A STRATEGIC TOPOLOGY ON INFORMATION STRUCTURES

By

Dirk Bergemann, Stephen Morris, and Rafael Veiel

November 2024

COWLES FOUNDATION DISCUSSION PAPER NO. 2413

COWLES FOUNDATION FOR RESEARCH IN ECONOMICS

YALE UNIVERSITY
Box 208281
New Haven, Connecticut 06520-8281

http://cowles.yale.edu/

A Strategic Topology on Information Structures*

Dirk Bergemann[†] Stephen Morris[‡] Rafael Veiel[§]

November 7, 2024

Abstract

Two information structures are said to be close if, with high probability, there is approximate common knowledge that interim beliefs are close under the two information structures. We define an "almost common knowledge topology" reflecting this notion of closeness. We show that it is the coarsest topology generating continuity of equilibrium outcomes. An information structure is said to be simple if each player has a finite set of types and each type has a distinct first-order belief about payoff states. We show that simple information structures are dense in the almost common knowledge topology and thus it is without loss to restrict attention to simple information structures in information design problems.

KEYWORDS: strategic topology, approximate common knowledge, information design.

JEL CLASSIFICATION: C72

*We acknowledge financial support through NSF Grant SES 2049744. We are grateful for comments of seminar participants at the Cowles Foundation Economic Theory conference, the Warwick economic theory conference, the Panorama of Mathematics II at the Hausdorff Center.

[†]Department of Economics, Yale University, dirk.bergemann@yale.edu

[‡]Department of Economics, Massachusetts Institute of Technology, semorris@mit.edu

[§]Department of Economics, Massachusetts Institute of Technology, veiel@mit.edu

1 Introduction

1.1 Motivation and Results

Players make choices in a game where payoffs depend on some unknown state of the world. Optimal strategic behavior will then depend on players' beliefs about the states, their beliefs about others' beliefs, and so on. The (common prior) information structure of the players is then a probability distribution over the players' beliefs and higher-order beliefs about the state. We are interested in how the information structure impacts equilibrium outcomes in such a game. It is known that equilibrium outcomes are sensitive to the infinite tails of higher-order beliefs (Rubinstein (1989) and Carlsson and Van Damme (1993)). Our main result is a characterization of the coarsest topology on information structures generating continuity of equilibrium outcomes.

Following Monderer and Samet (1989), an event is said to be *common p-belief* if everyone believes it with probability at least p, everybody believes with probability at least p that everyone believes it with probability at least p, and so on. An event is said to be *approximate common knowledge* if it is common p-belief with p close to 1. Now two information structures are close in the almost common knowledge topology if, with high ex ante probability, there is approximate common knowledge that their higher-order beliefs are close (in the product topology). Our main result (Theorem 1 in Section 1) establishes that the almost common knowledge topology is the coarsest topology generating continuity of equilibrium outcomes.

Our definition of information structures excludes any (payoff-irrelevant) correlating devices that players might have access to. In the language of Mertens and Zamir (1985), we restrict attention to *non-redundant* information structures. However, our main result allows the correlating device to appear within the equilibrium solution concept: we study *belief-invariant Bayes correlated equilibria* where players' action choices can be correlated but only when the correlation does not alter players' beliefs and higher-order beliefs about the state. While we think our notions of information structures and equilibrium are the most natural for our main exercise, we also spell out (in Section 4) how our results change if we allow more general information structures with redundancies and Bayes Nash Equilibrium (which rules out correlation in the solution concept).

We say that an information structure is simple if its support is finite and each type of a player has a distinct first-order belief. We show that simple information structures are dense. One implication is that it is without loss to focus on simple information structures in solving such information design problems.

1.2 Alternative Topologies and Related Literature

Monderer and Samet (1996) and Kajii and Morris (1998), building on Monderer and Samet (1989), identified topologies on information structures that generated continuity of equilibrium outcomes. And the topologies identified both have the same flavor as the almost common knowledge topology. The key difference is that both papers fix (different) sets of information structures with the restriction that each profile of types gives rise to a distinct state, and is not associated with higher-order beliefs. This makes the topology hard to interpret. It also makes both directions of the proof easier than in our problem. We describe, following our main result (Theorem 1 in Section 1), which steps of our proof relate to this early work and which are novel to this paper. We discuss the join measurability condition implicit in Monderer and Samet (1996) and Kajii and Morris (1998) in Section 6.3. Other differences are that the earlier work focuses on Bayes Nash equilibrium and restricts attention to countable information structures to ensure existence; we allow uncountable information structures (sets of the universal type space) and we ensure equilibrium existence by incorporating redundancies in our solution concept. But we also show how our results apply to Bayes Nash equilibrium and general information structures in Section 4. The earlier work measures closeness of equilibrium outcomes by players' expected payoffs in equilibrium; we note in Section 6.2 that our results would be unchanged if we used this approach.

Dekel et al. (2006) defined an interim strategic topology on hierarchies of beliefs under the solution concept of interim correlated rationalizability (ICR). Two belief hierarchies were said to be close in the interim strategic topology if, in any game, an action that was ICR at one hierarchy was approximately ICR at the other hierarchy. It is well-known (e.g., from the work of Rubinstein (1989), Carlsson and van Damme (1993) and Weinstein and Yildiz (2007)) that closeness in the product topology is not sufficient for close strategic behavior. Chen et al. (2017) characterized the interim strategic topology of Dekel et al. (2006) and showed that it requires closeness of beliefs about some tail events ("frames"). However, the connection to approximate common knowledge has been unclear. We show that two information structures are close in our almost common knowledge topology if and only if there is a high ex ante probability that hierarchies are close in the interim strategic topology.

2 Model

This section will introduce the model used in our main analysis, with four main ingredients.

1. We will hold fixed a finite set of players, a finite set of (payoff-relevant) *states* and a probability distribution over the states.

2. We define a *base game* to consist of players' possible actions and their payoff functions; i.e., how each player's payoff depends on the action profile chosen and the state.

3. An *information structure* will consist of a probability distribution over the state and players' beliefs and higher-order beliefs about the state, with the appropriate marginal over states.

4. The equilibrium solution concept will be the *belief-invariant Bayes correlated equilibrium* (BIBCE), see Definition 8 in Bergemann and Morris (2016). This is a joint distribution over the information structure and the players' actions such that players'

actions are best responses and measurable with respect to players' beliefs and higher-order beliefs about the state, but allowing payoff-irrelevant correlation of actions.

Our model of an information structure is restrictive in that it rules out players observing multiple signals giving rise to the same beliefs and higher-order beliefs. Equivalently, it rules out what Mertens and Zamir (1985) labelled *redundant* types, i.e., players observing payoff-irrelevant signals through which they can correlate their behavior. On the other hand, our equilibrium solution concept allows players to observe payoff-irrelevant correlating devices in choosing actions. Thus we have made a modelling choice to put correlation devices in the solution concept rather than the information structure.

In order to relate our work to the literature and to applications, we will later discuss (in Section 4) how our results easily translate to a setting with general information structures (allowing *redundancies*) and the more relaxed solution concept of Bayes Nash equilibrium, i.e., if we move correlation possibilities from the solution concept to the information structure.

2.1 Setting and Base Game

There is a finite set of players I and a finite set of (payoff) states Θ. Throughout the paper, we will fix a prior $\mu \in \Delta(\Theta)$. Without loss, we will maintain the assumption that μ has full support.

A *base game* then describes players' actions and payoffs: thus a base game is a tuple $\mathcal{G} = ((A_i)_{i \in I}, (u_i)_{i \in I})$, where A_i is a finite set of actions for player i and $u_i : A_i \times A_{-i} \times \Theta \to [-M, M]$ is a payoff function for player i, where $A_{-i} := \prod_{j \neq i} A_j$ and $M > 0$ is an exogenous payoff bound.

2.2 Information Structures

We follow Harsanyi (1967-68) and Mertens and Zamir (1985) in identifying signals or types with players' beliefs and higher-order beliefs, or hierarchies of beliefs. We first formally define the set of hierarchies of beliefs.

Definition 1. (Hierarchies of Beliefs) The profile of players' hierarchies of beliefs, $(\mathcal{T}_i)_{i\in I}$ is defined recursively as follows: For every i, let $\mathcal{T}_i^0 := \{*\}$ be a singleton and let $\mathcal{T}_i^1 := \Delta(\Theta)$. Given profiles $(\mathcal{T}_i^{m-1})_{i\in I}$ for $m > 1$, define for every i,

$$\mathcal{T}_i^m := \left\{ \left(\left(\tau_i^1, \ldots, \tau_i^{m-1} \right), \tau_i^m \right) \in \mathcal{T}_i^{m-1} \times \Delta(\mathcal{T}_{-i}^{m-1} \times \Theta) : \mathrm{marg}_{\Theta \times \mathcal{T}_{-i}^{m-2}}(\tau_i^m) = \tau_i^{m-1} \right\}.$$

Let $\mathcal{T}_i$ denote the set of sequences $\tau_i = (\tau_i^m)_m$ so that for every $\overline{m} \in \mathbb{N}$, the truncated sequence $(\tau_i^m)_{m\leq\overline{m}}$ belongs to $\mathcal{T}_i^{\overline{m}}$.

For simplicity, we will follow Mertens and Zamir (1985) in imposing the product topology on hierarchies $\mathcal{T}_i$. Let d_Π be a metric on $\Omega := \Theta \times \mathcal{T}$ inducing the product topology on $\mathcal{T} := \prod_{i\in I} \mathcal{T}_i$ and the discrete topology on Θ.[1] However, we later discuss (Section 6.1) why the use of the product topology on hierarchies is just for convenience and not important for our arguments. Mertens and Zamir (1985) show that for every $(\tau = (\tau_i)_{i\in I}, \theta) \in \Omega$ and every $i \in I$ there is a unique belief $\tau_i^* \in \Delta(\mathcal{T}_{-i} \times \Theta)$ so that for all $m \in \mathbb{N}$, $\tau_i^m = \mathrm{marg}_{\mathcal{T}_{-i}^{m-1}\times\Theta}(\tau_i^*)$, where $\tau \mapsto \tau^* = (\tau_i^*)_{i\in I}$ is a homeomorphism. Let $\mathscr{B}$ denote the Borel sigma-algebra on Ω.

We will refer to Ω as the "universal state space", with typical element $\omega = (\tau, \theta)$, where $\tau \in \mathcal{T}$. Now an information structure is just a probability distribution on the universal state space that respects the prior on states and the labelling of hierarchies.

Definition 2. (Information Structure). An information structure P is a Borel probability measure on Ω that satisfies two conditions: (i) [consistency] the marginal of P on Θ is μ;

[1] The product topology on $\mathcal{T}$ is the coarsest topology so that projections $\mathrm{proj}_{\mathcal{T}^m} : \mathcal{T} \to \mathcal{T}^m$ are continuous, where for every m, $\mathcal{T}^m$ is endowed with the weak topology. Dekel et al. (2006) provide a metric that induces the product topology on $\mathcal{T}$, which for any discount factor $\eta \in (0, 1)$ can be described by $d_\Pi((\theta, \tau), (\hat{\theta}, \hat{\tau})) = \mathbf{1}_{\theta\neq\hat{\theta}} + \sum_{n=1}^\infty \eta^n \max_i d_w^n(\tau_i^n, \hat{\tau}_i^n)$, where d_w^n is a product metric inducing the weak topology on $\mathcal{T}_i^n$.

(ii) [labelling] for every player i, there is a version of the conditional probability $P_i\colon \mathcal{T}_i \to \Delta(\Theta \times \mathcal{T}_{-i})$ of P so that

$$\tau_i^* = P_i(\tau_i), \;\; P\text{-a.s.}$$

We will write $\mathcal{P} \subseteq \Delta(\Omega)$ for the set of (common prior) information structures. We will sometimes describe these as *non-redundant* information structures because we do not allow multiple types or signals of a player giving rise to the same hierarchies of beliefs. This is to contrast them with the more general *redundant* information structures discussed in Section 4.

An information structure is *finite* if it has finite support. A hierarchy $\tau_i \in \mathcal{T}_i$ is *finite* if it is in the support of a finite information structure. The set of finite types in $\mathcal{T}_i$ is denoted by $\mathcal{T}_i^0$ and the set of finite states $\Omega^0 \subseteq \Omega$ is given by

$$\Omega^0 := \left\{(\tau, \theta) \in \Omega : \forall\, i, \;\; \tau_i^* \in \mathcal{T}_i^0\right\}.$$

2.3 Solution Concept: Belief-Invariant Bayes Correlated Equilibrium

Together, a base game $\mathcal{G}$ and an information structure P define a game of incomplete information $(\mathcal{G}, P)$. Now we define our main equilibrium solution concept. We will be allowing players' action choices to be correlated. So players' action choices will be described by a *decision rule,* mapping states and hierarchies to action profiles; thus for any incomplete information game $(\mathcal{G}, P)$, a decision rule is a measurable map $\sigma : \Theta \times \mathcal{T} \to \Delta(A)$, where $A := \prod_{i \in I} A_i$ and $\Delta(A)$ is endowed with the Euclidean topology.

An information structure P and decision rule σ jointly induce a measure $\sigma \circ P \in \Delta(A \times \Theta \times \mathcal{T})$ in the natural way. We will be interested in *outcomes* specifying a joint distribution over actions and states $\nu \in \Delta(A \times \Theta)$. Decision rule σ *induces* outcome ν_σ if ν_σ is the marginal of $\sigma \circ P$ on $A \times \Theta$. For every player i, a decision rule σ and hierarchy $\tau_i \in \mathcal{T}_i$ induce a belief $\sigma \circ \tau_i \in \Delta(A \times \Theta \times \mathcal{T}_{-i})$, which for every measurable set $E \subseteq \Theta \times \mathcal{T}_{-i}$ and action profile $a \in A$ satisfies $\sigma \circ \tau_i(\{a\} \times E) = \int_E \sigma(a|\theta, \tau_{-i}, \tau_i)\ \mathrm{d}\sigma \circ \tau_i^*$.

Our notion of equilibrium will be (one version of) incomplete information correlated equilibrium. Two properties will be key.

Definition 3. A decision rule σ is *ε-obedient* if for every player i, action a_i and deviation a_i',

$$\int \sum_{a_{-i} \in A_{-i}} (u_i(a_i, a_{-i}, \theta) - u_i(a_i', a_{-i}, \theta))\ \mathrm{d}\sigma \circ \tau_i(a_i, a_{-i}, \tau_{-i}, \theta) > -\varepsilon,\ a.s.$$

Thus if we interpret the decision rule $\sigma : \Theta \times \mathcal{T} \to \Delta(A)$ as a recommendation rule for a mediator in the game, ε-obedience requires that players will not lose more than ε by following the recommendation.

Importantly, this is an interim version of ε-obedience: we are requiring that the decision rule is almost always approximately optimal.[2]

Definition 4. A decision rule σ is *belief-invariant* if, for every $a_i \in A_i$, the marginal probability $\sigma(a_i \times A_{-i}|(\tau_i, \tau_{-i}), \theta) = \sigma_i(a_i|\tau_i)$ does not depend on (τ_{-i}, θ).[3]

Definition 5. A decision rule σ is an ε-belief-invariant Bayes correlated equilibrium (ε- BIBCE) of $(\mathcal{G}, P)$ if it satisfies ε-obedience and belief invariance.

We will say that a decision rule is a BIBCE if it is a 0-BIBCE.

[2]An ex-ante notion of ε-obedience would require interim ε-obedience only with probability $1-\varepsilon$, and thus allows players to choose actions that are not ε best responses with probability $1-\varepsilon$. The ex-ante notion of ε-obedience is much more permissive: Engl (1995) has established the lower hemi-continuity of ex ante ε-Bayes Nash equilibrium in games of incomplete information with respect to weak convergence of priors on a fixed type space.

[3]Lehrer et al. (2010) and Forges (2006) introduced the notion of belief-invariance to the study of incomplete information correlated equilibrium: Forges (1993) (implicitly) and Lehrer et al. (2010) and Forges (2006) (explicitly) imposed the restriction in their definitions of the belief-invariant Bayesian solution. This latter solution imposes (implicitly or explicitly) join feasibility, so that the decision rule depends on the prolile of players' signals. Liu (2015) showed that a subjective version of BIBCE is equivalent to interim correlated rationalizability. A belief-invariant decision rule can be interpreted as a payoff-irrelevant correlating device (where player i observes private signal a_i). Now BIBCE is equivalent to Bayes Nash equilibria if players are allowed to observe a correlated device.

3 Main Result

In this section, we first define an "approximate common knowledge" topology on information structures and then show it is the coarsest topology to generate continuity of BIBCE outcomes (our main result).

3.1 The Approximate Common Knowledge Topology

In words, we will say that two information structures are close if each assigns high ex ante probability to the event that there is approximate common knowledge that their interim beliefs are close. An event is "approximate common knowledge", if for some p close to 1, everyone believes the event with probability at least p, everyone believes with probability at least p that everyone believes the event with probability at least p, and so on.... (Monderer and Samet (1989)). The subtlety in formalizing this definition is how we define the event that interim beliefs are close.

We first define ε-neighborhoods, the set of universal states ε-close to a given universal state $\omega \in \Omega$.

Definition 6. For every $\varepsilon > 0$ and $\omega \in \Omega$, the ε-neighborhood of ω is given by

$$\mathcal{N}_\varepsilon(\omega) := \{\omega' \in \Omega : d_\Pi(\omega, \omega') < \varepsilon\}.$$

For any $\varepsilon > 0$, we define the ε-support of $P \in \mathcal{P}$, to be the ε-neighborhoods of universal states whose ε-neighborhoods are assigned positive probability by P:

$$\text{supp}_\varepsilon(P) := \bigcup_{\omega \in \Omega : P(\mathcal{N}_\varepsilon(\omega)) > 0} \mathcal{N}_\varepsilon(\omega).$$

Thus $\text{supp}_\varepsilon(P)$ is an ε-expansion of the support of P. For any two priors P, P', we define the intersection of their ε-supports:

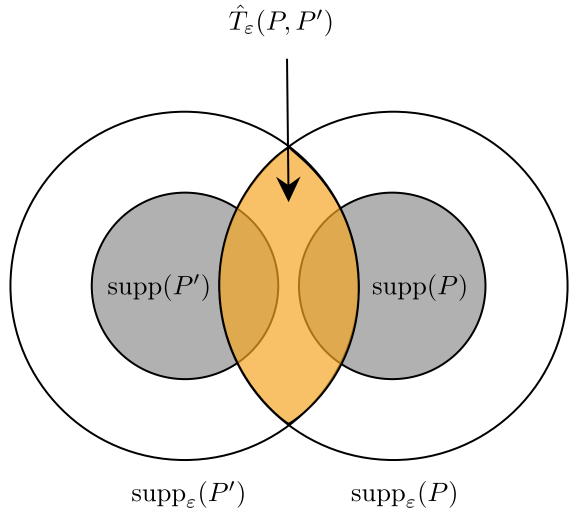

Figure 1: Sets $\hat{T}_\varepsilon(P, P')$ and ε-supports of P and P'.

$$\hat{T}_\varepsilon(P, P') := \text{supp}_\varepsilon(P) \cap \text{supp}_\varepsilon(P').$$

This is the set of universal states where interim beliefs are close to interim beliefs in the support of both priors. Following Monderer and Samet (1989), for every probability $p \in [0, 1]$ and event $E \in \mathscr{B}$, we define $B^p(E)$ to be the set of universal states where all players assign probability at least p to the event E. Thus

$$B^p(E) := \{(\tau, \theta) \in E : \forall i, \tau_i^*(E_{-i}) \geq p\}$$

where $E_{-i} := \text{proj}_{\Theta \times \mathcal{T}_{-i}}(E)$. For every $m \in \mathbb{N}$, let $[B^p]^m(E) := B^p \circ \cdots \circ B^p(E)$ denote the m-fold application of B^p; $[B^p]^m(E) := B^p \circ \cdots \circ B^p(E)$ is the set of universal states where all players assign probability at least p to all players assigning at least probability p.... (m times) to event E being true. Now an event is said to be *common p-belief* if this is true for all m.

Definition 7. (Common p-Belief) For any event $E \in \mathscr{B}$, the event that E is *common p-belief*, $C^p(E)$, is defined as

$$C^p(E) := \bigcap_{m \in \mathbb{N}} [B^p]^m (E).$$

An event is said to be approximate common knowledge if it is common p-belief for p close to 1. We can now define the approximate common knowledge (ACK) distance[4] between any two priors, $P, P' \in \mathcal{P}$.

Definition 8. (Approximate Common Knowledge Distance) For every $P, P' \in \mathcal{P}$, let

$$d^{ACK}(P, P') := \inf \left\{ \varepsilon \geq 0 : \begin{array}{c} P\left(C^{1-\varepsilon}\left(\hat{T}_\varepsilon(P, P')\right)\right) \geq 1-\varepsilon \\ P'\left(C^{1-\varepsilon}\left(\hat{T}_\varepsilon(P, P')\right)\right) \geq 1-\varepsilon \end{array} \right\}.$$

Thus two priors are close if, under both priors, there is high probability that there is approximate common knowledge that their interim beliefs are close.

Definition 9. (Approximate Common Knowledge Topology) The approximate common knowledge (ACK) topology is the topology generated by open sets $\left\{P' \in \mathcal{P} : d^{ACK}(P, P') < \varepsilon\right\}_{P \in \mathcal{P}}$.

While d^{ACK} fails the triangle inequality, the topology is metrizable.

Proposition 1. *The ACK topology is metrizable.*

The proof (in the Appendix) constructs a metric that generates the same topology. Because the topology is metrizable, it is characterized by the convergent sequences.

Definition 10. A sequence $(P^k)_k$ in $\mathcal{P}$ ACK-converges to P if and only if $d^{ACK}\left(P^k, P\right) \to 0$.

It is useful to consider what the topology looks like in some special cases. If there is only one payoff relevant state (i.e., $|\Theta| = 1$), then there is a unique universal state, where there is common knowledge of the state. If there is only one player, then an information structure is given by a distribution over common first order beliefs about Θ, and the topology reduces to

[4] A distance is a map $d : \mathcal{P} \times \mathcal{P} \to [0, \infty)$ which is zero on the diagonal.

the weak topology on $\Delta(\Delta(\Theta))$. Similarly, if there are many players but types are perfectly correlated. If types are independent, then a canonical information structure is given by a distribution in $\Delta(\Delta(\Theta)^I)$, and the topology again reduces to the weak topology. If we restrict to information structures with a fixed, finite supports, then an information structure is a probability distribution on a finite set and the topology reduces to the weak topology. So in order for the ACK topology to be interesting, there must be at least two states, at least two players, an unbounded number of types and neither independence or perfect correlation.

3.2 Continuity of Equilibrium Outcomes

We now define continuity of equilibrium outcomes (for BIBCE). For every $\varepsilon > 0$ and $(\mathcal{G}, P)$, let $\mathcal{B}^\varepsilon(\mathcal{G}, P)$ denote the collection of all ε-BIBCE under $(\mathcal{G}, P)$. Define the set of outcomes (i.e., elements of $\Delta(A \times \Theta)$) that are induced by a ε-BIBCE of $(\mathcal{G}, P)$

$$\mathcal{O}^\varepsilon(\mathcal{G}, P) := \{\nu_\sigma : \sigma \in \mathcal{B}^\varepsilon(\mathcal{G}, P)\}.$$

We write $\mathcal{O}_\varepsilon(\mathcal{G}, P)$ for the set of outcomes that are within ε of outcomes induced by an ε-BIBCE of $(\mathcal{G}, P)$, so

$$\mathcal{O}_\varepsilon(\mathcal{G}, P) := \{v \in \Delta(A \times \Theta) : \exists\, \sigma \in \mathcal{B}^\varepsilon(\mathcal{G}, P) \text{ s.t. } ||v - v_\sigma||_2 < \varepsilon\},$$

where $||\cdot||_2$ is the Euclidean norm on outcomes. Now we will say that two information structures are strategically close for base game $\mathcal{G}$ if the sets of BIBCE outcomes are close.

Definition 11. (Strategic Distance) For every $P, P' \in \mathcal{P}$ and base game $\mathcal{G}$, let

$$d^*(P, P'|\mathcal{G}) := \inf\left\{\varepsilon \geq 0 : \begin{array}{c} \mathcal{O}^0(\mathcal{G}, P) \subseteq \mathcal{O}_\varepsilon(\mathcal{G}, P') \\ \mathcal{O}^0(\mathcal{G}, P') \subseteq \mathcal{O}_\varepsilon(\mathcal{G}, P) \end{array}\right\}.$$

A topology on $\mathcal{P}$ generates continuity of equilibrium outcomes if for every $\varepsilon > 0$, every base game $\mathcal{G}$ and every $P \in \mathcal{P}$, the set $\{P' \in \mathcal{P} : d^* (P, P'|\mathcal{G}) < \varepsilon\}$ is open.

3.3 Main Result

The following is our main result.

Theorem 1. *The approximate common knowledge topology is the coarsest topology on $\mathcal{P}$ that gives rise to continuity of equilibrium outcomes.*

We will first sketch the main ideas in the proof, distinguishing between steps that are inherited from earlier work, and how, and which steps are and must be novel. To establish that approximate common knowledge convergence implies strategic convergence (i.e. that the ACK topology generates continuity of strategic outcomes), one must show that if two information structures are close enough in the approximate common knowledge topology, then in any game, the set of ε-BIBCE outcomes will be close. This builds on the arguments in Monderer and Samet (1996) and Kajii and Morris (1998) (which in turn build on an argument in Monderer and Samet (1989)). These two papers fix an equilibrium strategy profile under one information structure and show that there is an approximate equilibrium in any nearby information structure where the strategy profile is unchanged on the event where there is approximate common knowledge that interim beliefs are close, but can vary arbitrarily elsewhere. This proof strategy relies on a well-defined notion of holding the strategy profile fixed on the approximate common knowledge event. In our context, there is no well-defined notion of holding the strategy profile fixed on the approximate common knowledge event. In particular, if the two information structures are minimal (they do not have common knowledge subsets), the supports of distinct information structures will be disjoint, as illustrated in Figure 2. So instead we will continuously extend the equilibrium decision rule σ under one information structure P to its ε-support $\text{supp}_\varepsilon(P)$ and thus to the event:

$$\hat{T}_\varepsilon(P, P') := \text{supp}_\varepsilon(P) \cap \text{supp}_\varepsilon(P'),$$

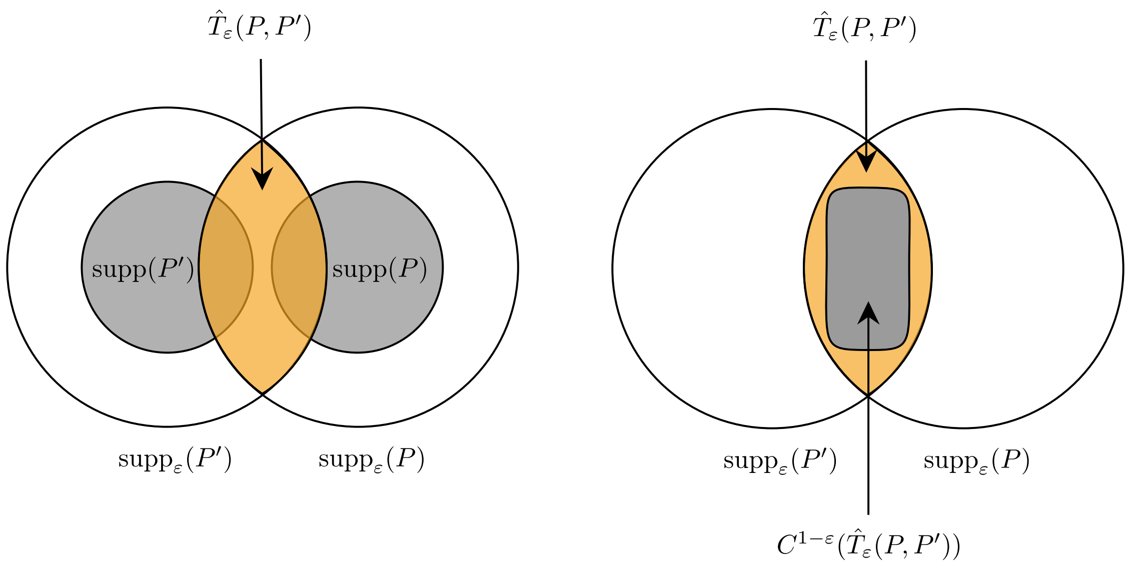

Figure 2: Sets $\hat{T}_\varepsilon(P,P')$,$C^{1-\varepsilon}(\hat{T}_\varepsilon(P,P'))$ and ε-supports of P and P'.

and thus the approximate common knowledge event:

$$C^{1-\varepsilon}\left(\hat{T}_\varepsilon(P,P')\right) \subseteq \hat{T}_\varepsilon(P,P').$$

This is one place where we are exploiting properties of the product topology on types to show that the extended decision rule allows us to find an approximate equilibrium under P' where the continuous extension of σ on $C^{1-\varepsilon}\left(\hat{T}_\varepsilon(P,P')\right)$ is held fixed. This argument goes through with any refinement of the product topology as we will discuss later.

To establish that strategic convergence implies approximate common knowledge convergence, it is enough to show that, for any two information structures that are not close in the approximate common knowledge topology, one can construct a game where an equilibrium outcome under one information structure is not close to any approximate equilibrium under the other information structure. Monderer and Samet (1996) and Kajii and Morris (1998) do this by showing that if two information structures P and P' are not close, there is an "infecting" event D^0 under one of the information structures, say P, such that there is no approximate common knowledge event on the complement of $C^{1-\varepsilon}\left(\hat{T}_\varepsilon(P,P')\right)$. One can then construct a binary coordination game in the spirit of the email game of Rubinstein (1989) where there is a unique equilibrium under P because there is a dominant strategy on the "infecting" event, but there are multiple equilibria under P'. This strategy is not available to us because we cannot assume that there is a dominant strategy on the "infecting" event because payoffs must be measurable with respect to the payoff states Θ.

Instead, we show that if two information structures are not close, there exists an "infecting" event D^m that is measurable with respect to m-order beliefs. For m large enough, this event has the property that some types in the support of P that are not in $\hat{T}_\varepsilon(P, P')$ are closer to any element in D^m than to any element in $\hat{T}_\varepsilon(P, P')$. Hence, for m large enough, some types can be excluded from $C^{1-\varepsilon}\left(\hat{T}_\varepsilon(P, P')\right)$ based on their beliefs on a m-order-measurable event. From event $D_\varepsilon^0 = D^m$ we obtain a cover $(D_\varepsilon^n)_{n\in\mathbb{N}}$ of the complement of $C^{1-\varepsilon}\left(\hat{T}_\varepsilon(P, P')\right)$ recursively, where

$$D_\varepsilon^n = \text{supp}(P) \setminus B^{1-\varepsilon}(D_\varepsilon^{n-1}),$$

for all $n \geq 1$. Now we can construct a first game where players are incentivized to announce their approximate m-th order beliefs on a finite grid, (for m large enough chosen as a function of ε) and build an email game style binary action coordination game on top of that. For the first component of the game, we can construct a game with an iterated scoring rule with the property that it is ε-rationalizable for players to truthfully report finite-order beliefs, from a finite grid which are closest in d_Π. This scoring rule is also used in Dekel et al. (2006) and Gossner and Mertens (2020). This game alone cannot be used to induce different outcomes in P and P' since on the set $\hat{T}_\varepsilon(P, P')$, reporting the same approximate finite-order beliefs is an ε-BIBCE for the types in both priors. As we cannot a priori rule out the possibility of $\hat{T}_\varepsilon(P, P')$ containing the support of P' the scoring rule is not suitable for separating outcomes. Players therefore also choose an additional action: Either action zero or action one. No matter what additional action is chosen by the opponent, action one is the

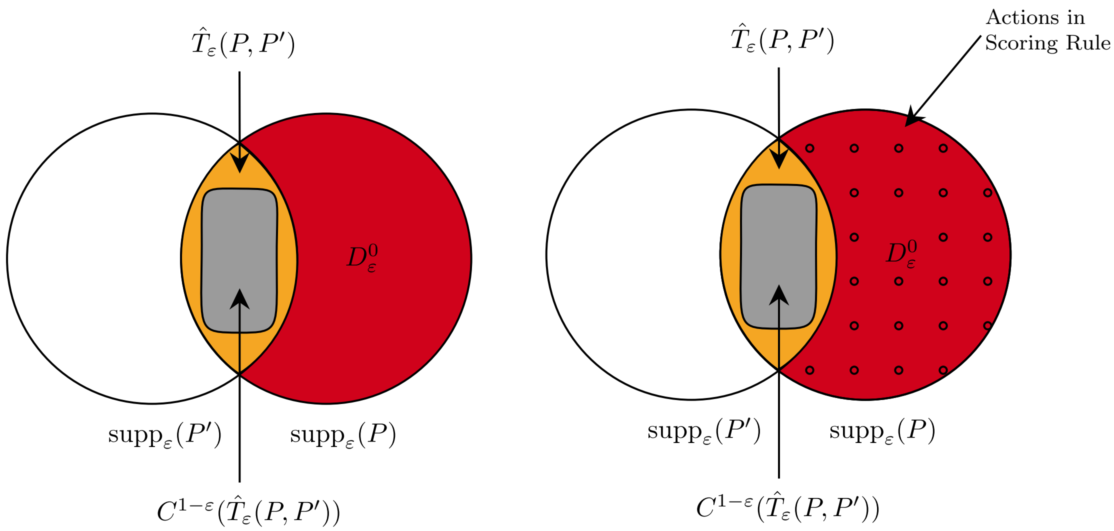

Figure 3: Infection argument.

unique, strict best-reply for players who reported themselves to be in D^0_ε. All other types will match action one if they believe with probability at least ε that their opponent also chose action one. Hence iterative deletion of dominated strategies implies that action one is played on D^n_ε. Figure 3 illustrates the region D^0_ε and the role of the scoring rule. Type profiles in the red region will play action one and an action represented by a circle in the right panel of the Figure. No type in the support of P' will play an action corresponding to a type in D^0_ε. Action one will infect all type profiles in the orange shaded region who are also in the support of P but will not infect types in the support of P'.

Hence, in every ε-BIBCE the types in orange and red shaded regions who are in the support of P will play action 1. There is a ε-BIBCE where all types play action 0 under P'. This establishes that the sets of outcomes of the priors are at least ε apart. In this second part of the proof, we exploit properties of the product topology to ensure that we can cover the red region D^0_ε with a finite grid and there is a finite game where players find it ε-optimal to report the closest element in the grid.

We now report two Lemmas before proceeding with the formal proof. Let $\mathcal{N}_\delta(E) := \bigcup_{\hat{\omega}\in E}\{\omega \in \Theta \times \mathcal{T} : d_\Pi(\hat{\omega},\omega) < \delta\}$ denote the δ-ball around an event $E \in \mathscr{B}$.

Lemma 1. *For every $\varepsilon > 0$ and any event $E \in \mathscr{B}$, there is a finite event $G_\varepsilon \subseteq \Omega \setminus \mathcal{N}_\varepsilon(E)$ so that for every $\omega \in \Omega$,*

$$\min_{\omega_g \in G_\varepsilon} d_\Pi(\omega, \omega_g) < \varepsilon \implies \omega \in \Omega \setminus E$$

and

$$\omega \in \Omega \setminus \mathcal{N}_\varepsilon(E) \implies \min_{\omega_g \in G_\varepsilon} d_\Pi(\omega, \omega_g) < \varepsilon.$$

Lemma 1 implies that no type in the support of P' will be ε-close (in the product topology) to a grid element in the red shaded region D^0_ε and every type in the support of P that is also in the red shaded region is ε-close (in the product topology) to a grid element in the red shaded region D^0_ε. Lemma 2 below then establishes that there exists a finite game, where it is uniquely ε-rationalizable for every type to report the closest grid element (which is also less than ε-close).

Lemma 2. *For every $\varepsilon > 0$, there is a finite set of action profiles $A \subseteq \mathcal{T}$ and payoffs $u^\varepsilon_i \colon A \times \Theta \to \mathbb{R}$ for every player i, so that for every information structure P, every ε-obedient decision rule σ satisfies*

$$\sigma(a|\theta, \tau) > 0 \implies a \in \arg\min_{a \in A} d_\Pi((\theta, \tau), (\theta, a)) \textit{ and } d_\Pi((\theta, \tau), (\theta, a)) < \varepsilon.$$

We now state two known results which we also use in our proof. The existence of BIBCE was established by Stinchcombe (2011), who also established it for general information structures, with redundancies as we later report in Proposition 7.

Proposition 2. *(Existence of BIBCE)* There exists a BIBCE for every $(\mathcal{G}, P)$.

Theorem A in Stinchcombe (2011) has established the existence of a more demanding notion of incomplete information correlated equilibrium, when one looks at correlated equilibrium of the agent normal form of the game of incomplete information. Forges (1993) called these "agent normal form correlated equilibria." Since every agent normal form correlated equilibrium induces an outcome equivalent BIBCE, existence of equilibria is guaranteed. We first establish that ACK-convergence is sufficient for continuity of equilibrium outcomes.

Proposition 3. *The set $\Omega^0 \subseteq \Omega$ is dense in the product topology on Ω.*

Proof. Lipman (2003) shows that finite common prior types are dense in the product topology on Ω. $\square$

Proposition 4. *For every base game $\mathcal{G}$ and every $\varepsilon > 0$, there is $\delta > 0$ so that if $d^{ACK}(P, P') < \delta$ then $d^*(P, P'|\mathcal{G}) < \varepsilon$.*

Proof. Let $\Omega_P := \cap_{\epsilon>0}\text{supp}_\varepsilon(P)$, $\Omega_{P'} := \cap_{\epsilon>0}\text{supp}_\varepsilon(P')$ and $\hat{\Omega}_\delta := C^{1-\delta}(\hat{T}_\delta(P, P'))$. It is without loss of generality to assume that $\Omega_P \cap \Omega_{P'} = \varnothing$. Fix a base game with payoffs given by $u_i \colon A \times \Theta \to [-M, M]$ for each player i. Suppose $d^{ACK}(P, P') < \delta = \frac{\varepsilon}{6M(u)}$, where $M(u) := \max\{M, |A \times \Theta|\} < \infty$ and recall that game $2M \geq \max_i \sup_{a_i, a_i', a_{-i}, \theta} |u_i(a_i, a_{-i}, \theta) - u_i(a_i', a_{-i}, \theta)|$. By Lemma 1, there is a finite set G_δ so that $\hat{\Omega}_\delta \subseteq \mathcal{N}_\varepsilon(G_\delta)$ and so for every $\omega \in \hat{\Omega}_\delta$, $\min_{g \in G_\delta} d_\Pi(\omega, g) \leq \delta$. Let $\zeta_\delta \colon \hat{\Omega}_\delta \to G_\delta$ be any map satisfying $d_\Pi(\omega, \zeta_\delta(\omega)) \leq \delta$, $\forall\ \omega \in \hat{\Omega}_\delta$. We show that every decision rule which is obedient under P, admits a decision rule arbitrarily close to it which is $6M\delta$-obedient under P'. Consider a decision rule $\sigma \colon \Omega_P \to \Delta(A)$ satisfying for every τ_i and for every $a_i' \in A_i$,

$$\int_{\Omega_P} \sum_{a \in A} \Delta u_i(a, a_i', \omega_\theta) \sigma(a|\omega) P(\mathrm{d}\omega|\tau_i) \geq 0,$$

where $\Delta u_i(a, a_i', \theta) = u_i(a, \theta) - u_i(a_i', a_{-i}, \theta)$ and $\omega_\theta = \text{proj}_\Theta(\omega)$. We have existence of such a decision rule from Proposition 2. We now extend σ to $\hat{\Omega}_P$ as follows: For every $\omega \in \hat{\Omega}_\delta \cup \Omega_P$, define the δ-extension

$$\sigma_\delta(\omega) := \begin{cases} \int \sigma(\omega')\ P(\mathrm{d}\omega'|\zeta_\delta(\omega)) & \text{if } \omega \in \hat{\Omega}_\delta, \\ \sigma(\omega) & \text{if } \omega \notin \hat{\Omega}_\delta. \end{cases}$$

For all i and $\tau \in \text{proj}_{\mathcal{T}}(\Omega_P)$,

$$\left| \int_{\hat{\Omega}_\delta \cup \Omega_P} \sum_{a \in A} \Delta u_i(a, a_i', \omega_\theta) \left(\sigma_\delta(a|\omega) - \sigma(a|\omega)\right) \; \tau_i(\mathrm{d}\omega) \right|$$

$$\leq \sum_{g \in G_\delta, a \in A} \Delta u_i(a, a_i', \omega_\theta))$$

$$\left| \left(\int \left(\int \sigma(a|\omega') \; P(\mathrm{d}\omega'|\zeta_\delta^{-1}(g), \tau_i') \right) \; P(\mathrm{d}\tau_i'|\zeta_\delta^{-1}(g)) \right) - \int \sigma(a|\omega') \; \tau_i(\mathrm{d}\omega|\zeta_\delta^{-1}(g)) \right| \tau_i(\zeta_\delta^{-1}(g)$$

$$= \sum_{g \in G_\delta, a \in A} \Delta u_i(a, a_i', \omega_\theta) \left| \int \left(\int \sigma(a|\omega') \left(\tau_i'(\mathrm{d}\omega'|\zeta_\delta^{-1}(g)) - \tau_i(\mathrm{d}\omega'|\zeta_\delta^{-1}(g))\right) \right) P(\mathrm{d}\tau_i'|\zeta_\delta^{-1}(g)) \right| P(\zeta_\delta^{-1}(g)|\tau_i)$$

$$\leq 2M\delta.$$

Since σ is obedient, we conclude that σ_δ is $2M\delta$-obedient. By construction, for every $(\theta', \tau') \in \hat{\Omega}_\delta \cap \Omega_{P'}$ there exists $(\theta, \tau) \in \Omega_P$ so that $\zeta_\delta(\theta', \tau') = \zeta_\delta(\theta, \tau)$ and $d_\Pi((\theta', \tau'), (\theta, \tau)) < \delta$. Since σ_δ is ζ_δ-measurable (hence finite valued) we have that

$$\int_{\hat{\Omega}_\delta \cup \Omega_P} \sum_{a \in A} \Delta u_i(a, a_i', \omega_\theta) \sigma_\delta(a|\omega) \left(\; \mathrm{d}P(\omega|\tau_i) - \; \mathrm{d}P'(\omega|\tau_i')\right) < 2M\delta.$$

$2M\delta$-obedience of σ_δ thus implies that

$$\textstyle\int_{\hat{\Omega}_\delta} \sum_{a \in A} \Delta u_i(a, a_i', \omega_\theta) \sigma_\delta(a|\omega) \; \mathrm{d}P'(\omega|\tau_i') \quad \geq -4M\delta \;\; .$$

Moreover, $(\theta', \tau') \in \hat{\Omega}_\delta$ implies that for any measurable $\hat{\sigma} \colon \Omega_{P'} \setminus \hat{\Omega}_\delta \to \Delta(A)$, we have $\int_{\Omega_{P'} \setminus \hat{\Omega}_\delta} \sum_{a \in A} \Delta u_i(a, a_i', \omega_\theta) \hat{\sigma}(a|\omega) \; P'(\mathrm{d}\omega|\tau_i') \geq -2M\delta$ and so

$$\int_{\Omega_{P'}} \sum_{a \in A} \Delta u_i(a, a_i', \omega_\theta) \sigma_\delta(a|\omega) \; P'(\mathrm{d}\omega|\tau_i') \geq -6M\delta. \tag{1}$$

So σ_δ satisfies $6M\delta$-obedience under P' if restricted to type profiles in $\hat{\Omega}_\delta$. We now argue that there exists a $6M\delta$-obedient decision rule under P' that agrees with the extension σ_δ

on $\hat{\Omega}_\delta$. For every player i, let $D_i := \left\{\tau \in \mathcal{T} : \tau_i(\hat{\Omega}_\delta) \leq 1-\delta\right\}$ and note that if $P'(\hat{\Omega}_\delta) < 1$ then $P'(\bigcup_{i\in I} D_i) > 0$.

Consider the auxiliary payoffs $\tilde{u}\colon \Omega \times A \to \mathbb{R}^I$, defined for every $(\theta,\tau) \in \Omega$ and $a \in A$ as follows

$$\tilde{u}_i((\theta,\tau),a) := \begin{cases} u_i(\theta, a_{-i}, a_i), & \text{if } \tau_i\left(\hat{\Omega}_\delta\right) \leq 1-\delta \\ \mathbf{1}_{\sigma_\delta(a|\theta,\tau)>0}, & \text{otherwise.} \end{cases}$$

By Proposition 2 we deduce that the incomplete information game $(\tilde{u}, P')$ admits an obedient decision rule $\bar{\sigma}$ that coincides with σ_δ on $\hat{\Omega}_\delta$. This induces a $6M\delta$-obedient decision of (u, P') defined on all of $\Omega_{P'}$: Indeed, by obedience of $\bar{\sigma}$ for every player i and $(\theta,\tau') \in \Omega_{P'}$ so that $\tau'_i(\hat{\Omega}_\delta) \leq 1-\delta$,

$$\int_{\Omega_{P'}} \sum_{a\in A} \Delta\tilde{u}_i(a, a'_i, \omega)\bar{\sigma}(a|\omega)\; P'(\mathrm{d}\omega|\tau'_i) = \int_{\Omega_{P'}} \sum_{a\in A} \Delta u_i(a, a'_i, \omega_\theta)\bar{\sigma}(a|\omega)\; P'(\mathrm{d}\omega|\tau'_i) \geq 0. \quad (2)$$

Moreover, for every $(\theta,\tau') \in \Omega_{P'}$ so that $\tau'_i(\hat{\Omega}_\delta) > 1-\delta$, there is $\tilde{\tau}' \in \hat{\Omega}_\delta$ so that $\tilde{\tau}'_i = \tau'_i$. So consider the combined decision rule,

$$\sigma'(\omega) := \begin{cases} \sigma_\delta(\omega) & , \text{ if } \omega \in \hat{\Omega}_\delta \\ \bar{\sigma}(\omega) & , \text{ if } \omega \in \Omega_{P'} \setminus \hat{\Omega}_\delta. \end{cases}$$

Combining (1) and (2) we deduce that for every τ

$$\int_{\Omega_{P'}} \sum_{a\in A} \Delta u_i(a, a'_i, \omega_\theta)\sigma'(a|\omega)\; P'(\mathrm{d}\omega|\tau'_i) \geq -6M\delta$$

We will now show that $\nu_{P',\sigma'}$ is close to $\nu_{P,\sigma}$. For any $(a,\theta) \in A \times \Theta$,

$$|\nu_{P,\sigma}(a,\theta) - \nu_{P',\sigma'}(a,\theta)| \leq \left| \sum_{g \in G_\delta} \left(\int \sigma(a|\tau',\theta) \; P(\theta, \mathrm{d}\tau'|g) \right) (P(\theta,\mathrm{d}\tau) - P'(\theta,\mathrm{d}\tau)) \right| + |P'(\hat{\Omega}_\delta|\theta) - P(\hat{\Omega}_\delta|\theta)| \leq 2\delta.$$

Hence $\sum_{a,\theta} |\nu_{P,\sigma}(a,\theta) - \nu_{P',\hat{\sigma}}(a,\theta)|^2 < |\Theta \times A| 4\delta^2$ and so $d^u(P,P') < 6M(u)\delta$. □

We now establish that failure of ACK-convergence implies a failure of convergence of equilibrium outcomes.

Proposition 5. *For every $\varepsilon > 0$, if $d^{ACK}(P,P') \geq \varepsilon$ then there is a game $\mathcal{G}$ so that $d^*(P,P'|\mathcal{G}) \geq \varepsilon$.*

Proof. We now establish that for all $\varepsilon \in (0,1/2)$ so that for all P, P' satisfy $d^{ACK}(P,P') \geq \varepsilon$, we also have $d^*(P,P'|\hat{\mathcal{G}}) \geq \varepsilon$ for some game $\hat{\mathcal{G}}$: If convergence fails in our metric, then there must be some game on which ex-ante strategic convergence fails. Then we must find such a game $\hat{\mathcal{G}}$. The condition $d^{ACK}(P,P') \geq \varepsilon$ means that $P\left(\hat{\Omega}_\varepsilon\right) \leq 1-\varepsilon$, $P'\left(\hat{\Omega}_\varepsilon\right) \leq 1-\varepsilon$ or both. Suppose that $P(\hat{\Omega}_\varepsilon) \leq 1-\varepsilon$ and $P'(\hat{\Omega}_\varepsilon) > 1-\varepsilon$. First, note that $P(\hat{T}_\varepsilon(P,P')) < 1$. Indeed, if $P(\hat{T}_\varepsilon(P,P')) = 1$, we also have that $P(\hat{\Omega}_\varepsilon) = 1$, which is a contradiction. Let $D_{\varepsilon,P} := \mathrm{supp}_\varepsilon(P) \setminus \hat{T}_\varepsilon(P,P')$ and $D^{\complement}_{\varepsilon,P} := \Omega \setminus \mathcal{N}_\varepsilon(D_{\varepsilon,P})$. From Lemmas 1 and 2 we conclude that there is m and z, an associated game $\hat{\mathcal{G}} = (A^{m,z}, (u_i^{m,z})_i)$ where the finite collection of action profiles takes the form $A^{m,z} = \hat{D}_{\varepsilon,P} \cup \hat{D}^{\complement}_{\varepsilon,P}$, with $\hat{D}_{\varepsilon,P} \subseteq D_{\varepsilon,P}$ a finite ε-grid, $\hat{D}^{\complement}_{\varepsilon,P} \subseteq D^{\complement}_{\varepsilon,P}$ a finite ε-grid, and so that for every ε-BIBCE, $\hat{\sigma}' \in \mathcal{B}^\varepsilon(\hat{\mathcal{G}},P')$ and every $\omega \in \Omega_{P'}$, $\hat{\sigma}'(\hat{D}_{\varepsilon,P}|\omega) = 0$. Moreover, for every $\omega \in D_{\varepsilon,P}$ and every ε-BIBCE, $\hat{\sigma} \in \mathcal{B}^\varepsilon(\hat{\mathcal{G}},P)$, $\hat{\sigma}(\hat{D}_{\varepsilon,P}|\omega) = 1$.

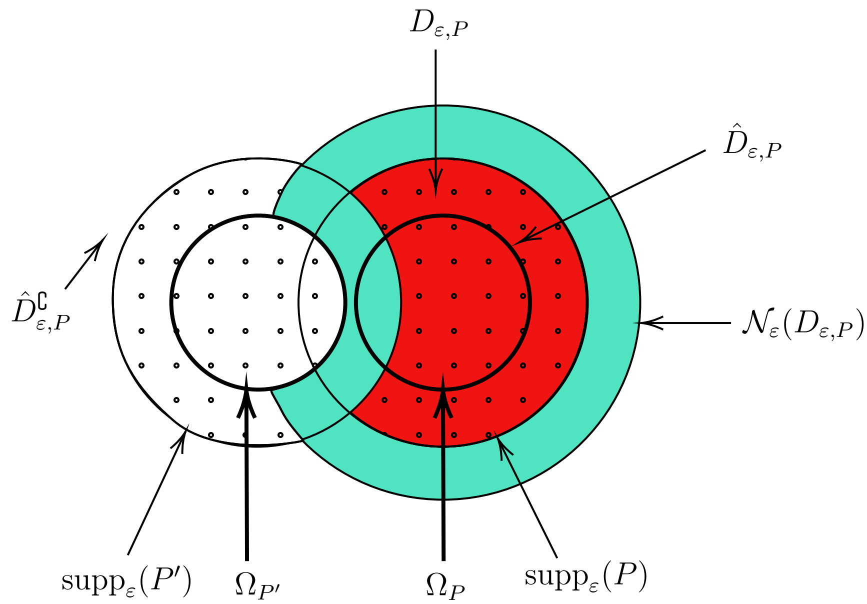

Figure 4: Domain of Grid Game.

Figure 4 above illustrates the sets we just defined. The area enclosed by the two bold circles represent the supports of P and P' respectively. The red shaded area represents the set $D_{\varepsilon,P}$ where the dots on top of the red shaded area represents $\hat{D}_{\varepsilon,P}$, the actions chosen by type profiles in $D_{\varepsilon,P}$. The green shaded area represents $\mathcal{N}_\varepsilon(D_{\varepsilon,P}) \setminus D_{\varepsilon,P}$, which contains no action in $A^{m,z}$. Finally, the dots with white background represent the remaining actions, i.e. the set $\hat{D}^{\complement}_{\varepsilon,P}$. Every type in $\Omega_{P'}$ will pick an action from that set.Let

$$D_P^m := \left\{\omega \in \Omega_P : \exists\ \hat{\sigma} \in \mathcal{B}^\varepsilon(\hat{\mathcal{G}}, P) \text{ s.t. } \hat{\sigma}(\hat{D}_{\varepsilon,P}|\omega) = 1\right\}.$$

For every player i, let $D_{P,i}^m := \{\tau_i : \exists\ (\theta, \hat{\tau}) \in \text{supp}(P),\ (\theta, (\tau_i, \hat{\tau}_{-i})) \in D_P^m\}$. Define the sets $D_{P'}^m$ and $D_{P',i}^m$ for prior P' analogously. Define the sequence $(D_P^{m+n})_{n\in\mathbb{N}}$, recursively for every $n \geq 1$,

$$D_{P,i}^{m+n} := \Omega_P \setminus B_i^{1-\varepsilon}(D_P^{m+n-1}),\ D_P^{m+n} := \prod_{i\in I} D_{P,i}^{m+n}.$$

Let $\hat{D}_{P,i} := \left\{\tau_i : \exists\ (\theta, \tau_{-i}) \in \Theta \times \mathcal{T}_{-i} \text{ s.t. } (\theta, \tau) \in \hat{D}_{\varepsilon,P}\right\}$. Consider the action set $A_i^* := \{a^*, a^{**}\}^I \times A^{m,z}$ and let payoffs of $\hat{\mathcal{G}}^*$ be given by

$$\hat{u}_i(a^0, d, \theta) = \begin{cases} \varepsilon \mathbf{1}_{a_i^0 = a^*} + u_i^{m,z}(d, \theta) & \text{if } d_i \in \hat{D}_{P,i} \\ \mathbf{1}_{a_i^0 = a^*} + u_i^{m,z}(d, \theta) & \text{if } d_i \notin \hat{D}_{P,i} \text{ and } \exists\ j \neq i \text{ s.t. } a_j^0 = a^* \end{cases}$$

In this game, playing $d_i \in \hat{D}_{P,i}$ and a^* is a dominant action for every type in $D_{P,i}^m$. Under payoffs $u^{m,z}$ and prior P', reporting a grid element in D_P^m has cost at least $\varepsilon > 0$. Hence every type under prior P' rationalizes both a^{**} and a^* while no type under prior P in the set $D_{P,i}^m$ rationalizes a^{**}. Now consider $\tau_i \in D_{P,i}^{m+1}$. Again, playing $a_i^0 = a^*$ is uniquely ε-rationalizable since all types in $D_{P,i}^m$ play a^*: Indeed, for every $\tau_i \in D_{P,i}^{m+1}$, there is a player i so that $\tau_i(D_{P,-i}^m) \geq \varepsilon$. For every $n \in \mathbb{N}$, there is a player i_n so that $D_{P,i_n}^{m+n} \neq \varnothing$. Proceeding inductively we obtain that $a_i^0 = a^*$ remains the unique ε-best-reply for some player at type profiles in $\Omega_P \setminus \hat{\Omega}_\varepsilon$. Deduce that there is a ε-BIBCE where all types of all players play a^{**} under P' while for all ε-BIBCE under P, P assigns at least probability ε to type profiles where some player's type plays a^*. Hence $d^*(P, P'|\hat{\mathcal{G}}^*) \geq \varepsilon$. A symmetric argument shows that whenever 1) $P(\hat{\Omega}_\varepsilon) > 1 - \varepsilon$ and $P'(\hat{\Omega}_\varepsilon) \leq 1 - \varepsilon$ or 2) $P(\hat{\Omega}_\varepsilon) > 1 - \varepsilon$ and $P'(\hat{\Omega}_\varepsilon) > 1 - \varepsilon$ implies that outcomes are ε apart under a similarly constructed game. Note that under $P(\hat{\Omega}_\varepsilon) > 1 - \varepsilon$ and $P'(\hat{\Omega}_\varepsilon) > 1 - \varepsilon$, the game with payoffs $u^{m,z}$ is enough to separate the outcomes of P and P'.

The grid game alone would not work for our proof: Indeed, we could construct a sequence of priors P^k for any prior P so that $d^{ACK}(P^k, P) > \varepsilon$ for all k but where $P^k(T_\varepsilon(P^k, P)) \uparrow 1$, i.e. the infecting event has diminishing ex-ante probability. If we used the grid game only, we would have that $d^*(P, P'|\hat{\mathcal{G}}) \downarrow 0$. Based on this game alone, the ACK topology would be too strong. □

Now Theorem 1 follows immediately from Propositions 4 and 5.

We will conclude this section by discussing a key property of ACK topology (denseness) and an extension of our main result (so we have continuity of exact BIBCE). Both will be important for our applications.

3.4 Denseness of Simple Types

An information structure P is *finite* if the support of P is finite. A finite information structure P is a *first-order belief* information structure if each type has a distinct first-order belief. An information structure P is *simple* if it is both a finite and a first-order belief information structure. We denote the collection of simple information structures

$$\mathcal{P}^* := \left\{ P \in \mathcal{P} : |\mathrm{supp}(P)| < \infty,\ \forall\ (\theta, \tau), (\hat{\theta}, \hat{\tau}) \in \mathrm{supp}(P), \forall\ i,\ \tau_i \neq \hat{\tau}_i \implies \tau_i^1 \neq \hat{\tau}_i^1 \right\}$$

It is often convenient to work with simple information structures. We have:

Proposition 6. *Simple information structures are dense in $\mathcal{P}$ under the ACK topology.*

Proof. Fix any information structure P. For any $\delta > 0$ we construct a finite information structure P_δ so that $d^*(P, P_\delta) < \delta$. To do so, we construct a finite grid G_δ of state and type pairs whose δ-neighborhood covers Ω_P, i.e. $\Omega_P \subseteq \mathcal{N}_\delta(G_\delta)$ and for all $\omega \in \Omega_P$, $\min_{g \in G_\varepsilon} d_\Pi(\omega, g) \leq \delta$. Consider any partition on Ω_P given by the pre-image of any map $\zeta \colon \Omega_P \to G_\delta$, where for all $\omega \in \Omega_P$, we have that $d_\Pi(\omega, \zeta(\omega)) \leq \delta$. Consider the information structure P_δ given by $P \circ \zeta^{-1}$. This information structure is finite and has the property that for any $\tau \in \mathrm{supp}(P_\delta)$ there exists $\tau' \in \mathrm{supp}(P)$ so that $d_\Pi(\tau, \tau') < \delta$ and vice versa; for all $\tau' \in \mathrm{supp}(P)$ there is $\tau \in \mathrm{supp}(P_\delta)$ so that $d_\Pi(\tau, \tau') < \delta$. Indeed, for every player i and $\omega \in \Omega_P$ and $(\theta, \tau) = \zeta(\omega)$, beliefs of any measurable event $E \in \mathscr{B}$ at τ_i are given by,

$$P \circ \zeta^{-1}(E|\tau_i) = \int_{\zeta_i^{-1}(\tau_i)} P(E|\hat{\tau}_i) P(\mathrm{d}\hat{\tau}_i | \zeta_i^{-1}(\tau_i)),$$

where $\zeta_i^{-1}(\tau_i) := \bigcup_{(\hat{\theta},\hat{\tau})\in G_\delta:\hat{\tau}_i=\tau_i} \left\{\tilde{\tau}_i : \exists\ (\tilde{\theta}, \tilde{\tau}_{-i}) \text{ s.t. } (\tilde{\theta}, \tilde{\tau}) \in \zeta^{-1}(\hat{\theta}, \hat{\tau})\right\}$. First order beliefs of $P\circ\zeta^{-1}$ take the form $P\circ\zeta^{-1}(\theta|\tau_i) = \int_{\zeta_i^{-1}(\tau_i)} P(\theta|\hat{\tau}_i)P(\mathrm{d}\hat{\tau}_i|\zeta_i^{-1}(\tau_i))$ and so for every $\tilde{\tau}_i \in \zeta_i^{-1}(\tau_i)$, first-order beliefs of $P \circ \zeta^{-1}(\cdot|\tau_i)$ and $P(\cdot|\tilde{\tau}_i)$ are no more than δ apart in the Euclidean topology on $\Delta(\Theta)$. Deduce that $P \circ \zeta^{-1}(\cdot|\tau_i)$ and $P(\cdot|\tilde{\tau}_i)$ have δ-close hierarchies of beliefs in the product topology on $\mathcal{T}_i$. Hence $\hat{T}_\delta(P, P_\delta) = \mathrm{supp}_\delta(P)$ and so $d^{ACK}(P, P_\delta) < \delta$. For δ small enough, there is $\hat{\zeta}\colon \mathcal{T} \to \mathcal{T}$,so that our choice of ζ above satisfies $\zeta(\theta, \tau) = (\theta, \hat{\zeta}(\tau))$ and so the marginal of $P\circ\zeta^{-1}$ on Θ coincides with that of P. Hence finite, canonical information structures are dense in $\mathcal{P}$ under the strategic topology. It remains to show that P_δ is close to a simple canonical information structure. Let $\Omega_{i,P_\delta}(\tau_i) := \{\hat{\tau}_i : P_\delta(\hat{\tau}_i) > 0, \hat{\tau}_i^1 = \tau_i^1\}$. For any $\epsilon > 0$ let $\rho_{i,\varepsilon}\colon \Theta \times \mathcal{T}_i \to \Delta(\{0, 1\})$ have the property that $\rho_{i,\varepsilon}(\theta, \hat{\tau}_i) = \rho_{i,\varepsilon}(\theta', \hat{\tau}_i)$ for all $\theta, \theta' \in \Theta$ and all $\hat{\tau}_i$ in the support of P_δ satisfying $|\Omega_{i,P_\delta}(\hat{\tau}_i)| = 1$. For any $\hat{\tau}_i$ in the support of P_δ satisfying $|\Omega_{i,P_\delta}(\hat{\tau}_i)| > 1$ let $||\rho_{i,\varepsilon}(\theta, \hat{\tau}_i) - \rho_{i,\varepsilon}(\theta', \hat{\tau}_i)||_2 < \epsilon$ so that for all distinct $\tilde{\tau}_i, \bar{\tau}_i \in \Omega_{i,P_\delta}(\tau_i)$ and $s_i, s_i' \in \{0, 1\}$,

$$\frac{\rho_{i,\varepsilon}(s_i|\theta, \tilde{\tau}_i)P_\delta(\theta|\tilde{\tau}_i)}{\sum_{\hat{\theta}} \rho_{i,\varepsilon}(s_i|\hat{\theta}, \tilde{\tau}_i)P_\delta(\hat{\theta}|\tilde{\tau}_i)} \neq \frac{\rho_{i,\varepsilon}(s_i'|\theta, \bar{\tau}_i)P_\delta(\theta|\bar{\tau}_i)}{\sum_{\hat{\theta}} \rho_{i,\varepsilon}(s_i'|\hat{\theta}, \bar{\tau}_i)P_\delta(\hat{\theta}|\bar{\tau}_i)}.$$

The prior $\hat{P}_\delta(\theta, \tau, s) = P_\delta(\theta, \tau)\prod_i \rho_{i,\varepsilon}(s_i|\theta, \tau)$ induces information structure $\tilde{P}_\delta$ so that $d^{ACK}(P, \tilde{P}_\delta) < \delta + \varepsilon$. Deduce that simple information structures are dense in $\mathcal{P}$. □

4 General Information Structures and Bayes Nash Equilibrium

Our main approach in this paper is to remove correlating devices from the information structure (and thus work with non-redundant information structures) and put correlating devices in the solution concept (BIBCE). However, to relate our work to the literature and discuss applications it is useful to discuss how our results apply when we allow correlating devices in the information structure (and thus allow for general *redundant information struc-*

tures) but remove correlating devices from the solution concept (and work with Bayes Nash equilibrium).

4.1 General Information Structures

A (common prior) *general information structure* describes a set of signals for each player and a joint distribution over states and profiles of signals: thus a general information structure is a tuple $\mathcal{S} = ((S_i)_{i \in I}, Q)$, where each S_i is a measurable space of signals[5] for player i and Q is a probability measure[6] in $\Delta(\Theta \times S)$ whose marginal on Θ is given by μ.

We dub this object a general information structure. Liu (2015) describes how one can always decompose a general information structure into a (non-redundant) information structure and a correlation device. While a general information structure is described by a signal space and a probability measure, we will adopt the convention of describing non-redundant information structures as just a measure, leaving it understood that the player's signal space is the universal space of hierarchies.

Every redundant information structure and can be naturally mapped to its non-redundant information structure by essentially integrating out redundant types. In particular, we map information structure $\mathcal{S} = ((S_i)_{i \in I}, Q)$ to a (non-redundant) information structure P as follows. For every i, and version of the conditional probability Q_i, first-order beliefs can be obtained for any player i, $\overline{\tau}_i^1(s_i) = \text{marg}_\Theta(Q_i(s_i))$ for every $s_i \in S_i$. For every $m > 1$ and $m-1$-order beliefs representation $\overline{\tau}_j^{m-1} \colon S_j \to \mathcal{T}_j^{m-1}$, for every j, obtain the m-order belief representation of player i, $\overline{\tau}_i^m(s_i) = Q_i(s_i) \circ (\text{id} \times \overline{\tau}_{-i}^{m-1})^{-1}$, where id is the identity on Θ and $\overline{\tau}_{-i}^{m-1} \colon s_{-i} \mapsto (\overline{\tau}_j^{m-1}(s_j))_{j \neq i}$. The representation of s_i in $\mathcal{T}_i$ is then given by $\overline{\tau}_i(s_i) = (\overline{\tau}_i^m(s_i))_m$. For every $s \in S$ let $\overline{\tau}(s) := (\overline{\tau}_i(s_i))_{i \in I}$ and we write $P_{\mathcal{S}}$ for the information structure thus induced by redundant information structure $\mathcal{S}$.

[5]We adopt the convention of referring to "signals" rather than "types" when describing general information structures in this section. We reserve the terminology "type" for hierarchies of beliefs, introduced in the next section.

[6]The product of measurable spaces is always endowed with the product sigma algebra.

4.2 Solution Concepts

Now we will say that a base game and a redundant information structure $(\mathcal{G}, \mathcal{S})$ together define a *Bayesian game.*

Belief-invariant Bayes correlated equilibrium will be defined as before for general information structures and Bayesian games. For completeness, we will spell out the definition allowing for general information structures, and also define Bayes Nash Equilibrium.

For any Bayesian game $(\mathcal{G}, \mathcal{S})$, a decision rule is a measurable map $\sigma : \Theta \times S \to \Delta(A)$, where $A := \prod_{i\in I} A_i$ and $\Delta(A)$ is endowed with the Euclidean topology. A general information structure $\mathcal{S}$ and decision rule σ jointly induce a measure $\sigma \circ Q \in \Delta(A \times \Theta \times S)$ in the natural way. We will be interested in *outcomes* specifying a joint distribution over actions and states $\nu \in \Delta(A \times \Theta)$. Decision rule σ *induces* outcome ν_σ if ν_σ is the marginal of $\sigma \circ Q$ on $A \times \Theta$. For every player i, a decision rule σ and a version of the conditional probability $Q_i \colon S_i \to \Delta(\Theta \times S_{-i})$ of Q induce a belief for every signal $s_i \in S_i$, $\sigma \circ Q_i \colon S_i \to \Delta(A \times \Theta \times S_{-i})$, which for every measurable set $E \subseteq A \times \Theta \times S_{-i}$ and every signal $s_i \in S_i$ satisfies $\sigma \circ Q_i(E|s_i) = \int_E \sigma(a|\theta, s_{-i}, s_i)\ \mathrm{d}Q_i(\theta, s_{-i}|s_i)$.

Now we have:

Definition 12. A decision rule σ is ε-obedient if, for every player i , there is a version of the conditional probability $Q_i \colon S_i \to \Delta(\Theta \times S_{-i})$ so that every action $a_i \in A_i$ and deviation a_i',

$$\int_{S\times\Theta} \sum_{a_{-i}\in A_{-i}} (u_i(a_i, a_{-i}, \theta) - u_i(a_i', a_{-i}, \theta))\ \mathrm{d}\sigma \circ Q_i(a_i, a_{-i}, s_{-i}, \theta|s_i) > -\varepsilon,\ \ a.s.$$

Definition 13. A decision rule σ is belief-invariant if, for every $a_i \in A_i$, the marginal probability $\sigma(a_i \times A_{-i}|(s_i, s_{-i}), \theta) = \sigma_i(a_i|s_i)$ does not depend on (s_{-i}, θ).

Definition 14. A decision rule σ is an ε-belief-invariant Bayes correlated equilibrium (ε- BIBCE) of $(\mathcal{G}, \mathcal{S})$ if it satisfies ε-obedience and belief invariance.

If we have a (non-redundant) information structure, these definitions reduce to those introduced earlier.

Now a decision rule σ is *conditionally independent* if $\sigma(a|(s_i, s_{-i}), \theta) = \prod_{i \in I} \sigma_i(a_i|s_i)$, for every $(a, s, \theta) \in A \times S \times \Theta$. Conditional independence requires that any randomization in a player's actions depends on their type only and is conditionally independent of others' types and the state. If a decision rule is conditionally independent, it gives a behavioral strategy for each player in the incomplete information game.

Definition 15. A decision rule σ is an ε-Bayes Nash equilibrium (ε-BNE) of $(\mathcal{G}, \mathcal{S})$ if it satisfies obedience (ε-obedience), belief-invariance and conditional independence.

We will say that a decision rule is a BIBCE or BNE if it is a 0-BIBCE or 0-BNE respectively. We will write $\mathcal{B}(\mathcal{G}, \mathcal{S})$ and $\mathcal{B}^{BNE}(\mathcal{G}, \mathcal{S})$ for the set of BIBCE and BNE decision rules, and ν_σ for the outcome in $\Delta(A \times \Theta)$induced by decision rule σ; we will write $\mathcal{O}(\mathcal{G}, \mathcal{S})$ for the set of BIBCE outcomes

$$\mathcal{O}(\mathcal{G}, \mathcal{S}) := \{\nu_\sigma \in \Delta(A \times \Theta) : \sigma \in \mathcal{B}(\mathcal{G}, \mathcal{S})\}$$

and $\mathcal{O}^{BNE}(\mathcal{G}, \mathcal{S})$ for the set BNE outcomes:

$$\mathcal{O}^{BNE}(\mathcal{G}, \mathcal{S}) := \left\{\nu_\sigma \in \Delta(A \times \Theta) : \sigma \in \mathcal{B}^{BNE}(\mathcal{G}, \mathcal{S})\right\}$$

4.3 Existence of Equilibria

We have existence of BIBCE.

Proposition 7. *(Existence of BIBCE)* There exists a BIBCE for every $(\mathcal{G}, \mathcal{S})$.

This was already established for (non-redundant) information structures in Section 2. For a redundant information structure, it is enough to find a BIBCE for its non-redundant version and extend the BIBCE to have players ignore redundancies.

However, strong conditions are required to ensure the existence of BNE, and this has been one obstacle to constructing topologies on information for BNE. Milgrom and Weber (1985) and Balder (1988) are apparently the best available (even if we restrict attention to finite action games). Two sufficient conditions from Milgrom and Weber (1985) are important. First, existence is guaranteed if information structures have countable support. Second, existence is guaranteed if the measure on signals is absolutely continuous with respect to the product of the marginal measures on individual player's signals. Monderer and Samet (1996) and Kajii and Morris (1998) therefore restricted attention countable information structures. However, existence of BNE or even ε-BNE fails when these properties fail: see Simon (2003), Hellman (2014) and Simon and Tomkowicz (2017) for examples. In particular, existence is not guaranteed on (non-redundant) information structures.[7] In the Appendix, we report an example from Hellman (2014) where BNE fails to exist and report a BIBCE for that example.

4.4 Varying Solution Concepts and Information Structures

We first discuss the connection between BNE and BIBCE outcomes. The next proposition states that the set of BIBCE outcomes depends only on the non-redundant information structure:

Proposition 8. *For any base game $\mathcal{G}$ and general information structure $\mathcal{S}$, $\mathcal{O}(\mathcal{G}, \mathcal{S}) = \mathcal{O}(\mathcal{G}, P_{\mathcal{S}})$.*

This is true because any need for redundancies / correlating devices is built into the solution concept. This observation parallels the observation of Dekel et al. (2007) that interim correlated rationalizability depends only on hierarchies of beliefs; Liu (2015) showed that (a subjective version of) BIBCE is equilibrium analogue of ICR and thus provides a proof. [8] For completeness, we give a proof in our notation in the Appendix. It is immediate from the definitions that BNE outcomes are BIBCE outcomes for any information structure.

[7] van Zandt (2010) establishes existence of BNE on the universal type space for supermodular games.
[8] See also Bergemann and Morris (2017) for a discussion of these issues.

Proposition 9. *For any game $\mathcal{G}$ and general information structure $\mathcal{S}$, $\mathcal{O}^{BNE}(\mathcal{G},\mathcal{S}) \subseteq \mathcal{O}(\mathcal{G},\mathcal{S})$.*

Now define the strategic distance between a pair of general information structures $\mathcal{S}$ and $\mathcal{S}'$ to be the strategic distance between their non-redundant representations, so $d^{**}(\mathcal{S},\mathcal{S}') := d^{*}(P_{\mathcal{S}}, P_{\mathcal{S}'})$.

4.5 Correlation and Non-Redundant Information Structures

On general information structures, players' ability to correlate their actions (using redundant correlating devices) matters for the set of BNE. However, non-redundant information structures, as long as there are at least two states, are very rich objects and, intuitively, there will be plenty of opportunity to approximate arbitrary correlating devices within them.

Our use of BIBCE as a solution concept allowed us to focus attention on (non-redundant) information structures and ensured existence of equilibrium. Thus we obtained cleaner results with this solution concept. But suppose one is interested in Bayes Nash equilibrium. In this case, redundant types / correlating devices potentially matter for equilibrium. And we potentially have problems with equilibrium existence. In this section, we will argue that there is a natural approach to dealing with these difficulties (maintaining BNE as the preferred solution concept) and that the same almost common knowledge topology is relevant for continuity of equilibrium outcomes.

Our main observation observation is that, since the universal type space is rich (as long as there at least two states), any correlating device can be embedded in a (non-redundant) information structure by perturbing types' first-order beliefs. The following proposition establishes that any BIBCE on any finite information structure can be approximated by an approximate BNE of some nearby simple information structure.

Proposition 10. *Let $|\Theta| \geq 2$. For any finite, general information structure $\mathcal{S}$, any $\sigma \in \mathcal{B}(\mathcal{G},\mathcal{S})$ and any $\varepsilon > 0$, there exists (i) a simple information structure $\mathcal{S}'$ such that $d^{**}(\mathcal{S},\mathcal{S}') \leq$*

ε ; (ii) a decision rule σ' such that (a) σ' is a ε-BNE of $(\mathcal{G}, \mathcal{S}')$ and (b) the outcome induced by $\mathcal{S}' \circ \sigma'$ is *ε-close to the outcome induced by $\mathcal{S} \circ \sigma$.*

Proof. Let $Q \circ \sigma \in \Delta (A \times T \times \Theta)$ be the extended outcome corresponding to $\sigma \in \mathcal{B}(\mathcal{G}, \mathcal{S})$. Consider the (non-canonical) information structure where each player's signal space was $S_i = A_i \times T_i$ and the prior was $Q \circ \sigma$. Note that since σ is an arbitrary BIBCE, this information structure will in general have redundancies. But under this information structure, there is a pure strategy BNE σ' where each player sets his action equal to his "recommendation" (i.e., the action component of his signal): $\sigma'(a_i|a_i, \tau_i) = \mathbf{1}_{a_i}$ and so

$$\sum_{\theta,\tau} \Delta u_i(a, a_i', \theta) \prod_i \sigma_i'(a_i|a_i, \tau_i)\sigma(a|\theta, \tau)Q(\theta, \tau_{-i}|\tau_i) = \sum_{\theta,\tau} \Delta u_i(a, a_i', \theta)\sigma(a|\theta, \tau)Q(\theta, \tau_{-i}|\tau_i) \geq 0.$$

We now apply the same pertubation as in the proof of Proposition 6: Let $\Omega_{i,Q\circ\sigma}(\tau_i, a_i) := \{(\hat{\tau}_i, \hat{a}_i) : P \circ \sigma(\hat{\tau}_i, \hat{a}_i) > 0, \hat{\tau}_i^1 = \tau_i^1\}$. Let $\rho_{i,\varepsilon} \colon \Theta \times \mathcal{T}_i \to \Delta(\{0, 1\})$ have the property that $\rho_{i,\varepsilon}(\theta, \hat{\tau}_i, \hat{a}_i) = \rho_{i,\varepsilon}(\theta', \hat{\tau}_i, \hat{a}_i)$ for all $\theta, \theta' \in \Theta$ and all $\hat{\tau}_i, \hat{a}_i$ in the support of $Q \circ \sigma$ satisfying $|\Omega_{i,P\circ\sigma}(\hat{\tau}_i, \hat{a}_i)| = 1$. For any $\hat{\tau}_i, \hat{a}_i$ in the support of $Q \circ \sigma$ satisfying $|\Omega_{i,Q\circ\sigma}(\hat{\tau}_i, \hat{a}_i)| > 1$ let $||\rho_{i,\varepsilon}(\theta, \hat{\tau}_i, \hat{a}_i) - \rho_{i,\varepsilon}(\theta', \hat{\tau}_i, \hat{a}_i)||_2 < \varepsilon$ so that for all distinct $(\tilde{\tau}_i, \tilde{a}_i), (\bar{\tau}_i, \bar{a}_i) \in \Omega_{i,P_\delta}(\tau_i, a_i)$ and $s_i, s_i' \in \{0, 1\}$,

$$\frac{\rho_{i,\varepsilon}(s_i|\theta, \tilde{\tau}_i, \tilde{a}_i)Q \circ \sigma(\theta|\tilde{\tau}_i, \tilde{a}_i)}{\sum_{\hat{\theta}} \rho_{i,\varepsilon}(s_i|\hat{\theta}, \tilde{\tau}_i, \tilde{a}_i)Q \circ \sigma(\hat{\theta}|\tilde{\tau}_i, \tilde{a}_i)} \neq \frac{\rho_{i,\varepsilon}(s_i'|\theta, \bar{\tau}_i, \bar{a}_i)Q \circ \sigma(\theta|\bar{\tau}_i, \bar{a}_i)}{\sum_{\hat{\theta}} \rho_{i,\varepsilon}(s_i'|\hat{\theta}, \bar{\tau}_i, \bar{a}_i)Q \circ \sigma(\hat{\theta}|\bar{\tau}_i, \bar{a}_i)}.$$

The prior $\hat{Q}(\theta, \tau, a, s) = Q \circ \sigma(\theta, \tau, a) \prod_i \rho_{i,\varepsilon}(s_i|\theta, \tau, a)$ induces canonical prior $Q' \in \mathcal{P}$ so that $d^{ACK}(P, P') < \varepsilon$ and σ' induces an outcome equivalent ε-BNE on P'. □

Proposition 10 shows that that any correlation device required for a BIBCE can be embedded in the universal state space at the expense of ε-slack in the obedience constraint. Notice that under complete information (i.e., if there was a single state), it would not be possible to do so. Now the denseness of simple information structures implies the immediate corollary that this is true for all information structures (not just canonical ones).

Corollary 1. *Let $|\Theta| \geq 2$. For any information structure $P \in \mathcal{P}$, any BIBCE $\sigma \in \mathcal{B}(\mathcal{G}, P)$ and any $\varepsilon > 0$, there exists (i) a simple information structure $P' \in \mathcal{P}^*$ such that $d^*(P, P') \leq \varepsilon$; and (ii) a decision rule σ' such that (a) σ' is a ε-BNE of* $(\mathcal{G}, P')$ and (b) the outcome induced by $P' \circ \sigma'$ is *ε-close to the outcome induced by $P \circ \sigma$.*

Proof. The denseness of simple information structures implies that there exists a simple information structure P'' with $d^*(P, P") \leq \varepsilon$. Now the corollary follows from applying Proposition 10 to P''. □

Corollary 1 implies that if we extend the notion of approximate BNE to allow not only only slack in the obedience constraints but also to allow nearby (in the ACK topology) information structures, we can first establish the existence of approximate BNE and then establish continuity of (approximate) BNE outcomes with respect to the ACK topology.

We first define an extended notion of ε-BNE outcomes:

$$\mathcal{O}_{\varepsilon}^{BNE^*}(\mathcal{G}, P) = \left\{\nu \in \Delta(A \times \Theta) : \exists\ P' \text{ with } d^*(P, P') \leq \varepsilon, \sigma \in \mathcal{B}_{\varepsilon}^{BNE}(\mathcal{G}, P) \text{ s.t. } ||v - v_{\sigma}|| < \varepsilon\right\}$$

Now we have existence:

Corollary 2. *For every game $\mathcal{G}$, prior $P \in \mathcal{P}$ and $\varepsilon > 0$, $\mathcal{O}_{\varepsilon}^{BNE}(\mathcal{G}, P) \neq \emptyset$.*

Proof. This result follows from Corollary 1 and the fact that $\mathcal{O}^{BNE}(\mathcal{G}, P) \neq \emptyset$ for all finite $P \in \mathcal{P}$. □

We now introduce a "richness" property of a base game.

Definition 16. (richness) A base game $\mathcal{G}$ is *rich* if for every action profile $a \in A$, there exists $\theta_a \in \Theta$ such that for every player i, every $a_i \in A_i$ and every $a_i' \neq a_i$,

$$u_i(a_i, a_{-i}, \theta_a) - u_i(a_i', a_{-i}, \theta_a) > 0.$$

With richness, we have continuity of approximate BNE outcomes:

Proposition 11. *Let* $|\Theta| \geq 2$. *Then for every rich base game* $\mathcal{G}$ *and any information structure* $P \in \mathcal{P}$,

$$\lim_{\varepsilon \downarrow 0} \bigcup_{\hat{P} \in \mathcal{P}^*: d^{ACK}(P,\hat{P}) \leq \varepsilon} \mathcal{O}^{BNE}(\mathcal{G}, \hat{P}) = \mathcal{O}(\mathcal{G}, P).$$

Proof. Let $\hat{P} \in \mathcal{P}^*$ satisfy $d^{ACK}(P, \hat{P}) \leq \varepsilon$. Then by Corollary 1, for every $\nu \in \mathcal{O}(\mathcal{G}, P)$ there exists a pure strategy ε-BNE, σ, of $(\mathcal{G}, \hat{P})$, so that $||\nu_\sigma - \nu||_2 \leq \varepsilon$. For every τ denote the associated action recommendation by $\alpha(\theta, \tau) = (\alpha_i(\tau_i))_i$, where $\sigma(\alpha(\theta, \tau)|\theta, \tau) = 1$. For any choice $\delta > 0$, define the stochastic map $\rho_\delta \colon \Theta \times \mathcal{T} \to \Delta(A)$

$$\rho_\delta(a|\theta,\tau) = \begin{cases} 1-\delta & \text{if } a = \alpha(\theta,\tau) \\ \delta & \text{if } \theta = \theta_a \\ 0 & \text{otherwise.} \end{cases}$$

For every $(\theta, \tau, a) \in \Theta \times \mathcal{T} \times A$, let $P_\delta(\theta, \tau, a) := \hat{P}(\theta, \tau)\rho_\delta(a|\theta, \tau)$ and note that P_δ has a canonical representation $\hat{P}_\delta \in \mathcal{P}$. Let $J_{\mathcal{G}} := \min_{i,a,a'} u_i(a_i, a_{-i}, \theta_a) - u_i(a'_i, a_{-i}, \theta_a)$. Then for every i and type τ_i in the support of $\hat{P}$, and any deviation a'_i,

$$\begin{aligned}\sum_{\theta,\tau,a} \Delta u_i(\alpha_i(\tau_i), \alpha_{-i}(\tau_{-i}), a'_i, \theta)\hat{P}_\delta(\theta, \tau_{-i}|\tau_i) &> -\varepsilon(1-\delta) + \delta \sum_{\theta,\tau,a} \Delta u_i(\alpha_i(\tau_i), \alpha_{-i}(\tau_{-i}), a'_i, \theta_{\alpha(\theta,\tau)})\hat{P}_\delta(\theta, \tau_{-i}|\tau_i) \\ &\geq -\varepsilon(1-\delta) + \delta J_{\mathcal{G}}\end{aligned}$$

Letting $\delta = \frac{\varepsilon}{J_{\mathcal{G}}+\varepsilon}$ implies that $\sigma \in \mathcal{B}_{BNE}(\mathcal{G}, \hat{P}_\delta)$.Moreover, $d^{ACK}(\hat{P}, \hat{P}_\delta) \leq \delta$ and so

$$\lim_{\varepsilon \downarrow 0} \bigcup_{\hat{P} \in \mathcal{P}^* \cap \mathcal{P}: d^{ACK}(P,\hat{P}) \leq \varepsilon} \mathcal{O}^{BNE}(\mathcal{G}, \hat{P}) \supseteq \mathcal{O}(\mathcal{G}, P).$$

The property of upper hemi-continuity established in Proposition 18 readily extends to the priors in the subset $\mathcal{P}$ and so we also have that

$$\lim_{\varepsilon\downarrow 0}\bigcup_{\hat{P}\in\mathcal{P}^*\cap\mathcal{P}:d^{ACK}(P,\hat{P})\leq\varepsilon}\mathcal{O}^{BNE}(\mathcal{G},\hat{P})\subseteq\lim_{\varepsilon\downarrow 0}\bigcup_{\hat{P}\in\mathcal{P}^*\cap\mathcal{P}:d^{ACK}(P,\hat{P})\leq\varepsilon}\mathcal{O}(\mathcal{G},\hat{P})\subseteq\mathcal{O}(\mathcal{G},P),$$

and so the result follows. □

4.6 Discussion

We conclude this section by discussing related literature about correlating devices and BNE.

A number of papers have highlighted the importance of redundant types, or correlating devices, for BNE, see for example Liu (2009) and Sadzik (2019). Our approach in this section is to observe that such correlation devices can be embedded in the universal state space in an almost payoff-irrelevant way, so it is natural to work with (non-redundant) information structures even if one is interested in BNE. Ely and Peski (2011) propose an alternative to the standard universal type space that embeds some correlation devices.

In the context of complete information games, Brandenburger and Friedenberg (2008) (see also Du (2012)) asked if correlation devices (supporting correlated equilibrium) could reflect higher-order strategic uncertainty (in this case, they said there is intrinsic correlation) or if extrinsic correlation is required. Their answer is that most correlated equilibria can be explained by intrinsic correlation alone. An analogous question (in an incomplete information context) to ask is which BIBCE could reflect higher-order uncertainty about strategic uncertainty and payoffs. The spirit of our results is that most BIBCE can be justified this way.

Gossner (2000) provided a partial order on correlating devices (for complete information games) capturing which correlating devices would support a larger set of correlated equilibria all games. A natural exercise would be to define a topology on correlating devices (generating continuity of the set of correlated equilibria) although as far as we know that has not been done. We could imagine decomposing a general information structure into a canonical information structure and a correlating device and defining a topology on canonical information structure / correlating device pairs that was sufficient for continuity of BNE outcomes. We have not pursued this approach.

5 Information Design

When studying information design problems, there will typically be many equilibria. In formulating information design problems, one must decide which equilibrium will be played. Two standard choices are to assume (i) the best equilibrium for the designer is played; or (ii) the worst equilibrium for the designer is played. In this section, we propose a formulation of information design problems that includes both those cases, but also allows for any continuous selection of equilibrium.

We will consider the following class of information design problems. A designer has a continuous (in the Hausdorff topology) objective function on *sets* of outcomes

$$V : 2^{\Delta(A\times\Theta)} \setminus \emptyset \to \mathbb{R}.$$

Recall that $\mathcal{P}^* \subseteq \mathcal{P}$ denotes the collection of simple (i.e., finite and first-order) information structures. Now we have:

Proposition 12. *For any rich $\mathcal{G}$ and any open set $\mathcal{P}' \subseteq \mathcal{P}$,*

$$\sup_{\mathcal{S}:P_{\mathcal{S}}\in\mathcal{P}'} V(\mathcal{O}^{BNE}(\mathcal{G},\mathcal{S})) = \sup_{P\in\mathcal{P}'\cap\mathcal{P}^*} V(\mathcal{O}^{BNE}(\mathcal{G},P)) = \sup_{P\in\mathcal{P}'\cap\mathcal{P}^*} V(\mathcal{O}(\mathcal{G},P)).$$

Thus to choose the optimal information structure within an open set, it is enough to focus on either BNE or BIBCE with simple information structures.

We will first describe how the information design problems described above fit within this class, and then describe how the result follows from denseness reported in Section 3.4 and the BNE results reported in Section 4.

5.1 Applications

Our assumption is that the designer cares about a set of outcomes. We are thinking that the designer cares about the outcomes consistent with a solution concept. This sub-section spells out the leading examples of designer objectives where the designer is interested in the best, or the worst, or some continuous selection from the equilibrium outcomes.

Suppose that the designer evaluates outcomes with utility function $u : A \times \Theta \to \mathbb{R}$.

The usual approach in information design is to assume that the designer can choose which equilibrium is played. In this case, if $O \subseteq \Delta(A \times \Theta)$ is the set of equilibrium outcomes, the designer's utility will be:

$$V_{MAX}(O) = \sup_{\nu \in O} \sum_{a,\theta} \nu(a,\theta) u(a,\theta)$$

This objective is continuous and Proposition 12 applies. However, the revelation principle applies to this problem, so we already know that we can restrict attention to finite information structures, without appeal to Proposition 12 and the machinary behind it.

An alternative assumption in information design is that there is "adversarial equilibrium selection," i.e., the designer expects the worst possible equilibrium (for her) to be played. In this case the designer's utility over sets of outcomes will be

$$V_{MIN}(O) = \inf_{\nu \in O} \sum_{a,\theta} \nu(a,\theta) u(a,\theta).$$

A few papers have studied this problem is recent years (Mathevet et al. (2020), Inostroza and Pavan (2023), Morris et al. (2024) and Li et al. (2023)). It is well known that the maximum is typically not attained in this design problem. However, while it is clear in the particular settings of these papers that the supremum can be approached using simple information structures, there is no existing general statement of this property. Thus Proposition 12 is a useful tool for this literature.

More generally, we could let the designer's objective correspond to an arbitrary continuous selection from BIBCE, so there exists $f\colon 2^{\Delta(A\times\Theta)}\backslash\emptyset \to \Delta(A\times\Theta)$ with the following property: for every $O' \subseteq O \subseteq \Delta(A\times\Theta)$,

$$f(O) \in O' \implies f(O') = f(O), \tag{3}$$

such that

$$V_f(O) = \sum_{a,\theta} f(a,\theta|O)u(a,\theta).$$

5.2 Proof of Proposition 12

Our denseness result implies:

Lemma 3. *For every rich game $\mathcal{G}$ and every information structure P, there exists $\varepsilon > 0$ and a simple information structure $P^\varepsilon \in \mathcal{P}^*$ so that $d^{ACK}(P, P^\varepsilon) < \varepsilon$ and*

$$|V(\mathcal{O}(P,\mathcal{G})) - V(\mathcal{O}^{BNE}(P^\varepsilon,\mathcal{G}))| < \varepsilon.$$

Proof. Suppose $\sigma \in \mathcal{B}(\mathcal{G}, P)$ satisfies $V_G(\mathcal{O}(\mathcal{G},P)) = \sum_{a,\theta} \nu_\sigma(a,\theta)\, u(a,\theta)$. Then there are sequences $(P^k,\varepsilon^k)_k$with $P^k \in \mathcal{P}^*$ for all k and $\varepsilon^k \downarrow 0$ so that $d^{ACK}(P,P^k) < \varepsilon_k$ for all k. Moreover, by upper hemi-continuity established in Proposition 18, there is a subset $\mathcal{O}^\infty(\mathcal{G},P) \subseteq \mathcal{O}(P,\mathcal{G})$ so that the sequence satisfies $\lim_{k\uparrow\infty} d_{\mathcal{H},\mathcal{G}}(\mathcal{O}^{BNE}(\mathcal{G},P^k), \mathcal{O}^\infty(\mathcal{G},P)) = 0$. Moreover, by the arguments in Proposition 11 we can pick the sequence so that $\nu_\sigma \in \mathcal{O}^\infty(\mathcal{G},P)$ and so by property (3) and continuity of V we have that $\lim_{k\uparrow\infty} V(\mathcal{O}^{BNE}(P^k,\mathcal{G})) = V(\mathcal{O}(P,\mathcal{G}))$ and so the result follows. □

Now the proof of Proposition 12 is completed as follows. By Lemma 3, we have that

$$\sup_{P\in\mathcal{P}'\cap\mathcal{P}^*} V_G(\mathcal{O}^{BNE}(\mathcal{G},P)) = \sup_{P\in\mathcal{P}'} V_G(\mathcal{O}(\mathcal{G},P)).$$

Since $\mathcal{P}^*\cap\mathcal{P}'\subseteq\mathcal{P}'$, we have that

$$\sup_{P\in\mathcal{P}'\cap\mathcal{P}^*} V_G(\mathcal{O}^{BNE}(\mathcal{G},P)) \leq \sup_{P\in\mathcal{P}'} V_G(\mathcal{O}^{BNE}(\mathcal{G},P)).$$

Moreover, by property (3) and the fact that $\mathcal{O}^{BNE}(\mathcal{G},P)\subseteq\mathcal{O}(\mathcal{G},P)$, we have that

$$\sup_{P\in\mathcal{P}'} V_G(\mathcal{O}^{BNE}(\mathcal{G},P)) \leq \sup_{P\in\mathcal{P}'} V_G(\mathcal{O}(\mathcal{G},P))$$

and so

$$\begin{aligned}\sup_{P\in\mathcal{P}'\cap\mathcal{P}^*} V_G(\mathcal{O}^{BNE}(\mathcal{G},P)) &\leq \sup_{P\in\mathcal{P}'} V_G(\mathcal{O}^{BNE}(\mathcal{G},P))\\ &\leq \sup_{P\in\mathcal{P}'} V_G(\mathcal{O}(\mathcal{G},P))\\ &= \sup_{P\in\mathcal{P}'\cap\mathcal{P}^*} V_G(\mathcal{O}^{BNE}(\mathcal{G},P)).\end{aligned}$$

6 Alternative Formulations and Related Literature

In this section, we will discuss a number of alternative topologies characterizing convergence of strategic outcomes. One purpose in doing so is that it will allow us to formally relate our work to the relevant related literatures.

6.1 Interim Topologies

We have defined and characterize an (ex ante) strategic topology on (common prior) information structures under an equilibrium solution concept (BIBCE). By contrast, Dekel et al. (2006) defined an interim strategic topology on hierarchies of beliefs under the solution concept of interim correlated rationalizability (ICR). Two belief hierarchies were said to be close in the interim strategic topology if, in any game, an action that was ICR at one hierarchy was approximately ICR at the other hierarchy. Chen et al. (2017) provide a characterization of the interim strategic topology in terms of belief hierarchies. Crucially, the interim strategic topology imposes restrictions on the tails of hierarchies of beliefs, unlike the product topology. We provide a formal statement of the characterization of Chen et al. (2017) in the Appendix.

Our definition of the almost common knowledge topology used the product topology in defining the event that interim beliefs were close. But we noted that the use of the product topology was not essential. A first purpose of this section is to record what properties the interim topology must satisfy in order to induce our almost common knowledge topology: it is enough that it is induced by a "nice" metric that (1) refines the product topology; and (2) has a countable dense subset.

The interim strategic topology satisfies these properties. But we also establish that if we had used the interim strategic topology as our initial interim topology, we could have dispensed with the requirement of approximate common knowledge in our definition of the ex ante topology. Intuitively, this is because interim strategic topology imposes enough restrictions on the tails of hierarchies of beliefs to generate the required approximate common knowledge.

Definition 17. (Nice Interim Metric) A metric d on Ω is nice if (1) its induced topology refines the product topology; (2) there is a countable subset of Ω^0 which is dense in Ω.

A topology on Ω is nice if it is induced by a nice metric. Our ACK topology would be the same if we replace the product interim topology with any nice interim topology in the definition. We used the product topology. We could have used total variation as a metric on interim beliefs, as Kajii and Morris (1998) do. But we could also have substituted the interim strategic topology as defined by Dekel et al. (2006) and characterized by Chen et al. (2017). Importantly, that topology (unlike the product topology) also imposes restrictions on infinite tails of hierarchies of beliefs which has implications for approximate common knowledge. In particular, the interim strategic topology the following property.

Definition 18. (Common Belief Invariance) A metric d_{CB} on Ω satisfies common belief invariance if for all events $E, E' \subseteq \Omega$ and every $\varepsilon > 0$,

$$C^{1-\varepsilon}\left(\mathcal{N}_{d_{CB},\varepsilon}(E) \cap \mathcal{N}_{d_{CB},\varepsilon}(E')\right) = \mathcal{N}_{d_{CB},\varepsilon}(E) \cap \mathcal{N}_{d_{CB},\varepsilon}(E'),$$

where $\mathcal{N}_{d_{CB},\varepsilon}(E)$ is the union of ε-neighborhoods around the points in E using metric d_{CB}.

For any metric $d : \Omega \times \Omega \to [0, \infty)$, $\varepsilon > 0$ and $P \in \mathcal{P}$, let $\text{supp}_{d,\varepsilon}(P) := \bigcup_{\omega \in \Omega : P(\mathcal{N}_{d,\varepsilon}(\omega))} \mathcal{N}_{d,\varepsilon}(\omega)$. Now we can consider the simplest and weakest natural way of translating an interim distance into an ex ante distance

Definition 19. The weak ex-ante distance induced by an interim distance d is defined as

$$d'(P, P') = \inf \left\{ \varepsilon \geq 0 : \begin{array}{c} P\left(\hat{T}_{d,\varepsilon}(P, P')\right) > 1 - \varepsilon \\ P'\left(\hat{T}_{d,\varepsilon}(P, P')\right) > 1 - \varepsilon \end{array} \right\},$$

where $\hat{T}_{d,\varepsilon}(P, P') := \text{supp}_{d,\varepsilon}(P) \cap \text{supp}_{d,\varepsilon}(P')$.

Proposition 13. *The weak ex ante distance induced by any nice interim metric satisfying common belief invariance induces the ACK topology.*

Proof. Appendix. □

Here, we have the special property that approximate common knowledge is for free.

Proposition 14. *The interim strategic topology is nice and common belief invariant.*

Proof. Follows from Propositions 16 and 17 in the Appendix. □

For example, Kajii and Morris (1998) define a topology on ex ante information structures (discussed below) but use total variation as a metric on interim beliefs.

6.2 Value-Based Topology

We could alternatively define our topology in terms of convergence of the ex ante expected equilibrium payoffs rather than equilibrium outcomes. This was the approach of Monderer and Samet (1996) and Kajii and Morris (1998) and also the recent work of Gensbittel et al. (2022) for zero sum games. This distinction is not important for our results.

Let $V(G,P)$ be the set of ex ante utilities of players (in $\mathbb{R}^n$) that can arise from some BIBCE of (G,P), and let $V_\varepsilon(G,P)$ be the set of ex ante utilities of players (in $\mathbb{R}^n$) that are within ε of some ε-BIBCE of (G,P). We can say that P,P' are ε-value close in game $\mathcal{G}$ if $V(\mathcal{G},P) \subseteq V_\varepsilon(G,P')$ and $V(\mathcal{G},P') \subseteq V_\varepsilon(G,P)$.

Definition 20. Let $d^V(P,P'|\mathcal{G})$ be the infimum of the set of ε such that P and P' are ε-value close in game $\mathcal{G}$.

Lemma 4. *Now* $d^*(P^k,P|\mathcal{G}) \to P$ *if and only if* $d^V(P^k,P|\mathcal{G}) \to 0$ for all $\mathcal{G}$.

Proof. One direction is immediate, because convergence of outcomes implies converges of values. In the other direction, it is enough to change payoffs so differences in outcome translate into large differences in payoffs. Consider two outcomes ν,ν' of P and P' respectively so that $V_i(\nu)=V_i(\nu')$ but $\nu \neq \nu'$. Consider the augmented payoffs $u_i(a,\theta)+h_i(a_{-i},\theta)$ and associated values $V_i^h(\nu) = V_i(\nu) + \sum_{a,\theta} \nu(a,\theta)h_i(a_{-i},\theta)$ and $V_i^h(\nu') = V_i(\nu') + \sum_{a,\theta} \nu'(a,\theta)h_i(a_{-i},\theta)$. Hence

$$V_i^h(\nu) - V_i^h(\nu') = \sum_{a,\theta} h_i(a_{-i},\theta)(v(a,\theta) - v'(a,\theta))$$

Consider the choice $h_i(a_{-i},\theta) = \mathbf{1}_{\{(a_{-i},\theta):\exists\ a_i \in A_i \text{ s.t. } v(a,\theta)>v'(a,\theta)\}} - \mathbf{1}_{\{(a_{-i},\theta):\exists\ a_i \in A_i \text{ s.t. } v(a,\theta)<v'(a,\theta)\}}$ and so $V_i^h(\nu) - V_i^h(\nu') > 0$. □

Thus there is little difference working with outcome-based strategic topologies and value-based strategic topologies.

However, in the case of zero-sum games, the value is uniquely defined although many outcomes might give rise to the same value. So it is convenient and natural to work with value-based strategic topologies in that case. Peski (2008) and Gossner and Mertens (2020) characterize changes in information structure that increase one player's payoff in all zero sum games. Gensbittel et al. (2022) study essentially the value-based strategic topology defined above but restricted to zero-sum games.

6.3 Join Measurability

Our approach in this paper has been to fix a set of payoff-relevant states Θ, and look at common prior information structures that describe beliefs and higher-order beliefs about those states. Then we characterize the coarsest topology under which equilibrium outcomes converge for any game where payoffs are measurable with respect to Θ. In particular, we do not allow games to depend in an arbitrary way on players' types (or signals).

An alternative approach would be to allow any game where payoffs were measurable with respect to the join of players' types. Equivalently, we could restrict attention to information structures where each payoff-relevant state could arise under only one profile of types; we call this a "join measurability" restriction on information structures. It was implicitly maintained in the works of Monderer and Samet (1996) and Kajii and Morris (1998). This restriction greatly simplifies the arguments. In particular, in the proofs of sufficiency analogous to Proposition 4 the join measurability approach allows a straightforward mapping of a strategy profile on one information structure to another. We were not able to do that, and required a continuous extension exploiting the structure of the universal type space. In the proofs of necessity analogous to Proposition 5, the join measurability approach requires only an email game like component where an infection argument operates and not also an iterated scoring rule to reveal finite levels of beliefs, as in this paper.

This is the most important difference between the work of Monderer and Samet (1996) and Kajii and Morris (1998), and this work. There are number of other differences that are less important. First, the earlier papers focused on BNE as a solution concept, while we focus on BIBCE. Second, they focused on countable information structures (to ensure existence of

BNE), whereas we do not impose that restriction. Third, their topologies were value-based whereas our topology is outcome-based (a difference that we argued was not important in the previous section) . Fourth, we restrict ourselves to a finite set of payoff-relevant states, but these papers must allow for countable payoff relevant states.

The set of information structures considered in Monderer and Samet (1996) and Kajii and Morris (1998), while both satisfying join measurability, were modelled differently. Monderer and Samet (1996) fixed a state space and prior probability. An information structure was then a profile of (countable) partitions of the state space. And payoffs could depend in arbitrary ways on the state space. On the other hand, Kajii and Morris (1998) fixed a countable set of "types" (or labels) for each player and allowed arbitrary probability distributions on the types space. The exact connection between the similar topologies defined on different classes of information structures was not known until recently, when Kambhampati (2023) showed an equivalence between the results.

6.4 Improper Priors and Completeness

In this paper, we have focused on common prior information structures. Our results imply that equilibrium outcomes converge along Cauchy sequences in the ACK topology. However, in general, Cauchy sequences may not have well-defined limits within the space of information structures. In this section, we show that if we enrich the class of information structures to include "improper" common prior information structures, and extend the ACK topology to this class of information structures, then all Cauchy sequences do converge to a well defined limit. This result is of independent interest, in the light of the importance of improper common prior limits in the literature on higher-order beliefs in games (discussed below).

An improper prior on the universal type space is simply a measure with perhaps infinite mass such that there is a conditional probability consistent with the interim beliefs on the universal type space. Formally, we have:

Definition 21. (Improper Prior) A measure on Ω, $Q\colon \mathscr{B} \to [0,\infty]$, is an improper prior if for every player i there is a measurable map $Q_i\colon \mathcal{T}_i \to \Delta(\Omega)$ so that

$$\tau_i^* = \operatorname{marg}_{\Theta\times\mathcal{T}_{-i}}\left(Q_i(\tau_i)\right), \ Q\text{-a.e.}$$

and for every measurable $E \in \mathscr{B}$,

$$\int_{\operatorname{proj}_{\mathcal{T}_i}(E)} Q_i(E|\tau_i)Q(\mathrm{d}\tau_i) = Q(E).$$

Let $\bar{\mathcal{P}}$ denote the set of improper priors and note that $\mathcal{P} \subseteq \bar{\mathcal{P}}$. The approximate common knowledge topology extends in a natural way to $\bar{\mathcal{P}}$:

Definition 22. (Approximate Common Knowledge Distance) For every $P, P' \in \bar{\mathcal{P}}$, let

$$d^{ACK}\left(P,P'\right) := \inf\left\{\varepsilon \geq 0 : \begin{array}{c} P\left(\Omega \setminus C^{1-\varepsilon}\left(\hat{T}_\varepsilon(P,P')\right)\right) < \varepsilon P\left(C^{1-\varepsilon}\left(\hat{T}_\varepsilon(P,P')\right)\right) \\ P'\left(\Omega \setminus C^{1-\varepsilon}\left(\hat{T}_\varepsilon(P,P')\right)\right) < \varepsilon P'\left(C^{1-\varepsilon}\left(\hat{T}_\varepsilon(P,P')\right)\right) \end{array}\right\}.$$

Notice that this Definition coincides with the earlier Definition 8 when applied to proper priors.

A decision rule σ is a BIBCE of an improper prior $Q \in \mathcal{Q}$, if σ is belief invariant and obedience holds almost everywhere. For any $\varepsilon > 0$, let the collection of ε-BIBCE on an improper prior Q be denoted by $\overline{\mathcal{B}}^{\varepsilon}(\mathcal{G}, Q)$.

Proposition 15. *The extended approximate common knowledge topology on $\overline{\mathcal{P}}$ is complete.*

Proof. Let $\mathscr{S} \subseteq \mathscr{B}$ be a semi-ring so that the sigma algebra it generates equals $\mathscr{B}$. Let $(P^k)_k$ be a Cauchy sequence in $\mathcal{P}$ and $(\varepsilon_k)_k$ a sequence so that for every k,

$$\varepsilon_k \leq \sup_{h>k} d^{ACK}(P^k, P^h).$$

Define a limit pre-measure on $\mathscr{S}$, which on any event $E \in \mathscr{S}$ takes the form

$$\xi_0(E) := \lim_{k\uparrow\infty} P^k(\mathcal{N}_{\varepsilon_k}(E_n)),$$

where we set $\mathcal{N}_\varepsilon(\varnothing) = \varnothing$. We now show that ξ_0 is a pre-measure on $\mathscr{S}$. Indeed, $\xi_0(\varnothing) = 0$. Since each P^k is a measure, they are all finitely additive and so for any finite, disjoint $(E_n)_{n\leq N}$, the limit and sum can be exchanged so that

$$\begin{aligned}\xi_0\left(\bigcup_{n\leq N} E_n\right) &= \lim_{k\uparrow\infty}\sum_{n\leq N} P^k(\mathcal{N}_{\varepsilon_k}(E_n)) \\ &= \sum_{n\leq N}\xi_0(E_n)\end{aligned}$$

Finally, countable monotonicity follows from Reverse Fatou's lemma (each P^k is bounded and non-negative):

$$\begin{aligned}\xi_0\left(\bigcup_{n\in\mathbb{N}} E_n\right) &= \lim_{k\uparrow\infty}\sum_{n\in\mathbb{N}} P^k(\mathcal{N}_{\varepsilon^k}(E_n)) \\ &\leq \sum_{n\in\mathbb{N}}\xi_0(E_n).\end{aligned}$$

By Caratheodory-Hahn's Extension Theorem, ξ_0 extends to a measure ξ on the sigma algebra generated by $\mathscr{S}$. It remains to show that ξ_0 is a Canonical Improper Prior. We need to prove that ξ_0 satisfies the consistency condition. For every event $E \in \mathscr{B}$, we have that

$$\begin{aligned}|P^k(E) - \xi(E)| &= \left|\int_{\mathcal{T}_i}\tau_i(\mathcal{N}_\varepsilon(E))P^k(\mathrm{d}\tau_i) - \lim_{k\uparrow\infty}\int_{\mathcal{T}_i}\tau_i(\mathcal{N}_{\varepsilon_k}(E))P^k(\mathrm{d}\tau_i)\right| \\ &\leq \varepsilon.\end{aligned}$$

Hence $\xi_0 \in \overline{\mathcal{P}}$. Deduce that $\overline{\mathcal{P}}$ is a complete metric space. □

7 Appendix

7.1 Proof of Proposition 2

In the set-up of Stinchcombe (2011) a prior is a countably additive probability $P \in \Delta(\prod_i \Omega_i)$, where each Ω_i is endowed with a sigma algebra $\mathcal{F}_i$ and P is defined on a sigma algebra containing the product sigma algebra. Payoffs are given by $U_i\colon \prod_i \Omega_i \to \mathbb{R}^A$. It is assumed that $\int_\Omega ||U_i(\omega)||_\infty dP(\omega) < \infty$. A behavior strategy is a $\mathcal{F}_i$-measurable map $b_i\colon \Omega_i \to \Delta(A_i)$. Let $\mathbb{B}_i$ be the set of behavior strategies and let $\mathbb{B} := \prod_i \mathbb{B}_i$. A measure $\nu \in \Delta(\prod_i \mathbb{B}_i)$ is a correlated equilibrium[9] if for every i and measurable deviation $\gamma_i\colon \Omega_i \times \Delta(A_i) \to \Delta(A_i)$,

$$\int_{\mathbb{B}} \int_\Omega \left(\sum_{a \in A} U_i(\omega)(a) \prod_i b_i(a|\omega) - \sum_{a \in A} U_i(\omega)(a) \cdot \gamma_i(\omega, b_i(a|\omega)) \cdot \prod_{j \neq i} b_j(a|\omega) \right) dP(\omega) d\nu(b) \geq 0.$$

We show that for every correlated equilibrium (P, ν) there is an outcome equivalent equilibrium (P, σ). We now translate this set-up into ours for any choice of general information structure $\mathcal{S}$: Ω_i must correspond to $\Theta \times T_i$, where $\mathcal{F}_i$ is generated by the projection of $\Theta \times T_i$ onto T_i . Payoffs translate into our set-up by setting $U_i(\theta, \tau)(a) = u_i(\theta, a)$. The condition that $\int_\Omega ||U_i(\omega)||_\infty dP(\omega) < \infty$ is then satisfied since Θ is finite and payoffs depend Θ and not T. A behavior strategy corresponds to a marginal decision rule, $\sigma(a_i|t_i)$.It remains to show that a distribution ν over profiles of marginal decision rules as in Stinchcombe (2011) is equivalent to a belief invariant decision rule. First, note that every measure $\nu \in \Delta(\prod_i \mathbb{B}_i)$ induces a belief invariant decision rule σ_ν where for every measurable event $E \subseteq A$

$$\sigma_\nu(a|\omega) = \int_{\mathbb{B}} \prod_i b_i(\tau_i)(a_i) \; d\nu(b).$$

Indeed, the resulting marginal probability $\sigma_\nu(a_i|\omega) = \int_{\mathbb{B}_i} b_i(\tau_i)(a_i) \; \mathrm{d}\nu(b_i)$ verifies belief invariance.

[9]In Stinchcombe (2011) this corresponds to a variation Strategy Correlated Equilibrium where deviation strategies depend action recommendations at realized types only.

7.2 Proof of Proposition 8

Let $(\mathcal{G}, P)$ be base game and an information structure, i.e. so that $P \in \mathcal{P}$. Consider an information structure $\hat{\mathcal{S}} = ((S_i, \hat{P}_i)_{i \in I}, \hat{P})$ satisfying $P = P_{\hat{\mathcal{S}}}$ and let σ' be a BIBCE of $(\mathcal{G}, \hat{\mathcal{S}})$. We construct an outcome equivalent BIBCE σ of $(\mathcal{G}, \hat{\mathcal{S}})$. For every $\theta \in \Theta$ and P-almost every $\tau \in \mathcal{T}$, let

$$\sigma(a|\theta, \tau) = \int_{\overline{\tau}^{-1}(\tau)} \sigma'(a|\theta, s') \, \mathrm{d}\hat{P}(s'|\theta, \overline{\tau}^{-1}(\tau)),$$

where $\tau \mapsto \hat{P}(\cdot|\theta, \overline{\tau}^{-1}(\tau)) \in \Delta(\overline{\tau}^{-1}(\tau))$ is a conditional probability on S. Then obedience constraints for any P-almost every type $\tau_i \in \mathcal{T}_i$ are satisfied

$$\sum_{\theta, a_{-i}} (u_i(a, \theta) - u_i(a_i', a_{-i}, \theta)) \int_{S_{-i}} \sigma(a|\theta, \tau) \, \mathrm{d}P(\tau_{-i}, \theta|\tau_i) =$$
$$\sum_{\theta, a_{-i}} (u_i(a, \theta) - u_i(a_i', a_{-i}, \theta)) \int_{S_{-i}} \int_{\overline{\tau}^{-1}(\tau)} \sigma'(a|\theta, s') \, \mathrm{d}\hat{P}(s'|\theta, \overline{\tau}^{-1}(\tau)) \, \mathrm{d}P(\tau_{-i}, \theta|\tau_i).$$

Noting that

$$\int_{\mathcal{T}_{-i}} \int_{\overline{\tau}^{-1}(\tau)} \sigma'(a|\theta, s') \, \mathrm{d}\hat{P}(s'|\theta, \overline{\tau}^{-1}(\tau)) \, \mathrm{d}P(\tau_{-i}, \theta|\tau_i) = \int_{S_{-i}} \sigma'(a|\theta, s') \, \mathrm{d}\hat{P}(s_{-i}', \theta|s_i),$$

we conclude that σ is a BIBCE of $(\mathcal{G}, P)$. Conversely, note that every BIBCE σ of $(\mathcal{G}, P)$ induces a $id \times \overline{\tau}$-measurable decision rule σ', where for every action profile $a \in A$, state $\theta \in \Theta$ and $\hat{P}$-almost every $s' \in S$,

$$\sigma'(a|\theta, s') = \sigma(a|\theta, \overline{\tau}(s')).$$

Performing a change of variables

$$\int_{S_{-i}} \sigma'(a|\theta, s') \, \mathrm{d}\hat{P}(s_{-i}', \theta|s_i) = \int_{\mathcal{T}_{-i}} \int_{\overline{\tau}^{-1}(\tau)} \sigma'(a|\theta, s') \, \mathrm{d}\hat{P}(s'|\theta, \overline{\tau}^{-1}(\tau)) \, \mathrm{d}P(\tau_{-i}, \theta|\tau_i)$$
$$= \int_{\mathcal{T}_{-i}} \sigma(a|\theta, \tau) \, \mathrm{d}P(\tau_{-i}, \theta|\tau_i).$$

and so obedience follows.

7.3 Proof of Proposition 9

Let $(\sigma_i\colon S_i \to \Delta(A_i))_{i\in I}$ be a BNE under $(\mathcal{G}, \mathcal{S})$. Consider the following representation of $\mathcal{S}$: Define the graph of $\overline{\tau}$, $\phi_{\overline{\tau}}\colon s \mapsto (\overline{\tau}(s), s)$ and consider the push forward probability $P^* := P \circ (id \times \phi_{\overline{\tau}})^{-1}$. Let $\sigma^R\colon \Theta \times \mathcal{T} \to \Delta(S)$ denote a $(id \times \overline{\tau})$-conditional probability of P^* so that for any measurable $E \subseteq S$ and state $\theta \in \Theta$ in the support of P,

$$P(E|\theta) = \int_{\mathcal{T}} \sigma^R(E|\theta, \tau) \mathrm{d}P^*(\tau|\theta).$$

We now construct a BIBCE σ^* on $\mathrm{marg}_{\Omega} P^*$ as follows:

$$\sigma^*(a|\theta, \tau) = \int_S \prod_{i\in I} \sigma_i(a|s_i)\ \mathrm{d}\sigma^R(s|\theta, \tau).$$

Since σ satisfies obedience, so does σ^*:

$$\sum_{\theta, a_{-i}} (u_i(a, \theta) - u_i(a_i', a_{-i}, \theta)) \int_{\mathcal{T}_{-i}} \sigma^*(a|\theta, \tau)\ \mathrm{d}\mathrm{marg}_{\Omega} P^*(\theta, \tau_{-i}|\tau_i)$$
$$= \sum_{\theta, a_{-i}} (u_i(a, \theta) - u_i(a_i', a_{-i}, \theta)) \int_S \prod_{i\in I} \sigma_i(a|s_i)\ \mathrm{d}P(\theta, s_{-i}|s_i).$$

7.4 Proof of Proposition 1

The ACK topology is metrizable by the metric

$$d^{f,g}(P, P') = \inf \left\{ \varepsilon \geq 0 : \begin{array}{l} P\left(C^{1-g(\varepsilon)}\left(\hat{T}_{g(\varepsilon)}(P, P')\right)\right) > f(\varepsilon) \\ P'\left(C^{1-g(\varepsilon)}\left(\hat{T}_{g(\varepsilon)}(P, P')\right)\right) > f(\varepsilon) \end{array} \right\}.$$

where $g(\varepsilon) := \varepsilon^2$ and $f(\varepsilon) := 1 - \frac{1}{2}\varepsilon(\varepsilon + 1)$.

Proof. First note that $d^{f,g}(P,P) = 0$ and $d^{f,g}(P,P') = d^{f,g}(P',P)$ are both immediate. Suppose now that $d^{f,g}(P,P') = 0$. Then we have that $\text{supp}(P) = \text{supp}(P')$, which by the definition of $\mathcal{P}$ means that $P = P'$. It remains to show that $d^{f,g}$ satisfies the triangle inequality. Let $P^1, P^2, P^3 \in \mathcal{P}$ satisfy $d^{f,g}(P^1,P^2) < \varepsilon_{1,2}$ and $d^{f,g}(P^3,P^2) < \varepsilon_{3,2}$, for $\varepsilon_{1,2}, \varepsilon_{3,2} \in [0,1]$. Note that for every $\omega \in \hat{T}_{g(\varepsilon_{1,2})}(P^1,P^2) \cap \hat{T}_{g(\varepsilon_{2,3})}(P^2,P^3) \subseteq \text{supp}_{g(\varepsilon_{1,2})}(P^1) \cap \text{supp}_{g(\varepsilon_{3,2})}(P^3)$ there is $\omega' \in \text{supp}(P^2)$ so that

$$d_\Pi(\omega,\omega') \leq \underline{\varepsilon} := \min\left\{g(\varepsilon_{1,2}), g(\varepsilon_{3,2})\right\}.$$

Moreover, $P^2(F) > 1 - (1 - f(\varepsilon_{1,2}) + 1 - f(\varepsilon_{3,2}))$, where

$$F := C^{1-g(\varepsilon_{1,2})}\left(\hat{T}_{g(\varepsilon_{1,2})}(P^1,P^2)\right) \cap C^{1-g(\varepsilon_{3,2})}\left(\hat{T}_{g(\varepsilon_{2,3})}(P^2,P^3)\right) \subseteq C^{1-g(\varepsilon_{1,2}+\varepsilon_{3,2})}\left(\hat{T}_{g(\varepsilon_{1,2}+\varepsilon_{3,2})}(P^1,P^3)\right),$$

and the last containment is due to the fact that $g(\varepsilon_{1,2}) + g(\varepsilon_{3,2}) = \varepsilon_{1,2}^2 + \varepsilon_{3,2}^2 \leq (\varepsilon_{1,2} + \varepsilon_{3,2})^2 = g(\varepsilon_{1,2} + \varepsilon_{3,2})$. Consider a P^1, P^3 and P^2 measurable, surjective map $\phi\colon F \to F \cap \text{supp}(P^2)$ so that for all $\omega \in F$, $d_\Pi(\phi(\omega),\omega) < \underline{\varepsilon}$. Letting $F_i := \text{proj}_{\mathcal{T}_i}(F)$,

$$\begin{aligned} |P^k(F) - P^2(F)| &= \left| \int_{F_i} \tau_i(F) P^k(\mathrm{d}\tau_i) - \int_{F_i} \tau_i(F) P^2(\mathrm{d}\tau_i) \right|. \\ &\leq \underline{\varepsilon} \end{aligned}$$

Hence $P^k(F) \geq f(\varepsilon_{1,2}) + f(\varepsilon_{3,2}) - 1 - \underline{\varepsilon}$. Then for any $\varepsilon_{1,2}, \varepsilon_{3,2} \in [0,1]$,

$$P^k\left(C^{1-g(\varepsilon_{1,2}+\varepsilon_{3,2})}\left(\hat{T}_{g(\varepsilon_{1,2}+\varepsilon_{3,2})}(P^1,P^3)\right)\right) \geq f(\varepsilon_{1,2}) + f(\varepsilon_{3,2}) - 1 - \underline{\varepsilon} \quad .$$

Given $g(\varepsilon) = \varepsilon^2$ and $f(\varepsilon) = 1 - \frac{1}{2}\varepsilon(\varepsilon+1)$. Then we have

$$\begin{aligned}
f(\varepsilon_{1,2}) + f(\varepsilon_{3,2}) - \underline{\varepsilon} - 1 &\geq f(\varepsilon_{1,2} + \varepsilon_{3,2}) \\
1 - \frac{1}{2}\varepsilon_{1,2}(\varepsilon_{1,2}+1) + 1 - \frac{1}{2}\varepsilon_{3,2}(\varepsilon_{3,2}+1) - \underline{\varepsilon} - 1 &\geq 1 - \frac{1}{2}\left(\varepsilon_{1,2} + \varepsilon_{3,2}\right)\left(\varepsilon_{1,2} + \varepsilon_{3,2} + 1\right) \\
-\frac{1}{2}\varepsilon_{1,2}(\varepsilon_{1,2}+1) - \frac{1}{2}\varepsilon_{3,2}(\varepsilon_{3,2}+1) - \underline{\varepsilon} &\geq -\frac{1}{2}\left(\varepsilon_{1,2} + \varepsilon_{3,2}\right)\left(\varepsilon_{1,2} + \varepsilon_{3,2} + 1\right) \\
(\varepsilon_{1,2}^2 + \varepsilon_{3,2}^2 + \varepsilon_{1,2} + \varepsilon_{3,2}) + 2\underline{\varepsilon} &\leq \left((\varepsilon_{1,2} + \varepsilon_{3,2})^2 + \varepsilon_{1,2} + \varepsilon_{3,2}\right) \\
(\varepsilon_{1,2}^2 + \varepsilon_{3,2}^2 + \varepsilon_{1,2} + \varepsilon_{3,2}) + 2\underline{\varepsilon} &\leq \left(\varepsilon_{1,2}^2 + \varepsilon_{3,2}^2 + \varepsilon_{1,2} + \varepsilon_{3,2}\right) + 2\varepsilon_{1,2}\varepsilon_{3,2} \\
\min_{k\in\{1,3\}} \varepsilon_{k,2}^2 &\leq \varepsilon_{1,2}\varepsilon_{3,2} \\
\min_{k\in\{1,3\}} \varepsilon_{k,2} &\leq \max_{k\in\{1,3\}} \varepsilon_{k,2}.
\end{aligned}$$

Hence $d^{f,g}(P^1, P^3) < \varepsilon_{1,2} + \varepsilon_{3,2}$ and so $d^{f,g}$ satisfies the triangle inequality. □

7.5 Proof of Lemma 1

Proof. For every precision $z \in \mathbb{N}$ define a finite grid approximation recursively: Consider the grid on first order beliefs, given by

$$A_i^{1,z} := \left\{ a \in \mathbb{R}^{\Theta} : a_\theta \in \left\{ \frac{n}{z} : 0 \leq n \leq z \right\}, \sum_\theta a_\theta = 1 \right\}$$

for any player i. Given a finite set $A_i^{m-1,z} \subseteq \mathcal{T}_i^{m-1}$ for every i, define

$$A_i^{m,z} := \left\{ a \in \mathbb{R}^{\prod_{n<m} A_{-i}^{n,z} \times \Theta} : \begin{array}{c} a_{a^1,\ldots,a^{m-1},\theta} \in \left\{ \frac{n}{z} : 0 \leq n \leq z \right\}, \\ \sum_{a^1,\ldots,a^{m-1},\theta} a_{a^1,\ldots,a^{m-1},\theta} = 1 \end{array} \right\}. \tag{4}$$

Then for any $m \in \mathbb{N}$, and any $t^m \in \mathcal{T}^m$,there is $a^m \in \prod_{i\in I} A_i^{m,z}$ so that $d_{Weak}^m(t^m, a^m) < 1/z$. Then there is m, z large enough so that for every, $\theta \in \Theta$, every $g \in \mathcal{T}$satisfying $(g^1, \ldots g^m) =$

a^m and $\tau \in \mathcal{T}$satisfying $(\tau^1,\ldots,\tau^m) = t^m$,

$$\begin{aligned} d_\Pi((\theta,\tau),(\theta,g)) =& \sum_{n=1}^{\infty} \eta^n d^m_{Weak}(\tau^1,\ldots,\tau^n,g^1,\ldots,g^n) \\ =& \frac{\eta}{z}\frac{1-\eta^{m+1}}{1-\eta} + \eta^m \frac{1}{1-\eta} \\ <& \varepsilon. \end{aligned}$$

Let $G_\varepsilon := \prod_{i\in I} A_i^{m,z} \cap (\Omega \setminus \mathcal{N}_\varepsilon(E))$. Suppose that $d_\Pi(\omega,\omega_g) < \varepsilon$ for some $\omega \in \Omega$ and $\omega_g \in G_\varepsilon$.Then clearly, it must be that $\omega \notin E$, since the opposite would imply that $\omega_g \in \mathcal{N}_\varepsilon(E) \cap G_\varepsilon$. Suppose now that $\omega \in \Omega \setminus \mathcal{N}_\varepsilon(E)$ and so by the above there is $\omega_g \in G_\varepsilon$ so that $d_\Pi(\omega',\omega_g) < \varepsilon$. □

7.6 Proof of Lemma 2 (Iterative Scoring Rule)

Proof. Fix m, z. and let $A^{m,z}$ be defined as in 4. Following Dekel et al. (2006), for every profile $(a^1,\ldots,a^m,\theta) \in A^{m,z} \times \Theta$ we define

$$\begin{aligned} u_i^m\left(a^1,\ldots,a^m,\theta\right) := 2a_i^1(\theta) - \sum_{\hat\theta}\left(a_i^1(\hat\theta)\right)^2 \\ + \sum_{n=2}^{m}\left(2a_i^n(a^1,\ldots,a^{n-1},\theta) - \sum_{\hat a^1,\ldots,\hat a^{n-1},\hat\theta}\left(a_i^n(\hat a^1,\ldots,\hat a^{n-1},\hat\theta)\right)^2\right). \end{aligned}$$

Dekel et al. (2006) show that the game where payoffs are given by u_i^m, the uniquely interim (correlated) rationalizable action profile of τ is the profile $a^1,\ldots,a^m$ which is $1/z$-close to τ in terms of distance d^m_{Weak} on $\mathcal{T}^m$. The result then follows from Claim 1.

$\square$

7.7 Open Set Definition of Strategic Topology

Definition 23. (Strategic Distance) A function $\overline{d}_{\mathcal{P}} \colon \mathcal{P} \times \mathcal{P} \to [0, \infty)$ is a strategic distance if,

1. for every game $\mathcal{G}$ and $\varepsilon > 0$, there exists $\delta_{\varepsilon,\mathcal{G}} > 0$ so that for all $P, P' \in \mathcal{P}$ satisfying $\overline{d}_{\mathcal{P}}(P, P') < \delta_{\varepsilon,u}$, we have that $d^u(P, P') < \varepsilon$,
2. for every $\varepsilon > 0$ and $P, P' \in \mathcal{P}$ satisfying $\sup_u d^u(P, P') < \varepsilon$, we also have $\overline{d}_{\mathcal{P}}(P, P') < \varepsilon$.

7.8 Proof of Proposition 13

Proof. Let d induce a nice, common belief invariant topology on Ω which is a stronger topology than the one induced by d_Π and we have that $\hat{T}_{d,\varepsilon}(P, P') \subseteq \hat{T}_\varepsilon(P, P')$. By Proposition 17 we deduce that $\hat{T}_{d,\varepsilon}(P, P') = C^{1-\varepsilon}\left(\hat{T}_{d,\varepsilon}(P, P')\right) \subseteq C^{1-\varepsilon}\left(\hat{T}_\varepsilon(P, P')\right)$ and so

$$\begin{aligned} \min\left\{P(\hat{T}_{d,\varepsilon}(P, P')), P(\hat{T}_{d,\varepsilon}(P, P'))\right\} > 1 - \varepsilon &\implies d^{ACK}(P, P') < \varepsilon \\ &\implies \forall\, \mathcal{G},\ \exists\, N > 0 \text{ s.t. } d^*(P, P'|\mathcal{G}) < N\varepsilon. \end{aligned}$$

We now show that for every $\varepsilon > 0$ there is $\delta > 0$ so that $P(C^{1-\delta}\left(\hat{T}_\delta(P, P')\right)) > 1 - \delta \implies P(\hat{T}_{d,\varepsilon}(P, P')) > 1-\varepsilon$. Suppose that $\text{supp}(P) \neq \text{supp}(P')$ (otherwise the condition is trivially satisfied). There is a countable subset $\Omega^0_d \subseteq \Omega^0$ which is dense under d. Suppose that for every $\varepsilon > 0$ there is a finite grid $G_{d,\varepsilon}$ whose ε-neighborhood covers a subset $H_\varepsilon \subseteq \hat{T}_{d,\varepsilon}(P, P')$ with $P(H_\varepsilon) > 1 - \varepsilon$. Since Ω is Hausdorff under the product topology, there is a δ so that $\min_{g,g' \in G_{d,\varepsilon} : g \neq g'} d_\Pi(g, g') > \delta$ and so the result follows. It remains to prove the existence of such a grid. We show that for any nice metric d, Ω is Polish. To establish that Ω is Polish it is enough to show that it is complete under d. Let $(\omega^k)_k$ be Cauchy. Since d refines the product topology, the sequence $(\omega^k)_k$ is Cauchy under the product topology. Since Ω is

Polish under the product topology (see Mertens et al. (2015)), we deduce that Ω is Polish under d. Then P is regular and so every set, in particular the set $\hat{T}_{d,\varepsilon}(P, P')$, admits a compact approximation H_ε from below. By compactness, H_ε admits such a grid. $\square$

7.9 Interim Strategic Topology

We follow Chen et al. (2017) in describing the interim strategic topology.

Definition 24. (Frame) A frame is a profile of maps $(\pi_i)_{i\in I}$ where for every $i \in I$, $\pi_i \colon \mathcal{T}_i \to F_i$, F_i is a finite set and for every prior $P \in \mathcal{P}$ and all types $\tau_i, \tau_i' \in \mathcal{T}_i$,

$$P(\cdot,\cdot|\tau_i) \circ (id \times \pi_{-i})^{-1} = P(\cdot,\cdot|\tau_i') \circ (id \times \pi_{-i})^{-1} \implies \pi_i(\tau_i) = \pi_i(\tau_i').$$

For any set $E \subseteq \Omega$ and any player i , let $E_i \subseteq \mathcal{T}_i$ denote the projection of E on $\mathcal{T}_i$. Let $\mathscr{F}$ denote the collection of events that are measurable with respect to a frame.

$$\mathscr{F} = \left\{E \in \mathscr{B} : \exists\ \pi \in \Pi, \tau \in E \text{ s.t. } \forall\ i \in I,\ \pi_i^{-1}(\pi_i(\tau)) = E_i\right\}.$$

The formal definition is recursive:

Definition 25. (Uniform Weak Distance on Frames) For all pairs $\omega = (\theta, \tau), \hat{\omega} = (\hat{\theta}, \hat{\tau}) \in \Omega$, define the Uniform Prokohorov Distance on Frames by $d^{\mathscr{F}}(\omega, \hat{\omega}) := \sup_m d^m(\omega, \hat{\omega})$, where the sequence of functions $(d^m)_m$ on $\Omega \times \Omega$ is defined recursively from a metric representing the weak* topology on profiles of first order beliefs $d^1(\omega, \hat{\omega}) = d_{\Delta(\Theta)^I}(\tau^1, \hat{\tau}^1)$, and for every $m > 1$,

$$d^m(\omega, \hat{\omega}) := d_\Theta(\theta, \hat{\theta}) + \inf\{\delta > 0 : \forall\ i,\ \tau_i(E) \leq \hat{\tau}_i(\mathcal{N}_{d^{m-1},\delta}(E)) + \delta,\ \forall\ E \in \mathscr{F}\}.$$

Definition 26. (Interim Strategic Topology) Define the uniform weak topology on frames as the topology on Ω generated by the sets $\left\{\omega' \in \Omega : d^{\mathscr{F}}(\omega, \omega') < \varepsilon\right\}$ for $\omega \in \Omega$ and $\varepsilon > 0$.

Proposition 16. *(Niceness) The interim strategic topology is nice.*

Proof. Dekel et al. (2006) show that a countable set of finite states is dense in Ω in the interim strategic topology. The interim strategic topology is a stronger topology than the product topology and is metrizable (see Dekel et al. (2006)). □

Proposition 17. *(Common Belief Invariance) For every canonical $P, P' \in \mathcal{P}$ and $\varepsilon > 0$,*

$$C_{1-\varepsilon}\left(\hat{T}_{d^{\mathscr{F}},\varepsilon}(P, P')\right) = \hat{T}_{d^{\mathscr{F}},\varepsilon}(P, P').$$

Proof. Consider $\omega = (\theta, \tau) \in \Omega_P$ and $\hat{\omega} = (\hat{\theta}, \hat{\tau}) \in \Omega_{P'}$ with $d^{\mathscr{F}}(\omega, \hat{\omega}) < \varepsilon$. Then it must be that $\omega, \hat{\omega} \in \hat{T}_{d^{\mathscr{F}},\varepsilon}(P, P')$. Since $\Omega_P \in \mathscr{F}$, the fact that $d^{\mathscr{F}}(\omega, \hat{\omega}) < \varepsilon$ implies that

$$\tau_i^*(\Omega_P) \leq \hat{\tau}_i^*(\mathcal{N}_{d^{\mathscr{F}},\varepsilon}(\Omega_P)) + \varepsilon.$$

Since $\tau_i^*(\Omega_P) = 1$ and $\hat{\tau}_i^*(\Omega_{P'}) = 1$ we have

$$\hat{\tau}_i^*(\hat{T}_{d^{\mathscr{F}},\varepsilon}(P, P')) = \hat{\tau}_i^*(\mathcal{N}_{d^{\mathscr{F}},\varepsilon}(\Omega_P)) \geq 1 - \varepsilon.$$

A symmetric argument implies that $\tau_i^*(\hat{T}_{d^{\mathscr{F}},\varepsilon}(P, P')) \geq 1 - \varepsilon$ and so $\hat{T}_{d^{\mathscr{F}},\varepsilon}(P, P') = B^{1-\varepsilon}(\hat{T}_{d^{\mathscr{F}},\varepsilon}(P, P'))$, which establishes the result. □

7.10 Continuity of Exact BIBCE in Rich Games

Our main result established that the ACK topology is the coarsest topology generating continuity in approximate BIBCE outcomes. For many applications of interest, we would like to make statements about continuity of exact BIBCE outcomes. We can extend our results to exact BIBCE if we allow for sufficiently rich games.

We introduce a "strong richness" property of a base game.

Definition 27. (strong richness) A base game $\mathcal{G}$ satisfies *strong richness* if there exists $\theta^* \in \Theta$ such that for every player i and every $a_i \in A_i$, there exists $a_{-i} \in A_{-i}$ such that (a_i, a_{-i}) is a strict Nash Equilibrium of $(u_j(\cdot, \cdot, \theta^*))_{j \in I}$.

Richness requires that the set of possible payoffs is sufficiently rich. It is in the spirit of but stronger than richness properties in the literature, e.g., Weinstein and Yildiz (2007). Recall that it is a maintained assumption that every state $\theta \in \Theta$ is assigned positive probability.

Lemma 5. *For every base game $\mathcal{G}$ satisfying strong richness, $\varepsilon > 0$; and simple information structure $P \in \mathcal{P}$, there exists a simple information structure $P^+ \in \mathcal{P}$ so that $d^{ACK}(P, P^+) < \varepsilon$ and*

$$d_{\mathcal{H}}(\mathcal{O}^{\varepsilon}(\mathcal{G}, P), \mathcal{O}(\mathcal{G}, P^+)) < 2M|A \times \Theta|\varepsilon,$$

where $d_{\mathcal{H}}(X, Y)$ is the Hausdorff distance between $X, Y \subseteq \Delta(A \times \Theta)$.

Proof. By richness part (2), for any player i and any action $a_i \in A_i$ there is $\alpha_{-i}(a_i) \in A_{-i}$ so that $(a_i, \alpha_{-i}(a_i))$ is a strict NE in the game with payoffs $(\overline{u}_j(\cdot, \cdot) := u_j(\cdot, \cdot, \theta^*))_{j \in I}$. We first perturb P so that every type assigns probability at least $\delta > 0$ to θ^* by adding a new profile of types $t^{\tau_i} = (t_j^{\tau_i})_{j \in I}$ for every player i and every type $\tau_i \in T_i := \{\tau_i : \tau \in \text{supp}(P)\}$. Define the extended prior P^+ with support given by

$$T^+ := \text{supp}(P) \cup \left(\bigcup_{i \in I} (T^i \cup \bar{T}^i) \right),$$

where $T^i := \left\{(\theta^*, \tau_i, t_{-i}^{\tau_i}) : \tau \in \text{supp}(P)\right\}$ and $\bar{T}^i := \{(\theta^*, t^{\tau_i}) : \tau_i \in T_i\}$. For every $(\theta, \tau) \in T^+$ and any given choices $\delta, \eta \in [0, 1)$, define

$$P^+(\theta, \hat{\tau}) := \begin{cases} (1-\eta)(1-\delta)P(\theta, \hat{\tau}) & \text{if } (\theta, \hat{\tau}) \in \text{supp}(P), \\ (1-\eta)\delta/|I| \sum_{i \in I} \mathbf{1}_{(\theta, \hat{\tau}) \in T^i} P(\hat{\tau}_i)\mu(\theta) & \text{if } (\theta, \hat{\tau}) \in \cup_{i \in I} T^i, \\ \eta/|I| \sum_{i \in I} \mathbf{1}_{(\theta, \hat{\tau}) \in \bar{T}^i} \sum_{\tau_i} \mathbf{1}_{\hat{\tau} = t^{\tau_i}} P(\tau_i)\mu(\theta) & \text{if } (\theta, \hat{\tau}) \in \cup_{i \in I} \bar{T}^i. \end{cases}$$

For every $\sigma \in \mathcal{B}^{\varepsilon}(\mathcal{G}, P)$ we construct an associated BIBCE $\sigma^+ \in \mathcal{B}(\mathcal{G}, P^+)$ as follows:

$$\sigma^+(a|\theta, \hat{\tau}) = \begin{cases} \sum_{i \in I} \mathbf{1}_{(\theta,\hat{\tau}) \in T^i \cup \overline{T}^i} \mathbf{1}_{a=(a_i, \alpha_{-i}(a_i))} \sigma(a_i|\hat{\tau}_i) & \text{if } (\theta, \hat{\tau}) \in \cup_{i \in I} \left(T^i \cup \bar{T}^i\right), \\ \sigma(a|\theta, \tau) & \text{otherwise.} \end{cases}$$

We first verify that $\sigma^+$is belief invariant:

$$\sum_{a_{-i}} \sigma^+(a|\theta, \hat{\tau}_{-i}, \hat{\tau}_i) = \begin{cases} \sigma(a_i|\hat{\tau}_i) & \text{if } (\theta, \hat{\tau}) \in \text{supp}(P), \\ \sigma(a_i|\hat{\tau}_i) & \text{if } (\theta, \hat{\tau}) \in T^i \cup \bar{T}^i, \\ \sum_j \mathbf{1}_{(\theta,\hat{\tau}) \in T^j \cup \bar{T}^j} \sigma(a_i|\hat{\tau}_j) & \text{if } (\theta, \hat{\tau}) \in \cup_{j \neq i} T^j \cup \bar{T}^j, \end{cases}$$

and so $(\theta, \hat{\tau}) \mapsto \sigma(a_i|\theta, \hat{\tau})$ is measurable with respect to the projection $(\theta, \hat{\tau}) \mapsto \hat{\tau}_i$ for all $a_i \in A_i$. This decision rule increases the expected payoff of σ for every type in the support of P: Let $J_{\mathcal{G}} := \min_{a_i, \hat{a}_i \in A_i, i \in I} \left(\bar{u}_i(a_i, \alpha_{-i}(a_i)) - \bar{u}_i(\hat{a}_i, \alpha_{-i}(a_i))\right) > 0$ and define $\delta := \frac{\varepsilon}{J_{\mathcal{G}}+\varepsilon} < \varepsilon$. Consider any $\tau_i \in T_i$ and deviation a'_i

$$\sum_{\theta,\tau} \sum_a \Delta u_i(a, a'_i, \theta) P^+ \circ \sigma^+(a, \theta, \tau_{-i}|\tau_i) > -\varepsilon(1-\delta) + \delta \sum_{a_i} \Delta\bar{u}_i(a_i, \alpha_{-i}(a_i), a'_i)\ \sigma(a_i|\tau_i)$$

$$\geq -\varepsilon(1-\delta) + J_{\mathcal{G}}\delta = 0.$$

It remains to check if obedience also holds for each type $t_i^{\tau_j}$ of player i and any associated $\tau_j \in T_j$ of player j:

$$\sum_{\theta,\tau} \sum_a \Delta u_i(a, a'_i, \theta)\ P^+ \circ \sigma^+(a, \theta, \tau_{-i}|t_i^{\tau_j}) = \sum_{a_j} \Delta\bar{u}_i(a_j, \alpha_{-j}(a_j), a'_i)\ \sigma(a_j|\tau_j)$$

$$\geq J_{\mathcal{G}} > 0.$$

Hence $\sigma^+ \in \mathcal{B}(\mathcal{G}, P^+)$ and outcomes $\nu_\sigma \in \mathcal{O}^{\varepsilon}(\mathcal{G}, P)$ and $\nu_{\sigma^+} \in \mathcal{O}(\mathcal{G}, P^+)$ satisfy $||\nu_\sigma - \nu'||_2 \leq 2M\frac{\varepsilon}{J_{\mathcal{G}}+\varepsilon} < 2M\varepsilon$. Let $g_i(\tau) := \overline{\tau}_i\left(P^+(\cdot, \cdot|\tau_i)\right)$ denote the new canonical type/belief hierarchy of player i, for each $\tau \in \text{supp}(P)$. Then for every event $E \subseteq \text{supp}(P)$ and any player i, $|g_i(\tau)(E) - \tau_i(E)| \leq \delta$. Since the total variation norm is stronger than the product topology, we conclude that $d_\Pi(g(\tau), \tau) \leq \delta < \varepsilon$. Hence $P^+(\mathcal{N}_\varepsilon(\text{supp}(P))) > 1 - \varepsilon$ and so $d^{ACK}(P, P^+) < \varepsilon$. Hence for every $\nu' \in \mathcal{O}(\mathcal{G}, P^+)$ there is $\nu \in \mathcal{O}^\varepsilon(\mathcal{G}, P)$ so that $||\nu - \nu'||_2 \leq 2M|A \times \Theta|\varepsilon$. □

Proposition 18. *For every base game $\mathcal{G}$ and any information structure $P \in \mathcal{P}$ there is a sequence $(\hat{P}^k)_k$ of simple information structures ACK-converging to P and a subset $\mathcal{O}^\infty(\mathcal{G}, P) \subseteq \mathcal{O}(\mathcal{G}, P)$ so that $\lim_{k\uparrow\infty} d_\mathcal{H}(\mathcal{O}(\mathcal{G}, \hat{P}^k), \mathcal{O}^\infty(\mathcal{G}, P)) = 0$. If $\mathcal{G}$ satisfies strong richness, then*

$$\lim_{k\uparrow\infty} d_\mathcal{H}(\mathcal{O}(\mathcal{G}, \hat{P}^k), \mathcal{O}(\mathcal{G}, P)) = 0.$$

Proof. By Proposition 6 there is a sequence of simple priors P^k converging strategically to P and a sequence of positive $(\varepsilon^k)_k$ converging monotonically to 0 so that for every k, $d^*(P^k, P) < \varepsilon^k$. We first prove upper hemi-continuity of BIBCE with respect to strategic convergence. Let $(\sigma^k)_k$ be a sequence of BIBCE so that for every $k \in \mathbb{N}$, $\sigma^k \in \mathcal{B}^0(\mathcal{G}, P^k)$ and the induced sequence of outcomes $(\nu_{\sigma^k})_k$ converges to some outcome ν_∞, i.e. $\lim_{k\uparrow\infty} ||\nu_{\sigma^k} - \nu_\infty||_2 = 0$. We have shown that for every k there is a ε^k-BIBCE $\sigma^{\varepsilon^k} \in \mathcal{B}^{\varepsilon^k}(\mathcal{G}, P)$ so that $||\nu_{\sigma^k} - \nu_{\sigma^{\varepsilon^k}}||_2 \leq 4M|A \times \Theta|\varepsilon^k$. Our construction in the proof of Proposition 4 has the feature that the sequence $\left(\sigma^{\varepsilon^k}\right)_k$ converges almost surely to an obedient decision rule (as it was constructed from a conditional expectation) and so there is a BIBCE $\sigma^\infty \in \mathcal{B}(\mathcal{G}, P)$ so that $\nu_{\sigma^\infty} = \nu_\infty$, which establishes upper hemi-continuity: There is a subset $\mathcal{O}^\infty(\mathcal{G}, P) \subseteq \mathcal{O}(\mathcal{G}, P)$ so that

$$\lim_{k\uparrow\infty} d_{\mathcal{H},\mathcal{G}}(\mathcal{O}(\mathcal{G}, P^k), \mathcal{O}^\infty(\mathcal{G}, P)) = 0.$$

Lower hemi-continuity is a consequence of Lemma 5 and richness part (2) of $\mathcal{G}$. Indeed, for every BIBCE $\sigma \in \mathcal{B}(\mathcal{G}, P)$ and every $k \in \mathbb{N}$, $d^*(P, P^k) < \varepsilon^k$ implies that there is $\sigma^{\varepsilon^k,k} \in \mathcal{B}^{\varepsilon^k}(\mathcal{G}, P^k)$ so that $||\nu_{\sigma^{\varepsilon^k,k}} - \nu_\sigma||_2 < \varepsilon^k$. Hence, there is a superset $\overline{\mathcal{O}}^\infty(\mathcal{G}, P) \supseteq \mathcal{O}(\mathcal{G}, P)$ so that $\lim_{k\uparrow\infty} d_{\mathcal{H},\mathcal{G}}(\mathcal{O}^{\varepsilon^k}(\mathcal{G}, P^k), \overline{\mathcal{O}}^\infty(\mathcal{G}, P)) = 0$. From the construction in the proof of Lemma

5 there is simple $\hat{P}^k \in \mathcal{P}$ so that $d^*(P^k, \hat{P}^k) < \varepsilon^k$ so that $d_{\mathcal{H},\mathcal{G}}(\mathcal{O}^{\varepsilon^k}(\mathcal{G}, P^k), \mathcal{O}(\mathcal{G}, \hat{P}^k)) < 2M(u)\varepsilon^k$. By the triangle inequality, the sequence $(\hat{P}^k)_k$ also converges strategically to P. By upper hemi-continuity established earlier and the triangle inequality there is a subset $\underline{\mathcal{O}}^\infty(\mathcal{G}, P) \subseteq \mathcal{O}(\mathcal{G}, P)$ so that $\lim_{k\uparrow\infty} d_{\mathcal{H},\mathcal{G}}(\mathcal{O}^{\varepsilon^k}(\mathcal{G}, P^k), \underline{\mathcal{O}}^\infty(\mathcal{G}, P)) = 0$. But then we have that $\overline{\mathcal{O}}^\infty(\mathcal{G}, P) = \underline{\mathcal{O}}^\infty(\mathcal{G}, P) = \mathcal{O}(\mathcal{G}, P)$ and so the result follows. □

Focusing attention on BIBCE as a solution concept, we have:

Lemma 6. *For any game $\mathcal{G}$ and any open $\mathcal{P}^* \subseteq \mathcal{P}$,*

$$\sup_{P \in \mathcal{P}^* \cap \mathcal{P}^{SIMPLE}} V(\mathcal{O}(\mathcal{G}, P)) \leq \sup_{P \in \mathcal{P}^*} V(\mathcal{O}(\mathcal{G}, P)).$$

Moreover, if $\mathcal{G}$ satisfies strong richness,

$$\sup_{P \in \mathcal{P}^* \cap \mathcal{P}^{SIMPLE}} V(\mathcal{O}(\mathcal{G}, P)) = \sup_{P \in \mathcal{P}^*} V(\mathcal{O}(\mathcal{G}, P)).$$

The first statement follows from our denseness result (6) while the second statement follows from our exact continuity results (Propositions 5and 18)

7.11 Example Existence of BIBCE

We consider the game by Hellman (2014) where BNE fails to exist and yet construct a straightforward BIBCE of this game. There are two players A and B and two actions L, R (see Hellman (2014) for the payoff matrix)

$$\sigma(y, \theta) = \begin{cases} \frac{1}{2}(L, L), \frac{1}{2}(R, R) & \text{, if } \theta \in \{(A, 1), (B, 1)\} \\ \frac{1}{2}(R, L), \frac{1}{2}(L, R) & \text{, if } \theta \in \{(A, -1), (B, -1)\} \end{cases}$$

This decision rule is independent of the type space and satisfies belief invariance: For any $a_i \in \{L, R\}$ and any $\theta \in \{(A,1),(B,1),(A,-1),(B,-1)\}$

$$\sum_{a_{-i} \in \{L,R\}} \sigma(a_i, a_{-i}|y,\theta) = \sigma(a_i, L|y,\theta) + \sigma(a_i, R|y,\theta) = \frac{1}{2}.$$

We do not need to report the full information structure considered in Hellman (2014), to verify obedience as σ is independent of types. Each player $i \in \{A, B\}$ has one of two first-order beliefs $P_i(\cdot|y)$ indexed by $y \in \{-1, 1\}$

$$P_A(\theta|y) = \begin{cases} \frac{1}{2} & \text{if } \theta = (A, y) \\ \frac{1}{4} & \text{if } \theta = (B, 1) \\ \frac{1}{4} & \text{if } \theta = (B, -1) \end{cases}, \quad P_B(\theta|y) = \begin{cases} \frac{1}{2} & \text{if } \theta = (B, y) \\ \frac{1}{4} & \text{if } \theta = (A, 1) \\ \frac{1}{4} & \text{if } \theta = (A, -1) \end{cases}.$$

So if recommended $a_i \in \{L, R\}$ player i's payoff increment from playing a different action $a_i' \neq a_i$ is given by when $i = A:$

For the case where $y = 1$ and $a_A = L$

$$\frac{1}{2}\left(u_A(L, L, (A, 1)) - u_A(R, L, (A, 1))\right) + \frac{1}{4}\left(u_A(L, L, (B, 1)) - u_A(R, L, (B, 1))\right) \\ + \frac{1}{4}\left(u_A(L, R, (B, -1)) - u_A(R, R, (B, -1))\right) = 0.35 > 0.$$

For the case where $y = 1$ and $a_A = R$

$$\frac{1}{2}\left(u_A(R, R, (A, 1)) - u_A(L, R, (A, 1))\right) + \frac{1}{4}\left(u_A(R, R, (B, 1)) - u_A(L, R, (B, 1))\right) \\ + \frac{1}{4}\left(u_A(R, L, (B, -1)) - u_A(L, L, (B, -1))\right) = 0.15 > 0.$$

For the case $y = -1$ and $a_A = L$,

$$\frac{1}{2}\left(u_A(L,R,(A,-1)) - u_A(R,R,(A,-1))\right) + \frac{1}{4}\left(u_A(L,R,(B,-1)) - u_A(R,R,(B,-1))\right) \\ + \frac{1}{4}\left(u_A(L,L,(B,1)) - u_A(R,L,(B,1))\right) = 0.35 > 0.$$

For the case $y = -1$ and $a_A = R$,

$$\frac{1}{2}\left(u_A(R,L,(A,-1)) - u_A(L,L,(A,-1))\right) + \frac{1}{4}\left(u_A(R,L,(B,-1)) - u_A(L,L,(B,-1))\right) \\ + \frac{1}{4}\left(u_A(R,R,(B,1)) - u_A(L,R,(B,1))\right) = 0.15 > 0.$$

Similarly for player $i = B$:

For the case where $y = 1$ and $a_B = L$

$$\frac{1}{2}\left(u_B(L,L,(B,1)) - u_B(L,R,(B,1))\right) + \frac{1}{4}\left(u_B(L,L,(A,1)) - u_B(L,R,(A,1))\right) \\ + \frac{1}{4}\left(u_B(R,L,(A,-1)) - u_B(R,R,(A,-1))\right) = 0.35 > 0.$$

For the case where $y = 1$ and $a_B = R$

$$\frac{1}{2}\left(u_B(R,R,(B,1)) - u_B(R,L,(B,1))\right) + \frac{1}{4}\left(u_B(R,R,(A,1)) - u_B(R,L,(A,1))\right) \\ + \frac{1}{4}\left(u_B(L,R,(A,-1)) - u_B(L,L,(A,-1))\right) = 0.15 > 0.$$

For the case $y = -1$ and $a_B = L$,

$$\frac{1}{2}\left(u_B(R,L,(B,-1)) - u_B(R,R,(B,-1))\right) + \frac{1}{4}\left(u_B(R,L,(A,-1)) - u_B(R,R,(A,-1))\right) \\ + \frac{1}{4}\left(u_B(L,L,(A,1)) - u_B(L,R,(A,1))\right) = 0.35 > 0.$$

For the case $y = -1$ and $a_B = R$,

$$\frac{1}{2}\left(u_B(L, R, (B, -1)) - u_B(L, L, (B, -1))\right) + \frac{1}{4}\left(u_B(L, R, (A, -1)) - u_B(L, L, (A, -1))\right) + \frac{1}{4}\left(u_B(R, R, (B, 1)) - u_A(L, R, (B, 1))\right) = 0.15 > 0,$$

and thus all the obedience conditions for a BIBCE are satisfied.

References

Balder, E. (1988): "Generalized Equilibrium Results for Games with Incomplete Information," *Mathematics of Operations Research*, 1988, 265–276.

Bergemann, D. and S. Morris (2016): "Bayes Correlated Equilibrium and the Comparison of Information Structures in Games," *Theoretical Economics*, 11, 487–522.

——— (2017): "Belief-Free Rationalizability and Informational Robustness," *Games and Economic Behavior*, 104, 744–759.

Brandenburger, A. and A. Friedenberg (2008): "Intrinsic Correlation in Games," *Journal of Economic Theory*, 141, 28–67.

Carlsson, H. and E. Van Damme (1993): "Global games and equilibrium selection," *Econometrica: Journal of the Econometric Society*, 989–1018.

Chen, Y.-C., A. Di Tillio, E. Faingold, and S. Xiong (2017): "Characterizing the Strategic Impact of Misspecified Beliefs," *Review of Economic Studies*, 84, 1424–1471.

Dekel, E., D. Fudenberg, and S. Morris (2006): "Topologies on Types," *Theoretical Economics*, 1, 275–309.

——— (2007): "Interim Correlated Rationalizability," *Theoretical Economics*.

Du, S. (2012): "Correlated Equilibrium and Higher-Order Beliefs about Play," *Games and Economic Behavior*, 76, 74–87.

Ely, J. and M. Peski (2011): "Critical Types," *Review of Economic Studies*, 78, 907–937.

Engl, G. (1995): "Lower Hemicontinuity of the Nash Equilibrium Correspondence," *Games and Economic Behavior*, 9, 151–160.

Forges, F. (1993): "Five Legitimate Definitions of Correlated Equilibrium in Games with Incomplete Information," *Theory and Decision*, 61, 329–344.

——— (2006): "Correlated Equilibrium in Games with Incomplete Information Re-Visited," *Theory and Decision*, 61, 329–324.

Gensbittel, F., M. Peski, and J. Renault (2022): "Value-Based Distance between Information Structures," *Theoretical Economics*.

Gossner, O. (2000): "Comparison of Information Structures," *Games and Economic Behavior*, 30, 44–63.

Gossner, O. and J.-F. Mertens (2020): "The Value of Information in Zero-Sum Games," Tech. rep., LSE.

Harsanyi, J. (1967-68): "Games with Incomplete Information Played by 'Bayesian' Players," *Management Science*, 14, 159–189, 320–334, 485–502.

Hellman, Z. (2014): "A Game with no Bayesian Approximate Equilibria," *Journal of Economic Theory*, 153, 138–151.

Inostroza, N. and A. Pavan (2023): "Adversarial Coordination and Public Information Design," Tech. rep., University of Toronto and Northwestern University.

Kajii, A. and S. Morris (1998): "Payoff Continuity in Incomplete Information Games," *Journal of Economic Theory*, 82, 267–276.

Kambhampati, A. (2023): "Payoff Continuity in Games of Incomplete Information Across Models of Knowledge," Tech. rep., United States Naval Academy.

Lehrer, E., D. Rosenberg, and E. Shmaya (2010): "Signaling and Mediation in Games with Common Interests," *Games and Economic Behavior*, 68, 670–682.

Li, F., Y. Song, and M. Zhao (2023): "Global Manipulation by Local Obfuscation," *Journal of Economic Theory*.

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

Liu, Q. (2009): "On Redundant Types and the Bayesian Formulation of Incomplete Information," *Journal of Economic Theory*, 144, 2115–2145.

——— (2015): "Correlation and Common Priors in Games with Incomplete Information," *Journal of Economic Theory*, 157, 49–75.

Mathevet, L., J. Perego, and I. Taneva (2020): "On Information Design in Games," *Journal of Political Economy*, 128, 1370–1404.

Mertens, J.-F., S. Sorin, and S. Zamir (2015): *Repeated games*, vol. 55, Cambridge University Press.

Mertens, J.-F. and S. Zamir (1985): "Formulation of Bayesian analysis for games with incomplete information," *International journal of game theory*, 14, 1–29.

Milgrom, P. and R. Weber (1985): "Distributional Strategies for Games with Incomplete Information," *Mathematics of Operations Research*, 10, 619–632.

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 Incomplete Information," *Mathematics of Operations Research*, 21, 707–725.

Morris, S., D. Oyama, and S. Takahashi (2024): "Implementation via Information Design in Binary-Action Supermodular Games," *Econometrica*.

Peski, M. (2008): "Comparison of Information Structures in Zero-Sum Games," *Games and Economic Behavior*, 62, 732–735.

Rubinstein, A. (1989): “The Electronic Mail Game: Strategic Behavior under ’Almost Common Knowledge’,” *American Economic Review.*

Sadzik, T. (2019): “Beliefs Revealed in Bayesian-Nash Equilibrium,” Tech. rep., UCLA.

Simon, R. S. (2003): “Games of Incomplete Information, Ergodic Theory, and the Measurability of Equilibria,” *Israel Journal of Mathematics*, 138, 73–92.

Simon, R. S. and G. Tomkowicz (2017): “A Bayesian Game without e-Equilibria,” Tech. rep., LSE.

Stinchcombe, M. B. (2011): “Correlated Equilibrium Existence for Infinite Games with Type-Dependent Strategies,” *Journal of Economic Theory*, 146, 638–655.

van Zandt, T. (2010): “Interim Bayesian Nash Equilibrium on Universal Type Spaces for Supermodular Games,” *Journal of Economic Theory*, 145, 249–263.

Weinstein, J. and M. Yildiz (2007): “A Structure Theorem for Rationalizability with Applications to Robust Predictions of Refinements,” *Econometrica*, 75, 365–400.