Topologies on Types: Connections*

Yi-Chun Chen[†] Siyang Xiong[‡]

May 28, 2008

Abstract

For different purposes, economists may use different topologies on types. We characterize the relationship among these various topologies. First, we show that for any general types, convergence in the uniform-weak topology implies convergence in both the strategic topology and the uniform strategic topology. Second, we explicitly construct a type which is not the limit of any finite types under the uniform strategic topology, showing that the uniform strategic topology is strictly finer than the strategic topology. With these results, we can linearly rank various topologies on the universal type space, which gives a clear picture of the relationship between the implication of types for beliefs and their implication for behaviors.

Keywords: the universal type space, the strategic topology; the uniform strategic topology; the uniform-weak topology; interim correlated rationalizable actions

JEL Classification: C70

*We are grateful to Eddie Dekel for his invaluable support, guidance, and detailed comments. We also thank Jeff Ely for helpful discussions and Aaron Sojourner for useful comments. All remaining errors are our own.

†Department of Economics, Northwestern University, 2001 Sheridan Road, Evanston, IL 60208, yichunchen2010@u.northwestern.edu

‡Department of Economics, Northwestern University, 2001 Sheridan Road, Evanston, IL 60208, xiong@u.northwestern.edu

1 Introduction

If two agents have similar beliefs on some payoff-relevant uncertainty, do they always make similar decisions? In the single-person-decision-making setup, where proximity of beliefs about uncertainty can be measured by the standard weak*-topology, the answer is clearly yes. The answer is much more complicated in the multi-person-game setup where agents' beliefs about each other's beliefs play a crucial role in determining the outcome of the interaction. In this setup, do similar beliefs still imply similar behaviors?

One compact way to formulate the complicated object of "beliefs about beliefs" is to use the notion of types developed by Harsanyi (1967/1968). A player's type specifies a belief over the set of unknown relevant parameters and opponents' type profiles. Mertens and Zamir (1985) shows that the set of all coherent hierarchies of beliefs is the universal type space and Harsanyi's idea suffers no loss of generality.

In practice, an ideal formulation of an economic situation may involve a very complicated type space which is hard to analyze directly. For reasons of tractability, economists may therefore replace this ideal formulation with some simpler type space in order to solve a simpler problem. To make this approximation meaningful, we should require that these simpler types be "close" to the true types so that they will exhibit "similar" behaviors in games. To measure the "closeness" of types, we need to define a topology on types.

Two kinds of information are encapsulated in a type. On the one hand, defined directly on hierarchies of beliefs, a type contains an agent's belief information. On the other hand, as shown in Dekel, Fudenberg, and Morris (2006) [hereafter DFM], any two different Mertens-Zamir types have different behaviors in some game, and hence, the behaviors of an agent are also implicitly encoded in a type. We can therefore define topologies on types according to these two aspects. Examples of "belief" topologies include the *product topology* introduced by Mertens and Zamir (1985) and the *uniform-weak topology* introduced by Di Tillio and Faingold (2007). Examples of "behavior" topologies are the *strategic topology* and the *uniform strategic topology* both introduced by DFM. A careful study of the connection among these topologies will reveal the relationship between the belief implication of types and the behavior implication of types, which is the main theme of this paper.

The product topology is used by Mertens and Zamir to achieve their homeomorphism result. A sequence of types t^n converges to a type t in the product topology if and only if the k^{th}-order belief of t^n converges to that of t for any k. Since proximity of tails of the hierarchies of t^n and t is not required, it is now well known that we may have two types which are close under the product topology but exhibit very different strategic behaviors (see Rubinstein (1989)).

This raises the question as to whether there exists a topology under which nearby types always have similar strategic behavior (see Monderer and Samet (1989), Monderer and Samet (1996), Kajii and Morris (1998), and Dekel, Fudenberg, and Morris (2006)). In particular, DFM propose the strategic topology which is just strong enough to guarantee that for every finite game, the correspondence which maps types into $\varepsilon-$interim-correlated-rationalizable (ε-ICR) actions is continuous. More precisely, DFM show that upper-hemicontinuity of the ε-ICR correspondence under strategic convergence is equivalent to product convergence, and it is precisely the lower-hemicontinuity property which makes strategic convergence a more stringent requirement.

DFM also introduce the notion of uniform strategic convergence, which adds to the notion of strategic convergence the additional requirement that the degree of similarity of ε-ICR actions is *uniform* over all bounded finite games. In contrast to the strategic topology, the uniform strategic topology has its own importance, especially when applied to mechanism design. We discuss this importance further in Section 5.1.

Though useful, DFM's definitions of strategic topologies are complex since they involve direct reference to best replies, which are tied to games. In particular, we are not able to tell whether a sequence of types converges in the strategic topology, unless we study the behaviors of all these types in all finite games. To address this issue, Di Tillio and Faingold (2007) propose the uniform-weak topology. A sequence of types t^n converges to a type t in the uniform-weak topology if and only if the k^{th}-order belief of t^n converges to that of t for any order k and the rate of convergence is uniform over k. Di Tillio and Faingold show that around any finite type, the strategic topology is fully characterized by the uniform-weak topology.

To characterize the relationship among these various topologies, we first study the implication of uniform-weak convergence around general types. Our first main result (Theorem 1) shows that uniform-weak convergence always guarantees uniform strategic convergence (hence also strategic convergence) around any types. This gives a sufficient condition for uniform strategic convergence (hence, also for strategic convergence), which is easy to use.[1] Coupled with Di Tillio and Faingold's result, we see that around finite types, uniform-weak convergence, strategic convergence and uniform strategic convergence are all equivalent.

In a recent paper, Ely and Peski (2007) propose an insightful partition of the universal type space into *regular* and *critical types*. Regular types are types around which the strategic topology is equivalent to the product topology. Around critical types the strategic topology is strictly finer. Ely and Peski (2007) offer this surprisingly concise characterization of critical types: a type is critical if and only if for some $p > 0$, it has common p-belief for some closed proper subset in the universal type space. Therefore, finite type spaces as well as other type spaces typically considered in applications consist entirely of critical types.

Our first main result leaves open the question as to whether Di Tillio and Faingold's equivalence result on finite types can be extended to all critical types, i.e., does strategic convergence also imply uniform-weak convergence around *any* critical type? To answer this question, we construct an infinite critical type which cannot be approximated by any sequence of finite types under the uniform strategic topology. This result, coupled with our first main result and DFM's denseness result for finite types under the strategic topology, shows that the Di Tillio and Faingold's equivalence result does not hold for a general critical type. Moreover, in sharp contrast to DFM's denseness result, our second main result (Theorem 2) shows that finite types are *nowhere dense* under the uniform strategic topology.[2] Hence, by our first main result, finite types are nowhere dense under the uniform-weak topology as well.

Given two topologies $\Im$ and $\Im'$, let "$\Im \succeq \Im'$" mean that $\Im$ is weakly finer than $\Im'$

[1] As shown in Di Tillio and Faingold (2007), the uniform-weak topology is a generalized notion of the common p-belief introduced by Monderer and Samet (1989).

[2] DFM conjecture that finite types are not dense under the uniform strategic topology. They aim to use their Proposition 2 to prove this conjecture. However, Chen and Xiong (2008) prove that their Proposition 2 is not true by constructing a counterexample. Hence, prior to the current paper, whether finite types were dense under the uniform strategic topology was an open question.

and let "$\Im \succ \Im'$" mean that $\Im$ is strictly finer than $\Im'$. Our main contribution can be summarized as follows. Our first main result shows that [uniform-weak topology]$\succeq$[uniform strategic topology].[3] Coupled with DFM's denseness result, our second main result shows that [uniform strategic topology]$\succ$[strategic topology]. Moreover, DFM show that [strategic topology]$\succ$[product topology]. Therefore, we show that these topologies are related as follows.

[uniform-weak topology] $\succeq$ [uniform strategic topology] $\succ$ [strategic topology] $\succ$ [product topology].

The next section of this paper contains basic definitions and notations. In Section 3, we present our first main result and a sketch of the proof. We turn to the issue of non-denseness of finite types in Section 4. In Section 5, we offer some discussion about related issues. All proofs are relegated to the Appendix.

2 Preliminaries

Throughout this paper, for any arbitrary separable metric space Y with metric d_Y, let $\Delta(Y)$ be the space of all probability measures on the Borel σ-algebra of Y endowed with the weak*-topology. It is well known that the weak*-topology is metrizable with the Prohorov distance ρ defined as

$$\rho(\mu, \mu') = \inf\{\gamma > 0 : \mu(E) \leq \mu'(E^{\gamma}) + \gamma \text{ for every Borel set } E \subseteq Y\}, \forall \mu, \mu' \in \Delta(Y)$$

where $E^{\gamma} \equiv \{y' : \inf_{y \in E} d_Y(y', y) < \gamma\}$. Unless explicitly noted, all product spaces will be endowed with the product topology and subspaces with the relative topology. Every finite or countable set is endowed with the discrete topology and denote the cardinality of a finite set E by $|E|$. Moreover, let suppμ denote the support of a measure μ defined on a finite set. Finally, for any $E \subseteq Y$ and $y \in Y$, let 1_E be the indicator function on E and δ_y be the point mass on y.

For simplicity, assume that there are two players, player 1 and player 2. Given a player $i \in \{1, 2\}$, let $-i$ denote the other player in $\{1, 2\}$. The basic uncertainty is a finite set which

[3] Whether [uniform-weak topology]$\succ$[uniform strategic topology] is true remains an open question, for which we do not have an answer.

is denoted by Θ. Let $Y^0 = \Theta$ and $Y^1 = Y^0 \times \Delta(Y^0)$. Then, for $k \geq 2$ define recursively

$$Y^k = \left\{ \left(\theta, \mu^1, ..., \mu^k\right) \in Y^0 \times \Delta\left(Y^0\right) \times \cdots \times \Delta\left(Y^{k-1}\right) : \mathrm{marg}_{Y^{l-2}} \mu^l = \mu^{l-1}, \forall l = 2, ..., k \right\}.$$

Then, the Mertens-Zamir universal type space is defined as

$$\mathcal{T} = \left\{ \left(\mu^1, \mu^2, ...\right) \in \times_{k=0}^{\infty} \Delta\left(Y^k\right) : \mathrm{marg}_{Y^{l-2}} \mu^l = \mu^{l-1}, \forall l \geq 2 \right\}.$$

For each $k \geq 1$, let $\pi^k : \mathcal{T} \to \Delta\left(Y^{k-1}\right)$ be the natural projection. For every player i and $k \geq 1$, let $\mathcal{T}_i$ and Y_i^k denote the copies of $\mathcal{T}$ and Y^k respectively, write $\pi_i^k : \mathcal{T}_i \to \Delta\left(Y_{-i}^{k-1}\right)$ for π^k, and define $\mathcal{T}_i^k = \pi_i^k(\mathcal{T}_i)$. An element $t_i \in \mathcal{T}_i$ is a type of player i. For simplicity, we will write t_i^k instead of $\pi_i^k(t_i)$ for the k^{th}-order belief of type t_i.[4] By the result of Mertens and Zamir (1985), $\mathcal{T}_i$ (endowed with product topology) is homeomorphic to $\Delta(\Theta \times \mathcal{T}_{-i})$. Let π_i^* denote this homeomorphism. In the Mertens-Zamir construction, for any type t_i, the marginal distribution of $\pi_i^*(t_i)$ on Y_{-i}^{k-1} agrees with the distribution t_i^k.

Let ρ^0 be the discrete metric on $Y^0 = \Theta$, i.e., $\rho^0(\theta, \theta') = 1$ if $\theta \neq \theta'$ and $\rho^0(\theta, \theta) = 0$. For $k \geq 1$, let ρ^k be the Prohorov metric on $\Delta\left(Y^{k-1}\right)$ with respect to the metric d^{k-1} on Y^{k-1} defined recursively as

$$d^{k-1}\left[\left(\theta, ..., \mu^{k-1}\right), \left(\theta', ..., \nu^{k-1}\right)\right] = \max\{\rho^0(\theta, \theta'), \rho^{k-1}(\mu^1, \nu^1), ..., \rho^{k-1}(\mu^{k-1}, \nu^{k-1})\}.$$

As defined by Di Tillio and Faingold (2007), the uniform-weak topology is generated by the metric

$$d^{uw}(t, s) \equiv \sup_{k \geq 1} \rho^k\left(t^k, s^k\right) \text{ for types } t \text{ and } s \text{ in } \mathcal{T}.$$

Following DFM, we assume that there is a fixed exogenous bound $M > 0$ for the payoffs of all finite games we consider. Let $G = (A_i, g_i)_{i=1,2}$ be a finite game where A_i is a finite set of actions for player i and $g_i : A_i \times A_{-i} \times \Theta \to [-M, M]$ is the payoff function. For $\gamma \geq 0$, we will use the following recursive definition of $\gamma-$interim-correlated-rationalizable set which is equivalent to the fixed-point definition (see Dekel, Fudenberg, and Morris (2007)). Let

$$R(G, \gamma) \equiv (R_i(G, \gamma))_{i=1,2} \equiv \left(\left(R_i(t_i, G, \gamma)_{t_i \in \mathcal{T}_i}\right)\right)_{i=1,2} \equiv \cap_{k=0}^{\infty} R^k(G, \gamma)$$

[4] Note that $\mathcal{T}_i^k = \pi_i^k(\mathcal{T}_i)$ and hence when we write $t_i^k \in \mathcal{T}_i^k$ without specifying the type t_i, t_i^k should be understood as the k^{th}-order belief of some type $t_i \in \mathcal{T}_i$.

where

$$\begin{aligned} R^0(G,\gamma) &\equiv \left((A_i)_{t_i\in\mathcal{T}_i}\right)_{i=1,2}, \\ R^k(G,\gamma) &\equiv \left(R_i^k(G,\gamma)\right)_{i=1,2} \equiv \left(\left(R_i^k(t_i,G,\gamma)_{t_i\in\mathcal{T}_i}\right)\right)_{i=1,2}, \ \forall k\geq 1, \end{aligned}$$

and $a_i \in R_i^k(t_i,G,\gamma)$ if and only if there exists a measurable function $\sigma_{-i}:\Theta\times\mathcal{T}_{-i}\to\Delta(A_{-i})$ such that[5]

$$\text{supp}\sigma_{-i}(\theta,t_{-i}) \subseteq R_{-i}^{k-1}(t_{-i},G,\gamma) \text{ for } \pi_i^*(t_i) - \text{almost surely } (\theta,t_{-i});$$

$$\int_{\Theta\times\mathcal{T}_{-i}} \left[\mathbf{g}_i(a_i,a_i',\theta)\bullet\sigma_{-i}(\theta,t_{-i})\right]\pi_i^*(t_i)\left[(\theta,dt_{-i})\right] \geq -\gamma \text{ for all } a_i'\in A_i\backslash\{a_i\} \text{ where}$$

$$\mathbf{g}_i(a_i,a_i',\theta) \equiv \left(g_i(a_i,a_{-i},\theta) - g_i(a_i',a_{-i},\theta)\right)_{a_{-i}\in A_{-i}}.$$

For any $\beta\in\Delta(A_i)$, let $\mathbf{g}_i(a_i,\beta,\theta)\in\Re^{|A_{-i}|}$ denote the vector of expected payoff difference under β, i.e.,

$$\mathbf{g}_i(a_i,\beta,\theta) \equiv \left(\sum_{a_i'\in A_i}\beta(a_i')\left[g_i(a_i,a_{-i},\theta) - g_i(a_i',a_{-i},\theta)\right]\right)_{a_{-i}\in A_{-i}}$$

Observe that $R_i^k(t_i,G,\gamma)$ depends only on the k^{th}-order belief of type t_i.

Through the following two lemmas, we will reach an alternative characterization of the $\gamma-$ICR set which will be used in the proof of Theorem 1. Their proofs can be found in Appendix A.1. First, in the definition of $R_i^k(t_i,G,\gamma)$ above, a conjecture σ_{-i} is a measurable function from $\Theta\times\mathcal{T}_{-i}$ to $\Delta(A_{-i})$. Since we will study the influence of k^{th}-order beliefs on $R_i^k(t_i,G,\gamma)$, the following definition and lemma offer a useful alternative definition of $R_i^k(t_i,G,\gamma)$.

Definition 1 $\overline{R}_i^0(t_i,G,\gamma) = A_i$ *for all* t_i *and* i. *For* $k\geq 1$, $a_i\in\overline{R}_i^k(t_i,G,\gamma)$ *if and only if there exists a measurable function* $\sigma_{-i}:\Theta\times\mathcal{T}_{-i}^{k-1}\to\Delta(A_{-i})$ *such that*

$$\textit{supp}\sigma_{-i}\left(\theta,t_{-i}^{k-1}\right) \subseteq \overline{R}_{-i}^{k-1}(t_{-i},G,\gamma) \text{ for } t_i^k - \textit{almost surely } \left(\theta,t_{-i}^{k-1}\right); \quad (1)$$

$$\int_{\Theta\times\mathcal{T}_{-i}^{k-1}} \left[\sigma_{-i}\left(\theta,t_{-i}^{k-1}\right)\bullet\mathbf{g}_i(a_i,a_i',\theta)\right] t_i^k\left[\left(\theta,dt_{-i}^{k-1}\right)\right] \geq -\gamma \textit{ for all } a_i'\in A_i, \quad (2)$$

where $\Theta\times\mathcal{T}_{-i}^0\equiv\Theta$.

[5] Throughout the paper "•" stands for the inner product of two vectors with the same dimension.

Lemma 1 $\overline{R}_i^k(t_i, G, \gamma) = R_i^k(t_i, G, \gamma)$ *for every integer* $k \geq 0$*, every* t_i*, and every player* i*.*

Hereafter, a measurable function σ_{-i} from $\Theta \times \mathcal{T}_{-i}^{k-1}$ to $\Delta(A_{-i})$ which satisfies condition (1) is said to be a *valid conjecture*. Moreover, an action a_i which satisfies condition (2) is said to be a $\gamma-$*best reply under* σ_{-i} *for type* t_i. Second, the following result shows that proving $a_i \in \overline{R}_i^k(t_i, G, \gamma)$ is equivalent to proving that for any mixed action $\beta \in \Delta(A_i \backslash \{a_i\})$, we can find a valid conjecture σ_{-i} (which may vary with β) such that playing a_i is γ-better than playing β. Note that this equivalent characterization reverses the quantifier of (2) which requires a valid conjecture σ_{-i} working for all a_i' (and hence for all $\beta \in \Delta(A_i \backslash \{a_i\})$).[6]

Lemma 2 *For any positive integer* k*, any* $\gamma \geq 0$*, any finite game* G*, and any type* $t_i \in \mathcal{T}_i$*,* $a_i \in R_i^k(t_i, G, \gamma)$ *if and only if for every* $\eta > 0$ *and* $\beta \in \Delta(A_i \backslash \{a_i\})$ *there is a valid conjecture* $\sigma_{-i}[\beta] : \Theta \times \mathcal{T}_{-i}^{k-1} \to \Delta(A_{-i})$ *for* t_i *under which*

$$\int_{\Theta \times \mathcal{T}_{-i}^{k-1}} \left[\mathbf{g}_i(a_i, \beta, \theta) \bullet \sigma_{-i}[\beta]\left(\theta, t_{-i}^{k-1}\right)\right] t_i^k\left[\left(\theta, dt_{-i}^{k-1}\right)\right] \geq -\gamma - \eta.$$

For each $t_i \in \mathcal{T}_i$, define $h_i(t_i|a_i, G)$ to be the minimal γ under which an action a_i is $\gamma-$rationalizable for t_i in G, i.e.,

$$h_i(t_i|a_i, G) = \min\{\gamma : a_i \in R_i(t_i, G, \gamma)\}.$$ [7]

Fix $\alpha \in (0, 1)$. Let $\mathcal{G}^m$ be the collection of all games where each player has m actions and the payoffs are bounded by M. Let $\mathcal{G}$ be the collection of all two-player finite games with payoffs bounded by M. Following DFM, we define the strategic topology and the uniform strategic topology on types to be the topologies generated by the metrics d^s and d^{us} respectively, where

$$\begin{aligned} d^s(t_i, s_i) &\equiv \sum_{m=1}^{\infty} \alpha^m \sup_{a_i \in A_i(G), G \in \mathcal{G}^m} |h_i(t_i|a_i, G) - h_i(s_i|a_i, G)| \text{ for types } t_i \text{ and } s_i \text{ in } \mathcal{T}_i; \\ d^{us}(t_i, s_i) &\equiv \sup_{a_i \in A_i(G), G \in \mathcal{G}} |h_i(t_i|a_i, G) - h_i(s_i|a_i, G)| \text{ for types } t_i \text{ and } s_i \text{ in } \mathcal{T}_i. \end{aligned}$$

Clearly, $d^{us}(t_i, s_i) \geq d^s(t_i, s_i)$.

[6] This minimax step is conceptually different from the standard equivalence between the never-best reply and a strictly dominated strategy, though the technical content is similar.

[7] DFM's Proposition 1 shows that the minimum exsits.

3 $d^{uw}-$convergence $\Rightarrow$ $d^{us}-$convergence

We are now ready to state our main result as follows.

Theorem 1 *For any $\varepsilon > 0$ and two types t_i and s_i in $\mathcal{T}_i$ with $d^{uw}(t_i, s_i) \leq \varepsilon$, we have $d^{us}(t_i, s_i) \leq 6M\varepsilon$ and hence $d^{s}(t_i, s_i) \leq 6M\varepsilon$.*

Theorem 1 is an immediate consequence of the following proposition.

Proposition 1 *For any finite game G, any $\varepsilon, \gamma \geq 0$, and any types t_i and s_i in $\mathcal{T}_i$ with $\rho^k\left(t_i^k, s_i^k\right) \leq \varepsilon$, we have $R_i^k(t_i, G, \gamma) \subseteq R_i^k(s_i, G, \gamma + 6M\varepsilon)$ for every integer $k \geq 0$.*

The formal proof of Proposition 1 is long and will be presented in Appendix A.2. Here we offer a sketch to highlight the essential ideas. We prove this proposition by induction on k and divide the proof of the induction step into five sub-steps. We now provide a roadmap by describing the role of each step.

First, for $a_i \in R_i^k(t_i, G, \gamma)$, there is some valid conjecture σ_{-i} under which a_i is a $\gamma-$best reply for type t_i. Our goal is to show that $a_i \in R_i^k(s_i, G, \gamma + 6M\varepsilon)$ by finding another valid conjecture σ'_{-i} under which a_i is a $(\gamma + 6M\varepsilon)-$best reply for type s_i. By Lemmas 1 and 2, it suffices to show that for every $\eta > 0$ and $\beta \in \Delta(A_i \backslash \{a_i\})$, there exists a valid conjecture σ'_{-i} such that

$$\int_{\Theta \times \mathcal{T}_{-i}^{k-1}} \left[\mathbf{g}_i(a_i, \beta, \theta) \bullet \sigma'_{-i}\left(\theta, t_{-i}^{k-1}\right)\right] t_i^k\left[\left(\theta, dt_{-i}^{k-1}\right)\right] \geq -\gamma - 6M\varepsilon - \eta. \tag{3}$$

To find the desired σ'_{-i}, we observe first that s_i^k must assign a large probability to $\left(\mathbf{T}_{-i}^{k-1}\right)^{\varepsilon}$ where $\mathbf{T}_{-i}^{k-1}$ is the set of $\left(\theta, t_{-i}^{k-1}\right)$s where σ_{-i} randomizes among $\gamma-$rationalizable actions of player $-i$ (step 1). Then, we define $\sigma'_{-i} = \sigma_{-i}$ on $\mathbf{T}_{-i}^{k-1}$ and it is immaterial how we define σ'_{-i} outside $\left(\mathbf{T}_{-i}^{k-1}\right)^{\varepsilon}$ so long as it is valid. To define σ'_{-i} on $\left(\mathbf{T}_{-i}^{k-1}\right)^{\varepsilon} \backslash \mathbf{T}_{-i}^{k-1}$, we first find a finite partition $\{\Phi^m\}$ of $\Delta(A_{-i})$ such that in each partition cell, there is some representative probability vector $\mathbf{q}^m$ which supnorm-approximates every $\mathbf{q}$ in Φ^m. We then use the pre-image of $\{\Phi^m\}$ under σ_i to induce a partition $\left\{F_m^{\theta}\right\}$ on $\mathbf{T}_{-i}^{k-1}$ (step 2).

Second, step 3 obtains a suitable measurable extension of σ'_{-i} from F_m^θ to $\left(F_m^\theta\right)^\varepsilon$ so that the induction hypothesis can be invoked to take care of the validity of σ'_{-i}. Third, in step 4, we solve the double-counting problem which arises when we specify σ'_{-i} on the intersection of $\left(F_m^\theta\right)^\varepsilon$ and $\left(F_{m'}^{\theta'}\right)^\varepsilon$. Here we will see the advantage of allowing σ'_{-i} to depend on β. Given β, there is one obvious way to achieve (3) by maximizing $\mathbf{g}_i\left(a_i,\beta,\theta\right)\bullet\sigma'_{-i}\left(\theta,t_{-i}^{k-1}\right)$—i.e., to assign σ'_{-i} according to the extension on $\left(F_m^\theta\right)^\varepsilon$ obtained in step 3 whenever $\mathbf{g}_i\left(a_i,\beta,\theta\right)\bullet\mathbf{q}^m\geq\mathbf{g}_i\left(a_i,\beta,\theta'\right)\bullet\mathbf{q}^{m'}$. Step 5 illustrates the idea of proving (3) under σ'_{-i}.

3.1 Step 1: focus on the support of t_i^k

Let $\mathbf{T}_{-i}^{k-1}\equiv\left\{\left(\theta,t_{-i}^{k-1}\right)\in\Theta\times\mathcal{T}_{-i}^{k-1}:\text{supp}\sigma_{-i}\left(\theta,t_{-i}^{k-1}\right)\subseteq R_i^{k-1}\left(t_{-i}^{k-1},G,\gamma\right)\right\}$. Since σ_{-i} is valid, $t_i^k\left(\mathbf{T}_{-i}^{k-1}\right)=1$. Moreover, with $\rho^k\left(t_i^k,s_i^k\right)\leq\varepsilon$, we have $s_i^k\left[\left(\mathbf{T}_{-i}^{k-1}\right)^\varepsilon\right]\geq1-\varepsilon$. Since the payoff has a uniform bound M, how to define σ'_{-i} on the complement of $\left(\mathbf{T}_{-i}^{k-1}\right)^\varepsilon$ is not important because the payoff resulted from this region is at most $M\varepsilon$. For $\left(\theta,t_{-i}^{k-1}\right)$ in $\mathbf{T}_{-i}^{k-1}$, we can take care of its validity by simply choosing σ'_{-i} to be identical to σ_{-i}. The key issue is to suitably define σ'_{-i} for $\left(\theta,t_{-i}^{k-1}\right)$ in $\left(\mathbf{T}_{-i}^{k-1}\right)^\varepsilon$ but outside $\mathbf{T}_{-i}^{k-1}$, which will be discussed in step 3.

3.2 Step 2: discretize $\Delta\left(A_{-i}\right)$

For the conjecture σ'_{-i} that we are about to define, we will prove $a_i\in R_i^k\left(s_i,G,\gamma+6M\varepsilon\right)$ by showing that (3) holds. However, it is not obvious how the condition $\rho^k\left(t_i^k,s_i^k\right)\leq\varepsilon$ can be applied to evaluate the payoff difference. To solve this problem, we discretize the simplex $\Delta\left(A_{-i}\right)$. Note that $\Delta\left(A_{-i}\right)$ is a $\left(\left|A_{-i}\right|-1\right)$-dimensional compact set. Therefore, for any positive integer h, we can find a finite partition $\{\Phi^m\}_{m=1}^{\overline{\Delta}}$ of $\Delta\left(A_{-i}\right)$ such that $\mathbf{q}^m\in\Phi^m$ is a representative element of Φ^m, and the supnorm of $\left(\mathbf{q}^m-\mathbf{q}\right)$ is no more than $\frac{1}{h}$ for any $\mathbf{q}\in\Phi^m$.

We can partition $\mathbf{T}_{-i}^{k-1}$ into $N=|\Theta|\times\overline{\Delta}$ sets $\left\{F_m^\theta\right\}$ where for each θ and m,

$$F_m^\theta\equiv\{\left(\theta,t_{-i}^{k-1}\right)\in\mathbf{T}_{-i}^{k-1}:\sigma_{-i}\left(\theta,t_{-i}^{k-1}\right)\in\Phi^m\}.$$

Then, we can approximate the expected payoff of t_i under σ_{-i} by replacing $\sigma_{-i}\left(\theta,t_{-i}^{k-1}\right)$ with

$\mathbf{q}^m$ if $\left(\theta, t_{-i}^{k-1}\right)$ belongs to F_m^θ. This approximation has at most an error of $\frac{2M|A_{-i}|}{h}$ which can be arbitrarily small if h is sufficiently large.

3.3 Step 3: measurable extension of the conjecture

We now specify σ'_{-i} for $\left(\theta, t_{-i}^{k-1}\right)$ in $\left(\mathbf{T}_{-i}^{k-1}\right)^\varepsilon$ but outside $\mathbf{T}_{-i}^{k-1}$. By step 1, we have defined $\sigma'_{-i} = \sigma_{-i}$ on $\mathbf{T}_{-i}^{k-1}$. By step 2, $\mathbf{T}_{-i}^{k-1} = \cup_{\theta,m} F_m^\theta$ and $\left(\mathbf{T}_{-i}^{k-1}\right)^\varepsilon = \cup_{\theta,m} \left(F_m^\theta\right)^\varepsilon$. Hence, one way to solve the problem is to extend the conjecture on F_m^θ to $\left(F_m^\theta\right)^\varepsilon$ for each (θ, m). We have to take care of two requirements to guarantee the validity of σ'_{-i}. First, σ'_{-i} must be measurable. Second, for any $\left(\theta, t_{-i}^{k-1}\right) \in \left(F_m^\theta\right)^\varepsilon$, $\sigma'_{-i}\left(\theta, t_{-i}^{k-1}\right)$ must be close to $\mathbf{q}^m$ so that the approximation in step 2 is still valid. These requirements are handled by the following lemma whose proof can be found in Appendix A.2.1.

Lemma 3 *Consider a separable metric space* (Y, d_Y)*, a Borel set* $F \subseteq Y$*, and* $\varepsilon > 0$*. Suppose* $f : F \to Z$ *is a measurable function from* F *to another measurable space* Z*. Then, there is a measurable function* $f^\varepsilon : F^\varepsilon \to Z$ *such that* $f^\varepsilon = f$ *on* F*, and for every* $y \in F^\varepsilon \backslash F$*,* $f^\varepsilon(y) = f(y')$ *for some* $y' \in F$ *with* $d_Y(y, y') < \varepsilon$*.*

The lemma guarantees that for $\left(\theta, t_{-i}^{k-1}\right)$ within $\left(F_m^\theta\right)^\varepsilon$ but outside F_m^θ, we can define $\sigma'_{-i}\left(\theta, t_{-i}^{k-1}\right)$ to be equal to $\sigma_{-i}\left(\theta', s_{-i}^{k-1}\right)$ for some $\left(\theta', s_{-i}^{k-1}\right)$ in F_m^θ and still preserve measurability. The validity of σ'_{-i} will then be granted by the induction hypothesis. Moreover, by doing so we reduce the comparison of the expected payoffs to the much more tractable task of evaluating the probability differences on the sets $\left(F_m^\theta\right)^\varepsilon$ and F_m^θ subject to a double counting problem to be discussed in the next step. That is, the expected payoffs for t_i under σ_{-i} and for s_i under σ'_{-i} can be approximated by $\sum_{\theta,m} \left[\mathbf{g}_i(a_i, \beta, \theta) \bullet \mathbf{q}^m\right] t_i^k\left(F_m^\theta\right)$ and $\sum_{\theta,m} \left[\mathbf{g}_i(a_i, \beta, \theta) \bullet \mathbf{q}^m\right] s_i^k\left[\left(F_m^\theta\right)^\varepsilon\right]$ respectively, for which the condition $\rho^k\left(t_i^k, s_i^k\right) \le \varepsilon$ is readily applied.

3.4 Step 4: the double-counting problem

The double-counting problem mentioned in step 3 arises because $\{(F_m^\theta)^\varepsilon\}_{\theta,m}$ may not partition $(\mathbf{T}_{-i}^{k-1})^\varepsilon$. Namely, for (θ, t_{-i}^{k-1}) in $(F_m^\theta)^\varepsilon \cap (F_{m'}^{\theta'})^\varepsilon$, should we define σ'_{-i} following the extension on $(F_m^\theta)^\varepsilon$ or the one on $(F_{m'}^{\theta'})^\varepsilon$? To solve this problem, it is important to recall Lemma 2 which allows us to fix $\beta \in \Delta(A_i \backslash \{a_i\})$ before finding the conjecture σ'_{-i} to prove (3). Then, for every (θ, t_{-i}^{k-1}) in $(F_m^\theta)^\varepsilon \cap (F_{m'}^{\theta'})^\varepsilon$, we can simply assign σ'_{-i} to follow the extension on $(F_m^\theta)^\varepsilon$ if $\mathbf{g}_i(a_i, \beta, \theta) \bullet \mathbf{q}^m \geq \mathbf{g}_i(a_i, \beta, \theta') \bullet \mathbf{q}^{m'}$. This will make the expected value of $[\mathbf{g}_i(a_i, \beta, \theta) \bullet \sigma'_{-i}(\theta, t_{-i})]$ as large as possible and help us to rationalize a_i.

Formally, we can relabel the elements in $\{F_m^\theta\}_{\theta,m}$ as follows. Let $\{F_m^\theta\}_{\theta,m} = \{F_{m_n}^{\theta_n}\}_{n=1}^N$ such that

$$\mathbf{g}_i(a_i, \beta, \theta_1) \bullet \mathbf{q}^{m_1} \geq \cdots \geq \mathbf{g}_i(a_i, \beta, \theta_N) \bullet \mathbf{q}^{m_N}.$$

For notational simplicity, we often write F_n instead of $F_{m_n}^{\theta_n}$ when no confusion may arise. By the way we order F_n, we should assign $\sigma'_{-i}(\theta, t_{-i}^{k-1})$ to follow the extension on $(F_n)^\varepsilon$ with n being the smallest number such that $(F_n)^\varepsilon$ contains (θ, t_{-i}^{k-1}). Formally, this amounts to modifying the sets $\{(F_n)^\varepsilon\}$ to the sets $\{E_n\}$ which are defined as follows, and define σ'_{-i} following the extension in step 3 on each E_n.

$$E_1 = (F_1)^\varepsilon \text{ and } E_n = (F_n)^\varepsilon \setminus \left(\cup_{l=1}^{n-1} E_l\right) \text{ for } n \geq 2.$$

With this modification, $\{E_n\}_{n=1}^N$ partitions $(\mathbf{T}_{-i}^{k-1})^\varepsilon$, and we can approximate the payoff of s_i under σ'_{-i} by $\sum_{n=1}^N [\mathbf{g}_i(a_i, \beta, \theta_n) \bullet \mathbf{q}^{m_n}]\, s_i^k[E_n]$ without double-counting.

3.5 Step 5: combining sets

The last step is to show that the difference of the two approximated payoffs is not large, i.e.,

$$\sum_{n=1}^N [\mathbf{g}_i(a_i, \beta, \theta_n) \bullet \mathbf{q}^{m_n}] \left[s_i^k(E_n) - t_i^k(F_n)\right] \geq -4M\varepsilon. \tag{4}$$

Let $A^n = \mathbf{g}_i(a_i, \beta, \theta_n) \bullet \mathbf{q}^{m_n}$ and $B^n = s_i^k(E_n) - t_i^k(F_n)$. For notational convenience, we also relabel $\{A^n\}_{n=1}^N$ and$\{B^n\}_{n=1}^N$ in a reverse order, i.e., let $C^n = A^{N-n+1}$ and $D^n = B^{N-n+1}$ for every $n = 1, ..., N$.

The following claim will be useful in showing (4) and is formally proved in Appendix A.2.2. It follows basically from our construction that $\cup_{n=1}^{l} E_n = \cup_{n=1}^{l} (F_n)^{\varepsilon}$ and the assumption $\rho^k \left(t_i^k, s_i^k\right) \leq \varepsilon$.

Claim 1 *We have* $\sum_{n=1}^{l} B^n \geq -\varepsilon$ *and* $\sum_{n=1}^{l} D^n \leq \varepsilon$ *for* $1 \leq l \leq N$.

We now sketch the idea for the proof of (4). First, recall from the previous step that

$$A^1 \geq A^2 \geq \cdots \geq A^N.$$

For heuristic purpose, assume that $N = 3$ and $A^3 \geq 0$. We prove that $\sum_{n=1}^{3} A^n B^n \geq -2M\varepsilon$ in two steps. If $\{B^n\}$ are all nonnegative, then $\sum_{n=1}^{3} A^n B^n \geq 0 > -2M\varepsilon$. If $\{B^n\}$ are all non-positive, then $\sum_{n=1}^{3} A^n B^n \geq 2M \sum_{n=1}^{3} B^n \geq -2M\varepsilon$ because of Claim 1 and our assumption that $|A^n| \leq 2M$ for all n.

The main problem is to deal with the situation in which B^1, B^2, and B^3 have different signs. For example, suppose $B^1 > 0$, $B^2 < 0$, and $B^3 < 0$. We use the following trick. Since $B^1 > 0$ and $A^1 \geq A^2$, we will not increase the value of $\sum_{n=1}^{3} A^n B^n$ if B^1 is "moved" from being multiplied by A^1 to being multiplied by A^2. That is, $\sum_{n=1}^{3} A^n B^n \geq A^2 (B^1 + B^2) + A^3 B^3$. Then, we check the sign of $B^1 + B^2$. If $(B^1 + B^2) \leq 0$, then

$$\sum_{n=1}^{3} A^n B^n \geq A^2 \left(B^1 + B^2\right) + A^3 B^3 \geq 2M \left(B^1 + B^2 + B^3\right) \geq -2M\varepsilon,$$

where second inequality follows because both $(B^1 + B^2)$ and B^3 are non-positive and $|A^n| \leq 2M$; the last inequality follows from Claim 1. If $(B^1 + B^2) > 0$, then the value of $A^2 (B^1 + B^2) + A^3 B^3$ decreases if $B^1 + B^2$ is further "moved" from being multiplied by A^2 to being multiplied by A^3, i.e. $A^2 (B^1 + B^2) + A^3 B^3 \geq A^3 (B^1 + B^2 + B^3)$. Hence,

$$\sum_{n=1}^{3} A^n B^n \geq A^2 \left(B^1 + B^2\right) + A^3 B^3 \geq A^3 \sum_{n=1}^{3} B^n \geq -2M\varepsilon,$$

where the last inequality follows from $|A^3| \leq 2M$ and Claim 1.

Our argument takes advantage of the property that A^n is decreasing in n so that "moving" a positive B toward being multiplied by a smaller A does not increase the value. After all these "moves" are done, only the last term may have a positive coefficient B. If the last term is negative, we go back to the single-signed case. If the last B is positive, since $\{A^n\}$ are all nonnegative, we can simply throw out $A^n B^n$ without increasing the value and also go back to the single-signed case.[8] The general proof involves dividing the summation $\sum_{n=1}^{N} A^n B^n$ into two groups. One group has all the nonnegative A^ns and the other group has all the negative A^ns. For both groups, we can invoke a similar trick and Claim **1** and conclude $\sum_{n=1}^{N} A^n B^n \geq -4M\varepsilon$ (see the proofs of Claims 2 and 3 in Appendix A.2 for details).

4 Non-denseness of finite types

In this section, we first show by an example that finite types are not dense under the uniform strategic topology. We achieve this goal by directly constructing an (infinite) type t_i^* such that $d^{us}(t_i^*, t_i) \geq \frac{M}{16}$ for any finite type t_i. Based upon the example, we go one step further to show that the set of finite types is nowhere dense in the universal type space, i.e., the complement of the uniform strategic closure of finite types is open and dense. Finally, we remark t_i^* is a critical type in the sense of Ely and Peski (2007) and comment on the implication of our example to the relationship between the strategic topology and the uniform-weak topology around critical types.

Throughout this section, consider the case that $\Theta = \{0, 1\}$. This simplification allows us to follow Morris (2002) to define the iterated expectations of a type $t = (\mu_1, \mu_2, ...) \in \mathcal{T}$ which will be used later. Let $\xi_0 : Y^0 \to [0,1]$ be defined as $\xi_0 = 1_{\{\theta=1\}}$ and

$$\widehat{\xi}_k(t) = \xi_k(\mu_k) \equiv \int_{Y^{k-1}} \xi_{k-1} d\mu_k, \; \forall k \geq 1.$$

Define $\widehat{\xi}(t) \equiv \left(\widehat{\xi}_k(t)\right)_{k=1}^{\infty} \in [0,1]^{\infty}$. Say a sequence of iterated expectations $x \in [0,1]^{\infty}$ *is generated by a type* t if $x = \widehat{\xi}(t)$. We know that every $x \in [0,1]^{\infty}$ can be generated by some type $t \in \mathcal{T}$ (see (Morris, 2002, Example)).

We are now ready to define the type t_i^*. Let t_i^* be the type which has the following

[8] Note that we cannot simply delete all $A^n B^n$ such that $B^n > 0$. Consider the special case we discuss here for example. While we have $\sum_{n=1}^{3} A^n B^n \geq A^2B^2 + A^3B^3$, we cannot apply Claim **1** to show that $A^2B^2 + A^3B^3 \geq -2M\varepsilon$. Recall that $B^2 + B^3 = s_i^k(E_2 \cup E_3) - t_i^k(F_2 \cup F_3)$. Since $(E_2 \cup E_3)$ may not be equal to $(F_2 \cup F_3)^{\varepsilon}$, we cannot get $B^2 + B^3 \geq -\varepsilon$ from $\rho^k\left(t_i^k, s_i^k\right) \leq \varepsilon$.

iterated expectations

$$\left(\widehat{\xi}_{\frac{n(n-1)}{2}+1}(t_i^*), ..., \widehat{\xi}_{\frac{n(n+1)}{2}}(t_i^*)\right) = \begin{cases} (1,1,....,1), & \text{if } n \text{ is odd;} \\ (0,0,....,0), & \text{if } n \text{ is even.} \end{cases}$$

That is, $\widehat{\xi}(t_i^*) = (1,0,0,1,1,1,0,0,0,0,...)$. For notational convenience, denote

$$\begin{aligned} a^0 &= (\frac{1}{2}, \frac{1}{2}, ..., \frac{1}{2}, ...); \\ a^n &= (\widehat{\xi}_n(t_i^*), \widehat{\xi}_{n+1}(t_i^*), \widehat{\xi}_{n+2}(t_i^*), ...) \text{ for } n = 1, 2, ... \end{aligned}$$

Note that $a^n \neq a^m$ for $n \neq m$. For $k = 1, 2, ...,$ let $(a^n)_k$ be the k^{th} element of the sequence a^n. One feature of a^n, which will be useful later in the proof, is that $(a^n)_k = (a^{n+k-1})_1$.

Let $\widehat{T}_i \times \widehat{T}_{-i} \subset \mathcal{T}_i \times \mathcal{T}_{-i}$ be the smallest (w.r.t. set-inclusion) belief-closed subset such that $t_i^* \in \widehat{T}_i$. Observe that 0 (and 1) is the minimum (and maximum) that a k^{th}-order expectation can achieve. Since $\widehat{\xi}(t_i^*) = (1,0,0,1,1,1,0,0,0,0,...)$, there is a unique type t_i^* in $\mathcal{T}_i$ which has the following hierarchy of beliefs:

- first-order belief: point mass on $\theta = 1$;
- second-order belief: point mass on [player $-i$ believes $\theta = 0$ with probability 1];
- third-order belief: point mass on [player $-i$ believes with probability 1 that player i believes with probability 1 that $\theta = 0$], and so on.

Let $t(1) = t_i^*$. Moreover, let $t(n)$ denote the unique type in $\mathcal{T}$ that generates the iterated expectations a^n, for $n \geq 2$. Hence,

$$\pi^*[t(n)]\left[\widehat{\xi}_1(t(n)), t(n+1)\right] = 1 \text{ for } n \geq 1. \tag{5}$$

That is, type t_i^* believes his opponent is $t(2)$ with probability 1; type $t(n)$ believes his opponent is $t(n+1)$ with probability 1 (cf. (Morris, 2002, Example) and Mertens and Zamir (1985)). Therefore, $\widehat{\xi}\left(\widehat{T}_i \cup \widehat{T}_{-i}\right) = \{a^1, a^2, ...\}$.

We now show that finite types are not dense under the uniform strategic topology. We provide an outline of our argument here, and the rigorous proof can be found in Appendix A.3.1. It is helpful to consider first the following modified version of the higher-order expectation (HOE) game due to Morris (2002).[9]

$$A_i = \{a^0, a^1, a^2, a^3, ...\};$$

$$g_i\left[(a_{i,n})_{n=1}^{\infty},\ (a_{-i,n})_{n=1}^{\infty},\ \theta\right] = \begin{cases} 0, & \text{if } (a_{i,n})_{n=1}^{\infty} = a^0; \\ M \times \inf \left\{ \begin{array}{c} -(a_{i,1}-\theta)^2, -(a_{i,2}-a_{-i,1})^2, \\ ..., -(a_{i,n}-a_{-i,n-1})^2, ... \end{array} \right\}, & \text{if } (a_{i,n})_{n=1}^{\infty} \neq a^0. \end{cases}$$

While $G = (A_i, g_i)_{i=1,2}$ is a game with infinitely many actions, we will show how it can be modified to a finite game in Appendix A.3.1. In this game, a player's actions are a^0, a^1, a^2,.... The action a^0 always generates the maximal payoff 0.[10] If the player chooses a^n with $n \geq 1$, he gets the infimum over the quadratic losses between the first coordinate of player i's action and θ, the quadratic loss between the second coordinate of player i's action and player $-i$'s first coordinate, and so on. That is, it is a coordination game in which a player tries to match the state of nature and his opponent's action. The players can achieve the maximal payoff 0 by either taking the safe action a^0 or having perfect coordination with nature as well as every coordinate of his opponent's action.

First, we have $a^1 \in R_i(t_i^*, 0)$ because for each k each player of type $t(n) \in \widehat{T}_i$ can rationalize a^n by holding the belief that $t(n+1)$ will choose the action a^{n+1}. Second, we show that for some positive but small enough γ, $a^1 \notin R_i(t_i, \gamma)$ for *every* finite type t_i. Let $T_i \times T_{-i} \subset \mathcal{T}_i \times \mathcal{T}_{-i}$ be the smallest belief-closed subset such that $t_i \in T_i$. Suppose instead that a^1 is γ–rationalizable for t_i. Then, (A) player i must believe most of his opponent's types (which are in the support of $\pi_i^*[t_i]$) are playing a^2 so as to get almost perfect coordination;

[9] One may wonder if we can use the original HOE game, in which the payoff is defined as,

$$g_i\left[(a_{i,n})_{n=1}^{\infty},\ (a_{j,n})_{n=1}^{\infty},\ \theta\right] = M \times \left[-\lambda_1 (a_{i,1}-\theta)^2 - \sum_{k=2}^{\infty} \lambda_k (a_{i,k}-a_{j,k-1})^2\right],$$

$$\text{such that } \lambda_k > 0 \text{ and } \sum_{k=1}^{\infty} \lambda_k = 1.$$

Consider two types t and t' such that their expectations differ only at the N^{th}-order. On the one hand, we should choose a λ_N large enough so that the game separates t and t' in the sense that the minimal γ to γ–rationalize some action under these types differs. On the other hand, the player chooses $a_{i,N} \approx \widehat{\xi}_N[t']$ only if both players almost truthfully report all their k^{th}-order expectation for every k up to N, which requires those λ_k's to be large enough. Since $\sum_{k=1}^{\infty} \lambda_k = 1$, it does not seem obvious to us how to resolve this tension for a large N.

[10] We add this safe action a^0 to make it easier to rule out certain actions as not being γ–rationalizable.

and (B) a^2 must be $\gamma-$rationalizable for these opponents. Pick one t_{-i} which is believed to play a^2 by t_i. Since a^2 is $\gamma-$rationalizable for t_{-i}, we can similarly get (A) player $-i$ must believe most of player i's types in the support of $\pi^*_{-i}[t_{-i}]$ are playing a^3; and (B) a^3 must be $\gamma-$rationalizable for these player i's types. Doing this inductively, we can get an infinite chain of $\gamma-$rationalization.

$$t_i = \widehat{t}(1) \to \widehat{t}(2) \to \widehat{t}(3) \to \widehat{t}(4) \to \widehat{t}(5) \to \cdots \text{such that} \tag{6}$$
$$\widehat{t}(k) \in T_i \text{ if } k \text{ is odd and } \widehat{t}(k) \in T_{-i} \text{ if } k \text{ is even;}$$
$$\widehat{t}(k+1) \text{ is in the support of } \pi^*\left[\widehat{t}(k)\right]; \text{ and } a^k \text{ is } \gamma-\text{rationalizable for } \widehat{t}(k).$$

Since t_i is a finite type, some type in T_i must recur in this infinite chain. That is, we can find $\widehat{t}(n) = \widehat{t}(m) = \widetilde{t}(1) \in T_i$ such that a^n and a^m are both $\gamma-$rationalizable for $\widetilde{t}(1)$ and $a^n \neq a^m$.

Recall that a player can always achieve the maximal payoff by taking the safe action a^0. Since γ is small enough, in order to make a^n $\gamma-$rationalizable for $\widetilde{t}(1)$, a^{n+1} must be $\gamma-$rationalizable for most of $\widetilde{t}(1)$'s opponent's types in the support of $\pi^*\left[\widetilde{t}(1)\right]$; to make a^m $\gamma-$rationalizable for $\widetilde{t}(1)$, a^{m+1} must be $\gamma-$rationalizable for most of $\widetilde{t}(1)$'s opponent's types in the support of $\pi^*\left[\widetilde{t}(1)\right]$. Hence, there exists some type $\widetilde{t}(2)$ in the support of $\pi^*\left[\widetilde{t}(1)\right]$ such that a^{n+1} and a^{m+1} are both $\gamma-$rationalizable for $\widetilde{t}(2)$. Similarly, we can pick $\widetilde{t}(k)$ in this way for all $k \geq 2$. Hence, we can construct another common chain of $\gamma-$rationalization for both a^n and a^m:

$$\widetilde{t}(1) \to \widetilde{t}(2) \to \widetilde{t}(3) \to \widetilde{t}(4) \to \widetilde{t}(5) \to \cdots \text{such that}$$
$$\widetilde{t}(k) \in T_i \text{ and } \left\{a^{n+k-1}, a^{m+k-1}\right\} \subset R_i\left(\widetilde{t}(k), \gamma\right) \text{ if } k \text{ is odd;}$$
$$\widetilde{t}(k) \in T_{-i} \text{ and } \left\{a^{n+k-1}, a^{m+k-1}\right\} \subset R_{-i}\left(\widetilde{t}(k), \gamma\right) \text{ if } k \text{ is even.}$$

Since the players are also trying to coordinate with nature in this game, we also have the property (C) for a small enough γ and any action a $\gamma-$rationalizable for a type t, $(a)_1 = 0$ iff $\widehat{\xi}_1(t) \in [0, \frac{1}{2}]$. We then reach a contradiction by applying property (C) to this common chain of rationalization for a^n and a^m. First, $a^n \neq a^m$ implies $(a^n)_{k^*} = 1 \neq 0 = (a^m)_{k^*}$ for some k^*. Second, by the definitions of a^n and a^m, $(a^{n+k^*-1})_1 = (a^n)_{k^*} = 1 \neq 0 = (a^m)_{k^*} = (a^{m+k^*-1})_1$, but property (C) implies,

$$
\begin{aligned}
(a^{n+k^*-1})_1 &= (a^{m+k^*-1})_1 = 0 \text{ if } \widehat{\xi}_1\left[\widetilde{t}(k^*)\right] \in [0, \frac{1}{2}]; \\
(a^{n+k^*-1})_1 &= (a^{m+k^*-1})_1 = 1 \text{ if } \widehat{\xi}_1\left[\widetilde{t}(k^*)\right] \notin [0, \frac{1}{2}],
\end{aligned}
$$

because both a^{n+k^*-1} and a^{m+k^*-1} are $\gamma-$rationalizable for $\widetilde{t}(k^*)$ in the common chain above. This is a contradiction. Therefore, $a^1 \notin R_i(t_i, \gamma)$.

In the rigorous proof for Proposition 2 in Appendix A.3.1, we formalize the argument above. We make the game finite by truncating a^n into its first N coordinates and proceed in four steps. In step 1, we show that if N is sufficiently large, we can still find a^n and a^m in the $\gamma-$rationalization chain in (6), such that the N-truncation of a^n and a^m are distinct. Step 2 proves properties (A) and (C). Step 3 constructs the common chain of $\gamma-$rationalization for a^n and a^m as above. Step 4 derives the contradiction. In the proof, $\gamma = \frac{M}{16}$ is small enough to achieve our goal.

Proposition 2 *$d^{us}(t_i^*, t_i) \geq \frac{M}{16}$ for any finite type t_i. Hence, finite types are not dense under d^{us}.*

In fact, Proposition 2 can be used to prove a stronger result.

Theorem 2 *Finite types are nowhere dense under d^{us}.*

Since the uniform-weak topology is finer than the uniform strategic topology by Theorem 1, finite types are nowhere dense under d^{uw} as well. The proofs of Proposition 2 and Theorem 2 are relegated to Appendix A.3.1 and A.3.2, respectively.

Recall that Ely and Peski (2007) define a critical type to be a type around which the strategic topology is strictly stronger than the product topology. Recall also that Di Tillio and Faingold (2007) show that the strategic topology is equivalent to the uniform-weak topology around finite types. As shown in Ely and Peski (2007), every finite type is critical but not conversely. In particular, the type t_i^* we construct is an infinite critical type. To see this, recall that $\widehat{T}_i \times \widehat{T}_{-i} \subset \mathcal{T}_i \times \mathcal{T}_j$ is the smallest belief-closed set such that $t_i^* \in \widehat{T}_i$. Clearly,

$\widehat{T}_i \times \widehat{T}_{-i}$ is of common 1-belief under t^*. By monotonicity of the common 1-belief operator, the closure of $\widehat{T}_i \times \widehat{T}_{-i}$ (under the product topology in $\mathcal{T}_i \times \mathcal{T}_j$) is of common 1-belief at t_i^*. Consider the types $(\bar{t}_1, \bar{t}_2) \in \mathcal{T}_i \times \mathcal{T}_j$ under which both $\bar{t}_1$ and $\bar{t}_2$ have the first-order belief $\Pr(\theta = 1) = \Pr(\theta = 0) = \frac{1}{2}$. For all $(t_1, t_2) \in \widehat{T}_i \times \widehat{T}_{-i}$, the first-order beliefs of t_1 and t_2 are either $\Pr(\theta = 1) = 1$ or $\Pr(\theta = 0) = 1$. Thus, $(\bar{t}_1, \bar{t}_2)$ does not belong to the closure of $\widehat{T}_i \times \widehat{T}_{-i}$. Hence, the closure of $\widehat{T}_i \times \widehat{T}_{-i}$ is a proper closed subset of $\mathcal{T}_i \times \mathcal{T}_j$. By (Ely and Peski, 2007, Theorem 3), we conclude t^* is a critical type. Since finite types are dense under the strategic topology by (Dekel, Fudenberg, and Morris, 2006, Theorem 3), Proposition 2 shows that strategic convergence to an infinite critical type does not imply uniform-weak convergence to this critical type.

5 Discussion

5.1 The uniform strategic topology

DFM study the uniform strategic topology in contrast to the strategic topology they propose. The denseness result from Section 4 demonstrates one difference between the two topologies. In particular, our example shows that it is sometimes hard to approximate complicated types with finite types when we require uniformity of this approximation among all finite games. However, such a uniform approximation may still be relevant.

Suppose we are facing a mechanism design problem where agents' information is modeled with a complicated type space T. DFM's denseness results on strategic topology states that *given* any fixed game G, we can find a simple type space T' to approximate T in terms of strategic behaviors in G. However, to solve a mechanism design problem is to search for a mechanism (game form) among all possible games. Hence, to ensure that the optimal solution on T' incurs approximately no loss of accuracy, it is crucial that the strategic behaviors under T' approximate those under T in *all* mechanisms instead of merely in some G. Thus, the uniform strategic topology is free of such a problem so long as the mechanism is searched within bounded finite games.[11]

[11] Suppose that G' is the optimal mechanism for the simpler type space T', which is close to the true type

5.2 Comparison of our proof to that of Di Tillio and Faingold (2007)

Di Tillio and Faingold (2007) prove that the uniform-weak topology is finer than the strategic topology around finite types. Our first main result extends this to all types. In Di Tillio and Faingold's proof, they exploit the fact that in a finite type space, there is a minimal uniform-weak distance between any two different types within this type space. Thus, when $d^{uw}(t_i^n, t_i)$ is sufficiently small relative to the minimum uniform-weak distance of the finite type space containing t_i, the two types t_i^n and t_i must believe in approximately the same set of opponents with approximately the same distribution, and no double-counting problem is involved. In our proof, we take advantage of the finiteness of games and solve the double-counting problem by applying the minimax argument.

5.3 Infinite order implications of Theorem 1

It is natural to doubt whether one can obtain information about the entire hierarchy of beliefs of types. Therefore, types which are close under the product topology may still be deemed indistinguishable from a practical viewpoint. In accordance with this idea, extensive studies have been carried out on finite-order implications of notions associated with a type. In particular, Lipman (2003) shows that common-prior types are dense under the product topology, while Weinstein and Yildiz (2007) show that in a fixed game without common-knowledge restrictions on payoffs, types with unique rationalizable actions are open and dense.

As pointed out by DFM, neither the result in Lipman (2003) nor the result in Weinstein and Yildiz (2007) holds when we consider the strategic topology instead of the product

space T in the strategic topology. The behaviors of T and T' in G' might still be quite different. This is especially true if G' has a lot of actions. Recall that the strategic distance $d^s(t_i, s_i)$ for types t_i and s_i in $\mathcal{T}_i$ is a weighted sum of the difference between the behaviors of t_i and s_i in all bounded finite game. If G' has m actions with m being large, $d^s(t_i, s_i)$ assigns a small weight α^m on the difference between the behaviors of t_i and s_i in G'. Therefore, even if $d^s(t_i, s_i)$ is small, the two types t_i and s_i may still exhibit quite different behaviors in G'. However, this problem can be avoided if we approximate T by T'' which is close to T in the uniform strategic topology.

topology. DFM show that there is an open set in the strategic topology which consists entirely of types with noncommon priors or multiple rationalizable actions. That is, in these cases strategic open sets are rich enough to separate noncommon-prior types from common-prior types, or types with multiple rationalizable actions from types with unique rationalizable actions. A straightforward consequence of Theorem 1 is that these strategic open sets identified in DFM's examples must contain uniform-weak open sets. Hence, neither Lipman's result nor that of Weinstein and Yildiz's holds in the uniform-weak topology.

5.4 The distance d^{**}

The uniform-weak topology is a natural way of strengthening the product topology. In DFM's Proposition 2, they propose a metric d^{**} which also implies uniform strategic convergence. Two types are close under d^{**} if they have uniformly close expectations on all bounded functions measurable with respect to the k^{th}-order beliefs for some k. However, this metric d^{**} is too strong in the sense that even when we restrict attention to finite-order beliefs, the topology it generates is still strictly stronger than the standard weak* topology.

In particular, consider an example in Chen and Xiong (2008). Suppose that $\Theta = \{0, 1\}$. Let t be a complete information type under which it is common 1-belief that "$\theta = 1$." Let $\{t^n\}$ be a sequence of types under which both players believe "$\theta = 1$" with probability $\left(1 - \frac{1}{n}\right)$ and it is common 1-belief that both players believe "$\theta = 1$" with probability $\left(1 - \frac{1}{n}\right)$ (cf. Monderer and Samet (1989)). Then, it is common $\left(1 - \frac{1}{n}\right)$-belief that "$\theta = 1$" under t^n, and moreover, $d^{uw}(t^n, t) \to 0$. By Theorem 1, we have $d^{us}(t^n, t) \to 0$. However, as shown in Chen and Xiong (2008), t^n does not converge to t under d^{**}. In contrast to the result of Di Tillio and Faingold (2007), this example also demonstrates that even if the limit type is a finite type, the (uniform) strategic convergence does not imply the $d^{**}-$convergence.

A Appendix

A.1 Alternative characterizations of the $\gamma-$ICR set

A.1.1 Proof of Lemma 1

Recall that $\overline{R}_i^0(t_i, G, \gamma) = R_i^0(t_i, G, \gamma) = A_i$ and for $k \geq 1$, $a_i \in R_i^k(t_i, G, \gamma)$ iff there exists a measurable function $\sigma_{-i} : \Theta \times \mathcal{T}_{-i} \to \Delta(A_{-i})$ such that

$$\text{supp}\sigma_{-i}(\theta, t_{-i}) \subseteq R_{-i}^{k-1}(t_{-i}, G, \gamma) \text{ for } \pi_i^*(t_i) - \text{almost surely } (\theta, t_{-i}); \tag{7}$$

$$\int_{\Theta \times \mathcal{T}_{-i}} [\mathbf{g}_i(a_i, a_i', \theta) \bullet \sigma_{-i}(\theta, t_{-i})] \, \pi_i^*(t_i)[(\theta, dt_{-i})] \geq -\gamma \text{ for all } a_i' \in A_i, \tag{8}$$

and $a_i \in \overline{R}_i^k(t_i, G, \gamma)$ iff there exists a measurable function $\sigma_{-i} : \Theta \times \mathcal{T}_{-i}^{k-1} \to \Delta(A_{-i})$ such that

$$\text{supp}\sigma_{-i}\left(\theta, t_{-i}^{k-1}\right) \subseteq \overline{R}_{-i}^{k-1}(t_{-i}, G, \gamma) \text{ for } t_i^k - \text{almost surely } \left(\theta, t_{-i}^{k-1}\right); \tag{9}$$

$$\int_{\Theta \times \mathcal{T}_{-i}^{k-1}} \left[\mathbf{g}_i(a_i, a_i', \theta) \bullet \sigma_{-i}\left(\theta, t_{-i}^{k-1}\right)\right] t_i^k \left[\left(\theta, dt_{-i}^{k-1}\right)\right] \geq -\gamma \text{ for all } a_i' \in A_i \tag{10}$$

where $\Theta \times \mathcal{T}_{-i}^0 \equiv \Theta$.

Lemma 1 $\overline{R}_i^k(t_i, G, \gamma) = R_i^k(t_i, G, \gamma)$ for every integer $k \geq 0$, every t_i, and every player i.

Proof. We prove this claim by induction on k. For $k = 0$, the lemma holds by definition. Now suppose that the lemma holds for some nonnegative integer $k - 1$.

($\overline{R}_i^k(t_i, G, \gamma) \subseteq R_i^k(t_i, G, \gamma)$) Suppose $a_i \in \overline{R}_i^k(t_i, G, \gamma)$. Then, there exists a measurable function $\sigma_{-i} : \Theta \times \mathcal{T}_{-i}^{k-1} \to \Delta(A_{-i})$ such that (9) and (10) hold. Now consider $\sigma_{-i}^* : \Theta \times \mathcal{T}_{-i} \to \Delta(A_{-i})$ such that $\sigma_{-i}^*(\theta, t_{-i}) \equiv \sigma_{-i}\left(\theta, t_{-i}^{k-1}\right)$ for all $(\theta, t_{-i}) \in \Theta \times \mathcal{T}_{-i}$. Note that σ_{-i}^* is measurable because σ_{-i} is measurable and $\Theta \times \mathcal{T}_{-i}$ is a second countable space endowed with the Borel σ-algebra (see (Aliprantis and Border, 1999, 4.43 Theorem)). First, (7) follows because the marginal distribution of $\pi_i^*(t_i)$ on $\Theta \times \mathcal{T}_{-i}^{k-1}$ agrees with t_i^k and both

(9) and the induction hypothesis hold. Second, for all $a_i' \in A_i$,

$$\begin{aligned}
&\int_{\Theta\times\mathcal{T}_{-i}} \left[\mathbf{g}_i\left(a_i, a_i', \theta\right)\bullet\sigma_{-i}^*\left(\theta, t_{-i}\right)\right]\pi_i^*\left(t_i\right)\left[\left(\theta, dt_{-i}\right)\right] \\
=& \int_{\Theta\times\mathcal{T}_{-i}^{k-1}} \left[\mathbf{g}_i\left(a_i, a_i', \theta\right)\bullet\sigma_{-i}\left(\theta, t_{-i}^{k-1}\right)\right]\pi_i^*\left(t_i\right)\left[\left(\theta, dt_{-i}^{k-1}\right)\right] \\
=& \int_{\Theta\times\mathcal{T}_{-i}^{k-1}} \left[\mathbf{g}_i\left(a_i, a_i', \theta\right)\bullet\sigma_{-i}\left(\theta, t_{-i}^{k-1}\right)\right]t_i^k\left[\left(\theta, dt_{-i}^{k-1}\right)\right] \geq -\gamma
\end{aligned}$$

where the first equality is due to the definition of σ_{-i}^* and the second is again because the marginal distribution of $\pi_i^*\left(t_i\right)$ on $\Theta\times\mathcal{T}_{-i}^{k-1}$ agrees with t_i^k. Therefore, (8) holds and hence $a_i \in R_i^k\left(t_i, G, \gamma\right)$.

($\overline{R}_i^k\left(t_i, G, \gamma\right) \supseteq R_i^k\left(t_i, G, \gamma\right)$) Suppose $a_i \in R_i^k\left(t_i, G, \gamma\right)$. Hence, there exists measurable function $\sigma_{-i} : \Theta\times\mathcal{T}_{-i} \to \Delta\left(A_{-i}\right)$ such that (7) and (8) hold. Since $\Theta\times\mathcal{T}_{-i}$ is a compact metric space, it is a standard Borel space. Hence, there is a regular conditional distribution of $\pi_i^*\left(t_i\right)$ on $\Theta\times\mathcal{T}_{-i}^{k-1}$ (see (Dudley, 2002, 10.2.2. Theorem)). Define $\sigma_{-i}^* : \Theta\times\mathcal{T}_{-i}^{k-1} \to \Delta\left(A_{-i}\right)$ as

$$\sigma_{-i}^*\left(\theta, t_{-i}^{k-1}\right) \equiv \int_{\Theta\times\mathcal{T}_{-i}} \sigma_{-i}\left(\widetilde{\theta}, s_{-i}\right)\pi_i^*\left(t_i\right)\left[\left(\widetilde{\theta}, ds_{-i}\right)|\widetilde{\theta}=\theta, s_{-i}^{k-1}=t_{-i}^{k-1}\right], \forall\left(\theta, t_{-i}^{k-1}\right)\in\Theta\times\mathcal{T}_{-i}^{k-1}.$$

Then, since σ_{-i} is a measurable function from $\Theta\times\mathcal{T}_{-i}$ to $\Re^{|A_{-i}|}$, by (Dudley, 2002, 10.2.5. Theorem), σ_{-i}^* is a version of the conditional expectation of σ_{-i} conditional on $\left(\theta, t_{-i}^{k-1}\right)$. Hence, σ_{-i}^* is measurable. Again, (9) follows because the marginal distribution of $\pi_i^*\left(t_i\right)$ on $\Theta\times\mathcal{T}_{-i}^{k-1}$ agrees with t_i^k and both (7) and the induction hypothesis hold. Moreover, for all $a_i' \in A_i$,

$$\begin{aligned}
&\int_{\Theta\times\mathcal{T}_{-i}^{k-1}} \left[\mathbf{g}_i\left(a_i, a_i', \theta\right)\bullet\sigma_{-i}^*\left(\theta, t_{-i}^{k-1}\right)\right]t_i^k\left[\left(\theta, dt_{-i}^{k-1}\right)\right] \\
=& \int_{\Theta\times\mathcal{T}_{-i}^{k-1}} \left[\int_{\Theta\times\mathcal{T}_{-i}} \left[\mathbf{g}_i\left(a_i, a_i', \theta\right)\bullet\sigma_{-i}\left(\theta, s_{-i}\right)\right]\pi_i^*\left(t_i\right)\left[\left(\theta, ds_{-i}\right)|\widetilde{\theta}=\theta, s_{-i}^{k-1}=t_{-i}^{k-1}\right]\right]\pi_i^*\left(t_i\right)\left[\left(\theta, dt_{-i}^{k-1}\right)\right] \\
=& \int_{\Theta\times\mathcal{T}_{-i}} \left[\mathbf{g}_i\left(a_i, a_i', \theta\right)\bullet\sigma_{-i}\left(\theta, t_{-i}\right)\right]\pi_i^*\left(t_i\right)\left[\left(\theta, dt_{-i}\right)\right] \geq -\gamma
\end{aligned}$$

where the first equality is because the marginal distribution of $\pi_i^*\left(t_i\right)$ on $\Theta\times\mathcal{T}_{-i}^{k-1}$ agrees with t_i^k and is also because of the definition of σ_{-i}^*, and the second equality follows from the law of iterated expectation (see (Dudley, 2002, 10.2.1. Theorem)). Therefore, (8) holds and hence $a_i \in \overline{R}_i^k\left(t_i, G, \gamma\right)$.

A.1.2 The proof of Lemma 2

Lemma 2 *For any positive integer k, any $\gamma \geq 0$, any finite game G, and any type $t_i \in \mathcal{T}_i$, $a_i \in R_i^k(t_i, G, \gamma)$ if and only if for every $\eta > 0$ and $\beta \in \Delta(A_i \backslash \{a_i\})$ there is a valid conjecture $\sigma_{-i}[\beta] : \Theta \times \mathcal{T}_{-i}^{k-1} \to \Delta(A_{-i})$ for t_i under which*

$$\int_{\Theta \times \mathcal{T}_{-i}^{k-1}} \left[\mathbf{g}_i(a_i, \beta, \theta) \bullet \sigma_{-i}[\beta]\left(\theta, t_{-i}^{k-1}\right)\right] t_i^k\left[\left(\theta, dt_{-i}^{k-1}\right)\right] \geq -\gamma - \eta. \tag{11}$$

Proof. Recall first a Minimax Theorem due to Fan (1952). Let f be a real-valued function defined on a product space $X \times Y$. Say f is convex-like on X if for every x, $x' \in X$ and $c \in [0,1]$, there is some $x'' \in X$ such that $f(x'', y) \leq cf(x, y) + (1-c) f(x', y)$ for all $y \in Y$. Say f is concave-like on Y if for every y, $y' \in Y$ and $c \in [0,1]$, there is some $y'' \in Y$ such that $f(x, y'') \geq cf(x, y) + (1-c) f(x, y')$ for all $x \in X$.

Fan's Minimax Theorem *Let X be a compact Hausdorff space and Y an arbitrary set (not topologized). Let f be a real-valued function on $X \times Y$ such that for every $y \in Y$, $f(\cdot, y)$ is lower semi-continuous on X. If f is convex-like on X and concave-like on Y, then*

$$\min_{x \in X} \sup_{y \in Y} f(x, y) = \sup_{y \in Y} \min_{x \in X} f(x, y).$$

To apply this result, define

$$\begin{aligned}
X &= \Delta(A_i \backslash \{a_i\}); \\
Y &= \left\{\sigma_{-i} : \Theta \times \mathcal{T}_{-i}^{k-1} \to \Delta(A_{-i}) : \sigma_{-i} \text{ is a valid conjecture}\right\}; \\
f(\beta, \sigma_{-i}) &= \int_{\Theta \times \mathcal{T}_{-i}^{k-1}} \left[\mathbf{g}_i(a_i, \beta, \theta) \bullet \sigma_{-i}(\theta, t_{-i})\right] t_i^k\left[\left(\theta, dt_{-i}^{k-1}\right)\right], \ \forall \beta \in X, \sigma_{-i} \in Y.
\end{aligned}$$

Obviously, f is convex-like on X, concave-like on Y, and $f(\cdot, \sigma_{-i})$ is lower semi-continuous on X for every σ_{-i}. Hence, by Fan's Minimax Theorem, we have $\min_{\beta \in X} \sup_{\sigma_{-i} \in Y} f(x, y) = \sup_{\sigma_{-i} \in Y} \min_{\beta \in X} f(x, y)$. First, $\sup_{\sigma_{-i} \in Y} \min_{\beta \in X} f(x, y) \geq -\gamma$ if and only if for any $\eta > 0$, there is a conjecture $\sigma_{-i} \in Y$ such that $f(\beta, \sigma_{-i}) \geq -\gamma - \eta$ for all $\beta \in X$. By Lemma 1 and DFM's Lemma 1, $\sup_{\sigma_{-i} \in Y} \min_{\beta \in X} f(x, y) \geq -\gamma$ is therefore equivalent to $a_i \in R_i^k(t_i, G, \gamma)$. Similarly, $\min_{\beta \in X} \sup_{\sigma_{-i} \in Y} f(x, y) \geq -\gamma$ is equivalent to for every $\eta > 0$ and $\beta \in X$, there is some valid conjecture $\sigma_{-i}[\beta] \in Y$ such that (11) hold.■

A.2 The proof of Proposition 1

A.2.1 The proof of Lemma 3

Lemma 3 *Consider a separable metric space* (Y, d_Y)*, a Borel set* $F \subseteq Y$*, and* $\varepsilon > 0$*. Suppose* $f : F \to Z$ *is a measurable function where* Z *is a measurable space. Then, there is a measurable function* $f^\varepsilon : F^\varepsilon \to Z$ *such that* $f^\varepsilon = f$ *on* F*, and for every* $y \in F^\varepsilon \backslash F$*,* $f^\varepsilon(y) = f(y')$ *for some* $y' \in F$ *with* $d_Y(y, y') < \varepsilon$*.*

Proof. Since F is a subset of a separable metric space, it is also separable under the relative topology (see (Dudley, 2002, p.32)). Let $\{y_1, y_2, ...\}$ be a countable dense subset of F. For any $y \in Y$, let $B(y, \varepsilon)$ denote the ε-open ball around y. First, we claim that $F^\varepsilon = \cup_{m=1}^\infty B(y_m, \varepsilon)$. Clearly, $F^\varepsilon \supseteq \cup_{m=1}^\infty B(y_m, \varepsilon)$. To see $F^\varepsilon \subseteq \cup_{m=1}^\infty B(y_m, \varepsilon)$, suppose $y \in F^\varepsilon$. Then, there is some $y' \in F$ such that $d_Y(y, y') < \varepsilon$. Hence, $B(y, \varepsilon) \cap F \neq \varnothing$. Since $B(y, \varepsilon) \cap F$ is relatively open in F and $\{y_1, y_2, ...\}$ is dense in F, there is m such that $y_m \in B(y, \varepsilon) \cap F$. Hence, $d_Y(y_m, y) < \varepsilon$ and therefore $y \in B(y_m, \varepsilon)$.

We modify the sets in $\{B(y_m, \varepsilon)\}_{m=1}^\infty$ to $\{\overline{B}_m\}_{m=1}^\infty$ such that

$$\overline{B}_1 = B(y_1, \varepsilon) \text{ and } \overline{B}_m = B(y_m, \varepsilon) \setminus \left[\cup_{l=1}^{m-1} \overline{B}_l\right] \text{ for } m \geq 2.$$

Observe that $\{\overline{B}_m\}_{m=1}^\infty$ partitions F^ε. Moreover, for any $m \in \mathbb{Z}_+$, since the sets in $\{B(y_m, \varepsilon)\}_{m=1}^\infty$ are open and hence measurable, $\overline{B}_m$ is also measurable. We now define f^ε as

$$f^\varepsilon(y) = \begin{cases} f(y), & \text{if } y \in F, \\ f(y_{m^*}) & \text{if } y \notin F \text{ and } y \in \overline{B}_{m^*}. \end{cases}$$

Then, by definition f^ε satisfies the property that $f^\varepsilon = f$ on F, and for every $y \in F^\varepsilon \backslash F$, $f^\varepsilon(y) = f(y')$ for some $y' \in F$ with $d_Y(y, y') < \varepsilon$.

Finally, we show that f^ε is measurable. Let $Z' \subseteq Z$ be an arbitrary measurable set. Then,

$$\begin{aligned} (f^\varepsilon)^{-1}(Z') &= f^{-1}(Z') \cup \bigcup_{\{m \in \mathbb{Z}_+ : f(y_m) \in Z'\}} \left[\overline{B}_m \backslash F\right], \\ \text{where } \overline{B}_m \backslash F &= \{y \in Y : y \in \overline{B}_m \text{ and } y \notin F\}. \end{aligned}$$

Since f is a measurable function and Z' is measurable, $f^{-1}(Z')$ is also measurable. For each m, the set $\left[\overline{B}_m \backslash F\right]$ is measurable. Hence, the set $\cup_{\{m \in Z_+ : f(y_m) \in Z'\}} \left[\overline{B}_m \backslash F\right]$ is also measurable, because it is a countable union of measurable sets. Therefore, $(f^{\varepsilon})^{-1}(Z')$ is measurable, and f^{ε} is a measurable function.■

A.2.2 The proof of Proposition 1

Proposition 1 *For any finite game G, any $\varepsilon, \gamma \geq 0$, and any types t_i and s_i in $\mathcal{T}_i$ with $\rho^k\left(t_i^k, s_i^k\right) \leq \varepsilon$, we have $R_i^k(t_i, G, \gamma) \subseteq R_i^k(s_i, G, \gamma + 6M\varepsilon)$ for every integer $k \geq 0$.*

Proof. Let G be a finite game. Since G is fixed, we will drop hereafter the explicit reference of G from our notation for expositional ease. We prove the proposition by induction. Suppose $a_i \in R_i^k(t_i, \gamma)$. By Lemma 1, there is a valid conjecture $\sigma_{-i} : \Theta \times \mathcal{T}_{-i}^{k-1} \to \Delta(A_{-i})$ for t_i such that

$$\int_{\Theta \times \mathcal{T}_{-i}^{k-1}} \left[\mathbf{g}_i(a_i, a_i', \theta) \bullet \sigma_{-i}\left(\theta, t_{-i}^{k-1}\right)\right] t_i^k\left[\left(\theta, dt_{-i}^{k-1}\right)\right] \geq -\gamma \text{ for all } a_i' \in A_i \backslash \{a_i\}. \qquad (12)$$

Our goal is to show that $a_i \in R_i^k(s_i, \gamma + 6M\varepsilon)$.

Consider first the case of $k = 1$. By Definition 1, $\Theta \times \mathcal{T}_{-i}^{k-1} = \Theta$ and hence σ_{-i} is a measurable function from Θ to $\Delta(A_{-i})$. Let $\sigma'_{-i} = \sigma_{-i}$. Since $R_{-i}^0(t_{-i}, \gamma) = A_{-i}$, σ'_{-i} is trivially valid. If $\varepsilon \geq 1$, the claim trivially holds because the payoff is bounded by M and hence $R_i^k(s_i, G, \gamma + 6M\varepsilon) = A_i$. Now suppose $\varepsilon < 1$. Then, since we endow Θ with the discrete metric, $(\Theta')^{\varepsilon} = \Theta'$ for any $\Theta' \subseteq \Theta$. Hence, $|s_i^1(\Theta') - t_i^1(\Theta')| \leq \varepsilon$ since $\rho^1(t_i^1, s_i^1) \leq \varepsilon$. Let $A^{\theta} = \mathbf{g}_i(a_i, a_i', \theta) \bullet \sigma'_{-i}(\theta)$ and $B^{\theta} = s_i^1(\theta) - t_i^1(\theta)$. Then,

$$\sum_{\theta \in \Theta} A^{\theta} B^{\theta} = \sum_{\left\{\theta \in \Theta : B^{\theta} < 0\right\}} A^{\theta} B^{\theta} + \sum_{\left\{\theta \in \Theta : B^{\theta} \geq 0\right\}} A^{\theta} B^{\theta} \geq -4M\varepsilon$$

where the last inequality is due to $\left|A^{\theta}\right| \leq 2M$ and $|s_i^1(\Theta') - t_i^1(\Theta')| \leq \varepsilon$ in particular for $\Theta' = \left\{\theta \in \Theta : B^{\theta} < 0\right\}$ and for $\Theta' = \left\{\theta \in \Theta : B^{\theta} \geq 0\right\}$. Therefore, by (12), $\sum_{\theta \in \Theta} A^{\theta} s_i^1(\theta) \geq -\gamma - 4M\varepsilon$. Hence, $a_i \in R_i^k(s_i, \gamma + 6M\varepsilon)$.

Now consider the induction step. By Lemma 2, to prove $a_i \in R_i^k(s_i, \gamma + 6M\varepsilon)$, it suffices to show that for every $\eta > 0$ and $\beta \in \Delta(A_i \backslash \{a_i\})$, there is a valid conjecture σ'_{-i}

under which we have

$$\int_{\Theta\times\mathcal{T}_{-i}^{k-1}}\left[\mathbf{g}_i\left(a_i,\beta,\theta\right)\bullet\sigma'_{-i}\left(\theta,t_{-i}^{k-1}\right)\right]s_i^k\left[\left(\theta,dt_{-i}^{k-1}\right)\right]\geq-\gamma-6M\varepsilon-\eta. \tag{$\bigstar$}$$

Since player $-i$ has $|A_{-i}|$ actions, let $\|\mathbf{q}-\mathbf{q}'\|_{|A_{-i}|}\equiv\max\left\{|q_1-q_1'|,...,\left|q_{|A_{-i}|}-q'_{|A_{-i}|}\right|\right\}$ for any $\mathbf{q}$ and $\mathbf{q}'$ in $\Delta\left(A_{-i}\right)$. Pick an arbitrary positive integer h. We can discretize $\Delta\left(A_{-i}\right)$ with a finite partition $\{\Phi^m\}_{m=1}^{\overline{\Delta}}$ such that for each m there is some $\mathbf{q}^m\in\Phi^m$ with $\|\mathbf{q}-\mathbf{q}^m\|_{|A_{-i}|}\leq 1/h$ for any $\mathbf{q}\in\Phi^m$. Consider

$$\mathbf{T}_{-i}^{k-1}\equiv\left\{\left(\theta,t_{-i}^{k-1}\right)\in\Theta\times\mathcal{T}_{-i}^{k-1}:\operatorname{supp}\sigma_{-i}\left(\theta,t_{-i}^{k-1}\right)\subseteq R_i^{k-1}\left(t_{-i}^{k-1},\gamma\right)\right\}.$$

We can induce from $\{\Phi^m\}$ and σ_{-i} a partition $\left\{F_m^\theta\right\}$ on $\mathbf{T}_{-i}^{k-1}$ such that for each $\theta\in\Theta$ and $m=1,...,\overline{\Delta}$,

$$F_m^\theta\equiv\left\{\left(\theta,t_{-i}^{k-1}\right)\in\mathbf{T}_{-i}^{k-1}:\sigma_{-i}\left(\theta,t_{-i}^{k-1}\right)\in\Phi^m\right\}.$$ [12]

Label $\left\{F_m^\theta\right\}$ as $\{F_{m_n}^{\theta_n}\}_{n=1}^N$ where $N=|\Theta|\times\overline{\Delta}$ such that

$$\mathbf{g}_i\left(a_i,\beta,\theta_1\right)\bullet\mathbf{q}^{m_1}\geq\cdots\geq\mathbf{g}_i\left(a_i,\beta,\theta_N\right)\bullet\mathbf{q}^{m_N}.$$

Hereafter, we write F_n instead of $F_{m_n}^{\theta_n}$ whenever no confusion may arise. Hence, $\mathbf{T}_{-i}^{k-1}=\cup_{n=1}^N F_n$ and $\left(\mathbf{T}_{-i}^{k-1}\right)^\varepsilon=\cup_{n=1}^N\left(F_n\right)^\varepsilon$.

Define $E_1=\left(F_1\right)^\varepsilon$ and $E_n=\left(F_n\right)^\varepsilon\setminus\left(\cup_{l=1}^{n-1}E_l\right)$ for $n\geq 2$. Observe that $\{E_n\}_{n=1}^N$ partitions $\left(\mathbf{T}_{-i}^{k-1}\right)^\varepsilon$, and moreover, for any $1\leq l\leq N$, we have

$$\bigcup_{n=1}^{l}E_n = \bigcup_{n=1}^{l}\left(F_n\right)^\varepsilon; \tag{13}$$

$$\bigcup_{n=l+1}^{N}E_n = \left(\mathbf{T}_{-i}^{k-1}\right)^\varepsilon\diagdown\left[\bigcup_{n=1}^{l}\left(F_n\right)^\varepsilon\right]. \tag{14}$$

We now proceed to define the conjecture σ'_{-i}. We divide $\Theta\times\mathcal{T}_{-i}^{k-1}$ into three areas: (I) $\mathbf{T}_{-i}^{k-1}$; (II) $\left(\mathbf{T}_{-i}^{k-1}\right)^\varepsilon\setminus\mathbf{T}_{-i}^{k-1}$; (III) $\left(\Theta\times\mathcal{T}_{-i}^{k-1}\right)\setminus\left(\mathbf{T}_{-i}^{k-1}\right)^\varepsilon$, and define σ'_{-i} on these three areas respectively.

First, for area (I), let $\sigma'_{-i}=\sigma_{-i}$. Second, since $t_{-i}^{k-1}\mapsto R_{-i}^{k-1}\left(t_{-i},\gamma+6M\varepsilon\right)$ is upper hemi-continuous under the product topology on $\mathcal{T}_{-i}^{k-1}$, by Kuratowski-Ryll-Nardzewski

[12] We can make each Φ^m measurable, so that each F_m^θ is also measurable.

Theorem (see Aliprantis and Border (1999)), there is a measurable selection $r(\cdot)$ with $r\left(t_{-i}^{k-1}\right) \in R_{-i}^{k-1}(t_{-i}, \gamma + 6M\varepsilon)$ for all $t_{-i}^{k-1} \in \mathcal{T}_{-i}^{k-1}$. Then, we define σ'_{-i} as $r(\cdot)$ on area (III). Third, we extend the definition of σ'_{-i} from area (I) to area (II) by Lemma 3. Recall that $\{F_n\}$ is a partition of $\mathbf{T}_{-i}^{k-1}$. Since σ_{-i} is valid, it is measurable on F_n for every n. By Lemma 3, there is a measurable function $\lambda_{-i}^n(\cdot)$ on $(F_n)^\varepsilon$ such that $\lambda_{-i}^n = \sigma_{-i}$ on F_n and for every $\left(\theta, t_{-i}^{k-1}\right) \in (F_n)^\varepsilon$ there is some $\left(\theta', s_{-i}^{k-1}\right) \in F_n$ such that

$$d^{k-1}\left(\left(\theta, t_{-i}^{k-1}\right), \left(\theta', s_{-i}^{k-1}\right)\right) < \varepsilon \text{ and } \lambda_{-i}^n\left(\theta, t_{-i}^{k-1}\right) = \sigma_{-i}\left(\theta', s_{-i}^{k-1}\right).$$

Recall also that $E_n \subseteq (F_n)^\varepsilon$, and moreover, $\{E_n\}$ forms a partition of $\left(\mathbf{T}_{-i}^{k-1}\right)^\varepsilon$. In sum, we define the conjecture $\sigma'_{-i} : \Theta \times \mathcal{T}_{-i}^{k-1} \to \Delta(A_{-i})$ as

$$\sigma'_{-i}\left(\theta, t_{-i}^{k-1}\right) = \begin{cases} \lambda_{-i}^n\left(\theta, t_{-i}^{k-1}\right), & \text{if } \left(\theta, t_{-i}^{k-1}\right) \in E_n; \\ \delta_{r\left(t_{-i}^{k-1}\right)}, & \text{if } \left(\theta, t_{-i}^{k-1}\right) \notin \left(\mathbf{T}_{-i}^{k-1}\right)^\varepsilon. \end{cases}$$

Observe that σ'_{-i} is valid in areas (I) and (III) by the definition of $r(\cdot)$ and the validity of σ_{-i}. In area (II), by the induction hypothesis and the extension λ_{-i}^n defined above, σ'_{-i} is also valid.

It remains to show that the inequality (★) holds under σ'_{-i}. It is a direct consequence of the following three lemmas which we will prove later.

Lemma 4 *We have*

$$\sum_{n=1}^{N}\left[\mathbf{g}_i(a_i, \beta, \theta_n) \bullet \mathbf{q}^{m_n}\right] t_i^k[F_n] \geq -\gamma - \frac{2M|A_{-i}|}{h}. \tag{15}$$

Lemma 4 says that replacing on the left-hand side of (12) a_i' by β and σ_{-i} by $\mathbf{q}^{m_n}$ on each F_n would induce at most a loss $\frac{2M|A_{-i}|}{h}$.

Lemma 5 *We have*

$$\begin{aligned} &\int_{\Theta \times \mathcal{T}_{-i}^{k-1}}\left[\mathbf{g}_i(a_i, \beta, \theta) \bullet \sigma'_{-i}\left(\theta, t_{-i}^{k-1}\right)\right] s_i^k\left[\left(\theta, dt_{-i}^{k-1}\right)\right] \\ \geq\ &\sum_{n=1}^{N}\left[\mathbf{g}_i(a_i, \beta, \theta_n) \bullet \mathbf{q}^{m_n}\right] s_i^k[E_n] - 2M\varepsilon - \frac{2M|A_{-i}|}{h}. \end{aligned} \tag{16}$$

Lemma 5 says that by using $\sum_{n=1}^{N}\left[\mathbf{g}_i\left(a_i,\beta,\theta_n\right)\bullet\mathbf{q}^{m_n}\right]s_i^k\left[E_n\right]$ to approximate the left-hand side of (16), we may incur two kinds of losses and both are small. One is due to the error outside $\left(\mathbf{T}_{-i}^{k-1}\right)^{\varepsilon}$, which is at most $2M\varepsilon$, and the other results from the approximation of σ'_{-i} by $\mathbf{q}^{m_n}$ on E_n, which is at most $\frac{2M|A_{-i}|}{h}$.

Lemma 6 *We have*

$$\sum_{n=1}^{N}\left[\mathbf{g}_i\left(a_i,\beta,\theta_n\right)\bullet\mathbf{q}^{m_n}\right]\left[s_i^k\left(E_n\right)-t_i^k\left(F_n\right)\right]\geq -4M\varepsilon. \tag{17}$$

Lemma 6 is a generalization of step 5 in Section 3 which says that we only have to compare the beliefs of t_i^k and s_i^k on the probabilities of sets $\{E_n\}$ and $\{F_n\}$. We will show that the difference of the two approximated payoffs is at most $4M\varepsilon$.

By adding up (15)–(17), we get

$$\int_{\Theta\times T_{-i}^{k-1}}\left[\mathbf{g}_i\left(a_i,\beta,\theta\right)\bullet\sigma'_{-i}\left(\theta,t_{-i}^{k-1}\right)\right]s_i^k\left[\left(\theta,dt_{-i}^{k-1}\right)\right]\geq -\gamma-6M\varepsilon-\frac{4M\left|A_{-i}\right|}{h}.$$

Since $\frac{4M|A_{-i}|}{h}\to 0$ as $h\to\infty$, we can choose h large enough so that (★) holds. Hence, for every $\eta>0$ and $\beta\in\Delta\left(A_i\backslash\{a_i\}\right)$, there is a valid conjecture σ'_{-i} under which (★) holds. Thus, by Lemma 2, $a_i\in R_i^k\left(s_i,\gamma+6M\varepsilon\right)$.■

We now prove Lemmas 4–6.

Proof of Lemma 4 Recall that a_i is a $\gamma-$best reply under σ_i for t_i. Hence,

$$-\gamma\leq\int_{\mathbf{T}_{-i}^{k-1}}\left[\mathbf{g}_i\left(a_i,\beta,\theta\right)\bullet\sigma_{-i}\left(\theta,t_{-i}^{k-1}\right)\right]t_i^k\left[\left(\theta,dt_{-i}^{k-1}\right)\right]. \tag{18}$$

Since $\{F_n\}_{n=1}^{N}$ is a partition of $\mathbf{T}_{-i}^{k-1}$, we can write

$$\begin{aligned}&\int_{\mathbf{T}_{-i}^{k-1}}\left[\mathbf{g}_i\left(a_i,\beta,\theta\right)\bullet\sigma_{-i}\left(\theta,t_{-i}^{k-1}\right)\right]t_i^k\left[\left(\theta,dt_{-i}^{k-1}\right)\right] \qquad (19)\\ =&\sum_{n=1}^{N}\int_{\left(\theta,t_{-i}^{k-1}\right)\in F_n}\left[\mathbf{g}_i\left(a_i,\beta,\theta\right)\bullet\sigma_{-i}\left(\theta,t_{-i}^{k-1}\right)\right]t_i^k\left[\left(\theta,dt_{-i}^{k-1}\right)\right].\end{aligned}$$

Since $\left\|\sigma_{-i}\left(\theta, t_{-i}^{k-1}\right)-\mathbf{q}^{m_n}\right\|_{|A_{-i}|} \leq 1/h$ for every $\left(\theta, t_{-i}^{k-1}\right) \in F_n$ and $\left|\mathbf{g}_i\left(a_i, \beta, \theta\right)\right| \leq 2M$, we have

$$
\begin{aligned}
& \sum_{n=1}^{N} \int_{\left(\theta, t_{-i}^{k-1}\right) \in F_n} \left[\mathbf{g}_i\left(a_i, \beta, \theta\right) \bullet \sigma_{-i}\left(\theta, t_{-i}^{k-1}\right)\right] t_i^k\left[\left(\theta, d t_{-i}^{k-1}\right)\right] \\
\leq \; & \sum_{n=1}^{N}\left[\mathbf{g}_i\left(a_i, \beta, \theta_n\right) \bullet \mathbf{q}^{m_n}\right] t_i^k\left[F_n\right]+\frac{2M\left|A_{-i}\right|}{h}.
\end{aligned} \tag{20}
$$

By combining (18)–(20), we get (15).■

Proof of Lemma 5 With $\rho^k\left(t_i^k, s_i^k\right) \leq \varepsilon$ and $t_i^k\left(\mathbf{T}_{-i}^{k-1}\right)=1$, we have $s_i^k\left[\left(\mathbf{T}_{-i}^{k-1}\right)^{\varepsilon}\right] \geq 1-\varepsilon$ and hence $s_i^k\left[\left(\Theta \times \mathcal{T}_{-i}^{k-1}\right) \setminus \left(\mathbf{T}_{-i}^{k-1}\right)^{\varepsilon}\right] \leq \varepsilon$. Since $\left|\mathbf{g}_i\left(a_i, \beta, \theta\right)\right| \leq 2M$, we have

$$
\begin{aligned}
& \int_{\Theta \times \mathcal{T}_{-i}^{k-1}} \left[\mathbf{g}_i\left(a_i, \beta, \theta\right) \bullet \sigma_{-i}^{\prime}\left(\theta, t_{-i}^{k-1}\right)\right] s_i^k\left[\left(\theta, d t_{-i}^{k-1}\right)\right] \\
\geq \; & -2M\varepsilon+\int_{\left(\mathbf{T}_{-i}^{k-1}\right)^{\varepsilon}} \left[\mathbf{g}_i\left(a_i, \beta, \theta\right) \bullet \sigma_{-i}^{\prime}\left(\theta, t_{-i}^{k-1}\right)\right] s_i^k\left[\left(\theta, d t_{-i}^{k-1}\right)\right].
\end{aligned} \tag{21}
$$

Recall that $\sigma_{-i}^{\prime}\left(\theta, t_{-i}^{k-1}\right)=\lambda_{-i}^n\left(\theta, t_{-i}^{k-1}\right)$ for all $\left(\theta, t_{-i}^{k-1}\right) \in E_n$ and $\left\{E_n\right\}_{n=1}^N$ is a partition of $\left(\mathbf{T}_{-i}^{k-1}\right)^{\varepsilon}$. Therefore, we can write

$$
\begin{aligned}
& \int_{\left(\mathbf{T}_{-i}^{k-1}\right)^{\varepsilon}} \left[\mathbf{g}_i\left(a_i, \beta, \theta\right) \bullet \sigma_{-i}^{\prime}\left(\theta, t_{-i}^{k-1}\right)\right] s_i^k\left[\left(\theta, d t_{-i}^{k-1}\right)\right] \\
= \; & \sum_{n=1}^{N} \int_{E_n} \left[\mathbf{g}_i\left(a_i, \beta, \theta\right) \bullet \lambda_{-i}^n\left(\theta, t_{-i}^{k-1}\right)\right] s_i^k\left[\left(\theta, d t_{-i}^{k-1}\right)\right].
\end{aligned} \tag{22}
$$

By the definition of λ_{-i}^n, for every $\left(\theta, t_{-i}^{k-1}\right) \in E_n$, $\lambda_{-i}^n\left(\theta, t_{-i}^{k-1}\right)=\sigma_{-i}\left(\theta^{\prime}, s_{-i}^{k-1}\right)$ for some $\left(\theta^{\prime}, s_{-i}^{k-1}\right) \in F_n$. Hence, $\left\|\lambda_{-i}^n\left(\theta, t_{-i}^{k-1}\right)-\mathbf{q}^{m_n}\right\|_{|A_{-i}|} \leq 1/h$. Since $\left|\mathbf{g}_i\left(a_i, \beta, \theta\right)\right| \leq 2M$, we have

$$
\begin{aligned}
& \sum_{n=1}^{N} \int_{E_n} \left[\mathbf{g}_i\left(a_i, \beta, \theta\right) \bullet \lambda_{-i}^n\left(\theta, t_{-i}^{k-1}\right)\right] s_i^k\left[\left(\theta, d t_{-i}^{k-1}\right)\right] \\
\geq \; & \sum_{n=1}^{N}\left[\mathbf{g}_i\left(a_i, \beta, \theta_n\right) \bullet \mathbf{q}^{m_n}\right] t_i^k\left[E_n\right]-\frac{2M\left|A_{-i}\right|}{h}.
\end{aligned} \tag{23}
$$

Then, (16) follows by combining (21)–(23).■

Proof of Lemma 6 We want to show that

$$
\sum_{n=1}^{N}\left[\mathbf{g}_i\left(a_i, \beta, \theta_n\right) \bullet \mathbf{q}^{m_n}\right]\left[s_i^k\left(E_n\right)-t_i^k\left(F_n\right)\right] \geq -4M\varepsilon. \tag{24}
$$

Recall that $A^n = \mathbf{g}_i(a_i, \beta, \theta_n) \bullet \mathbf{q}^{m_n}$ and $B^n = s_i^k(E_n) - t_i^k(F_n)$. Moreover, for notational convenience, let $C^n = A^{N-n+1}$ and $D^n = B^{N-n+1}$ for $n = 1, ..., N$.

Let L be the integer such that $A^n \geq 0$ if and only if $n \leq L$. Then, (24) becomes

$$\sum_{n=1}^{L} A^n B^n + \sum_{n=1}^{N-L} C^n D^n \geq -4M\varepsilon. \tag{25}$$

We prove this lemma by establishing three claims. Claims 2 and 3 show that $\sum_{n=1}^{L} A^n B^n \geq -2M\varepsilon$ and $\sum_{n=1}^{N-L} C^n D^n \geq -2M\varepsilon$, which concludes the proof of (24) and the lemma. Claim 1 is a technical intermediate step which will be used in the proofs for Claims 2 and 3.

Claim 1 *We have* $\sum_{n=1}^{l} B^n \geq -\varepsilon$ *and* $\sum_{n=1}^{l} D^n \leq \varepsilon$ *for* $1 \leq l \leq N$.

Proof. To see $\sum_{n=1}^{l} B^n \geq -\varepsilon$, observe that

$$\sum_{n=1}^{l} B^n = \sum_{n=1}^{l} s_i^k[E_n] - \sum_{n=1}^{l} t_i^k[F_n] = s_i^k\left[\cup_{n=1}^{l} E_n\right] - t_i^k\left[\cup_{n=1}^{l} F_n\right] \geq -\varepsilon. \tag{26}$$

where the first equality is by the definition of B^n; the second equality follows from our construction that sets in $\{F_n\}_{n=1}^{N}$ are pair-wise disjoint and sets in $\{E_n\}_{n=1}^{N}$ are pair-wise disjoint; the last inequality follows because $\cup_{n=1}^{l} E_n = \left(\cup_{n=1}^{l} F_n\right)^{\varepsilon}$ by (13) and $\rho^k\left(t_i^k, s_i^k\right) \leq \varepsilon$. To see $\sum_{n=1}^{l} D^n \leq \varepsilon$, note that by (14) and $t_i^k\left(\mathbf{T}_{-i}^{k-1}\right) = 1$, $\sum_{n=N-l+1}^{N} t_i^k[F_n] = 1 - \sum_{n=1}^{N-l} t_i^k[F_n]$ and $\sum_{n=N-l+1}^{N} s_i^k[E_n] \leq 1 - \sum_{n=1}^{N-l} s_i^k[E_n]$. Hence,

$$\sum_{n=1}^{l} D^n = \sum_{n=N-l+1}^{N} s_i^k[E_n] - \sum_{n=N-l+1}^{N} t_i^k[F_n] \leq \left(1 - \sum_{n=1}^{N-l} s_i^k[E_n]\right) - \left(1 - \sum_{n=1}^{N-l} t_i^k[F_n]\right).$$

Hence, $\sum_{n=1}^{l} D^n \leq \varepsilon$ by (26).■

Claim 2 $\sum_{n=1}^{L} A^n B^n \geq -2M\varepsilon$.

Proof. We establish this claim in two steps:

Step 1 *For every integer* l *such that* $1 \leq l \leq L$, *there exists a vector* $\overline{B} \in \Re^l$ *which has the following properties: (1)* $\sum_{n=1}^{l} \overline{B}^n = \sum_{n=1}^{l} B^n$; *(2)* $\sum_{n=1}^{l'} \overline{B}^n \geq -\varepsilon$ *for every* $1 \leq l' \leq l$; *(3)* $\sum_{n=1}^{l} A^n B^n \geq \sum_{n=1}^{l} A^n \overline{B}^n$; *(4)* $\overline{B}^n \leq 0$ *for* $n < l$.

We prove the existence of $\overline{B}$ by induction on l. If $l=1$, let $\overline{B}=B^1$. Then, properties (1) and (3) are obvious. Property (2) is due to Claim 1. Property (4) is vacuously true. Now suppose for a positive integer l there is some $\overline{Z}=\left(\overline{Z}^1,...,\overline{Z}^l\right)\in\Re^l$ which satisfies properties (1)–(4). We now proceed to define $\overline{B}$ as follows so that $\overline{B}$ is a $(l+1)$-vector which makes the statement true for the case $(l+1)$. Let $\overline{B}=\left(\overline{B}^1,...,\overline{B}^l,\overline{B}^{l+1}\right)$ where $\overline{B}^n=\overline{Z}^n$ for every $n=1,...,l-1$, and

$$\left(\overline{B}^l,\overline{B}^{l+1}\right)=\begin{cases}\left(\overline{Z}^l,B^{l+1}\right), & \text{if } \overline{Z}^l\leq 0;\\ \left(0,\overline{Z}^l+B^{l+1}\right), & \text{if } \overline{Z}^l>0.\end{cases}$$

Property (1) follows from the induction hypothesis. To see this, observe that

$$\sum_{n=1}^{l+1}\overline{B}^n=\sum_{n=1}^{l-1}\overline{B}^n+\overline{B}^l+\overline{B}^{l+1}=\sum_{n=1}^{l-1}\overline{Z}^n+\overline{Z}^l+B^{l+1}=\sum_{n=1}^{l}B^n+B^{l+1}=\sum_{n=1}^{l+1}B^n$$

where the second equality follows because $\overline{B}^l+\overline{B}^{l+1}=\overline{Z}^l+B^{l+1}$ by our definition of $\overline{B}$; the third equality follows from the induction hypothesis that $\sum_{n=1}^{l}\overline{Z}^n=\sum_{n=1}^{l}B^n$.

To see property (2), suppose first $l'\leq l$. Then,

$$\sum_{n=1}^{l'}\overline{B}^n=\begin{cases}\sum_{n=1}^{l'}\overline{Z}^n, & \text{if } [l'\leq l-1] \text{ or } [l'=l \text{ and } \overline{Z}^l\leq 0];\\ \sum_{n=1}^{l'-1}\overline{Z}^n, & \text{if } l'=l \text{ and } \overline{Z}^l>0.\end{cases}$$

Since $\sum_{n=1}^{l'}\overline{Z}^n\geq-\varepsilon$ and $\sum_{n=1}^{l'-1}\overline{Z}^n\geq-\varepsilon$ by the induction hypothesis, we have $\sum_{n=1}^{l'}\overline{B}^n\geq-\varepsilon$. Second, suppose $l'=l+1$. Then, $\sum_{n=1}^{l+1}\overline{B}^n=\sum_{n=1}^{l+1}B^n$ by property (1) proved above. Hence, $\sum_{n=1}^{l+1}\overline{B}^n=\sum_{n=1}^{l+1}B^n\geq-\varepsilon$ by Claim 1.

We now prove property (3). Observe first that

$$A^l\overline{B}^l+A^{l+1}\overline{B}^{l+1}=\begin{cases}A^l\overline{Z}^l+A^{l+1}B^{l+1}, & \text{if } \overline{Z}^l\leq 0;\\ A^{l+1}\left(\overline{Z}^l+B^{l+1}\right), & \text{if } \overline{Z}^l>0.\end{cases}$$

Moreover, $A^{l+1}\left(\overline{Z}^l+B^{l+1}\right)\leq A^l\overline{Z}^l+A^{l+1}B^{l+1}$ when $\overline{Z}^l>0$, because $A^{l+1}\leq A^l$. Hence, $A^l\overline{B}^l+A^{l+1}\overline{B}^{l+1}\leq A^l\overline{Z}^l+A^{l+1}B^{l+1}$. Then,

$$\sum_{n=1}^{l+1}A^n\overline{B}^n=\sum_{n=1}^{l-1}A^n\overline{B}^n+A^l\overline{B}^l+A^{l+1}\overline{B}^{l+1}\leq\sum_{n=1}^{l}A^n\overline{Z}^n+A^{l+1}B^{l+1}\leq\sum_{n=1}^{l+1}A^nB^n,$$

where the first inequality follows from the definition of $\overline{B}$ and the fact that $A^l\overline{B}^l + A^{l+1}\overline{B}^{l+1} \leq A^l\overline{Z}^l + A^{l+1}B^{l+1}$; the second inequality is due to the induction hypothesis. Hence, property (3) is satisfied.

To see property (4), observe that

$$\left(\overline{B}^1, ..., \overline{B}^l\right) = \begin{cases} \left(\overline{Z}^1, ..., \overline{Z}^{l-1}, \overline{Z}^l\right), & \text{if } \overline{Z}^l \leq 0; \\ \left(\overline{Z}^1, ..., \overline{Z}^{l-1}, 0\right), & \text{if } \overline{Z}^l > 0. \end{cases}$$

Since $\overline{Z}^n \leq 0$ for every $1 \leq n \leq l-1$ by the induction hypothesis, property (4) is also satisfied.

Step 2 $\sum_{n=1}^{L} A^n B^n \geq -2M\varepsilon$.

By step 1, we can find a L-vector $\overline{B}$ which satisfies properties (1)–(4). First, suppose $\overline{B}^L \leq 0$. Then, by property (4), $\overline{B}^n \leq 0$ for every $n \leq L$. Hence,

$$\sum_{n=1}^{L} A^n B^n \geq \sum_{n=1}^{L} A^n \overline{B}^n \geq 2M \sum_{n=1}^{L} \overline{B}^n \geq -2M\varepsilon,$$

where the first inequality follows from property (3); the second inequality follows because $\overline{B}^n \leq 0$ and $|A^n| \leq 2M$ for every $n \leq L$; the last inequality follows from property (2).

Second, suppose $\overline{B}^L > 0$. Then,

$$\sum_{n=1}^{L} A^n B^n \geq \sum_{n=1}^{L} A^n \overline{B}^n \geq \sum_{n=1}^{L-1} A^n \overline{B}^n \geq 2M \sum_{n=1}^{L-1} \overline{B}^n \geq -2M\varepsilon,$$

where the first inequality follows from property (3); the second inequality follows because A^L and $\overline{B}^L$ are both nonnegative, i.e., $A^L\overline{B}^L \geq 0$; the third inequality follows from property (4) and $|A^n| \leq 2M$ for every n; the last inequality follows from property (2).

Claim 3 $\sum_{n=1}^{N-L} C^n D^n \geq -2\varepsilon$.

Proof. Observe that $\sum_{n=1}^{N-L} C^n D^n = \sum_{n=1}^{N-L} (-C^n)(-D^n)$. Since C^n is increasing in n and $C^n \leq 0$ for all $N-L \geq n \geq 1$, it follows that $(-C^n)$ is decreasing in n and $(-C^n) \geq 0$ for all $N-L \geq n \geq 1$. Moreover, Claim 1 implies that $\sum_{n=1}^{l} (-D^n) \geq -\varepsilon$ for $1 \leq l \leq N$. Hence, following the proof of Claim 2, and replacing A^n, B^n, and L in the proof with $(-C^n)$, $(-D^n)$, and $(N-L)$ respectively, we get $\sum_{n=1}^{N-L} C^n D^n \geq -2\varepsilon$.■

A.3 Proofs of Proposition 2 and Theorem 2

A.3.1 Proof of Proposition 2

Proposition 2 *$d^{us}(t_i^*, t_i) \geq \frac{M}{16}$ for any finite type t_i.* Hence, finite types are not dense under d^{us}.

Proof. Pick any finite type t_i. Let $T_i \times T_{-i} \subset \mathcal{T}_i \times \mathcal{T}_{-i}$ be the smallest belief-closed set such that $t_i \in T_i$. Since t_i is a finite type, $|T_i \times T_{-i}| < \infty$. For simplicity, we abbreviate $|T_i \times T_{-i}|$ as $|T|$. Henceforth, we fix

$$N = 20\,|T|\,.$$

We will consider a finite game with every player's action space being the N-truncation of a^0, a^1, a^2,... where for every $n \geq 0$, a^n is defined as in Section 4. From now on, for every $n \geq 0$, we will abuse notation and denote the finite truncation of $a^n = ((a^n)_1, ..., (a^n)_N)$ also by a^n.

Consider the following finite game, $G = (A_i, g_i)_{i=1,2}$ where for $i = 1, 2$, $A_i = \{a^0, a^1, a^2, ...\}$ and for any $a_i = ((a_i)_1, ..., (a_i)_N) \in A_i$, $a_{-i} = ((a_{-i})_1, ..., (a_{-i})_N) \in A_{-i}$ and $\theta \in \Theta$,

$$g_i\left[a_i,\ a_{-i},\ \theta\right] = \begin{cases} 0, & \text{if } a_i = a^0; \\ M \times \min\left\{ \begin{array}{c} -\left[(a_i)_1 - \theta\right]^2, -\left[(a_i)_2 - (a_{-i})_1\right]^2, \\ ..., -\left[(a_i)_N - (a_{-i})_{N-1}\right]^2 \end{array} \right\}, & \text{if } a_i \neq a^0. \end{cases} \tag{27}$$

Since the game is now fixed, we will simplify the notation in this proof by writing $R_i\,(t_i, \gamma)$ instead of $R_i\,(t_i, G, \gamma)$. We prove that $d^{us}(t_i^*, t_i) \geq \frac{M}{16}$ by establishing the following two claims:

Claim 4 *$a^1 \in R_i\,(t_i^*, 0)$.*

Proof. Let $t\,(1) = t_i^*$. Define $\widehat{R}\,(0) \in \left(\left(2^{A_i}\right)^{\widehat{T}_i}\right)_{i=1,2}$ by $\widehat{R}_i\,(t\,(k)\,, 0) = \left\{a^k\right\}$ if k is odd and $\widehat{R}_{-i}\,(t\,(k)\,, 0) = \left\{a^k\right\}$ if k is even. By equation (5), for any odd k, $\sigma_{-i}^k : \Theta \times \mathcal{T}_{-i} \to \Delta\,(A_{-i})$ such that $\sigma_{-i}^k\,(\theta, t_{-i}) \equiv \delta_{a^{k+1}}$ for all (θ, t_{-i}), a^k is a $0-$best reply for type $t\,(k)$ under σ_{-i}^k. The case with an even k is similar. Hence, $\overline{R}\,(0)$ satisfies the $0-$best reply property in the type space $\widehat{T}_i \times \widehat{T}_{-i}$. Hence, $\overline{R}_i\,(t\,(k)\,, 0) \subseteq R_i\,(t\,(k)\,, 0)$.■

Claim 5 $a^1 \notin R_i\left(t_i, \frac{M}{16}\right)$.

Proof. In this proof, it is convenient for us to denote by $\nu_{s_i,\sigma_{-i}} \in \Delta\left(A_{-i} \times \Theta \times \mathcal{T}_{-i}\right)$ the measure induced from a conjecture σ_{-i} and a type s_i, i.e., for any measurable set $E \subseteq \mathcal{T}_{-i}$ and $\left(\overline{a}_{-i}, \overline{\theta}\right) \in A_{-i} \times \Theta$,

$$\nu_{s_i,\sigma_{-i}}\left(E \times \{(\overline{a}_{-i}, \theta)\}\right) \equiv \int_E \sigma_{-i}\left(\overline{\theta}, t_{-i}\right)[\overline{a}_{-i}]\, \pi_i^*\left(s_i\right)\left[\left(\theta = \overline{\theta}, dt_{-i}\right)\right].$$

Clearly, $\text{marg}_{\Theta \times \mathcal{T}_{-i}} \nu_{s_i,\sigma_{-i}} = \pi_i^*\left[s_i\right]$. We prove the result in four steps.

Step 1 *$[(a^m)_1, ..., (a^m)_C] \neq [(a^n)_1, ..., (a^n)_C]$ for any $C \geq 18$, $n \neq m$, and $n \leq C$.*

We denote the k^{th} block in a^1 as

$$B_k = \begin{cases} \underbrace{(1, 1, ..., 1)}_{k \text{ times}} & \text{if } k \text{ is odd;} \\ \underbrace{(0, 0, ..., 0)}_{k \text{ times}} & \text{if } k \text{ is even.} \end{cases}$$

Hence, $a^1 = (B_1, B_2, B_3, ...)$. Say a block is *interior in a^l* if $B_k = \left(\left(a^l\right)_{l_1}, ..., \left(a^l\right)_{l_2}\right)$ for some $2 \leq l_1 < l_2 \leq C - 1$, and $\left(a^l\right)_{l_1 - 1} \neq \left(a^l\right)_{l_1}$, $\left(a^l\right)_{l_2} \neq \left(a^l\right)_{l_2+1}$. Observe that since $n \neq m$, a^n and a^m cannot have an interior block at the same position. That is, there is no l_1 and l_2 such that $2 \leq l_1 < l_2 \leq C - 1$ and for some B_k,

$$\left[(a^m)_{l_1}, ..., (a^m)_{l_2}\right] = \left[(a^n)_{l_1}, ..., (a^n)_{l_2}\right] = B_k,$$
$$\text{and } (a^m)_{l_1-1} = (a^n)_{l_1-1} \neq (a^n)_{l_1}, (a^m)_{l_2} = (a^n)_{l_2} \neq (a^n)_{l_2+1}.$$

Hence, to prove $[(a^m)_1, ..., (a^m)_C] \neq [(a^n)_1, ..., (a^n)_C]$, it suffices to show $[(a^n)_1, ..., (a^n)_C]$ contains an interior block. For $n \leq C$, suppose that a^n starts with part of B_{h+1}. That is, we have deleted $(B_1, ..., B_h)$ from a^1 to get a^n. Also, to get a^n, we need to delete exactly the first $n - 1$ coordinates of a^1. Hence,

$$\frac{h(h+1)}{2} \leq n - 1 \Rightarrow h^2 \leq 2n - 2,$$

where $\frac{h(h+1)}{2}$ corresponds to the number of coordinates in $(B_1, ..., B_h)$.

The first two blocks in a^n are (parts of) B_{h+1} and B_{h+2}, so the combined coordinates are at most $2h+3$. If $2h+3 \leq C-1$, a^n contains B_{h+2} as an interior block. Since $h \leq \sqrt{2n-2}$ and $n \leq C$, we have $2h+3 \leq 2\sqrt{2C-2}+3$. With $C \geq 18$, we have $2\sqrt{2C-2}+3 \leq C-1$. Therefore, a^n contains B_{h+2} as an interior block if $C \geq 18$. Since a^n and a^m cannot have an interior block at the same position, we conclude $[(a^m)_1, ..., (a^m)_C] \neq [(a^n)_1, ..., (a^n)_C]$.

Step 2 *For any positive integer $1 \leq n \leq N-1$, $i = 1, 2$, $\bar{t}_i \in \mathcal{T}_i$, and $a^n \in R_i\left(\bar{t}_i, \frac{M}{16}\right)$ which is a $\frac{M}{16}-$best reply to a valid conjecture σ_{-i}, we have*

1. $(a^n)_1 = 0$ *iff* $\widehat{\xi}_1\left(\bar{t}_i\right) \in [0, \frac{1}{2}]$;
2. $\nu_{\bar{t}_i,\sigma_{-i}}\left(a^{n+1}\right) \geq \frac{3}{4}$.

To see (1), let $a_{-i} = ((a_{-i})_1, ..., (a_{-i})_N) \in A_{-i}$. Let $\nu \equiv \nu_{\bar{t}_i,\sigma_{-i}}$ and $\mathbb{E}_\nu(\cdot) \equiv \int (\cdot)\, d\nu$. For $n \geq 1$, we have $\mathbb{E}_\nu\left[g_i\left(a^n, a_{-i}, \theta\right)\right] \leq \mathbb{E}_\nu\left[-M \times [(a^n)_1 - \theta]^2\right]$. Since $\theta \in \{0,1\}$, $\theta^2 = \theta$. Since $\text{marg}_{\Theta\times\mathcal{T}_{-i}}\nu = \pi_i^*\left[\bar{t}_i\right]$, $\mathbb{E}_\nu(\theta) = \widehat{\xi}_1\left(\bar{t}_i\right)$. Hence,

$$\mathbb{E}_\nu\left[-M \times [(a^n)_1 - \theta]^2\right] = \begin{cases} -\widehat{\xi}_1\left(\bar{t}_i\right) M & \text{if } (a^n)_1 = 0; \\ \left(\widehat{\xi}_1\left(\bar{t}_i\right) - 1\right) M & \text{if } (a^n)_1 = 1. \end{cases}$$

However, $g_i\left(a^0, a_{-i}, \theta\right) = 0$ for any a_{-i}. Therefore, to make $a^n \in R_i\left(\bar{t}_i, \frac{M}{16}\right)$, we must have

$$\begin{aligned} \widehat{\xi}_1\left(\bar{t}_i\right) &\in \left[0, \frac{1}{16}\right] \subset \left[0, \frac{1}{2}\right], \text{ if } (a^n)_1 = 0; \\ \widehat{\xi}_1\left(\bar{t}_i\right) &\in \left[\frac{15}{16}, 1\right] \subset \left(\frac{1}{2}, 1\right], \text{ if } (a^n)_1 = 1. \end{aligned}$$

We now prove (2). Suppose $\nu\left(a^{n+1}\right) = \alpha$. We have

- $\mathbb{E}_\nu\left[g_i\left(a^n, a^{n+1}, \theta\right)\right] - \mathbb{E}_\nu\left[g_i\left(a^0, a^{n+1}, \theta\right)\right] \leq 0$;
- $\mathbb{E}_\nu\left[g_i\left(a^n, a^0, \theta\right)\right] - \mathbb{E}_\nu\left[g_i\left(a^0, a^0, \theta\right)\right] \leq -\frac{M}{4}$, which is because $g_i\left(a^n, a^0, \theta\right) \leq -\frac{M}{4}$ and $g_i\left(a^0, a^0, \theta\right) = 0$;
- $\mathbb{E}_\nu\left[g_i\left(a^n, a^m, \theta\right)\right] - \mathbb{E}_\nu\left[g_i\left(a^0, a^m, \theta\right)\right] \leq -M$ for any $m \notin \{0, n+1\}$. To see this, note that $N-1 \geq 18$ and hence by step 1,

$$\left[(a^m)_1, ..., (a^m)_{N-1}\right] \neq \left[\left(a^{n+1}\right)_1, ..., \left(a^{n+1}\right)_{N-1}\right] = \left[(a^n)_2, ..., (a^n)_N\right].$$

Therefore,

$$\mathbb{E}_{\nu}\left[g_i\left(a^n, a_{-i}, \theta\right)\right] - \mathbb{E}_{\nu}\left[g_i\left(a^0, a_{-i}, \theta\right)\right] \leq \alpha \times (0) + (1-\alpha) \times (-\frac{M}{4}).$$

To make a^n a $\frac{M}{16}-$rationalizable action for t, we must have

$$(1-\alpha) \times (-\frac{M}{4}) \geq -\frac{M}{16}.$$

That is, $\alpha \geq \frac{3}{4}$.

Step 3 *Suppose that for distinct $m, n \leq N-1$ and some finite type $\bar{t}_i$, we have $\{a^n, a^m\} \subset R_i\left(\bar{t}_i, \frac{M}{16}\right)$. Then, there is some finite type $\bar{t}_{-i}$ such that $\{a^{n+1}, a^{m+1}\} \subset R_{-i}\left(\bar{t}_{-i}, \frac{M}{16}\right)$.*

Since $a^n \in R_i\left(\bar{t}_i, \frac{M}{16}\right)$, a^n is a $\frac{M}{16}-$best reply for $\bar{t}_i$ under some valid conjecture σ^n_{-i}. Let $\nu^n \equiv \nu_{\bar{t}_i, \sigma^n_{-i}}$. By property (2) in step 2, $\nu^n\left(a^{n+1}\right) \geq \frac{3}{4}$. Similarly, $a^m \in R_i\left(\bar{t}_i, \frac{M}{16}\right)$ and a^m is a $\frac{M}{16}-$best reply for $\bar{t}_i$ under some valid conjecture σ^m_{-i}. Let $\nu^m \equiv \nu_{\bar{t}_i, \sigma^m_{-i}}$. By property (2) in step 2, $\nu^m\left(a^{m+1}\right) \geq \frac{3}{4}$. Since $\text{marg}_{\Theta \times \mathcal{T}_{-i}} \nu^n = \text{marg}_{\Theta \times \mathcal{T}_{-i}} \nu^m = \pi^*_i\left[\bar{t}_i\right]$, $\nu^n\left(a^{n+1}\right) \geq \frac{3}{4}$, and $\nu^m\left(a^{m+1}\right) \geq \frac{3}{4}$, we must have

$$\begin{aligned}
\pi^*_i\left[\bar{t}_i\right]\left(\left\{t_{-i} : \nu^n(a^{n+1}, t_{-i}) > 0\right\}\right) &= \nu^n\left[\left\{(a_{-i}, t_{-i}) : \nu^n(a^{n+1}, t_{-i}) > 0\right\}\right] \\
&\geq \nu^n\left[\left\{\left(a^{n+1}, t_{-i}\right) : \nu^n(a^{n+1}, t_{-i}) > 0\right\}\right] \geq \frac{3}{4};
\end{aligned}$$

Similarly, $\pi^*_i\left[\bar{t}_i\right]\left(\{t_{-i} : \nu^m(a^{m+1}, t_{-i}) > 0\}\right) \geq \frac{3}{4}$. Hence,

$$\left\{t_{-i} : \nu^n(a^{n+1}, t_{-i}) > 0\right\} \cap \left\{t_{-i} : \nu^m(a^{m+1}, t_{-i}) > 0\right\} \neq \varnothing.$$

Let $\bar{t}_{-i} \in \{t_{-i} : \nu^n(a^{n+1}, t_{-i}) > 0\} \cap \{t_{-i} : \nu^m(a^{m+1}, t_{-i}) > 0\}$. Since $\nu^n(a^{n+1}, t_{-i}) > 0$ only if $a^{n+1} \in R_{-i}\left(t_{-i}, \frac{M}{16}\right)$ and $\nu^m(a^{m+1}, t_{-i}) > 0$ only if $a^{m+1} \in R_{-i}\left(t_{-i}, \frac{M}{16}\right)$, we have $\{a^{n+1}, a^{m+1}\} \subset R_{-i}\left(\bar{t}_{-i}, \frac{M}{16}\right)$.

Step 4 $a^1 \notin R_i\left(t_i, \frac{M}{16}\right)$.

Suppose instead that $a^1 \in R_i\left(t_i, \frac{M}{16}\right)$. By property (2) in step 2, we can construct a chain of rationalization where the first $|T|+1$ elements are as follows

$$t_i = \widehat{t}(1) \to \widehat{t}(2) \to ... \to \widehat{t}(|T|) \to \widehat{t}(|T|+1) \to \cdots$$

$$\text{such that } \widehat{t}(k) \in T_i \text{ if } n \text{ is odd and } \widehat{t}(k) \in T_{-i} \text{ if } k \text{ is even;}$$

$$\widehat{t}(k+1) \text{ is in the support of } \pi^*\left[\widehat{t}(k)\right]; \text{ and } a^k \text{ is } \frac{M}{16} - \text{rationalizable for } \widehat{t}(k).$$

Recall that $|T| = |T_i \times T_{-i}|$, there exists $\widehat{t}(n) = \widehat{t}(m) \equiv \widetilde{t}(1) \in T_i$ such that $n < m \leq |T|+1$. Thus, $a^n \in R_i\left(\widetilde{t}(1), \frac{M}{16}\right)$ and $a^m \in R_i\left(\widetilde{t}(1), \frac{M}{16}\right)$. By repeatedly applying step 3, we can construct two chains of rationalization (for a^n and a^m, respectively), which have the common first $19|T|-1$ elements as follows,

$$\widetilde{t}(1) \to \widetilde{t}(2) \to ... \to \widetilde{t}(19|T|-2) \to \widetilde{t}(19|T|-1) \to \cdots$$

$$\text{such that for } k \leq 19|T|-1,$$

$$\widetilde{t}(k) \in T_i \text{ and } \left\{a^{n+k-1}, a^{m+k-1}\right\} \subset R_i\left(\widetilde{t}(k), \frac{M}{16}\right) \text{ if } k \text{ is odd;}$$

$$\widetilde{t}(k) \in T_{-i} \text{ and } \left\{a^{n+k-1}, a^{m+k-1}\right\} \subset R_{-i}\left(\widetilde{t}(k), \frac{M}{16}\right) \text{ if } k \text{ is even.}$$

By the definition of a^n and a^m, we have

$$\begin{aligned}
\left[(a^n)_1, (a^n)_2, ..., (a^n)_{19|T|-2}, (a^n)_{19|T|-1}\right] &= \left[(a^n)_1, \left(a^{n+1}\right)_1, ..., \left(a^{n+19|T|-3}\right)_1, \left(a^{n+19|T|-2}\right)_1\right]; \\
\left[(a^m)_1, (a^m)_2, ..., (a^m)_{19|T|-2}, (a^m)_{19|T|-1}\right] &= \left[(a^m)_1, \left(a^{m+1}\right)_1, ..., \left(a^{m+19|T|-3}\right)_1, \left(a^{m+19|T|-2}\right)_1\right].
\end{aligned}$$

Since $19|T|-1 \geq 18$, $n \neq m$, and $n \leq 19|T|-1$, by step 1 $(a^n)_{k^*} \neq (a^m)_{k^*}$ for some $1 \leq k^* \leq 19|T|-1$, i.e. $\left(a^{n+k^*-1}\right)_1 = 0 < 1 = \left(a^{m+k^*-1}\right)_1$. By Property (1) in step 2, $\left(a^{n+k^*-1}\right)_1 = 0$ implies $\widehat{\xi}_1\left(\widetilde{t}(k^*-1)\right) \in [0, \frac{1}{2}]$, but $\left(a^{m+k^*-1}\right)_1 = 1$ implies $\widehat{\xi}_1\left(\widetilde{t}(k^*-1)\right) \notin [0, \frac{1}{2}]$, which is a contradiction. Hence, $a^1 \notin R_i\left(t_i, \frac{M}{16}\right)$.■

A.3.2 Proof of Theorem 2

Theorem 2 *Finite types are nowhere dense under d^{us}.*

Proof. Let $\overline{T^F}$ denote the closure of the set of finite types under d^{us}. It suffices to show that for any finite type $\bar{t}_i$, there is a sequence of types $\{t_i(n)\}_{n=1}^{\infty} \subset \mathcal{T}_i \backslash \overline{T^F}$ such that $d^{us}\left(t_i(n), \bar{t}_i\right) \to 0$. Let s_i^* be the unique type whose iterated expectation is $y^* = (0,1,0,0,1,1,1,0,0,0,0,...)$. Recall $t_{-i}^* \in \mathcal{T}_{-i}$ is the unique type which generates the iterated expectations $x^* = (1,0,0,1,1,1,0,0,0,0,...)$. Hence, $\pi_i^*(s_i^*)\left[\left\{\theta = 0, t_{-i} = t_{-i}^*\right\}\right] = 1$. For each positive integer n, consider $\mu^n \in \Delta(\Theta \times \mathcal{T}_{-i})$ such that for any Borel set $E \subseteq \Theta \times \mathcal{T}_{-i}$,

$$\mu^n[E] \equiv \left(1 - \frac{1}{n}\right)\pi_i^*(\bar{t}_i)[E] + \frac{1}{n}\pi_i^*(s_i^*)[E]. \tag{28}$$

Since π_i^* is a homeomorphism between $\mathcal{T}_i$ and $\Delta(\Theta \times \mathcal{T}_{-i})$, there is some type $t_i(n)$ with $\pi_i^*(t_i(n)) = \mu^n$. Recall that for any $s_i \in \mathcal{T}_i$, the marginal distribution of $\pi_i^*(s_i)$ on Y_{-i}^{k-1} agrees with $\widetilde{t}_i^k$. Hence, by (28), for any $k \geq 1$ and Borel set $E \subseteq Y_{-i}^{k-1}$, we have

$$(t_i(n))^k[E] = \left(1 - \frac{1}{n}\right)\bar{t}_i^k[E] + \frac{1}{n}(s_i^*)^k[E]. \tag{29}$$

We now divide the rest of the proof into the following two steps.

Step 1 $d^{us}\left(t_i(n), \bar{t}_i\right) \to 0$ *as* $n \to \infty$.

We show that $d^{uw}\left(t_i(n), \bar{t}_i\right) \to 0$ as $n \to \infty$, which implies $d^{us}\left(t_i(n), \bar{t}_i\right) \to 0$ by Theorem 1. To see that $\rho^k\left(t_i(n)^k, \bar{t}_i^k\right) \leq \frac{1}{n}$ for every k, observe that by (29), we have

$$(t_i(n))^k[E] = \left(1 - \frac{1}{n}\right)\bar{t}_i^k[E] + \frac{1}{n}(s_i^*)^k[E] \leq (\bar{t}_i)^k\left[E^{1/n}\right] + \frac{1}{n}$$

for any Borel set $E \subseteq Y^{k-1}$. Hence, $\rho^k\left(t_i(n)^k, \bar{t}_i^k\right) \leq \frac{1}{n}$.

Step 2 $t_i(n) \in \mathcal{T} \backslash \overline{T^F}$ *for every* $n \geq 8$.

We prove this claim by showing that $d^{us}(t_i(n), t_i) \geq \frac{M}{2n}$ for any finite type t_i and $n \geq 8$. Pick any finite type t_i. Let $T_i \times T_{-i} \subset \mathcal{T}_i \times \mathcal{T}_{-i}$ be the smallest belief-closed set such that $t_i \in T_i$. Since t_i is a finite type, $|T_i \times T_{-i}| < \infty$. Following the proof of Proposition 2 in Appendix A.3.1, we can construct a finite game $G = (A_j, g_j)_{j=i,-i}$ (as defined in (27)) with an action a^1 for player $-i$ such that $a^1 \in R_{-i}\left(t_{-i}^*, G, 0\right)$, but $a^1 \notin R_{-i}\left(t_{-i}, G, \frac{M}{16}\right)$ for any $t_{-i} \in T_{-i}$. Moreover, the payoffs in G are always between $-M$ and 0. Based upon G, we define another finite game $G' = \left(A_j', g_j'\right)_{j=i,-i}$ where $A_i' = A_i \times \{z_1, z_2\}$, $A_{-i}' = A_{-i}$, $g_{-i}'[(a_i, z), a_{-i}, \theta] = g_{-i}(a_i, a_{-i}, \theta)$, and

$$g_i'[(a_i, z), a_{-i}, \theta] = g_i(a_i, a_{-i}, \theta) + \begin{cases} \frac{M}{n}, & \text{if } z = z_1; \\ M, & \text{if } z = z_2,\ a_{-i} = a^1; \\ 0, & \text{if } z = z_2,\ a_{-i} \neq a^1. \end{cases}$$

Since the payoffs in G are between $-M$ and 0, it follows that the payoffs in G' are between $-M$ and M. We now divide the proof of step 2 into the following three substeps.

Step 2.1 $a^1 \in R_{-i}\left(t_{-i}^*, G', 0\right)$.

Let $\widehat{T}_i \times \widehat{T}_{-i} \subset \mathcal{T}_i \times \mathcal{T}_{-i}$ be the smallest belief-closed subset such that $t_{-i}^* \in \widehat{T}_{-i}$. Since $a^1 \in R_{-i}\left(t_{-i}^*, G, 0\right)$, it suffices to show that $R_{-i}(s_{-i}, G, 0) \subseteq R_{-i}(s_{-i}, G', 0)$ for all $s_{-i} \in \widehat{T}_{-i}$.

Let $a_i \in R_i(s_i, G, 0)$. Then, there is some valid conjecture σ_{-i} such that a_i is a $0-$best reply to σ_{-i}. Hence, there is a function $z_{s_i} : A_i \to \{z_1, z_2\}$ such that $(a_i, z_{s_i}(a_i))$ is a $0-$best reply for s_i under σ_{-i} in G'. Define $\widehat{R}(0) \in \left(\left(2^{A_j}\right)^{\widehat{T}_j}\right)_{j=i,-i}$ as

$$\widehat{R}_i(s_i, 0) = \{(a_i, z) : a_i \in R_i(s_i, G, 0) \text{ and } z = z_{s_i}(a_i)\}, \forall s_i \in \widehat{T}_i,$$
$$\widehat{R}_{-i}(s_{-i}, 0) = R_{-i}(s_{-i}, G, 0), \forall t_{-i} \in \widehat{T}_{-i}.$$

We claim that $\widehat{R}(0)$ has the $0-$best reply property in G' for the type space $\widehat{T}_i \times \widehat{T}_{-i}$. By our construction, it suffices to check player $-i$. Let $a_{-i} \in R_{-i}(s_{-i}, G, 0)$. Hence, there is a valid conjecture σ_i such that a_{-i} is a $0-$best reply under σ_i for type s_{-i} in G. We define a new conjecture σ_i' such that $\sigma_i'(\theta, s_i)[a_i, z] = \sigma_i(\theta, s_i)[a_i]$ if $z = z_{s_i}(a_i)$; $\sigma_i'(\theta, s_i)[a_i, z] = 0$ otherwise. Then, by choosing any action a_{-i}', player $-i$ gets exactly the same payoff under the conjecture σ_i' in G' as the payoff under the conjecture σ_i in G. Hence, a_{-i} remains a $0-$best reply to σ_i' in G'. Therefore, $R_{-i}(s_{-i}, G, 0) \subseteq R_{-i}(s_{-i}, G', 0)$.

Step 2.2 $a^1 \notin R_{-i}\left(t_{-i}, G', \frac{M}{16}\right)$ *for any* $t_{-i} \in T_{-i}$.

Since $a^1 \notin R_{-i}\left(t_{-i}, G, \frac{M}{16}\right)$, it suffices to show that $R_{-i}\left(t_{-i}, G', \frac{M}{16}\right) \subseteq R_{-i}\left(t_{-i}, G, \frac{M}{16}\right)$ for $t_{-i} \in T_{-i}$. Define $\widetilde{R}\left(\frac{M}{16}\right) \in \left(\left(2^{A_i}\right)^{T_i}\right)_{i=1,2}$ as

$$\widetilde{R}_i\left(t_i, \frac{M}{16}\right) = \left\{a_i : (a_i, z) \in R_i\left(t_i, G', \frac{M}{16}\right)\right\}, \forall t_i \in T_i;$$
$$\widetilde{R}_{-i}\left(t_{-i}, \frac{M}{16}\right) = R_{-i}\left(t_{-i}, G', \frac{M}{16}\right), \forall t_{-i} \in T_{-i}.$$

We show that $\widetilde{R}\left(\frac{M}{16}\right)$ has the $\frac{M}{16}-$best reply property in G for the type space $T_i \times T_{-i}$. First, for $a_i \in \widetilde{R}_i\left(t_i, \frac{M}{16}\right)$, there is (a_i, z) belonging to $R_i\left(t_i, G', \frac{M}{16}\right)$. Hence, there is a valid conjecture σ_{-i} such that (a_i, z) is a $\frac{M}{16}-$best reply under σ_{-i} for type t_i in G'. Then, a_i is a $\frac{M}{16}-$best reply under σ_{-i} in G, because for any a_i', we have

$$\begin{aligned}
&\int_{\Theta \times T_{-i}} \left[\mathbf{g}_i(a_i, a_i', \theta) \bullet \sigma_{-i}(\theta, t_{-i})\right] \pi_i^*(t_i)\left[(\theta, dt_{-i})\right] \\
&= \int_{\Theta \times T_{-i}} \left[\mathbf{g}_i'([a_i, z], [a_i', z], \theta) \bullet \sigma_{-i}(\theta, t_{-i})\right] \pi_i^*(t_i)\left[(\theta, dt_{-i})\right] \\
&\geq -\frac{M}{16}.
\end{aligned}$$

Second, for $a_{-i} \in \widetilde{R}_{-i}\left(t_{-i}, \frac{M}{16}\right)$, there is a valid conjecture σ_i' such that a_{-i} is a $\frac{M}{16}-$best reply under σ_i' for type t_{-i} in G'. Define a new conjecture σ_i in G such that

$$\sigma_i(\theta, t_i)[a_i] = \sum_{z \in \{z_1, z_1\}} \sigma_i'(\theta, t_i)[a_i, z].$$

Note that by choosing any action a_{-i}', player $-i$ gets exactly the same payoff under the conjecture σ_i' in G' as the payoff under the conjecture σ_i in G. Therefore, a_{-i} remains a $\frac{M}{16}-$best reply under σ_i for type t_{-i} in G. Therefore, $R_{-i}\left(t_{-i}, G', \frac{M}{16}\right) \subseteq R_{-i}\left(t_{-i}, G, \frac{M}{16}\right)$.

Step 2.3 $d^{us}(t_i(n), t_i) \geq \frac{M}{2n}$ *for every* $n \geq 8$.

Consider the following conjecture for player i.

$$\widehat{\sigma}_{-i}(\theta, t_{-i}) \equiv \begin{cases} \delta_{a^1}, & \text{if } t_{-i} = t_{-i}^*; \\ \text{some measurable selection from } R_{-i}(t_{-i}, G, 0), & \text{otherwise.} \end{cases}$$

By step 2.1, $\widehat{\sigma}_{-i}$ is valid in G'. Let a_i be a $0-$best reply under $\widehat{\sigma}_{-i}$ for type $t_i(n)$ in G. We now show that for $n \geq 8$, $(a_i, z_2) \in R_i(t_i(n), G', 0)$ and $(a_i, z_2) \notin R_i\left(t_i, G', \frac{M}{2n}\right)$ to conclude that $d^{us}(t_i(n), t_i) \geq \frac{M}{2n}$.

First, we show that $(a_i, z_2) \in R_i(t_i(n), G', 0)$. Since a_i is a $0-$best reply under $\widehat{\sigma}_{-i}$ for type $t_i(n)$ in G and $\widehat{\sigma}_{-i}$ is valid in G', it remains to verify that

$$\int_{\Theta \times \mathcal{T}_{-i}} [\mathbf{g}_i'([a_i, z_2], [a_i, z_1], \theta) \bullet \widehat{\sigma}_{-i}(\theta, t_{-i})] \, \pi_i^*(t_i(n))[(\theta, dt_{-i})] \geq 0.$$

Recall that $\pi_i^*(t_i(n)) = \mu^n$ and $\pi_i^*(s_i^*)\left[\left\{\theta = 0, t_{-i} = t_{-i}^*\right\}\right] = 1$. Hence, $\pi_i^*(t_i(n))\left[\left\{\theta = 0, t_{-i} = t_{-i}^*\right\}\right] \geq 1/n$ by (28). By our definition of $\widehat{\sigma}_{-i}$,

$$\begin{aligned} & \int_{\Theta \times \mathcal{T}_{-i}} [\mathbf{g}_i'([a_i, z_2], [a_i, z_1], \theta) \bullet \widehat{\sigma}_{-i}(\theta, t_{-i})] \, \pi_i^*(t_i(n))[(\theta, dt_{-i})] \\ \geq\ & \frac{1}{n} \times M + \left(1 - \frac{1}{n}\right) \times 0 - \frac{M}{n} = 0. \end{aligned}$$

Second, we show that $(a_i, z_2) \notin R_i\left(t_i, G', \frac{M}{2n}\right)$ for $n \geq 8$. For any $t_{-i} \in T_{-i}$, since $a^1 \notin R_{-i}\left(t_{-i}, G', \frac{M}{16}\right)$ by step 2.2, we have $a^1 \notin R_{-i}\left(t_{-i}, G', \frac{M}{2n}\right)$ for $n \geq 8$. Since $\pi_i^*(t_i)[T_{-i}] = 1$, any valid conjecture σ_{-i}' which $\frac{M}{2n}-$rationalizes a_i for t_i must satisfy $\sigma_{-i}'(\theta, t_{-i})[a^1] = 0$ for any (θ, t_{-i}) with a positive probability under $\pi_i^*(t_i)$. Thus,

$$\int_{\Theta \times \mathcal{T}_{-i}} \left[\mathbf{g}_i'([a_i, z_2], [a_i, z_1], \theta) \bullet \sigma_{-i}'(\theta, t_{-i})\right] \pi_i^*(t_i(n))[(\theta, dt_{-i})] = 0 - \frac{M}{n} < -\frac{M}{2n}.$$

Therefore, $(a_i, z_2) \notin R_i\left(t_i, G', \frac{M}{2n}\right)$ for $n \geq 8$.■

References

ALIPRANTIS, C., AND K. BORDER (1999): *Infinite Dimensional Analysis.* Berlin: Springer-Verlag.

CHEN, Y.-C., AND S. XIONG (2008): “"Topologies on types": Correction,” *Theoretical Economics*, forthcoming.

DEKEL, E., D. FUDENBERG, AND S. MORRIS (2006): “Topologies on Types,” *Theoretical Economics*, 1, 275–309.

——— (2007): “Interim Correlated Rationalizability,” *Theoretical Economics*, 2, 15–40.

DI TILLIO, A., AND E. FAINGOLD (2007): “Uniform Topology on Types and Strategic Convergence,” mimeo, Yale University.

DUDLEY, R. (2002): *Real Analysis and Probability.* Cambridge University Press.

ELY, J., AND M. PESKI (2007): “Critical Types,” mimeo, Northwestern University.

FAN, K. (1952): “Minimax Theorems,” *Proceedings of the National Academy of Sciences of the United States of America*, 39, 42–47.

HARSANYI, J. (1967/1968): “Games with incomplete information played by "Bayesian" players, I-III,” *Management Science*, 14, 159–182, 320–334, 486–502.

KAJII, A., AND S. MORRIS (1998): “Payoff continuity in incomplete information games,” *Journal of Economic Theory*, 82, 267–276.

LIPMAN, B. L. (2003): “Finite order implications of common priors,” *Econometrica*, 71, 1255–1267.

MERTENS, J.-F., AND S. ZAMIR (1985): “Formulation of Bayesian Analysis for Games with Incomplete Information,” *International Journal of Game Theory*, 14, 1–29.

MONDERER, D., AND D. SAMET (1989): “Approximating common knowledge with common beliefs,” *Games and Economic Behavior*, 1, 170–190.

——— (1996): “Proximity of information in games with common beliefs,” *Mathematics of Operations Research*, 21, 707–725.

Morris, S. (2002): "Typical Types," mimeo, Princeton University.

Rubinstein, A. (1989): "The Electronic Mail Game: the Strategic Behavior under Almost Common Knowledge," *American Economic Review*, 79, 385–391.

Weinstein, J., and M. Yildiz (2007): "A structure theorem for rationalizability with application to robust predictions of refinements," *Econometrica*, 75, 365–400.