiz (2004), "Finite order implications of any equilibrium." Unpublished paper, Department of Economics, MIT. [50]