The Strategic Impact of Higher-Order Beliefs

Yi-Chun Chen† Alfredo Di Tillio‡ Eduardo Faingold§ Siyang Xiong¶

August 26, 2015

Abstract

Previous research has established that the predictions of game theory are quite sensitive to the assumptions made about the players’ beliefs. We evaluate the severity of this robustness problem by characterizing conditions on the primitives of the model—the players’ beliefs and higher-order beliefs about the payoff-relevant parameters—for the behavior of a given Harsanyi type to be approximated by the behavior of (a sequence of) perturbed types. This amounts to providing belief-based characterizations of the strategic topologies of Dekel, Fudenberg, and Morris (2006). We apply our characterizations to a variety of questions concerning robustness to perturbations of higher-order beliefs, including genericity of types that are consistent with a common prior, and we investigate the connections between our notions of robustness and the notion of common p-belief of Monderer and Samet (1989).

Keywords: Games with incomplete information, rationalizability, higher-order beliefs, robustness

JEL Classification: C70, C72.
1 Introduction

A major concern with non-cooperative game theory is its reliance on details. The formal description of a strategic situation as a game requires informational assumptions that are often not verifiable in full detail by the analyst in real life, such as the players’ beliefs about the precise order of moves, the actions available to the players when they move, and their exact payoff functions. Unfortunately, game theoretic solutions are known to depend sensitively on those assumptions. For example, in an exchange economy with uncertainty, where agents share a common prior on the underlying state of the world, there is no feasible trade that is commonly known to be mutually acceptable (Milgrom and Stokey, 1982), whereas such a trade can exist if the common prior and common knowledge assumptions are slightly violated (Morris, 1994). In auction environments, bidders with independent values retain information rents (Myerson, 1981), while the auctioneer can fully extract their surplus when their values are slightly correlated (Crémer and McLean, 1988). In the alternating-offers bargaining model of Rubinstein (1982), players reach an immediate agreement in the unique (subgame perfect) equilibrium, but delay to agreement is possible in equilibrium when the players have heterogeneous beliefs about who will make an offer in each period (Yildiz, 2003). In all these disparate and prominent economic scenarios, a model that makes unwarranted informational assumptions may deliver predictions that are not robust.

What form of misspecification error can be allowed in a model of the players’ beliefs, to ensure the model will deliver robust predictions across a wide range of economic situations? Focusing on strategic-form games with incomplete information, we take the point of view of the analyst who posits a type space (Harsanyi, 1967-68) to model the players’ uncertainty, and recognizes that his model may be misspecified. For example, he may assume that there is common knowledge of the true payoff-relevant parameter, but understands that such common knowledge assumption can be at best an approximation of reality. Or, as is often the case in practice, he may posit a finite type space, or a type space with a common prior, but realizes that the true type space may be larger, or that the true common prior distribution may be slightly different from the one assumed, or even that the players may have slightly different priors. To analyze the impact of such kinds of misspecifications, we study the tail properties of the hierarchies of beliefs encoded in the Harsanyi types—a player’s beliefs about the payoff-relevant parameters, his beliefs about the other players’ beliefs about the payoff-relevant parameters, and so on, ad infinitum—and their implications for behavior. Our main finding is an exact characterization of what it takes for a pair of types to display similar strategic behaviors. Thus, we measure the minimum level of precision of the analyst’s information model that is required for accurate predictions of strategic play.

To explain our results we first need to be precise about what we mean by “strategic behavior.” Our behavioral assumption is that players play a Bayesian equilibrium on a type space (possibly

---

1Full surplus extraction has also been itself the subject of a debate concerning whether, and in what sense, it is a robust result. See Heifetz and Neeman (2006) and Chen and Xiong (2013).
without a common prior). Thus, from the perspective of the analyst, who does not know the true type space of the players and has a concern for robustness, the relevant solution concept is (interim correlated) rationalizability (Dekel, Fudenberg, and Morris, 2006). Indeed, the set of actions that are rationalizable for a type $t$ coincides with the set of actions that can be played in some Bayesian equilibrium on some type space, by some type that has the same hierarchy of beliefs as $t$ (Dekel, Fudenberg, and Morris, 2007, Remark 2). A similar perspective is taken by Bergemann and Morris (2009) in the context of robust mechanism design. See also Aumann (1987) and Brandenburger and Dekel (1987) for early papers pioneering this approach.

Formally, our main results are characterizations of the strategic topology and the uniform strategic topology of Dekel, Fudenberg, and Morris (2006). The former is the coarsest topology on the universal type space (Mertens and Zamir, 1985)—the space of all hierarchies of beliefs—under which the correspondence that maps each type of a player into his set of rationalizable actions displays the same kind of continuity properties that the best-reply, Nash equilibrium and rationalizability correspondences exhibit in complete information games.\footnote{See the introduction of Dekel, Fudenberg, and Morris (2006) for a precise analogy.} Thus, for any player, a sequence of types $t^n$ converges in the strategic topology to a type $t$ if and only if, for every finite game and every action $a$ of the player in the game, the following conditions are equivalent: (a) action $a$ is rationalizable for type $t$; (b) for every $\varepsilon > 0$ and sufficiently large $n$, action $a$ is $\varepsilon$-rationalizable for type $t^n$, where $\varepsilon$ is a size of sub-optimization allowed in the incentive constraints. Convergence in the uniform strategic topology adds the requirement that the rate of convergence in (b) be uniform across all finite games (with uniformly bounded payoffs).

As shown by Dekel, Fudenberg, and Morris (2006), a sequence of types converges in the strategic topology only if it converges in the product topology: for every integer $k \geq 1$, the sequence of $k$-order beliefs must converge weakly. However, the Electronic Mail game of Rubinstein (1989) and, more generally, the structure theorem of Weinstein and Yildiz (2007), show that convergence in the product topology does not imply strategic convergence. Our characterizations are based on a strengthening of product convergence that requires $k$-order beliefs to converge at a rate that is uniform in $k$.

We first explain the characterization of the uniform strategic topology, as it is simpler to state and can serve as a benchmark for the other characterization result. For each $k$, endow the space of $k$-order beliefs with the Prohorov distance, which is a standard distance that metrizes the topology of weak convergence of probability measures (Billingsley, 1999). Say that a sequence of types $t^n$ converges uniform-weakly to a type $t$ if the $k$-order belief of $t^n$ converges to the $k$-order belief of $t$ and the rate of convergence is uniform in $k$. Our first main result, Theorem 1, states that uniform strategic convergence is equivalent to uniform weak convergence.\footnote{The partial result that uniform weak convergence implies uniform strategic convergence was proved in Chen, Di Tillio, Faingold, and Xiong (2010). The reverse implication is new to the present paper.} To interpret, consider an analyst who would like to make predictions with some minimal level of accuracy, and wants to
achieve this level of accuracy uniformly across all strategic situations that the players might face. A necessary and sufficient condition for such uniformly robust prediction is that the analyst’s model of the players’ beliefs and higher-order beliefs be sufficiently precise, with the required degree of precision, as measured by the Prohorov distance, binding uniformly over all levels of the belief hierarchy.

The content of Theorem 1 can be dissected in two parts. First, the theorem underscores the role of uniform convergence of hierarchies of beliefs as a requirement for robustness. In light of the structure theorem of Weinstein and Yildiz (2007), which shows that the tails of the hierarchies of beliefs can have a large impact on strategic behavior, the role of uniform convergence should not come as a surprise. Second, the theorem quantifies the impact of a misspecification at each order of the hierarchy by the Prohorov distance. We view this as a nontrivial part of the theorem. Indeed, the Prohorov distance, on which the notion of uniform weak convergence is based, is but one of many equivalent distances that metrize the topology of weak convergence of probability measures. For any such distance, one can consider the associated uniform distance over infinite hierarchies of beliefs. It turns out that these distances may generate different topologies over infinite hierarchies, even though the induced topologies over $k$-order beliefs coincide for each finite $k$.\footnote{In Chen, Di Tillio, Faingold, and Xiong (2010) we report an example of a sequence of types that converges uniform weakly but fails to converge in the uniform topology associated with a distance (different from Prohorov) that metrizes the topology of weak convergence of probability measures.} Theorem 1 identifies one of these uniform distances that ultimately characterizes the uniform-strategic topology.

The characterization of the strategic topology (Theorem 2) is also based on uniform convergence and the Prohorov metric, but is more subtle. The relevant class of events for uniform weak convergence, and a fortiori, uniform strategic convergence, is the entire Borel $\sigma$-algebra of the universal type space. By contrast, our characterization of the strategic topology highlights the role of coarser information structures called frames. A frame is a profile of finite partitions of the universal type space—one partition for each player—that satisfies a measurability condition: each player’s belief concerning the events in the frame must pin down a unique atom of that player’s partition. (We discuss the meaning of this condition below.) For any frame $\mathcal{P}$ and any positive integer $k$, we define a distance over types, $d^k_{\mathcal{P}}$, that is analogous to the Prohorov distance over $k$-order beliefs, but restricts the events for which the proximity is measured to those in the frame $\mathcal{P}$. Say that a sequence of types $t^n$ converges to a type $t$ uniform-weakly on $\mathcal{P}$ if, for every positive integer $k$, $t^n$ converges to $t$ under $d^k_{\mathcal{P}}$ and the rate of convergence is uniform in $k$. Our second main result, Theorem 2, states that a sequence of types converges strategically if and only if it converges uniform-weakly on every frame.

A frame can be interpreted as a “self-contained” coarsening of the canonical information structure of the universal type space. Each event in the partition of player $i$ must be measurable with respect to the partition generated by identifying types that share the same beliefs about the param-
eter and the events in the partition of \(-i\). In other words, each event in the partition of player \(i\) can be interpreted as a statement regarding the belief of player \(i\) about the parameters and the events in the partition of player \(-i\), which can in turn be interpreted as statements about player \(-i\)’s beliefs about the parameters and the events in player \(i\)’s partition, and so on. If the strategic behavior of player \(i\) in a game is measurable with respect to player \(i\)’s partition, it is intuitive that the strategic behavior of player \(-i\) must in turn be measurable with respect to player \(-i\)’s partition. This is the fundamental intuition for why strategic convergence implies uniform weak convergence only on frames and not on all information structures.

To shed further light on the impact of higher-order beliefs, we use our characterization to investigate the connection between strategic convergence and a natural notion of uniform convergence based on common \(p\)-beliefs (Monderer and Samet, 1989). Say that a sequence of types \(t^n\) converges in common beliefs to a type \(t\) if \(t^n\) converges to \(t\) in the product topology and, for every event \(E\) and every \(p \geq 0\), the following conditions are equivalent: (i) \(E\) is common \(p\)-belief for type \(t\); (ii) for every \(\varepsilon > 0, k \geq 1\) and sufficiently large \(n\), type \(t^n\) has common \((p - \varepsilon)\)-belief on the event that the players have \(k\)-order beliefs that are \(\varepsilon\)-close to those from \(E\). This is the interim analogue of the ex ante notion of convergence based on common \(p\)-beliefs, which the seminal papers of Monderer and Samet (1996) and Kajii and Morris (1998) have shown to characterize the ex ante strategic topology for Bayesian equilibrium on countable common prior type spaces.

We establish, as a corollary of the characterization of Theorem 2, that strategic convergence implies convergence in common beliefs (Corollary 1). However, somewhat surprisingly, we find that the converse fails. These results highlight a fundamental difference between the common prior, equilibrium, ex ante framework of the early literature and our non-common prior, non-equilibrium, interim framework. Nonetheless, when the limit is assumed a finite type—a type that belongs to a finite type space—we show that convergence in common beliefs is equivalent to uniform weak convergence, and hence, a fortiori, to both uniform strategic and strategic convergence (Theorem 3).

Finally, we use our characterizations to revisit, and reverse, two important genericity results concerning the structure of the universal type space. The first result, due to Ely and Pęski (2011), shows that critical types—those types which display discontinuous rationalizable behavior—form a meager set under the product topology. By way of contrast, we show that they form an open and dense set under the strategic topology (Corollary 2). Second, Lipman (2003) shows that types consistent with a common prior are dense in the universal type space under the product topology. Instead, we show that those types are nowhere dense under the strategic topology (Theorem 4). Furthermore, we show that an even stronger lack of robustness affects common prior type spaces. When the family of type spaces is endowed with the Hausdorff topology derived from the underlying product topology on types, the set of common prior type spaces is nowhere dense. Moreover,

\footnote{An event \(E\) is common \(p\)-belief for a given type if that type assigns probability at least \(p\) to \(E\), assigns probability at least \(p\) to the event that \(E\) obtains and the other players assign probability at least \(p\) to \(E\), and so forth, ad infinitum.}
the analogous claim holds for the Hausdorff topology derived from the strategic topology on types (Theorem 5).

The rest of the paper is organized as follows. Section 2 sets up the incomplete information model and reviews the basic definitions and properties of type spaces, hierarchies of beliefs, common \( p \)-beliefs and the solution concept of interim correlated rationalizability. Section 3 presents the strategic topologies of Dekel, Fudenberg, and Morris (2006) and, after a review of the complete information benchmark, their characterizations in terms of beliefs. Section 3 also studies the relationship of our characterizations with the notion of common \( p \)-belief of Monderer and Samet (1989), and concludes with the analysis of critical types. Section 4 examines the implications of our characterization theorems on finite and common prior models. Section 5 discusses some further interpretations of our results and relations to the literature, as well as some extensions. Finally, the appendix collects the proofs of all results and other technical discussions.

2 Basic Setup

Given a measurable space \( X \), we write \( \Delta(X) \) for the space of probability measures on \( X \), equipped with the \( \sigma \)-algebra generated by the sets of the form \( \{ v \in \Delta(X) : v(E) \geq p \} \), where \( E \subseteq X \) is a measurable set and \( p \in (0, 1] \). Unless otherwise stated, product spaces are endowed with the product \( \sigma \)-algebra, subspaces with the relative \( \sigma \)-algebra, and finite spaces with the discrete \( \sigma \)-algebra. With slight abuse of notation, we often denote one-point sets \( \{ x \} \) simply by \( x \), and cylinders \( E \times Y \) in product spaces \( X \times Y \) by their bases \( E \subseteq X \).

2.1 Games with Incomplete Information

In a game with incomplete information, payoffs depend on the profile of actions as well as on an exogenous parameter that may not be common knowledge among the players. The Bayesian approach to modeling games with incomplete information, due to Harsanyi (1967-68), requires the specification of a type space to model this possible lack of common knowledge. Such spaces describe the players’ uncertainty about the payoff-relevant parameter, but also their uncertainty about each other’s uncertainty about the parameter, and any higher-order uncertainty.

We consider two-player games with incomplete information and denote the set of players by \( I = \{1, 2\} \).\(^6\) The space \( \Theta \) of payoff-relevant parameters is assumed to be finite and to contain at least two points. A game is a profile \( G = (A_i, g_i)_{i \in I} \) where \( A_i \) is a finite set of actions of player \( i \) and \( g_i : A_i \times A_{-i} \times \Theta \to [-M, M] \) is his payoff function, where \( M > 0 \) is a uniform bound on payoffs that is fixed throughout.\(^7\) A type space is a profile \( T = (T_i, \mu_i)_{i \in I} \) where \( T_i \)

\(^6\)We focus on two-player games for ease of notation. All our results extend to the general \( N \)-player case.

\(^7\)Following standard notation, for each player \( i \) we let \(-i\) designate the other player in \( I \).
is a measurable space of types of player $i$ and $\mu_i : T_i \rightarrow \Delta(\Theta \times T_{-i})$ is a measurable function that associates, with each type $t_i$ of player $i$, his belief $\mu_i(t_i)$ about the payoff-relevant parameters and the types of player $-i$. The details of how type spaces encode higher-order uncertainty are reviewed in section 2.3 below.

At this point, it is worthwhile digressing to discuss how the general formulation of type spaces above relates to the one commonly used in applications. Readers who are familiar with our formulation may skip the next two paragraphs without incurring any loss.

In most applications of games with incomplete information, uncertainty is modeled through the standard interdependent values model: the space of payoff-relevant parameters is assumed to have a product structure $\Theta = \prod_i \Theta_i$; there is a probability distribution $f \in \Delta(\Theta)$ according to which the profile $\theta = (\theta_i)_{i \in I}$ is drawn; upon the realization of $\theta$ each player $i$ learns his coordinate $\theta_i$ and forms a belief over $\Theta_{-i}$ by updating the prior distribution $f$ via Bayes’ formula. This model can be translated into our formulation as follows: for each $i \in I$, we have $T_i = \Theta_i$ and $\mu_i : T_i \rightarrow \Delta(\Theta \times T_{-i})$ is given by

$$\mu_i(t_i)(\theta, t_{-i}) = \begin{cases} f(t_i, t_{-i}) / \sum_{\theta_{-i} \in \Theta_{-i}} f(t_i, \theta_{-i}) & \text{if } \theta = (t_1, t_2) \\ 0 & \text{if } \theta \neq (t_1, t_2). \end{cases}$$

The standard interdependent values model is flexible in some dimensions: it nests both the private values model and the pure common values model as particular cases, and it allows types to be either independent or correlated. Moreover, the model can be easily generalized to one in which players have different prior distributions $f_i \in \Delta(\Theta)$, an extension that is desirable in some applications. Most of the mechanism design literature is built on the standard interdependent values model.

However, the standard interdependent values model is restrictive for the purpose of our robustness exercise. It embodies a strong common knowledge assumption: the belief of a type concerning the payoff relevant parameter only, called his first-order belief, completely pins down his joint belief concerning the parameter and the type of the other player. As a consequence, the second-order belief of a type, defined as his joint belief concerning the parameter and the first-order belief of the other player, is completely pinned down by his first-order belief, and likewise for all higher-order beliefs. In other words, all higher-order uncertainty is trivial in the standard interdependent values model. But, as discussed in the introduction, in this paper we are interested in quantifying the behavioral impact of perturbing the players’ types, so it makes sense to allow perturbations to be as general as possible, rather than constrained by common knowledge assumptions. In other words, while in many cases it might make perfect sense to adopt the interdependent values model for the unperturbed model, it is desirable to allow the perturbations to relax the common knowledge assumptions assumed in the unperturbed model, as the very nature of the exercise is to quantify the degree with which strategic behavior is sensitive to such common knowledge assumptions. For this reason, we adopt the general formulation of type spaces above.
2.2 Solution Concept

Our analysis is based on the solution concept of *interim correlated rationalizability* due to Dekel, Fudenberg, and Morris (2007), which is an extension of the complete information rationalizability of Bernheim (1984) and Pearce (1984) to games with incomplete information. The definition follows a recursive process of elimination of non-best-replies in which the elimination is performed on a type-by-type basis. An action is rationalizable for a type if it survives infinitely many rounds of elimination of non-best-replies; at each round an action is declared a non-best-reply for a type, if it is not a best-reply to any conjecture that is consistent with the type’s beliefs and assigns probability one to each type of the other player choosing actions that have not been eliminated thus far. The formal definition below allows for \( \varepsilon \)-best replies, which leads to the concept of \( \varepsilon \)-rationalizability.

Let a game \( G = (A_i, g_i)_{i \in I} \), a type space \( T = (T_i, \mu_i)_{i \in I} \) and \( \varepsilon \geq 0 \) be given. For every \( i \in I \) and \( t_i \in T_i \) define the set of 0-order \( \varepsilon \)-rationalizable actions as \( R_i^0(t_i, G, T, \varepsilon) = A_i \). Proceeding recursively, for \( k \geq 1 \), define the set of \( k \)-order \( \varepsilon \)-rationalizable actions of type \( t_i \), written \( R_i^k(t_i, G, T, \varepsilon) \), as the set of actions \( a_i \in A_i \) for which there exists a conjecture \( \nu \in \Delta(\Theta \times T_{-i} \times A_{-i}) \) satisfying the following three properties:

(a) Action \( a_i \) is an \( \varepsilon \)-best-reply to the conjecture: for every \( a_i' \in A_i \),

\[
\sum_{(\theta, a_{-i}) \in \Theta \times A_{-i}} [g_i(\theta, a_i, a_{-i}) - g_i(\theta, a_i', a_{-i})] \nu(\theta \times T_{-i} \times a_{-i}) \geq -\varepsilon.
\]

(b) The conjecture assigns probability one to the event that each type of player \(-i\) plays an action that is \((k - 1)\)-order \( \varepsilon \)-rationalizable:

\[
\nu(\Theta \times \{ (t_{-i}, a_{-i}) : a_{-i} \in R_{-i}^{k-1}(t_{-i}, G, T, \varepsilon) \}) = 1.
\]

(c) The conjecture is consistent with the belief of type \( t_i \): for every \( \theta \in \Theta \) and measurable \( E_{-i} \subseteq T_{-i} \),

\[
\mu_i(t_i)(\theta \times E_{-i}) = \nu(\theta \times E_{-i} \times A_{-i}).
\]

The set of \( \varepsilon \)-rationalizable actions of type \( t_i \), written \( R_i(t_i, G, T, \varepsilon) \), is the set of actions that are \( k \)-order \( \varepsilon \)-rationalizable for every \( k \):

\[
R_i(t_i, G, T, \varepsilon) = \bigcap_{k \geq 1} R_i^k(t_i, G, T, \varepsilon).
\]

2.3 Higher-Order Beliefs

Type spaces contain implicit descriptions of each player’s beliefs about the payoff-relevant parameters, his beliefs about the other player’s beliefs about the payoff-relevant parameters, and so
on. Any such higher-order uncertainty can be modeled in a canonical type space, called *universal type space*, whose construction, due to Mertens and Zamir (1985), we now review. An equivalent formulation is found in Brandenburger and Dekel (1993).

A *first-order belief* of player $i$ is a probability distribution over payoff-relevant parameters, that is, an element of $T_{i,1} = \Delta(\Theta)$. Recursively, for $k \geq 2$, a *$k$-order belief* of player $i$ is a joint probability distribution over payoff-relevant parameters and $(k-1)$-order beliefs of the other player, that is, an element of $T_{i,k} = \Delta(\Theta \times T_{i,k-1})$. A *hierarchy of beliefs* of player $i$ is an infinite sequence in $T_{i,1} \times T_{i,2} \times \cdots$ satisfying the condition that higher-order beliefs are compatible with lower-order beliefs, that is, an element of the set

$$T^*_i := \left\{ (t_{i,1}, t_{i,2}, \ldots) \in \bigotimes_{k=1}^{\infty} T_{i,k} : \rho_{i,k}(t_{i,k+1}) = t_{i,k} \forall k \geq 1 \right\},$$

where $\rho_{i,k}$ is the function that associates each $(k+1)$-order belief with its induced $k$-order belief. 8

Given a type space $(T_i, \mu_i)_{i \in I}$, the hierarchy of beliefs $t_i(t_i) = (t_{i,1}(t_i), t_{i,2}(t_i), \ldots)$ induced by a type $t_i \in T_i$ is defined as follows. The first-order belief $t_{i,1}(t_i)$ is the marginal of $\mu_i(t_i)$ on $\Theta$, that is, $t_{i,1}(t_i)(\theta) = \mu_i(t_i)(\theta \times T_{i-})$ for every $\theta \in \Theta$. For $k \geq 2$, recursively, the $k$-order belief $t_{i,k}(t_i)$ is defined by letting

$$t_{i,k}(t_i)(\theta \times E) = \mu_i(t_i)(\theta \times T_{i,k-1}^{-1}(E))$$

for every $\theta \in \Theta$ and measurable $E \subseteq T_{i,k-1}$. Thus, every type from an arbitrary type space is naturally mapped to a hierarchy of beliefs.

Conversely, every hierarchy of beliefs can be viewed as a type: there exists a unique measurable function

$$\mu^*_i: T^*_i \rightarrow \Delta(\Theta \times T^*_{i-})$$

with the property that the probability measure $\mu^*_i(t_{i,1}, t_{i,2}, \ldots)$ extends each of the measures $t_{i,1}, t_{i,2}, \ldots$. Moreover, $\mu^*_i$ is an isomorphism. 9 The type space $(T^*_i, \mu^*_i)_{i \in I}$ is called the universal type space. 10 To ease notation, for each event $E \subseteq \Theta \times T_{i-}$ we shall often write $\mu^*_i(E \mid t_i)$ instead of the more cumbersome $\mu^*_i(t_i)(E)$.

---

8The first-order belief $\rho_{i,1}(t_{i,2})$ induced by a second-order belief $t_{i,2}$ is the marginal of $t_{i,2}$ on $\Theta$. Proceeding recursively, the $k$-order belief induced by a $(k+1)$-order belief $t_{i,k+1}$ is the probability distribution $\rho_{i,k}(t_{i,k+1})$ on $\Theta \times T_{i,k-1}$ such that $\rho_{i,k}(t_{i,k+1})(\theta \times E) = t_{i,k+1}(\theta \times \rho_{i,k-1}^{-1}(E))$ for every $\theta \in \Theta$ and measurable $E \subseteq T_{i,k-1}$. 9That is, a measurable bijection with measurable inverse. 10The term universal is borrowed from category theory. It means that every abstract type space $(T_i, \mu_i)_{i \in I}$ is mapped into the universal type space $(T^*_i, \mu^*_i)$ by a unique belief-preserving morphism. In general, a belief-preserving morphism between arbitrary type spaces $(T_i, \mu_i)$ and $(T'_i, \mu'_i)$ is defined as a profile of measurable functions $\varphi_i: T_i \rightarrow T'_i$ such that $\mu'_i(\varphi(t_i))(\theta \times E) = \mu_i(t_i)(\theta \times \varphi^{-1}_i(E))$ for every $\theta \in \Theta$, measurable $E \subseteq T'_{i-}$ and $t_i \in T_i$. The unique belief-preserving morphism that maps an arbitrary type space $(T_i, \mu_i)$ into the universal type space $(T^*_i, \mu^*_i)$ is the profile of mappings $\tau_i$ defined above.
Dekel, Fudenberg, and Morris (2007) show that the set of $\varepsilon$-rationalizable actions of any type $t_i$ from an arbitrary type space $T$ is pinned down by the hierarchy of beliefs of $t_i$, that is,

$$R_i(t_i, G, T, \varepsilon) = R_i(t_i(t_i), G, T^*, \varepsilon).$$

This fact has an important consequence for the framing of our robustness exercise. It implies that we can, and henceforth will, identify types with their induced hierarchy of beliefs. We therefore drop the reference to the type space from the notation for $\varepsilon$-rationalizability, and write $R(t_i, G, \varepsilon)$ for the set of $\varepsilon$-rationalizable actions of type $t_i$. To further ease notation, for $\varepsilon = 0$ we denote $R(t_i, G, 0)$ simply by $R(t_i, G)$, and call it the set of rationalizable actions of type $t_i$.

### 2.4 Common Beliefs

The notion of common belief, due to Monderer and Samet (1989), is a useful tool to study the robustness of strategic behavior in many settings (Monderer and Samet, 1996; Kajii and Morris, 1997, 1998; Ely and Peski, 2011). To define common belief in our context, we consider events that are subsets of the space $\Omega = \Theta \times T_1^* \times T_2^*$. Each point in $\Omega$ is a complete description of the “state of the world,” in that it corresponds to a complete resolution of all the relevant uncertainty: the payoff-relevant parameter and the type of each player.

For each measurable set $E \subseteq \Omega$ and each type $t_i \in T_i^*$, let $E_{t_i}$ designate the section of $E$ over $t_i$:

$$E_{t_i} = \{(\theta, t_{i-}) : (\theta, t_1, t_2) \in E\},$$

which is a measurable set by standard arguments. For each $p \in [0, 1]$, the event that player $i$ assigns probability at least $p$ to $E$ is

$$B_i^p(E) = \{t_i \in T_i^* : \mu_i^*(E_{t_i} | t_i) \geq p\},$$

which is also a measurable set.\footnote{The measurability of $B_i^p(E)$ follows from the measurability of the map $t_i \mapsto \mu_i(E_{t_i} | t_i)$. The class of events $\{E \subseteq \Omega : E$ is a measurable set such that $t_i \mapsto \mu_i(E_{t_i} | t_i)$ is measurable$\}$ can be readily verified to be a monotone class containing the algebra of finite disjoint unions of measurable rectangles, which generates the product $\sigma$-algebra on $\Omega$. It follows that the map $t_i \mapsto \mu_i(E_{t_i} | t_i)$ is measurable for every measurable $E$.} Then, for each $p = (p_1, p_2) \in [0, 1]^2$, define the event that $E$ is mutual $p$-belief as

$$B^p(E) = \Theta \times B_1^{p_1}(E) \times B_2^{p_2}(E),$$

and the event that $E$ is common $p$-belief as

$$C^p(E) = B^p(E) \cap B^p(E \cap B^p(E)) \cap B^p(E \cap B^p(E \cap B^p(E))) \cap \cdots.$$
Then, define the event that \( E \) is common \( p \)-belief for player \( i \), written \( C^p_i(E) \), as the projection of \( C^p(E) \) on \( T^*_i \), which is a measurable set because \( C^p(E) \) is a rectangle.\(^{12}\) Thus, we have

\[
C^p_i(E) = B^p_{i}(E) \cap B^p_{i}(E \cap B^p_{i-1}(E)) \cap B^p_{i}(E \cap B^p_{i-1}(E \cap B^p_{i}(E))) \cap \ldots.
\]

Note that for \( p_{-i} = 0 \) we have \( C^p_i(E) = B^p_{i}(E) \), while for \( p = (1, 1) \) the set \( C^{(1,1)}_i(\theta) := C_i^{(1,1)}(\theta \times T^*_1 \times T^*_2) \) contains only one type of player \( i \).

Finally, common belief has the following well known fixed-point characterization:

\[
C^p(E) = B^p(E \cap C^p(E)) \quad \text{and} \quad C^p_i(E) = B^p_{i}(E \cap C^p_{i-1}(E)).
\]

(1)

### 3 Characterization of Robustness

To what kinds of misspecification of the players’ uncertainty are the predictions of rationalizability robust? That is, given a player \( i \) and a type \( t_i \), what sequences of perturbed types \( t^n_i \) converge to \( t_i \) in the sense that the strategic behaviors of \( t^n_i \) approach those of \( t_i \)? To formalize this question, we use the notions of strategic convergence proposed by Dekel, Fudenberg, and Morris (2006), which we now recall.

**Definition 1** (Strategic topology). A sequence of types \( t^n_i \) converges strategically to a type \( t_i \) if for every game \( G \) and every action \( a_i \) of player \( i \) in \( G \), the following conditions are equivalent:

(a) \( a_i \) is rationalizable for \( t_i \) in \( G \);

(b) for every \( \varepsilon > 0 \) there exists \( N \) such that for every \( n \geq N \), \( a_i \) is \( \varepsilon \)-rationalizable for \( t^n_i \) in \( G \).

The strategic topology is the topology of strategic convergence on \( T^*_i \).\(^{14}\)

The implication from (b) to (a) is a form of upper hemicontinuity of the rationalizable correspondence: if for every \( \varepsilon > 0 \) an action is \( \varepsilon \)-rationalizable for every type in the tail of the sequence, then the action is also rationalizable for the limit type. In particular, every action that is rationalizable for every type in the tail of the sequence remains rationalizable for the limit type. In non-technical terms, this means that the unperturbed model never fails to predict a behavior that is predicted by the perturbed model. It is not a strong requirement: Dekel, Fudenberg, and Morris

\(^{12}\)The definition of \( C^p(E) \) is analogous to the common repeated belief of Monderer and Samet (1996), which differs from the original definition of Monderer and Samet (1989). A similar definition appears in Ely and Pęski (2011) for the case where \( E \) is a rectangle. We allow the event \( E \) to be any measurable set.

\(^{13}\)Recall that for notational convenience we often denote cylinders by their bases. Thus, here we identify \( B^p_{i}(E) \) and \( C^p_i(E) \) with \( \Theta \times B^p_{i}(E) \times T^*_i \) and \( \Theta \times C^p_i(E) \times T^*_i \), respectively.

\(^{14}\)This definition follows Ely and Pęski (2011). The original definition of Dekel, Fudenberg, and Morris (2006) is stated a bit differently, but the two are equivalent, as shown in the working paper version of Ely and Pęski (2011).
(2006, Theorems 1 and 2) show that the implication from (b) to (a) (holding for every game and action) is equivalent to convergence in the product topology, that is, weak convergence of $k$-order beliefs for every $k \geq 1$.\footnote{The definition of the product topology is recursive: endow $T_{i,1}$ with the Euclidean topology, and for each $k \geq 2$, endow $T_{i,k} = \Delta(\Theta \times T_{i,k-1})$ with the topology of weak convergence of probability measures relative to the product topology on $\Theta \times T_{i,k-1}$. A sequence of types $t^n_i$ converges to $t_i$ in the product topology if for every $k \geq 1$, the $k$-order belief of $t^n_i$ converges to the $k$-order belief of $t_i$. Note that, since $\Theta$ is finite, the Borel $\sigma$-algebra of the product topology on $T^*_i$ coincides with the $\sigma$-algebra assumed in the topology-free formulation of section 2.3. Likewise, the Borel $\sigma$-algebra on $\Delta(\Theta \times T^*_i)$ generated by the topology of weak convergence of probability measures (induced by the product topology on $T^*_i$) coincides with the $\sigma$-algebra of our topology-free formulation. Finally, if we endow each $T^*_i$ with the product topology and $\Delta(\Theta \times T^*_i)$ with the topology of weak convergence of probability measures, then $T^*_i$ and $\Delta(\Theta \times T^*_i)$ are compact metrizable, and $\mu^*_i$ is a homeomorphism (Mertens and Zamir, 1985).}

Similarly, the implication from (a) to (b) is a form of lower hemicontinuity: if an action is rationalizable for the limit type, then for every $\varepsilon > 0$ the action is $\varepsilon$-rationalizable for every type in the tail of the sequence. To interpret, it is easier to examine an equivalent condition that does not involve $\varepsilon$-rationalizability: every action that is strictly rationalizable for the limit type remains rationalizable for all types in the tail of the sequence.\footnote{An action is strictly rationalizable if it is $\varepsilon$-rationalizable for some negative $\varepsilon$. In other words, all incentive constraints hold with a slack that is bounded away from zero. The working paper version of Ely and Pęski (2011) shows that the two definitions are equivalent.} In other words, the unperturbed model never makes predictions that are not valid for the perturbed model, unless those predictions rely on non-strict incentives (i.e. indifferences).\footnote{Imposing the condition also for rationalizable actions that rely on indifferences is way too strong, as it would lead to the discrete topology.}

As the electronic mail game example of Rubinstein (1989) demonstrates, this condition is not implied by convergence in the product topology. Finally, if the implication from (a) to (b) holds for every action in every game, then the implication from (b) to (a) also holds (Dekel, Fudenberg, and Morris, 2006, Corollary 1).

The definition of strategic convergence allows the rate of convergence $N$ to depend on the game $G$. An alternative, stronger notion of convergence, similar to the one adopted in the early literature on robustness (Monderer and Samet, 1989, 1996; Kajii and Morris, 1997, 1998), requires the rate of convergence $N$ to be independent of $G$ (across all games with uniformly bounded payoffs), which leads to the following definition.

**Definition 2** (Uniform strategic topology). A sequence of types $t^n_i$ converges uniform-strategically to a type $t_i$ if for every $\varepsilon > 0$ there exists $N$ such that for every game $G$ and every action $a_i$ of player $i$ in $G$, the following conditions are equivalent:

(a) $a_i$ is rationalizable for $t_i$ in $G$;

(b) for every $\varepsilon > 0$ and $n \geq N$, $a_i$ is $\varepsilon$-rationalizable for $t^n_i$ in $G$.

The uniform strategic topology is the topology of uniform strategic convergence on $T^*_i$. 

\footnote{Imposing the condition also for rationalizable actions that rely on indifferences is way too strong, as it would lead to the discrete topology.}
The main question we address in the paper is:

Under what conditions on primitives (i.e., beliefs and higher-order beliefs) does a sequence of types \( t^n_i \) converge strategically, or uniform strategically, to a type \( t_i \)?

### 3.1 Complete Information Benchmark

Before turning to our general analysis, it is worth reviewing a particular case for which the robustness question above has a well-known, and simple, answer: the case of complete information, where some payoff-relevant parameter is common knowledge for the unperturbed type \( t_i \). In this case, we have the following well known result:\(^{18}\)

**Proposition 1** (Convergence to complete information). Let \( \Theta \in \Theta \) and \( t_i = C_i^{(1,1)}(\theta) \). For every sequence of types \( t^n_i \), the following statements are equivalent:

(i) for every \( p \in (0,1) \) there exists \( N \) such that for every \( n \geq N \), \( t^n_i \in C_i^{(p,p)}(\theta) \);

(ii) \( t^n_i \) converges uniform-strategically to \( t_i \);

(iii) \( t^n_i \) converges strategically to \( t_i \).

The two canonical examples below illustrate the application of condition (i) as a criterion for strategic (and uniform strategic) approximation of common knowledge. In both examples, we assume that \( \Theta = \{\theta_0, \theta_1\} \).

**Example 1** (Public announcements with mistakes, Monderer and Samet, 1989). One situation in which common knowledge of a payoff-relevant parameter arises is when there is an informed third party (assumed to be non-strategic) who publicly announces the true value of the parameter before a game is played. Moving slightly away from common knowledge, suppose that each player (independently) hears an erroneous announcement with a small probability \( \varepsilon > 0 \). Suppose also that both players assign prior probability \( \frac{1}{2} \) to each parameter value. We thus have a type space with two types, \( t^{\varepsilon}_{i,0} \) and \( t^{\varepsilon}_{i,1} \), for each player \( i \). Type \( t^{\varepsilon}_{i,0} \) heard announcement “the parameter is \( \theta_0 \),” while type \( t^{\varepsilon}_{i,1} \) heard announcement “the parameter is \( \theta_1 \).” Their beliefs, derived from Bayes’ rule, are represented in the diagrams of Figure 1.

Thus, each type assigns probability \( p_{\varepsilon} = (1 - \varepsilon)^2 \) to the event that both his own announcement and the other player’s announcement are correct. As a result, for each player \( i \) we have \( t^{\varepsilon}_{i,0} \in C_i^{(p_{\varepsilon},p_{\varepsilon})}(\theta_0) \) and \( t^{\varepsilon}_{i,1} \in C_i^{(p_{\varepsilon},p_{\varepsilon})}(\theta_1) \). By Proposition 1, as \( \varepsilon \to 0 \), types \( t^{\varepsilon}_{i,0} \) and \( t^{\varepsilon}_{i,1} \) converge strategically to \( t_{i,0} = C_i^{(1,1)}(\theta_0) \) and \( t_{i,1} = C_i^{(1,1)}(\theta_1) \), respectively. This implies that, irrespective of the payoff structure of the game, the predictions of the model under common

---

\(^{18}\)This equivalence is the interim analogue of the classic result of Monderer and Samet (1989), who characterize an ex ante notion of strategic approximation of common knowledge. The result is a corollary of Theorem 3 below.
knowledge must be robust to the small amounts of incomplete information that arise in the public announcements protocol when players are subject to a small probability of hearing mistakes.

**Example 2** (Faulty message exchange, Rubinstein, 1989). The players assign prior probability $\frac{1}{2}$ to $\theta_1$. Player 1 then learns the true value of the parameter. If and only if it is $\theta_1$, the players engage in an automated exchange. Player 1 sends a message to player 2. If the message arrives, player 2 automatically replies with a confirmation to player 1. If player 1 receives it, he automatically replays with a new message, and so on and so forth. Each time a message or confirmation is sent, it gets lost with probability $q > 0$, independently of all other events. The protocol gives rise to a countably infinite set of types $\{t^0_i, t^1_i, t^2_i, \ldots\}$ for each player $i$. Type $t^0_i$ of player 1 knows the parameter is $\theta_0$ and, hence, no message was exchanged. Type $t^0_2$ of player 2 received no messages and is thus uncertain whether the parameter is $\theta_0$ or $\theta_1$, although he knows that in the first case no message was ever sent, while in the second exactly one was. For each $m \geq 1$, type $t^m_1$ of player 1 knows the parameter is $\theta_1$ and knows that exactly $m$ messages were sent, so he is uncertain whether player 2 sent $m - 1$ or $m$ confirmations, while type $t^m_2$ of player 2 knows the parameter is $\theta_1$ and knows that exactly $m$ confirmations were sent, so he is uncertain whether player 1 sent $m$ or $m + 1$ messages. The diagrams in Figure 2, where $q' = q/(1 + q)$ and $q'' = 1/(2 - q)$, illustrate the beliefs of the various types, which are derived from Bayes’ rule.

As $m \to \infty$, the sequence of types $t^m_i$ converges to $t_i = C(1,1)(\theta_1)$ in the product topology, since for every $m \geq 1$ the $(2m - 2 + i)$-order beliefs of $t^m_i$ exactly match those of $t_i$. That is, the first order belief of $t^m_i$ assigns probability 1 to the event that the parameter is $\theta_1$, the second-order belief of $t^m_i$ assigns probability 1 to the event that the parameter is $\theta_1$ and player $-i$ assigns probability 1 to $\theta_1$, and so on, up to order $(2m - 2 + i)$. However, $t^m_i \neq C_i^{(p,p)}(\theta_1)$ for every $m \geq 0$ and $p > \max\{1 - q', 1 - q''\}$.

It follows, by Proposition 1, that $t^m_i$ does not converge strategically to

---

19 Type $t^0_2$ assigns probability $1 - q' < p$ to $\theta_1$, while $t^1_2$ assigns probability $q'' > 1 - p$ to $t^0_2$. Thus, $t^0_2 \neq C_2^{(p,p)}(\theta_1)$ and, by (1), $t^1_2 \neq C_2^{(p,p)}(\theta_1)$. Proceeding by induction, suppose that $t^{m-1}_2 \neq C_2^{(p,p)}(\theta_1)$ and $t^m_2 \neq C_2^{(p,p)}(\theta_1)$ for some $m \geq 1$. Then, since $t^m_2$ assigns probability $q''$ to $t^m_1$ and $t^{m+1}_2$ assigns probability $q''$ to $t^m_2$, again by (1) we have

---

![Figure 1: Type space for Example 1](image-url)
$t_i = C^{(1,1)}(\theta_1)$. The conclusion is that, for some payoff structures, the predictions under common knowledge of $\theta_1$ are not robust to the perturbation given by the sequence $t_i^m$, despite the fact that the perturbation preserves arbitrarily high orders of mutual knowledge of $\theta_1$.20

Finally, as we will see later, strategic and uniform-strategic convergence are no longer equivalent when the unperturbed type $t_i$ is not a complete information type. In the remainder of this section, we provide two generalizations of convergence criterion (i), which characterize strategic and uniform-strategic convergence in the general case. We begin with uniform-strategic convergence, as its characterization takes a simpler form.

### 3.2 Characterization of the Uniform Strategic Topology

Perturbations of the tails of belief hierarchies can have a large impact on strategic behavior. In the complete information case, the characterization of Proposition 1 quantifies the impact of such perturbations by identifying a condition on the tails of the perturbed hierarchies—common $p$-belief, for $p$ converging to 1—that is just enough to ensure the robustness of rationalizable behavior. This suggests that in order to characterize strategic (and uniform strategic) convergence in the general case, the necessary condition that $k$-order beliefs converge for every $k$ (convergence in the product topology) should be strengthened by a requirement that the rate of convergence be uniform in $k$.

In order to define uniformity across the orders of the belief hierarchy, we first need to fix a distance on the space of $k$-order beliefs that metrizes the topology of weak convergence of probability measures.21 The choice of distance matters: different distances, even if equivalent at every

---

20 The payoff structure for which the lack of robustness obtains is a version of the well known coordinated attack problem.
21 Recall that a necessary condition for strategic convergence is that $k$-order beliefs converge weakly.
order $k$, may yield different uniformities. The characterization of uniform-strategic convergence that we provide is based on the Prohorov distance (see e.g. Billingsley, 1999, p. 72). The definition is recursive: for each integer $k \geq 1$, define the distance $d_{i}^{k}$ on $T_{i}^{*}$ as the Prohorov distance over $k$-order beliefs, assuming the space of $(k - 1)$-order beliefs of player $-i$ is metrized by $d_{-i}^{k-1}$.

Thus, for each player $i$, we set $d_{i}^{0} = 0$ and, for each integer $k \geq 0$ and types $t_{i}$ and $t_{i}'$,

$$d_{i}^{k+1}(t_{i}, t_{i}') = \inf \left\{ \delta > 0 : \mu_{i}^{*}(E|t_{i}) \leq \mu_{i}^{*}(E_{\delta,k}|t_{i}') + \delta \text{ for every measurable } E \subseteq \Theta \times T_{-i}^{*} \right\}$$

where

$$E_{\delta,k} = \left\{ (\theta, t_{-i}') \in \Theta \times T_{-i}^{*} : d_{-i}^{k}(t_{-i}, t_{-i}') < \delta \text{ for some } t_{-i} \text{ with } (\theta, t_{-i}) \in E \right\}.$$  

That is, types $t_{i}$ and $t_{i}'$ are at a distance at most $\delta$ at order $k + 1$ if for every event $E$ concerning the parameter and the type of player $-i$, the probability that $t_{i}$ assigns to $E$ is at most $\delta$ above the probability that $t_{i}'$ assigns to a $\delta$-neighborhood of $E$ at order $k$.

The following notion of convergence of types, introduced in Chen, Di Tillio, Faingold, and Xiong (2010), is the uniform counterpart of product convergence when the topology of weak convergence of $k$-order beliefs is metrized by the Prohorov distance.

**Definition 3** (Uniform weak convergence). A sequence of types $t_{i}^{n}$ converges uniform-weakly to a type $t_{i}$ if

$$d_{i}^{\text{uw}}(t_{i}, t_{i}^{n}) := \sup_{k \geq 1} d_{i}^{k}(t_{i}, t_{i}^{n}) \to 0 \text{ as } n \to \infty.$$ 

Thus, under uniform-weak convergence, the tails of the belief hierarchies converge at the same rate as the lower-order beliefs, where the rate of convergence is measured relative to the Prohorov distance at each order. The examples below illustrate the definition.

**Example 3** (Convex combinations with convergent weights). Consider a sequence of convex combinations of a fixed pair of types in which the weight on one of the types converges to one. That is, given a pair of types $t_{i}, t_{i}' \in T_{i}^{*}$, for each $n \geq 1$, let $t_{i}^{n}$ be the type whose belief is:

$$\mu_{i}^{*}(\cdot|t_{i}^{n}) = \left( 1 - \frac{1}{n} \right) \mu_{i}^{*}(\cdot|t_{i}) + \frac{1}{n} \mu_{i}^{*}(\cdot|t_{i}').$$

Thus, for each $n \geq 1$, $k \geq 1$ and measurable $E \subseteq \Theta \times T_{-i}^{*}$,

$$\mu_{i}^{*}(E^{1/n,k}|t_{i}^{n}) \geq \mu_{i}^{*}(E|t_{i}^{n}) \geq \left( 1 - \frac{1}{n} \right) \mu_{i}^{*}(E|t_{i}) \geq \mu_{i}^{*}(E|t_{i}) - 1/n,$$

and hence $d_{i}^{k}(t_{i}, t_{i}^{n}) \leq 1/n$. Therefore, $t_{i}^{n} \to t_{i}$ uniform-weakly.

---

22Viewed as a distance on $T_{i}^{*}$, $d_{i}^{k}$ is only a pseudo-distance—as opposed to a standard distance—, since there exist distinct types with the same $k$-order beliefs (and hence different $\ell$-order beliefs, for some $\ell > k$).
Example 4 (Faulty message exchange with convergent hazard rates). Consider the type space \(\{\mu^0_t, \mu^1_t, \mu^2_t \ldots\}\) of Example 2 and, for each \(n \geq 1\), let \(\{\mu^0_{i,n}, \mu^1_{i,n}, \mu^2_{i,n} \ldots\}\) be the analogous type space where the probability of a message getting lost is \(q_n\) rather than \(q\). We thus view each \(\mu^0_t\) as a type in the unperturbed model, and \(\mu^m_{i,n}\) as a perturbation of type \(\mu^m_t\). If the sequence \(q_n\) converges to \(q\), then for each \(m \geq 0\) the sequence \(\mu^m_{i,n}\) converges to \(\mu^m_t\) uniform-weakly as \(n \to \infty\). To see this, let \(q^*_n = q_n/(1 + q_n)\), \(q^*_0 = 1/(2 - q_n)\) and \(\delta_n = \max\{|q^*_n - q'|, |q^*_n - q''|\}\). Types \(\mu^0_t\) and \(\mu^0_{i,n}\) are both certain that the parameter is \(\theta_0\). Types \(\mu^0_t\) and \(\mu^0_{i,n}\) assign probabilities \(q'\) and \(q^*_n\), respectively, to \(\theta_0\). For each player \(i\) and \(m \geq 1\), types \(\mu^m_{i,n}\) and \(\mu^m_t\) are both certain the parameter is \(\theta_1\). Thus, for each player \(i\) and \(m \geq 0\), the first-order distance between \(\mu^m_{i,n}\) and \(\mu^m_t\) is at most \(\delta_n\), that is, for every \(\delta > \delta_n\), type \(\mu^m_{i,n}\) is in the first-order \(\delta\)-neighborhood \(\{\mu^m_t\}_{\delta,1}\) of type \(\mu^m_t\). Proceeding inductively, let \(k \geq 1\) and suppose that \(\mu^m_{i,n} \in \{\mu^m_t\}_{\delta,k}\) for each player \(i\), \(m \geq 0\) and \(\delta > \delta_n\). This immediately implies that for each \(i\), parameter \(\theta\) and indices \(\ell, m \geq 0\),

\[
\mu^*_t(\theta \times \ell_i^m, \mu^m_t) > 0 \quad \Rightarrow \quad \mu^*_t(\theta \times \{\ell_i^m\}_{\delta,k}, \mu^m_{i,n}) \geq \mu^*_t(\theta \times \ell_i^m, \mu^m_t) - \delta \quad \forall \delta > \delta_n.
\]

Thus, the \((k + 1)\)-order distance between \(\mu^m_t\) and \(\mu^m_{i,n}\) is at most \(\delta_n\). Since \(\delta_n \to 0\), it follows that \(\mu^m_{i,n} \to \mu^m_t\) uniform-weakly for every player \(i\) and \(m \geq 0\).

In Chen, Di Tillio, Faingold, and Xiong (2010), we showed that uniform-weak convergence implies uniform-strategic convergence. Here we prove the reverse implication, thus establishing the equivalence between the two notions. The proof of the result is in appendix A.2.

Theorem 1 (Characterization of uniform strategic convergence). A sequence of types converges uniform-strategically if and only if it converges uniform-weakly.

Ever since Rubinstein’s (1989) seminal paper, misspecifications of higher-order beliefs have been recognized to have a potentially large impact on strategic predictions. The systematic treatment of Weinstein and Yildiz (2007) exposed the pervasiveness of this sensitivity by showing that the phenomenon is not peculiar to the Electronic Mail game, and hence advocated wider scrutiny of the assumptions one makes about the players’ subjective beliefs. Theorem 1 quantifies the strategic impact of such assumptions (uniformly over games) by identifying an appropriate measure of proximity of hierarchies of beliefs. In effect, the role of the Prohorov distance in the definition of uniform weak convergence, and hence in the characterization result, is nontrivial. Given any distance that metrizes the topology of weak convergence of probability measures, one can always define an associated uniform distance over infinite hierarchies of beliefs. However, an example in Chen, Di Tillio, Faingold, and Xiong (2010, Section 5.2) shows that these distances may generate different topologies over infinite hierarchies, even though the induced topologies over \(k\)-order beliefs coincide for each finite \(k\). Theorem 1 identifies one of these uniform distances that ultimately characterizes the uniform-strategic topology.

Theorem 1 serves as a benchmark for the characterization of the strategic topology in the next section, which is also based on uniform convergence in the orders of the hierarchy of beliefs.
Uniform weak convergence requires convergence of beliefs concerning all events in the universal type space. Relaxing this requirement is the key to our characterization of the strategic topology below.

### 3.3 Characterization of the Strategic Topology

The definition of uniform weak convergence embodies two kinds of uniformity. The first kind, emphasized in the previous section, requires the rate of convergence of \( k \)-order beliefs to be uniform in \( k \). The second kind, which applies to any fixed order \( k \), is the uniformity in events that is implicit in the definition of the Prohorov distance: the \( k \)-order distance between types \( t_i \) and \( t'_i \) is less than \( \delta \) if and only if

\[
\sup_E \left[ \mu^*_i(E|t_i) - \mu^*_i(E|t'_i) \right] \leq \delta,
\]

where the supremum ranges over all measurable subsets \( E \subseteq \Theta \times T^*_i \). Such uniformity in events is unnecessarily strong if the goal is to characterize the strategic topology, which does not require the rate of strategic convergence to be uniform in games. Indeed, given a fixed game, not all events concerning player \(-i\) are strategically relevant for player \(i\).

The main result of this section, Theorem 2, characterizes the strategic topology in terms of a notion of convergence of types that maintains the uniformity in the orders of the belief hierarchy, but only considers events from certain coarse information structures, which we now define.

**Definition 4 (Frames).** A frame is a profile \( \mathcal{P} = (\mathcal{P}_i)_{i \in I} \), where each \( \mathcal{P}_i \) is a finite measurable partition of \( T^*_i \) such that for each \( t_i, t'_i \in T^*_i \),

\[
\mu^*_i(\theta \times E|t_i) = \mu^*_i(\theta \times E|t'_i) \quad \forall \theta \in \Theta, \quad \forall E \in \mathcal{P}_{-i} \quad \implies \quad t'_i \in \mathcal{P}_i(t_i).
\]

Thus, a partition profile is a frame if any two types of player \( i \) that agree on their beliefs concerning the parameter and the events in the partition of player \(-i\) must belong to the same element of \( i \)'s partition. Equivalently, each event in the partition of player \( i \) must be measurable with respect to the partition generated by identifying types that share the same beliefs about the parameter and the events in the partition of player \(-i\). A frame is thus a coarse information model that is “self-contained:” each event in the partition of player \( i \) can be interpreted as a statement regarding the belief of player \( i \) about the parameters and the events in the partition of player \(-i\), which can in turn be interpreted as statements about player \(-i\)'s beliefs about the parameters and the events in player \( i \)'s partition, and so on.

---

23For instance, consider a game where an action \( a^*_i \) is rationalizable for a type \( t_i \) if and only if he attaches probability greater than a threshold to player \(-i\) playing a certain action \( a^*_{-i} \). Then, in order for type \( t_i \) to verify whether he can rationalize \( a^*_i \), he only needs to check whether his belief about the event that \( a^*_{-i} \) is rationalizable for player \(-i\) lies above or below the threshold. In particular, the probability that type \( t_i \) assigns to any measurable proper subset of the latter event does not matter.
As the notion of frames will play a central role in our characterization, it is helpful to go over a few examples to illustrate the definition.

**Example 5** (Finite-order frames). A \( k \)-order frame is a frame whose elements are \( k \)-order measurable events, that is, any two types with the same \( k \)-order beliefs must belong to the same partition element. Note that any profile of first-order measurable partitions is a (first-order) frame, as it can be readily verified from Definition 4. Examples of higher order frames can be constructed recursively, beginning with an arbitrary first-order frame. For instance, fix \( \theta \in \Theta \) and, for each player \( i \), fix a Borel measurable subset \( B_i \subseteq [0, 1] \). Then, partition \( T_i^* \) according to whether the probability that \( i \) assigns to \( \theta \) lies in \( B_i \). That is, consider the first-order frame \( \mathcal{P}_i = \{F_i, T_i^* \setminus F_i\} \), where

\[
F_i = \{t_i : \mu_i^*(\theta | t_i) \in B_i\}.
\]

Next, fix a Borel measurable subset \( B_i' \subseteq [0, 1] \) for each player \( i \) and consider the partition \( \mathcal{P}_i' = \{F_i', T_i^* \setminus F_i'\} \), where \( F_i' \) is the event that player \( i \) assigns a probability in \( B_i' \) to the event that the parameter is \( \theta \) and player \( -i \) assigns to \( \theta \) a probability in \( B_{-i} \). That is,

\[
F_i' = \{t_i : \mu_i^*(\theta \times F_{-i} | t_i) \in B_i'\}.
\]

Then, the join of the partitions \( \mathcal{P}_i \) and \( \mathcal{P}_i' \), which we denote by \( \mathcal{P}_i'' \), partitions \( T_i^* \) into four second-order measurable events, according to whether a type belongs to each of the events \( F_i \) and \( F_i' \). The profile \( \mathcal{P}'' = (\mathcal{P}_i'')_{i \in I} \) is an example of a second-order frame. Examples of higher order frames can be constructed in a similar fashion.

**Example 6** (Common belief frames). Partitioning types according to whether or not a given subset of parameters is common belief gives rise to a frame. Formally, for each \( E \subseteq \Theta \) and \( p \in [0, 1]^2 \), the profile of bi-partitions \( \mathcal{P}_i = \{C_i^p(E), T_i^* \setminus C_i^p(E)\} \), \( i \in I \), is a frame. Indeed, if any two types of player \( i \) agree on the probabilities assigned to the elements of \( \Theta \times \mathcal{P}_{-i} \), then they must agree on the probability of the event \( E \times C_i^p(E) \). Thus, by property (1) of common beliefs, either both types belong to \( C_i^p(E) \), or neither does. More generally, given a frame \( \mathcal{P} = (\mathcal{P}_i)_{i \in I} \), a collection of events \( E_1, \ldots, E_m \in 2^\Theta \otimes \mathcal{P}_1 \otimes \mathcal{P}_2 \) and a collection of pairs \( p_1, \ldots, p_m \in [0, 1]^2 \), the profile of joined partitions

\[
\mathcal{P}_i \cup \{C_i^{p_1}(E_1), T_i^* \setminus C_i^{p_1}(E_1)\} \cup \cdots \cup \{C_i^{p_m}(E_m), T_i^* \setminus C_i^{p_m}(E_m)\}, \quad i \in I,
\]

is a frame, called a common belief frame.

**Example 7** (Strategic frames). Given a game \( G = (A_i, g_i)_{i \in I} \), the strategic frame associated with \( G \) is the profile of partitions \( \mathcal{P} = (\mathcal{P}_i)_{i \in I} \) where the elements of each \( \mathcal{P}_i \) are the equivalence classes of types with the same rationalizable actions in \( G \), that is,

\[
\mathcal{P}_i = \\{\{t_i : R_i(t_i, G) = A_i' \} : \emptyset \neq A_i' \subseteq A_i\}.
\]

\(^{24}\)Recall that the join of a pair of partitions \( \mathcal{P}_i \) and \( \mathcal{P}_i' \), written \( \mathcal{P}_i \cup \mathcal{P}_i' \) is the coarsest partition that is finer than both \( \mathcal{P}_i \) and \( \mathcal{P}_i' \).
The proof that $\mathcal{P}$ is indeed a frame, given in Appendix A.3, is based on the observation that an action is rationalizable for a type of player $i$ only if it can be rationalized (at each order) by a conjecture $v \in \Delta(\Theta \times T_{-i}^* \times A_{-i})$ that assumes that all types of player $-i$ that share the same rationalizable actions behave in the same way. That is, for each $A_{-i}' \subseteq A_{-i}$ and $a_{-i} \in A_{-i}'$, the ratio $\nu(\theta \times E_{-i} \times a_{-i})/\nu(\theta \times E_{-i} \times A_{-i}')$ is the same for each parameter $\theta$ and each event $E_{-i} \subseteq \{t_{-i} : R_{-i}(t_{-i}, G) = A_{-i}'\}$. This observation implies that, in order for a type of player $i$ to rationalize an action in $G$, what matters is the probability assigned to events of the form $\theta \times \{t_{-i} : R_{-i}(t_{-i}, G) = A_{-i}'\}$, not how this probability is distributed within those events. As a result, all types of player $i$ who agree on the probabilities assigned to the elements of $\Theta \times \mathcal{P}_{-i}$ must have the same rationalizable actions and thus belong to the same element of $\mathcal{P}_i$.

We now define the notion of convergence that is used in the characterization of the strategic topology below. Given a frame $\mathcal{P} = (\mathcal{P}_i)_{i \in I}$, for each player $i$ set $d_{i,\mathcal{P}}^{0} = 0$ and, recursively, for each integer $k \geq 0$ and types $t_i$ and $t'_i$, define

$$d_{i,\mathcal{P}}^{k+1}(t_i, t'_i) = \inf \left\{ \delta > 0 : \mu_i^*(E | t_i) \leq \mu_i^*(E_{\mathcal{P}}^k | t'_i) + \delta \right\} \quad \forall E \in 2^{\Theta} \otimes \mathcal{P}_{-i},$$

where $E_{\mathcal{P}}^k$ denotes the $\delta$-neighborhood of event $E$ at order $k$ relative to $\mathcal{P}$, that is,

$$E_{\mathcal{P}}^k = \{(\theta, t'_{-i}) \in \Theta \times T_{-i}^* : d_{i,\mathcal{P}}(t_{-i}, t'_{-i}) < \delta \text{ for some } t_{-i} \text{ with } (\theta, t_{-i}) \in E\}.$$

Thus, the definition of $d_{i,\mathcal{P}}^k$ is similar to that of the $k$-order Prohorov distance $d_{i}^k$, but restricts the events for which the proximity is measured to those in the frame $\mathcal{P}$.\(^{26}\)

**Definition 5** (Uniform weak convergence relative to a frame). Let $\mathcal{P} = (\mathcal{P}_i)_{i \in I}$ be a frame. A sequence of types $t^n_i$ converges to a type $t_i$ \textit{uniformlyweakly} on $\mathcal{P}$ if

$$d_{i,\mathcal{P}}(t_i, t^n_i) = \sup_{k \geq 1} d_{i,\mathcal{P}}^k(t_i, t^n_i) \to 0 \quad \text{as} \quad n \to \infty.$$

With this definition in place, we are ready to state our main result, whose proof is presented in Appendix A.4.

**Theorem 2** (Characterization of strategic convergence). A sequence of types converges strategically if and only if it converges uniformlyweakly on every frame.
draws attention to a particular form of uniform convergence across the levels of the belief hierarchy as a condition for robustness. While it is expected that some form of uniform convergence should play a role, it is much less clear at the outset what kind of uniformity would ultimately lead to a characterization. The theorem highlights the role of frames as the coarse information structures that are relevant for strategic convergence.

Second, the characterization provides a practical criterion for robustness, similar to the role of common $p$-beliefs in the complete information case. From the applied perspective, the if direction of the theorem is more useful. Indeed, it is often the case that an analyst is interested in a particular game, or a small parametrized family of games. If the perturbations that the analyst contemplates pass the robustness criterion provided by the theorem, then the analyst is reassured that the predictions of the unperturbed model are robust in the game(s) of interest. However, if the perturbation fails the criterion, then the theorem only implies that there is some payoff structure for which the predictions of the unperturbed model fail to be robust. Failure of the criterion thus “raises a flag” to the analyst, but is ultimately inconclusive: further examination is required to determine whether the predictions of interest are robust. Nonetheless, the only if direction of the result is still useful, as it ensures the criterion is sharp: among all sufficient conditions for robustness that do not make use of the payoff structure, the criterion of Theorem 2 is one that minimizes the instances when the “flag is raised” unnecessarily.

The applied perspective put forward above leads naturally to a question: How easy is it to verify that a sequence of perturbations passes the robustness criterion of Theorem 2? This is ultimately a complexity question and it is hard to give a precise answer. Instead, we provide an example below to illustrate a relatively simple application of the criterion in a nontrivial case where the stronger (but easier-to-check) condition of uniform weak convergence fails. The economics of the example is interesting in its own right: the unperturbed type space is the one generated by the automated message exchange protocol of Rubinstein (1989), while the perturbed type space is generated by a variation of the same protocol in which the players, whom we interpret as boundedly rational, become “overwhelmed by evidence” once the number of messages exchanged reaches an arbitrarily large (but finite) threshold, in which case they believe the parameter is common knowledge.

Example 8 (Faulty message exchange with “boundedly rational” players). Suppose the unperturbed model is the message exchange type space \( \{t_i^0, t_i^1, t_i^2, \ldots\} \) of Example 2. Recall that in that type space there is no type that has common knowledge of \( \theta_1 \), regardless of the number of messages exchanged. The perturbed model is a variation of this type space in which the players mistakenly believe there is common knowledge of \( \theta_1 \) as soon as they have received \( n \) messages, where \( n \geq 1 \) is interpreted as the level of sophistication of the players.\(^{27}\) Formally, for each \( n \geq 1 \)

\(^{27}\)Here the words “mistakenly” and “sophistication” should be interpreted with care. To present the perturbed model we have used a bounded rationality story: unlike in the unperturbed model, the players are not able to apply Bayes’ rule exactly; once the number of messages exchanged exceeds a finite threshold, they believe there is common knowledge
we consider a finite type space \( \{s_{1}^{0,n}, s_{1}^{1,n}, \ldots, s_{1}^{n,n}\} \) where \( s_{1}^{n,n} \) has common knowledge of \( \theta_{1} \), and the beliefs of the remaining types mimic those of the type space in Example 2, as illustrated in Figure 3. For each player \( i \) and \( m \geq 0 \), does the sequence \( s_{i}^{m,n} \) converge strategically to \( t_{i}^{m} \) as \( n \to \infty \)?

\[
\begin{align*}
\delta_{1}^{0,n} & \xrightarrow{1} \theta_{0} \times s_{2}^{0,n} \\
\delta_{1}^{m,n} & \xrightarrow{q''} \theta_{1} \times s_{2}^{m-1,n} \\
\delta_{1}^{n,n} & \xrightarrow{1} \theta_{1} \times s_{2}^{n,n} \\
\end{align*}
\]

\[
\begin{align*}
s_{2}^{0,n} & \xrightarrow{q'} \theta_{0} \times s_{1}^{0,n} \\
s_{2}^{m,n} & \xrightarrow{q''} \theta_{1} \times s_{1}^{m,n} \\
s_{2}^{n,n} & \xrightarrow{1} \theta_{1} \times s_{1}^{n,n} \\
\end{align*}
\]

Figure 3: Type space for Example 8

The sequence \( s_{i}^{m,n} \) fails the uniform weak converge criterion. Indeed, it can be readily verified that the \((2n - 1 + i)\)-order Prohorov distance between \( s_{i}^{n,n} \) and \( t_{i}^{n} \) must exceed \( \min\{q', q''\} \) and hence, by an induction argument, for each \( 0 \leq m \leq n - 1 \) the \((4n - 2m - i + 1)\)-order Prohorov distance between \( s_{i}^{m,n} \) and \( t_{i}^{m} \) is also at least \( \min\{q', q''\} \).

Nonetheless, we show that the sequence converges strategically by verifying the criterion of Theorem 2: the sequence \( s_{i}^{m,n} \) converges to \( t_{i}^{m} \) uniform-weakly on every frame. In effect, for every frame \( \mathcal{P} = (\mathcal{P}_{i})_{i \in I} \) we show that \( \delta_{i}^{\text{uw}}(\mathcal{P}_{i}, \mathcal{P}_{i}^{m,n}, t_{i}^{m}) = 0 \) for every \( n \) sufficiently large. To see why, first note that, since each \( \mathcal{P}_{i} \) is finite, there exists \( N \) large enough that, for each \( n \geq N - 1 \) and player \( i \), the intersection \( \mathcal{P}_{i}(t_{i}^{n}) \cap \{t_{i}^{0}, t_{i}^{1}, t_{i}^{2}, \ldots\} \) is infinite.\(^{28}\) This implies that for every \( i \in I, n \geq N, k \geq 1 \) and \( \delta > 0 \), type \( s_{-i}^{n,n} = C_{-i}^{1,1}(\theta_{1}) \) belongs to the \( k \)-order Prohorov \( \delta \)-neighborhood of each of the frame elements

\[
\mathcal{P}_{-i}(t_{-i}^{n-1}), \quad \mathcal{P}_{-i}(t_{-i}^{n}) \quad \text{and} \quad \mathcal{P}_{-i}(t_{-i}^{n+1}),
\]

as each of these sets, having an infinite intersection with \( \{t_{-i}^{0}, t_{-i}^{1}, t_{-i}^{2}, \ldots\} \), must contain some type \( t_{-i}^{m} \) that has the same \( k \)-order beliefs as \( s_{-i}^{n,n} \).\(^{29}\) As the support of \( \mu_{i}^{*}(\cdot|t_{i}^{n}) \) is contained in of \( \theta_{1} \), rather than just mutual knowledge up to a finite order. However, bounded rationality is only a story here, as the perturbed model is still a standard type space.

\(^{28}\)Following standard notation, \( \mathcal{P}_{i}(t_{i}) \) designates the element of \( \mathcal{P}_{i} \) containing \( t_{i} \).

\(^{29}\)Recall from Example 2 that \( s_{i}^{n,n} \) and \( t_{i}^{m} \) are at zero \((2m - 2 + i)\)-order Prohorov distance.
the strategic frame of $G$

each player’s partition must contain an element that is open in the product topology, namely, any

converge uniform weakly on the strategic frame of $G$

must also converge strategically in game $G$

hold: convergence of rationalizable behavior in a given game $G$

behavior of $t$

game-by-game version of the

if $d$

if $d$

Therefore,

if $d$

demonstrated above), imply that $d^k_{i,\mathcal{D}}(t^m_i, s^m_i) = 0$ for every $k \geq 1$ and $n \geq N$. Next, since

the beliefs of $s^{0,n}_i, \ldots, s^{n-1,n}_i$ mimic those of $t^0_i, \ldots, t^{n-1}_i$ (compare Figures 2 and 3), for every

$0 \leq m \leq n - 1$ we have $d^1_{i,\mathcal{D}}(t^m_i, s^m_i) = 0$, and for every $k \geq 1$,

$$d^{k+2}_{i,\mathcal{D}}(t^0_i, s^0_i) \leq \max \left\{ d^k_{i,\mathcal{D}}(t^1_i, s^1_i), d^k_{i,\mathcal{D}}(t^0_i, s^0_i) \right\},$$

and for every $1 \leq m \leq n - 1$,

$$d^{k+2}_{i,\mathcal{D}}(t^m_i, s^m_i) \leq \max \left\{ d^k_{i,\mathcal{D}}(t^{m+1}_i, s^{m+1}_i), d^k_{i,\mathcal{D}}(t^m_i, s^m_i), d^k_{i,\mathcal{D}}(t^{m-1}_i, s^{m-1}_i) \right\}.$$ 

These conditions, combined with the fact that $d^k_{i,\mathcal{D}}(t^n_i, s^n_i) = 0$ for every $k \geq 1$ and $n \geq N$

(demonstrated above), imply that $d^k_{i,\mathcal{D}}(t^m_i, s^m_i) = 0$ for every $k \geq 1$, $n \geq N$ and $0 \leq m \leq n$.

Therefore, $d^{uw}_{i,\mathcal{D}}(t^m_i, s^m_i) = 0$ for every $n \geq N$ and $0 \leq m \leq n$. Since the frame $\mathcal{D}$ is arbitrary,

we conclude that $t^{m,n}_i$ converges to $t^m_i$ uniform-weakly on every frame.

It is worthwhile discussing the role of strategic frames in the characterization. They play a central role in the if direction of the result. In fact, the proof in Appendix A.4 implies a stronger, game-by-game version of the if direction: given any game $G$ and any type $t_i$, the rationalizable behavior of $t_i$ in game $G$ is robust to any perturbation $t^n_i$ that converges to $t_i$ uniform-weakly on the strategic frame of $G$.

However, the game-by-game version of the only if direction fails to hold: convergence of rationalizable behavior in a given game $G$ does not imply uniform weak convergence on the strategic frame of $G$. To see why, pick any game $G$ in which the payoff of each player depends on his own action and on the parameter only, so that the strategic frame of $G$ is a nontrivial first-order frame. Then, any sequence of types whose first order beliefs converge must also converge strategically in game $G$. But this clearly does not imply that the sequence will converge uniform weakly on the strategic frame of $G$, as it does not even imply convergence at second order.

A related comment is that uniform weak convergence on all strategic frames is equivalent to strategic convergence. This follows directly from the stronger version of the if direction discussed above, and the only if direction of Theorem 2. However, we find such a characterization less interesting than the one stated in the theorem, as we do not have a characterization of strategic frames that does not make explicit reference to the rationalizable correspondence, hence we regard strategic frames as “endogenous” objects. By contrast, general frames are defined in terms of a purely epistemic condition. In fact, not every frame is a strategic frame. In a strategic frame, each player’s partition must contain an element that is open in the product topology, namely, any

\[ \theta_1 \times \{ t^{n-1}_{i-1}, t^n_{i-1}, t^{n+1}_{i-1} \}, \]

it follows that

$$\mu^{\star}_i(E^{\delta,k}_i | x^{n,n}_i) \geq \mu^{\star}_i(E^{\delta,k}_i | x^{n,n}_i) \geq \mu^{\star}_i(E | x^n_i) - \delta \quad \forall E \in 2^\Theta \otimes \mathcal{P}_1 \otimes \mathcal{P}_2,$$

and since $\delta > 0$ is arbitrary, we have $d^k_{i,\mathcal{D}}(t^n_i, s^n_i) = 0$ for every $k \geq 1$ and $n \geq N$. Next, since

the beliefs of $s^{0,n}_i, \ldots, s^{n-1,n}_i$ mimic those of $t^0_i, \ldots, t^{n-1}_i$ (compare Figures 2 and 3), for every

$0 \leq m \leq n - 1$ we have $d^1_{i,\mathcal{D}}(t^m_i, s^m_i) = 0$, and for every $k \geq 1$,

$$d^{k+2}_{i,\mathcal{D}}(t^0_i, s^0_i) \leq \max \left\{ d^k_{i,\mathcal{D}}(t^1_i, s^1_i), d^k_{i,\mathcal{D}}(t^0_i, s^0_i) \right\},$$

and for every $1 \leq m \leq n - 1$,

$$d^{k+2}_{i,\mathcal{D}}(t^m_i, s^m_i) \leq \max \left\{ d^k_{i,\mathcal{D}}(t^{m+1}_i, s^{m+1}_i), d^k_{i,\mathcal{D}}(t^m_i, s^m_i), d^k_{i,\mathcal{D}}(t^{m-1}_i, s^{m-1}_i) \right\}.$$ 

By this we mean that (a) implies (b) in Definition 1.
element consisting of types whose set of rationalizable actions is minimal (w.r.t. set inclusion). General frames, however, need not have elements that are open sets. Letting $B_i$ be the set of rational numbers in the first-order frame of Example 5 illustrates this fact.

Finally, Theorem 2 enables a connection with the early literature on robustness and, in particular, with the notion of common $p$-beliefs. This connection is the subject of the next section.

### 3.4 Connection to Common Beliefs

Previous work on robustness of Bayesian equilibrium in common prior settings (Monderer and Samet, 1989, 1996; Kajii and Morris, 1998) has proved equivalences between ex ante notions of strategic convergence and convergence of common beliefs. In this section, we reexamine this connection in our interim, non-common prior framework.

**Definition 6** (Common belief convergence). A sequence of types $t^n_i$ converges in common beliefs to a type $t_i$ if for every $\delta > 0$ and $k \geq 1$ there exists $N \geq 1$ such that for every $n \geq N$, $p \in [0, 1]^2$ and measurable $E \subseteq \Omega$,

$$t_i \in C^p_i(E) \implies t^n_i \in C^{p-(\delta, \delta)}_i (E^{\delta,k}).$$

Thus convergence in common beliefs requires each event that is common $p$-belief for the limit type to have, at each order $k$, a $\delta$-neighborhood that is common-$(p - (\delta, \delta))$-belief for all types that lie sufficiently far in the tail of the sequence, where “sufficiently far” depends on $\delta$ and on the order $k$, but is independent of the event $E$ and the degree of common belief $p$.

To shed light on the definition, it is useful to begin with a formal analogy with product convergence. Recall that a sequence of types $t^n_i$ converges to a type $t_i$ in the product topology if and only if for every $\delta > 0$ and $k \geq 1$ there exists $N \geq 1$ such that for every $n \geq N$, $p \in [0, 1]$ and measurable set $E \subseteq \Omega$,

$$t_i \in B^p_i(E) \implies t^n_i \in B^{p-\delta}_i (E^{\delta,k}).$$

Thus, the definition of convergence in common beliefs formally replaces the belief operator $B^p_i$ in the characterization of product convergence above, by the common belief operator $C^p_i$. This

---

31That such an element is product-open follows from minimality and the upper hemi-continuity of the rationalizable correspondence in the product topology.

32If we modify the definition by restricting the class of events $E$ to be the class of sets that are closed in the product topology, then the notion of convergence remains the same. Indeed, letting $\overline{E}$ denote the product-topology closure of $E$, we have $C^p_i(E) \subseteq C^p_i(\overline{E})$ and $E^{\delta,k} = (\overline{E})^{\delta,k}$.

33This follows directly from the following two facts: (i) the Mertens-Zamir isomorphism $\mu_i : T^*_j \rightarrow \Delta(\Theta \times T^*_j)$ becomes a homeomorphism when each $T^*_j$ is endowed with the product topology and $\Delta(\Theta \times T^*_j)$ is endowed with the topology of weak convergence (as remarked in footnote 15); (ii) the Prohorov metric on $\Delta(\Theta \times T^*_j)$ metrizes the topology of weak convergence.
strengthens product convergence by imposing discipline on the tails of the belief hierarchy, which are irrelevant for product convergence. Indeed, notice that convergence in common beliefs implies product convergence, because setting $p = (p, 0)$ yields $B_i^p = C_i^p$. As for the connection with strategic convergence, Theorem 2 yields the following corollary, proved in appendix A.5. The result follows from the fact that uniform weak convergence in every common belief frame (Example 6) implies convergence in common beliefs, which we establish in appendix A.5 through a sequence of lemmas.

**Corollary 1** (Common belief convergence). Strategic convergence implies convergence in common beliefs.

However, *the converse does not hold*, as we show with a striking example in appendix A.6. In particular, we construct a type $t_1$ that, for some $0 < p < q < 1$, has common $(q, p)$-belief of no nontrivial event. We then exploit this fact to construct a sequence $t_1^n$ such that the probability assigned to a certain pair $(\theta_0, t_2)$ drops to $q - \Delta \geq p$ for some $\Delta > 0$ under $t_1^n$, while ensuring that all events commonly $p$-believed at $t_1$ for some $p \in (0, 1]^2$ remain so at $t_1^n$. As the probability assigned to $(\theta_0, t_2)$ under $t_1^n$ and $t_1$ differ by $\Delta$, we are then able to construct a frame on which $t_1^n$ fails to converge to $t_1$ uniform-weakly, and hence conclude by Theorem 2 that $t_1^n$ does not converge strategically to $t_1$.

The corollary and the counterexample to its converse shed further light on the robustness problem by providing a necessary condition for strategic convergence that, albeit not sufficient, is easier to interpret than the more abstract notion of uniform weak convergence on frames. Furthermore, they clarify the connection between the strategic topology for rationalizability on the universal type space and the ex ante strategic topologies for Bayesian equilibrium in common prior type spaces of the early literature. In particular, the example shows that the equivalence between strategic robustness and common $p$-beliefs breaks down when we adopt the interim approach.

Finally, the result has implications for the analysis of critical types, which Ely and Pęski (2011) define as those types to which product convergence fails to imply strategic convergence. Indeed, Ely and Pęski (2011, Theorem 1) show that a type $t_i \in T_i^*$ is critical if and only if $t_i \in C_i^{p1}(E)$ for some $p > 0$ and some product-closed proper subset $E_i \subseteq T_i^*$. Given this characterization, Corollary 1 yields the following result concerning the genericity of critical types, which we prove in appendix A.7.

**Corollary 2** (Genericity of Critical Types). The set of critical types is open and dense in the universal type space under the strategic topology.

Ely and Pęski (2011) highlight a tension concerning critical types. On one hand, they are

---

34 The analogous definition for uniform-strategic convergence is: $t_i$ is uniformly critical if there is some sequence that converges to $t_i$ in the product topology but fails to converge uniform-strategically. An immediate implication of our Theorem 1 is that all types in the universal type space are uniformly critical, since for every type $t_i$ there is always a sequence that converges to $t_i$ in the product topology but does not converge uniform-weakly.
pervasive in applications—for instance, types in finite type spaces and common prior type spaces (both of which we discuss in the next section) are examples of critical types. On the other hand, they are topologically small: Ely and Pęski (2011, Theorem 2) show that their complement is a residual subset of the universal type space under the product topology. The result above shows, however, that this tension disappears when one considers strategic genericity. The criticality property is robust to strategic perturbations, while by positing a non-critical type one is bound to make strategic predictions that are not robust in some game. We defer a more detailed discussion of the interpretation of strategic genericity to the next section, where we present our genericity results concerning common prior types.

4 Implications for Finite Models and Common Priors

In this section we examine some implications of our results for two commonly used classes of models, namely finite type spaces and common prior type spaces.

4.1 Finite Types

It is common, in both theoretical and applied work, to model incomplete information scenarios by positing a finite type space. In addition to advantages in terms of tractability, this simplifying assumption has received some justification since Dekel, Fudenberg, and Morris (2006) show that finite types, those belonging to a finite type space, are dense in the universal type space under the strategic topology. That is, the strategic behaviors of any type can be approximated by the strategic behaviors of finite types.35

As we have seen above, strategic convergence is a strictly stronger condition than convergence in common beliefs. However, when the limit type is a finite type, we show that common belief convergence implies uniform weak convergence. We thus have the following theorem, proved in appendix A.8.

**Theorem 3.** Given a finite type \( t_i \) and a sequence of (possibly infinite) types \( t^n_i \), the following statements are equivalent:

(a) \( t^n_i \rightarrow t_i \) uniform-weakly;

(b) \( t^n_i \rightarrow t_i \) uniform-strategically;

(c) \( t^n_i \rightarrow t_i \) strategically;

(d) \( t^n_i \rightarrow t_i \) uniform-weakly on every frame;

35Chen, Di Tillio, Faingold, and Xiong (2010, Theorem 3) show, however, that they are nowhere dense in the uniform strategic topology.
(e) $t^*_i$ \rightarrow t_i$ in common beliefs.

Thus, when the unperturbed model is finite, any distinction between the uniform strategic and the strategic topology disappears. Furthermore, we obtain equivalence with common belief convergence, closing the gap with the literature on ex ante robustness in the case of perturbations of finite types. In other words, by assuming finite types, an analyst can guarantee robustness of his predictions to perturbations, provided the perturbations that he contemplates as possible preserve approximate common beliefs in the sense of Definition 6.

4.2 Common Prior Types

The common prior assumption, according to which the beliefs of all players are derived from the same prior by conditioning on their private information, is a cornerstone of virtually all models of information economics. However, its interpretation and methodological justification are controversial—see Morris (1995), Gul (1998) and Aumann (1998)—, and its positive implications often puzzling, as witnessed by the celebrated agreement theorem (Aumann, 1976) and the no-trade theorem that followed it (Milgrom and Stokey, 1982). Taking a more fundamental approach to the issue, Lipman (2003) investigates the extent to which models satisfying the assumption approximate models where it is violated, focusing on finite-order beliefs approximations. In this section we adopt a similar perspective, and apply our characterizations to revisit the problem, considering also strategic approximations.

We recall the following definition from Mertens and Zamir (1985).

**Definition 7.** A common prior is a probability measure $\pi \in \Delta(\Theta \times T^*_i \times T^*_j)$ such that, for each player $i$ and measurable sets $E \subseteq T^*_i$ and $F \subseteq \Theta \times T^*_{-i}$,

$$
\pi(E \times F) = \int_E \mu^*_i(F|t_i)\pi(dt_i).
$$

Thus, $\pi$ is a common prior if, for each player $i$, the function $t_i \mapsto \mu^*_i(F|t_i)$ is (a version of) the conditional probability under $\pi$ of the event $F$, given the type of player $i$. When $\pi$ assigns positive probability to a particular type $t_i$, Bayes rule applies given the event $E = \{t_i\}$, and the equation above can be stated in terms of the usual formula for conditional probability, that is,

$$
\mu^*_i(F|t_i) = \pi(F|t_i) \frac{\pi(t_i \times F)}{\pi(t_i)}.
$$

As noted in Lipman (2003), there are two standard, related approaches to common priors in the literature. The first follows an ex ante perspective, as it formalizes the common prior assumption as a property of a type space. Let $\mathcal{T}$ denote the family of all type spaces $(T_i)_{i \in I}$ such that each set $T_i$ is closed in the product topology on $T_i^*$. 

27
Definition 8. A type space \((T_i)_{i \in I}\) is a common prior type space if there exists a common prior \(\pi\) such that, for each player \(i\), the set \(T_i\) is the support of the marginal of \(\pi\) on \(T_i^*\). Every type space in \(\mathcal{T}\) that is not a common prior type space is a non-common prior type space.

The support condition formalizes in a meaningful way the requirement that the beliefs of the types in \(T_i\) are consistent with a common prior. Indeed, as shown in our online appendix Chen, Di Tillio, Faingold, and Xiong (2015), there exist common priors whose marginals on \(T_i^*\) have the whole set \(T_i^*\) as their supports, hence without the support condition every type space is a common prior type space.

The second approach, adopted by Lipman (2003), follows the analogous interim perspective and applies to types.

Definition 9. A type \(t_i \in T_i^*\) is a common prior type if there is a common prior \(\pi\) with \(\pi(t_i) > 0\). Every type in \(T_i^*\) that is not a common prior type is a non-common prior type.

The positivity condition is needed for the definition of a common prior type to have a minimal bite—without the condition, every type is a common prior type (see again our online appendix Chen, Di Tillio, Faingold, and Xiong, 2015). Observe that a finite type space \(T = (T_i)_{i \in I}\) is a common prior type space if and only if, for each player \(i\), each type \(t_i \in T_i\) is a common prior type. That is, in the case of a finite type space, the ex ante and the interim approach deliver the same formalization of the common prior assumption.

Lipman (2003) shows that the set of finite common prior types is dense in the product topology, but explicitly warns that the latter should not be interpreted as providing a foundation for the common prior assumption: Although every type can be approximated by a (finite) common prior type in the product topology, the strategic behavior of that type can be very different from the strategic behavior of any approximating common prior type. Our next theorem shows that this lack of robustness is, in effect, a generic phenomenon in the universal type space. In particular, the denseness result of Lipman (2003) is reversed, once we consider the strategic rather than the product topology. The proof of the result is in appendix A.9.

Theorem 4 (Genericity of non-common prior types). The set of non-common prior types contains a set of types that is open and dense in \(T_i^*\) under the strategic topology.

To explain the implications of the result, let us first interpret and elaborate on the openness and denseness conditions. The fact that the set of non-common prior types contains a (strategically) open set means that no non-common prior type in that set can be approximated, in terms of strategic behaviors, by common prior types. Thus, if we cannot exclude that the “true” type lies in

---

36This means that \(T_i\) is the smallest subset of \(T_i^*\) that is closed in the product topology on \(T_i^*\) and has probability 1 under \(\pi\). The faulty message exchange type space of Example 2 becomes a common prior type space, once for each player \(i\) we add the type \(C_i^{-1}(1,1)(\theta_i)\) to the set of types of player \(i\).
that set, then by assuming a common prior type we cannot guarantee in advance (i.e. whatever the
game is) that our predictions will be close to those regarding the true type. Now, by denseness,
*every* type is arbitrarily close (in the strategic topology) to a non-common prior type in the set.
Thus, unless we have sufficiently detailed knowledge of the true type’s strategic behavior in each
game, then indeed we cannot exclude that the true type is a non-common prior type. Strategic
genericity of non-common prior types can then be stated in non-technical terms as follows: By
assuming a common-prior type, a modeler has—from his own, imperfect view of the true type’s
strategic behaviors—no advance guarantee that his behavioral predictions will be close to those
regarding the true type that he is trying to approximate.

It is important to note that the above interpretation does not rely on assuming that the modeler’s
notion of proximity of types is the one formalized by strategic convergence. Indeed, assume,
in the spirit of Weinstein and Yildiz (2007), that the modeler’s imprecise knowledge concerns
the true type’s higher-order beliefs, rather than strategic behaviors. Since strategic convergence
implies product convergence, denseness in the strategic topology implies denseness in the product
topology. Thus, the following interpretation also holds: By assuming a common-prior type, the
modeler has—from his own, imperfect view of the true type’s higher-order beliefs or even just his
strategic behaviors—no advance guarantee that his behavioral predictions will be close to those
regarding the true type that he is trying to approximate.

A caveat of Theorem 4 is that many economic models assume an uncountable common prior
type space with a nonatomic prior, so that no type is a common prior type according to our defini-
tion. Since Theorem 4 does not apply to such cases, it is natural to wonder whether genericity of
non-common prior types can then be stated in non-technical terms as follows: By
assuming a common-prior type, the
topology. Thus, unless we have sufficiently detailed knowledge of the true type’s strategic behavior in each
game, then indeed we cannot exclude that the true type is a non-common prior type. Strategic
genericity of non-common prior types can then be stated in non-technical terms as follows: By
assuming a common-prior type, a modeler has—from his own, imperfect view of the true type’s
strategic behaviors—no advance guarantee that his behavioral predictions will be close to those
regarding the true type that he is trying to approximate.

A caveat of Theorem 4 is that many economic models assume an uncountable common prior
type space with a nonatomic prior, so that no type is a common prior type according to our defini-
tion. Since Theorem 4 does not apply to such cases, it is natural to wonder whether genericity of
non-common priors critically depends on the interim perspective. To address this issue, we now
turn to the ex ante perspective and consider topologies on type spaces rather than types. Specifi-
cally, let \( T^* = \times_i \subset T_i^* \) and identify each type space \( (T_i)_{i \in I} \) with the rectangle \( \times_i \subset T_i \subseteq T^* \).
Then \( T \) can be viewed as a family of (product)-closed subsets of \( T^* \). Let \( d_i^p \) and \( d_i^s \) be any
distances on \( T_i^* \) that metrize the product and the strategic topology, respectively. For each
t = (t_i)_{i \in I} and \( t' = (t'_i)_{i \in I} \) in \( T^* \), let

\[
\begin{align*}
    d^p(t, t') &= \max_{i \in I} d_i^p(t_i, t'_i) \quad \text{and} \quad d^s(t, t') = \max_{i \in I} d_i^s(t_i, t'_i).
\end{align*}
\]

The topologies on \( T \) that we consider are the Hausdorff topologies on \( T \) corresponding to the
distances \( d^p \) and \( d^s \) on \( T^* \). For each \( T, T' \in T \) and \( \bar{d} \in \{d^p, d^s\} \), let

\[
    h_{\bar{d}}(T, T') = \max \left\{ \sup_{t \in T} \inf_{t' \in T'} d(t, t'), \sup_{t' \in T'} \inf_{t \in T} d(t, t') \right\}.
\]

**Definition 10.** The *product topology on type spaces* is the topology on \( T \) induced by the metric
\( h_{d^p} \). The *strategic topology on type spaces* is the topology on \( T \) induced by the metric
\( h_{d^s} \).

---

Dekel, Fudenberg, and Morris (2006) show that the strategic topology on \( T_i^* \) is metrizable by the distance \( d_i^s \)
declared as follows. For each game \( G = (A_j, g_j)_{j \in I}, a_i \in A_i \) and \( t_i \in T_i^* \), let \( e_i(t_i, a_i, G) = \inf \{ e : a_i \in \mathcal{R}_i(t_i, G, e) \} \). For each \( m \geq 1 \), write \( G^m \) for the set of games where, for every \( j \in I \), the set \( A_j \) has at most \( m \) elements.
Then, for each \( t_i, t'_i \in T_i^* \),

\[
    d_i^s(t_i, t'_i) = \sum_{m \geq 1} 2^{-m} \sup_{G=(A_j, g_j)_{j \in I} \in G^m} \max_{a_i \in A_i} | e_i(t_i, a_i, G) - e_i(t'_i, a_i, G) |.
\]
In other words, a sequence of type spaces $T^n$ converges to a type space $T$ in the product topology on type spaces if for every $\delta > 0$, $k \geq 1$ and sufficiently large $n$, the $k$-order belief of each type in $T$ is $\delta$-close to the $k$-order belief of some type in $T^n$, and vice versa, the $k$-order belief of each type in $T^n$ is $\delta$-close to the $k$-order belief of some type in $T$. Similarly, $T^n$ converges to $T$ in the strategic topology on type spaces if for every $\varepsilon > 0$, game $G$ and sufficiently large $n$, the rationalizable actions of each type in $T$ are $\varepsilon$-rationalizable for some type in $T^n$, and vice versa, the rationalizable actions of each type in $T^n$ are $\varepsilon$-rationalizable for some type in $T$.

We are now ready to present our last theorem, which we prove in appendix A.9.

Theorem 5 (Genericity of non-common prior type spaces). The set of non-common prior type spaces contains a set of type spaces that is open and dense in $T$ under both the product topology and the strategic topology on type spaces.

The theorem highlights a fundamental difference between the ex ante and the interim approach to common priors. If we change the perspective and consider type spaces as opposed to types, we obtain a reversal of Lipman’s (2003) denseness result even if we still take product convergence as the underlying notion of proximity of types. Moreover, the strategic genericity of non-common prior types extends to non-common prior type spaces. Thus, common prior models only serve to approximate a nongeneric set of incomplete information scenarios, whether we consider the approximation from the interim perspective (Theorem 4) or the ex ante perspective (Theorem 5). As a consequence of these results, we conclude that there is no general, principled approach to justify common priors based on strategic approximations. While there are strategic contexts where a common prior model does provide a good approximation, this needs to be assessed on a case-by-case basis.

5 Discussion: Interpretation of Results, and Extensions

5.1 Universal Quantifier over Payoffs

The definition of the strategic topology, on which our robustness exercise is based, has a universal quantifier over payoff structures. This has important implications for both the interpretation and the applicability of our results. As discussed in section 3.3, Theorem 2 provides a practical criterion for robustness that can be useful in situations where solving for the rationalizability correspondence globally in the universal type space is impractical. Although the condition is not necessary for robustness relative to a fixed payoff structure, the theorem shows that the condition is sharp: there is no way to improve it unless one looks for joint conditions on hierarchies of beliefs and payoffs. Of course, such a context-free approach comes with a caveat: when the analyst verifies that the condition fails, he needs to undertake a more in-depth examination of the robustness problem in the particular strategic context of interest.
In light of this caveat, it is natural to wonder whether it is possible to obtain characterizations that make more use of the payoff information, while maintaining a reasonable degree of tractability. First, consider the extreme opposite of our approach, namely to look for a characterization of robust predictions that has complete freedom to depend on the payoff functions of the players. In this case, no such a characterization can be insightful or useful, as it must boil down to a reformulation of the definition of robustness in the particular game. This issue is not peculiar to multi-agent settings. Already in a decision theoretic framework à la Savage (one-person game against nature), how does one characterize the robustness of a decision maker’s choices to sequences of perturbations of his subjective beliefs? A sharp sufficient condition (when the state space is compact and utilities are continuous) is that the sequence of perturbations converges to the unperturbed belief of the decision maker in the topology of weak convergence of probability measures. But if the goal is to find a necessary and sufficient condition among all those that are allowed to depend on the decision maker’s utility directly, then no such a characterization exists that will not amount to rephrasing the definition of robustness relative to the given utility function.

What does seem possible, and extremely desirable, is to develop characterizations of robustness for classes of games. That is, to pursue characterizations that lie between our context-free approach and the other extreme of full context dependence. One would fix a class of games satisfying a certain property, and ask what is the weakest sufficient condition (on sequences of perturbations) for robust predictions in all games satisfying the property. Thus, the characterization is allowed to rely on a property of payoffs, but not on further details of the payoff functions.\(^{38}\)

While developing characterizations for classes of games is important, we anticipate some challenges. For illustration, consider the following single-agent problem. Assume that the agent has two actions to choose: high or low. The agent’s payoffs depend on a fundamental variable \(\theta\) that takes finitely many possible values in the real line. A modeler believes that a single-crossing condition holds: \(\text{high}\) is strictly better for the agent if \(\theta\) is no less than a fixed cutoff \(\bar{\theta}\), while \(\text{low}\) is strictly better if \(\theta\) is below \(\bar{\theta}\). Thus, only games that satisfy this single-crossing condition (relative to a fixed order on \(\Theta\) and a fixed threshold \(\bar{\theta}\)) would concern the modeler. Denote this class of games by \(\mathcal{G}\). Say that a sequence of types/beliefs \(\mu^n\) converges to \(\mu\) \(\mathcal{G}\)-strategically if for every game \(G \in \mathcal{G}\) and every action \(a\) in \(G\), \(a\) is a best reply for \(\mu\) in \(G\) if and only if for every \(\epsilon > 0\) there exists \(N\) such that for every \(n \geq N\), \(a\) is an \(\epsilon\)-best response for \(\mu^n\) in \(G\). For this class of decision problems, we can show that \(\mathcal{G}\)-strategic convergence is equivalent to the following condition: \(\mu^n(\{\theta : \theta \geq \bar{\theta}\}) \to \mu(\{\theta : \theta \geq \bar{\theta}\})\) if \(\mu(\{\theta : \theta \geq \bar{\theta}\}) = 0\) or 1, and \(\mu^n(\theta) \to \mu(\theta)\) for every \(\theta\) if \(\mu(\{\theta : \theta \geq \bar{\theta}\}) \in (0, 1)\).\(^{39}\) While this suggests that characterizing robustness in terms

\(^{38}\)Such an approach has been pursued recently by Morris, Shin, and Yildiz (2015) in the context of global games.

\(^{39}\)A proof is available upon request. Note that belief convergence, i.e., \(\mu^n(\theta)\) converges to \(\mu(\theta)\) for every \(\theta\), implies strategic convergence but it is unnecessarily strong, whereas convergence of \(\mu^n(\{\theta : \theta \geq \bar{\theta}\})\) to \(\mu(\{\theta : \theta \geq \bar{\theta}\})\) and \(\mu^n(\{\theta : \theta < \bar{\theta}\})\) to \(\mu(\{\theta : \theta < \bar{\theta}\})\) is a necessary requirement, but it is too weak to ensure strategic convergence. To see the former, note that for \(\theta'\) and \(\theta''\) larger than \(\bar{\theta}\), the sequence of point masses at \(\theta''\) converges to the point mass at \(\theta'\) \(\mathcal{G}\)-strategically but not in beliefs; to see the latter, consider \(\Theta = \{-1, 0, 1\}\) and \(\bar{\theta} = 0\) such that \(u(a, 1) - u(a', 1) = 2\)
of beliefs is indeed possible for some classes of games, we note that even for this very simple class of one-player games satisfying a single-crossing condition the characterization is somewhat unwieldy, in that it rests sometimes on beliefs over coarse events, and sometimes on beliefs over fine events, depending on properties of the limit belief. In general, we do not know what properties of classes of games are likely to lead to tractable characterizations of robustness in terms of beliefs. This is an important direction for future research.

Finally, we remark that the trade-off between the sharpness of a sufficient condition and the complexity of its verification process is common in economic theory. A notable example, which we think serves as a good analogy to explain our methodological approach, is Blackwell’s celebrated theorem on the comparison of statistical experiments (Blackwell, 1951, 1953). If two experiments, $A$ and $B$, are ranked by Blackwell’s criterion so that $B$ is a garbling of $A$, then the decision maker must earn a higher expected utility from experiment $A$ than from $B$. But what if $B$ is not a garbling of $A$? Blackwell’s theorem only asserts that there will be some payoff function under which the decision maker’s payoff from $B$ is higher than from $A$. Nonetheless, Blackwell’s theorem is useful, as the garbling condition is both sharp and easy to check.

The literature on comparison of experiments has also found it useful to complement Blackwell’s theorem with analogous results that focus on subclasses of payoffs functions, as in Lehmann (1988) and, more recently, Quah and Strulovici (2009).

5.2 Robustness under Non-Strict Incentives

The definition of strategic convergence, on which our robustness exercise is based, requires that in every game, every action that is rationalizable for a type remains, for $\varepsilon > 0$ arbitrarily small, $\varepsilon$-rationalizable for any type sufficiently far in the tail of the sequence of perturbations. This condition is equivalent to requiring strictly rationalizable actions for a type to remain rationalizable for small perturbations. It is thus natural to wonder what can be said regarding stronger notions of robustness that require rationalizable actions that rely on indifferences to also be robust. The answer will depend on the notion of robustness. For instance, if we ask what sequences of types

$u(a, 0) - u(a', 0) = 1$, and $u(a, -1) - u(a', -1) = -2$. Let $\mu''(\theta = 0)$, $\mu'(\theta = 1)$, $\mu''(\theta = -1)$, and $\mu'(\theta = -1)$ be all equal to $1/2$. Then, $\mu''$ and $\mu'$ assign the same probability on the events $\{\theta \geq \bar{\theta}\}$ and $\{\theta < \bar{\theta}\}$; however, the sequence $(\mu^n)$ with $\mu_n = \mu''$ for every $n$ does not converge to $\mu'$ $\mathcal{G}$-strategically.

\footnote{Other examples of similar approaches in economic theory are stochastic orders such as first-order and second-order stochastic dominance. Also, the revealed preference exercise in decision theory is predicated on similar ideas: Even though the analyst might be interested on a very small class of decision problems, the axioms typically assume knowledge of the decision maker’s behavior in a universal domain of decisions that might go well beyond the context of interest. Finally, the celebrated folk theorem for repeated games with imperfect public monitoring states that pairwise identifiability is a sufficient condition for the patient limit of the set of equilibrium payoffs to equal the set of all feasible individually rational payoffs (when that set has nonempty interior). The condition is sharp, in that when it fails there is some game with full dimensional payoffs such that the folk theorem fails. But, for many games, pairwise identifiability is stronger than necessary.}
converge in the sense that every action is robustly rationalizable in every game, then nothing interesting can be said: only constant sequences converge, that is, we get the discrete topology.

Nonetheless, we do think it would be fruitful to study weaker notions of robustness of non-strictly rationalizable actions, and we believe that the strategic topology should play a central role in such analysis. What we have in mind is an analogy with the refinements literature on Nash equilibrium, which aims at characterizing strategies (or connected components of strategies) that are robust to small perturbations of behavior (the "trembles"), or equivalently, perturbations of payoffs (Kohlberg and Mertens, 1986). In the refinements literature, the space of perturbations is finite-dimensional, so perturbations simply mean convergence in the Euclidean topology, which also turns out to be the weak topology generated by continuity of the strict Nash equilibrium correspondence. In other words, in the refinements literature, strict equilibria are, by definition, robust. And the only question is: in a given game, which non-strict equilibria are robust to perturbations in the topology that renders strict equilibria robust in every game (i.e. the Euclidean topology). A similar question can be asked in the context of our exercise, which considers perturbations of the players’ beliefs and higher-order beliefs. By analogy, it seems natural to expect that progress in understanding which (possibly non-strictly) rationalizable actions are robust in a given game, we first need to understand the weak topology on types generated by the continuity of the strictly rationalizable correspondence.

5.3 Extension to Compact-Continuous Games

Interim correlated rationalizability is well defined when the space \( \Theta \) and the action sets \( A_i \) are compact metrizable and the payoff functions are continuous: the recursive and fixed-point definitions coincide, and are non-empty valued.

The characterization of the strategic topology in terms of uniform weak convergence on frames (Theorem 2) continues to hold in the compact-continuous case. The characterization of uniform strategic convergence (Theorem 1) also remains valid, provided the definition of uniform strategic convergence is modified to require the rate of convergence to be uniform on all classes of compact continuous games whose payoffs are not only uniformly bounded, but also equicontinuous in \( \Theta \).\(^{41}\)

A Appendix

In the appendix we use the following notations. Given two measurable spaces \( X, Y \) and a probability measure \( \mu \in \Delta(X \times Y) \), we write \( \text{marg}_X \mu \) for the marginal of the distribution \( \mu \) on \( X \), that is, for every measurable \( E \subseteq X \), \( (\text{marg}_X \mu)(E) = \mu(E \times Y) \). Given a measurable correspondence

\(^{41}\)Naturally, the definitions of \( d_i^{\Theta, p} \) and \( d_i^q \) should no longer be the discrete metric on \( \Theta \), but the Prohorov metric on \( \Theta \) instead.
\(\varsigma : X \Rightarrow Y\), we define graph \(\varsigma = \{ (x, y) \in X \times Y : y \in \varsigma(x) \}\). If \(X\) is finite and \(\nu\) is a probability measure on \(X\), we write \(\text{supp} \nu\) to denote the support of \(\nu\), that is, \(\text{supp} \nu = \{ x \in X : \nu(x) > 0 \}\). Finally, given a game \(G = (A_i, g_i)_{i \in I}\), a probability distribution \(\beta \in \Delta(\Theta \times A_{-i})\) and \(\epsilon \geq 0\), we write \(BR_i(\beta, G, \epsilon)\) for the set of \(\epsilon\)-best replies to \(\beta\), that is, the set of all \(a_i \in A_i\) such that \(g_i(a_i, \beta) \geq g_i(a'_i, \beta) - \epsilon\) for all \(a'_i \in A_i\).

### A.1 Properties of rationalizability

Our proofs below use the following properties of interim correlated rationalizability. First, in the definition of the \(k\)-order \(\epsilon\)-rationalizable actions of a type \(t_i\) we can replace the conjectures of the form \(\nu \in \Delta(\Theta \times T_{-i}^* \times A_{-i})\) with those of the form \(\sigma_{-i} : \Theta \times T_{-i}^* \rightarrow \Delta(A_{-i})\). The two kinds of conjectures are related by the disintegration formula:

\[
v(\theta \times E_{-i} \times a_{-i}) = \int_{E_{-i}} \sigma_{-i}((\theta, t_{-i})[a_{-i}] \mu^*_i(\theta \times dt_{-i}[t_i]),
\]

for every \(\theta \in \Theta\), \(a_{-i} \in A_{-i}\) and measurable subset \(E_{-i} \subseteq T_{-i}^*\).\(^{42}\) Thus, \(R_i^k(t_i, G, \epsilon)\) is the set of all \(a_i \in A_i\) for which there is a measurable function \(\sigma_{-i} : \Theta \times T_{-i}^* \rightarrow \Delta(A_{-i})\) that satisfies:

\[
\text{supp} \sigma_{-i}(\theta, t_{-i}) \subseteq R_{-i}^{k-1}(t_{-i}, G, \epsilon) \quad \forall (\theta, t_{-i}) \in \Theta \times T_{-i}^*,
\]

\[
\int_{\Theta \times T_{-i}^*} \left[ g_i(a_i, \sigma_{-i}(\theta, t_{-i}), \theta) - g_i(a'_i, \sigma_{-i}(\theta, t_{-i}), \theta) \right] d\mu^*_i(\theta, t_{-i}[t_i]) \geq -\epsilon \quad \forall a'_i \in A_i.
\]

(See Corollary 1 in Dekel, Fudenberg, and Morris, 2007.) For future reference, a measurable function \(\sigma_{-i}\) that satisfies (2) is called a \((k-1)\)-order \(\epsilon\)-rationalizable conjecture; a measurable function \(\sigma_{-i}\) such that \(\text{supp} \sigma_{-i}(\theta, t_{-i}) \subseteq R_{-i}^{k-1}(t_{-i}, G, \epsilon)\) for every \((\theta, t_{-i}) \in \Theta \times T_{-i}^*\) is called an \(\epsilon\)-rationalizable conjecture.

Second, we can define rationalizability in terms of iterated eliminations of strongly dominated actions. For each \(i \in I\) and \(t_i \in T_{i}^*\), let \(S_i^0(t_i, G, \epsilon) = A_i\) and, recursively for \(k \geq 1\), define \(S_i^k(t_i, G, \epsilon)\) as the set of all \(a_i \in A_i\) such that for every mixed deviation \(\alpha_i \in \Delta(A_i)\) there is a conjecture \(\sigma_{-i} : \Theta \times T_{-i}^* \rightarrow \Delta(A_{-i})\) that satisfies:

\[
\text{supp} \sigma_{-i}(\theta, t_{-i}) \subseteq S_{-i}^{k-1}(t_{-i}, G, \epsilon) \quad \forall (\theta, t_{-i}) \in \Theta \times T_{-i}^*,
\]

\[
\int_{\Theta \times T_{-i}^*} \left[ g_i(a_i, \sigma_{-i}(\theta, t_{-i}), \theta) - g_i(\alpha_i, \sigma_{-i}(\theta, t_{-i}), \theta) \right] d\mu^*_i(\theta, t_{-i}[t_i]) \geq -\epsilon.
\]

Then,

\[
R_i^k(t_i, G, \epsilon) = S_i^k(t_i, G, \epsilon),
\]

---

\(^{42}\)Given a \(\nu\) that satisfies \(\text{marg}_{\Theta \times T_{-i}^*} \nu = \mu^*_i(t_i)\), the disintegration formula only pins down \(\sigma_{-i}\) up to a set of \(\mu^*_i(t_i)\)-probability zero. But, outside this null set, we can set \(\sigma_{-i}\) equal to a measurable selection from the correspondence \(R_i^{k-1}(\cdot, G, \epsilon)\), thus ensuring that (2) is satisfied everywhere provided \(\nu(\Theta \times \text{graph} R_i^{k-1}) = 1\) (as opposed to almost everywhere). The fact that such a measurable selection exists follows from the upper hemi-continuity of \(R_i^{k-1}(\cdot, G, \epsilon)\) (in the product topology) and the Kuratowski-Nyll-Nardzewski Selection Theorem.
and thus,

\[ R_i(t_i, G, \varepsilon) = \bigcap_{k \geq 1} S^k_i (t_i, G, \varepsilon). \]

(See Chen, Di Tillio, Faingold, and Xiong (2010), Proposition 1.) Likewise, \( a_i \in R_i(t_i, G, \varepsilon) \) if and only if, for every \( \alpha_i \in \Delta(A_i) \), there is an \( \varepsilon \)-rationalizable conjecture \( \sigma_{-i} : \Theta \times T_{-i}^* \rightarrow \Delta(A_i) \) that satisfies (3).

Finally, rationalizability has a characterization in terms of best-reply sets. Say a profile of correspondences \( \zeta_i : T_i^* \Rightarrow A_i, i \in I, \) has the \( \varepsilon \)-best-reply property if for each \( i \in I, t_i \in T_i^* \) and \( a_i \in \zeta_i(t_i) \) there exists \( v \in \Delta(\Theta \times T_i^* \times A_{-i}) \) such that the following conditions hold:

\[
\text{marg}_{\Theta \times T_i^*} v = \mu_i^*(t_i). \quad (\text{marg}_{T_i^* \times A_{-i}} v)[\text{graph } \zeta_{-i}] = 1, \quad a_i \in BR_i(\text{marg}_{A_{-i} \times \Theta} v, G, \varepsilon).
\]

Then \( R_i(\cdot, G, \varepsilon) \) is the greatest (w.r.t. pointwise set inclusion) profile of correspondences with the \( \varepsilon \)-best-reply property.\(^{44}\) (See Claim 2 in Dekel, Fudenberg, and Morris, 2007.)

### A.2 Proof of Theorem 1

The if direction is proved in Chen, Di Tillio, Faingold, and Xiong (2010). The proof of the only if direction relies on Lemma 1, Corollary 3 and Lemma 2 below.

We begin with some useful definitions and notations. Given a game \( G = (A_i, g_i)_{i \in I} \), say that an action \( a_i^0 \in A_i \) is a zero action for player \( i \) if \( g_i(a_i^0, a_{-i}, \theta) = 0 \) for all \( \theta \) and \( a_{-i} \). An action profile \( a^e = (a_i^e, a_{-i}^e) \) is a coordination pair if \( g_i(a_i^e, a_{-i}^e, \theta) = \max \theta g_i \) for all \( \theta \) and \( i \).

(In particular, any action of player \( i \) that is part of a coordination pair is rationalizable for any type of player \( i \).) Finally, let \( \beta_{t_i, \sigma_{-i}} \in \Delta(\Delta_{-i} \times \Theta) \) denote the belief of type \( t_i \) over actions of player \(-i\) and payoff-relevant parameters, when he has conjecture \( \sigma_{-i} : \Theta \times T_{-i}^* \rightarrow \Delta(A_{-i}) \), that is,

\[
\beta_{t_i, \sigma_{-i}}(a_{-i}, \theta) : = \int_{T_{-i}^*} \sigma_{-i}(\theta, t_{-i})[a_{-i}] \mu_i^*(\theta \times dt_{-i}|t_i) \quad \forall (a_{-i}, \theta) \in A_{-i} \times \Theta.
\]

The only if part of the theorem is a direct implication of Lemma 2 below. Lemma 1 and Corollary 3 are intermediate results.

**Lemma 1.** For every \( \varepsilon > 0 \), integer \( k \geq 1 \), player \( j \) and finite set of finite types \( \{t_{j,1}, t_{j,2}, \ldots, t_{j,N}\} \subset T_j^* \), there is a game \( G = (A_i, g_i)_{i \in I} \) with payoffs in the interval \([-5, 3]\), and a set of actions \( \{a_{j,1}^*, a_{j,2}^*, \ldots, a_{j,N}^*\} \subset A_j \), such that:

(i) every player \( i \) has a zero action \( a_i^0 \in A_i \);

---

\(^{43}\)When graph \( \zeta_i \) is not measurable, this expression is taken to mean that \( \text{marg}_{T_{-i} \times A_{-i}} v_i \) assigns probability one to a measurable subset of graph \( \zeta_i \).

\(^{44}\)Such greatest profile of correspondences is well defined, because the pointwise union of any family of profiles of correspondences with the \( \varepsilon \)-best-reply property must also have the \( \varepsilon \)-best-reply property.
(ii) there is a coordination pair $a^c \in A_1 \times A_2$ such that $d_j^c \not\in \{a^c_{j,1}, a^c_{j,2}, \ldots, a^c_{j,N}\}$ and $g_i (a_i, a^c_{-i}, \theta) \geq -2$ for every $a_i \in A_i$, $\theta \in \Theta$ and $i \in I$.

(iii) $R_j (\cdot, G, \gamma) = R_k^k (\cdot, G, \gamma)$ for every $i \in I$ and $\gamma \in [0, \frac{1}{2})$;

(iv) for every $1 \leq n \leq N$, $a^*_{j,n} \in R_j (t_{j,n}, G)$;

(v) for every $1 \leq n \leq N$ and $s_j \in T^*_j$ with $d_j^j (t_{j,n}, s_j) > \varepsilon$, $a^*_{j,n} \not\in R_j (s_j, G, \frac{1}{2})$.

Proof. The proof is by induction on $k$. Consider first $k = 1$ and $j = 1$ ($j = 2$ can be similarly proved). Fix a finite set of player 1’s types $\{t_{1,1}, t_{1,2}, \ldots, t_{1,N}\}$. Enumerate the nonempty subsets of $\Theta$ as $E_1, E_2, \ldots, E_L$. For each $(n, \ell) \in \{1, 2, \ldots, N\} \times \{0, 1, \ldots, L\}$ consider the function $\phi_{n,\ell} : \Theta \to [-1, 1]$ described in the following table:

<table>
<thead>
<tr>
<th></th>
<th>$\theta \in E_\ell$</th>
<th>$\theta \not\in E_\ell$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\ell = 0$</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>$\ell \geq 1$</td>
<td>$-(1 - \mu_1 (E_\ell</td>
<td>t_{1,n}))$</td>
</tr>
</tbody>
</table>

Thus, the functions $\phi_{n,\ell}$ define an auxiliary game between player 1 and Nature, where $\ell = 0$ is a safe bet for player 1, and $\ell \geq 1$ is a risky bet on the event $\theta \not\in E_\ell$. The rewards of the risky bets are such that:

- any type that has the same first-order beliefs as type $t_{1,n}$ is exactly indifferent between $\ell = 0$ and any $\ell \geq 1$;
- any type whose first-order belief is different from that of $t_{1,n}$ strictly prefers some risky bet $\ell \geq 1$ than the safe bet $\ell = 0$.

We use the functions $\phi_{n,\ell}$ to construct a game $G = (A_i, g_i)_{i \in I}$ to prove our claim for $k = 1$. In this game,

$$A_1 = (\{1, 2, \ldots, N\} \times \{0, 1, \ldots, L\}) \cup \{a^0_1, a^c_1\} \quad \text{and} \quad A_2 = \{1, 2, \ldots, N\} \cup \{a^0_2, a^c_2\}.$$  

Player 1’s payoffs are specified as follows:

- $a^0_1$ is a zero action for player 1;
- if player 1 chooses $a^c_1$, she gets 3 if player 2 chooses $a^c_2$, and gets 0 otherwise (regardless of $\theta$);
- if player 1 chooses $(n, \ell) \in \{1, \ldots, N\} \times \{0, \ldots, L\}$ and the state is $\theta$, she gets $\phi_{n,\ell} (\theta)$ if player 2 chooses $n$, and she gets $\phi_{n,\ell} (\theta) - 1$ if player 2 chooses any action different from $n$. 

36
Player 2’s payoffs are specified as follows:

- Player 2 gets 3 if \((a^e_1, a^e_2)\) is chosen (regardless of \(\theta\)), and gets 0 otherwise.

Thus, when the state is \(\theta\), we can draw the payoff matrix \(g(\cdot, \cdot, \theta)\) as follows:

\[
\begin{array}{|c|c|c|c|c|c|}
\hline
 & 1 & 2 & \cdots & N & a^0_j, a^e_j \\
\hline
(1, \ell) & \phi_{1,\ell}(\theta), 0 & \phi_{1,\ell}(\theta) - 1, 0 & \cdots & \phi_{1,\ell}(\theta) - 1, 0 & \phi_{1,\ell}(\theta) - 1, 0 \\
(2, \ell) & \phi_{2,\ell}(\theta) - 1, 0 & \phi_{2,\ell}(\theta), 0 & \cdots & \phi_{2,\ell}(\theta), 0 & \phi_{2,\ell}(\theta) - 1, 0 \\
\vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\
(N, \ell) & \phi_{N,\ell}(\theta) - 1, 0 & \phi_{N,\ell}(\theta), 0 & \cdots & \phi_{N,\ell}(\theta), 0 & \phi_{N,\ell}(\theta) - 1, 0 \\
\hline
a^0_1 & 0, 0 & 0, 0 & \cdots & 0, 0 & 0, 0 \\
a^e_1 & 0, 0 & 0, 0 & \cdots & 0, 0 & 3, 3 \\
\hline
\end{array}
\]

Since \(\phi_{n,\ell}(\theta) \in [-1, 1]\) for all \(n, \ell\) and \(\theta\), game \(G\) has payoffs bounded between \(-2\) and 3.

We claim that \(G\), along with the actions \(a^*_1, n := (n, 0), n = 1, \ldots, N\), satisfy properties (i)--(v) for \(k = 1\). Properties (i) and (ii) clearly hold. To prove (iii), first note that \(R_2(\cdot, G, \gamma) = R_2(\cdot, G, \gamma) = A_2\) for all \(\gamma \geq 0\). Indeed, the profile of correspondences \((\xi_t)_{t \in T}\), where \(\xi_2(t_2) = A_2\) and \(\xi_1(t_1) = \{a^0_1, a^e_1\}\) for all \(t_1\) and \(t_2\), has the \(\gamma\)-best-reply property. It follows that \(R_1(\cdot, G, \gamma) = R_1(\cdot, G, \gamma)\) for all \(\gamma \geq 0\). Hence, (iii) holds.

It remains to prove (iv) and (v). First, \((n, 0)\) is rationalizable for \(t_{1,n}\), since given the conjecture that player 2 plays \(n\), type \(t_{1,n}\) gets 0 by playing \((n, \ell)\) for any \(\ell\):

\[
\mu_1(E_{\ell}|t_{1,n}) \cdot \left(1 - \mu_1(E_{\ell}|t_{1,n})\right) + (1 - \mu_1(E_{\ell}|t_{1,n})) \cdot \mu_1(E_{\ell}|t_{1,n}) = 0,
\]

and gets at most 0 by playing any action not in \((n, 0), \ldots, (n, L)\). Thus, (iv) holds for \(a^*_n = (n, 0)\). Second, consider any type \(s_1\) with \(d_1^1(t_{1,n}, s_1) > \varepsilon\). Then, there exists some \(1 \leq \ell \leq L\) such that \(\mu_1(E_{\ell}|t_{1,n}) > \mu_1(E_{\ell}|s_1) + \varepsilon = \mu_1(E_{\ell}|s_1) + \varepsilon\). \(^{45}\) Then, given any conjecture about the behavior of player 2, the difference in expected payoffs between \((n, \ell)\) and \((n, 0)\) for type \(s_1\) is

\[
\mu_1(E_{\ell}|s_1) \cdot \left(1 - \mu_1(E_{\ell}|t_{1,n})\right) + (1 - \mu_1(E_{\ell}|s_1)) \cdot \mu_1(E_{\ell}|t_{1,n})
\]

\[
= \mu_1(E_{\ell}|t_{1,n}) - \mu_1(E_{\ell}|s_1) > \varepsilon.
\]

Hence, \(a^*_n = (n, 0)\) is not \(\varepsilon\)-rationalizable for type \(s_1\), which proves (v).

We now prove our claim for \(k + 1\) assuming that it holds for \(k\). Again, we assume \(j = 1\), and the proof for \(j = 2\) is similar. Let \(t_{1,1}, \ldots, t_{1,N}\) be arbitrary finite types of player 1.

Consider the finite set \(T_2 = \{t_{2,1}, t_{2,2}, \ldots, t_{2,N}\}\) of all types of player 2 that are assigned positive probability by some \(t_{1,n}\), for \(n = 1, \ldots, N\). By the induction hypothesis, we can find a game \(G = (A_i, g_i)_{i \in I}\), a set of actions \(\{a^*_2, a^*_{2,1}, \ldots, a^*_{2,N}\} \subset A_2\), and action profiles \(a^0\) and \(a^e\) in \(G\), that satisfy properties (i)--(v) relative to the finite set of finite types \(T_2\).

\(^{45}\)The equality follows because \(a^0 = 0\) and \(\Theta\) is endowed with the discrete metric.
Let $T^k_2 = \{t^k_{2,1}, t^k_{2,2}, \ldots, t^k_{2,N}\}$ be the set of $k$-order beliefs of types in $T_2$. Enumerate the nonempty subsets of $\Theta \times T^k_2$ as $E_1, E_2, \ldots, E_L$. For each $1 \leq \ell \leq L$, define

$$F_\ell = \{(a^*_{2,n'}, \theta) : 1 \leq n' \leq N', (\theta, t^k_{2,n'}) \in E_\ell\}.$$ 

For each $(n, \ell) \in \{1, \ldots, N\} \times \{0, \ldots, L\}$, define a function $\phi_{n,\ell} : A_2 \times \Theta \rightarrow [-1, 1]$ as in the following table:

<table>
<thead>
<tr>
<th>$\ell$</th>
<th>$(a_2, \theta) \in F_\ell$</th>
<th>$(a_2, \theta) \notin F_\ell$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\ell = 0$</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>$\ell \geq 1$</td>
<td>$- (1 - \mu_1(E_\ell</td>
<td>t_{1,n}))$</td>
</tr>
</tbody>
</table>

We use the game $G$ and the functions $\phi_{n,\ell}$ to define a new game $\tilde{G} = (\tilde{A}_i, \tilde{s}_i)$ to prove our claim for $k + 1$. In this game, 

$$\tilde{A}_1 = A_1 \cup \{1, \ldots, N\} \times \{0, 1, \ldots, L\} \cup \{a^0_{11}\} \text{ and } \tilde{A}_2 = A_2 \times \{0, 1, \ldots, N\}.$$ 

Player 1’s payoffs are specified as follows (see also the following table):

- $a^0_1$ is a zero action for player 1;
- if player 1 chooses $a_1 \in A_1$ and the state is $\theta$, he gets $g_1(a_1, a_2, \theta)$ if player 2 chooses $(a_2, 0) \in A_2 \times \{0\}$, and gets $g_1(a_1, a_2, \theta) - 3$ otherwise.
- if player 1 chooses $(n, \ell)$ and the state is $\theta$, he gets $\frac{\phi_{n,\ell}(a_2, \theta)}{2} - 1$ if player 2 chooses $(a_2, 0) \in A_2 \times \{0\}$; he gets $\phi_{n,\ell}(a_2, \theta)$ if player 2 chooses $(a_2, n) \in A_2 \times \{n\}$; and he gets $\phi_{n,\ell}(a_2, \theta) - 1$ if player 2 chooses $(a_2, m) \in A_2 \times \{m\}$ with $m \neq n$ and $m \neq 0$.

<table>
<thead>
<tr>
<th>$(N, \ell)$</th>
<th>$A_2 \times {0}$</th>
<th>$A_2 \times {1}$</th>
<th>$A_2 \times {2}$</th>
<th>\ldots</th>
<th>$A_2 \times {N}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\phi_{N,\ell}(a_2, \theta)$</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>\ldots</td>
<td>0</td>
</tr>
</tbody>
</table>

Player 2’s payoffs are specified as follows:

- If player 2 chooses $(a_2, m)$, he gets $g_2(a_2, a_1, \theta)$ if player 1 chooses $a_1 \in A_1$, and he gets $g_2(a_2, a_1, \theta)$ otherwise.
By the induction hypothesis, game $G$ satisfies property (ii) and has payoffs in the interval $[-5, 3]$; it follows that game $G$ also has payoffs in the interval $[-5, 3]$.

We now prove that game $G$, along with the actions $a^*_{1,n} \equiv (n, 0)$, $n = 1, \ldots, N$, satisfy properties (i)–(v) for $k + 1$. First, (i) follows because $a^1_1$ and $(a^2_1, n)$ are zero actions for players 1 and 2, respectively. Second, (ii) is satisfied by the coordination pair $(a^c_1, (a^c_2, 0))$: for any $\bar{a}_1 \in \bar{A}_1$ and $\theta \in \Theta$,

$$
\bar{g}_1 (\bar{a}_1, (a^c_2, 0), \theta) \geq \min \left\{ g_1 (a_1, a^c_2, \theta), \min_{n, \ell} \phi_{n, \ell} (a^c_2, \theta) / 2 - 1, 0 \right\} \geq -2;
$$

moreover, $\bar{g}_2 ((a_2, n), a^c_1, \theta) = g_2 (a_2, a^c_1, \theta) \geq -2$ for any $(a_2, n) \in \bar{A}_2$ and $\theta \in \Theta$, and $\bar{g}_1 (a^c_1, (a^c_2, 0), \theta) = \bar{g}_2 ((a^c_2, 0), a^c_1, \theta) = 3$ for any $\theta \in \Theta$.

The proof of (iii)–(v) relies on the following claim, whose proof is postponed.

**Claim 1.** For every integer $r \geq 0$ and $\gamma \in [0, 1/2)$,

1. $R^r_1 (\cdot, G, \gamma) \cap A_1 = R^r_1 (\cdot, G, \gamma)$;
2. $R^r_2 (\cdot, G, \gamma) = R^r_2 (\cdot, G, \gamma) \times \{0, 1, 2, \ldots, N\}$.

We now prove that Claim 1 implies properties (iii)–(v).

(iii): By the induction hypothesis, $R^r_2 (\cdot, G, \gamma) = R^r_2 (\cdot, G, \gamma)$. Thus, Claim 1 implies that $R^r_2 (\cdot, G, \gamma) = R^r_2 (\cdot, G, \gamma)$. This, in turn, implies $R^r_1 (\cdot, G, \gamma) = R^{r+1}_1 (\cdot, G, \gamma)$ and hence (iii).

(iv): Given any $1 \leq n \leq N$, consider the conjecture $\sigma_2 : \Theta \times T_2 \to \Delta(\bar{A}_2)$ such that $\sigma_2 (\theta, t_{2,n'}) [a^*_{2,n'}, n] = 1$ for each $n' = 1, \ldots, N'$ and $\theta \in \Theta$. By part 1 of Claim 1 and the fact that $G$ satisfies property (iv) (by the induction hypothesis), $\sigma_2$ is a rationalizable conjecture in $\bar{G}$. Moreover, given such a conjecture, $t_{1,n}$ gets an expected payoff of 0 by playing $(n, \ell)$ for any $\ell$, and gets at most 0 by playing any action in $\bar{A}_1 \setminus \{(n, 0), \ldots, (n, L)\}$:

$$
\beta_{t_{1,n}, \sigma_2} (F_\ell) \cdot \left( 1 - \mu_1 (E_\ell | t_{1,n}) \right) + (1 - \beta_{t_{1,n}, \sigma_2} (F_\ell)) \cdot \mu_1 (E_\ell | t_{1,n})
= \mu_1 (E_\ell | t_{1,n}) \cdot \left( 1 - \mu_1 (E_\ell | t_{1,n}) \right) + (1 - \mu_1 (E_\ell | t_{1,n})) \cdot \mu_1 (E_\ell | t_{1,n}) = 0.
$$

In particular, $a^*_{1,n} = (n, 0)$ is a best reply for $t_{1,n}$.

(v): Fix $1 \leq n \leq N$ and consider any type $s_1$ with $d^k_{1,n} (t_{1,n}, s_1) > \varepsilon$. Then, there exists some $1 \leq \ell \leq L$ such that

$$
\mu_1 (E_\ell | t_{1,n}) > \mu_1 \left( \left( E_\ell | s_{1} \right)^{\varepsilon, k} \right) + \varepsilon.
$$

It follows that, given any $\xi_{2}$-rationalizable conjecture $\sigma_2 : \Theta \times T_2^* \to \Delta(\bar{A}_2)$, the difference in expected payoffs between actions $(n, \ell)$ and $(n, 0)$ for type $s_1$ is at least $\varepsilon / 2$. To prove this, we consider two cases separately: when player 2 chooses $m = n$; and when player 2 chooses $m \neq n$. 39
First, conditional on player 2 choosing \( n \), the expected payoff difference between actions \( (n, \ell) \) and \( (n, 0) \) for type \( s_1 \), given an arbitrary \( \frac{\epsilon}{2} \)-rationalizable conjecture \( \sigma_2 \), is

\[
\sum_{a_2, \theta} \phi_{n, \ell}(a_2, \theta) \beta_{s_1, \sigma_2}(a_2, \theta | n) = \beta_{s_1, \sigma_2}(F_\ell | n) \cdot \left( - (1 - \mu_1(E_\ell | l_1, n)) \right) + \left( 1 - \beta_{s_1, \sigma_2}(F_\ell | n) \right) \cdot \mu_1(E_\ell | l_1, n).
\]

But, for any \( (\theta, s_2) \in \Theta \times T_2^* \) and \( n' = 1, \ldots, N' \),

\[
\sigma_2(\theta, s_2)(a^*_{2,n'}, n) > 0 \Rightarrow (a^*_{2,n'}, n) \in R_2(s_2, \tilde{G}, \frac{\epsilon}{2}) \quad \text{(since } \sigma_2 \text{ is } \frac{\epsilon}{2} \text{-rationalizable)}
\]

\[
\Rightarrow a^*_{2,n'} \in R_2(s_2, G, \frac{\epsilon}{2}) \quad \text{(by Claim 1)}
\]

\[
\Rightarrow a^*_2(t_2, n', s_2) \leq \epsilon, \quad \text{(by the induction hypothesis)}
\]

and thus, if \( \beta_{s_1, \sigma_2}(n) > 0 \),

\[
\beta_{s_1, \sigma_2}(F_\ell | n) \leq \mu_1 \left( (E_\ell)^{\epsilon, k} | s_1 \right),
\]

which implies

\[
\sum_{a_2, \theta} \phi_{n, \ell}(a_2, \theta) \beta_{s_1, \sigma_2}(a_2, \theta | n) \geq \mu_1 \left( (E_\ell)^{\epsilon, k} | s_1 \right) \cdot \left( - (1 - \mu_1(E_\ell | l_1, n)) \right) + \left( 1 - \mu_1 \left( (E_\ell)^{\epsilon, k} | s_1 \right) \right) \cdot \mu_1(E_\ell | l_1, n)
\]

\[
= \mu_1(E_\ell | l_1, n) - \mu_1 \left( (E_\ell)^{\epsilon, k} | s_1 \right) > \epsilon.
\]

Second, conditional on player 2 choosing \( m \neq n \), the expected payoff difference between actions \( (n, \ell) \) and \( (n, 0) \) for type \( s_1 \) (given any conjecture) is at least \( \phi_{n, \ell}(a^*_2, \theta)/2 = \mu_1(E_\ell | l_1, n)/2 \geq \epsilon/2 \). (This is because \( a^*_2 \neq a^*_{2,n'} \) for all \( n' \), and hence \( a^*_2(\theta) \neq F_\ell \) for every \( \theta \).) We have thus shown that, given any \( \epsilon/2 \)-rationalizable conjecture, and conditional on any choice of \( m = 0, \ldots, N \) by player 2 with \( \beta_{s_1, \sigma_2}(m) > 0 \), type \( s_1 \) gains at least \( \epsilon/2 \) by deviating from \( (n, 0) \) to \( (n, \ell) \). Thus, he also gains \( \epsilon/2 \) unconditionally on \( m \), and hence property (v) follows.

To conclude the proof, it remains to prove Claim 1. We prove it by induction on \( r \geq 0 \). First, the claim is trivially true for \( r = 0 \). We now consider \( r \geq 1 \), assume that the claim holds for any \( 0 \leq r' < r \), and prove that it also holds for \( r \).

\[
R^*_1(t_1, G, \gamma) \supset R^*_1(t_1, \tilde{G}, \gamma) \cap A_1: \text{Let } a_1 \in R^*_1(t_1, \tilde{G}, \gamma) \cap A_1. \text{ Then, there is an } (r - 1)\text{-order } \gamma\text{-rationalizable conjecture } \tilde{\sigma}_2 : \Theta \times T_2^* \rightarrow \Delta(A_2) \text{ in } \tilde{G} \text{ such that for any } a'_1 \in A_1,
\]

\[
\int_{\Theta \times T_2^*} d\mu_1(\theta, t_2 | t_1) \sum_{(a_2, n) \in A_2} \left[ \tilde{g}_1(a_1, (a_2, n), \theta) - \tilde{g}_1(a'_1, (a_2, n), \theta) \right] \tilde{\sigma}_2(\theta, t_2) [(a_2, n)] \geq -\gamma.
\]

(4)
Consider the mapping \( \varphi_2 : \tilde{A}_2 \to A_2 \),

\[
\varphi_2 (a_2, n) = \begin{cases} 
  a_2, & \text{if } n = 0; \\
  a'_2, & \text{if } n \neq 0.
\end{cases}
\]

Define \( \sigma_2 \) as the conjecture in \( G \) such that \( \sigma_2 (\theta, t_2) [a_2] = \tilde{\sigma}_2 (\theta, t_2) \left[ \varphi_2^{-1} (a_2) \right] \) for each \((\theta, t_2) \in \Theta \times T^*_2 \) and \( a_2 \in A_2 \). Since \( \tilde{\sigma}_2 \) is \((r - 1)\)-order \( \gamma \)-rationalizable in \( \tilde{G} \), \( \tilde{\sigma}_2 (\theta, t_2) [\{(a_2, n) \in \tilde{A}_2 : n > 0\}] \) implies \((a_2, 0) \in R^{-1}_2 (t_2, \tilde{G}, \gamma) \), and by the induction hypothesis, \( a_2 \in R^{-1}_2 (t_2, G, \gamma) \). Moreover, \( a'_2 \) is part of a coordination pair in \( G \), hence \( a'_2 \) is rationalizable in \( G \) for any type. Thus, \( \sigma_2 \) is an \((r - 1)\)-order \( \gamma \)-rationalizable conjecture in \( G \). Moreover, for each \( a'_1 \in A_1 \),

\[
\int_{\Theta \times T^*_2} d\mu_1 (\theta, t_2 | t_1) \sum_{(a_2, n) \in \tilde{A}_2} \left[ \tilde{g}_1 (a_1, (a_2, n), \theta) - g_1 (a'_1, (a_2, n), \theta) \right] \tilde{\sigma}_2 (\theta, t_2) [(a_2, n)]
\]

\[
= \int_{\Theta \times T^*_2} d\mu_1 (\theta, t_2 | t_1) \left( \sum_{a_2 \in A_2} \left[ g_1 (a_1, a_2, \theta) - g_1 (a'_1, a_2, \theta) \right] \tilde{\sigma}_2 (\theta, t_2) [(a_2, 0)] + \left[ (g_1 (a_1, a'_2, \theta) - 3) - (g_1 (a'_1, a'_2, \theta) - 3) \right] \tilde{\sigma}_2 (\theta, t_2) \left[ \{(a_2, n) \in \tilde{A}_2 : n > 0\} \right] \right)
\]

\[
= \int_{\Theta \times T^*_2} d\mu_1 (\theta, t_2 | t_1) \sum_{a_2 \in A_2} \left[ g_1 (a_1, a_2, \theta) - g_1 (a'_1, a_2, \theta) \right] \sigma_2 (\theta, t_2) [a_2] \tag{5}
\]

Then, (4) and (5) imply \( a_1 \in R^*_1 (t_1, G, \gamma) \).

\( R^*_1 (t_1, G, \gamma) \subset R^*_1 (t_1, \tilde{G}, \gamma) \cap A_1 \): Let \( a_1 \in R^*_1 (t_1, G, \gamma) \). Then, there exists an \((r - 1)\)-order \( \gamma \)-rationalizable conjecture \( \sigma_2 : \Theta \times T^*_2 \to \Delta (A_2) \) in \( G \) such that for each \( a'_1 \in A_1 \),

\[
\int_{\Theta \times T^*_2} d\mu_1 (\theta, t_2 | t_1) \sum_{a_2 \in A_2} \left[ g_1 (a_1, a_2, \theta) - g_1 (a'_1, a_2, \theta) \right] \sigma_2 (\theta, t_2) [a_2] \geq -\gamma. \tag{6}
\]

Define \( \tilde{\sigma}_2 \) as the conjecture in \( \tilde{G} \) such that \( \tilde{\sigma}_2 (\theta, t_2) [(a_2, 0)] = \sigma_2 (\theta, t_2) [a_2] \) for each \((\theta, t_2) \in \Theta \times T^*_2 \) and \( a_2 \in A_2 \) (and thus \( \tilde{\sigma}_2 (\theta, t_2) [(a_2, n)] = 0 \) for each \( n > 0 \)). Since \( \sigma_2 \) is \((r - 1)\)-order \( \gamma \)-rationalizable in \( G \), \( \sigma_2 (\theta, t_2) [a_2] > 0 \) implies \( a_2 \in R^{-1}_2 (t_2, G, \gamma) \), and by the induction hypothesis, \( a_2 \in R^{-1}_2 (t_2, G, \gamma) \). Hence, \( \tilde{\sigma}_2 \) is \((r - 1)\)-order \( \gamma \)-rationalizable in \( G \). We will now show that \( a_1 \) is a \( \gamma \)-best reply to \( \tilde{\sigma}_2 \) for \( t_1 \) in \( \tilde{G} \). First, by (6) and the definition of \( \tilde{\sigma}_2 \),

\[
\int_{\Theta \times T^*_2} d\mu_1 (\theta, t_2 | t_1) \sum_{a_2 \in A_2} \left[ g_1 (a_1, a_2, \theta) - g_1 (a'_1, a_2, \theta) \right] \tilde{\sigma}_2 (\theta, t_2) [(a_2, 0)] \geq -\gamma \quad \forall a'_1 \in A_1. \tag{7}
\]

Second, setting \( a'_1 = a'^*_1 \) in (6) and recalling that \( \gamma < 1/2 \),

\[
\int_{\Theta \times T^*_2} d\mu_1 (\theta, t_2 | t_1) \sum_{a_2 \in A_2} g_1 (a_1, a_2, \theta) \sigma_2 (\theta, t_2) [a_2] \geq -\gamma > -1/2.
\]

Since \( \phi_n, \ell \left( a'^*_2, \gamma \right) \) \( 2 - 1 \leq -1/2 \) for all \( n \) and \( \ell \),

\[
\int_{\Theta \times T^*_2} d\mu_1 (\theta, t_2 | t_1) \sum_{a_2 \in A_2} \left[ \tilde{g}_1 (a_1, a_2, \theta) - \tilde{g}_1 ((n, \ell), a_2, \theta) \right] \tilde{\sigma}_2 (\theta, t_2) [(a_2, 0)] \geq 0 \quad \forall n, \ell. \tag{8}
\]
By (7) and (8), \( a_1 \) is a \( \gamma \)-best reply to \( \tilde{\sigma}_2 \) for \( t_1 \) in \( \tilde{G} \), and hence \( a_1 \in R^*_1 (t_1, \tilde{G}, \gamma) \).

\[
R^*_2 (t_2, \tilde{G}, \gamma) \subset R^*_2 (t_2, G, \gamma) \times \{0, 1, \ldots, N\}: \text{ Suppose } (a_2, m) \in R^*_2 (t_2, \tilde{G}, \gamma). \text{ Then, there is an } (r - 1) \text{-order } \gamma \text{-rationalizable conjecture } \tilde{\sigma}_1 : \Theta \times T^*_1 \to \Delta (\tilde{A}_1) \text{ in } \tilde{G} \text{ such that for each } (a_2', m') \in \tilde{A}_2,
\]

\[
\int_{\Theta \times T^*_1} d\mu_2 (\theta, t_1 | t_2) \sum_{\tilde{a}_1 \in \tilde{A}_1} \left[ \tilde{g}_2 ((a_2, m), \tilde{a}_1, \theta) - \tilde{g}_2 ((a_2', m'), \tilde{a}_1, \theta) \right] \tilde{\sigma}_1 (\theta, t_1) [\tilde{a}_1] \geq -\gamma.
\]

Consider the map \( \varphi_1 : \tilde{A}_1 \to A_1 \) such that

\[
\varphi_1 (\tilde{a}_1) = \begin{cases} 
\tilde{a}_1 & \text{if } \tilde{a}_1 \in A_1, \\
a^*_i & \text{if } \tilde{a}_1 \notin A_1.
\end{cases}
\]

Let \( \sigma_1 \) be the conjecture in \( G \) such that \( \sigma_1 (\theta, t_1) [a_1] = \tilde{\sigma}_1 (\theta, t_1) [\varphi_1^{-1} (a_1)] \) for each \((\theta, t_1) \in \Theta \times T^*_1 \) and \( a_1 \in A_1 \). For each \( \tilde{a}_1 \in \tilde{A}_1 \), since \( \tilde{\sigma}_1 \) is \((r - 1)\)-order \( \gamma \)-rationalizable in \( \tilde{G} \), \( \tilde{\sigma}_1 (\theta, t_1) [\tilde{a}_1] > 0 \) implies \( \tilde{a}_1 \in R^*_1 (t_1, \tilde{G}, \gamma) \), and by the induction hypothesis, \( \tilde{a}_1 \in R^*_1 (t_1, G, \gamma) \). Moreover, \( a^*_i \) is part of a coordination pair in \( G \), hence it is rationalizable for any type. Thus, \( \sigma_1 \) is an \((r - 1)\)-order \( \gamma \)-rationalizable conjecture in \( G \). Moreover, (4) implies that for each \( a'_2 \in A_2 \),

\[
-\gamma \leq \int_{\Theta \times T^*_1} d\mu_2 (\theta, t_1 | t_2) \sum_{\tilde{a}_1 \in \tilde{A}_1} \left[ \tilde{g}_2 ((a_2, m), \tilde{a}_1, \theta) - \tilde{g}_2 ((a'_2, m'), \tilde{a}_1, \theta) \right] \tilde{\sigma}_1 (\theta, t_1) [\tilde{a}_1]
\]

\[
= \int_{\Theta \times T^*_1} d\mu_2 (\theta, t_1 | t_2) \left( \sum_{a_1 \in A_1} \left[ g_2 (a_2, a_1, \theta) - g_2 (a'_2, a_1, \theta) \right] \tilde{\sigma}_1 (\theta, t_1) [a_1] \right)
\]

\[
+ \left[ g_2 (a_2, a^*_i, \theta) - g_2 (a'_2, a^*_i, \theta) \right] \tilde{\sigma}_1 (\theta, t_1) [\tilde{A}_1 \setminus A_1]
\]

\[
= \int_{\Theta \times T^*_1} d\mu_2 (\theta, t_1 | t_2) \sum_{a_1 \in A_1} \left[ g_2 (a_2, a_1, \theta) - g_2 (a'_2, a_1, \theta) \right] \sigma_1 (\theta, t_1) [a_1]
\]

Therefore, \( a_2 \in R^*_2 (t_2, G, \gamma) \).

\[
R^*_2 (t_2, \tilde{G}, \gamma) \supseteq R^*_2 (t_2, G, \gamma) \times \{0, 1, \ldots, N\}: \text{ Let } (a_2, m) \in R^*_2 (t_2, G, \gamma) \times \{0, 1, \ldots, N\}. \text{ Then, there is an } (r - 1) \text{-order } \gamma \text{-rationalizable conjecture } \sigma_1 : \Theta \times T^*_1 \to \Delta (A_1) \text{ in } G \text{ such that for each } a'_2 \in A_2,
\]

\[
\int_{\Theta \times T^*_1} d\mu_2 (\theta, t_1 | t_2) \sum_{a_1 \in A_1} \left[ g_2 (a_2, a_1, \theta) - g_2 (a'_2, a_1, \theta) \right] \sigma_1 (\theta, t_1) [a_1] \geq -\gamma. \quad (9)
\]

Let \( \tilde{\sigma}_1 \) be the conjecture in \( \tilde{G} \) such that \( \tilde{\sigma}_1 (\theta, t_1) [a_1] = \sigma_1 (\theta, t_1) [a_1] \) for each \((\theta, t_1) \in \Theta \times T^*_1 \) and \( a_1 \in A_1 \) (and thus \( \tilde{\sigma}_1 (\theta, t_1) [a_1] = 0 \) for any \( a_1 \notin A_1 \)). Since \( \sigma_1 \) is \((r - 1)\)-order \( \gamma \)-rationalizable in \( G \), \( \sigma_1 (\theta, t_1) [a_1] > 0 \) implies \( a_1 \in R^*_1 (t_1, G, \gamma) \), and by the induction hypothesis, \( a_1 \in R^*_1 (t_1, \tilde{G}, \gamma) \). Hence, \( \tilde{\sigma}_1 \) is \((r - 1)\)-order \( \gamma \)-rationalizable in \( \tilde{G} \). Since \( \tilde{g}_2 ((a_2, m), a_1, \theta) = \tilde{g}_2 ((a_2, m'), a_1, \theta) \) for each \( m \) and \( m' \), it follows from (9) that \((a_2, m) \in R^*_2 (t_2, \tilde{G}, \gamma) \).
**Corollary 3.** For every $\varepsilon > 0$, player $i$, positive integer $k$ and finite type $t_i \in T_i^*$, there exists a game $G$ and an action $a_i$ of player $i$ in $G$, such that:

(i) $R_j(\cdot, G, \gamma) = R_{ij}^k(\cdot, G, \gamma)$ for every $\gamma \in [0, M/10)$ and $j \in I$;

(ii) $a_i \in R_i(t_i, G)$;

(iii) $a_i \neq R_{ij}^k(s_i, G, M\varepsilon/10)$ for every $s_i \in T_i^*$ with $d_{ij}^k(s_i, t_i) > \varepsilon$.

**Proof.** Immediate implication of Lemma 1, upon rescaling the payoffs by a factor of $M/5$. ■

**Lemma 2.** For every $\varepsilon > 0$ there exists $\delta > 0$ such that for every $i \in I$ and $t_i, t'_i \in T_i^*$ with $d_{ij}^k(t_i, t'_i) > \varepsilon$ there is a game $G$ such that $R_i(t_i, G) \not\subset R_i(t'_i, G, \delta)$.

**Proof.** Fix an $\varepsilon > 0$, a player $i$, an integer $k \geq 1$ and types $t_i, t'_i \in T_i^*$ with $d_{ij}^k(t_i, t'_i) > \varepsilon$. Fix $0 < \delta < M\varepsilon/10$ and choose $\rho > 0$ small enough that

$$\frac{M(\varepsilon - \rho)}{10} - 4M\rho > \delta. \quad (10)$$

Since finite types are dense in the product topology, there is a finite type $t''_i$ such that $d_{ij}^k(t_i, t''_i) < \rho$. Then, $d_{ij}^k(t''_i, t'_i) > \varepsilon - \rho$. By Corollary 3, there is a game $G' = (A'_i, g'_i)_{i \in I}$ and some action $a'_i$ of player $i$ in $G'$, such that $a'_i \in R_i(t''_i, G', 0)$ and $a'_i \neq R_{ij}^k(t', G', M(\varepsilon - \rho)/10)$. By Proposition 2 in Chen, Di Tillio, Faingold, and Xiong (2010), $a'_i \in R_{ij}^k(t_i, G', 4M\rho)$. Then, it follows from (i) of Corollary 3 that $a'_i \in R_i(t_i, G', 4M\rho)$ and $a'_i \neq R_i(t'_i, G', M(\varepsilon - \rho)/10)$.

To conclude, consider the game $G = (A_j, g_j)_{j \in I}$, defined as follows:

$$A_i = A_i', \quad A_{-i} = A_{-i}' \times A_i',$$

$$g_i(a_i, (a_{-i}', a_{-i}'), \theta) = \begin{cases} g_i'(a_i, a_{-i}', \theta) + 4M\rho & : \, a_i = a_{-i}' \\ g_i'(a_i, a_{-i}', \theta) & : \, \text{otherwise}, \end{cases}$$

and

$$g_{-i}(a_{-i}', a_{-i}'), \theta) = g_{-i}'(a_{-i}', a_{-i}, \theta).$$

In game $G$, player $-i$ is indifferent among all actions $a_{-i}^i$; moreover, player $i$ gets an additional payoff of $4M\rho$ whenever his action matches player $-i$’s choice of $a_{-i}^i$. Therefore, $R_i(\cdot, G, \gamma) = R_i(\cdot, G', \gamma + 4M\rho)$ for every $\gamma \geq 0$. In particular, we have that $a_i^j \in R_i(t_i, G, 0)$ and $a_i^j \notin R_i(t'_i, G, M(\varepsilon - \rho)/10 - 4M\rho) \supseteq R_i(t'_i, G, \delta)$, where the inclusion follows from (10).\(^{46}\) ■

The only if direction of Theorem 1 then follows directly from Lemma 2.

\(^{46}\)If necessary, rescale the payoffs from $G$ to ensure $|g_j| \leq M$ for every $j \in I$, and rescale $\delta$ by the same factor.
A.3 Strategic frames

For each player $i$ and $\emptyset \neq A'_i \subseteq A_i$, let

$$[A'_i] = \{t_i : R_i(t_i, G) = A'_i\}.$$ 

Fix any two types $t_i, t'_i \in T^*_i$ and suppose that $\mu^*_i(\theta \times [A'_i] | t_i) = \mu^*_i(\theta \times [A'_i] | t'_i)$ for each $\theta \in \Theta$ and $\emptyset \neq A'_i \subseteq A_i$. We must show that $R_i(t_i, G) = R_i(t'_i, G)$. If $a_i \in R_i(t_i, G)$ then, by the characterization of rationalizability in terms of best-reply sets (Appendix A.1), there is a conjecture $v \in \Delta(\Theta \times T^*_i \times A_{-i})$ such that $\mu^*_i(\theta \times (E \setminus [A'_i]) | t_i) = 0$. Since $v$ is well defined, we have

$$v(\theta \times (E \setminus [A'_i]) | t_i) = 0.$$

By construction, we have $\mu^*_i(\theta \times (E \setminus [A'_i]) | t'_i) > 0$. Thus, we must show that $\mu^*_i(\theta \times [A'_i] | t_i) = \mu^*_i(\theta \times [A'_i] | t'_i)$, which follows by interchanging the roles of $t_i$ and $t'_i$ in the argument above. This proves that the profile of partitions is indeed a frame.

A.4 Proof of Theorem 2

We begin with the following auxiliary result about the structure of $\varepsilon$-rationalizability.

**Lemma 3.** Fix a game $G = (A_i, g_i)_{i \in I}$ and $\varepsilon \geq 0$. For every integer $k \geq 1$, $i \in I$, and $t_i \in T^*_i$, we have $a_i \in R^k_i(t_i, G, \varepsilon)$ if and only if, for every $\alpha_i \in \Delta(A_i)$,

$$\sum_{\theta \in \Theta, B \subseteq A_{-i}} \max_{a_{-i} \in B} \left[ g_i(a_i, a_{-i}, \theta) - g_i(\alpha_i, a_{-i}, \theta) \right] \times \mu_i(\theta \times \{t_i : R^k_{-i}(t_{-i}, G, \varepsilon) = B\} | t_i) \geq -\varepsilon. \quad (12)$$
Likewise, \( a_i \in R_i(t_i, G, \varepsilon) \) if and only if, for every \( \alpha_i \in \Delta(A_i) \),

\[
\sum_{\theta \in \Theta, B \subseteq A_{-i}} \max_{a_{-i} \in B} \left[ g_i(a_i, a_{-i}, \theta) - g_i(\alpha_i, a_{-i}, \theta) \right] \\
\times \mu_i \left( \theta \times \left\{ t_{-i} : R_{-i}(t_{-i}, G, \varepsilon) = B \right\} | t_i \right) \geq -\varepsilon. \tag{13}
\]

**Proof.** The first part is equivalent to the iterative definition of interim correlated rationalizability in terms of dominance, given by condition (3) in Appendix A.1. The second part follows directly from the first. \( \blacksquare \)

The if direction of the theorem is an immediate consequence of the following lemma:

**Lemma 4.** Fix \( \delta > 0 \) and a game \( G = (A_i, g_i)_{i \in I} \) and let \( \mathcal{P} \) denote the strategic frame associated with \( G \). For every integer \( k \geq 0 \), \( i \in I \) and \( t_i, t_i' \in T_{i^*} \),

\[
d_{i, \mathcal{P}}^k(t_i, t_i') < \delta \implies R_i(t_i, G) \subseteq R_i^k(t_i', G, 4M\delta).
\]

In particular, for every \( i \in I \) and \( t_i, t_i' \in T_{i^*} \),

\[
d_{i, \mathcal{P}}^w(t_i, t_i') < \delta \implies R_i(t_i, G) \subseteq R_i(t_i', G, 4M\delta).
\]

**Proof.** We need only prove the first result, as the second result is a straightforward implication of the first one. For \( k = 0 \) the result is trivially true, as \( R_i^0 \equiv A_i \). Proceeding by induction, we assume the result is true for \( k \geq 0 \) and show that it remains true for \( k + 1 \). Consider two types \( t_i, t_i' \) with \( d_{i, \mathcal{P}}^{k+1}(t_i, t_i') < \delta \). Fix an arbitrary \( a_i \in R_i(t_i, G) \) and let us show that \( a_i \in R_i^{k+1}(t_i', G, 4M\delta) \). Given \( \alpha_i \in \Delta(A_i) \), by Lemma 3 we have

\[
\sum_{\theta \in \Theta, B \subseteq A_{-i}} \Delta g_i(\theta, B) \mu_i(\theta \times [B] | t_i) \geq 0, \tag{14}
\]

where, for each \( \theta \in \Theta \) and nonempty \( B \subseteq A_{-i} \),

\[
[B] = \left\{ t_{-i} : R_{-i}(t_{-i}, G) = B \right\}, \\
\Delta g_i(\theta, B) = \max_{a_{-i} \in B} \left[ g_i(a_i, a_{-i}, \theta) - g_i(\alpha_i, a_{-i}, \theta) \right].
\]

By Lemma 3, in order to prove that \( a_i \in R_i^{k+1}(t_i', G, 4M\delta) \) we need only show that

\[
\sum_{\theta \in \Theta, B \subseteq A_{-i}} \Delta g_i(\theta, B) \mu_i(\theta \times [[B]] | t_i') \geq -4M\delta,
\]

where

\[
[[B]] = \left\{ t_{-i} : R_{-i}^k(t_{-i}, G, 4M\delta) = B \right\} \quad \forall B \subseteq A_{-i}.
\]
To prove this, first note that the induction hypothesis implies
\[ [B]_{C,\delta,k}^{\delta,k} \subseteq \bigcup_{C \supseteq B} [C] \quad \forall B \subseteq A_{-i}. \tag{15} \]

Second, enumerate the elements of the finite set \( \Theta \times \{ B : \emptyset \neq B \subseteq A_{-i} \} \) as \( \{ (\theta_n, B_n) \}_{n=1}^{N} \) \((N = |\Theta|(|2^{A_{-i}}| - 1))\) so that
\[ \Delta g_i(\theta_n, B_n) \geq \Delta g_i(\theta_{n+1}, B_{n+1}) \quad \forall n = 1, \ldots, N - 1 \tag{16} \]
and
\[ (\theta_m = \theta_n \text{ and } B_m \supseteq B_n) \implies m \leq n. \tag{17} \]

Then,
\[
\sum_{\theta \in \Theta, B \subseteq A_{-i}} \Delta g_i(\theta, B) \mu_i(\theta \times [B]|_{t_i'}) \geq \sum_{n=1}^{N} \Delta g_i(\theta_n, B_n) \left( \mu_i(\theta_n \times [B_n]|_{t_i'}) - \mu_i(\theta_n \times [B_n]|_{t_i}) \right) \tag{18} \\
= \sum_{n=1}^{N-1} \left( \Delta g_i(\theta_n, B_n) - \Delta g_i(\theta_{n+1}, B_{n+1}) \right) \times \sum_{m=1}^{n} \left( \mu_i(\theta_m \times [B_m]|_{t_i'}) - \mu_i(\theta_m \times [B_m]|_{t_i}) \right) \tag{19} \\
= \sum_{n=1}^{N-1} \left( \Delta g_i(\theta_n, B_n) - \Delta g_i(\theta_{n+1}, B_{n+1}) \right) \times \left( \mu_i\left( \bigcup_{m=1}^{n} \theta_m \times [B_m]|_{t_i'} \right) - \mu_i\left( \bigcup_{m=1}^{n} \theta_m \times [B_m]|_{t_i} \right) \right) \tag{20} \\
\geq \sum_{n=1}^{N-1} \left( \Delta g_i(\theta_n, B_n) - \Delta g_i(\theta_{n+1}, B_{n+1}) \right) \times \left( \mu_i\left( \bigcup_{m=1}^{n} \theta_m \times [B_m]|_{t_i'} \right) - \mu_i\left( \bigcup_{m=1}^{n} \theta_m \times [B_m]|_{t_i} \right) \right) \tag{21} \\
\geq \sum_{n=1}^{N-1} \left( \Delta g_i(\theta_n, B_n) - \Delta g_i(\theta_{n+1}, B_{n+1}) \right) (\delta - \delta) \tag{22} \\
= -\left( \Delta g_i(\theta_1, B_1) - \Delta g_i(\theta_N, B_N) \right) \delta \\
\geq -4M \delta. \tag{23} \]

\[ \text{Such enumeration is possible because } \Delta g_i(\theta, B') \geq \Delta g_i(\theta, B) \text{ whenever } B' \supseteq B. \]
where (18) follows from (14); (19) follows by a standard “integration by parts” argument; (20) follows from the fact that \( \{ (\theta_m, [B_m]) \}_{m=1}^N \) and \( \{(\theta_m, [[B_m]]) \}_{m=1}^N \) are partitions of \( \Theta \times T^- \); (21) follows from (15), (16) and (17); (22) follows from the assumption that \( d_{i,j}^{k+1}(t_i, t'_j) \leq \delta \), the fact that \( [B_m] \in \mathcal{P}_m \) and (16); and (23) follows from \( |\Delta g_i| \leq 2M \).

Before turning to the proof of the only if direction of Theorem 2, we introduce some notation. First, given a frame \( \mathcal{P} \) and a type \( t_i \), let \( \mu^*_i(t_i) \) denote the belief of type \( t_i \) over the events in \( \Theta \times \mathcal{P}_i \). We view \( \mu^*_i(t_i) \) as an element of the finite-dimensional simplex \( \Delta(\Theta \times \mathcal{P}_i) \) (viewed as a subset of the Euclidean space \( \mathbb{R}^{|\Theta| \cdot |\mathcal{P}_i|} \)). Second, let \( \varphi_i : \Delta(\Theta \times \mathcal{P}_i) \rightarrow \mathcal{P}_i \) designate the function that maps each \( q \in \Delta(\Theta \times \mathcal{P}_i) \) into \( \mathcal{P}_i(t_i) \), where \( t_i \) is some type with \( \mu^*_i(t_i) = q \). Since \( \mathcal{P}_i \) is a frame, the definition of \( \varphi_i \) is independent of the choice of \( t_i \).

Third, given a game \( G = (A_i, g_i)_{i \in I} \), for each \( \beta_i \in \Delta(A_i \times \Theta) \) define a probability distribution \( \beta_i^\mathcal{P} \in \Delta(\Theta \times \mathcal{P}_i) \) as follows:

\[
\beta_i^\mathcal{P}(\theta, E) = \sum_{a_i \in \varphi_i^{-1}(E)} \beta_i(a_i, \theta) \quad \forall (\theta, E) \in \Theta \times \mathcal{P}_i.
\]

The proof of the only if direction of Theorem 2 relies on Lemmas 5 and 6 and Corollary 4 below.

**Lemma 5.** Fix a frame \( \mathcal{P} \). For each \( \zeta > 0 \) there exist \( \epsilon > 0 \) and a game \( G = (A_i, g_i)_{i \in I} \) such that for each \( i \in I \),

(i) \( A_i \) is a finite subset of \( \Delta(A_i \times \mathcal{P}_i) \) that is \( \sqrt{\epsilon} \)-dense (relative to the Euclidean norm \( \| \cdot \| \)) in every element of the partition \( \{ \varphi_i^{-1}(E) : E \in \mathcal{P}_i \} \);

(ii) for each \( a_{i-}, a'_{i-} \in A_{i-} \),

\[
\varphi_{i-}(a_{i-}) = \varphi_{i-}(a'_{i-}) \implies g_i(a_{i-}, a_{i-}, \theta) = g_i(a_{i-}, a'_{i-}, \theta) \quad \forall \theta \in \Theta, a_{i-} \in A_i;
\]

(iii) for each \( \beta_{i-} \in \Delta(A_{i-} \times \Theta) \),

\[
BR_i(\beta_{i-}, G, \epsilon) \supseteq \{ a_{i-} : a_{i-} \in A_i, \| a_i - \beta_{i-}^\mathcal{P} \| < \sqrt{\epsilon} \};
\]

(iv) for each \( \beta_{i-}, \beta'_{i-} \in \Delta(A_{i-} \times \Theta) \) with \( \| \beta_{i-}^\mathcal{P} - \beta'_{i-}^\mathcal{P} \| > \zeta \),

\[
BR_i(\beta'_{i-}, G, 2\epsilon) \cap \{ a_{i-} : a_{i-} \in A_i \} \ni a_{i-} \in A_i, \| a_i - \beta'_{i-}^\mathcal{P} \| < \sqrt{\epsilon} \} = \emptyset.
\]

**Proof.** Fix \( \zeta > 0 \) and a frame \( \mathcal{P} \). Let \( 0 < \epsilon < \zeta^2/(1 + \sqrt{3})^2 \). Cover the finite-dimensional simplex \( \Delta(\Theta \times \mathcal{P}_i) \) by a finite union of open balls \( B_1, \ldots, B_N \) of diameter \( \sqrt{\epsilon} \). Select one point from \( B_n \cap \varphi_i^{-1}(E) \), for each \( n = 1, \ldots, N \) and \( E \in \mathcal{P}_i \), and let \( A_i \) denote the set of selected points. By construction, \( A_i \) satisfies (i).
Consider the quadratic score \( s_i : \Delta(\Theta \times P_{-i}) \times \Theta \times P_{-i} \to \mathbb{R} \),
\[
s_i(q, \theta, E) = 2q(\theta, E) - \|q\|^2 \quad \forall (q, \theta, E) \in \Delta(\Theta \times P_{-i}) \times \Theta \times P_{-i},
\]
which can be readily shown to satisfy
\[
s_i(q, q) - s_i(q', q) = \|q' - q\|^2 \quad q, q' \in \Delta(\Theta \times P_{-i}). \tag{24}
\]
Then, define \( g_i : A_i \times A_{-i} \times \Theta \to \mathbb{R} \),
\[
g_i(a_i, a_{-i}, \theta) = s_i(a_i, \theta, \varphi_i^{-1}(a_{-i})) \quad \forall (a_i, a_{-i}, \theta) \in A_i \times A_{-i} \times \Theta,
\]
which clearly satisfies (ii).

Given \( \beta_{-i} \in \Delta(A_{-i} \times \Theta) \) and \( a_i \in A_i \) with \( \|a_i - \beta_{-i}^P\| < \sqrt{\varepsilon} \), for any \( a_i' \in A_i \) we have
\[
g_i(a_i, \beta_{-i}) - g_i(a_i', \beta_{-i})
= s_i(a_i, \beta_{-i}^P) - s_i(a_i', \beta_{-i}^P) \tag{by (ii)}
= s_i(a_i, \beta_{-i}^P) - s_i(\beta_{-i}^P, \beta_{-i}^P) + s_i(\beta_{-i}^P, \beta_{-i}^P) - s_i(a_i', \beta_{-i}^P)
= -\|a_i - \beta_{-i}^P\|^2 + \|a_i' - \beta_{-i}^P\|^2 \tag{by (24)}
> -\varepsilon, \tag{by \( \|a_i - \beta_{-i}^P\| < \sqrt{\varepsilon} \)}
\]
hence \( a_i \in BR_i(\beta_{-i}, G, \varepsilon) \), and this proves (iii).

Turning to (iv), let \( \beta_{-i}, \beta_{-i}' \in \Delta(A_{-i} \times \Theta) \) with \( \|\beta_{-i}' - \beta_{-i}^P\| > \zeta \) and \( a_i \in A_i \) with \( \|a_i - \beta_{-i}^P\| < \sqrt{\varepsilon} \). Then, \( \|a_i - \beta_{-i}^P\| > \zeta - \sqrt{\varepsilon} > \sqrt{3\varepsilon} \). By (i) we can find some \( a_i' \in A_i \) with \( \|a_i' - \beta_{-i}^P\| < \sqrt{\varepsilon} \). Thus,
\[
g_i(a_i, \beta_{-i}) - g_i(a_i', \beta_{-i}')
= s_i(a_i, \beta_{-i}'^P) - s_i(a_i', \beta_{-i}'^P) \tag{by (ii)}
= s_i(a_i, \beta_{-i}^P) - s_i(\beta_{-i}^P, \beta_{-i}'^P) + s_i(\beta_{-i}^P, \beta_{-i}'^P) - s_i(a_i', \beta_{-i}^P)
= -\|a_i - \beta_{-i}^P\|^2 + \|a_i' - \beta_{-i}^P\|^2 \tag{by (24)}
< -3\varepsilon + \varepsilon = -2\varepsilon, \tag{by \( \|a_i - \beta_{-i}^P\| > \sqrt{3\varepsilon} \) and \( \|a_i' - \beta_{-i}^P\| < \sqrt{\varepsilon} \)}
\]
and hence \( a_i \notin BR_i(\beta_{-i}', G, 2\varepsilon) \), as required.

\[\textbf{Lemma 6.} \text{ Fix } \delta > 0 \text{ and a frame } P. \text{ Let } \varepsilon > 0 \text{ and } G = (A_i, g_i)_{i \in I} \text{ satisfy the properties (i)--(iv) of Lemma 5 relative to } \zeta = \delta/(|\Theta| \cdot \max_i |P_i|). \text{ Then the following statements hold:}
\]

(a) for every \( i \in I \) and \( t_i \in T_i^* \),
\[
R(t_i, G, \varepsilon) \supseteq \left\{ a_i \in A_i \cap \varphi_i^{-1}(P_i(t_i)) : \|a_i - \mu_i^*(t_i)\|_{\Theta \times P_{-i}} < \sqrt{\varepsilon} \right\}.
\]
\]

48
(b) for every integer \( k \geq 0 \), \( i \in I \) and \( t_i, t'_i \in T_i^* \) with \( d_{i,\mathcal{P}}^k(t_i, t'_i) > \delta \),

\[
R_i^k(t'_i, G, 2\varepsilon) \cap \left\{ a_i \in A_i \cap \varphi_i^{-1}(\mathcal{P}_i(t_i)) : \| a_i - \mu_i^*(t_i) \|_{\Theta \times \mathcal{P}_i} < \sqrt{\varepsilon} \right\} = \emptyset.
\]

Proof. To prove (a), we show that the pair of correspondences \( \varsigma_i : T_i^* \Rightarrow A_i, \ i \in I \), defined by

\[
\varsigma_i(t_i) = \left\{ a_i \in A_i \cap \varphi_i^{-1}(\mathcal{P}_i(t_i)) : \| a_i - \mu_i^*(t_i) \|_{\Theta \times \mathcal{P}_i} < \sqrt{\varepsilon} \right\} \quad \forall t_i \in T_i^*.
\]

which is nonempty-valued by (i) of Lemma 5, has the \( \varepsilon \)-best-reply property. Indeed, given any \( a_i \in \varsigma_i(t_i) \) and any conjecture \( \sigma_{-i} : \Theta \times T_{-i}^* \to \Delta(A_{-i}) \) with \( \text{supp} \sigma_{-i}(\theta, t_{-i}) \subseteq \varsigma_i(t_{-i}) \) for all \((\theta, t_{-i}) \in \Theta \times T_{-i}^*\), we must have

\[
\beta_{t_i, \sigma_{-i}}\left(\theta \times (A_{-i} \cap \varphi_{-i}^{-1}(E))\right) = \mu_i^*(\theta \times E|t_i) \quad \forall \theta \in \Theta, E \in \mathcal{P}_{-i},
\]

and hence, \( \| a_i - \beta_{t_i, \sigma_{-i}}^P \| = \| a_i - \mu_i^*(t_i) \|_{\Theta \times \mathcal{P}_i} < \sqrt{\varepsilon} \), by (i) of Lemma 5. It then follows by (iii) of Lemma 5 that \( a_i \in BR_i(\beta_{t_i, \sigma_{-i}} G, \varepsilon) \). We have thus shown that the profile of correspondences \( \varsigma_i \) has the \( \varepsilon \)-best-reply property.

Turning to (b), since \( d_{i,\mathcal{P}}^0 \equiv 0 \) the result is true for \( k = 0 \) (vacuously). Proceeding by induction, assume the result is true for \( k \geq 0 \) and let us show that it remains true for \( k + 1 \). Fix \( i \in I \) and \( t_i, t'_i \in T_i^* \) with \( d_{i,\mathcal{P}}^{k+1}(t_i, t'_i) > \delta \). Then, there exist \( \theta \in \Theta \) and \( E \in \mathcal{P}_{-i} \) such that

\[
\mu_i^*(\theta \times E_{\mathcal{P}}^{\delta, k} | t'_i) < \mu_i^*(\theta \times E|t_i) - \delta / (|\Theta| \cdot |\mathcal{P}_{-i}|).
\]  

(25)

Consider an arbitrary \( k \)-order \( 2\varepsilon \)-rationalizable conjecture \( \sigma_{-i} : \Theta \times T_{-i}^* \to \Delta(A_{-i}) \). By the induction hypothesis, for every \( t_{-i}, t'_{-i} \in T_{-i}^* \), every \( \theta \in \Theta \) and every \( a_{-i} \in A_{-i} \) with \( \| a_{-i} - \mu_{-i}^*(t_{-i}) \|_{\Theta \times \mathcal{P}_{-i}} < \sqrt{\varepsilon} \), we can have \( \sigma_{-i}(\theta, t'_{-i})[a_{-i}] > 0 \) only if \( d_{i,\mathcal{P}}^k(t_i, t'_i) \leq \delta \). In particular, for every \( t_{-i}, t'_{-i} \in T_{-i}^* \), \( \theta \in \Theta \) and \( a_{-i} \in A_{-i} \),

\[
(\mu_{-i}^*(t_{-i})|_{\Theta \times \mathcal{P}_{-i}} = a_{-i} \quad \text{and} \quad d_{i,\mathcal{P}}^k(t_i, t'_i) > \delta) \quad \implies \quad \sigma_{-i}(\theta, t'_{-i})[a_{-i}] = 0.
\]

(26)

It follows that for every \( (\theta, t'_{-i}) \in \Theta \times T_{-i}^* \),

\[
t'_i \notin E_{\mathcal{P}}^{\delta, k} \quad \implies \quad \sigma_{-i}(\theta, t'_{-i})[A_{-i} \cap \varphi_{-i}^{-1}(E)] = 0.
\]

since for every \( a_{-i} \in A_{-i} \cap \varphi_{-i}^{-1}(E) \) and every \( t_{-i} \in T_{-i}^* \) with \( \mu_{-i}^*(t_{-i})|_{\Theta \times \mathcal{P}_{-i}} = a_{-i} \) we must have \( t_{-i} \in E \). Hence, for every \( k \)-order \( 2\varepsilon \)-rationalizable conjecture \( \sigma_{-i} \) we have

\[
\beta_{t'_i, \sigma_{-i}}\left(\theta \times (A_{-i} \cap \varphi_{-i}^{-1}(E))\right)
\]

\[
= \int_{T_{-i}^*} \sigma_{-i}(\theta, t'_{-i})[A_{-i} \cap \varphi_{-i}^{-1}(E)] \mu_i^*(\theta \times dt'_{-i} | t'_i)
\]

\[
= \int_{E_{\mathcal{P}}^{\delta, k}} \sigma_{-i}(\theta, t'_{-i})[A_{-i} \cap \varphi_{-i}^{-1}(E)] \mu_i^*(\theta \times dt'_{-i} | t'_i) \quad \text{(by (26))}
\]

\[
\leq \mu_i^*(\theta \times E_{\mathcal{P}}^{\delta, k} | t'_i)
\]

\[
< \mu_i^*(\theta \times E|t_i) - \delta / (|\Theta| \cdot |\mathcal{P}_{-i}|), \quad \text{(by (25))}
\]

49
and thus, \( \| \beta_{i_1, \sigma_{-1}}^{\theta} - \mu_i^*(t_i) \|_{\Theta \times \mathcal{P}_{-i}} > \zeta \). It follows from part (iv) of Lemma 5 that

\[
R_{i_1}^{k+1}(t_i', G, 2\varepsilon) \cap \{ a_i : \| a_i - \mu_i^*(t_i) \|_{\Theta \times \mathcal{P}_{-i}} < \sqrt{\varepsilon} \} = \emptyset.
\]
as was to be shown.

**Corollary 4.** For every frame \( \mathcal{P} \) and \( \delta > 0 \) there exists \( \varepsilon > 0 \) and a game \( G \) such that, for every \( i \in I \) and \( t_i, t_i' \in T_i^* \), if \( d_{i_1, \sigma}^{\mathcal{P}}(t_i, t_i') > \delta \) then \( R_i(t_i, G) \not\subseteq R_i(t_i', G, \varepsilon) \).

**Proof.** A straightforward implication of Lemma 6 is that there exist \( \varepsilon > 0 \) and a game \( G' \) such that, for every \( i \in I \) and \( t_i, t_i' \in T_i^* \), if \( d_{i_1, \sigma}^{\mathcal{P}}(t_i, t_i') > \delta \) then \( R_i(t_i, G', \varepsilon) \not\subseteq R_i(t_i', G', 2\varepsilon) \). To conclude, we can use a construction similar to the last part of the proof of Lemma 2 to obtain a game \( G \) with \( R_i(\cdot, G, \gamma) = R_i(\cdot, G', \gamma + \varepsilon) \) for all \( \gamma \geq 0 \).

The only if direction of Theorem 2 is an immediate implication of Corollary 4.

### A.5 Proof of Corollary 1

We begin with the following definition. Given a frame \( \mathcal{P} \) and a finite measurable partition \( \Pi_i \) of the finite-dimensional simplex \( \Delta(\Theta \times \mathcal{P}_{-i}) \), define the partition on \( T_i^* \) induced by \( \Pi_i \), written \( T_i^* / \Pi_i \), as follows: any two types of player \( i \) belong to the same element of \( T_i^* / \Pi_i \) if and only if their beliefs over \( \Theta \times \mathcal{P}_{-i} \) belong to the same element of \( \Pi_i \). The following lemma is straightforward from the definitions, and the proof is omitted:

**Lemma 7.** Each \( T_i^* / \Pi_i \) is a measurable partition of \( T_i^* \), and the join \( (\mathcal{P}_i \lor (T_i^* / \Pi_i))_{i \in I} \) is a frame.\(^{48}\)

We now need the following piece of notation. Given an integer \( m \geq 1 \), a measurable subset \( E \subseteq \Omega \) and \( p \in [0,1]^2 \), define the event that \( E \) is \( m \)-order \( p \)-belief recursively as follows:

\[
[B^p]^m(E) \overset{\text{def}}{=} B^p(E \cap [B^p]^{m-1}(E)),
\]

where \( [B^p]^0(E) \overset{\text{def}}{=} \Omega \). Then, the event that \( E \) is \( m \)-order \( p \)-belief for player \( i \), written \( [B^p]^m(E) \), is the projection of \( [B^p]^m(E) \) onto \( T_i^* \). We these definitions in place, we have:

\[
C^p(E) = \bigcap_{m \geq 1} [B^p]^m(E) \quad \text{and} \quad C_i^p(E) = \bigcap_{m \geq 1} [B_i^p]^m(E).
\]

**Lemma 8.** For every integer \( k \geq 1 \) and \( \delta > 0 \) there exists a \( k \)-order frame \( \mathcal{P} \) such that, for every \( i \in I \), every atom of \( \mathcal{P}_i \) has \( d^k_{i_1} \)-diameter at most \( \delta \).

---

\(^{48}\)Recall that the join of a pair of partitions, denoted by the symbol \( \lor \), is the coarsest partition that is finer than both partitions in the pair.
Proof. For \( k = 1 \) the result is trivial, as any profile of first-order measurable partitions is a frame. Proceeding by induction, consider \( k \geq 1 \), fix \( \delta > 0 \) and let \( \mathcal{P} \) be a \( k \)-order measurable frame whose atoms all have \( d_i^k \)-diameter less than \( \delta/2 \). Let \( \Pi_i \) be a finite partition of the simplex \( \Delta(\Theta \times \mathcal{P}_{-i}) \) (viewed as a subset of the Euclidean space \( \mathbb{R}^{|\Theta| \cdot |\mathcal{P}_{-i}|} \)) into finitely many Borel measurable subsets with Euclidean diameter less than \( \delta/\sqrt{|\Theta| \cdot |\mathcal{P}_{-i}|} \). By Lemma 7, the join \((\mathcal{P}_i \vee (\Pi_i/T_i^*))_{i \in I}\) is a \((k + 1)-\)order frame. We claim that every atom of \( \mathcal{P}_i \vee (\Pi_i/T_i^*) \) has \( d_i^{k+1} \)-diameter less than \( \delta \). Let \( t_i \) and \( t'_i \) be two types with \( t'_i \in (T_i^* / \Pi_i) (t_i) \), let \( E \) be a measurable subset of \( \Theta \times T_i^* \) and let us show that \( \mu_i^* (E_{\delta,k}|t'_i) \geq \mu_i^* (E|t_i) - \delta \). Since each atom of \( \mathcal{P}_{-i} \) has \( d_i^k \)-diameter less than \( \delta/2 \), there is some \( F \in 2^\Theta \otimes \mathcal{P}_{-i} \) with \( E \subseteq F \subseteq \mathcal{P}_{-i} \). Then, \( |\mu_i^* (F|t'_i) - \mu_i^* (F|t_i)| < \delta \), because each atom of \( \Pi_i \) has Euclidean diameter less than \( \delta/\sqrt{|\Theta| \cdot |\mathcal{P}_{-i}|} \) and we have \( t'_i \in (T_i^* / \Pi_i) (t_i) \). Thus,

\[
\mu_i^* (E_{\delta,k}|t'_i) \geq \mu_i^* (F|t'_i) \geq \mu_i^* (F|t_i) - \delta \quad \text{by } |\mu_i^* (F|t'_i) - \mu_i^* (F|t_i)| < \delta \quad \text{by } E \subseteq F \]

as claimed. \( \square \)

Lemma 9. Fix \( \delta > 0 \), an integer \( k \geq 1 \) and a \( k \)-order frame \( \mathcal{P} \) whose atoms all have \( d_i^k \)-diameter less than \( \delta \) for every \( i \in I \). Then, for every \( m = 0, \ldots, k \),

\[
d_i^m (t_i, t'_i) \leq d_{i,\mathcal{P}}^m (t_i, t'_i) + m\delta \quad \forall i \in I, \forall t_i, t'_i \in T_i^*.
\]

Proof. Fix \( \delta > 0 \), an integer \( k \geq 1 \) and a \( k \)-order frame \( \mathcal{P} \) whose atoms all have \( d_i^k \)-diameter less than \( \delta \), for every \( i \in I \). (Such a frame exists by Lemma 8.) For \( m = 0 \) the conclusion of the lemma is trivial, as \( d_i^0 = d_{i,\mathcal{P}}^0 = 0 \). Consider \( 1 \leq m \leq k \) and assume the conclusion of the lemma holds for \( m - 1 \). Let \( t_i, t'_i \in T_i^* \) and \( \eta > d_{i,\mathcal{P}}^m (t_i, t'_i) \) and let us show that \( d_i^m (t_i, t'_i) \leq \eta + m\delta \). Fix \( E \in 2^\Theta \otimes \mathcal{P}_{-i} \). Since all the atoms of \( \mathcal{P}_{-i} \) have \( d_{-i}^k \)-diameter less than \( \delta \), there is some \( F \in 2^\Theta \otimes \mathcal{P}_{-i} \) with \( E \subseteq F \subseteq \mathcal{P}_{-i} \). Then,

\[
\mu_i^* (E^{\eta + m\delta, m-1}|t'_i) \geq \mu_i^* (E^{\eta + (m-1)\delta, m-1}|t'_i) \geq \mu_i^* (F^\eta_{\mathcal{P}}|t'_i) \geq \mu_i^* (F|t_i) - \eta \geq \mu_i^* (E|t_i) - \eta - m\delta, \quad \text{by } \eta > d_{i,\mathcal{P}}^m (t_i, t'_i) \quad \text{by } E \subseteq F
\]

and hence, \( d_i^m (t_i, t'_i) \leq \eta + m\delta \). But since our choice of \( \eta > d_{i,\mathcal{P}}^m (t_i, t'_i) \) was arbitrary, we have shown that \( d_i^m (t_i, t'_i) \leq d_{i,\mathcal{P}}^m (t_i, t'_i) + m\delta \), as required. \( \square \)
Lemma 10. Fix $\delta > 0$, a frame $\mathcal{P}$ and a finite set $P \subset [0,1]^2$. Let $\mathcal{P}$ denote the common belief frame

$$
(\mathcal{P}_i \cup \bigvee_{p, E} \{C_i^p(E), T_i^* \setminus C_i^p(E)\})_{i \in I},
$$

where the join ranges over all $p \in P$ and $E \in 2^\Theta \otimes \mathcal{P}_1 \otimes \mathcal{P}_2$. Then, for every pair of integers $k \geq \ell \geq 1$, every $p \in P$ and every $E \in 2^\Theta \otimes \mathcal{P}_1 \otimes \mathcal{P}_2$,

$$(C_i^p(E))_{\mathcal{P}}^{\delta, \ell} \subseteq \left[ B_i^{p,-(\delta, \delta)} \right]^{\ell-k} \left( E_{\mathcal{P}}^{\delta, k} \right) \quad \forall i \in I.
$$

Proof. We fix $k \geq 1$, $E \in 2^\Theta \otimes \mathcal{P}_1 \otimes \mathcal{P}_2$ and $p \in P$ and prove the result by induction on $\ell \geq k$. First, the result is trivial for $\ell = k$, as $\left[ B_i^{p,-(\delta, \delta)} \right]^{0} (\cdot) = T_i^*$ for every $i$. Next, suppose the result is true for $\ell \geq k$ and let us show that it remains true for $\ell + 1$. Pick $t_i \in C_i^p(E)$ and $t'_i \in T_i^*$ with $d_{\mathcal{P}, \mathcal{P}}^{\ell+1}(t_i, t'_i) < \delta$ and let us show that

$$
t'_i \in B_i^{p,-\delta} \left( E_i^{\delta, k} \cap \left[ B_i^{p,-(\delta, \delta)} \right]^{\ell-k} \left( E_i^{\delta, k} \right) \right) = \left[ B_i^{p,-(\delta, \delta)} \right]^{\ell+1-k} \left( E_i^{\delta, k} \right). \quad (27)
$$

Indeed,

$$
\mu_i^* \left( \left( E_i^{\delta, k} \cap \left[ B_i^{p,-(\delta, \delta)} \right]^{\ell-k} \left( E_i^{\delta, k} \right) \right)_{t'_i} \right) \geq \mu_i^* \left( \left( E \cap C_i^p(E) \right)_{t'_i} \right) \geq \mu_i^* \left( \left( E \cap C_i^p(E) \right)_{t_i} \right) - \delta \geq p_i - \delta,
$$

where the first inequality follows from the induction hypothesis and the fact that $d_{i, \mathcal{P}}^k \leq d_{i, \mathcal{P}}^{\ell}$, the second from $d_{i, \mathcal{P}}^{\ell}(t_i, t'_i) < \delta$, the third from $d_{i, \mathcal{P}}^{\ell+1}(t_i, t'_i) < \delta$, and the last from $t_i \in C_i^p(E) = B_i^{p}(E \cap C_i^p(E))$. This proves $(27)$.

Lemma 11. For every $\delta > 0$ and integer $k \geq 1$ there exists a common belief frame $\mathcal{P}$ such that for every $t_i, t'_i \in T_i^*$, if $d_{i, \mathcal{P}}^{uw}(t_i, t'_i) < \delta$, then for every $p \in [0,1]^2$ and measurable $E \subseteq \Omega$,

$$
t_i \in C_i^p(E) \implies t'_i \in C_i^{p-(\delta(k+2),k)}(E^{\delta(k+2),k}).
$$

Thus, uniform-weak convergence in every common belief frame implies convergence in common beliefs.

Proof. Fix $\delta > 0$, an integer $k \geq 1$ and a finite subset $P \subset [0,1]^2$ with the property that for every $p \in [0,1]^2$ there exists $q \in P$ with $p \geq q \geq p - (\delta, \delta)$. Let $\mathcal{P}$ be a $k$-order frame whose atoms

---

49Recall that, for any measurable subset $E \subseteq \Omega$ and any type $t_i$ of player $i$, $E_{t_i}$ denotes the section of $E$ over $t_i$. See Section 2.4.
have $d_i^k$-diameter at most $\delta$, for every player $i$. Such a frame exists by Lemma 8. Consider the associated common belief frame $\mathcal{P}$ constructed as in Lemma 10:

$$\left(\mathcal{P}_i \vee \bigvee_{p \in \mathcal{P}, E \in \mathcal{P}} \{C_i^p(E), T_i^* \setminus C_i^p(E)\}\right)_{i \in I},$$

where the join ranges over all $p \in P$ and $E \in 2^\Theta \otimes \mathcal{P}_1 \otimes \mathcal{P}_2$.

Fix $i \in I$, a pair of types $t_i, t'_i \in T_i^*$ with $d_{i,\mathcal{P}}(t_i, t'_i) < \delta$, an event $E \subseteq \Omega$ and $p \in [0, 1]^2$ with $t_i \in C_i^p(E)$. Pick $q \in P$ with $q \geq p \geq p - (\delta, \delta)$ and $F \in 2^\Theta \otimes \mathcal{P}_1 \otimes \mathcal{P}_2$ such that $E \subseteq F \subseteq E^{\delta,k}$. (Such an event $F$ exists, as the atoms of $\mathcal{P}$ have $d_i^k$-diameter at most $\delta$.) Then,

$$t'_i \in \bigcap_{\ell \geq k} \left( C_i^p(F) \right)_{\mathcal{P}}^{\delta,\ell} \subseteq \bigcap_{m \geq 1} \left[ B_i^{q-\delta 1} \right]^m \left( F_{\mathcal{P}}^{\delta,k} \right) \quad \text{(by $d_{i,\mathcal{P}}(t_i, t'_i) < \delta, q \leq p$ and Lemma 10)}$$

$$= C_i^{q-\delta 1}(F_{\mathcal{P}}^{\delta,k}) \quad \text{(by the definition of common belief)}$$

$$\subseteq C_i^{q-\delta 1}(E^{(k+1)\delta,k}) \quad \text{(by Lemma 9)}$$

$$\subseteq C_i^{p-2\delta 1}(E^{(k+2)\delta,k}), \quad \text{(by $F \subseteq E^{\delta,k}$ and $q \geq p - (\delta, \delta)$)}$$

and hence $t'_i \in C_i^{p-(k+2)\delta 1}(E^{(k+2)\delta,k})$.

It follows from Lemma 11 that uniform-weak convergence relative to every frame implies convergence in common beliefs. Thus, by Theorem 2, strategic convergence implies convergence in common beliefs, as was to be shown.

### A.6 A sequence converging in common beliefs but not strategically

We construct an example of a sequence of types $t_i^n$ that converges in common beliefs to a type $t_1$, but does not converge uniform-weakly on a frame and hence, by Theorem 2, does not converge strategically. To construct the sequence, fix $\theta_0 \in \Theta$ and $0 < p < q < 1$. For each player $i$, pick a type $r_i$ that satisfies the following two conditions:

(i) $r_i$ assigns probability zero to state $\theta_0$;

(ii) for every product-closed proper subset $E \subset \Omega$ and $\eta > 0$, $r_i \notin C_i^{\eta 1}(E)$.\(^{50}\)

Let $s_i$ and $t_i$ be the types who assign probability one to $\theta_0$ and whose beliefs about the other player’s types are as specified in Figure 4 below.

\(^{50}\)Types that satisfy these conditions exist. Ely and Pęski (2011, Theorem 1) show that the types that satisfy property (ii) are precisely those types to which product convergence is equivalent to strategic convergence, called regular types. They show that the set of regular types is a residual subset of the universal type space (in the product topology), in particular a non-empty set. As for condition (i), note that any type of player $i$ that assigns probability one to some $\theta, u_{-i}$, where $u_{-i}$ is a type of player $-i$ that satisfies property (ii), must also satisfy property (ii). This implies the existence of types satisfying both (i) and (ii).
Since $r_1$ and $r_2$ satisfy property (ii) above, for every $p \in (0, 1]^2$, every product-closed proper subset $E \subseteq \Omega$, and every $i \in I$,\[ t_i \in C_i^p(E) \iff p \leq (p, p) \quad \text{and} \quad E \supseteq \theta_0 \times \{s_1, t_1\} \times \{s_2, t_2\}. \tag{28} \]

In particular, no nontrivial event is common $(q, p)$-belief at $t_1$.

The construction of the sequence mimics the structure in Figure 4. Fix $0 < \Delta \leq q - p$ and for each player $i$ define $s^n_i$ and $t^n_i$ as follows: let $t^n_1 = r_1$ and, for each $n \geq 1$, let $s^n_2$, $s^n_1$, $t^n_2$ and $t^{n+1}_1$ be the types who assign probability one to $\theta_0$ and whose beliefs about the other player’s types are as described in Figure 5 below.

The sequence $t^n_i$ converges to $t_1$ in common beliefs. To see why, first note that $t^n_1 \to t_1$ in the product topology: by the construction in Figures 4 and 5, $s^1_2$ has the same first-order belief as $s_2$, hence $s^1_1$ has the same second-order belief as $s_1$, which implies $t^n_1$ has the same third-order belief as $t_2$, and so forth. Second, given an arbitrary $\delta > 0$, by the construction in Figure 5 and the product-convergence $t^n_1 \to t_1$, for each integer $k \geq 1$ we have $t^n_i \in B^\Delta_{1,k}(\theta_0 \times \{t_1\}^\delta \times \{t_2\})$ for all $n$ large enough. Since $t_2 \in C_{2}^{(p,p)}(\theta_0 \times \{s_1, t_1\} \times \{s_2, t_2\})$ and $q - \Delta \geq p$, it follows that for all $n$ large enough,

$$t^n_1 \in B^p_1(\theta_0 \times \{t_1\}^\delta \times C_{2}^{(p,p)}(\theta_0 \times \{s_1, t_1\} \times \{s_2, t_2\})) \subseteq C_{1}^{(p,p)}((\theta_0 \times \{s_1, t_1\} \times \{s_2, t_2\})^\delta \times \{t_2\}).$$

It follows, by (28) and footnote 6, that $t^n_i \to t_1$ in common beliefs, as claimed.

It remains to exhibit a frame $\mathcal{P}$ such that $t^n_i \not\to t_1$ uniform-weakly on $\mathcal{P}$. For each $r = (r_1, \ldots, r_4) \in (0, 1]^4, i \in I$ and measurable subset $E \subseteq \Omega$, let $[M^r_i](E) = \Omega$, and for each $k \geq 1$, define recursively

$$[M^r_i]^k(E) = B^r_{i} \left( E \cap B^r_{i} \left( E \cap B^r_{i} \left( E \cap B^r_{i}(E \cap [M^r_i]^{k-1}(E)) \right) \right) \right).$$
Note that

\[
\left[ M^k_i \right] (E) \subseteq B_i^k \left( E \cap \left[ M^{(r_2,r_3,r_4,r_1)}_{i_1} \right] ^{k-1} \right) \subseteq \left[ M^k_i \right] ^{k-1} (E) \quad \forall k \geq 1. \tag{29}
\]

For each \( k \geq 1 \), let

\[
P^k_1 = \left[ M^{(q,p,p,p)}_1 \right] (\theta_0), \quad Q^k_1 = \left[ M^{(p,p,q,p)}_1 \right] (\theta_0),
\]
\[
P^k_2 = \left[ M^{(p,p,p,q)}_2 \right] (\theta_0), \quad Q^k_2 = \left[ M^{(p,q,p,p)}_2 \right] (\theta_0)
\]

and finally,

\[
P_i = \bigcap_{k=1}^{\infty} P^k_i \quad \text{and} \quad Q_i = \bigcap_{k=1}^{\infty} Q^k_i \quad \forall i \in I.
\]

Let \( \mathcal{P} \) be the profile of partitions \( (\mathcal{P}_i)_{i \in I} \),

\[
\mathcal{P}_i = \{ P_i \setminus Q_i, Q_i \setminus P_i, P_i \cap Q_i, T^*_i \setminus (P_i \cup Q_i) \} \quad \forall i \in I.
\]

To verify that \( \mathcal{P} \) is a frame, note that (29) implies

\[
P^k_1 \subseteq B^q_1 (\theta_0 \times P^k_2) \subseteq P^k_1 \quad \forall k \geq 1
\]

and hence

\[
P_1 = \bigcap_{k=2}^{\infty} P^k_1 \subseteq \bigcap_{k=2}^{\infty} B^q_1 (\theta_0 \times P^k_2) = B^q_1 (\theta_0 \times P_2) \subseteq \bigcap_{k=2}^{\infty} P^k_1 = P_1.
\]

Therefore, \( P_1 = B^q_1 (\theta_0 \times P_2) \). By analogous arguments,

\[
Q_1 = B^p_1 (\theta_0 \times Q_2), \quad P_2 = B^q_2 (\theta_0 \times Q_1), \quad \text{and} \quad Q_2 = B^p_2 (\theta_0 \times P_1).
\]

Thus, for all \( i \in I \) and \( t_i \in T^*_i \), the element of \( \mathcal{P}_i \) containing \( t_i \) is determined by the values \( \mu^*_i (\theta_0 \times P_{-i} | t_i) \) and \( \mu^*_i (\theta_0 \times Q_{-i} | t_i) \), and hence, a fortiori, by the restriction of \( \mu^*_i (\cdot | t_i) \) to \( \Theta \times \mathcal{P}_{-i} \). This shows that \( \mathcal{P} \) is a frame.

We now prove that \( t^n_1 \) does not converge to \( t_1 \) uniform weakly on \( \mathcal{P} \). Fix \( 0 < \delta < \min\{ p, \Delta \} \).

It is enough to show that for every positive integer \( n \),

\[
d^3_{1,\mathcal{P}} (t^n_1, t^n_1) > \delta, \quad d^2_{2,\mathcal{P}} (s_2, s_2^n) > \delta, \quad d^1_{1,\mathcal{P}} (s_1, s_1^n) > \delta, \quad d^1_{2,\mathcal{P}} (t_2, t_2^n) > \delta. \tag{30}
\]

To prove this, we show that the first inequality holds for \( n = 1 \), that for each \( n \geq 1 \), if the first inequality holds for \( n \), then all others also hold for \( n \), and finally that if the last inequality holds for \( n \), then the first holds for \( n + 1 \). In the proof we use the following facts, which are immediate from the definition of \( t_1, t_2, s_1, s_2 \) and the fact that \( \mu^*_i (\theta_0 | r_i) = 0 \) and \( P_i \cup Q_i \subseteq B^p_i (\theta_0) \) for each \( i \in I \):

\[
t_1 \in P_1, \quad s_1 \in Q_1, \quad s_2 \in Q_2, \quad t_2 \in P_2, \quad r_i \notin (P_i \cup Q_i)^{\delta,1}_\mathcal{P} \quad \forall i \in I. \tag{31}
\]
Since \( \mu_1^*(\theta_0|t_1) = 1 \) and \( \mu_1^*(\theta_0|t_1^1) = 0 \), we have \( d_{1,p}^1(t_1, t_1^1) = 1 > \delta \), which proves the first inequality in (30) for \( n = 1 \). Now fix any \( n \geq 1 \) and assume the first inequality in (30) holds for \( n \). Then, by (31), \( \{t_1^n, r_1\} \cap (P_1)^{\delta,4n-3} = \emptyset \). Since \( \mu_2^*(\theta_0 \times \{t_1^n, r_1\}|s_1^n) = 1 \) and, by (31), \( \mu_2^*(\theta_0 \times P_1|s_2) = \mu_2^*(\theta_0, t_1|s_2) = p \), it follows that \( d_{2,p}^{4n-2}(s_2, s_2^n) \geq p > \delta \). Thus, the second inequality in (30) holds for \( n \). The latter implies, by (31), \( \{s_2^n, r_2\} \cap (P_2)^{\delta,4n-2} = \emptyset \), and since \( \mu_1^*(\theta_0 \times \{s_2^n, r_2\}|s_1^n) = 1 \) and, by (31), \( \mu_1^*(\theta_0 \times Q_2|s_1) = \mu_1^*(\theta_0, s_2|s_1) = p \), we also have \( d_{1,p}^{4n-1}(s_1, s_1^n) \geq p > \delta \), that is, the third inequality in (30) holds for \( n \). This in turn implies, by (31), \( \{s_1^n, r_1\} \cap (P_1)^{\delta,4n-1} = \emptyset \), and since \( \mu_2^*(\theta_0 \times \{s_1^n, r_1\}|t_1^n) = 1 \) and, by (31), \( \mu_2^*(\theta_0 \times Q_1|t_2) = \mu_2^*(\theta_0, s_1|t_2) = p \), we obtain \( d_{2,p}^{4n}(t_2, t_2^n) \geq p > \delta \). This proves that the fourth inequality in (30) holds for \( n \), and hence, by (31), that \( \{t_2^n, r_2\} \cap (P_2)^{\delta,4n} = \emptyset \). It remains to show that the latter implies that the first inequality in (30) holds for \( n + 1 \). Indeed, since \( \mu_1^*(\theta_0 \times \{t_2^n, r_2\}|t_1^{n+1}) = 1 - q + \Delta \) and, by (31), \( \mu_1^*(\theta_0 \times P_2|t_1) = \mu_1^*(\theta_0, r_2|t_1) = q \), we have \( d_{1,p}^{4n+1}(t_1, t_1^{n+1}) \geq \Delta > \delta \), as claimed.

A.7 Proof of Corollary 2

An immediate implication of Ely and Pęski (2011, Theorem 1) is that every finite type is critical. This fact, together with the denseness of finite types in the strategic topology (Dekel, Fudenberg, and Morris, 2006), implies that the set of critical types is dense in the strategic topology.

Next, we show that the set of critical types is open in the strategic topology, or equivalently, that the set of regular types is closed. Suppose not. Then, there is a sequence of regular types \( t_i^n \) that converges to some critical type \( t_i \). By Ely and Pęski (2011, Theorem 1), there is some \( p > 0 \) and some product-closed, proper subset \( E \subset \Omega \) with \( t_i \in C_i^{p,1}(E) \). Then, there is an integer \( k \geq 1 \) and \( \delta \in (0, p) \) such that the \( d_{k}^{p} \)-closure of \( E^{\delta,k} \) is a proper subset of \( \Omega \). Moreover, by Corollary 1, \( t_i^n \rightarrow t_i \) in common beliefs, and hence \( t_i^n \in C_i^{(p-\delta),1}(E^{\delta,k}) \) for all \( n \) large enough. It follows, again by Ely and Pęski (2011, Theorem 1), that \( t_i^n \) is a critical type for all \( n \) large enough, and this is a contradiction. The contradiction shows that the set of regular types is closed.

A.8 Proof of Theorem 3

Since the implications \( (a) \iff (b) \iff (c) \iff (d) \iff (e) \) follow from Theorems 1, 2 and 1, it is enough to prove the implication \( (e) \implies (a) \). The latter is an immediate consequence of the following lemma.

Lemma 12. Let \((T_1, T_2)\) be a finite type space. There exist \( \eta > 0 \) and \( k \geq 1 \) such that for each \( \delta \in (0, \eta), \epsilon \in [0,1), m \geq 0, i \in I \) and \((t_i, t'_i) \in T_i \times T_i^*\) with \( d_{i}^{k}(t_i, t'_i) < \delta \), one has

\[
t'_i \in \left[ B_i^{(1-\epsilon,1-\epsilon)} \right]^m (\Theta \times T_1^{\delta,k} \times T_2^{\delta,k}) \quad \Rightarrow \quad d_{i}^{k+m}(t_i, t'_i) < \delta + \epsilon.
\]  

(32)
Proof. Choose \( k \geq 1 \) and \( \eta > 0 \) so that, for each \( i \in I \) and \( t_i, t_i' \in T_i \), if \( d^{k-1}_i(t_i, t_i') < 2\eta \) then \( t_i' = t_i \). Since \( T_1 \) and \( T_2 \) are finite, such \( k \) and \( \eta \) exist. Fix \( \delta \in (0, \eta) \). Thus, for each \( n \geq 0, i \in I \), \( t_i \in T_i \) and \( E \subseteq \Theta \times T_{-i} \),

\[
E_{\delta,k} \cap (\Theta \times T_{-i})_{\delta,k} \subseteq E_{\delta,k}.
\]

The proof of \((32)\) is by induction in \( m \). Obviously, \((32)\) holds for \( m = 0 \). Now assume that it holds for some \( m \geq 0 \) and fix \( i \in I \), \( t_i \in T_i \) and \( t_i' \in T_i^* \) with \( t_i' \in \bigcap B_i^{(1-\varepsilon,1-\varepsilon)} \) and \( d^k_i(t_i, t_i') < \delta \). Since \( \mu_i^*(\Theta \times T_{-i}|t_i) = 1 \), in order to prove that \( d^k_i+m(t_i, t_i') < \delta + \varepsilon \) it suffices to show that \( \mu_i^*(E^{\delta+s,k+m}|t_i') \geq \mu_i^*(E|t_i) - \delta - \varepsilon \) for each \( E \subseteq \Theta \times T_{-i} \). Indeed,

\[
\begin{align*}
\mu_i^*(E^{\delta+k+m}|t_i') &\geq \mu_i^*(E^{\delta,k} \cap \bigcap B_i^{(1-\varepsilon,1-\varepsilon)} \Theta \times T_{1}^{\delta,k} \times T_{2}^{\delta,k})|t_i') \\
&\geq \mu_i^*(E^{\delta,k-1} \cap (\Theta \times T_{-i})_{\delta,k} \cap \bigcap B_i^{(1-\varepsilon,1-\varepsilon)} \Theta \times T_{1}^{\delta,k} \times T_{2}^{\delta,k})|t_i') \\
&\geq \mu_i^*(E^{\delta,k-1}|t_i') - \varepsilon \geq \mu_i^*(E|t_i) - \delta - \varepsilon,
\end{align*}
\]

where the first inequality follows from the induction hypothesis, the second from \((33)\), the third from \( t_i' \in \bigcap B_i^{(1-\varepsilon,1-\varepsilon)} \) and the fourth from \( d^k_i(t_i, t_i') < \delta \).

A.9 Proofs of Theorems 4 and 5

The proofs of the theorems rely on the four lemmas below. Before presenting the results, we need some preliminaries. First, in the proofs we use the following notation: for each \( \delta > 0, i \in I \) and measurable \( E \subseteq \Theta \times T_i^* \), we let \( E^{\delta,\infty} \) denote the uniform-weak \( \delta \)-neighborhood of \( E \), that is,

\[
E^{\delta,\infty} = \bigcap_{k \geq 1} E_{\delta,k}.
\]

Second, we say a finite type space \( T = (T_i)_{i \in I} \) has full support if \( \mu_i^*(E|t_i) > 0 \) for every \( i \in I \), \( t_i \in T_i \) and \( E \subseteq \Theta \times T_{-i} \). Third, generalizing our notations \( h_{d^p} \) and \( h_{d^s} \), we define the metric \( d^{uw} \) on \( T^* \) by letting \( d^{uw}(t, t') \) be the minimal \( d_i^w(t_i, t_i') \) for every \( t = (t_i)_{i \in I} \) and \( t' = (t_i')_{i \in I} \) in \( T^* \), and given two type spaces \( T = (T_i)_{i \in I} \) and \( T' = (T'_i)_{i \in I} \) in \( T \) we write \( h_{d^{uw}}(T, T') \) for the Hausdorff distance between \( T \) and \( T' \) under \( d^{uw} \), that is,

\[
h_{d^{uw}}(T, T') = \max \left\{ \sup_{t \in T} \inf_{t' \in T'} d^{uw}(t, t'), \sup_{t' \in T'} \inf_{t \in T} d^{uw}(t, t') \right\}.
\]

Since for each \( i \in I \) the uniform-weak topology on \( T_i^* \) is finer than the product topology on \( T_i^* \), and hence product-closed sets in \( T_i^* \) are also uniform-weakly closed, \( h_{d^{uw}} \) is also a metric on \( T \).

Lemma 13. Let \( T = (T_i)_{i \in I} \) be a finite type space with full support. Let \( j \in I \) and \( t_j \in T_j \) and suppose there is a sequence of common prior types \( t_j^n \) such that \( d_j^{uw}(t_j, t_j^n) \to 0 \). Then there is a sequence of finite common prior type spaces \( T^n \) such that \( h_{d^{uw}}(T, T^n) \to 0 \).
Proof. Since $T$ is a finite type space with full support, there exists $\eta > 0$ such that, for each $i \in I$, both of the following hold: 

$$d_i^{uw}(t_i, t_i') < 2\eta \quad \Rightarrow \quad t_i' = t_i \quad \forall t_i, t_i' \in T_i, \quad (34)$$

$$\eta < \mu_i^*(E|t_i) \quad \forall t_i \in T_i, \quad \forall \emptyset \neq E \subseteq \Theta \times T_{-i}. \quad (35)$$

Fix $\delta \in (0, \eta)$ and choose $n$ such that $d_j^{uw}(\tilde{t}_j, \tilde{t}_j') < \delta$. Let $\pi^n$ be a common prior with $\pi^n(t^n_j) > 0$. Since $\pi^n$ is a common prior, for each $i \in I$ we have

$$\pi^n\left(\{|t_i\}^{\delta, \infty}\right) > 0 \quad \Rightarrow \quad \left|\pi^n\left(E^{\delta, \infty}|\{|t_i\}^{\delta, \infty}\right) - \mu_i^*(E|t_i)\right| < \delta \quad \forall t_i \in T_i, \quad \forall E \subseteq \Theta \times T_{-i}. \quad (36)$$

Since $\pi^n\left(\{|\tilde{t}_j\}^{\delta, \infty}\right) \geq \pi^n(t^n_j) > 0$, it follows from (35) and (36) that, for each $i \in I$,

$$\pi^n\left(\{t_i\}^{\delta, \infty}\right) > 0 \quad \text{and} \quad \left|\pi^n\left(E^{\delta, \infty}|\{t_i\}^{\delta, \infty}\right) - \mu_i^*(E|t_i)\right| < \delta \quad \forall t_i \in T_i, \quad \forall E \subseteq \Theta \times T_{-i}. \quad (37)$$

Let $(T^n_i)_{i\in I}$ be the type space defined as follows. For each $i \in I$, $T^n_i = \{f^n_i(t_i) : t_i \in T_i\}$, where

$$\mu_i^*(\theta \times f^{-n}_{-i}(t_{-i})|f^n_i(t_i)) = \frac{\pi^n(\theta \times \{t_i\}^{\delta, \infty} \times \{t_{-i}\}^{\delta, \infty})}{\pi^n(\Theta \times \{t_i\}^{\delta, \infty} \times T_{-i}^{\delta, \infty})} \quad \forall \theta \in \Theta, \quad \forall t_i \in T_i, \quad \forall t_{-i} \in T_{-i},$$

which is well defined by (34). Clearly, $(T^n_i)_{i\in I}$ is a common prior type space, with the common prior $\tilde{\pi}^n$ defined as follows:

$$\tilde{\pi}^n(\theta, f^n_1(t_1), f^n_2(t_2)) = \frac{\pi^n(\theta \times \{t_1\}^{\delta, \infty} \times \{t_{-1}\}^{\delta, \infty})}{\pi^n(\Theta \times T_{-1}^{\delta, \infty})} \quad \forall (\theta, t_1, t_2) \in \Theta \times T_1 \times T_2.$$

Moreover, for each $i \in I, t_i \in T_i$ and $k \geq 1$ we have $d_i^k(t_i, f^n_i(t_i)) < \delta$. To prove the claim for $k = 1$, note that for each $E \subseteq \Theta$ we have

$$\mu_i^*(E|f^n_i(t_i)) = \frac{\pi^n\left(E \times \{t_i\}^{\delta, \infty} \times T_{-i}^{\delta, \infty}\right)}{\pi^n(\Theta \times \{t_i\}^{\delta, \infty} \times T_{-i}^{\delta, \infty})} \geq \pi^n\left(E \times T_{-i}^{\delta, \infty}|\{t_i\}^{\delta, \infty}\right) \geq \mu_i^*(E|t_i) - \delta,$$

where the second inequality follows from (37). Proceeding by induction, suppose the claim is true for some $k \geq 1$. Then, for each $E \subseteq \Theta \times T_{-i}$,

$$\mu_i^*(E^{\delta, k}|f^n_i(t_i)) \geq \mu_i^*(\{(\theta, f^{-k}_{-i}(t_{-i})) : (\theta, t_{-i}) \in E\}|f^n_i(t_i))$$

$$= \frac{\pi^n\left(E^{\delta, \infty} \times \{t_i\}^{\delta, \infty}\right)}{\pi^n(\Theta \times \{t_i\}^{\delta, \infty} \times T_{-i}^{\delta, \infty})} \geq \pi^n\left(E^{\delta, \infty}|\{t_i\}^{\delta, \infty}\right) \geq \mu_i^*(E|t_i) - \delta,$$

where the first inequality follows from the induction hypothesis, and the third from (37). Thus, $d_i^{uw}(t_i, f^n_i(t_i)) < \delta$ for each $i \in I$ and $t_i \in T_i$, and hence $h_{d^{uw}}(T, T^n) < \delta$. \hfill \blacksquare

\textsuperscript{51}By definition of conditional probability, if $\pi$ is a common prior, then, for each $i \in I$ and measurable $E \subseteq T_i^*$ with $\pi(E) > 0$, we have $\pi(F|E) = (1/\pi(E)) \int_E \mu_i^*(F|t_i) \pi(dt_i)$ for each measurable $F \subseteq \Theta \times T_{-i}$.  

58
Lemma 14. For each ε > 0 and finite type space T there exists a finite non-common prior type space with full support T' such that \( h_{d^{uw}}(T, T') < ε \).

Proof. Fix ε > 0 and a finite type space \( T = (T_i)_{i \in I} \). In order to construct the type space \( T' \), let \( δ ∈ (0, \frac{ε}{2}) \) and consider first the type space \( T'' = (T_i'')_{i \in I} \) defined as follows. For each \( i \in I \), \( T_i'' = \{ f_i(t_i) : t_i ∈ T_i \} \), where for each \( t_i \in T_i \) and \( E ⊆ Θ × T_{-i} \),

\[
μ_i^*(\{(θ, f_{-i}(t_{-i})) : (θ, t_{-i}) ∈ E\} | f_i(t_i)) = (1 - δ) μ_i^*(E | t_i) + \frac{δ |E|}{|Θ × T_{-i}|}.
\]

Clearly, \( T'' \) has full support, and an immediate induction argument shows that \( d_i^{uw}(t_i, f_i(t_i)) < \frac{ε}{2} \) for each \( i \in I \) and \( t_i \in T_i \). Thus, \( h_{d^{uw}}(T, T'') < \frac{ε}{2} \). If \( T'' \) is a non-common prior type space, then setting \( T' = T'' \) completes the proof. Otherwise, assume that \( T'' \) admits a common prior \( π \), fix \((\tilde{t}_1, \tilde{t}_2) \in T''\) and two distinct \( θ', θ'' \in Θ \), let \( 0 < δ' < \min_{i \in I, t_i'' ∈ T_i''} |E_π^*(E | t_i, T_i''')| \), and let \( T' = (T_i')_{i \in I} \) be the type space defined as follows. For each \( i \in I \), \( T_i' = \{ f'_i(t_i') : t_i'' ∈ T_i'' \} \), where

\[
μ_i^*(θ', f'_i(\tilde{t}_1)) | f'_i(\tilde{t}_2)) = μ_i^*(θ', f'_i(\tilde{t}_1)) | f'_i(\tilde{t}_2)) = μ_i^*(θ'', f'_i(\tilde{t}_1)) | f'_i(\tilde{t}_2)) = μ_i^*(θ'', f'_i(\tilde{t}_1)) | f'_i(\tilde{t}_2)) = μ_i^*(θ', f'_i(\tilde{t}_1)) | f'_i(\tilde{t}_2)) - δ'.
\]

and, for every \( i \in I \) and \( (θ, t_i''') \in Θ × T_i'' × T_{-i}'' \) for which \( μ_i^*(θ, f'_i(t_i''')) | f'_i(t_i''')) \) is not yet defined,

\[
μ_i^*(θ, f'_i(t_i''')) | f'_i(t_i''')) = μ_i^*(θ, t_i'' | t_i''').
\]

Clearly, \( T' \) has full support, and an immediate induction argument shows that \( h_{d^{uw}}(T'', T') < \frac{ε}{2} \). Thus, \( h_{d^{uw}}(T, T') \leq h_{d^{uw}}(T, T'') + h_{d^{uw}}(T'', T') < ε \). Moreover, \( T' \) does not admit a common prior, because

\[
\frac{μ_i^*(θ', f'_i(\tilde{t}_2)) | f'_i(\tilde{t}_1))}{μ_i^*(θ''', f'_i(\tilde{t}_2)) | f'_i(\tilde{t}_1))} = \frac{μ_i^*(θ', f'_i(\tilde{t}_2)) | f'_i(\tilde{t}_1))}{μ_i^*(θ''', f'_i(\tilde{t}_2)) | f'_i(\tilde{t}_1))} = \frac{π(θ', \tilde{t}_1, \tilde{t}_2)}{π(θ'', \tilde{t}_1, \tilde{t}_2)} = \frac{μ_i^*(θ', f'_i(\tilde{t}_2)) | f'_i(\tilde{t}_1))}{μ_i^*(θ''', f'_i(\tilde{t}_2)) | f'_i(\tilde{t}_1))} < \frac{μ_i^*(θ', f'_i(\tilde{t}_2)) | f'_i(\tilde{t}_1))}{μ_i^*(θ'', f'_i(\tilde{t}_2)) | f'_i(\tilde{t}_1))},
\]

where the second and third equality follow from the fact that \( π \) is a common prior for \( T'' \).

Lemma 15. For every finite non-common prior type space with full support \( T = (T_i)_{i \in I} \) there exists \( ε > 0 \) such that \( h_{d^{uw}}(T, T') > ε \) for every common prior type space \( T' \). Furthermore, \( d_i^{uw}(t_i, t'_i) > ε \) for every \( i \in I, t_i \in T_i \) and common prior type \( t'_i \in T_i^* \).

Proof. The second claim is an immediate consequence of the first claim and the proof of Lemma 13. To prove the first claim, first note that since \( T \) is finite, there exists \( η > 0 \) such that

\[
d_i^{uw}(t_i, t'_i) < 2η \quad ⇒ \quad t'_i = t_i \quad ∀ t_i, t'_i ∈ T_i.
\]

(38)
Moreover, since $T$ is finite and has no common prior, there exists $f : \Theta \times T \to \mathbb{R}$ such that
\[
\sum_{(\theta, t') \in \Theta \times T_2} f(\theta, t_1, t'_2) \mu^*_1(\theta, t'_2 | t_1) > 0
\]
\[
> \sum_{(\theta, t') \in \Theta \times T_1} f(\theta, t'_1, t_2) \mu^*_2(\theta, t'_1 | t_2) \quad \forall (t_1, t_2) \in T. \quad (39)
\]

(See, for instance, Samet, 1998) Now suppose, contrary to our claim, that $h_{d_{uw}}(T, T^n) \to 0$ for a sequence of common prior type spaces $T^n$. Choose $N$ so that $h_{d_{uw}}(T, T^n) < \eta$ for each $n \geq N$, and for each such $n$ define $f^n : \Theta \times T^n \to \mathbb{R}$ by letting
\[
f^n(\theta, t^n_1, t^n_2) = f(\theta, t_1, t_2) \quad \forall (t_1, t_2) \in T, \forall (t^n_1, t^n_2) \in T_1^{n,\infty} \times T_2^{n,\infty}, \quad (40)
\]
which is well defined by (38) and by the fact that $T_i^n \subseteq T_i^{n,\infty}$ for each $i \in I$. Then, it follows from (39) and (40) that, for each sufficiently large $n$,
\[
\sum_{(\theta, t^n) \in \Theta \times T^n_2} f^n(\theta, t^n_1, t^n_2) \mu^*_1(\theta, t'_2 | t^n_1) > 0
\]
\[
> \sum_{(\theta, t^n) \in \Theta \times T^n_1} f(\theta, t'_1, t^n_2) \mu^*_2(\theta, t'_1 | t^n_2) \quad \forall (t^n_1, t^n_2) \in T^n. \quad (41)
\]

But, letting $\pi^n$ denote the common prior for the type space $T^n$, we then reach the contradiction that $\sum_{(\theta, t^n) \in \Theta \times T^n} \pi^n(\theta, t^n)$ is both positive and negative. \hfill \blacksquare

**Lemma 16.** Let $T = (T_i)_{i \in I}$ be a finite type space, and let $T^n = (T_i^n)_{i \in I}$ be a sequence of type spaces in $T$. Then, $h_{d_{uw}}(T, T^n) \to 0$ if and only if $h_{d_{uw}}(T, T^n) \to 0$.

**Proof.** Since uniform-weak convergence implies product convergence, it suffices to prove the “only if” part. First note that, since $(T_i)_{i \in I}$ is a finite type space, by Lemma 12 there exist $\eta > 0$ and $k \geq 1$ such that for every $\delta \in (0, \eta)$, $i \in I$ and $t_i \in T_i$,
\[
\{t_i\}_{i \in I} \cap C_i^1(\Theta \times T_1^\delta \times T_2^\delta) \subseteq \{t_i\}_{i \in I}. \quad (41)
\]

Now fix any $\delta \in (0, \eta)$. Since $h_{d_{uw}}(T, T^n) \to 0$, there exists $N \geq 1$ such that
\[
\sup_{t_i \in T_i, t^n_1 \in T^n_1} d_i^\delta(t_i, t^n_1) < \delta \quad \text{and} \quad \sup_{t^n_1 \in T^n_1} \inf_{t_i \in T_i} d_i^\delta(t_i, t^n_1) < \delta \quad \forall i \in I, \forall n \geq N. \quad (42)
\]

Then, for each $i \in I$ and $n \geq N$, we have $T^n_i \subseteq T_i^\delta$ and hence, using the fact that $T_i^n \subseteq C_i^1(\Theta \times T_1^n \times T_2^n)$, also $T_i^n \subseteq C_i^1(\Theta \times T_1^\delta \times T_2^\delta)$. It follows from (41) that for each $i \in I$, $n \geq N$, $t_i \in T_i$ and $t^n_1 \in T^n_1$, if $d_i^\delta(t_i, t^n_1) < \delta$ then $d_{i,uw}(t_i, t^n_1) < \delta$. Thus, by (42), $h_{d_{uw}}(T, T^n) < \delta$. \hfill \blacksquare

60
We can now prove Theorems 4 and 5. Dekel, Fudenberg, and Morris (2006, Theorem 3) show that finite types are dense under the strategic topology. Since the uniform weak topology is finer than the strategic topology, it follows from Lemma 14 that the set of types that belong to some finite non-common prior type space with full support are also dense under the strategic topology. By Lemma 15 and Theorem 3, each such type belongs to a strategically open set that does not contain any common prior type. Thus, Theorem 4 follows.

Mertens and Zamir (1985, Theorem 3.1) show that the set of finite type spaces is dense in $T$ under the product topology on type spaces. Since the uniform weak topology is finer than the product topology, it follows from Lemma 14 that the set of finite non-common prior type spaces with full support is also dense in $T$ under the product topology on type spaces. By Lemma 15 and Lemma 16, each finite non-common prior type space with full support also belongs to a subset of $T$ that is open under the product topology on type spaces and does not contain any common prior type space. Thus, Theorem 5 follows.

References


