International Journal of Game Theory (2021) 50:311–360
https://doi.org/10.1007/s00182-020-00744-y

ORIGINAL PAPER

Equilibria in infinite games of incomplete information

Oriol Carbonell-Nicolau[1]

Accepted: 11 October 2020 / Published online: 29 April 2021

Abstract
The notion of *communication equilibrium* extends Aumann's (J Math Econ 1:67–96, 1974, https://doi.org/10.1016/0304-4068(74)90037-8) correlated equilibrium concept for complete information games to the case of incomplete information. This paper shows that this solution concept has the following property: for the class of incomplete information games with compact metric type and action spaces, and with payoff functions jointly measurable and continuous in actions, limits of Bayes-Nash equilibria of finite approximations to an infinite game are communication equilibria (and, in general, *not* Bayes-Nash equilibria) of the limit game. Stinchcombe's (J Econ Theory 146:638–655, 2011b, https://doi.org/10.1016/j.jet.2010.12.006) extension of Aumann's (J Math Econ 1:67–96, 1974, https://doi.org/10.1016/0304-4068(74)90037-8) solution concept to the case of incomplete information fails to satisfy this condition.

Keywords Infinite games of incomplete information · Bayes-Nash equilibrium · Communication equilibrium · Correlated equilibrium · Strategic approximation of an infinite game

JEL Classification C72

1 Introduction

The aim of this paper is to understand which solution concepts for incomplete information games with infinitely many actions and types (henceforth infinite games) are generally "good" predictors of Bayes-Nash equilibrium behavior in "nearby" games with finitely many strategies. It is shown that, from this perspective, the notion of *communication equilibrium*, which extends Aumann's (1974) *correlated equilibrium*

I would like to thank Philip Reny for very useful discussions, and Maxwell Stinchcombe, Richard McLean, and two anonymous referees for their valuable comments.

✉ Oriol Carbonell-Nicolau
carbonell-nicolau@rutgers.edu

1 Department of Economics, Rutgers University, New Brunswick, USA

concept for complete information games to the case of incomplete information, is generally more appropriate than the Bayes-Nash solution concept or the notion of correlated equilibrium formulated in Stinchcombe (2011b).

A communication equilibrium is a particular type of correlated strategy (i.e., a mixture over action profiles for every type profile), interpreted as a mixture over action profiles recommended by a mediator for each reported type profile. A player can be dishonest, misreporting her type, and, in addition, a player can be disobedient, playing some mixture over the player's actions (conditional on the player's type) instead of the action recommended by the mediator. A communication equilibrium is a correlated strategy that is immune to misreporting and disobedience.

The notion of communication equilibrium employed here extends that used in Myerson (1991, Sect. 6.3, p. 258) for finite games (see also Myerson (1982) and Forges (1986, 1990, 1993)) and differs from Stinchcombe's (2011b) correlated equilibrium, in the sense that there are correlated equilibria that fail to be communication equilibria, and vice versa.

Roughly speaking, a strategic approximation of an infinite game of incomplete information is defined as a countable set of behavioral strategy profiles with the following property: given any sequence of games whose finite sets of behavioral strategy profiles eventually include every member of the countable set, limits of Bayes-Nash equilibria of the finite games are "equilibria" of the infinite game. This definition is based on a notion introduced by Reny (2011b) for normal-form games.[1]

Of course, the definition of a strategic approximation must specify what it means for a sequence of behavioral strategy profiles to converge to a point. This paper identifies a topology on the space of correlated strategies that guarantees the existence of strategic approximations, and argues that coarser topologies are too weak to warrant the existence of a strategic approximation.

If one requires that limits of Bayes-Nash equilibria of approximating games be Bayes-Nash equilibria of the infinite game, then strategic approximations do not generally exist. Indeed, in this case, one can find simple games for which there are convergent sequences of Bayes-Nash equilibria whose limit points are not Bayes-Nash equilibria (see Sect. 4). A similar problem arises if one requires that limits of Bayes-Nash equilibria of approximating games be correlated equilibria of the limit game, according to the notion of correlated equilibrium defined in Stinchcombe (2011b) (see Sect. 4).

There are two ways around this problem. The first is to use finer topologies for the notion of convergence in the definition of a strategic approximation. The second is to modify the equilibrium concept for the limit game. We pursue the second idea, using the communication equilibrium concept, which allows us to prove the existence of a strategic approximation for a wider class of topologies.

This paper confines attention to the class $\mathfrak{G}$ of all the incomplete information games with compact, metric type and action spaces and with payoff functions jointly measurable and continuous in actions. The main result identifies a topology for which all the members of $\mathfrak{G}$ admit a strategic approximation. This topology can be argued to

[1] Alternative notions of an approximation to an infinite game via a sequence of finite games have been considered by Stinchcombe (2005) and Stinchcombe (2005, 2011a), who shows that discretizing action and type spaces, rather than spaces of behavioral strategies, is a rather delicate matter.

be the "weakest" possible topology, in the sense that, for weaker topologies, there are games in $\mathfrak{G}$ that do not admit a strategic approximation.

Strategic approximations lead naturally to the notion of "robust" communication equilibrium profiles (*i.e.*, robust to the finite perturbations considered in this paper), and a corollary of our main result identifies sufficient conditions for existence of this refinement.[2]

2 Preliminaries

Throughout the paper, the following definitions will be adopted. If Y is a metric space, then $\mathscr{B}(Y)$ will denote the σ-algebra of the Borel subsets of Y, $\Delta(Y)$ will represent the set of probability measures on $(Y, \mathscr{B}(Y))$, and $C^b(Y)$ will denote the set of all bounded and continuous real-valued functions on Y.

Definition 1 The w-***topology*** on $\Delta(Y)$ is defined as the coarsest topology for which all the functionals in

$$\left\{\mu \in \Delta(Y) \mapsto \int_Y f(y)\mu(dy) \in \mathbb{R} : f \in C^b(Y)\right\}$$

are continuous.

We shall refer to the notion of convergence of measures in $\Delta(Y)$ with respect to the w-topology as ***weak convergence*** of measures and we shall write $\mu^\alpha \underset{w}{\rightarrow} \mu$ to indicate that the net of measures (μ^α) converges weakly to μ.

If Y is a complete, separable metric space, the w-topology on $\Delta(Y)$ is metrizable, and the Prokhorov metric defines a compatible metric (see Prokhorov 1956, Theorem 1.11). The Prokhorov metric on $\Delta(Y)$ is defined by the map $\varrho_{\Delta(Y)} : \Delta(Y)\times\Delta(Y) \rightarrow \mathbb{R}$ given by

$$\varrho_{\Delta(Y)}(\mu, \nu) := \inf\left\{\epsilon : \forall B \in \mathscr{B}(Y), \mu(B) \leq \nu(N_\epsilon(B)) + \epsilon\right\}, \tag{1}$$

where $N_\epsilon(B)$ denotes the ϵ-neighborhood of B, *i.e.*, $N_\epsilon(B) := \bigcup_{b\in B} N_\epsilon(b)$, and $N_\epsilon(b)$ denotes the ϵ-neighborhood of b in Y. An equivalent formulation (see *e.g.*, Dudley 1968, p. 1564) is

$$\varrho_{\Delta(Y)}(\mu, \nu) := \inf\left\{\epsilon : \forall \text{closed } B \subseteq Y, \mu(B) \leq \nu(N_\epsilon(B)) + \epsilon\right\}. \tag{2}$$

[2] Stinchcombe (2011b) and Cotter (1991) establish existence of correlated equilibrium within the class $\mathfrak{G}$. Other authors (see *e.g.*, Milgrom and Weber 1985; Balder 1988; Carbonell-Nicolau and McLean 2018, 2019; He and Yannelis 2016) have proven existence of Bayes-Nash equilibria (and hence communication equilibria) under the additional assumption of diffuse joint information of the players. See Simon (2003) for a proof of the fact that equilibria need not exist if one drops the diffuseness assumption. In related frameworks, such as the state-space framework of Yannelis and Rustichini (1991), Hellman and Levy (2017), and Carbonell-Nicolau and McLean (2020), and the lattice framework of Athey (2001), McAdams (2003) and Reny (2011a), existence results can be proven in which the requirement of diffuse information is replaced by assumptions we do not make here.

2.1 Games and strategies

Definition 2 A ***normal-form game*** (or simply a ***game***) is a collection $G = (Z_i, f_i)_{i=1}^N$, where N is a finite number of players, Z_i is a nonempty set of actions for player i, and $f_i : Z \to \mathbb{R}$ represents player i's payoff function, defined on the set of action profiles $Z := \times_{i=1}^N Z_i$.

Throughout the sequel, given N sets $Z_1, \dots, Z_N$, we adhere to the following conventions, which are standard in the literature, even though they sometimes entail abuses of notation: for $i \in \{1, \dots, N\}$, $Z_{-i} := \times_{j \neq i} Z_j$; given i, the set $\times_{j=1}^N Z_j$ is sometimes represented as $Z_i \times Z_{-i}$, and $z = (z_i, z_{-i}) \in Z_i \times Z_{-i}$ is used for a member z of $\times_{j=1}^N Z_j$.

Definition 3 A ***Bayesian game*** is a collection

$$\Gamma = (T_i, X_i, u_i, p)_{i=1}^N ,$$

where

- $\{1, \dots, N\}$ is a finite set of players;
- T_i is a nonempty, compact, metric space of types for player i;
- X_i is a nonempty, compact, metric space of actions for player i;
- u_i is a real-valued map on $T \times X$, where $T := \times_{i=1}^N T_i$ and $X := \times_{i=1}^N X_i$; it represents player i's payoff function, and it is assumed bounded and $(\mathscr{B}(T \times X), \mathscr{B}(\mathbb{R}))$-measurable; and
- p is a probability measure on $(T, \mathscr{B}(T))$, describing the players' common priors over type profiles.

This paper is concerned with Bayesian games $\Gamma = (T_i, X_i, u_i, p)_{i=1}^N$ such that $u_i(t, \cdot) : X \to \mathbb{R}$ is continuous for each $t \in T$ and i. The set of all such Bayesian games will be denoted by $\mathfrak{G}$.

Definition 4 Suppose that $\Gamma = (T_i, X_i, u_i, p)_{i=1}^N$ is a Bayesian game. A ***behavioral strategy*** for player i in Γ is a $(\mathscr{B}(T_i), \mathscr{B}(\Delta(X_i)))$-measurable map $\mu_i : T_i \to \Delta(X_i)$.

Given a Bayesian game $\Gamma = (T_i, X_i, u_i, p)_{i=1}^N$, the set of behavioral strategies for player i in Γ is denoted by $\mathscr{T}_i$, and we define $\mathscr{T} := \times_{i=1}^N \mathscr{T}_i$; the dependence of $\mathscr{T}_i$ and $\mathscr{T}$ on Γ is not made explicit and will (hopefully) be clear from the context.

A behavioral strategy $\mu_i \in \mathscr{T}_i$ describes the mixture $\mu_i(\cdot|t_i) \in \Delta(X_i)$ over the actions in X_i employed by the type t_i of player i.

Given a Bayesian game $\Gamma = (T_i, X_i, u_i, p)_{i=1}^N$, define the normal-form game

$$\mathfrak{G}_\Gamma := (\mathscr{T}_i, U_i)_{i=1}^N , \tag{3}$$

where $U_i : \mathscr{T} \to \mathbb{R}$ is defined by

$$U_i(\mu_1, \dots, \mu_N) := \int_T \int_{X_N} \cdots \int_{X_1} u_i(t, x) \mu_1(dx_1|t_1) \cdots \mu_N(dx_N|t_N) p(dt).$$

Definition 5 Suppose that $\Gamma = (T_i, X_i, u_i, p)_{i=1}^N$ is a Bayesian game. A ***correlated strategy*** in Γ is a $(\mathscr{B}(T), \mathscr{B}(\Delta(X)))$-measurable map $\mu : T \to \Delta(X)$.

Given a Bayesian game $\Gamma = (T_i, X_i, u_i, p)_{i=1}^N$, the set of correlated strategies in Γ is denoted by $\mathscr{M}$ (here again the dependence of $\mathscr{M}$ on Γ is not explicitly indicated).

A correlated strategy $\mu \in \mathscr{M}$ specifies a mixture $\mu(t) \in \Delta(X)$ over action profiles in X conditional on every type profile t in T.

A strategy profile $(\mu_1, \ldots, \mu_N) \in \mathscr{T}$ induces a correlated strategy μ in a natural way. Indeed, given a strategy profile $(\mu_1, \ldots, \mu_N) \in \mathscr{T}$, the map $\mu : T \to \Delta(X)$ defined by

$$\mu(t) := \bigotimes_{i=1}^{N} \mu_i(t_i)$$

is a correlated strategy in $\mathscr{M}$.

2.2 Equilibrium

Definition 6 Suppose that $G = (Z_i, f_i)_{i=1}^N$ is a normal-form game. A strategy profile $z = (z_i, z_{-i})$ in $\times_{i=1}^N Z_i$ is a ***Nash equilibrium*** of G if $f_i(y_i, z_{-i}) \leq f_i(z)$ for every $y_i \in Z_i$ and i.

Definition 7 A ***Bayes-Nash equilibrium*** of a Bayesian game $\Gamma = (T_i, X_i, u_i, p)_{i=1}^N$ is a Nash equilibrium of the game $\mathfrak{G}_\Gamma$ defined in (3), *i.e.*, a profile $(\mu_1, \ldots, \mu_N) \in \mathscr{T}$ such that for each i,

$$U_i(\mu_i, \mu_{-i}) \geq U_i(\nu_i, \mu_{-i}), \quad \text{for all } \nu_i \in \mathscr{T}_i.$$

Suppose that $\Gamma = (T_i, X_i, u_i, p)_{i=1}^N$ is a Bayesian game. For each i, let $\mathscr{A}_i$ be the set of all $(\mathscr{B}(T_i \times X_i), \mathscr{B}(\Delta(X_i)))$-measurable maps $\alpha_i : T_i \times X_i \to \Delta(X_i)$, and let $\mathscr{D}_i$ be the set of all $(\mathscr{B}(T_i), \mathscr{B}(\Delta(T_i)))$-measurable maps $\eta_i : T_i \to \Delta(T_i)$.

Definition 8 Suppose that $\Gamma = (T_i, X_i, u_i, p)_{i=1}^N$ is a Bayesian game. A correlated strategy $\mu \in \mathscr{M}$ is a ***communication equilibrium*** of Γ if for each i and $(\alpha_i, \eta_i) \in \mathscr{A}_i \times \mathscr{D}_i$,

$$\begin{aligned}
&\int_T \int_{T_i} \int_X \int_{X_i} u_i(t, y_i, x_{-i}) \alpha_i(dy_i | t_i, x_i) \mu(dx | \tau_i, t_{-i}) \eta_i(d\tau_i | t_i) p(dt) \\
&\quad \leq \int_T \int_X u_i(t, x) \mu(dx | t) p(dt).
\end{aligned}$$

A correlated strategy $\mu \in \mathscr{M}$ can be viewed as a mixture $\mu(t) \in \Delta(X)$ recommended by a mediator for each given reported type profile $t \in T$. A player i can be dishonest, misreporting her type according to η_i (which specifies a mixture over T_i, $\eta_i(t_i)$, for each type $t_i \in T_i$), and, in addition, a player can be disobedient, playing the mixture $\alpha_i(t_i, x_i) \in \Delta(X_i)$, when her type is t_i, instead of the action x_i recommended by the mediator. A communication equilibrium is a correlated strategy that is immune to misreporting and disobedience.

Definition 8 extends Aumann's (1974) notion of correlated equilibrium to games of incomplete information. In the special case of Bayesian games with finitely many types and actions, Definition 8 coincides with the equilibrium concept defined in Myerson (1991, Sect. 6.3, p. 258).

The next definition requires some terminology.

Let $([0,1], \mathscr{B}([0,1]), \lambda)$ be the measure space of the unit interval with the σ-algebra of the Borel subsets of $[0,1]$ and the normalization of the Lebesgue measure over $[0,1]$.

Suppose that $\Gamma = (T_i, X_i, u_i, p)_{i=1}^N$ is a Bayesian game. For each i, let $\mathscr{X}_i$ be the set of all $(\mathscr{B}(T_i \times [0,1]), \mathscr{B}(\Delta(X_i)))$-measurable maps $\varphi_i : T_i \times [0,1] \to \Delta(X_i)$.

The following definition is introduced in Stinchcombe (2011b).

Definition 9 Suppose that $\Gamma = (T_i, X_i, u_i, p)_{i=1}^N$ is a Bayesian game. A profile $\varphi = (\varphi_1, \ldots, \varphi_N) \in \times_{i=1}^N \mathscr{X}_i$ is a ***correlated equilibrium*** of Γ if

$$\int_{T\times[0,1]} \int_X u_i(t,x) \left[\bigotimes_{j=1}^N \varphi_j(t_j,a) \right](dx)[p\otimes\lambda](d(t,a)) \geq \int_{T\times[0,1]} \int_X u_i(t,x) \left[\psi_i\,(t_i, \varphi_i(t_i,a)) \otimes \left[\bigotimes_{j\neq i} \varphi_j(t_j,a) \right] \right](dx)[p\otimes\lambda](d(t,a))$$

for each i and each $(\mathscr{B}(T_i \times \Delta(X_i)), \mathscr{B}(\Delta(X_i)))$-measurable map $\psi_i : T_i \times \Delta(X_i) \to \Delta(X_i)$.

Definition 9 is also an extension of Aumann's (1974) notion of correlated equilibrium to games of incomplete information.[3]

A correlated equilibrium (Definition 9), viewed as a correlated strategy (Definition 5), need not be a communication equilibrium. Conversely, communication equilibria need not exhibit the specific kind of correlation required in Definition 9, as illustrated in Sect. 4.

To see that a correlated equilibrium need not be a communication equilibrium, note first that a profile $\varphi = (\varphi_1, \ldots, \varphi_N) \in \times_{i=1}^N \mathscr{X}_i$ induces a correlated strategy (Definition 5) $\mu : T \to \Delta(X)$ defined as follows:

$$\mu(B|t) := \int_{[0,1]} \left[\bigotimes_{i=1}^N \varphi_i(t_i,a) \right](B)\lambda(da). \tag{4}$$

We claim that a correlated equilibrium φ of a Bayesian game $\Gamma = (T_i, X_i, u_i, p)_{i=1}^N$ need not induce (via (4)) a communication equilibrium of Γ. To see this, consider the following game $\Gamma = (T_i, X_i, u_i, p)_{i=1}^N$, with $N = 2$, payoff-irrelevant type spaces $T_1 = T_2 := \{0, 1\}$, action spaces $X_1 = X_2 := \{A, B\}$, and payoff bi-matrix

[3] There are alternative ways of defining the notion of correlated equilibrium (see, *e.g.*, Bergemann and Morris 2016), which are not considered here.

	A	B
A	2,1	1,1
B	1,1	2,1

Assume that $p = p_1 \otimes p_2$, where each p_i assigns $\frac{1}{2}$ probability to each type. Now define the profile $\varphi = (\varphi_1, \varphi_2) \in \mathscr{X}_1 \times \mathscr{X}_2$ as follows:

$$\varphi_1(t_1, a) := \begin{cases} \delta_A & \text{if } t_1 = 1, \\ \delta_A & \text{if } t_1 = 0 \text{ and } a \in \left[0, \frac{1}{2}\right], \\ \delta_B & \text{if } t_1 = 0 \text{ and } a \in \left(\frac{1}{2}, 1\right], \end{cases} \tag{5}$$

and

$$\varphi_2(t_2, a) := \begin{cases} \delta_A & \text{if } a \in \left[0, \frac{1}{2}\right], \\ \delta_B & \text{if } a \in \left(\frac{1}{2}, 1\right], \end{cases} \tag{6}$$

where δ_A (resp. δ_B) denotes the Dirac measure on $\{A, B\}$ with support $\{A\}$ (resp. $\{B\}$).

The profile φ is a correlated equilibrium of Γ. Indeed, it is clear that player 2 cannot profitably deviate, and, in addition, for each $\psi_1 : T_1 \times \Delta(X_1) \to \Delta(X_1)$, we have

$$\begin{aligned} &\int_{T\times[0,1]} \int_X u_1(t, x)\, [\varphi_1(t_1, a) \otimes \varphi_2(t_2, a)]\,(dx)[p \otimes \lambda](d(t, a)) \\ &\qquad = \frac{7}{4} \geq \int_{T\times[0,1]} \int_X u_1(t, x) \\ &\qquad [\psi_1\,(t_1, \varphi_1(t_1, a)) \otimes \varphi_2(t_2, a)]\,(dx)[p \otimes \lambda](d(t, a)). \end{aligned}$$

To see that the last inequality holds, note that the only "events" (t, a) for which the strategy φ_1 does not attain the maximum payoff for player 1 (*i.e.*, 2), given that player 2's strategy is φ_2, are those in the set $\left\{(t, a) : t_1 = 1 \text{ and } a \in \left(\frac{1}{2}, 1\right]\right\}$. Since $\varphi_1|_{\{(t_1,a):t_1=1\}} = \delta_A$, player 1 can only improve her payoff via a deviation of the form $\psi_1\,(t_1, \varphi_1(t_1, a))$ if $\psi_1(1, \delta_A)$ assigns positive probability to the action B, *i.e.*, if $\psi_1(B|1, \delta_A) > 0$. But, for any such ψ_1,

$$\begin{aligned} &\int_{T\times[0,1]} \int_X u_1(t, x)\, [\psi_1\,(t_1, \varphi_1(t_1, a)) \otimes \varphi_2(t_2, a)]\,(dx)[p \otimes \lambda](d(t, a)) \\ &\quad = \frac{1}{4}(2\psi_1(A|1, \delta_A) + \psi_1(B|1, \delta_A)) + \frac{1}{4}2 \\ &\qquad + \frac{1}{4}(\psi_1(A|1, \delta_A) + 2\psi_1(B|1, \delta_A)) + \frac{1}{4}2 \\ &\quad = \frac{1}{4}(2\psi_1(A|1, \delta_A) + 1 - \psi_1(A|1, \delta_A)) + \frac{1}{4}2 \\ &\qquad + \frac{1}{4}(\psi_1(A|1, \delta_A) + 2[1 - \psi_1(A|1, \delta_A)]) + \frac{1}{4}2 \end{aligned}$$

$$= \frac{1}{4}(1 + \psi_1(A|1, \delta_A)) + \frac{1}{4}2 + \frac{1}{4}(2 - \psi_1(A|1, \delta_A)) + \frac{1}{4}2$$
$$= \frac{7}{4}.$$

While φ is a correlated equilibrium of Γ, it is *not* a communication equilibrium of Γ. Indeed, given φ (as defined by (5) and (6)), the corresponding correlated strategy μ defined via (4) satisfies

$$\int_T \int_X u_1(t, x)\mu(dx|t)p(dt) = \frac{7}{4},$$

and the misreporting strategy $\eta_1 \in \mathscr{D}_1$ defined by $\eta_1(t_1) := \delta_0$ for all $t_1 \in T_1$ (where δ_0 denotes the Dirac measure in $\Delta(T_1)$ with support $\{0\}$) yields

$$\int_T \int_{T_1} \int_X u_1(t, x)\mu(dx|\tau_1, t_2)\eta_1(d\tau_1|t_1)p(dt) = 2 > \frac{7}{4} = \int_T \int_X u_1(t, x)\mu(dx|t)p(dt).$$

Note that the scope for profitable deviations is less restrictive for the notion of communication equilibrium *vis-à-vis* the correlated equilibrium concept.

2.3 Strategic approximations

The archetypal approach to the analysis of robustness of equilibrium points in infinite games of complete information is based on the classic closed graph theorem for the Nash equilibrium correspondence when the payoff functions are the parameters. This classic result and its subsequent generalizations rely on continuity of the payoff functions.[4] Similar approximation results based on continuity of the *expected* payoff functions have been developed for games of incomplete information by Milgrom and Weber (1985, Theorem 2).[5] In the presence of payoff discontinuities, "good" approximations to an infinite game must eventually include strategies that are of particular strategic significance to the players. This issue, which is pointed out in Simon (1987) and Reny (2011b), does not arise in the context of continuous games. In fact, when payoff functions are smooth, any strategy can be reasonably approximated by an arbitrary, nearby strategy. Thus, the notion of a "well-defined" approximating sequence of games is necessarily more nuanced when the limit game exhibits payoff discontinuities. These considerations motivate Reny's (2011b) concept of a finite approximation to an infinite normal-form game of complete information (Definition 2).

Definition 10 (Reny 2011b) Suppose that $G = (Z_i, f_i)_{i=1}^N$ is a normal-form game, and let $Z := \times_{i=1}^N Z_i$ be a metric space. A ***strategic approximation*** of G is a countable set of strategies $Z^\infty = \times_{i=1}^N Z_i^\infty$ contained in Z satisfying the following: if for each player i, $Z_i^1 \subseteq Z_i^2 \subseteq \cdots$ is an increasing sequence of finite subsets of Z_i whose union contains

[4] See, *e.g.*, Lucchetti and Patrone (1986), Stinchcombe (2005), and Gürkan and Pang (2007).

[5] See also the recent extensions in Prokopovych and Yannelis (2019) and He and Sun (2019).

Z_i^∞, and if for each n, z^n is a Nash equilibrium of the game $(Z_j^n, f_j|_{Z_1^n\times\cdots\times Z_N^n})_{j=1}^N$, then any limit point of the sequence (z^n) is a Nash equilibrium of G.

In our setting, the *expected* payoff functions exhibit marked discontinuities (see Stinchcombe (2011a, b)). Consequently, we adopt Reny's (2011b) approach.[6]

In light of Definition 10, the reader may be tempted to define a strategic approximation of a Bayesian game as a strategic approximation of the normal-form game defined in (3). That is, given a Bayesian game $\Gamma = (T_i, X_i, u_i, p)_{i=1}^N$, a strategic approximation of Γ could be defined as a countable set of strategies $\mathscr{T}^\infty = \times_{i=1}^N \mathscr{T}_i^\infty$ contained in $\mathscr{T} = \times_{i=1}^N \mathscr{T}_i$ satisfying the following: if for each player i, $\mathscr{T}_i^1 \subseteq \mathscr{T}_i^2 \subseteq \cdots$ is an increasing sequence of finite subsets of $\mathscr{T}_i$ whose union contains $\mathscr{T}_i^\infty$, if for each n, μ^n is a Nash equilibrium of the game $(\mathscr{T}_j^n, U_j|_{\mathscr{T}_1^n\times\cdots\times\mathscr{T}_N^n})_{j=1}^N$, and if the sequence (μ^n) "converges" to a point μ, then μ is a Bayes-Nash equilibrium of Γ. Of course, this definition is not precise enough, for it does not specify the notion of convergence for the sequence (μ^n). This paper introduces a topology for the space of correlated strategies (which, as explained at the end of Sect. 2.1, contains the space of behavioral strategy profiles for the Bayesian game Γ). For this topology, defining a strategic approximation of a Bayesian game via Definition 10 (*i.e.*, applying Definition 10 directly to the normal-form game defined in (3)) is problematic. Indeed, as illustrated in Sect. 4 below, limits of Bayes-Nash equilibria of sequences of finite approximating games need not be Bayes-Nash equilibria of the limit game. Thus, an alternative definition is needed in which the solution concept for the limit game is weakened. This paper proposes the following definition.

Definition 11 Suppose that $\Gamma = (T_i, X_i, u_i, p)_{i=1}^N$ is a Bayesian game, and let $\mathfrak{G}_\Gamma = (\mathscr{T}_i, U_i)_{i=1}^N$ be its corresponding normal form as defined in (3). A ***strategic approximation*** of Γ is a countable set of strategies $\mathscr{T}^\infty = \times_{i=1}^N \mathscr{T}_i^\infty$ contained in $\mathscr{T} = \times_{i=1}^N \mathscr{T}_i$ satisfying the following: if for each player i, $(\mathscr{T}_i^\alpha)$ is an increasing net of finite subsets of $\mathscr{T}_i$ whose union contains $\mathscr{T}_i^\infty$, if for each α, μ^α is a Nash equilibrium of the game $(\mathscr{T}_j^\alpha, U_j|_{\mathscr{T}_1^\alpha\times\cdots\times\mathscr{T}_N^\alpha})_{j=1}^N$, and if the net (μ^α) "converges" to a point μ, then μ is a communication equilibrium of Γ.

This definition raises three issues. First, the reader may wonder whether it would be more appropriate to replace, in Definition 11, "communication equilibrium" by "correlated equilibrium," as formulated in Definition 9. Section 4 illustrates that strategic approximations defined in terms of correlated equilibrium limit points are problematic.

Second, what can be said about the existence of Nash equilibria in the finite games $(\mathscr{T}_j^\alpha, U_j|_{\mathscr{T}_1^\alpha\times\cdots\times\mathscr{T}_N^\alpha})_{j=1}^N$? The following result provides an answer.

Proposition *Suppose that* $\Gamma = (T_i, X_i, u_i, p)_{i=1}^N$ *is a Bayesian game in* $\mathfrak{G}$*, and let* $\mathfrak{G}_\Gamma = (\mathscr{T}_i, U_i)_{i=1}^N$ *represent its corresponding normal form, as defined in* (3). *Given finite sets* $\mathscr{S}_1, \ldots, \mathscr{S}_N$*, where* $\mathscr{S}_i \subseteq \mathscr{T}_i$ *for each* i*, there are finite supersets* $\mathscr{S}_1' \supseteq \mathscr{S}_1, \ldots, \mathscr{S}_N' \supseteq \mathscr{S}_N$*, where* $\mathscr{S}_i' \subseteq \mathscr{T}_i$ *for each* i*, such that the game* $(\mathscr{S}_j', U_j|_{\mathscr{S}_1'\times\cdots\times\mathscr{S}_N'})_{j=1}^N$ *possesses a Nash equilibrium.*

[6] Thanks to an anonymous referee for pointing out the limitations of standard approximation results in the context of discontinuous games.

Proof By Nash's Theorem, the mixed extension of $(\mathscr{S}_i, U_i|_{\mathscr{S}_1\times\cdots\times\mathscr{S}_N})_{i=1}^N$ has a Nash equilibrium $(q_1, \ldots, q_N) \in \times_{i=1}^N \Delta(\mathscr{S}_i)$, each q_i induces a member $\mu_i \in \mathscr{T}_i$ defined by

$$\mu_i(B_i|t_i) := \sum_{\mu_i \in \mathscr{S}_i} q_i(\mu_i)\mu_i(B_i|t_i),$$

and $\mu = (\mu_1, \ldots, \mu_N)$, so defined, is a Nash equilibrium of $(\mathscr{S}_i \cup \{\mu_i\}, U_i|_{(\mathscr{S}_1\cup\{\mu_1\})\times\cdots\times(\mathscr{S}_N\cup\{\mu_N\})})_{i=1}^N$. □

Third, note that the notion of convergence for the net (μ^α) in Definition 11 has not been specified. Each profile of behavioral strategies $(\mu_1, \ldots, \mu_N) \in \mathscr{T}$ in Γ can be identified with a correlated strategy $\mu : T \to \Delta(X)$ in $\mathscr{M}$ defined by

$$\mu(t) := \bigotimes_{i=1}^N \mu_i(t_i). \tag{7}$$

Thus, if one views the elements of the net (μ^α) and the limit μ in Definition 11 as members of $\mathscr{M}$, a topology on $\mathscr{M}$ fully determines the notion of convergence in Definition 11.

This paper considers a topology on the space $\mathscr{M}$ of correlated strategies of a Bayesian game $\Gamma = (T_i, X_i, u_i, p)_{i=1}^N$, defined as follows.

Given a $(\mathscr{B}(T_i), \mathscr{B}(T_i))$-measurable map $g_i : T_i \to T_i$, define $p * g_i \in \Delta(T)$ by

$$[p * g_i](A_i \times A_{-i}) := \int_{T_i\times A_{-i}} \delta_{g_i(t_i)}(A_i)p(dt) \tag{8}$$

for all measurable rectangles $A_i \times A_{-i} \subseteq T_i \times T_{-i}$ in $\mathscr{B}(T)$, where $\delta_{g_i(t_i)}$ denotes the Dirac measure in $\Delta(T_i)$ with support $\{g_i(t_i)\}$.

Let $\boldsymbol{P}_i$ denote the subset of $\Delta(T)$ defined by

$$\boldsymbol{P}_i := \big\{p * g_i \in \Delta(T) : g_i : T_i \to T_i \text{ is } (\mathscr{B}(T_i), \mathscr{B}(T_i))\text{-measurable}\big\},$$

and define

$$\boldsymbol{P} := \bigcup_{i=1}^N \boldsymbol{P}_i. \tag{9}$$

The map g_i can be viewed as a "misreporting rule," assigning a "reported type" $g_i(t_i) \in T_i$ to each type $t_i \in T_i$ of player i, and the compound measure $p * g_i$ defined in (8) describes the distribution over type profiles induced by the misreporting rule g_i, a distribution whereby Nature first chooses a type profile $(t_1, \ldots, t_N) \in T$ using the prior p and then player i "switches" her type from t_i to $g_i(t_i)$. The set of all such distributions is denoted by $\boldsymbol{P}_i$. The members of $\boldsymbol{P} := \bigcup_{i=1}^N \boldsymbol{P}_i$ can then be thought of as "distorted priors" in the sense that, for each $\hat{p} \in \boldsymbol{P}$, one and only one player i is misreporting her type according to some rule g_i.

Given $\hat{p} \in \Delta(T)$ and $\mu \in \mathscr{M}$, define the probability measure $\hat{p} \otimes \mu \in \Delta(T \times X)$ by

$$[\hat{p} \otimes \mu](A \times B) := \int_A \mu(B|t)\hat{p}(dt) \tag{10}$$

for all measurable rectangles $A \times B \subseteq T \times X$ in $\mathscr{B}(T \times X)$. The compound measure $\hat{p} \otimes \mu$ is a Borel probability measure on the Cartesian product $T \times X$ of type-actions profiles whereby, first, a type profile $t \in T$ is selected according to the "distorted prior" $\hat{p}$, and then, conditional on t, the correlated device $\mu(t) \in \Delta(X)$ is implemented to choose an action profile from X.

Now define the equivalence relation $\sim \subseteq \mathscr{M} \times \mathscr{M}$ as follows:

$$\mu \sim \nu \quad \Longleftrightarrow \quad \forall \hat{p} \in \boldsymbol{P}, \exists S \in \mathscr{B}(T) : \hat{p}(S) = 1 \& \forall t \in S, \mu(t) = \nu(t).$$

In words, μ and ν are equivalent if, for all $\hat{p} \in \boldsymbol{P}$, μ and ν differ only on a $\hat{p}$-null subset of T. Note that, for $\hat{p} \in \boldsymbol{P}$, $\hat{p} \otimes \mu = \hat{p} \otimes \nu$ whenever $\mu \sim \nu$.

Let $\mathscr{M}/\sim$ be the set of equivalence classes of elements of $\mathscr{M}$ generated by $\sim$,

$$\mathscr{M}/\sim := \{[\mu] : \mu \in \mathscr{M}\} = \{\{\nu \in \mathscr{M} : \nu \sim \mu\} : \mu \in \mathscr{M}\}.$$

Two correlated strategies in $\mathscr{M}$ belong to the same equivalence class if, for each $\hat{p} \in \boldsymbol{P}$, they coincide on a $\hat{p}$-full measure subset of T.

Next, endow $\Delta(T \times X)$ with the weak topology (Definition 1), and define, for each $\hat{p} \in \boldsymbol{P}$, the map $\vartheta_{\hat{p}} : \mathscr{M}/\sim \to \Delta(T \times X)$ by

$$\vartheta_{\hat{p}}([\mu]) := \hat{p} \otimes \mu.$$

The *initial topology* on $\mathscr{M}/\sim$ generated by the family of maps $\{\vartheta_{\hat{p}}\}_{\hat{p} \in \boldsymbol{P}}$, denoted by Υ, is the weakest topology on $\mathscr{M}/\sim$ that makes all the functions $\vartheta_{\hat{p}}$ continuous, and a net $([\mu^\alpha])$ in $\mathscr{M}/\sim$ Υ-converges to a point $[\mu] \in \mathscr{M}/\sim$, denoted as

$$[\mu^\alpha] \underset{\Upsilon}{\to} [\mu],$$

if and only if $\vartheta_{\hat{p}}([\mu^\alpha]) \underset{w}{\to} \vartheta_{\hat{p}}([\mu])$ for all $\hat{p} \in \boldsymbol{P}$ (see, *e.g.*, Aliprantis and Border 2006, Lemma 2.52), *i.e.*, if and only if

$$\hat{p} \otimes \mu^\alpha \underset{w}{\to} \hat{p} \otimes \mu, \quad \text{for all } \hat{p} \in \boldsymbol{P}.$$

We sometimes write $\mu^\alpha \underset{\Upsilon}{\to} \mu$ for $[\mu^\alpha] \underset{\Upsilon}{\to} [\mu]$, hoping that no confusion will arise.

2.3.1 Remarks about the topology Υ

Some remarks about the topology Υ are in order.

To begin, we define two natural topologies on the set of correlated strategies, $\mathscr{M}$, and compare them with the topology Υ.

First, consider the set

$$\{p \otimes \mu : \mu \in \mathscr{M}\}$$

of compound probability measures in $\Delta(T \times X)$ (recall the definition in (10)), endowed with the relativization of the weak topology on $\Delta(T \times X)$ (Definition 1). Next, consider the set $\mathfrak{M}$ of all equivalence classes in $\mathscr{M}$ of correlated strategies that only differ on a p-null subset of T (*i.e.*, two elements μ and ν in $\mathscr{M}$ are in the same equivalence class if there is a set $S \in \mathscr{B}(T)$ such that $p(S) = 1$ and $\mu(t) = \nu(t)$ for all $t \in S$). Let the set $\mathfrak{M}$ be provided with the initial topology on $\mathfrak{M}$ generated by the map $\mu \in \mathfrak{M} \mapsto p \otimes \mu$, so that a net $([\mu^\alpha])$ converges to $[\mu]$ in $\mathfrak{M}$ if and only if

$$p \otimes \mu^\alpha \underset{w}{\to} p \otimes \mu$$

(see, *e.g.*, Aliprantis and Border 2006, Lemma 2.52). Note that the map $[\mu] \mapsto p \otimes \mu$ is a homeomorphism between $\mathfrak{M}$ and $\{p \otimes \mu : \mu \in \mathscr{M}\}$, so that the relative weak topology on $\{p \otimes \mu : \mu \in \mathscr{M}\}$ can be viewed as a topology on (equivalence classes in) $\mathscr{M}$.

Clearly, the topology Υ is stronger than the weak topology on $\{p \otimes \mu : \mu \in \mathscr{M}\}$, *i.e.*, $[\mu^\alpha] \underset{\Upsilon}{\to} [\mu]$ implies $p \otimes \mu^\alpha \underset{w}{\to} p \otimes \mu$.[7] In addition, the topology Υ is weaker than the topology of uniform convergence on $\mathscr{M}$, *i.e.*, if the net (μ^α) converges uniformly to μ in $\mathscr{M}$ (so that for each $\epsilon > 0$, there exists α^* such that, for all $\alpha \geq \alpha^*$,

$$\varrho_{\Delta(X)}(\mu^\alpha(t), \mu(t)) < \epsilon, \quad \text{for all } t \in T$$

(recall the definition of $\varrho_{\Delta(X)}$ in (1))), then $[\mu^\alpha] \underset{\Upsilon}{\to} [\mu]$.[8] To see this, suppose that (μ^α) converges uniformly to μ in $\mathscr{M}$. It will be shown that

$$\hat{p} \otimes \mu^\alpha \underset{w}{\to} \hat{p} \otimes \mu, \quad \text{for all } \hat{p} \in \boldsymbol{P},$$

which, recall, is equivalent to Υ-convergence of $[\mu^\alpha]$ to $[\mu]$. By the Portmanteau Theorem (see, *e.g.*, Aliprantis and Border 2006, Theorem 15.3), it suffices to show that, for all $\hat{p} \in \boldsymbol{P}$,

$$\int_{T \times X} f(t, x)[\hat{p} \otimes \mu^\alpha](d(t, x)) \to \int_{T \times X} f(t, x)[\hat{p} \otimes \mu](d(t, x)), \tag{11}$$

[7] While the implication is mathematically correct, the first statement is not, for the sets $\mathscr{M}/\sim$ and $\mathfrak{M}$ differ from one another. Note that, for any $\mu \in \mathscr{M}$, the equivalence class $[\mu]$, viewed as a member of $\mathscr{M}/\sim$, is contained in the corresponding equivalence class from $\mathfrak{M}$, but the reverse containment does not hold.

[8] Again, the first assertion is an abuse of terminology.

for all bounded continuous maps $f : T \times X \to \mathbb{R}$. Fix $\hat{p} \in \boldsymbol{P}$ and a bounded continuous map $f : T \times X \to \mathbb{R}$. We claim that the net of maps

$$\left(t \in T \mapsto \int_X f(t,x)\mu^\alpha(dx|t)\right) \tag{12}$$

converges uniformly to the map $t \in T \mapsto \int_X f(t,x)\mu(dx|t)$. The proof of this fact is relegated to "Appendix". Because the net in (12) converges uniformly to the map $t \in T \mapsto \int_X f(t,x)\mu(dx|t)$, the Lebesgue Dominated Convergence Theorem for nets (see, *e.g.*, Dunford and Schwartz 1958, Theorem 7, p. 124) implies that (11) holds, as we sought.

Next, we consider a standard topology on $\mathscr{T}$, the set of behavioral strategy profiles. This topology, which is used in Balder (1988) and in Carbonell-Nicolau and McLean (2018), *inter alia*, is defined as the *product narrow quotient topology on* $\mathscr{T}$, *i.e.*, the product topology on $\mathscr{T}$ induced by the quotient topology for the narrow topology (see Balder 2001, Definition 1.3) on each factor $\mathscr{T}_i$. More precisely, let p_i be the marginal projection of p into $\Delta(X_i)$ (*i.e.*, $p_i \in \Delta(X_i)$ and $p_i(B) := p(B \times T_{-i})$ for all $B \in \mathscr{B}(X_i)$), consider the narrow quotient topology on the equivalence classes in $\mathscr{T}_i$ of transition probabilities that only differ on a p_i-null set, and endow $\mathscr{T}$ with its corresponding product topology. Letting $p_i \otimes \mu_i$ ($\mu_i \in \mathscr{T}_i$) be the compound measure in $\Delta(T_i \times X_i)$ defined by

$$[p_i \otimes \mu_i](A \times B) := \int_A \mu_i(B|t_i)p_i(dt_i)$$

for all measurable rectangles $A \times B \subseteq T_i \times X_i$ in $\mathscr{B}(T_i \times X_i)$, this product topology can be shown to be equivalent to the product weak topology on the set of distributional strategy profiles, $\times_{i=1}^N \mathscr{D}_i$, where $\mathscr{D}_i := \{p_i \otimes \mu_i : \mu_i \in \mathscr{T}_i\}$ (see Carbonell-Nicolau and McLean 2018, Sect. 5.2).[9]

The relativization of the topology Υ on $\mathscr{T}$ (recall that each element $(\mu_1, \ldots, \mu_N)$ of $\mathscr{T}$ is identified with a correlated strategy $\mu : T \to \Delta(X)$ in $\mathscr{M}$ defined by (7)) is fundamentally different from the product narrow quotient topology on $\mathscr{T}$. Indeed, it is possible for a sequence in $\mathscr{T}$ to converge, with respect to both topologies, to different limit points. This is illustrated in Sect. 4, which presents an example in which Υ-convergence (and even convergence with respect to the relative weak topology on $\{p \otimes \mu : \mu \in \mathscr{M}\}$) induces correlation of actions across players in the limit, while the product narrow quotient topology on $\mathscr{T}$ exhibits independent randomization over actions across players in the limit.

3 The main results

Recall that $\mathfrak{G}$ denotes the space of all Bayesian games $(T_i, X_i, u_i, p)_{i=1}^N$ such that $u_i(t,\cdot) : X \to \mathbb{R}$ is continuous for each $t \in T$ and i. The first main result of this paper

[9] See Castaing et al. (2004, ch. 2) for alternative formulations of these topologies.

asserts that the topology Υ defined in the previous section guarantees the existence of a strategic approximation of Γ (according to Definition 11) for all $\Gamma \in \mathfrak{G}$. We also illustrate the fact that Υ is the "weakest" possible topology ensuring that all the games in $\mathfrak{G}$ admit a strategic approximation, in the sense that, for weaker topologies, there are games in $\mathfrak{G}$ that do not admit a strategic approximation.

Theorem 1 *The topology Υ guarantees that every Bayesian game in $\mathfrak{G}$ admits a strategic approximation.*

The formal proof of Theorem 1 is provided in Sect. 6. The idea of the proof is as follows. Let $\Gamma = (T_i, X_i, u_i, p)_{i=1}^N$ be a Bayesian game in $\mathfrak{G}$. For each i, let $\mathscr{C}_i$ be the set of all the continuous behavioral strategies in $\mathscr{T}_i$. The space $\mathscr{C}_i$, endowed with the topology of uniform convergence of functions, is separable, and so a countable dense subset $\mathscr{Q}_i$ may be selected from $\mathscr{C}_i$. The set $\mathscr{Q} := \times_{i=1}^N \mathscr{Q}_i$ is a countable set of strategies contained in $\mathscr{T} = \times_{i=1}^N \mathscr{T}_i$, and it can be shown that $\mathscr{Q}$ is a strategic approximation of Γ (in the sense of Definition 11). Specifically, if, for each player i, $(\mathscr{T}_i^\alpha)$ is an increasing net of finite subsets of $\mathscr{T}_i$ whose union contains $\mathscr{Q}_i$, *i.e.*, $\mathscr{T}_i^\alpha \subseteq \mathscr{T}_i^\beta$ whenever $\alpha \leq \beta$ and $\bigcup_\alpha \mathscr{T}_i^\alpha \supseteq \mathscr{Q}_i$; if, for each α, $(\mu_1^\alpha, \ldots, \mu_N^\alpha)$ is a Nash equilibrium of the game

$$(\mathscr{T}_i^\alpha, U_i|_{\mathscr{T}_1^\alpha \times \cdots \times \mathscr{T}_N^\alpha})_{i=1}^N;$$

if, for each α, $\mu^\alpha : T \to \Delta(X)$ denotes the correlated strategy in $\mathscr{M}$ defined by

$$\mu^\alpha(t) := \bigotimes_{i=1}^N \mu_i^\alpha(t_i);$$

and if $[\mu^\alpha] \underset{\Upsilon}{\to} [\mu]$ for some $\mu \in \mathscr{M}$, then μ is a communication equilibrium of Γ.

The proof that μ is a communication equilibrium of Γ proceeds by contradiction, *i.e.*, it is shown that the assumption that μ is not a communication equilibrium, so that there exist i and $(\alpha_i, \eta_i) \in \mathscr{A}_i \times \mathscr{D}_i$ such that

$$\begin{aligned}&\int_{T\times X} u_i(t,x)[p\otimes\mu](d(t,x))\\ &\quad< \int_T \int_{T_i} \int_X \int_{X_i} u_i(t, y_i, x_{-i})\alpha_i(dy_i|t_i, x_i)\mu(dx|\tau_i, t_{-i})\eta_i(d\tau_i|t_i)p(dt),\end{aligned}$$

leads to an impossibility. Using the Υ-convergence of $([\mu^\alpha])$ to $[\mu]$, it is possible to extract sequences $(\mathscr{T}_i^n)$, $i \in \{1, \ldots, N\}$, and $(\mu_1^n, \ldots, \mu_N^n)$ such that, for large enough n and for some $\rho_i^* \in \mathscr{T}_i^n$, one has $U_i(\rho_i^*, \mu_{-i}^n) > U_i(\mu_1^n, \ldots, \mu_N^n)$. This gives the desired contradiction, since $(\mu_1^n, \ldots, \mu_N^n)$ is a Nash equilibrium of the game $(\mathscr{T}_\iota^n, U_\iota|_{\mathscr{T}_1^n \times \cdots \times \mathscr{T}_N^n})_{\iota=1}^N$.

We now show that the topology Υ is necessary for the games in $\mathfrak{G}$ to admit a strategic approximation.[10] To this end, we consider a very simple Bayesian game, denoted by

[10] The argument here is based on an example provided by an anonymous referee.

$\Gamma = (T_i, X_i, u_i, p)_{i=1}^N$. There are two players (*i.e.*, $N = 2$), and type spaces are identical doubletons, $T_1 = T_2 := \{0, 1\}$. The common prior p is uniform on the diagonal $\{(0, 0), (1, 1)\}$. Player 1 has one action, A, and player 2 has two actions, A and B, so that $X_1 := \{A\}$ and $X_2 := \{A, B\}$. Types are payoff-irrelevant, and the payoff bi-matrix is as follows:

	A	B
A	1,1	2,0

There is a unique communication equilibrium $\mu : T \to \Delta(X)$ in this trivial game, given by $\mu(t) := \delta_{(A,A)}$ for all $t \in T$, where $\delta_{(A,A)}$ denotes the Dirac measure in $\Delta(X)$ with support $\{(A, A)\}$. To see this, note that $X = \{(A, A), (A, B)\}$ and suppose that $\hat{\mu}$ is a correlated strategy in $\mathscr{M}$ such that $\hat{\mu}(\{(A, B)\}|t) > 0$ for some $t \in T$. If $t \in \{(0, 0), (1, 1)\}$, then it is clear that player 2 has an incentive to be disobedient, playing A with probability 1 upon receiving the signal t_2. If $t \in \{(0, 1), (1, 0)\}$, then player 1 can improve her payoff by being dishonest, lying about her type (if $t = (0, 1)$, she reports $\tau_1 = 0$ upon receiving the signal $t_1 = 1$, and if $t = (1, 0)$, she reports $\tau_1 = 1$ upon receiving the signal $t_1 = 0$).

Now let $\mathscr{T}_i^1 \subseteq \mathscr{T}_i^2 \subseteq \cdots$ be an increasing sequence of finite subsets of $\mathscr{T}_i$, $i \in \{1, 2\}$. (Here $\mathscr{T}_i$ is the set of all maps $\nu_i : \{0, 1\} \to \Delta(X_i)$.) Let (μ_1^*, μ_2^*) be the unique Nash equilibrium of the normal form of Γ, $(\mathscr{T}_i, U_i)_{i=1}^2$, *i.e.*, $\mu_1^*(t_1) := \delta_A$ and $\mu_2^*(t_2) := \delta_A$. Then, for each n, $(\mu_1^n, \mu_2^n) := (\mu_1^*, \mu_2^*)$ is a Nash equilibrium of the game $(\mathscr{T}_i^n \cup \{\mu_i^*\}, U_i|_{(\mathscr{T}_1^n \cup \{\mu_1^*\}) \times (\mathscr{T}_2^n \cup \{\mu_2^*\})})_{i=1}^2$.

Define $\mu^n : T \to \Delta(X)$ by $\mu^n(t) := \mu_1^n(t_1) \otimes \mu_2^n(t_2)$ and suppose that (μ^n) does not $\mathfrak{T}$-converge to ν in $\mathscr{M}$. Then there exist i and $\hat{p} \in \boldsymbol{P}_i$ such that $\hat{p} \otimes \mu^n$ does not converge weakly to $\hat{p} \otimes \nu$. But then there exists $\tau \in T$ such that $\nu(\tau) \neq \mu_1^*(\tau_1) \otimes \mu_2^*(\tau_2) = \mu(\tau)$, implying that ν is not a communication equilibrium of Γ. Thus, if one employs a notion of convergence weaker than $\mathfrak{T}$-convergence in Definition 11, the game Γ does not admit a strategic approximation.

3.1 On the existence of communication equilibrium

While the general existence of communication equilibria for the class $\mathfrak{G}$ of Bayesian games is an open question, Theorem 1 can be used to identify a subclass of $\mathfrak{G}$ for which "robust" communication equilibria exist (in the sense of Definition 11).[11]

[11] Finding conditions under which the topology $\mathfrak{T}$ is compact would be useful to establish the general existence of communication equilibria within the class of games $\mathfrak{G}$. As per Exercise 2.48 in Megginson (1998), the topology $\mathfrak{T}$ is compact if and only if (i) $\vartheta_{\hat{p}}(\mathscr{M}/\sim)$ is compact in $\Delta(T \times X)$ for every $\hat{p} \in \boldsymbol{P}$; and (ii) the image of $\mathscr{M}/\sim$ in $\prod_{\boldsymbol{P}} \Delta(T \times X)$ under the map $[\mu] \in \mathscr{M}/\sim \mapsto (\hat{p} \otimes \mu)_{\hat{p} \in \boldsymbol{P}}$ is closed. While the first condition can be shown to hold, we have not been able to establish the second condition, which requires the following: if $([\mu^\alpha])$ is a net in $\mathscr{M}/\sim$ such that

$$\hat{p} \otimes \mu^\alpha \underset{w}{\to} \hat{p} \otimes \mu_{\hat{p}} \quad \text{for all } \hat{p} \in \boldsymbol{P},$$

Strategic approximations add a sense of robustness to the notion of communication equilibrium. Indeed, if there is a sequence of Bayes-Nash equilibria of games with finite, successively larger spaces of behavioral strategies, and if the sequence converges, the limit point is, by virtue of Theorem 1, a communication equilibrium. This equilibrium is "robust" in the sense that it describes Bayes-Nash equilibrium behavior in "nearby" finite Bayesian games. Of course, such a "strategic approximation" is vacuous if such a sequence of approximating Bayes-Nash equilibria does not exist, and so a natural question is whether the games in $\mathfrak{G}$ can be shown to have "robust" approximate communication equilibria. The following result provides, in certain cases, an answer in the affirmative.

Let $\mathfrak{G}^*$ be the set of all Bayesian games $(T_i, X_i, u_i, p)_{i=1}^N$ in $\mathfrak{G}$ satisfying the following condition: Given an increasing sequence of finite subsets of $\mathscr{T}_i$, $\mathscr{T}_i^1 \subseteq \mathscr{T}_i^2 \subseteq \cdots$ ($i \in \{1, \ldots, N\}$), there exists (passing to a subsequence if necessary) a corresponding sequence $(\mu_1^n, \ldots, \mu_N^n)$, where each $(\mu_1^n, \ldots, \mu_N^n)$ is a Nash equilibrium of $(\mathscr{T}_i^n, U_i|_{\mathscr{T}_1^n \times \cdots \times \mathscr{T}_N^n})_{i=1}^N$, such that the sequence of correlated strategies (μ^n) defined by

$$\mu^n(t) := \bigotimes_{i=1}^{N} \mu_i^n(t_i) \tag{13}$$

satisfies

$$\frac{1}{m}\sum_{n=1}^{m} \mu^n(t) \xrightarrow[w]{m\to\infty} \mu(t), \quad \text{for every } t \in T, \tag{14}$$

for some $\mu \in \mathscr{M}$.

Corollary (to Theorem 1) *Suppose that $\Gamma = (T_i, X_i, u_i, p)_{i=1}^N$ is a Bayesian game in $\mathfrak{G}^*$. Then there is (i) an increasing sequence of finite subsets of $\mathscr{T}_i$, $\mathscr{T}_i^1 \subseteq \mathscr{T}_i^2 \subseteq \cdots$ ($i \in \{1, \ldots, N\}$), and (ii) a Nash equilibrium $(\mu_1^n, \ldots, \mu_N^n)$ of $(\mathscr{T}_i^n, U_i|_{\mathscr{T}_1^n \times \cdots \times \mathscr{T}_N^n})_{i=1}^N$, for each n, such that the sequence of correlated strategies (μ^n) defined by (13) $\mathfrak{T}$-converges in $\mathscr{M}$, and the limit point is a communication equilibrium of Γ.*

Proof By Theorem 1, Γ admits a strategic approximation $\mathscr{T}^\infty = \times_{i=1}^N \mathscr{T}_i^\infty$, and so, because $\Gamma \in \mathfrak{G}^*$, and given an increasing sequence $\mathscr{T}_i^1 \subseteq \mathscr{T}_i^2 \subseteq \cdots$ of finite subsets of $\mathscr{T}_i$ whose union contains $\mathscr{T}_i^\infty$ ($i \in \{1, \ldots, N\}$), there exists (passing to a subsequence if necessary) a corresponding sequence $(\mu_1^n, \ldots, \mu_N^n)$, where each $(\mu_1^n, \ldots, \mu_N^n)$ is a Nash equilibrium of $(\mathscr{T}_i^n, U_i|_{\mathscr{T}_1^n \times \cdots \times \mathscr{T}_N^n})_{i=1}^N$, such that the sequence of correlated strategies (μ^n) defined by (13) satisfies (14) for some $\mu \in \mathscr{M}$. Applying Theorem 2.6 in Balder (2001), it follows that $\hat{p} \otimes \mu^n \xrightarrow[w]{} \hat{p} \otimes \mu$ for all $\hat{p} \in \Delta(T)$, implying,

---

where, for each $\hat{p} \in \boldsymbol{P}$, $\mu_{\hat{p}}$ is an element of $\mathscr{M}$, then there exists $\mu^* \in \mathscr{M}$ such that

$$\hat{p} \otimes \mu_{\hat{p}} = \hat{p} \otimes \mu^*, \quad \text{for all } \hat{p} \in \boldsymbol{P}.$$

in particular, that $\mu^n \underset{\Upsilon}{\to} \mu$. Because $\mathscr{T}^\infty$ is a strategic approximation of Γ, it follows that μ is a communication equilibrium of Γ. □

4 Discussion

To begin, we consider the existence—or lack thereof—of strategic approximations of the normal form $\mathfrak{G}_\Gamma$ (defined in (3)) of a Bayesian game Γ, in the sense of applying Definition 10 directly to $\mathfrak{G}_\Gamma$. The following example illustrates that the normal form of a Bayesian game may be approximated by a sequence of finite "subgames" for which there is a corresponding sequence of Nash equilibria converging to a non-Nash equilibrium profile in the limit game. Two modes of convergence for the sequence of Nash equilibrium profiles are considered. The first convergence mode derives from the topology on behavioral strategy profiles used in Balder (1988) and in Carbonell-Nicolau and McLean (2018), *inter alia*, while the second is weaker than Υ-convergence.

Consider the following two-player Bayesian game taken from Milgrom and Weber (1985, Example 2). Suppose that each player's type is a member of the [0, 1] interval, and let the action set of each player be a doubleton, $\{1, 2\}$. The payoffs are independent of the types, and are given by the standard "Battle of the Sexes" payoff bi-matrix:

	1	2
1	2,1	0,0
2	0,0	1,2

Suppose that type profiles (t_1, t_2) are uniformly distributed on the 45° line in $[0, 1] \times [0, 1]$.

Let $\Gamma = (T_i, X_i, u_i, p)_{i=1}^N$ denote the corresponding Bayesian game, and let $\mathfrak{G}_\Gamma = (\mathscr{T}_i, U_i)_{i=1}^N$ represent its normal form, as defined in (3).

For each player i and each $n \in \mathbb{N}$, let $s_i^n(t_i)$ be the strategy defined as[12]

$$s_i^n(t_i) := \begin{cases} 1 & \text{if the integer part of } nt_i \text{ is odd,} \\ 2 & \text{otherwise.} \end{cases}$$

Now for each player i, let $\mathscr{T}_i^1 \subseteq \mathscr{T}_i^2 \subseteq \cdots$ be any increasing sequence of finite behavioral strategy sets, and define $\mathscr{Y}_i^n := \mathscr{T}_i^n \cup \{s_i^n\}$ for each i and n. Clearly, for each n, the strategy profile (s_1^n, s_2^n) is a Nash equilibrium of the normal form in which the players' strategy spaces are $\mathscr{Y}_1^n$ and $\mathscr{Y}_2^n$.

There are number of topologies that one may consider when applying Definition 10. For example, one may assume that the sequence (s_1^n, s_2^n)—or, more precisely, the sequence $(\delta_{s_1^n}, \delta_{s_2^n})$ in $\mathscr{T} = \mathscr{T}_1 \times \mathscr{T}_2$, where $\delta_{s_i^n}$ denotes the map $t_i \in T_i \mapsto \delta_{s_i^n(t_i)} \in$

[12] It is clear that these functions can be viewed as behavioral strategies that assign a Dirac probability measure to each type.

$\Delta(X_i)$, and where $\delta_{s_i^n(t_i)}$ represents the Dirac measure in $\Delta(X_i)$ with support $\{s_i^n(t_i)\}$—converges to a point (μ_1, μ_2) in $\mathscr{T}$ if and only if the sequence $(p_1 \otimes \delta_{s_1^n}, p_2 \otimes \delta_{s_2^n})$ converges weakly to $(p_1 \otimes \mu_1, p_2 \otimes \mu_2)$, *i.e.*, if and only if $p_i \otimes \delta_{s_i^n}$ converges weakly to $p_i \otimes \mu_i$ for each i, where each p_i is the marginal projection of p into $\Delta(X_i)$ (*i.e.*, $p_i \in \Delta(X_i)$ and $p_i(B) := p(B \times T_{-i})$ for all $B \in \mathscr{B}(X_i)$), and where $p_i \otimes \nu_i$ ($\nu_i \in \mathscr{T}_i$) is defined as the compound measure in $\Delta(T_i \times X_i)$ defined by

$$[p_i \otimes \nu_i](A \times B) := \int_A \nu_i(B|t_i) p_i(dt_i)$$

for all measurable rectangles $A \times B \subseteq T_i \times X_i$ in $\mathscr{B}(T_i \times X_i)$. Accordingly, $\mathscr{T}_i$ is viewed as a subspace of $\Delta(T_i \times X_i)$ with the w-topology (Definition 1), which renders $\mathscr{T}_i$ metric.[13]

Using this convergence mode, the sequence (s_i^n) converges to a strategy in which player i ignores her type and plays each action (1 or 2) with equal probability, and the limit point for the sequence (s_1^n, s_2^n) is clearly not a Nash equilibrium of $\mathfrak{G}_\Gamma$. Consequently, the game $\mathfrak{G}_\Gamma$ does not admit a strategic approximation in the sense of Definition 10.

In terms of topologizing $\mathscr{T}$, another possibility is to identify each member (μ_1, μ_2) of $\mathscr{T}$ with the measure $p \otimes \mu$ in $\Delta(T \times X)$, where $\mu : T \to \Delta(X)$ is defined by $\mu(t) := \mu_1(t_1) \otimes \mu_2(t_2)$ and where $p \otimes \mu$ is the compound measure defined by

$$[p \otimes \mu](A \times B) := \int_A \mu(B|t) p(dt)$$

for all measurable rectangles $A \times B \subseteq T \times X$ in $\mathscr{B}(T \times X)$. Accordingly, $\mathscr{T}$ is regarded as a subspace of $\Delta(T \times X)$ with the w-topology (Definition 1). Note that, because $p \in \boldsymbol{P}$, the associated notion of convergence is weaker than $\mathfrak{T}$-convergence.

In this case, the sequence (s_1^n, s_2^n) converges to a measure ϱ in $\Delta(T \times X)$ that chooses (t_1, t_2) uniformly from the diagonal $\{(\tau_1, \tau_2) \in [0, 1]^2 : \tau_1 = \tau_2\}$, and then, conditional on (t_1, t_2), the action profiles $(1, 1)$ and $(2, 2)$ are selected equiprobably. This limit point cannot possibly be generated by a measure of the form $p \otimes \mu$, where $\mu(t) = \mu_1(t_1) \otimes \mu_2(t_2)$ for all $t \in T$ and $(\mu_1, \mu_2) \in \mathscr{T}$, and so it is not a Nash equilibrium of $\mathfrak{G}_\Gamma$. The conclusion is therefore the same as before: the game $\mathfrak{G}_\Gamma$ does not admit a strategic approximation in the sense of Definition 10.

In light of this example, a natural next question is whether progress can be made by weakening the solution concept for the limit game. This is precisely what Definition 11—which uses the weaker communication equilibrium concept—does, and the main results from Sect. 3 provide an affirmative answer. However, an equally valid question is whether one can replace, in Definition 11, "communication equilibrium" by "correlated equilibrium" (in the sense of Definition 9), and consider the resulting

[13] An equivalent topology is the product narrow topology on $\mathscr{T}$ (see Balder 2001, Definition 1.3), or, more precisely, the product topology induced by the narrow quotient topology on the equivalence classes from each factor $\mathscr{T}_i$ of transition probabilities that only differ on a p_i-null set. This product topology is equivalent to the product weak topology on $\times_i \mathscr{D}_i$, where $\mathscr{D}_i := \{p_i \otimes \mu_i : \mu_i \in \mathscr{T}_i\}$ (see Carbonell-Nicolau and McLean 2018, Sect. 5.2).

notion of strategic approximation in lieu of that in Definition 11. In the remainder of this section, it is shown that this alternative to Definition 11 is problematic. Specifically, it is shown that, for a slight variation of the example considered above, and viewing $\mathscr{T}$ as a subspace of $\Delta(T \times X)$ with the w-topology (which yields a convergence mode weaker than $\mathfrak{T}$-convergence), and for *any* sequence of finite versions of a Bayesian game that includes a particular sequence of behavioral strategy profiles in $\mathscr{T}$, there is a corresponding sequence of Nash equilibria converging to a correlated strategy in $\mathscr{M}$ that is *not* a correlated strategy profile in $\times_{i=1}^{N} \mathscr{X}_i$ (recall the definition of $\mathscr{X}_i$ introduced immediately before Definition 9).

First, observe that the limit measure ϱ from the previous example is expressible as a measure of the form

$$\sigma(A \times B) = \int_{A\times[0,1]} [\varphi_1(t_1, a) \otimes \varphi_2(t_2, a)]\,(B)[p \otimes \lambda](d(t, a)) \tag{15}$$

for all measurable rectangles $A \times B \subseteq T \times X$ in $\mathscr{B}(T \times X)$. Indeed, it suffices to define, for each $(t_1, t_2) \in T$, $\varphi_1(\cdot|(t_1, a))$ and $\varphi_2(\cdot|(t_2, a))$ as the Dirac probability measure supported on $\{1\}$ if $a \in \left[0, \frac{1}{2}\right)$, and otherwise let $\varphi_1(\cdot|(t_1, a))$ and $\varphi_2(\cdot|(t_2, a))$ be the Dirac probability measure supported on $\{2\}$. Thus, the limit measure ϱ may be viewed as a correlated profile in $\times_{i=1}^{N} \mathscr{X}_i$ (see the definition of $\mathscr{X}_i$ introduced immediately before Definition 9).

However, this is not true in general. Specifically, consider a variant of the above game in which Nature chooses the type profiles $\left(\frac{1}{3}, \frac{1}{3}\right)$, $\left(1, \frac{1}{3}\right)$, and $(1, 1)$, each with $\frac{1}{4}$ probability, and randomizes uniformly over the diagonal $\left\{(\tau_1, \tau_2) \in [0, 1]^2 : \tau_1 = \tau_2\right\}$ with $\frac{1}{4}$ probability. Suppose that the payoff bimatrix corresponding to the type profile $\left(1, \frac{1}{3}\right)$ is given by

	1	2
1	1,1	1,1
2	1,1	1,1

For each n, (s_1^n, s_2^n) is a Nash equilibrium of the normal form in which the players' strategy spaces are $\mathscr{Y}_1^n$ and $\mathscr{Y}_2^n$. In addition, the sequence (s_1^n, s_2^n) converges to a measure in $\Delta(T \times X)$ that selects, conditional on (t_1, t_2), the action profiles $(1, 1)$, $(1, 2)$, $(2, 1)$, and $(2, 2)$ with respective probabilities $\frac{1}{3}$, $\frac{1}{6}$, $\frac{1}{6}$, and $\frac{1}{3}$ if $(t_1, t_2) = \left(1, \frac{1}{3}\right)$, and $(1, 1)$ and $(2, 2)$ equiprobably otherwise. Note that in this case the conditional distribution on actions for the limit measure is not constant, as in the previous example, but rather depends on the type profile selected by Nature. Because $p \otimes \lambda$ is a product measure (so that the conditional distribution of a does not vary with t), the limit measure is not expressible as a measure of the form σ as defined in (15), and, consequently, it cannot be viewed as a correlated equilibrium in the sense of Definition 9. The induced limit correlated strategy is, as can be easily verified, a communication equilibrium.

Similar arguments apply if one uses instead the topology Υ from Sect. 2.3 in Definition 11.[14]

5 Sketch of the proof of Theorem 1

The details of the proof of Theorem 1 are relegated to Sect. 6. In this section, we present a sketch of the proof, outlining the main argument.

Theorem 1 asserts that the topology Υ guarantees that every Bayesian game in $\mathfrak{G}$ admits a strategic approximation.

Fix a game $\Gamma = (T_i, X_i, u_i, p)_{i=1}^N$ in $\mathfrak{G}$. For each i, let $\mathscr{C}_i$ represent the set of all the continuous members of the function space $\Delta(X_i)^{T_i}$, and put $\mathscr{C} := \times_{i=1}^N \mathscr{C}_i$. The space $\mathscr{C}_i$, endowed with the topology of uniform convergence of functions, is separable, and so a countable dense subset $\mathscr{Q}_i$ may be selected from $\mathscr{C}_i$. The set $\mathscr{Q} := \times_{i=1}^N \mathscr{Q}_i$ is a countable set of strategies contained in $\mathscr{T} = \times_{i=1}^N \mathscr{T}_i$ and we claim that $\mathscr{Q}$ is a strategic approximation of Γ (in the sense of Definition 11).

For each player i, let $(\mathscr{T}_i^\alpha)$ be an increasing net of finite subsets of $\mathscr{T}_i$ whose union contains $\mathscr{Q}_i$, *i.e.*, $\mathscr{T}_i^\alpha \subseteq \mathscr{T}_i^\beta$ whenever $\alpha \leq \beta$ and $\bigcup_\alpha \mathscr{T}_i^\alpha \supseteq \mathscr{Q}_i$. Suppose that for each α, $(\mu_1^\alpha, \ldots, \mu_N^\alpha)$ is a Nash equilibrium of the game

$$(\mathscr{T}_i^\alpha, U_i|_{\mathscr{T}_1^\alpha \times \cdots \times \mathscr{T}_N^\alpha})_{i=1}^N.$$

For each α, let $\mu^\alpha : T \to \Delta(X)$ be the correlated strategy in $\mathscr{M}$ defined by

$$\mu^\alpha(t) := \bigotimes_{i=1}^N \mu_i^\alpha(t_i).$$

Suppose that $[\mu^\alpha] \underset{\Upsilon}{\to} [\mu]$ for some $\mu \in \mathscr{M}$. We must show that μ is a communication equilibrium of Γ. To this end, we suppose that there exist i and profitable deviations $(\alpha_i, \eta_i) \in \mathscr{A}_i \times \mathscr{D}_i$ such that

$$\int_{T\times X} u_i(t,x)[p\otimes\mu](d(t,x))$$
$$< \int_T \int_{T_i} \int_X \int_{X_i} u_i(t, y_i, x_{-i})\alpha_i(dy_i|t_i, x_i)\mu(dx|\tau_i, t_{-i})\eta_i(d\tau_i|t_i)p(dt),$$

[14] For $t = (t_1, t_2) \in T$ with $t_1 = t_2$, the sequence of measures

$$\begin{aligned} &\delta_{s_1^1(t_1)} \otimes \delta_{s_2^1(t_2)}, \frac{1}{2}(\delta_{s_1^1(t_1)} \otimes \delta_{s_2^1(t_2)}) + \frac{1}{2}(\delta_{s_1^2(t_1)} \otimes \delta_{s_2^2(t_2)}), \\ &\frac{1}{3}(\delta_{s_1^1(t_1)} \otimes \delta_{s_2^1(t_2)}) + \frac{1}{3}(\delta_{s_1^2(t_1)} \otimes \delta_{s_2^2(t_2)}) + \frac{1}{3}(\delta_{s_1^3(t_1)} \otimes \delta_{s_2^3(t_2)}), \ldots \end{aligned} \tag{16}$$

converges weakly to the measure $\mu(t) \in \Delta(X)$ defined by $\mu(\{1,1\}|t) = \mu(\{2,2\}|t) = \frac{1}{2}$. In general, for every $t = (t_1, t_2) \in T$, the sequence $(\mu^n(t))$ given in (16) converges weakly to some measure $\mu(t) \in \Delta(X)$. Applying Theorem 2.6 in Balder (2001), it follows that $\hat{p} \otimes \mu^n \underset{w}{\to} \hat{p} \otimes \mu$ for all $\hat{p} \in \Delta(T)$, and so, in particular, $\mu^n \underset{\Upsilon}{\to} \mu$.

and derive a contradiction.

We now outline the steps leading to the desired contradiction. The proofs of the assertions made here can be found in Sect. 6.

1. There is no loss of generality in assuming that η_i satisfies the following: there exists a $(\mathscr{B}(T_i), \mathscr{B}(T_i))$-measurable map $g_i : T_i \to T_i$ such that $\eta_i(t_i) = \delta_{g_i(t_i)}$ for each $t_i \in T_i$, where $\delta_{g_i(t_i)}$ denotes the Dirac measure in $\Delta(T_i)$ with support $\{g_i(t_i)\}$. (See Claim A.)
2. There are sequences $(\mathscr{T}_1^n, \ldots, \mathscr{T}_N^n)$ and $(\mu_1^n, \ldots, \mu_N^n)$ satisfying the following: for each j, $\mathscr{T}_j^1 \subseteq \mathscr{T}_j^2 \subseteq \cdots$ and $\bigcup_n \mathscr{T}_j^n \supseteq \mathscr{Q}_j$; for each j and n, $\mathscr{T}_j^n$ is a finite subset of $\mathscr{T}_j$ and $(\mu_1^n, \ldots, \mu_N^n)$ is a Nash equilibrium of the game $(\mathscr{T}_\iota^n, U_\iota|_{\mathscr{T}_1^n \times \cdots \times \mathscr{T}_N^n})_{\iota=1}^N$; and
$$p \otimes \mu^n \underset{w}{\to} p \otimes \mu \quad \text{and} \quad [p * g_i] \otimes \mu^n \underset{w}{\to} [p * g_i] \otimes \mu,$$
where $\mu^n : T \to \Delta(X)$ is the correlated strategy in $\mathscr{M}$ defined by $\mu^n(t) := \otimes_{\iota=1}^N \mu_\iota^n(t_\iota)$ and $p*g_i$ denotes the compound measure defined in (8). (See Claim B.)
3. Define the correlated strategy $\mu^* : T \to \Delta(X)$ obtained from μ when player i misreports according to η_i and uses the deviation plan α_i:
$$\mu^*(B_i \times B_{-i}|t) := \int_{T_i} \int_{X_i \times B_{-i}} \alpha_i(B_i|t_i, x_i)\mu(dx|\tau_i, t_{-i})\eta_i(d\tau_i|t_i)$$
for all $B_i \times B_{-i} \subseteq X_i \times X_{-i}$ in $\mathscr{B}(X_i \times X_{-i})$. The correlated strategy μ^* can be "approximated" by an analogous transformation of the sequence (μ^n), in the following sense:
$$p \otimes \rho^n \underset{w}{\to} p \otimes \mu^*,$$
where $\rho^n : T \to \Delta(X)$ is defined by
$$\rho^n(t) := \rho_i^n(t_i) \otimes \left[\bigotimes_{j \neq i} \mu_j^n(t_j) \right]$$
and $\rho_i^n \in \mathscr{T}_i$ is obtained from μ_i^n when player i misreports according to η_i and uses the deviation plan α_i:
$$\rho_i^n(B|t_i) := \int_{T_i} \int_{X_i} \alpha_i(B|t_i, x_i)\mu_i^n(dx_i|\tau_i)\eta_i(d\tau_i|t_i).$$
(See Claim C.)
4. For large enough n, the behavioral strategy ρ_i^n from the previous item can be "approximated" by a behavioral strategy ρ_i in the following sense: there exists $\rho_i \in \mathscr{T}_i$ such that some subsequence of $(\hat{\rho}^n)$, denoted again by $(\hat{\rho}^n)$, satisfies
$$p \otimes \hat{\rho}^n \underset{w}{\to} p \otimes \mu^*,$$

where $\hat{\rho}^n : T \to \Delta(X)$ is defined by

$$\hat{\rho}^n(t) := \rho_i(t_i) \otimes \left[\underset{j \neq i}{\otimes} \mu_j^n(t_j) \right].$$

(See Claim D.)

5. There is no loss of generality in assuming that the behavioral strategy ρ_i from the previous item is a member of $\mathscr{C}_i$, in the following sense: there exists $\rho_i^* \in \mathscr{C}_i$ such that

$$p \otimes \tilde{\rho}^n \underset{w}{\to} p \otimes \mu^{**},$$

where $\tilde{\rho}^n : T \to \Delta(X)$ is defined by

$$\tilde{\rho}^n(t) := \rho_i^*(t_i) \otimes \left[\underset{j \neq i}{\otimes} \mu_j^n(t_j) \right],$$

and where μ^{**} satisfies

$$\int_{T \times X} u_i(t,x)[p \otimes \mu](d(t,x)) < \int_{T \times X} u_i(t,x)[p \otimes \mu^{**}](d(t,x)). \quad (17)$$

(See Claim E.)

6. There exists a sequence (ν_i^n) with $\nu_i^n \in \mathscr{T}_i^n$ for each n such that

$$p \otimes \nu^n \underset{w}{\to} p \otimes \mu^{**},$$

where $\nu^n : T \to \Delta(X)$ is defined by

$$\nu^n(t) := \nu_i^n(t_i) \otimes \left[\underset{j \neq i}{\otimes} \mu_j^n(t_j) \right].$$

(See the proof of Claim F.) Consequently, by Theorem 3.1 in Balder (2001),

$$\int_{T \times X} u_i(t,x)[p \otimes \nu^n](d(t,x)) \to \int_{T \times X} u_i(t,x)[p \otimes \mu^{**}](d(t,x)).$$

Similarly, because $p \otimes \mu^n \underset{w}{\to} p \otimes \mu$ (item 2), one obtains

$$\int_{T \times X} u_i(t,x)[p \otimes \mu^n](d(t,x)) \to \int_{T \times X} u_i(t,x)[p \otimes \mu](d(t,x)).$$

Consequently (using (17)),

$$\int_{T \times X} u_i(t,x)[p \otimes \mu^n](d(t,x)) \to \int_{T \times X} u_i(t,x)[p \otimes \mu](d(t,x))$$

$$
\begin{aligned}
&< \int_{T\times X} u_i(t,x)[p\otimes\mu^{**}](d(t,x)) \\
&\leftarrow \int_{T\times X} u_i(t,x)[p\otimes\nu^{n}](d(t,x)),
\end{aligned}
$$

and so it follows that there exists n^{**} such that

$$
\begin{aligned}
U_i(\nu_i^{n^{**}}, \mu_{-i}^{n^{**}}) &= \int_{T\times X} u_i(t,x)[p\otimes\nu^{n^{**}}](d(t,x)) \\
&> \int_{T\times X} u_i(t,x)[p\otimes\mu^{n^{**}}](d(t,x)) \\
&= U_i(\mu_1^{n^{**}},\ldots,\mu_N^{n^{**}}), \quad \text{for all } n \geq n^{**}.
\end{aligned}
$$

Since $\nu_i^{n^{**}} \in \mathscr{T}_i^{n^{**}}$, this gives the desired contradiction, since $(\mu_1^n,\ldots,\mu_N^n)$ is a Nash equilibrium of the game $(\mathscr{T}_\iota^n, U_\iota|_{\mathscr{T}_1^n\times\cdots\times\mathscr{T}_N^n})_{\iota=1}^N$.

6 Proof of Theorem 1

In preparation for the proof of Theorem 1, we introduce some terminology and develop a series of lemmas. To keep the flow of the main argument, the proofs of most of the lemmas are relegated to "Appendix".

Let Y and Z be metric spaces, and let $\Delta(Y\times Z)$ denote the set of all probability measures on $(Y\times Z, \mathscr{B}(Y)\otimes\mathscr{B}(Z))$. The set of all bounded and continuous real-valued functions on Z is denoted by $C^b(Z)$.

Definition 12 The ***ws-topology*** on $\Delta(Y\times Z)$ is the coarsest topology for which all the functionals in

$$
\left\{\mu\in\Delta(Y\times Z)\mapsto \int_{S\times Z} f(z)\mu(d(y,z))\in\mathbb{R} : (S,f)\in\mathscr{B}(Y)\times C^b(Z)\right\}
$$

are continuous.

We sometimes write $\nu^n \underset{ws}{\longrightarrow} \nu$ to indicate that the sequence of measures (ν^n) converges to ν with respect to the *ws*-topology.

Definition 13 The ***s-topology*** on $\Delta(Y)$ is the coarsest topology for which all the functionals in

$$
\{\mu\in\Delta(Y)\mapsto\mu(S)\in\mathbb{R} : S\in\mathscr{B}(Y)\}
$$

are continuous.

Suppose that Y and Z are compact metric spaces. Given $p\in\Delta(Y)$ and a $(\mathscr{B}(Y),\mathscr{B}(\Delta(Z)))$-measurable map $\mu : Y\to\Delta(Z)$, define $p\otimes\mu\in\Delta(Y\times Z)$

by

$$[p \otimes \mu](A \times B) := \int_A \mu(B|y)p(dy)$$

for all $A \times B \subseteq Y \times Z$ in $\mathscr{B}(Y \times Z)$.

Let $\mathscr{P}^p(Y \times Z)$ be the set of all ν in $\Delta(Y \times Z)$ that take the form $\nu = p \otimes \mu$ for some $\mu : Y \to \Delta(Z)$.

Lemma 1 *Suppose that Y and Z are compact metric spaces, and let $p \in \Delta(Y)$. Then $\mathscr{P}^p(Y \times Z)$ is compact.*

Proof The assertion is established in the proof of Theorem 1 in Milgrom and Weber (1985, p. 626). □

Weak convergence of measures in $\mathscr{P}^p(Y \times Z)$ is equivalent to so-called *weak-strong (ws) convergence*. The weak-strong topology was introduced by Schäl (1975), and this paper utilizes results for this topology found in Balder (2001).

Lemma 2 *Suppose that Y and Z are compact metric spaces. Given $p \in \Delta(Y)$, a sequence (ν^n) in $\mathscr{P}^p(Y \times Z)$ is weakly convergent with limit point $\nu \in \mathscr{P}^p(Y \times Z)$ if and only if*

$$\int_{Y \times Z} f(y, z)\nu^n(d(y, z)) \to \int_{Y \times Z} f(y, z)\nu(d(y, z))$$

for every bounded $(\mathscr{B}(Y \times Z), \mathscr{B}(\mathbb{R}))$-measurable map $f : Y \times Z \to \mathbb{R}$ such that $f(y, \cdot) : Z \to \mathbb{R}$ is continuous for each $y \in Y$.

Proof Suppose that the sequence (ν^n) in $\mathscr{P}^p(Y \times Z)$ is weakly convergent with limit point $\nu \in \mathscr{P}^p(Y \times Z)$. Then the sequence $(\nu^n(\cdot \times Z))$ converges to $\nu(\cdot \times Z)$ in the s-topology (Definition 13), and so, applying Theorem 3.7(viii) in Schäl (1975), it follows that (ν^n) converges to ν in the ws-topology. Conversely, if (ν^n) ws-converges to ν in $\mathscr{P}^p(Y \times Z)$, then, by Theorem 3.7(viii) in Schäl (1975), it is clearly the case that $\nu^n \underset{w}{\to} \nu$. Thus, within $\mathscr{P}^p(Y \times Z)$, weak convergence of measures is equivalent to weak-strong convergence of measures. It only remains to observe that, by Theorem 3.1(b) in Balder (2001), $\nu^n \underset{ws}{\longrightarrow} \nu$ is equivalent to the following condition:

$$\int_{Y \times Z} f(y, z)\nu^n(d(y, z)) \to \int_{Y \times Z} f(y, z)\nu(d(y, z))$$

for every bounded $(\mathscr{B}(Y \times Z), \mathscr{B}(\mathbb{R}))$-measurable map $f : Y \times Z \to \mathbb{R}$ such that $f(y, \cdot) : Z \to \mathbb{R}$ is continuous for each $y \in Y$. □

The proofs of the following lemmas are relegated to "Appendix".

Lemma 3 *Suppose that* $\Gamma = (T_i, X_i, u_i, p)_{i=1}^N$ *is a Bayesian game in* $\mathfrak{G}$. *Suppose that* $(\mu, \alpha_i, \eta_i) \in \mathscr{M} \times \mathscr{A}_i \times \mathscr{D}_i$ *and*

$$\begin{aligned}&\int_T \int_X u_i(t,x)\mu(dx|t)p(dt)\\&\quad < \int_T \int_{T_i} \int_X \int_{X_i} u_i(t, y_i, x_{-i})\alpha_i(dy_i|t_i, x_i)\mu(dx|\tau_i, t_{-i})\eta_i(d\tau_i|t_i)p(dt).\end{aligned} \tag{18}$$

Then there exist $\alpha_i^* \in \mathscr{A}_i$ *and* $\eta_i^* \in \mathscr{D}_i$ *such that*

$$\begin{aligned}&\int_T \int_X u_i(t,x)\mu(dx|t)p(dt)\\&\quad < \int_T \int_{T_i} \int_X \int_{X_i} u_i(t, y_i, x_{-i})\alpha_i^*(dy_i|t_i, x_i)\mu(dx|\tau_i, t_{-i})\eta_i^*(d\tau_i|t_i)p(dt)\end{aligned} \tag{19}$$

and the following conditions are satisfied: η_i^* *is a simple function and there exists a* $(\mathscr{B}(T_i), \mathscr{B}(T_i))$*-measurable map* $g_i : T_i \to T_i$ *such that* $\eta_i^*(t_i) = \delta_{g_i(t_i)}$ *for each* $t_i \in T_i$;[15] *the function* $t_i \in T_i \mapsto \alpha_i^*(t_i, \cdot) \in \Delta(X_i)^{X_i}$ *is simple; and, for each* $t_i \in T_i$, *the map* $x_i \in X_i \mapsto \alpha_i^*(t_i, x_i) \in \Delta(X_i)$ *is continuous.*

Lemma 4 *Suppose that* $\Gamma = (T_i, X_i, u_i, p)_{i=1}^N$ *is a Bayesian game. Suppose that* (μ_i^n) *and* (ν_i^n) *are sequences in* $\mathscr{T}_i$. *Suppose that*

$$\varrho_{\Delta(X_i)}(\mu_i^n(t_i), \nu_i^n(t_i)) \to 0, \quad \textit{for every } t_i \in T_i. \tag{20}$$

Suppose further that (μ_{-i}^n) *is a sequence in* $\mathscr{T}_{-i}$. *Then, for every subsequence* (n_k) *of* (n),

$$\begin{aligned}&\varrho_{\Delta(X)}\left(\frac{1}{m}\sum_{k=1}^m \left[\mu_i^{n_k}(t_i) \otimes \left[\bigotimes_{j\neq i} \mu_j^{n_k}(t_j)\right]\right], \frac{1}{m}\sum_{k=1}^m \left[\nu_i^{n_k}(t_i) \otimes \left[\bigotimes_{j\neq i} \mu_j^{n_k}(t_j)\right]\right]\right) \xrightarrow{m\to\infty} 0,\\&\textit{for every } t \in T.\end{aligned} \tag{21}$$

We are now ready to prove Theorem 1.

Proof of Theorem 1 Fix a game $\Gamma = (T_i, X_i, u_i, p)_{i=1}^N$ in $\mathfrak{G}$. For each i, let $\mathscr{C}_i$ represent the set of all the continuous members of the function space $\Delta(X_i)^{T_i}$, and put $\mathscr{C} := \times_{i=1}^N \mathscr{C}_i$. The space $\mathscr{C}_i$, endowed with the topology of uniform convergence of functions, is separable (see, *e.g*., Aliprantis and Border 2006, Lemma 3.99), and so a countable dense subset $\mathscr{Q}_i$ may be selected from $\mathscr{C}_i$. The set $\mathscr{Q} := \times_{i=1}^N \mathscr{Q}_i$ is a countable set of strategies contained in $\mathscr{T} = \times_{i=1}^N \mathscr{T}_i$ and we claim that $\mathscr{Q}$ is a strategic approximation of Γ (in the sense of Definition 11).

[15] Recall that $\delta_{g_i(t_i)}$ denotes the Dirac measure in $\Delta(T_i)$ with support $\{g_i(t_i)\}$.

For each player i, let $(\mathcal{T}_i^\alpha)$ be an increasing net of finite subsets of $\mathcal{T}_i$ whose union contains $\mathcal{Q}_i$, *i.e.*, $\mathcal{T}_i^\alpha \subseteq \mathcal{T}_i^\beta$ whenever $\alpha \leq \beta$ and $\bigcup_\alpha \mathcal{T}_i^\alpha \supseteq \mathcal{Q}_i$. Suppose that for each α, $(\mu_1^\alpha, \ldots, \mu_N^\alpha)$ is a Nash equilibrium of the game

$$(\mathcal{T}_i^\alpha, U_i|_{\mathcal{T}_1^\alpha \times \cdots \times \mathcal{T}_N^\alpha})_{i=1}^N.$$

For each α, let $\mu^\alpha : T \to \Delta(X)$ be the correlated strategy in $\mathcal{M}$ defined by

$$\mu^\alpha(t) := \bigotimes_{i=1}^N \mu_i^\alpha(t_i).$$

Suppose that $[\mu^\alpha] \underset{\Upsilon}{\to} [\mu]$ for some $\mu \in \mathcal{M}$. We must show that μ is a communication equilibrium of Γ. To this end, we suppose that there exist i and $(\alpha_i, \eta_i) \in \mathcal{A}_i \times \mathcal{D}_i$ such that

$$\begin{aligned}&\int_{T\times X} u_i(t,x)[p\otimes\mu](d(t,x))\\ &\quad < \int_T \int_{T_i} \int_X \int_{X_i} u_i(t, y_i, x_{-i})\alpha_i(dy_i|t_i, x_i)\mu(dx|\tau_i, t_{-i})\eta_i(d\tau_i|t_i)p(dt)\end{aligned} \tag{22}$$

and derive a contradiction. □

The proof proceeds in a series of claims.

Claim A *There is no loss of generality in assuming that α_i and η_i satisfy the following: η_i is a simple function and there exists a $(\mathcal{B}(T_i), \mathcal{B}(T_i))$-measurable map $g_i : T_i \to T_i$ such that $\eta_i(t_i) = \delta_{g_i(t_i)}$ for each $t_i \in T_i$; the function $t_i \in T_i \mapsto \alpha_i(t_i, \cdot) \in \Delta(X_i)^{X_i}$ is simple; and, for each $t_i \in T_i$, the map $x_i \in X_i \mapsto \alpha_i(t_i, x_i) \in \Delta(X_i)$ is continuous.*

Proof of Claim A The assertion follows immediately from Lemma 3. □

Claim B *There are sequences $(\mathcal{T}_1^n, \ldots, \mathcal{T}_N^n)$ and $(\mu_1^n, \ldots, \mu_N^n)$ satisfying the following: for each j, $\mathcal{T}_j^1 \subseteq \mathcal{T}_j^2 \subseteq \cdots$ and $\bigcup_n \mathcal{T}_j^n \supseteq \mathcal{Q}_j$; for each j and n, $\mathcal{T}_j^n$ is a finite subset of $\mathcal{T}_j$ and $(\mu_1^n, \ldots, \mu_N^n)$ is a Nash equilibrium of the game $(\mathcal{T}_\iota^n, U_\iota|_{\mathcal{T}_1^n \times \cdots \times \mathcal{T}_N^n})_{\iota=1}^N$; and*

$$p \otimes \mu^n \underset{w}{\to} p \otimes \mu \quad \textit{and} \quad [p * g_i] \otimes \mu^n \underset{w}{\to} [p * g_i] \otimes \mu,$$

where $\mu^n : T \to \Delta(X)$ is the correlated strategy in $\mathcal{M}$ defined by $\mu^n(t) := \otimes_{\iota=1}^N \mu_\iota^n(t_\iota)$.

Proof of Claim B Endow $\mathcal{M}$ with the metric $d_{\mathcal{M}} : \mathcal{M} \times \mathcal{M} \to \mathbb{R}$ defined by

$$d_{\mathcal{M}}(\nu, \theta) := \max\left\{\varrho_{\Delta(X)}(p \otimes \nu, p \otimes \theta), \varrho_{\Delta(X)}([p * g_i] \otimes \nu, [p * g_i] \otimes \theta)\right\}.$$

More precisely, the metric space $(\mathcal{M}, d_{\mathcal{M}})$ is the space of all equivalence classes of members of $\mathcal{M}$ that are identical on a subset of T of full p-measure and a subset of

T of full $p * g_i$-measure, *i.e.*, the correlated strategies ν and θ in $\mathscr{M}$ are in the same equivalence class if there exist S and S' in $\mathscr{B}(T)$ with $p(S) = 1 = [p * g_i](S')$ such that $\nu(t) = \theta(t)$ for all $t \in S \cup S'$.

Because $[\mu^\alpha] \underset{\Upsilon}{\rightarrow} [\mu]$, it follows that

$$p \otimes \mu^\alpha \underset{w}{\rightarrow} p \otimes \mu \quad \text{and} \quad [p * g_i] \otimes \mu^\alpha \underset{w}{\rightarrow} [p * g_i] \otimes \mu.$$

Consequently, $d_{\mathscr{M}}(\mu^\alpha, \mu) \rightarrow 0$. For each j, let $\{q_j^1, q_j^2, \ldots\}$ be an enumeration of $\mathscr{Q}_j$. Note that there exist α_1 and α_{j1} ($j \in \{1, \ldots, N\}$) such that $\mu^\alpha \in N_1(\mu)$ for all $\alpha \geq \alpha_1$ and $q_j^1 \in \mathscr{T}_j^\alpha$ for all $\alpha \geq \alpha_{j1}$ and all j. Since there exists α_1^* with $\alpha_1^* \geq \alpha_1$ and $\alpha_1^* \geq \alpha_{j1}$ for all j, it follows that $\mu^\alpha \in N_1(\mu)$ and $q_j^1 \in \mathscr{T}_j^\alpha$ for all j and all $\alpha \geq \alpha_1^*$. Next, note that there exist α_2 and α_{j2} ($j \in \{1, \ldots, N\}$) such that $\mu^\alpha \in N_{\frac{1}{2}}(\mu)$ for all $\alpha \geq \alpha_2$ and $q_j^2 \in \mathscr{T}_j^\alpha$ for all $\alpha \geq \alpha_{j2}$ and all j. Since there exists α_2^* with $\alpha_2^* \geq \alpha_2$ and $\alpha_2^* \geq \alpha_{j2}$ for all j and $\alpha_2^* \geq \alpha_1^*$, it follows that $\mu^\alpha \in N_{\frac{1}{2}}(\mu)$ and $q_j^1, q_j^2 \in \mathscr{T}_j^\alpha$ for all j and all $\alpha \geq \alpha_2^*$. Proceeding inductively in this fashion gives a sequence (α_n^*) such that the sequences $(\mu_1^n, \ldots, \mu_N^n) := (\mu_1^{\alpha_n^*}, \ldots, \mu_N^{\alpha_n^*})$ and $(\mathscr{T}_1^n, \ldots, \mathscr{T}_N^n) := (\mathscr{T}_1^{\alpha_n^*}, \ldots, \mathscr{T}_N^{\alpha_n^*})$ have the desired properties, *i.e.*, for each j, $\mathscr{T}_j^1 \subseteq \mathscr{T}_j^2 \subseteq \cdots$ and $\bigcup_n \mathscr{T}_j^n \supseteq \mathscr{Q}_j$; for each j and n, $\mathscr{T}_j^n$ is a finite subset of $\mathscr{T}_j$ and $(\mu_1^n, \ldots, \mu_N^n)$ is a Nash equilibrium of the game $(\mathscr{T}_\iota^n, U_\iota|_{\mathscr{T}_1^n \times \cdots \times \mathscr{T}_N^n})_{\iota=1}^N$; and

$$p \otimes \mu^n \underset{w}{\rightarrow} p \otimes \mu \quad \text{and} \quad [p * g_i] \otimes \mu^n \underset{w}{\rightarrow} [p * g_i] \otimes \mu. \qquad \square$$

Define $\mu^* : T \rightarrow \Delta(X)$ by

$$\mu^*(B_i \times B_{-i}|t) := \int_{T_i} \int_{X_i \times B_{-i}} \alpha_i(B_i|t_i, x_i)\mu(dx|\tau_i, t_{-i})\eta_i(d\tau_i|t_i) \tag{23}$$

for all $B_i \times B_{-i} \subseteq X_i \times X_{-i}$ in $\mathscr{B}(X_i \times X_{-i})$.

Claim C *We have $p \otimes \rho^n \underset{w}{\rightarrow} p \otimes \mu^*$, where $\rho^n : T \rightarrow \Delta(X)$ is defined by*

$$\rho^n(t) := \rho_i^n(t_i) \otimes \left[\bigotimes_{j \neq i} \mu_j^n(t_j) \right] \tag{24}$$

and $\rho_i^n \in \mathscr{T}_i$ is defined by

$$\rho_i^n(B|t_i) := \int_{T_i} \int_{X_i} \alpha_i(B|t_i, x_i)\mu_i^n(dx_i|\tau_i)\eta_i(d\tau_i|t_i). \tag{25}$$

Proof of Claim C Define $\hat{\mu}^n : T \to \Delta(X)$ and $\hat{\mu} : T \to \Delta(X)$ by

$$\hat{\mu}^n(t) := \hat{\mu}_i^n(t_i) \otimes \left[\underset{j \neq i}{\otimes} \mu_j^n(t_j) \right] \quad \text{and} \quad \hat{\mu}(B|t) := \int_{T_i} \mu(B|\tau_i, t_{-i}) \eta_i(d\tau_i|t_i),$$

where $\hat{\mu}_i^n \in \mathscr{T}_i$ is defined by

$$\hat{\mu}_i^n(B|t_i) := \int_{T_i} \mu_i^n(B|\tau_i) \eta_i(d\tau_i|t_i).$$

Note that $p \otimes \hat{\mu}^n = [p * g_i] \otimes \mu^n$ and $p \otimes \hat{\mu} = [p * g_i] \otimes \mu$. Consequently, since $[p * g_i] \otimes \mu^n \underset{w}{\to} [p * g_i] \otimes \mu$ (Claim B), it follows that $p \otimes \hat{\mu}^n \underset{w}{\to} p \otimes \hat{\mu}$. Now Theorem 2.6 in Balder (2001) gives the following:

(I) Every subsequence of $(\hat{\mu}^n)$ has a further subsequence $(\hat{\mu}^{n_k})$ satisfying the following: for every subsequence $(\hat{\mu}^{n_{k_l}})$ of $(\hat{\mu}^{n_k})$ there is a p-null set $S \in \mathscr{B}(T)$ such that

$$\frac{1}{m} \sum_{l=1}^{m} \hat{\mu}^{n_{k_l}}(t) \xrightarrow[w]{m \to \infty} \hat{\mu}(t), \quad \text{for every } t \in T \setminus S. \tag{26}$$

It will now be shown that (I) implies the following:

(II) Every subsequence of (ρ^n) has a further subsequence (ρ^{n_k}) satisfying the following: for every subsequence $(\rho^{n_{k_l}})$ of (ρ^{n_k}) there is a p-null set $S' \in \mathscr{B}(T)$ such that

$$\frac{1}{m} \sum_{l=1}^{m} \rho^{n_{k_l}}(t) \xrightarrow[w]{m \to \infty} \mu^*(t), \quad \text{for every } t \in T \setminus S'$$

(recall the definition of μ^* given in (23)).

Given a subsequence of (n), there is, by virtue of (I), a subsequence $(\hat{\mu}^{n_k})$ such that, for a given subsequence $(\hat{\mu}^{n_{k_l}})$ of $(\hat{\mu}^{n_k})$, there is a p-null set $S \in \mathscr{B}(T)$ such that (26) holds. To establish (II), it suffices to show that

$$\frac{1}{m} \sum_{l=1}^{m} \rho^{n_{k_l}}(t) \xrightarrow[w]{m \to \infty} \mu^*(t), \quad \text{for every } t \in T \setminus S. \tag{27}$$

For each m, define $\nu^m : T \to \Delta(X)$ by

$$\nu^m(t) := \frac{1}{m} \sum_{l=1}^{m} \left[\overset{N}{\underset{i=1}{\otimes}} \hat{\mu}_i^{n_{k_l}}(t_i) \right].$$

Let $\alpha : T \times X \to \Delta(X)$ be defined by

$$\alpha(B_i \times B_{-i}|t, x) := \alpha_i(B_i|t_i, x_i) \otimes \left[\underset{j \neq i}{\otimes} \delta_{x_j}(B_j) \right].$$

For each $t \in T$, $\nu^m(t) \otimes \alpha(t, \cdot)$ is the measure in $\Delta(X \times X)$ defined by

$$[\nu^m(t) \otimes \alpha(t, \cdot)](A \times B) := \int_A \alpha(B|t, x)\nu^m(dx|t)$$

for all measurable rectangles $A \times B \subseteq X \times X$ in $\mathscr{B}(X \times X)$. Because (26) holds and, for each $t_i \in T_i$, the map $x_i \in X_i \mapsto \alpha_i(t_i, x_i) \in \Delta(X_i)$ is continuous (Claim A), so that, for each $t \in T$, the map $x \in X \mapsto \alpha(t, x) \in \Delta(X)$ is continuous, Theorem 4 in Kawabe (1994) gives

$$\nu^m(t) \otimes \alpha(t, \cdot) \underset{w}{\to} \hat{\mu}(t) \otimes \alpha(t, \cdot), \quad \text{for all } t \in T \setminus S, \tag{28}$$

where $\hat{\mu}(t) \otimes \alpha(t, \cdot)$ is defined analogously to $\nu^m(t) \otimes \alpha(t, \cdot)$.

Let $\sigma^m : T \to \Delta(X)$ and $\sigma : T \to \Delta(X)$ be defined by

$$\sigma^m(B|t) := [\nu^m(t) \otimes \alpha(t, \cdot)](X \times B) \quad \text{and} \quad \sigma(B|t) := [\hat{\mu}(t) \otimes \alpha(t, \cdot)](X \times B).$$

By Theorem 2.8(i) in Billingsley (1999), (28) implies that

$$\sigma^m(t) \underset{w}{\to} \sigma(t), \quad \text{for all } t \in T \setminus S. \tag{29}$$

Note that

$$\sigma^m = \frac{1}{m} \sum_{l=1}^{m} \rho^{n_{k_l}} \quad \text{and} \quad \sigma = \mu^* \tag{30}$$

(see the definitions of ρ^n and μ^* in (24) and (23), respectively). The second equality is straightforward. To see that the first equality holds, note that, for any measurable rectangle $B_i \times B_{-i} \subseteq X_i \times X_{-i}$ in $\mathscr{B}(X)$,

$$\begin{aligned}
&\sigma^m(B_i \times B_{-i}|t) = [\nu^m(t) \otimes \alpha(t, \cdot)](X \times (B_i \times B_{-i})) \\
&= \int_X \alpha(B_i \times B_{-i}|t, x)\nu^m(dx|t) \\
&= \int_X \left[\alpha_i(B_i|t_i, x_i) \otimes \left[\underset{j \neq i}{\otimes} \delta_{x_j}(B_j) \right] \right] \left[\frac{1}{m} \sum_{l=1}^{m} \left[\hat{\mu}_i^{n_{k_l}}(t_i) \otimes \left[\underset{j \neq i}{\otimes} \mu_j^{n_{k_l}}(t_j) \right] \right] \right](dx) \\
&= \int_{X_i \times B_{-i}} \alpha_i(B_i|t_i, x_i) \left[\frac{1}{m} \sum_{l=1}^{m} \left[\hat{\mu}_i^{n_{k_l}}(t_i) \otimes \left[\underset{j \neq i}{\otimes} \mu_j^{n_{k_l}}(t_j) \right] \right] \right](dx)
\end{aligned}$$

$$= \frac{1}{m}\sum_{l=1}^{m}\left[\int_{X_i \times B_{-i}} \alpha_i(B_i|t_i, x_i)\left[\hat{\mu}_i^{n_{k_l}}(t_i) \otimes \left[\underset{j\neq i}{\otimes} \mu_j^{n_{k_l}}(t_j)\right]\right](dx)\right]$$

$$= \frac{1}{m}\sum_{l=1}^{m}\left[\int_{X_i} \alpha_i(B_i|t_i, x_i)\hat{\mu}_i^{n_{k_l}}(dx_i|t_i)\left[\underset{j\neq i}{\otimes} \mu_j^{n_{k_l}}(t_j)\right](B_{-i})\right]$$

$$= \frac{1}{m}\sum_{l=1}^{m}\left[\rho_i^{n_{k_l}}(B_i|t_i)\left[\underset{j\neq i}{\otimes} \mu_j^{n_{k_l}}(t_j)\right](B_{-i})\right]$$

$$= \frac{1}{m}\sum_{l=1}^{m}\rho^{n_{k_l}}(B_i \times B_{-i}|t).$$

The desired convergence in (27) follows immediately from (29) and (30).

We conclude that (II) holds. Consequently, Theorem 2.6 in Balder (2001) implies that $p \otimes \rho^n \underset{w}{\rightarrow} p \otimes \mu^*$, as we sought. □

Claim D *There exists $\rho_i \in \mathscr{T}_i$ such that some subsequence of $(\hat{\rho}^n)$, denoted again by $(\hat{\rho}^n)$, satisfies $p \otimes \hat{\rho}^n \underset{w}{\rightarrow} p \otimes \mu^*$, where $\hat{\rho}^n : T \rightarrow \Delta(X)$ is defined by*

$$\hat{\rho}^n(t) := \rho_i(t_i) \otimes \left[\underset{j\neq i}{\otimes} \mu_j^n(t_j)\right].$$

Proof of Claim D Recall the definition of ρ^n in (24). Because $p \otimes \rho^n \underset{w}{\rightarrow} p \otimes \mu^*$ (Claim C), Theorem 2.6 in Balder (2001) gives the following:

(III) Every subsequence of (ρ^n) has a further subsequence (ρ^{n_k}) satisfying the following: for every subsequence $(\rho^{n_{k_l}})$ of (ρ^{n_k}) there is a p-null set $S^* \in \mathscr{B}(T)$ such that

$$\frac{1}{m}\sum_{l=1}^{m}\rho^{n_{k_l}}(t) \xrightarrow[w]{m\rightarrow\infty} \mu^*(t), \quad \text{for every } t \in T \setminus S^*.$$

Recall that the functions $\eta_i : T_i \rightarrow \Delta(T_i)$ and $t_i \in T_i \mapsto \alpha_i(t_i, \cdot) \in \Delta(X_i)^{X_i}$ are simple (Claim A). This implies that the behavioral strategy $\rho_i^n : T_i \rightarrow \Delta(X_i)$ defined in (25) is a simple function and there is a finite partition of T_i such that each ρ_i^n is constant on each partition element. Consequently, since $\Delta(X_i)$ is compact, there exists a subsequence of (ρ_i^n), denoted again by (ρ_i^n), that converges uniformly to some $\rho_i \in \mathscr{T}_i$. Hence

$$\varrho_{\Delta(X_i)}(\rho_i^n(t_i), \rho_i(t_i)) \rightarrow 0, \quad \text{for every } t_i \in T_i.$$

Applying Lemma 4 gives, for every subsequence (n_k) of (n),

$$\varrho_{\Delta(X)}\left(\frac{1}{m}\sum_{k=1}^{m}\left[\rho_i^{n_k}(t_i) \otimes \left[\underset{j\neq i}{\otimes} \mu_j^{n_k}(t_j)\right]\right], \frac{1}{m}\sum_{k=1}^{m}\left[\rho_i(t_i) \otimes \left[\underset{j\neq i}{\otimes} \mu_j^{n_k}(t_j)\right]\right]\right)$$

$$\xrightarrow{m\to\infty} 0, \quad \text{for every } t \in T.$$

This, together with (III), implies the following:

(IV) Every subsequence of $(\hat{\rho}^n)$ has a further subsequence $(\hat{\rho}^{n_k})$ satisfying the following: for every subsequence $(\hat{\rho}^{n_{k_l}})$ of $(\hat{\rho}^{n_k})$ there is a p-null set $S'' \in \mathscr{B}(T)$ such that

$$\frac{1}{m}\sum_{l=1}^{m} \hat{\rho}^{n_{k_l}}(t) \xrightarrow[w]{m\to\infty} \mu^*(t), \quad \text{for every } t \in T \setminus S''.$$

Given (IV), Theorem 2.6 in Balder (2001) implies that $p \otimes \hat{\rho}^n \underset{w}{\to} p \otimes \mu^*$. □

Claim E *There exists $\rho_i^* \in \mathscr{C}_i$ such that $p \otimes \tilde{\rho}^n \underset{w}{\to} p \otimes \mu^{**}$, where $\tilde{\rho}^n : T \to \Delta(X)$ is defined by*

$$\tilde{\rho}^n(t) := \rho_i^*(t_i) \otimes \left[\bigotimes_{j\neq i} \mu_j^n(t_j)\right], \tag{31}$$

and where μ^{**} satisfies

$$\int_{T\times X} u_i(t,x)[p\otimes\mu](d(t,x)) < \int_{T\times X} u_i(t,x)[p\otimes\mu^{**}](d(t,x)). \tag{32}$$

Proof of Claim E Because $p \otimes \hat{\rho}^n \underset{w}{\to} p \otimes \mu^*$ (Claim D), Theorem 2.6 in Balder (2001) gives the following:

(V) Every subsequence of $(\hat{\rho}^n)$ has a further subsequence $(\hat{\rho}^{n_k})$ satisfying the following: for every subsequence $(\hat{\rho}^{n_{k_l}})$ of $(\hat{\rho}^{n_k})$ there is a p-null set $S'' \in \mathscr{B}(T)$ such that

$$\frac{1}{m}\sum_{l=1}^{m} \hat{\rho}^{n_{k_l}}(t) \xrightarrow[w]{m\to\infty} \mu^*(t), \quad \text{for every } t \in T \setminus S''. \tag{33}$$

Next, we show that

$$\frac{1}{m}\sum_{n=1}^{m} \hat{\rho}^n(t) = \rho_i(t_i) \otimes \left(\frac{1}{m}\sum_{n=1}^{m}\left[\bigotimes_{j\neq i} \mu_j^n(t_j)\right]\right), \quad \text{for all } t \in T. \tag{34}$$

Fix $t \in T$ and a measurable rectangle $B_i \times B_{-i} \subseteq X_i \times X_{-i}$ in $\mathscr{B}(X)$. Then

$$\begin{aligned}\frac{1}{m}\sum_{n=1}^{m} \hat{\rho}^n(B_i \times B_{-i}|t) &= \frac{1}{m}\sum_{n=1}^{m}\left(\rho_i(B_i|t_i)\cdot\left[\bigotimes_{j\neq i} \mu_j^n(t_j)\right](B_{-i})\right)\\ &= \rho_i(B_i|t_i)\left(\frac{1}{m}\sum_{n=1}^{m}\left(\left[\bigotimes_{j\neq i} \mu_j^n(t_j)\right](B_{-i})\right)\right),\end{aligned}$$

implying (34).

In light of (34), (33) is expressible as

$$\rho_i(t_i) \otimes \left(\frac{1}{m} \sum_{l=1}^{m} \left[\underset{j \neq i}{\otimes} \mu_j^{n_{k_l}}(t_j) \right] \right) \xrightarrow[w]{m \to \infty} \mu^*(t), \quad \text{for every } t \in T \setminus S''.$$

Applying Theorem 2.8 in Billingsley (1999), this implies that

$$\rho_i(t_i) \otimes \left(\frac{1}{m} \sum_{l=1}^{m} \left[\underset{j \neq i}{\otimes} \mu_j^{n_{k_l}}(t_j) \right] \right) \xrightarrow[w]{m \to \infty} \rho_i(t) \otimes \mu_{-i}^*(t), \quad \text{for every } t \in T \setminus S'',$$

where $\mu_{-i}^*(t)$ denotes the marginal projection of $\mu^*(t)$ into $\Delta(X_{-i})$ (*i.e.*, $\mu_{-i}^*(t) \in \Delta(X_{-i})$ and $\mu_{-i}^*(B_{-i}|t) = \mu^*(X_i \times B_{-i}|t)$). Consequently, defining $\nu^* : T \to \Delta(X)$ by

$$\nu^*(t) := \rho_i(t_i) \otimes \mu_{-i}^*(t),$$

one obtains the following:

(VI) Every subsequence of $(\hat{\rho}^n)$ has a further subsequence $(\hat{\rho}^{n_k})$ satisfying the following: for every subsequence $(\hat{\rho}^{n_{k_l}})$ of $(\hat{\rho}^{n_k})$ there is a p-null set $S'' \in \mathscr{B}(T)$ such that

$$\begin{aligned} \frac{1}{m} \sum_{l=1}^{m} \hat{\rho}^{n_{k_l}}(t) &= \rho_i(t_i) \otimes \left(\frac{1}{m} \sum_{l=1}^{m} \left[\underset{j \neq i}{\otimes} \mu_j^{n_{k_l}}(t_j) \right] \right) \\ &\xrightarrow[w]{m \to \infty} \nu^*(t) = \rho_i(t_i) \otimes \mu_{-i}^*(t), \text{ for all } t \in T \setminus S''. \end{aligned}$$

By Theorem 2.6 in Balder (2001), this implies that $p \otimes \hat{\rho}^n \underset{w}{\to} p \otimes \nu^*$, and since $p \otimes \hat{\rho}^n \underset{w}{\to} p \otimes \mu^*$ (Claim D), it follows that $\mu^*(t) = \nu^*(t)$ for p-a.e. $t \in T$.

Now Luzin's Theorem gives a sequence (A_i^n) of compact subsets of T_i such that $p(A_i^n \times T_{-i}) \to 1$ and $\rho_i|_{A_i^n}$ is continuous for each n (here ρ_i is the measure given in Claim D). By Theorem 4.1 in Dugundji (1951), each $\rho_i|_{A_i^n}$ can be extended to a map $\theta_i^n \in \mathscr{C}_i$. Define $\theta^n : T \to \Delta(X)$ by

$$\theta^n(t) := \theta_i^n(t_i) \otimes \mu_{-i}^*(t),$$

and observe that

$$\begin{aligned} \int_{T \times X} u_i(t,x)[p \otimes \theta^n](d(t,x)) &= \int_{A_i^n \times T_{-i}} \int_X u_i(t,x) \left[\theta_i^n(t_i) \otimes \mu_{-i}^*(t)\right](dx) p(dt) \\ &+ \int_{[T_i \setminus A_i^n] \times T_{-i}} \int_X u_i(t,x) \left[\theta_i^n(t_i) \otimes \mu_{-i}^*(t)\right](dx) p(dt) \end{aligned}$$

$$
\begin{aligned}
&= \int_{A_i^n \times T_{-i}} \int_X u_i(t,x) \left[\theta_i^n(t_i) \otimes \mu_{-i}^*(t)\right](dx)p(dt) \\
&\quad + \int_{[T_i \setminus A_i^n] \times T_{-i}} \int_X u_i(t,x) \left[\theta_i^n(t_i) \otimes \mu_{-i}^*(t)\right](dx)p(dt) \\
&\quad + \int_{[T_i \setminus A_i^l] \times T_{-i}} \int_X u_i(t,x) \left[\rho_i(t_i) \otimes \mu_{-i}^*(t)\right](dx)p(dt) \\
&\quad - \int_{[T_i \setminus A_i^n] \times T_{-i}} \int_X u_i(t,x) \left[\rho_i(t_i) \otimes \mu_{-i}^*(t)\right](dx)p(dt) \\
&= \int_{T \times X} u_i(t,x)[p \otimes \nu^*](d(t,x)) \\
&\quad + \int_{[T_i \setminus A_i^n] \times T_{-i}} \int_X u_i(t,x) \left[\theta_i^n(t_i) \otimes \mu_{-i}^*(t)\right](dx)p(dt) \\
&\quad - \int_{[T_i \setminus A_i^n] \times T_{-i}} \int_X u_i(t,x) \left[\rho_i(t_i) \otimes \mu_{-i}^*(t)\right](dx)p(dt) \\
&\to \int_{T \times X} u_i(t,x)[p \otimes \nu^*](d(t,x)) \\
&= \int_{T \times X} u_i(t,x)[p \otimes \mu^*](d(t,x)).
\end{aligned}
$$

Consequently, in light of (22), it follows that (32) holds for μ^{**} defined by $\mu^{**} := \theta^{n^*}$ for some (sufficiently large) n^*.

Now let $\tilde{\rho}^n$ be defined as in (31), where $\rho_i^* := \theta_i^{n^*} \in \mathscr{C}_i$. Note that the proof will be complete if we show that $p \otimes \tilde{\rho}^n \underset{w}{\to} p \otimes \mu^{**}$. By Theorem 2.6 in Balder (2001), it suffices to show the following:

(VII) Every subsequence of $(\tilde{\rho}^n)$ has a further subsequence $(\tilde{\rho}^{n_k})$ satisfying the following: for every subsequence $(\tilde{\rho}^{n_{k_l}})$ of $(\tilde{\rho}^{n_k})$ there is a p-null set $\tilde{S} \in \mathscr{B}(T)$ such that

$$
\begin{aligned}
\frac{1}{m} \sum_{l=1}^{m} \tilde{\rho}^{n_{k_l}}(t) &= \rho_i^*(t_i) \otimes \left(\frac{1}{m} \sum_{l=1}^{m} \left[\bigotimes_{j \neq i} \mu_j^{n_{k_l}}(t_j) \right] \right) \\
&\xrightarrow[w]{m \to \infty} \mu^{**}(t) = \rho_i^*(t) \otimes \mu_{-i}^*(t), \ \text{ for all } t \in T \setminus \tilde{S}.
\end{aligned}
$$

But (VII) follows from (VI). Indeed, given a subsequence of (n), (VI) gives a further subsequence (n_k) such that, for every subsequence (n_{k_l}), there is a p-null set $S'' \in \mathscr{B}(T)$ such that

$$
\rho_i(t_i) \otimes \left(\frac{1}{m} \sum_{l=1}^{m} \left[\bigotimes_{j \neq i} \mu_j^{n_{k_l}}(t_j) \right] \right) \xrightarrow[w]{m \to \infty} \rho_i(t_i) \otimes \mu_{-i}^*(t), \ \text{ for all } t \in T \setminus S'',
$$

implying (by Theorem 2.8 in Billingsley (1999)) that

$$\rho_i^*(t_i) \otimes \left(\frac{1}{m} \sum_{l=1}^{m} \left[\bigotimes_{j \neq i} \mu_j^{n_{k_l}}(t_j) \right] \right) \xrightarrow[w]{m \to \infty} \rho_i^*(t) \otimes \mu_{-i}^*(t), \text{ for all } t \in T \setminus S''. \quad \square$$

Claim F *There exist n and $\rho_i^* \in \mathscr{T}_i^n$ such that $U_i(\rho_i^*, \mu_{-i}^n) > U_i(\mu_1^n, \dots, \mu_N^n)$.*

Proof of Claim F Let $\rho_i^* \in \mathscr{C}_i$ be the behavioral strategy given by Claim E. Recall that the sequence $(\mathscr{T}_1^n, \dots, \mathscr{T}_N^n)$ satisfies $\mathscr{T}_j^1 \subseteq \mathscr{T}_j^2 \subseteq \cdots$ and $\bigcup_n \mathscr{T}_j^n \supseteq \mathscr{Q}_j$ for each j (Claim B), and that each $\mathscr{Q}_j$ is a countable, dense subset of $\mathscr{T}_j$ (with respect to the topology of uniform convergence). Consequently, there exists a sequence (ν_i^n) with $\nu_i^n \in \mathscr{T}_i^n$ for each n such that (ν_i^n) converges to ρ_i^* uniformly. Therefore,

$$\varrho_{\Delta(X_i)}(\nu_i^n(t_i), \rho_i^*(t_i)) \to 0, \quad \text{for every } t_i \in T_i,$$

and so Lemma 4 gives, for every subsequence (n_k) of (n),

$$\varrho_{\Delta(X)} \left(\frac{1}{m} \sum_{k=1}^{m} \left[\nu_i^{n_k}(t_i) \otimes \left[\bigotimes_{j \neq i} \mu_j^{n_k}(t_j) \right] \right], \frac{1}{m} \sum_{k=1}^{m} \left[\rho_i^*(t_i) \otimes \left[\bigotimes_{j \neq i} \mu_j^{n_k}(t_j) \right] \right] \right) \xrightarrow{m \to \infty} 0, \tag{35}$$

for every $t \in T$.

Now since $p \otimes \tilde{\rho}^n \underset{w}{\rightarrow} p \otimes \mu^{**}$ (Claim E), Theorem 2.6 in Balder (2001) gives the following:

(VIII) Every subsequence of $(\tilde{\rho}^n)$ has a further subsequence $(\tilde{\rho}^{n_k})$ satisfying the following: for every subsequence $(\tilde{\rho}^{n_{k_l}})$ of $(\tilde{\rho}^{n_k})$ there is a p-null set $\tilde{S} \in \mathscr{B}(T)$ such that

$$\frac{1}{m} \sum_{l=1}^{m} \tilde{\rho}^{n_{k_l}}(t) \xrightarrow[w]{m \to \infty} \mu^{**}(t) = \rho_i^*(t) \otimes \mu_{-i}^*(t), \text{ for all } t \in T \setminus \tilde{S}.$$

Define $\nu^n : T \to \Delta(X)$ by

$$\nu^n(t) := \nu_i^n(t_i) \otimes \left[\bigotimes_{j \neq i} \mu_j^n(t_j) \right].$$

Combining (VIII) and (35) gives the following:

(IX) Every subsequence of (ν^n) has a further subsequence (ν^{n_k}) satisfying the following: for every subsequence $(\nu^{n_{k_l}})$ of (ν^{n_k}) there is a p-null set $\hat{S} \in \mathscr{B}(T)$ such that

$$\frac{1}{m} \sum_{l=1}^{m} \nu^{n_{k_l}}(t) \xrightarrow[w]{m \to \infty} \mu^{**}(t) = \rho_i^*(t) \otimes \mu_{-i}^*(t), \text{ for all } t \in T \setminus \hat{S}.$$

Again applying Theorem 2.6 in Balder (2001), we see that $p \otimes \nu^n \underset{w}{\to} p \otimes \mu^{**}$. Consequently, by Theorem 3.1 in Balder (2001),

$$\int_{T\times X} u_i(t,x)[p\otimes\nu^n](d(t,x)) \to \int_{T\times X} u_i(t,x)[p\otimes\mu^{**}](d(t,x)).$$

Similarly, because $p \otimes \mu^n \underset{w}{\to} p \otimes \mu$ (Claim B), one obtains

$$\int_{T\times X} u_i(t,x)[p\otimes\mu^n](d(t,x)) \to \int_{T\times X} u_i(t,x)[p\otimes\mu](d(t,x)).$$

Consequently (using (32)),

$$\begin{aligned}\int_{T\times X} u_i(t,x)[p\otimes\mu^n](d(t,x)) &\to \int_{T\times X} u_i(t,x)[p\otimes\mu](d(t,x))\\ &< \int_{T\times X} u_i(t,x)[p\otimes\mu^{**}](d(t,x))\\ &\leftarrow \int_{T\times X} u_i(t,x)[p\otimes\nu^n](d(t,x)),\end{aligned}$$

and so it follows that there exists n^{**} such that

$$\begin{aligned}U_i(\nu_i^{n^{**}}, \mu_{-i}^{n^{**}}) &= \int_{T\times X} u_i(t,x)[p\otimes\nu^{n^{**}}](d(t,x))\\ &> \int_{T\times X} u_i(t,x)[p\otimes\mu^{n^{**}}](d(t,x))\\ &= U_i(\mu_1^{n^{**}}, \ldots, \mu_N^{n^{**}}), \quad \text{for all } n \geq n^{**}.\end{aligned}$$

Since $\nu_i^{n^{**}} \in \mathscr{T}_i^{n^{**}}$, the proof is complete. □

Claim F gives the desired contradiction, since $(\mu_1^n, \ldots, \mu_N^n)$ is a Nash equilibrium of the game $(\mathscr{T}_\iota^n, U_\iota|_{\mathscr{T}_1^n\times\cdots\times\mathscr{T}_N^n})_{\iota=1}^N$ (see Claim B).

We have shown that the topology $\mathfrak{T}$ guarantees that every Bayesian game Γ in $\mathfrak{G}$ admits a strategic approximation. This finishes the proof of Theorem 1. □

Appendix

To begin, we establish an unproven claim made in Sect. 2.3.1.

Let $\Gamma = (T_i, X_i, u_i, p)_{i=1}^N$ be a Bayesian game, and suppose that the net (μ^α) converges uniformly to μ in $\mathscr{M}$. Fix $\hat{p} \in \boldsymbol{P}$ and a bounded continuous map $f : T \times X \to \mathbb{R}$. We claim that the net of maps

$$\left(t \in T \mapsto \int_X f(t,x)\mu^\alpha(dx|t)\right) \tag{36}$$

converges uniformly to the map $t \in T \mapsto \int_X f(t,x)\mu(dx|t)$.

Prior to proving this claim, we state and prove the following lemma.

Lemma 5 *Suppose that Y, Z, and E are metric spaces with Y and Z compact. If (h^α) is a uniformly convergent net of maps $h^\alpha : Y \to Z$ with limit point $h : Y \to Z$, and if $g : Y \times Z \to E$ is a continuous map, then the net $(y \in Y \mapsto g(y, h^\alpha(y)))$ converges uniformly to the map $y \in Y \mapsto g(y, h(y))$.*

Proof Fix $\epsilon > 0$. We must show that there exists α^* such that, for all $\alpha \geq \alpha^*$,

$$d_E(g(y, h^\alpha(y)), g(y, h(y))) < \epsilon, \quad \text{for all } y \in Y. \tag{37}$$

Because g is continuous and $Y \times Z$ is compact, g is uniformly continuous. Consequently, there exists $\delta > 0$ such that, for (y, z) and (y', z') in $Y \times Z$, $d_Y(y, y') < \delta$ and $d_Z(z, z') < \delta$ imply that $d_E(g(y, z), g(y', z')) < \epsilon$.

From the uniform convergence of the net (h^α) to h, one can choose α^* such that, for $\alpha \geq \alpha^*$, $d_Z(h^\alpha(y), h(y)) < \delta$ for all $y \in Y$. Consequently, for $\alpha \geq \alpha^*$, one obtains (37). □

To see that the net of maps in (36) converges uniformly to the map $t \in T \mapsto \int_X f(t,x)\mu(dx|t)$, we apply Lemma 5 with $Y = T$, $Z = \Delta(X)$, $E = \mathbb{R}$, $h^\alpha = \mu^\alpha$, $h = \mu$, and g defined by $g(t, \mu) := \int_X f(t,x)\mu(dx)$. The map g is continuous. Indeed, suppose that (t^n, μ^n) is a convergent sequence in $T \times \Delta(X)$ with limit point $(t, \mu) \in T \times \Delta(X)$. For each n, let δ_{t^n} denote the Dirac measure in $\Delta(T)$ with support $\{t^n\}$. Because the sequence (t^n, μ^n) converges to (t, μ), the sequence (δ_{t^n}, μ^n) converges to (δ_t, μ) in $\Delta(T) \times \Delta(X)$ (see, *e.g.*, Aliprantis and Border 2006, Theorem 15.8). Consequently, by Theorem 2.8(ii) in Billingsley (1999), $\delta_{t^n} \otimes \mu^n \underset{w}{\to} \delta_t \otimes \mu$, and so the Portmanteau Theorem (see, *e.g.*, Aliprantis and Border 2006, Theorem 15.3) yields

$$\begin{aligned} g(t^n, \mu^n) &= \int_{T \times X} f(\tau, x)[\delta_{t^n} \otimes \mu^n](d(\tau, x)) \\ &\to \int_{T \times X} f(\tau, x)[\delta_t \otimes \mu](d(\tau, x)) = g(t, \mu). \end{aligned}$$

Since (μ^α) converges uniformly to μ, and since g is continuous, Lemma 5 implies that the net of maps in (36) converges uniformly to the map $t \in T \mapsto \int_X f(t,x)\mu(dx|t)$, as we sought.

The remainder of this appendix contains the proofs of the lemmas stated in Sect. 6. For the convenience of the reader, each proof is preceded by a restatement of its corresponding lemma.

Proof of Lemma 4

Lemma 3 *Suppose that* $\Gamma = (T_i, X_i, u_i, p)_{i=1}^N$ *is a Bayesian game in* $\mathfrak{G}$. *Suppose that* $(\mu, \alpha_i, \eta_i) \in \mathscr{M} \times \mathscr{A}_i \times \mathscr{D}_i$ *and*

$$\begin{aligned}&\int_T \int_X u_i(t,x)\mu(dx|t)p(dt)\\ &\quad< \int_T \int_{T_i} \int_X \int_{X_i} u_i(t, y_i, x_{-i})\alpha_i(dy_i|t_i, x_i)\mu(dx|\tau_i, t_{-i})\eta_i(d\tau_i|t_i)p(dt).\end{aligned} \tag{18}$$

Then there exist $\alpha_i^* \in \mathscr{A}_i$ *and* $\eta_i^* \in \mathscr{D}_i$ *such that*

$$\begin{aligned}&\int_T \int_X u_i(t,x)\mu(dx|t)p(dt)\\ &\quad< \int_T \int_{T_i} \int_X \int_{X_i} u_i(t, y_i, x_{-i})\alpha_i^*(dy_i|t_i, x_i)\mu(dx|\tau_i, t_{-i})\eta_i^*(d\tau_i|t_i)p(dt)\end{aligned} \tag{19}$$

and the following conditions are satisfied: η_i^* *is a simple function and there exists a* $(\mathscr{B}(T_i), \mathscr{B}(T_i))$*-measurable map* $g_i : T_i \to T_i$ *such that* $\eta_i^*(t_i) = \delta_{g_i(t_i)}$ *for each* $t_i \in T_i$;[16] *the function* $t_i \in T_i \mapsto \alpha_i^*(t_i, \cdot) \in \Delta(X_i)^{X_i}$ *is simple; and, for each* $t_i \in T_i$, *the map* $x_i \in X_i \mapsto \alpha_i^*(t_i, x_i) \in \Delta(X_i)$ *is continuous.*

Proof The proof is organized in four steps.

Step 1 *There is no loss of generality in assuming that* α_i *is continuous.*

Proof of Step 1 Define $p' \in \Delta(T)$ by

$$p'(A_i \times A_{-i}) := \int_{T_i \times A_{-i}} \eta_i(A_i|t_i)p(dt)$$

for all $A_i \times A_{-i} \subseteq T_i \times T_{-i}$ in $\mathscr{B}(T)$, and define $\rho \in \Delta(T \times X)$ by

$$\rho(A \times B) := \int_A \mu(B|t)p'(dt)$$

for all $A \times B \subseteq T \times X$ in $\mathscr{B}(T \times X)$. By Luzin's Theorem, there is a sequence (K^n) of compact subsets of $T_i \times X_i$ such that $\rho(K^n \times T_{-i} \times X_{-i}) \to 1$ and $\alpha_i|_{K^n}$ is continuous for each n. Applying Theorem 4.1 of Dugundji (1951), each $\alpha_i|_{K^n}$ can be extended to a continuous map $\widehat{\alpha}_i^n : T_i \times X_i \to \Delta(X_i)$. Define $\vartheta^n : T \times X \to \mathbb{R}$ and $\vartheta : T \times X \to \mathbb{R}$ by

$$\vartheta^n(t,x) := \int_{X_i} u_i(t, y_i, x_{-i})\widehat{\alpha}_i^n(dy_i|t_i, x_i)$$

[16] Recall that $\delta_{g_i(t_i)}$ denotes the Dirac measure in $\Delta(T_i)$ with support $\{g_i(t_i)\}$.

and

$$\vartheta(t,x) := \int_{X_i} u_i(t, y_i, x_{-i})\alpha_i(dy_i|t_i, x_i).$$

Because $\vartheta^n = \vartheta$ on $K^n \times T_{-i} \times X_{-i}$, for each n, and since $\rho(K^n \times T_{-i} \times X_{-i}) \to 1$, it follows that

$$\begin{aligned}
&\int_{T\times X} \vartheta^n(t,x)\rho(d(t,x)) \\
&= \int_T \int_{T_i} \int_X \int_{X_i} u_i(t, y_i, x_{-i})\widehat{\alpha}_i^n(dy_i|t_i, x_i)\mu(dx|\tau_i, t_{-i})\eta_i(d\tau_i|t_i)p(dt) \\
&\to \int_T \int_{T_i} \int_X \int_{X_i} u_i(t, y_i, x_{-i})\alpha_i(dy_i|t_i, x_i)\mu(dx|\tau_i, t_{-i})\eta_i(d\tau_i|t_i)p(dt) \\
&= \int_{T\times X} \vartheta(t,x)\rho(d(t,x)),
\end{aligned}$$

and in light of (18) we conclude that

$$\begin{aligned}
&\int_T \int_X u_i(t,x)\mu(dx|t)p(dt) \\
&\quad < \int_T \int_{T_i} \int_X \int_{X_i} u_i(t, y_i, x_{-i})\alpha_i'(dy_i|t_i, x_i)\mu(dx|\tau_i, t_{-i})\eta_i(d\tau_i|t_i)p(dt)
\end{aligned}$$

for some continuous $\alpha_i' \in \mathscr{A}_i$. □

Step 2 *There exists $\alpha_i^* \in \mathscr{A}_i$ such that*

$$\begin{aligned}
&\int_T \int_X u_i(t,x)\mu(dx|t)p(dt) \\
&\quad < \int_T \int_{T_i} \int_X \int_{X_i} u_i(t, y_i, x_{-i})\alpha_i^*(dy_i|t_i, x_i)\mu(dx|\tau_i, t_{-i})\eta_i(d\tau_i|t_i)p(dt)
\end{aligned} \tag{38}$$

and the following conditions are satisfied: the function $t_i \in T_i \mapsto \alpha_i^*(t_i, \cdot) \in \Delta(X_i)^{X_i}$ is simple, and, for each $t_i \in T_i$, the map $x_i \in X_i \mapsto \alpha_i^*(t_i, x_i) \in \Delta(X_i)$ is continuous.

Proof of Step 2 Let $\mathscr{C}(X_i, \Delta(X_i))$ represent the set of all the continuous functions from X_i into $\Delta(X_i)$, and endow the space $\mathscr{C}(X_i, \Delta(X_i))$ with the supremum metric. Then $\mathscr{C}(X_i, \Delta(X_i))$ is a separable metric space (see, *e.g.*, Aliprantis and Border 2006, Lemma 3.99). Define $\widetilde{\boldsymbol{\alpha}}_i : T_i \to \mathscr{C}(X_i, \Delta(X_i))$ by

$$[\widetilde{\boldsymbol{\alpha}}_i(t_i)](x_i) := \alpha_i(t_i, x_i)$$

(recall that α_i can be taken continuous by Step 1). Because α_i is continuous, Theorem 4.55 in Aliprantis and Border (2006) implies that the map $\widetilde{\boldsymbol{\alpha}}_i : T_i \to \mathscr{C}(X_i, \Delta(X_i))$ is $(\mathscr{B}(T_i), \mathscr{B}(\mathscr{C}(X_i, \Delta(X_i))))$-measurable. Consequently, applying Theorem 4.38 in

Aliprantis and Border (2006), it follows that $\widetilde{\boldsymbol{\alpha}}_i$ is the pointwise limit of a sequence $(\widetilde{\boldsymbol{\alpha}}_i^n)$ of $(\mathscr{B}(T_i), \mathscr{B}(\mathscr{C}(X_i, \Delta(X_i))))$-measurable simple functions. Now, for each n, define $\alpha_i^n : T_i \times X_i \to \Delta(X_i)$ by

$$\alpha_i^n(t_i, x_i) := [\widetilde{\boldsymbol{\alpha}}_i^n(t_i)](x_i).$$

Note that it suffices to show that there exists n for which $\alpha_i^* := \alpha_i^n$ satisfies (38).

Applying Theorem 4.55 in Aliprantis and Border (2006), we see that (α_i^n) is a sequence of $(\mathscr{B}(T_i \times X_i), \mathscr{B}(\Delta(X_i)))$-measurable functions. We claim that (α_i^n) converges to α_i pointwise. To see this, fix $(t_i, x_i) \in T_i \times X_i$. It must be shown that $\alpha_i^n(t_i, x_i) \underset{w}{\to} \alpha_i(t_i, x_i)$. We know that the sequence $(\widetilde{\boldsymbol{\alpha}}_i^n)$ converges to $\widetilde{\boldsymbol{\alpha}}_i$ pointwise. Consequently, the sequence $(\widetilde{\boldsymbol{\alpha}}_i^n(t_i))$ of maps in $\mathscr{C}(X_i, \Delta(X_i))$ converges uniformly to $\widetilde{\boldsymbol{\alpha}}_i(t_i) \in \mathscr{C}(X_i, \Delta(X_i))$, *i.e.*, for each $\epsilon > 0$, there exists M such that, for all $n \geq M$ and $x_i \in X_i$, we have

$$\alpha_i^n(t_i, x_i) = [\widetilde{\boldsymbol{\alpha}}_i^n(t_i)](x_i) \in N_\epsilon([\widetilde{\boldsymbol{\alpha}}_i(t_i)](x_i)) = N_\epsilon(\alpha_i(t_i, x_i)),$$

implying that $\alpha_i^n(t_i, x_i) \underset{w}{\to} \alpha_i(t_i, x_i)$.

Next, define $\theta : T \times X \to \Delta(X)$ and $\theta^n : T \times X \to \Delta(X)$ by

$$\theta(B_i \times B_{-i}|t, x) := \alpha_i(B_i|t_i, x_i) \otimes \left[\bigotimes_{j \neq i} \delta_{x_j}(B_j)\right] \text{ and}$$

$$\theta^n(B_i \times B_{-i}|t, x) := \alpha_i^n(B_i|t_i, x_i) \otimes \left[\bigotimes_{j \neq i} \delta_{x_j}(B_j)\right]$$

for all $B_i \times B_{-i} \subseteq X_i \times X_{-i}$ in $\mathscr{B}(X)$, where δ_{x_j} denotes the Dirac measure in $\Delta(X_j)$ with support $\{x_j\}$. Define $\eta : T \to \Delta(T)$ by

$$\eta(B_i \times B_{-i}|t) := \eta_i(B_i|t_i) \otimes \left[\bigotimes_{j \neq i} \delta_{t_j}(B_j)\right]$$

for all $B_i \times B_{-i} \subseteq T_i \times T_{-i}$ in $\mathscr{B}(T)$, and $\mu^* : T \to \Delta(X)$ by

$$\mu^*(B|t) := \int_T \mu(B|\tau)\eta(d\tau|t).$$

Because the sequence (α_i^n) converges to α_i pointwise, it follows from Theorem 2.8(ii) in Billingsley (1999) that the sequence (θ^n) converges to θ pointwise. Consequently, applying Theorem 2.6 in Balder (2001), it follows that

$$p \otimes \mu^* \otimes \theta^n \underset{w}{\to} p \otimes \mu^* \otimes \theta,$$

where $p \otimes \mu^* \otimes \theta \in \Delta(T \times X \times X)$ is defined by

$$[p \otimes \mu^* \otimes \theta](A \times B \times B') := \int_{A \times B} \theta(B'|(t, x))[p \otimes \mu^*](d(t, x))$$

for all $A \times B \times B' \subseteq T \times X \times X$ in $\mathscr{B}(T \times X \times X)$, and each $p \otimes \mu^* \otimes \theta^n$ is defined similarly.

Let $\nu \in \Delta(T \times X)$ (resp. $\nu^n \in \Delta(T \times X)$) be defined by $\nu(A \times B) := [p \otimes \mu^* \otimes \theta](A \times X \times B)$ (resp. $\nu^n(A \times B) := [p \otimes \mu^* \otimes \theta^n](A \times X \times B)$). By Theorem 2.8(i) in Billingsley (1999), $\nu^n \underset{w}{\to} \nu$. Therefore, since ν^n and ν are members of $\mathscr{P}^{p'}(T \times X)$, Lemma 2 gives

$$\begin{aligned}
&\int_T \int_{T_i} \int_X \int_{X_i} u_i(t, y_i, x_{-i}) \alpha_i^n(dy_i|t_i, x_i) \mu(dx|\tau_i, t_{-i}) \eta_i(d\tau_i|t_i) p(dt) \\
&\quad = \int_{T \times X} u_i(t, x) \nu^n(d(t, x)) \\
&\quad \to \int_{T \times X} u_i(t, x) \nu(d(t, x)) \\
&\quad = \int_T \int_{T_i} \int_X \int_{X_i} u_i(t, y_i, x_{-i}) \alpha_i(dy_i|t_i, x_i) \mu(dx|\tau_i, t_{-i}) \eta_i(d\tau_i|t_i) p(dt).
\end{aligned}$$

This, together with (18), gives (38). □

Next, observe that

$$\begin{aligned}
&\int_T \int_{T_i} \int_X \int_{X_i} u_i(t, y_i, x_{-i}) \alpha_i^*(dy_i|t_i, x_i) \mu(dx|\tau_i, t_{-i}) \eta_i(d\tau_i|t_i) p(dt) \\
&\quad = \int_{T_i} \int_{T_{-i}} \int_{T_i} \int_X \int_{X_i} u_i(t, y_i, x_{-i}) \alpha_i^*(dy_i|t_i, x_i) \\
&\quad \times \mu(dx|\tau_i, t_{-i}) \eta_i(d\tau_i|t_i) p(dt_{-i}|t_i) p_i(dt_i) \\
&\quad = \int_{T_i} \int_{T_i} \int_{T_{-i}} \int_X \int_{X_i} u_i(t, y_i, x_{-i}) \alpha_i^*(dy_i|t_i, x_i) \\
&\quad \times \mu(dx|\tau_i, t_{-i}) p(dt_{-i}|t_i) \eta_i(d\tau_i|t_i) p_i(dt_i),
\end{aligned}$$

where p_i represents the marginal projection of p into $\Delta(T_i)$. Define $\zeta : T_i \times T_i \to \mathbb{R}$ by

$$\zeta(t_i, \tau_i) := \int_{T_{-i}} \int_X \int_{X_i} u_i(t, y_i, x_{-i}) \alpha_i^*(dy_i|t_i, x_i) \mu(dx|\tau_i, t_{-i}) p(dt_{-i}|t_i). \quad (39)$$

Step 3 *There exists a $(\mathscr{B}(T_i), \mathscr{B}(T_i))$-measurable map $g : T_i \to T_i$ such that*

$$\int_T \int_X u_i(t, x) \mu(dx|t) p(dt) < \int_{T_i} \zeta(t_i, g(t_i)) p_i(dt_i).$$

Proof of Step 3 Because the map ζ is $(\mathscr{B}(T_i \times T_i), \mathscr{B}(\mathbb{R}))$-measurable, Theorem 2 in Brown and Purves (1973) gives, for each n, a $(\mathscr{B}(T_i), \mathscr{B}(T_i))$-measurable map

$g^n : T_i \to T_i$ such that for every $t_i \in T_i$,

$$\zeta(t_i, g^n(t_i)) \geq \sup_{\tau_i \in T_i} \zeta(t_i, \tau_i) - \frac{1}{n}.$$

This, together with the fact that

$$\sup_{\tau_i \in T_i} \zeta(t_i, \tau_i) - \frac{1}{n} \geq \int_{T_i} \zeta(t_i, \tau_i)\eta_i(d\tau_i|t_i) - \frac{1}{n}, \quad \text{for all } t_i \in T_i,$$

gives

$$\zeta(t_i, g^n(t_i)) \geq \int_{T_i} \zeta(t_i, \tau_i)\eta_i(d\tau_i|t_i) - \frac{1}{n}, \quad \text{for all } t_i \in T_i,$$

Consequently,

$$\int_{T_i} \zeta(t_i, g^n(t_i))p_i(dt_i) \geq \int_{T_i}\int_{T_i} \zeta(t_i, \tau_i)\eta_i(d\tau_i|t_i)p_i(dt_i) - \frac{1}{n}.$$

Because

$$\int_{T_i}\int_{T_i} \zeta(t_i, \tau_i)\eta_i(d\tau_i|t_i)p_i(dt_i) > \int_T\int_X u_i(t, x)\mu(dx|t)p(dt)$$

(Step 2), it follows that there exists a large enough n for which

$$\int_{T_i} \zeta(t_i, g^n(t_i))p_i(dt_i) > \int_T\int_X u_i(t, x)\mu(dx|t)p(dt).$$

□

Step 4 *There exists a simple map $\eta_i^* \in \mathscr{D}_i$ such that*

$$\int_T\int_X u_i(t, x)\mu(dx|t)p(dt) < \int_{T_i}\int_{T_i} \zeta(t_i, \tau_i)\eta_i^*(d\tau_i|t_i)p_i(dt_i).$$

Proof of Step 4 Define $\lambda \in \Delta(T_i \times T_i)$ by

$$\lambda(A \times B) := \int_A \delta_{g(t_i)}(B)p_i(dt_i) \tag{40}$$

for all $A \times B \subseteq T_i \times T_i$ in $\mathscr{B}(T_i \times T_i)$, where g is the map from Step 3. Because the map ζ defined in (39) is $(\mathscr{B}(T_i \times T_i), \mathscr{B}(\mathbb{R}))$-measurable, Luzin's Theorem gives a sequence (S^n) of compact subsets of $T_i \times T_i$ such that $\lambda(S^n) \to 1$ and each $\zeta^n := \zeta|_{S^n}$ is continuous. Since S^n is compact, $\zeta^n : S^n \to \mathbb{R}$ is uniformly continuous, and so there exists $\delta_n > 0$ such that $d_i(t_i, \hat{t}_i) < \delta_n$ and $d_i(\tau_i, \hat{\tau}_i) < \delta_n$ implies that

$|\zeta^n(t_i, \tau_i) - \zeta^n(\hat{t}_i, \hat{\tau}_i)| < \frac{1}{n}$, where d_i is a compatible metric for T_i. In addition, since S^n is compact, there exists a finite $\frac{\delta_n}{2}$-partition $\{P^{(n,1)}, \ldots, P^{(n,k_n)}\}$ of S^n (*i.e.*, a partition such that each $P^{(n,k)}$ has radius less than $\frac{\delta_n}{2}$) consisting of sets in $\mathscr{B}(T_i \times T_i)$.

For each n and $k \in \{1, \ldots, k_n\}$, let

$$P_1^{(n,k)} := \left\{ t_i \in T_i : \exists \tau_i : (t_i, \tau_i) \in P^{(n,k)} \right\}. \tag{41}$$

Step 4.1 *The partition $\{P^{(n,1)}, \ldots, P^{(n,k_n)}\}$ of S^n can be chosen to satisfy the following: if $(t_i^k, \tau_i^k) \in P^{(n,k)}$ and $(t_i^\kappa, \tau_i^\kappa) \in P^{(n,\kappa)}$ for $\kappa \neq k$, and if $\tilde{t}_i^k \in P_1^{(n,k)}$ and $\tilde{t}_i^\kappa \in P_1^{(n,\kappa)}$, then $(\tilde{t}_i^k, \tau_i^k) \neq (\tilde{t}_i^\kappa, \tau_i^\kappa)$.*

Proof of Step 4.1 Note that there exists a finite set $\{(t_i^1, \tau_i^1), \ldots, (t_i^{k_n}, \tau_i^{k_n})\} \subseteq S^n$ such that

$$S^n \subseteq \bigcup_{k=1}^{k_n} \left(N_{\delta_n/2}(t_i^k) \times N_{\delta_n/2}(\tau_i^k) \right).$$

(Here the $\delta_n/2$-neighborhoods are neighborhoods in T_i.) Now define $A^{(n,1)}, \ldots, A^{(n,k_n)}$ as follows:

- $A^{(n,1)} := N_{\delta_n/2}(t_i^1) \times N_{\delta_n/2}(\tau_i^1)$;
- $A^{(n,2)} := \left(N_{\delta_n/2}(t_i^2) \times N_{\delta_n/2}(\tau_i^2)\right) \setminus \left(N_{\delta_n/2}(t_i^1) \times N_{\delta_n/2}(\tau_i^1)\right)$;
- $A^{(n,3)} := \left(N_{\delta_n/2}(t_i^3) \times N_{\delta_n/2}(\tau_i^3)\right) \setminus \left[\left(N_{\delta_n/2}(t_i^1) \times N_{\delta_n/2}(\tau_i^1)\right) \cup \left(N_{\delta_n/2}(t_i^2) \times N_{\delta_n/2}(\tau_i^2)\right)\right]$; and so on.

Letting $P^{(n,k)} := A^{(n,k)} \cap S^n$, one obtains a $\delta_n/2$-partition $\{P^{(n,1)}, \ldots, P^{(n,k_n)}\}$ of S^n. To see that this partition has the desired property, fix $(\hat{t}_i^k, \hat{\tau}_i^k) \in P^{(n,k)}$ and $(\hat{t}_i^\kappa, \hat{\tau}_i^\kappa) \in P^{(n,\kappa)}$ for $\kappa > k$, and choose $\tilde{t}_i^k \in P_1^{(n,k)}$ and $\tilde{t}_i^\kappa \in P_1^{(n,\kappa)}$. Then the construction of the partition entails that if $\tilde{t}_i^k = \tilde{t}_i^\kappa$ then $\hat{\tau}_i^k \neq \hat{\tau}_i^\kappa$. □

For each n and $k \in \{1, \ldots, k_n\}$, let

$$\hat{P}^{(n,k)} := \left\{ (t_i, \tau_i) \in P^{(n,k)} : \tau_i = g(t_i) \right\} \quad \text{and}$$
$$\hat{P}_1^{(n,k)} := \left\{ t_i \in T_i : \exists \tau_i : (t_i, \tau_i) \in \hat{P}^{(n,k)} \right\}. \tag{42}$$

Step 4.2 *For each n and $k \in \{1, \ldots, k_n\}$, the set $\hat{P}^{(n,k)}$ belongs to $\mathscr{B}(T_i \times T_i)$.*

Proof of Step 4.2 Because the map g from Step 3 is $(\mathscr{B}(T_i), \mathscr{B}(T_i))$-measurable, the graph of g,

$$\mathrm{Gr}(g) := \{(t_i, g(t_i)) : t_i \in T_i\} \subseteq T_i \times T_i,$$

belongs to $\mathscr{B}(T_i \times T_i)$ (see, *e.g.*, Aliprantis and Border 2006, Theorem 12.28). Consequently, since $P^{(n,k)} \in \mathscr{B}(T_i \times T_i)$ and

$$\hat{P}^{(n,k)} = \mathrm{Gr}(g) \cap P^{(n,k)},$$

it follows that $\hat{P}^{(n,k)} \in \mathscr{B}(T_i \times T_i)$. $\square$

Step 4.3 *For each n and $k \in \{1, \ldots, k_n\}$, the set $\hat{P}_1^{(n,k)}$ belongs to $\mathscr{B}(T_i)$.*

Proof of Step 4.3 The assertion follows from Theorem 18.10 in Kechris (1995), together with the facts that $\hat{P}^{(n,k)} \in \mathscr{B}(T_i \times T_i)$ (Step 4.2) and that, for each $t_i \in T_i$, the set $\{\tau_i : (t_i, \tau_i) \in \hat{P}^{(n,k)}\}$ is finite (in fact, a singleton). $\square$

For each n and $k \in \{1, \ldots, k_n\}$, choose $t_i^{(n,k)} \in \hat{P}_1^{(n,k)}$ such that

$$\zeta^n(t_i^{(n,k)}, g(t_i^{(n,k)})) \geq \sup_{t_i \in \hat{P}_1^{(n,k)}} \zeta^n(t_i, g(t_i)) - \frac{1}{n}$$

and $t_i^* \in T_i$, and define $f^n : T_i \to T_i$ by

$$f^n(t_i) := \begin{cases} g(t_i^{(n,k)}) & \text{if there exists } k \text{ such that } t_i \in \hat{P}_1^{(n,k)}, \\ t_i^* & \text{otherwise.} \end{cases} \tag{43}$$

Step 4.4 *The map $f^n : T_i \to T_i$ defined in (43) is $(\mathscr{B}(T_i), \mathscr{B}(T_i))$-measurable.*

Proof of Step 4.4 The assertion follows from the following facts: f^n has finite range and (by Step 4.3) the inverse images of the members of the range belong to $\mathscr{B}(T_i)$. $\square$

Step 4.5 *For every $t_i \in \hat{P}_1^{(n,k)}$,*

$$\zeta^n(t_i, f^n(t_i)) \geq \zeta^n(t_i, g(t_i)) - \frac{2}{n}.$$

Proof of Step 4.5 First, recall that $d_i(t_i, \hat{t}_i) < \delta_n$ and $d_i(\tau_i, \hat{\tau}_i) < \delta_n$ implies that

$$|\zeta^n(t_i, \tau_i) - \zeta^n(\hat{t}_i, \hat{\tau}_i)| < \frac{1}{n}.$$

Now, given $t_i \in \hat{P}_1^{(n,k)}$, one has

$$\begin{aligned} \zeta^n(t_i, f^n(t_i)) = \zeta^n(t_i, g(t_i^{(n,k)})) &\geq \zeta^n(t_i^{(n,k)}, g(t_i^{(n,k)})) - \frac{1}{n} \\ &\geq \sup_{\tau_i \in \hat{P}_1^{(n,k)}} \zeta^n(\tau_i, g(\tau_i)) - \frac{2}{n} \geq \zeta^n(t_i, g(t_i)) - \frac{2}{n}. \quad \square \end{aligned}$$

Next, define $\lambda^n \in \Delta(T_i \times T_i)$ by

$$\lambda^n(A \times B) := \int_A \delta_{f^n(t_i)}(B) p_i(dt_i) \tag{44}$$

for all $A \times B \subseteq T_i \times T_i$ in $\mathscr{B}(T_i \times T_i)$, where $\delta_{f^n(t_i)} \in \Delta(T_i)$ denotes the Dirac measure on T_i with support $\{f^n(t_i)\}$.

Now recall the definition of $\hat{P}^{(k,n)}$ in (42) and define

$$\tilde{P}^{(n,k)} := \mathrm{Gr}(f^n) \cap \left(\hat{P}_1^{(n,k)} \times T_i\right) \quad \text{and} \quad \tilde{P}_1^{(n,k)} := \left\{t_i \in T_i : \exists \tau_i : (t_i, \tau_i) \in \tilde{P}^{(n,k)}\right\},$$

where $\mathrm{Gr}(f^n)$ denotes the graph of f^n in $T_i \times T_i$.

Step 4.6 *For each n and $k \in \{1, \ldots, k_n\}$, the set $\tilde{P}^{(n,k)}$ belongs to $\mathscr{B}(T_i \times T_i)$.*

Proof of Step 4.6 Analogous to the proof of Step 4.2. □

Step 4.7 *For each n and $k \in \{1, \ldots, k_n\}$, the set $\tilde{P}_1^{(n,k)}$ belongs to $\mathscr{B}(T_i)$.*

Proof of Step 4.7 Analogous to the proof of Step 4.3. □

Step 4.8 *For each n and $k \neq \kappa$, $\tilde{P}^{(n,k)} \cap \tilde{P}^{(n,\kappa)} = \emptyset$.*

Proof of Step 4.8 Choose $(t_i, \tau_i) \in \tilde{P}^{(n,k)}$ and $(\hat{t}_i, \hat{\tau}_i) \in \tilde{P}^{(n,\kappa)}$. It suffices to show that $(t_i, \tau_i) \neq (\hat{t}_i, \hat{\tau}_i)$.

Because $(t_i, \tau_i) \in \tilde{P}^{(n,k)}$, one has $(t_i, \tau_i) = (t_i, f^n(t_i))$ and $t_i \in \hat{P}_1^{(n,k)}$, implying (by (43)) that $(t_i, \tau_i) = (t_i, g(t_i^{(n,k)}))$. In addition, since $t_i \in \hat{P}_1^{(n,k)}$, one has $t_i \in P_1^{(n,k)}$ (recall the definition of $P_1^{(n,k)}$ in (41)). Summarizing, one has $t_i \in P_1^{(n,k)}$ and $(t_i, \tau_i) = (t_i, g(t_i^{(n,k)}))$.

Similarly, one can show that $(\hat{t}_i, \hat{\tau}_i) = (\hat{t}_i, g(t_i^{(n,\kappa)}))$ and $\hat{t}_i \in P_1^{(n,\kappa)}$.

Since $(t_i^{(n,k)}, g(t_i^{(n,k)})) \in P^{(n,k)}$, $(t_i^{(n,\kappa)}, g(t_i^{(n,\kappa)})) \in P^{(n,\kappa)}$, $t_i \in P_1^{(n,k)}$, and $\hat{t}_i \in P_1^{(n,\kappa)}$, it follows from Step 4.1 that $(t_i, \tau_i) = (t_i, g(t_i^{(n,k)})) \neq (\hat{t}_i, g(t_i^{(n,\kappa)})) = (\hat{t}_i, \hat{\tau}_i)$.
□

Step 4.9 *For each n and $k \in \{1, \ldots, k_n\}$, $\hat{P}_1^{(n,k)} = \tilde{P}_1^{(n,k)}$.*

Proof of Step 4.9 Suppose that $t_i \in \hat{P}_1^{(n,k)}$. Then $(t_i, f^n(t_i)) \in \tilde{P}^{(n,k)}$, and so $t_i \in \tilde{P}_1^{(n,k)}$. Hence, $\hat{P}_1^{(n,k)} \subseteq \tilde{P}_1^{(n,k)}$. Conversely, suppose that $t_i \in \tilde{P}_1^{(n,k)}$. Then $(t_i, \tau_i) \in \tilde{P}^{(n,k)}$ for some τ_i, implying that $t_i \in \hat{P}_1^{(n,k)}$, and so $\hat{P}_1^{(n,k)} \supseteq \tilde{P}_1^{(n,k)}$. □

Step 4.10 *For each n and $k \neq \kappa$, $\hat{P}_1^{(n,k)} \cap \hat{P}_1^{(n,\kappa)} = \emptyset$.*

Proof of Step 4.10 Suppose that $t_i \in \hat{P}_1^{(n,k)} \cap \hat{P}_1^{(n,\kappa)}$. Then $(t_i, \tau_i) \in \hat{P}^{(n,k)} \subseteq P^{(n,k)}$ and $(t_i, \hat{\tau}_i) \in \hat{P}^{(n,\kappa)} \subseteq P^{(n,\kappa)}$ for some τ_i and $\hat{\tau}_i$, and so $\tau_i = g(t_i) = \hat{\tau}_i$. Hence, $(t_i, \tau_i) = (t_i, \hat{\tau}_i) \in P^{(n,k)} \cap P^{(n,\kappa)}$, a contradiction. □

Step 4.11 *For each n,*

$$\lambda^n \left(\bigcup_{k=1}^{k_n} \tilde{P}^{(n,k)}\right) = \lambda(S^n).$$

Proof of Step 4.11 Fix n. Then,

$$\lambda^n\left(\bigcup_{k=1}^{k_n}\tilde{P}^{(n,k)}\right)=\sum_{k=1}^{k_n}\lambda^n(\tilde{P}^{(k,n)})=\sum_{k=1}^{k_n}p_i(\tilde{P}_1^{(k,n)})=\sum_{k=1}^{k_n}p_i(\hat{P}_1^{(k,n)})$$
$$=\sum_{k=1}^{k_n}\lambda(\hat{P}^{(k,n)})=\sum_{k=1}^{k_n}\lambda(P^{(k,n)})=\lambda(S^n),$$

where the first equality follows from Step 4.8; the second equality uses Step 4.9 and Step 4.10; and the third equality uses Step 4.9; □

Step 4.12 *We have*

$$\liminf_n\int_{T_i}\zeta(t_i,f^n(t_i))p_i(dt_i)\geq\int_{T_i}\zeta(t_i,g(t_i))p_i(dt_i). \tag{45}$$

Proof of Step 4.12 We have

$$\begin{aligned}
&\int_{T_i}\zeta(t_i,f^n(t_i))p_i(dt_i)\\
&\quad=\int_{T_i\times T_i}\zeta(t_i,\tau_i)\lambda^n(d(t_i,\tau_i))\\
&\quad=\sum_{k=1}^{k_n}\int_{\tilde{P}^{(k,n)}}\zeta^n(t_i,\tau_i)\lambda^n(d(t_i,\tau_i))+\int_{(T_i\times T_i)\setminus\bigcup_k\tilde{P}^{(k,n)}}\zeta(t_i,\tau_i)\lambda^n(d(t_i,\tau_i))\\
&\quad=\sum_{k=1}^{k_n}\int_{\tilde{P}_1^{(k,n)}}\zeta^n(t_i,f^n(t_i))p_i(dt_i)+\int_{(T_i\times T_i)\setminus\bigcup_k\tilde{P}^{(k,n)}}\zeta(t_i,\tau_i)\lambda^n(d(t_i,\tau_i))\\
&\quad=\sum_{k=1}^{k_n}\int_{\hat{P}_1^{(k,n)}}\zeta^n(t_i,f^n(t_i))p_i(dt_i)+\int_{(T_i\times T_i)\setminus\bigcup_k\tilde{P}^{(k,n)}}\zeta(t_i,\tau_i)\lambda^n(d(t_i,\tau_i))\\
&\quad\geq\sum_{k=1}^{k_n}\int_{\hat{P}_1^{(k,n)}}\left[\zeta^n(t_i,g(t_i))-\frac{2}{n}\right]p_i(dt_i)+\int_{(T_i\times T_i)\setminus\bigcup_k\tilde{P}^{(k,n)}}\zeta(t_i,\tau_i)\lambda^n(d(t_i,\tau_i))\\
&\quad=\sum_{k=1}^{k_n}\int_{\hat{P}_1^{(k,n)}}\zeta^n(t_i,g(t_i))p_i(dt_i)-\frac{2}{n}\sum_{k=1}^{k_n}p_i(\hat{P}_1^{(n,k)})\\
&\qquad+\int_{(T_i\times T_i)\setminus\bigcup_k\tilde{P}^{(k,n)}}\zeta(t_i,\tau_i)\lambda^n(d(t_i,\tau_i))\\
&\quad=\sum_{k=1}^{k_n}\int_{\hat{P}^{(k,n)}}\zeta(t_i,\tau_i)\lambda(d(t_i,\tau_i))-\frac{2}{n}\sum_{k=1}^{k_n}\lambda(\hat{P}^{(n,k)})\\
&\qquad+\int_{(T_i\times T_i)\setminus\bigcup_k\tilde{P}^{(k,n)}}\zeta(t_i,\tau_i)\lambda^n(d(t_i,\tau_i))
\end{aligned}$$

$$
\begin{aligned}
&= \sum_{k=1}^{k_n} \int_{P^{(k,n)}} \zeta(t_i, \tau_i)\lambda(d(t_i, \tau_i)) - \frac{2}{n} \sum_{k=1}^{k_n} \lambda(P^{(n,k)}) \\
&\quad + \int_{(T_i \times T_i) \setminus \bigcup_k \tilde{P}^{(k,n)}} \zeta(t_i, \tau_i)\lambda^n(d(t_i, \tau_i)) \\
&= \int_{S^n} \zeta(t_i, \tau_i)\lambda(d(t_i, \tau_i)) - \frac{2}{n}\lambda(S^n) \\
&\quad + \int_{(T_i \times T_i) \setminus \bigcup_k \tilde{P}^{(k,n)}} \zeta(t_i, \tau_i)\lambda^n(d(t_i, \tau_i)) \\
&= \int_{S^n} \zeta(t_i, \tau_i)\lambda(d(t_i, \tau_i)) - \frac{2}{n}\lambda(S^n) \\
&\quad + \int_{(T_i \times T_i) \setminus \bigcup_k \tilde{P}^{(k,n)}} \zeta(t_i, \tau_i)\lambda^n(d(t_i, \tau_i)) \\
&\quad + \int_{(T_i \times T_i) \setminus S^n} \zeta(t_i, \tau_i)\lambda(d(t_i, \tau_i)) - \int_{(T_i \times T_i) \setminus S^n} \zeta(t_i, \tau_i)\lambda(d(t_i, \tau_i)) \\
&= \int_{T_i \times T_i} \zeta(t_i, \tau_i)\lambda(d(t_i, \tau_i)) - \frac{2}{n}\lambda(S^n) + \int_{(T_i \times T_i) \setminus \bigcup_k \tilde{P}^{(k,n)}} \zeta(t_i, \tau_i)\lambda^n(d(t_i, \tau_i)) \\
&\quad - \int_{(T_i \times T_i) \setminus S^n} \zeta(t_i, \tau_i)\lambda(d(t_i, \tau_i)) \\
&= \int_{T_i} \zeta(t_i, g(t_i))p_i(dt_i) - \frac{2}{n}\lambda(S^n) + \int_{(T_i \times T_i) \setminus \bigcup_k \tilde{P}^{(k,n)}} \zeta(t_i, \tau_i)\lambda^n(d(t_i, \tau_i)) \\
&\quad - \int_{(T_i \times T_i) \setminus S^n} \zeta(t_i, \tau_i)\lambda(d(t_i, \tau_i)) \\
&\xrightarrow{n \to \infty} \int_{T_i} \zeta(t_i, g(t_i))p_i(dt_i),
\end{aligned}
$$

implying (45). Here, the first equality uses the definition of λ^n in (44); the second equality uses the definition of ζ^n from the first paragraph of the proof of Step 4, Step 4.6, and Step 4.8; the third equality uses Step 4.7, Step 4.9, and Step 4.10; the fourth equality follows from Step 4.9; the inequality follows from Step 4.5; the sixth and seventh equalities follow from Step 4.2 and the definition of λ in (40); and the limit at the end follows from Step 4.11, together with the boundedness of ζ and the fact that $\lambda(S^n) \to 1$. □

Step 4.13 *There exists n^* such that*

$$
\int_T \int_X u_i(t, x)\mu(dx|t)p(dt) < \int_{T_i} \zeta(t_i, f^{n^*}(t_i))p_i(dt_i). \tag{46}
$$

Proof of Step 4.13 The assertion follows immediately from Step 3 and Step 4.12. □

Letting $\eta_i^* \in \mathscr{D}_i$ be defined by

$$
\eta_i^*(B|t_i) := \delta_{f^{n^*}(t_i)}(B),
$$

where $\delta_{f^{n^*}(t_i)}$ represents the Dirac measure in $\Delta(T_i)$ with support $f^{n^*}(t_i)$, and where n^* is the natural number from Step 4.13, it follows from (46) that

$$\int_T \int_X u_i(t,x)\mu(dx|t)p(dt) < \int_{T_i} \zeta(t_i, f^{n^*}(t_i))p_i(dt_i)$$
$$= \int_{T_i}\int_{T_i} \zeta(t_i,\tau_i)\eta_i^*(d\tau_i|t_i)p_i(dt_i).$$

This finishes the proof of Step 4. □

Combining Step 4 and the fact that

$$\int_{T_i}\int_{T_i} \zeta(t_i,\tau_i)\eta_i^*(d\tau_i|t_i)p_i(dt_i)$$
$$= \int_T \int_{T_i}\int_X \int_{X_i} u_i(t,y_i,x_{-i})\alpha_i^*(dy_i|t_i,x_i)\mu(dx|\tau_i,t_{-i})\eta_i^*(d\tau_i|t_i)p(dt)$$

yields (19). This finishes the proof of Lemma. □

Proof of Lemma 5

Lemma 4 *Suppose that $\Gamma = (T_i, X_i, u_i, p)_{i=1}^N$ is a Bayesian game. Suppose that (μ_i^n) and (ν_i^n) are sequences in $\mathscr{T}_i$. Suppose that*

$$\varrho_{\Delta(X_i)}(\mu_i^n(t_i), \nu_i^n(t_i)) \to 0, \quad \textit{for every } t_i \in T_i. \tag{20}$$

Suppose further that (μ_{-i}^n) is a sequence in $\mathscr{T}_{-i}$. Then, for every subsequence (n_k) of (n),

$$\varrho_{\Delta(X)}\left(\frac{1}{m}\sum_{k=1}^m \left[\mu_i^{n_k}(t_i)\otimes\left[\bigotimes_{j\neq i}\mu_j^{n_k}(t_j)\right]\right], \frac{1}{m}\sum_{k=1}^m \left[\nu_i^{n_k}(t_i)\otimes\left[\bigotimes_{j\neq i}\mu_j^{n_k}(t_j)\right]\right]\right)$$
$$\xrightarrow{m\to\infty} 0, \quad \textit{for every } t \in T. \tag{21}$$

Proof Suppose that (21) does not hold for some subsequence (n_k) of (n). Then, for some $t \in T$, and extracting a subsequence if necessary,

$$\varrho_{\Delta(X)}\left(\frac{1}{m}\sum_{k=1}^m \left[\mu_i^{n_k}(t_i)\otimes\left[\bigotimes_{j\neq i}\mu_j^{n_k}(t_j)\right]\right],\right.$$
$$\left.\frac{1}{m}\sum_{k=1}^m \left[\nu_i^{n_k}(t_i)\otimes\left[\bigotimes_{j\neq i}\mu_j^{n_k}(t_j)\right]\right]\right) \xrightarrow{m\to\infty} \gamma$$

for some $\gamma > 0$. Therefore, there exist $\gamma' > 0$ and M such that for each $m \geq M$,

$$\varrho_{\Delta(X)}\left(\frac{1}{m}\sum_{k=1}^{m}\left[\mu_i^{n_k}(t_i)\otimes\left[\underset{j\neq i}{\otimes}\mu_j^{n_k}(t_j)\right]\right], \frac{1}{m}\sum_{k=1}^{m}\left[\nu_i^{n_k}(t_i)\otimes\left[\underset{j\neq i}{\otimes}\mu_j^{n_k}(t_j)\right]\right]\right) \geq \gamma',$$

i.e. (recall (2)),

$$\begin{aligned}\inf\Bigg\{\epsilon : \forall\text{closed } B \subseteq X, \frac{1}{m}\sum_{k=1}^{m}\left[\int_{X_{-i}}\mu_i^{n_k}((B)_{x_{-i}}|t_i)\left[\underset{j\neq i}{\otimes}\mu_j^{n_k}(t_j)\right](dx_{-i})\right]\\ \leq \epsilon + \frac{1}{m}\sum_{k=1}^{m}\left[\int_{X_{-i}}\nu_i^{n_k}((N_\epsilon(B))_{x_{-i}}|t_i)\left[\underset{j\neq i}{\otimes}\mu_j^{n_k}(t_j)\right](dx_{-i})\right]\Bigg\} \geq \gamma',\end{aligned} \tag{47}$$

where $(B)_{x_{-i}}$ (resp. $(N_\epsilon(B))_{x_{-i}}$) denotes the section of B (resp. $N_\epsilon(B)$) in X_i at x_{-i}.[17] Now (20) implies the following:

$$\varrho_{\Delta(X_i)}(\mu_i^{n_k}(t_i), \nu_i^{n_k}(t_i)) \xrightarrow{k\to\infty} 0.$$

Therefore,

$$\inf\left\{\epsilon : \forall\text{closed } B \subseteq X_i, \mu_i^{n_k}(B|t_i) \leq \nu_i^{n_k}(N_\epsilon(B)|t_i) + \epsilon\right\} \xrightarrow{k\to\infty} 0,$$

implying that there exist $\gamma'' \in (0, \gamma')$ and K such that for each $k \geq K$, and for each B closed in X_i,

$$\mu_i^{n_k}(B|t_i) \leq \gamma'' + \nu_i^{n_k}(N_{\gamma''}(B)|t_i).$$

Consequently, for each $k \geq K$, and for each B closed in X and $x_{-i} \in X_{-i}$,[18]

$$\mu_i^{n_k}((B)_{x_{-i}}|t_i) \leq \gamma'' + \nu_i^{n_k}(N_{\gamma''}((B)_{x_{-i}})|t_i) \leq \gamma'' + \nu_i^{n_k}((N_{\gamma''}(B))_{x_{-i}}|t_i), \tag{48}$$

implying that for each $k \geq K$ and B closed in X,

$$\begin{aligned}\int_{X_{-i}}\mu_i^{n_k}((B)_{x_{-i}}|t_i)\left[\underset{j\neq i}{\otimes}\mu_j^{n_k}(t_j)\right](dx_{-i}) &\leq \gamma''\\ + \int_{X_{-i}}\nu_i^{n_k}((N_{\gamma''}(B))_{x_{-i}}|t_i)\left[\underset{j\neq i}{\otimes}\mu_j^{n_k}(t_j)\right](dx_{-i}).\end{aligned}$$

[17] The section of a closed (resp. open) subset of a product space is closed (resp. open) (see, *e.g.*, Bourbaki (1989, p. 46, Corollary)).

[18] The last inequality in (48) follows from the fact that $N_{\gamma''}((B)_{x_{-i}}) \subseteq (N_{\gamma''}(B))_{x_{-i}}$. To see that this containment holds, suppose that $x_i \in N_{\gamma''}((B)_{x_{-i}})$. Then $x_i \in N_{\gamma''}(y_i)$ for some $y_i \in (B)_{x_{-i}}$, implying that $(y_i, x_{-i}) \in B$. Since $x_i \in N_{\gamma''}(y_i)$, it follows that $(x_i, x_{-i}) \in N_{\gamma''}(B)$ and so $x_i \in (N_{\gamma''}(B))_{x_{-i}}$.

Consequently, there is an M' such that for each $m \geq M'$ and each B closed in X,

$$\frac{1}{m}\sum_{k=1}^{m}\left[\int_{X_{-i}} \mu_i^{n_k}((B)_{x_{-i}}|t_i)\left[\underset{j\neq i}{\otimes}\mu_j^{n_k}(t_j)\right](dx_{-i})\right]$$
$$\leq \gamma''' + \frac{1}{m}\sum_{k=1}^{m}\left[\int_{X_{-i}} \nu_i^{n_k}((N_{\gamma''}(B))_{x_{-i}}|t_i)\left[\underset{j\neq i}{\otimes}\mu_j^{n_k}(t_j)\right](dx_{-i})\right],$$

for some $\gamma''' \in (0, \gamma')$. But this implies that for $m \geq \max\{M, M'\}$, the left-hand side of (47) must be strictly less than γ', a contradiction. □

References

Aliprantis CD, Border KC (2006) Infinite dimensional analysis: A Hitchhiker's guide. Springer, Berlin. https://doi.org/10.1007/3-540-29587-9

Athey S (2001) Single crossing properties and the existence of pure strategy equilibria in games of incomplete information. Econometrica 69:861–889. https://doi.org/10.1111/1468-0262.00223

Aumann RJ (1974) Subjectivity and correlation in randomized strategies. J Math Econ 1:67–96. https://doi.org/10.1016/0304-4068(74)90037-8

Balder EJ (1988) Generalized equilibrium results for games with incomplete information. Math Oper Res 13:265–276. https://doi.org/10.1287/moor.13.2.265

Balder EJ (2001) On ws-convergence of product measures. Math Oper Res 26:494–518. https://doi.org/10.1287/moor.26.3.494.10581

Bergemann D, Morris S (2016) Bayes correlated equilibrium and the comparison of information structures in games. Theoret Econ 11:487–522. https://doi.org/10.3982/te1808

Billingsley P (1999) Convergence of probability measures: Wiley-Interscience. https://doi.org/10.1002/9780470316962

Nicolas B (1989) General topology, vol 1–4. Springer, Berlin. https://doi.org/10.1007/978-3-642-61701-010.1007/978-3-642-61701-0

Brown LD, Purves RA (1973) Measurable selections of extrema. Ann Stat 1:902–912. https://doi.org/10.1214/aos/1176342510

Carbonell-Nicolau O, McLean RP (2018) On the existence of Nash equilibrium in Bayesian games. Math Oper Res 43:100–129. https://doi.org/10.1287/moor.2017.0849

Carbonell-Nicolau O, McLean RP (2019) Nash and Bayes-Nash equilibria in strategic-form games with intransitivities. Econ Theor 68:935–965. https://doi.org/10.1007/s00199-018-1151-7

Carbonell-Nicolau O, McLean RP (2020) On the existence of equilibrium in Bayesian games with correlated types, Mimeograph

Charles C, de Fitte PR, Valadier M (2004) Young measures on topological spaces. Springer, Netherlands. https://doi.org/10.1007/1-4020-1964-5

Cotter KD (1991) Correlated equilibrium in games with type-dependent strategies. J Econ Theory 54:48–68. https://doi.org/10.1016/0022-0531(91)90104-c

Dudley RM (1968) Distances of probability measures and random variables. Ann Math Stat 39:1563–1572. https://doi.org/10.1214/aoms/1177698137

Dugundji J (1951) An extension of Tietze's theorem. Pac J Math 1:353–367. https://doi.org/10.2140/pjm.1951.1.353

Dunford N, Schwartz JT (1958) Linear operators: general theory. Pure and applied mathematics. Interscience Publishers, New York

Forges F (1986) An approach to communication equilibria. Econometrica 54:1375. https://doi.org/10.2307/1914304

Françoise F (1990) Universal mechanisms. Econometrica 58:1341. https://doi.org/10.2307/2938319

Forges F (1993) Five legitimate definitions of correlated equilibrium in games with incomplete information. Theor Decis 35:277–310. https://doi.org/10.1007/bf01075202

Gürkan G, Pang J-S (2007) Approximations of Nash equilibria. Math Program 117:223–253. https://doi.org/10.1007/s10107-007-0156-y
He W, Sun Y (2019) Pure-strategy equilibria in Bayesian games. J Econ Theory 180:11–49. https://doi.org/10.1016/j.jet.2018.11.007
He W, Yannelis NC (2016) Existence of equilibria in discontinuous Bayesian games. J Econ Theory 162:181–194. https://doi.org/10.1016/j.jet.2015.12.009
Hellman Z, Levy YJ (2017) Bayesian games with a continuum of states. Theoret Econ 12:1089–1120. https://doi.org/10.3982/te1544
Kawabe J (1994) A criterion for weak compactness of measures on product spaces with applications. Yokohama Math J 42:159–169
Kechris AS (1995) Classical descriptive set theory. Springer, New York. https://doi.org/10.1007/978-1-4612-4190-4
Lucchetti R, Patrone F (1986) Closure and upper semicontinuity results in mathematical programming. Nash and economic equilibria. Optimization 17:619–628. https://doi.org/10.1080/02331938608843178
McAdams D (2003) Isotone equilibrium in games of incomplete information. Econometrica 71:1191–1214. https://doi.org/10.1111/1468-0262.00443
Megginson Robert E (1998) An introduction to Banach space theory. Springer, New York. https://doi.org/10.1007/978-1-4612-0603-3
Milgrom PR, Weber RJ (1985) Distributional strategies for games with incomplete information. Math Oper Res 10:619–632. https://doi.org/10.1287/moor.10.4.619
Myerson RB (1982) Optimal coordination mechanisms in generalized principal–agent problems. J Math Econ 10:67–81. https://doi.org/10.1016/0304-4068(82)90006-4
Myerson RB (1991) Game theory: analysis of conflict. Harvard University Press, Cambridge
Prokhorov YV (1956) Convergence of random processes and limit theorems in probability theory. Theory Prob Appl 1:157–214. https://doi.org/10.1137/1101016
Prokopovych P, Yannelis NC (2019) On monotone approximate and exact equilibria of an asymmetric first-price auction with affiliated private information. J Econ Theory 184:1–29. https://doi.org/10.1016/j.jet.2019.07.012
Reny PJ (2011a) On the existence of monotone pure-strategy equilibria in Bayesian games. Econometrica 79:499–553. https://doi.org/10.3982/ecta8934
Reny PJ (2011b) Strategic approximations of discontinuous games. Econ Theor 48:17–29. https://doi.org/10.1007/s00199-010-0518-1
Schäl M (1975) On dynamic programming: compactness of the space of policies. Stoch Processes Appli 3:345–364. https://doi.org/10.1016/0304-4149(75)90031-9
Simon LK (1987) Games with discontinuous payoffs. Rev Econ Stud 54:569. https://doi.org/10.2307/2297483
Simon RS (2003) Games of incomplete information, ergodic theory, and the measurability of equilibria. Isr J Math 138:73–92. https://doi.org/10.1007/bf02783420
Stinchcombe MB (2005) Nash equilibrium and generalized integration for infinite normal form games. Games Econ Behav 50:332–365. https://doi.org/10.1016/j.geb.2004.09.007
Stinchcombe MB (2011a) Balance and discontinuities in infinite games with type-dependent strategies. J Econ Theory 146:656–671. https://doi.org/10.1016/j.jet.2010.12.009
Stinchcombe MB (2011b) Correlated equilibrium existence for infinite games with type-dependent strategies. J Econ Theory 146:638–655. https://doi.org/10.1016/j.jet.2010.12.006
Yannelis NC, Aldo R (1991) Equilibrium points of non-cooperative random and Bayesian games. Positive operators, Riesz spaces, and economics. Springer, Berlin, pp 23–48. https://doi.org/10.1007/978-3-642-58199-1_2