Bayesian games with nested information*

Royi Jacobovic[†] Yehuda John Levy[‡] Eilon Solan[§]

August 13, 2025

Abstract

A Bayesian game is said to have nested information if the players are ordered, and each player knows the types of all players that follow her in that order. We prove that all multiplayer Bayesian games with finite actions spaces, bounded payoffs, Polish type spaces, and nested information admit a Bayesian equilibrium.

Keywords: Bayesian games, nested information, Bayesian equilibrium, selection theorems.

JEL Classification: C72.

1 Introduction

Although models of incomplete information are abundant in economic modeling, general results on existence of equilibrium are hard to come by. In particular, while often

*We thank three anonymous referees for their valuable input, which helped us significantly improve the paper. We also thank Raffi Hassin, Ziv Hellman, Ehud Lehrer, Dov Samet, and Omri Solan for useful discussions. Jacobovic acknowledges the support of the Israel Science Foundation, grant #3739/24. Solan acknowledges the support of the Israel Science Foundation, grant #211/22.

†School of Mathematical Sciences, Tel-Aviv University, Tel-Aviv, Israel, 6997800, E-mail: royi-jacobo@tauex.tau.ac.il.

‡Adam Smith Business School, University of Glasgow, 2 Discovery Place, Glasgow, G11 6EY, UK, E-mail: john.levy@glasgow.ac.uk

§School of Mathematical Sciences, Tel-Aviv University, Tel-Aviv, Israel, 6997800, E-mail: eilons@tauex.tau.ac.il.

it is natural to model the possible rewards and beliefs of agents using a continuum of possibilities, there are relatively few general existence results at our disposal, and they use fairly restrictive assumptions on the players' information. In this paper, we examine games under a naturally arising information structure, namely, *nested information*, where the players can be ordered according to the amount of information they possess, from the most knowledgeable to the least knowledgeable.

Harsanyi [20] laid the foundation of games of incomplete information, also known as Bayesian games, which have greatly influenced the development of game theory. In that model, each agent has a *type*, which includes her belief about payoffs, others' beliefs about the payoffs, others' beliefs about others' beliefs about the payoffs, and so forth. There is a prior over the possible type profiles that may occur. Each agent is informed of her own type, and must choose her policy as a function of it. Payoffs are a function of types and actions, and agents try to maximize their expected payoffs, given the strategies of the others, leading to the notion of *Bayesian equilibrium*, the natural generalization of *Nash equilibrium* to the incomplete-information setup.

While games in which agents may have only finitely many types pose no difficulty for equilibrium existence, when there is a continuum of types the situation becomes much thornier. These are standards frameworks in economic modeling, as it is natural and convenient to allow, e.g., prices, quantities, and profits, to assume any value (within some range). However, the use of a continuum of types makes it extremely difficult to show that Bayesian equilibrium must exist.

One of the very few general existence results is Milgrom and Weber [43], who assumed that the prior belief of the agents is either independent across types, or at least absolutely continuous with respect to some independent prior (i.e., absolutely continuous with respect to the product of the marginals). It remained for some time an open question as to whether, failing this condition, equilibria could fail to exist. Simon [52] showed that this was indeed the case by constructing an example of a Bayesian game with a continuum of types and no Bayesian equilibrium. Hellman [21] provided an example of a two-player Bayesian game with finite action spaces and no Bayesian ε-equilibrium for all $\varepsilon > 0$ sufficiently small; that is, for every strategy profile, a positive probability of types of one of the players can profit more than ε by deviating. Simon and Tomkowicz [53, 54] provided examples of, respectively, three-player and two-player Bayesian games that do not admit a Harsanyi ε-equilibrium for $\varepsilon > 0$ sufficiently small; that is, for every strategy profile, at least one player can profit more than ε by deviating at the ex-ante stage game.

The information structure we examine in this paper is that of *nested information*; that is, the players can be ordered from most knowledgeable to least knowledgeable. The most knowledgeable player (say, Player 1) knows everything Player 2 knows (and possibly more), Player 2 knows everything Player 3 knows (and possibly more), etc. Such structures have been modeled in hierarchical organization paradigms, financial market games, persuasion models, and others; we recall some of these works and more below. Such games generally do not satisfy the absolute continuity condition of Milgrom and Weber [43]. For instance, if there are three agents, two of which are informed of a value v in some range $[\underline{v}, \overline{v}]$ which distributes continuously while the third is not, the possible type profiles distribute continuously along a diagonal, and do not satisfy the absolute continuity condition.

In this work, we study these games when players have finitely many actions at their disposal. Our main result shows that in such games, Bayesian equilibria do exist, thereby exhibiting an additional class of incomplete information games possessing equilibria. (A discussion on models with a continuum of actions appear in Section 7.1.) Our proof introduces two new tools to the study of Bayesian games, which may prove useful also for other classes of games as well as other questions on Bayesian games.

The first tool, used for establishing existence of Bayesian ε-equilibrium, is a finite approximation of the belief hierarchy. As is well known, the players' belief about the payoffs, the others' beliefs, the others' beliefs about the others beliefs, and so forth, form an infinite hierarchy. When information is nested, as we will elaborate below, this infinite hierarchy is determined by the first n orders of the ladder, where n is the number of players. This finiteness of the relevant levels will allow us to construct a finite approximation of the space of infinite hierarchies that is sufficient for our purpose: We will define an approximating Bayesian game whose finite type spaces are induced by this finite approximation, and show that a Bayesian equilibrium of this game, which exists by [20], yields a Bayesian ε-equilibrium of the original game.

The second tool, used for establishing existence of Bayesian equilibrium, is the Measurable "Measurable Choice" Theorem by Mertens [40], a tool which had previously been used in the study of stochastic games but is novel in its application to Bayesian games.[1] To construct a Bayesian equilibrium we would like to take a limit of Bayesian ε-equilibria as ε goes to 0. However, it is well known that in the limit, correlation may be introduced; see Stinchcombe [56]. Conceptually, we construct the equilibrium among the accumulation points of a sequence of Bayesian $\frac{1}{n}$-equilibria, step-by-step, starting from the least knowledgeable player, and for each player we need to use a purification result to guarantee appropriate consistency with the selections already chosen. Not only is the purification done repeatedly, but it needs to be done a continuum-many times at each stage, all in a measurable fashion; this is precisely where the Measurable "Measurable Choice" Theorem comes into play.

Structure of the paper. The paper is organized as follows. Section 2 discusses related literature on Bayesian games and on nested information structures. Section 3 presents the model and the main result. Section 4 gives heuristic overviews of the proofs. Discussion and open problems appear in Section 5. The proof of the main result appears in Section 6. Section 7 presents two extensions of the main result.

2 Literature on Bayesian Equilibrium and on Nested Information

Bayesian Equilibria Since Bayesian (and even Harsanyi) ε-equilibria need not exist in Bayesian games, it is important to find sufficient conditions on the parameters of the game that ensure they exist. A large literature expanded the sufficient conditions identified by Harsanyi [20] and Milgrom and Weber [43].

Stinchcombe and White [55] proved that when all players share the same information, or when there are two players and information is nested, a Harsanyi equilibrium exists. Ui [58] proved the existence and uniqueness of Bayesian equilibrium in Bayesian games where the payoff function is continuously differentiable on the action

[1] An application of the Measurable "Measurable Choice" Theorem to a one sender - many receivers game can be found in Zeng [64].

space for each vector of types, and its gradient satisfies certain conditions. Hellman and Levy [22, 23] studied Bayesian games with purely atomic types; they showed that a Bayesian equilibrium exists when the common knowledge relation is smooth, namely, the common knowledge classes are level sets of a Borel function. Moreover, for any common knowledge relation that is not smooth, there exists a type space that yields this common knowledge relation and payoff functions such that the resulting Bayesian game does not have a Bayesian ε-equilibrium, provided ε is sufficiently small. Carbonell-Nicolau and McLean [9] extended the result of Milgrom and Weber [43] to Bayesian games with general action sets, by requiring that the payoff functions are upper semi-continuous and satisfy a condition related to Reny's uniform payoff security (Reny [49]). Olszewski and Siegel [46] simplified the application of Reny's [49] better-reply security to Bayesian games where players' types are independent, and used this condition to prove the existence of Harsanyi equilibria for classes of games in which payoff discontinuities arise only at "ties."

Several papers provided sufficient conditions that guarantee the existence of a *pure* Bayesian equilibrium, see, e.g., Radner and Rosenthal [48], Vives [61], Khan and Sun [30], Reny [50], and Barelli and Duggan [8]. While [48] assumed the players' types are independent, and [8] made the more general assumption of [43] regarding the absolute continuity of the joint distribution of types, the other works do not make these assumptions. Existence of equilibria in Bayesian games with *infinitely* many players was studied by, e.g., Kim and Yannelis [32], Balbus et al. [7], and Yang [63].

It is interesting to note that Bayesian games with nested information can be recast as regular projective games (see Myerson and Reny [44], Section 9). It follows from Theorem 9.3 in Myerson and Reny [44] that for every $\varepsilon > 0$, Bayesian games with nested information admit a Bayesian ε-equilibrium under proper technical conditions, which include the continuity of the payoff function over the type space and the fact that the type distribution has a continuous density function. Our paper strengthens [44] by (i) proving the result for $\varepsilon = 0$, and (ii) weakening the conditions required to derive the existence result (while requiring that the set of actions is finite rather than general) by dropping the requirement that the prior has a density with respect to a product distribution on types, which, would not be satisfied in many nested information structures of interest.

Nested Information Nested information arises naturally in strictly hierarchical organizations, where higher-level managers have more information than lower-level managers and workers. For example, Mathevet and Taneva [39] considered a game where these information structures arise endogenously. In their case, before the agents simultaneously take their payoff-relevant action, there is cheap-talk communication between the players following the strict order prescribed by the hierarchy; that is, each agent sends messages to the agent immediately below her. It can be shown that regardless of the exact information transmitted, the information structure endogenously generated in the cheap-talk stage will be nested.

Nested information also arises naturally in situations where players obtain or are exposed to different levels of information. For example, managers of firms are more informed about the firm's financial situation than large investors, who in turn are more informed than small investors. Experiments on financial market games where investors can predict future dividends or future value of a certain stock for different spans of time have been reported by, e.g., Toth [57] and Huber [26].

Another model with nested information is when players are divided into two subsets: those who obtain symmetric information about the state of the world, and those who are completely ignorant about it. For example, Debo et al. [14] and Kremer and Debo [33] study service system that can provide service in various qualities. Customers have two possible types: some know the service quality, while the others are not exposed to this information. Additional literature on service systems with similar features can be found in Hassin [19].

A more general information structure is considered in Wu et al. [62], who study a routing model where the players are divided into groups, and the players in each group obtain the same information on the state of nature. When the number of groups is 2 and the players in one of the groups obtain no information, or when the signals that the groups share are nested, this model exhibits nested information as well.

Finally, nested information arises in two-player models, where one player is more informed than the other, like dynamic games with asymmetric information (e.g., Aumann and Maschler [5, 6], Cardaliaguet and Rainer [10], Grün [18], De Angelis et al. [12], and Jacobovic [27]), and Bayesian persuasion models (e.g., Kamenica and Gentzkow [29] and Kamenica [28]).

3 The Model and Main result

Notations. Let $\mathcal{N} = \{1, 2, \dots, n\}$, with n finite. Whenever $(X_i)_{i\in\mathcal{N}}$ is a collection of sets, we denote their Cartesian product by $X \equiv \prod_{i\in\mathcal{N}} X_i$; for $j \in \mathcal{N}$ we denote $X_{-j} \equiv \prod_{i\in\mathcal{N}\setminus\{j\}} X_i$. For any $1 \le j_1 \le j_2 \le n$ denote $[j_1 : j_2] \equiv \{j_1, j_1 + 1, j_1 + 2, \dots, j_2\}$ and $X_{j_1:j_2} \equiv \prod_{i=j_1}^{j_2} X_i$. A product of measurable spaces will always be considered a measurable space with the product σ-field. Whenever $x = (x_i)_{i\in\mathcal{N}}$ is a vector and $j \in \mathcal{N}$, we set $x_{-j} \equiv (x_i)_{i\in\mathcal{N}\setminus\{j\}}$, and for any $1 \le j_1 \le j_2 \le n$ we set $x_{j_1:j_2} \equiv (x_{j_1}, x_{j_1+1}, \dots, x_{j_2})$. When $(U_i)_{i=1}^n$ are real-valued functions, we denote by $U_{j_1:j_2}$ the vector-valued function $(U_{j_1}, U_{j_1+1}, \dots, U_{j_2})$.

For every measurable set X, we denote by $\Delta(X)$ the set of probability distributions on X. We consider $\Delta(X)$ as a topological space, e.g., by endowing it with some metric like the total variation metric or the Prokhorov metric. When X and Y are two random variables, we say that Y *is determined by* X if there exists a measurable function $\kappa(\cdot)$ such that $\kappa(X) = Y$ with probability one.

Definition 1 (Bayesian game) *A* Bayesian game Γ *is given by*

- *A finite set of players* $\mathcal{N} \equiv \{1, 2, \dots, n\}$, *for some* $n \ge 2$.
- *For each* $i \in \mathcal{N}$, *a Polish*[2] *space* $\mathcal{T}_i$.
- *A common prior distribution* $\mathbb{P}$ *on* $\mathcal{T} \equiv \prod_{i\in\mathcal{N}} \mathcal{T}_i$.
- *For each* $i \in \mathcal{N}$, *a finite set* $\mathcal{A}_i$ *of actions. Recall that* $\mathcal{A} = \prod_{i\in\mathcal{N}} \mathcal{A}_i$.
- *For each* $i \in \mathcal{N}$, *a bounded and measurable payoff function* $R_i : \mathcal{T} \times \mathcal{A} \to \mathbb{R}$. *For each* $i \in \mathcal{N}$, *we denote by* $R_i(a) : \mathcal{T} \to \mathbb{R}$ *the* a*-section of* R_i, *for each*

[2] Recall that a *Polish space* is a separable completely metrizable topological space; that is, a space homeomorphic to a complete metric space that has a countable dense subset.

$a \in \mathcal{A}$; and by $R_i(t) : \mathcal{A} \to \mathbb{R}$ the t-section of R_i, for each $t \in \mathcal{T}$. We also set $R \equiv (R_i)_{i \in \mathcal{N}}$.

We will denote by $t = (t_1, t_2, \ldots, t_n)$ a random type profile, so that t_i is the random type of Player i.

For every $i \in \mathcal{N}$, denote by $\mathcal{X}_i \equiv \Delta(\mathcal{A}_i)$ the set of mixed actions of Player i. A (behavior) *strategy* of Player i is a measurable function $s_i : \mathcal{T}_i \to \mathcal{X}_i$. This definition indicates the interpretation of the type spaces: each player $i \in \mathcal{N}$ knows her own type, and is not told the types of the other players. Denote by $\mathcal{S}_i$ the set of strategies of Player i, so that $\mathcal{S} \equiv \prod_{i \in \mathcal{N}} \mathcal{S}_i$ is the set of all strategy profiles.

Every strategy profile $s \in \mathcal{S}$ induces a probability distribution over $\mathcal{T} \times \mathcal{A}$, denoted $\mathbb{P}_s$, which satisfies

$$\mathbb{P}_s(T \times B_1 \times \cdots \times B_n) = \int_T \prod_{i \in \mathcal{N}} s_i(t_i)(B_i) \mathrm{d}\mathbb{P}(t), \tag{1}$$

for every Borel set $T \subseteq \mathcal{T}$ and every measurable sets $B_i \subseteq \mathcal{A}_i$ for $i \in \mathcal{N}$. Denote by $\mathbb{E}_s$ the corresponding expectation operator. For every $s \in \mathcal{S}$, whenever the expectation of R_i with respect to $\mathbb{P}_s$ is well defined, Player i's expected payoff under the strategy profile s is the real number

$$U_i(s) \equiv \mathbb{E}_s[R_i],$$

and her conditional payoff given her information is the random variable (determined by t_i)[3]

$$U_i(s \mid t_i) \equiv \mathbb{E}_s[R_i \mid t_i].$$

The solution concept we will concentrate on in this paper is Bayesian ε-equilibrium.

Definition 2 (Bayesian ε-equilibrium) *Given $\varepsilon \geq 0$, a strategy profile $s^* \in \mathcal{S}$ is a* Bayesian ε-equilibrium *if for every player $i \in \mathcal{N}$ and every strategy $s_i \in \mathcal{S}_i$,*

$$U_i(s_i, s^*_{-i} \mid t_i) \leq U_i(s^* \mid t_i) + \varepsilon, \quad \mathbb{P}\text{-}a.s. \tag{2}$$

[3] Here for simplicity we abuse notation. Formally, $U_i(s \mid t_i)$ is the conditional expectation of R_i given the sigma-field $\mathcal{B}(\mathcal{T}_i) \times \prod_{j \neq i} \{\mathcal{T}_j, \emptyset\}$.

Remark 1 (The relation between Bayesian and Harsanyi equilibria) Given $\varepsilon \geq 0$, a strategy profile $s^* \in \mathcal{S}$ is a *Harsanyi ε-equilibrium* if for every player $i \in \mathcal{N}$ and every strategy $s_i \in \mathcal{S}_i$,

$$U_i\left(s_i, s_{-i}^*\right) \leq U_i\left(s^*\right) + \varepsilon. \tag{3}$$

Standard conditioning implies that every Bayesian ε-equilibrium is a Harsanyi ε-equilibrium. When $(\mathcal{T}, \mathcal{B}(\mathcal{T}), \mathbb{P})$ is complete, a Harsanyi 0-equilibrium is also a Bayesian 0-equilibrium. When $\varepsilon > 0$, a Harsanyi ε-equilibrium is not necessarily a Bayesian ε-equilibrium. Indeed, modifying a Bayesian 0-equilibrium on a set of types of sufficiently small measure arbitrarily will generically yield such an example. In fact, the example provided by Hellman [21] shows that when $\varepsilon > 0$ is sufficiently small, Harsanyi ε-equilibria may exist while Bayesian ε-equilibria do not. See Hellman and Levy [24] for further discussion on this issue.

Harsanyi [20] first presented the model of Bayesian games, and proved that when all sets that define the game are finite, a Bayesian equilibrium exists. Milgrom and Weber [43] studied Bayesian games with general type spaces, and proved that a Harsanyi equilibrium exists in distributional strategies when $\mathbb{P}$ is absolutely continuous w.r.t. the product of its marginals, that is, w.r.t. $\mathbb{P}_1 \otimes \mathbb{P}_2 \otimes \cdots \otimes \mathbb{P}_n$, where $\mathbb{P}_i$ is defined by $\mathbb{P}_i(B_i) \equiv \mathbb{P}\left(\left(\prod_{j \neq i} \mathcal{T}_j\right) \times B_i\right)$ for every $i \in \mathcal{N}$ and every $B_i \in \mathcal{B}(\mathcal{T}_i)$.[4] As mentioned in the introduction, Simon [52] and Hellman [21] (resp., Simon and Tomkowicz [53, 54]) provided examples of Bayesian games with finite action spaces and no Bayesian ε-equilibria (resp., no Harsanyi ε-equilibria), for $\varepsilon > 0$ sufficiently small. Additional sufficient conditions on the parameters of the game that ensure the existence of a 0-equilibrium have already been reviewed in the introduction.

[4] Milgrom and Weber [43] include an assumption, denoted there (R1), which requires that for each Player i and each $\varepsilon > 0$, there is a subset $E \subseteq \mathcal{T}$ such that the collection $\{R_i(t, \cdot)\}_{t \in E}$ is equicontinuous. When actions sets are finite, like in our model, (R1) holds vacuously; indeed, any collection of functions on a finite set is equicontinuous. In fact, that assumption holds vacuously under the more general model of [43] – a model we discuss later in Section 7.1 – allowing for compact action sets and payoffs that are Borel, bounded, and continuous in actions for each fixed type profile. This observation follows from Scorza-Dragoni type theorems (see, e.g., [34, Theorem 1]), of which the following classical version is a particular case: Let T, X, Y be Polish spaces, μ a Borel measure on T, and $f : T \times X \to Y$ a Borel function such that for each fixed $t \in T$, $f(t, \cdot)$ is continuous. Then there is a compact set $E \subseteq T$ such that $f|_{E \times X}$ is (jointly) continuous.

In this paper we concentrate on Bayesian games where the information of the players is nested.

Definition 3 (Nested information) *We say that the information of the players in a Bayesian game is* nested *if t_{i+1} is determined by t_i, for every $i = 1, 2, \ldots, n-1$; that is, if for each $i < n$ there is a mapping $\kappa_i : \mathcal{T}_i \to \mathcal{T}_{i+1}$ such that*[5]

$$\mathbb{P}\big(t_{i+1} = \kappa_i(t_i), \quad \textit{for every } 1 \le i < n\big) = 1. \tag{4}$$

Remark 2 (Players possessing the same information) Note that the definition allows for two or more players to possess the same information. Indeed, players i and $i+1$ have the same information if the function κ_i in Eq. (4) is a bijection.

Remark 3 ($\mathbb{P}$-a.s. versus everywhere in Eq. (4)) Nested information requires that $t_{i+1} = \kappa_i(t_i)$, $\mathbb{P}$-a.s. and not everywhere. This distinction is irrelevant for our purposes. Indeed, let $\mathbb{Q}$ be the measure on $\prod_{i \in I} \mathcal{T}_i$ whose marginal on $\mathcal{T}_1$ coincides with that under $\mathbb{P}$, and that is determined by its marginal on $\mathcal{T}_1$ and the functions $\kappa_1, \ldots, \kappa_{i-1}$. We then have $\mathbb{P} = \mathbb{Q}$.

Remark 4 (Nested information and absolute continuity of information) We here show that a Bayesian game with nested information may not satisfy the requirement of absolute continuity of information structure, as studied by Milgrom and Weber [43]. Indeed, suppose that $\mathcal{T}_i = [0,1]$ for each $i \in \mathcal{N}$, and $\mathbb{P}$ is the uniform distribution on the diagonal $\{t_1 = t_2 = \cdots = t_n\}$. The resulting measure $\bigotimes_{i=1}^n \mathbb{P}_i$ is the Lebesgue measure on $[0,1]^n$, and hence $\mathbb{P}$ is concentrated on a set of $(\bigotimes_{i=1}^n \mathbb{P}_i)$-measure zero. Therefore, this information structure does not satisfy the absolute continuity condition of [43], yet the players have nested information.

Since Bayesian games that satisfy absolute continuity of information structures do not necessarily have nested information (see, e.g., Example 2 in [43]), it follows that nested information is unrelated to absolute continuity of information structure.

The main result of the present work is the following.

[5] More generally, we could say that the information of the players is nested if the condition in Definition 3 holds after a permutation of the players.

Theorem 1 (Existence of 0-equilibrium) *Every Bayesian game with nested information admits a Bayesian* 0*-equilibrium.*

Remark 5 (Extensions of Theorem 1) Below we will discuss three extensions of Theorem 1, to games with inconsistent beliefs (Section 5), compact action sets (Section 7.1), and tree-like information structure (Section 7.2).

Remark 6 (Games with two players and nested information) Example 3.3 in Stinchcombe and White [55] implies that in the presence of two players who have nested information (and concave payoffs in their actions) a Harsanyi 0-equilibrium exists. As mentioned in Remark 1, when $(\mathcal{T}, \mathcal{B}(\mathcal{T}), \mathbb{P})$ is complete, this implies the existence of a Bayesian 0-equilibrium. Thus, Theorem 1 extends the result of Stinchcombe and White [55] to any number of players.

Remark 7 (Comparison with Levy [38]) Levy [38] examines Bayesian games in which the type space can be partitioned into a collection, generally a continuum, of components, such that each component is a common knowledge and the game on each component possesses an equilibrium. That paper introduces conditions under which an equilibrium can be selected on each component in a measurable manner to induce an equilibrium in the entire game. (That work generalizes Hellman and Levy [22], which establishes a similar result under the additional assumption that each common knowledge component is countable.) In the framework of nested information, for every $t'_n \in \mathcal{T}_n$ the set $\{t \in \mathcal{T} \colon t_n = t'_n\}$ is a common knowledge component, yet there is no result that guarantees the existence of a Bayesian 0-equilibrium on each connected component.

4 The Driving Force Behind our Proofs

This section gives a heuristic explanation of our methodology. The proof is divided into two main parts: The first establishes the existence of ε-equilibria, while the second uses the Measurable "Measurable Choice" Theorem to construct an appropriate limit of approximate equilibria which constitute an exact equilibrium. As remarked earlier, we point out that it is a well-known problem that in games with a continuum of states, limits of approximate equilibria do not, in general, naturally induce exact equilibria, as the limiting process can induce correlation; for an elaboration on this point, see Stinchcombe [56].

Belief Hierarchies In Bayesian games, to determine her action, on top of her own information on the players' types, a player needs to take into account also:

- her information on the information the other players have on the players' types,
- her information on the information each Player i has on the information each Player $j \neq i$ has on the players' types,
- her information on the information each Player i has on the information each Player $j \neq i$ has on the information each Player $k \neq j$ has on the players' types,
- etc.

In general, the information encapsulated in higher levels cannot be deduced from the information encapsulated in lower levels. This gives rise to an *infinite belief hierarchy*, which typically depends on the players' types.

The belief hierarchy of a player identifies the set of type profiles that should be taken into account when determining the player's action. In fact, the belief hierarchies of the players divide the set of type profiles into disjoint subsets, called *minimal belief subspaces*, such that the type profiles in each subspace are closed, in the sense that when the actual type profile is in a given subspace, only type profiles in that subspace need to be considered to determine the players' actions in equilibrium. As showed by Simon [52], even if the game restricted to each of the minimal belief subspaces has an equilibrium, the amalgamation of these equilibria need not be measurable.

When information is nested, the infinite belief hierarchy can be deduced from its first n levels, where n is the number of players. Indeed, if, say, there are two players and Player 1 is more informed than Player 2, then Player 2's infinite belief hierarchy can be deduced from her own information on the players' types and her information on

Player 1's information on the players' types; and Player 1's infinite belief hierarchy can be deduced from her own information on the players' types, and Player 2's information on the players' types and on Player 1's information on the players' types. For example, the next level in the belief hierarchy of Player 2 corresponds to Player 2's information on Player 1's information on Player 2's information on the players' types, which coincide with Player 2's information on the players' types. Similarly, the next level in the belief hierarchy of Player 1 corresponds to Player 1's information on Player 2's information on Player 1's information on the players' types, which coincide with Player 2's information on Player 1's information on the players' types. Thus, when information is nested, there is no need to consider infinite belief hierarchies, and the game has a finite structure.

ε-Equilibrium: Approximating Belief Hierchies The observation made in the previous paragraph leads us to define a finite approximation of the belief hierarchy when information is nested, which is useful in proving the existence of a Bayesian ε-equilibrium. Let us explain this approximation.

Assume that the action spaces are finite, and suppose again that there are two players, where Player 1 is more informed than Player 2. When Player 1 observes the type profile realization $t = (t_1, t_2)$, Player 1 has a belief over the matrix game that is being played. As at present the payoffs are bounded and we are interested in an ε-equilibrium, we can assume that the collection $\mathcal{R}$ of all possible matrix games is finite. Fixing $\delta > 0$, we can choose a δ-dense subset $\mathcal{D}_1$ of the set $\Delta(\mathcal{R})$ of probability distributions over $\mathcal{R}$, and approximate Player 1's belief at t by the closest point in $\mathcal{D}_1$, denoted $\varphi_1(t)$. We can then consider the mapping ψ_2 that assigns to each t the pair $(\varphi_1(t), R(t))$, namely, Player 1's approximated belief at t and the payoff matrix at t, and consider the distribution of this vector given t_2 which is Player 2's information at t. Since the mapping φ_1 takes only finitely many values, and the number of possible payoff matrices is finite, the range of ψ_2 is finite dimensional, and hence can be in turn δ-approximated by a mapping φ_2 with finitely many values; the range of φ_2 is a δ-dense subset $\mathcal{D}_2$ of $\Delta(\mathcal{D}_1 \times \mathcal{R})$. The mapping φ_2 represents the approximated information Player 2 has at t, on both the payoff matrix and on Player 1's information on the payoff matrix. Finally, we say that Player 1's approximated belief is composed by the pair (φ_1, φ_2), and Player 2's approximated belief is composed solely of φ_2.

Since the approximating information divides the state space into finitely many sets, the resulting game admits a Bayesian 0-equilibrium. The properties of the approximation then imply that this Bayesian 0-equilibrium is a Bayesian ε-equilibrium, provided δ is sufficiently small.

Correlation of Limits. To construct a Bayesian 0-equilibrium, we would like to consider an accumulation point of a sequence of Bayesian $\frac{1}{n}$-equilibria as $n \to \infty$. Unfortunately, as mentioned above and discussed in [43] (see also [56]), when the type space is general, a limit of strategy profiles may be a correlated strategy profile. We will illustrate this issue using a variation of Example 2 of [43].

There are three players, and the incomplete information concerns the value of a state variable that is uniformly distributed in $[0, 1]$: Players 1 and 2 know the state, while Player 3 does not obtain any information on the state. Thus, the game exhibits nested information. Formally, $\mathcal{T}_1 = \mathcal{T}_2 = [0, 1]$, $\mathcal{T}_3$ is the singleton $\emptyset$, and $\mathbb{P}(t_1 = t_2) = 1$. Each player has two actions, L and R, and the payoff function, which is independent of the state, is given in Figure 1.

	L	R
L	$1, 1, 1$	$0, 0, 0$
R	$0, 0, 0$	$1, 1, 1$

L

	L	R
L	$1, 1, 0$	$0, 0, 2$
R	$0, 0, 2$	$1, 1, 0$

R

Figure 1: A three-player game: Player 1 (resp., 2, 3) selects a row (resp, column, matrix).

For every $k \in \mathbb{N}$, the following strategy profile s^k is a Bayesian 0-equilibrium: Player 1 (resp., Player 2) selects L for every t_1 (resp., t_2) such that the integer part of kt_1 (resp., kt_2) is odd, and R otherwise (i.e., the two players alternate their action according to the k'th Rademacher function); Player 3 selects L.

The limit of the strategies $(s_1^k)_{k\in\mathbb{N}}$, $(s_2^k)_{k\in\mathbb{N}}$, and $(s_3^k)_{k\in\mathbb{N}}$ in the weak-* topology on $L^\infty([0, 1], \Delta(\{L, R\}))$ are the strategies in which at every t_1 (resp., t_2), Player 1 (resp., Player 2) selects $[\frac{1}{2}(L), \frac{1}{2}(R)]$, and Player 3 selects L. However, this limit strategy profile is not a Bayesian 0-equilibrium, since Player 3 can profit by deviating to R.

The only reasonable limit of the profiles $(s^k)_{k\in\mathbb{N}}$ is the correlated strategy which always mixes between (L, L, L) and (R, R, L) with equal probabilities, which is their limit in the weak-* topology on $L^\infty([0,1], \Delta(\{L,R\}^3))$. This resulting correlation hints at the need for purification tools.

We note that this difficulty is not helped by working with distribution strategies, as in [43]; the weak-$*$ convergence on $L^\infty([0,1], \Delta(\{L,R\}))$ (resp., $L^\infty([0,1], \Delta(\{L,R\}^3))$) is equivalent to the weak convergence of the measures induced by the Lebesgue measure and these strategies (resp., profiles of strategies) on $[0,1] \times \Delta(\{L,R\})$ (resp., $[0,1] \times \Delta(\{L,R\}^3)$). The difficulty runs much deeper, and requires the use of properties of the information structure – in this case, the nested information. Indeed, in [52], ε-equilibria exist for each $\varepsilon > 0$, but no exact equilibria exist.[6]

Exact Equilibrium: Using the Measurable "Measurable Choice" Theorem. To overcome the difficulty of correlations in the limit, we use iteratively an extension of Mertens' [40] Measurable "Measurable Choice" Theorem.

To illustrate the construction, consider first a three-player Bayesian game where, as above, $\mathcal{T}_1 = \mathcal{T}_2$, $\mathcal{T}_3 = \{\emptyset\}$, the prior $\mathbb{P}$ is concentrated on the set $\{t_1 = t_2\}$, each player has two actions, L and R, and the payoff function is bounded and measurable. In this framework, equilibrium existence had been previously an open question.

Fix a sequence of approximate equilibria $(s^k)_{k\in\mathbb{N}}$, where s^k is a Bayesian $\frac{1}{k}$-equilibrium for every $k \in \mathbb{N}$. As described above, an accumulation point s^* of $(s^k)_{k\in\mathbb{N}}$ may fail to be a Bayesian 0-equilibrium, because, e.g., the payoff $U_3(s^*_{1:2}, L \mid \emptyset)$ may differ from $\lim_{l\to\infty} U_3(s^{k_l}_{1:2}, L \mid \emptyset)$, where $(k_l)_{l\in\mathbb{N}}$ is some sequence such that $s^* = \lim_{l\to\infty} s^{k_l}$.

To avoid this problem, we need $(k_l)_{l\in\mathbb{N}}$ to be such that not only does $(s^{k_l}_3(\emptyset))_{l\in\mathbb{N}}$ converge, but so do the sequences $\big(U_3(s^{k_l}_{1:2}, L \mid \emptyset)\big)_{l\in\mathbb{N}}$ and $\big(U_3(s^{k_l}_{1:2}, R \mid \emptyset)\big)_{l\in\mathbb{N}}$. Denote the corresponding limits by $\rho_3[L]$ and $\rho_3[R]$. The selection of $s^*_{1:2}(t_1)$ over $t_1 \in \mathcal{T}_1$ is done among the accumulation points of $(s^{k_l}_{1:2}(t_1))_{l\in\mathbb{N}}$ so that, in the aggregate, $U_3(s^*_{1:2}, a_3 \mid \emptyset) = \rho_3[a_3]$ for $a_3 \in \{L, R\}$. The resulting selections together with s^*_3 can

[6] See [24] for a thorough discussion on this issue.

be shown to be an equilibrium.

To be more precise, and of particular relevance as the examples become more complex and we look towards a general technique, is that we do *not* actually fix a specific subsequence of indices $(k_l)_{l\in\mathbb{N}}$; what we do is stipulate that in the latter stage, when we select among accumulation points of $(s^k_{1:2}(t_1))_{k\in\mathbb{N}}$, we only select among limits of subsequences whose indices $(k_l)_{l\in\mathbb{N}}$ ensure not only the convergence $s_3^{k_l}(\emptyset) \to s_3^*(\emptyset)$, but *also* the convergence $U_3(s^{k_l}_{1:2}, L \mid \emptyset) \to \rho_3[L]$ and $U_3(s^{k_l}_{1:2}, R \mid \emptyset) \to \rho_3[R]$. There may be many such subsequences $(k_l)_{l\in\mathbb{N}}$, and we do not select a particular one. Rather, we select the limits we desire, and make sure that future selections are consistent with these limits.

In the example above on Page 14, for every $t \in [0, 1]$, the set of accumulation points of $(s^k(t))_{k\in\mathbb{N}}$ is $\{(L, L, L), (R, R, L)\}$, and any strategy profile s^* in which on a set $T \subseteq [0, 1]$ of Lebesgue measure $\frac{1}{2}$ both Players 1 and 2 select L, on T^c they both select R, and Player 3 selects L, is a selection that satisfies our requirements.

Now, let us spice up the example. Suppose there are four players: Player 4 knows nothing, Player 3 knows something, and Players 1 and 2 know everything. Formally, $\mathcal{T}_1 = \mathcal{T}_2$ and a.s. $t_1 = t_2$ (that is, $\kappa_1(t_1) = t_1$), $\mathcal{T}_3$ is non-trivial (there is some $\kappa_2 : \mathcal{T}_2 \to \mathcal{T}_3$), while $\mathcal{T}_4$ is a singleton $\emptyset$ ($\kappa_3(t_3) = \emptyset$). Again, assume players have two actions L and R. We fix a sequence of strategy profiles $(s^k)_{k\in\mathbb{N}}$ where s^k is a Bayesian $\frac{1}{k}$-equilibrium. Once again, starting with the least knowledgeable player, we select some accumulation point of the triplet $\big(s^k_4(\emptyset), U_4(s^k_{1:3}, L \mid \emptyset), U_4(s^k_{1:3}, R \mid \emptyset)\big)_{k\in\mathbb{N}}$ to $\big(s^*_4(\emptyset), \rho_4[L], \rho_4[R]\big)$.

Move on to the second-least informed player, Player 3, who has partial knowledge. We need to take care that the construction of s^*_3 would leave open the door for construction of $s^*_{1:2}$ with $\rho_4[L] = U_4(s^*_{1:3}, L \mid \emptyset)$ and $\rho_4[R] = U_4(s^*_{1:3}, R \mid \emptyset)$. To this end, for each $t_3 \in \mathcal{T}_3$, we choose an accumulation point that is consistent with the convergence we have already established of the 9-tuple of $s^k_3(t_3)$ and $U_j(s^k_{1:2}, a_3, a_4 \mid t_3)$ for each $j = 3, 4$ and each $a_3, a_4 \in \{L, R\}$, where $U_4(\cdot \mid t_3)$ means $U_4(\cdot \mid \kappa_3(t_3))$. That is, we keep track not only of Player 3's strategy, but also of the expected payoffs to Player 3 *and* Player 4 for each action *profile* of these players. Denote a chosen accumulation point of this 9-tuple as $s^*_3(t_3)$ and $\rho_3[j, a_3, a_4](t_3)$ for $j = 3, 4$ and $a_3, a_4 \in$

$\{L, R\}$. Like in the previous example, it is not sufficient to choose any accumulation points across different $t_3 \in \mathcal{T}_3$; we must take care that these accumulation points are chosen so that if Player 4 imagines the setup in which it is only her and Player 3, and payoffs are given by ρ_3, then her expected payoff for an action $a_4 \in \{L, R\}$ when Player 3 uses s_3^* is precisely $\rho_4[a_4]$, i.e., for $a_4 \in \{L, R\}$, $\rho_4[a_4] = \int_{\mathcal{T}_3} \rho_3[4, s_3^*(t_3), a_4]\mathbb{P}(dt_3)$. This establishes a certain consistency between ρ_4, ρ_3, and s_3^*.

In the next step, consider Player 1 and 2, who have full knowledge because $t_3 = \kappa_2(t_1)$ and $\kappa_3(t_3) = \emptyset$. We need to choose for each t_1 an accumulation point of $(s_{1:2}^k(t_1))_{k\in\mathbb{N}}$ that is consistent with previous selections for this particular $\kappa_2(t_1)$, i.e., along subsequences of indices which give the chosen accumulation points $s_4^*(\emptyset)$, $s_3^*(\kappa_2(t_1))$, and $\rho_3[j, a_3, a_4]$ of $s_4^k(\emptyset)$, $s_3^k(\kappa_2(t_1))$, and $U_j(s_{1:2}^k(t_1), a_3, a_4 \mid \kappa_2(t_1))$ for $j = 3, 4$ and $a_3, a_4 \in \{L, R\}$, respectively. Furthermore, we need that the selection is done across all $t_1 \in \mathcal{T}_1 = \mathcal{T}_2$, so that, for each $t_3 \in \mathcal{T}$, the expected payoffs to Player $j = 3, 4$ under $s_{1:2}^*$ for any pair of actions $a_3, a_4 \in \{L, R\}$ they may play, $U_j(s_{1:2}^*, a_3, a_4 \mid t_3)$, agrees with $\rho_3[j, a_3, a_4](t_3)$. To do it for all $t_3 \in \mathcal{T}_3$ in parallel in a measurable fashion, we appeal to the Measurable "Measurable Choice" Theorem of [40].

The proof in general is a formalization of the ideas above, although we note that we go player-by-player, without 'bunching together' players who have identical information.

5 Discussion

Our study raises several extensions and open problems, which we present in this section.

Inconsistent beliefs. We assumed that all players share the same belief $\mathbb{P}$ on $\mathcal{T}$. Our result holds also when beliefs are inconsistent, namely, each player $i \in \mathcal{N}$ holds a different belief $\mathbb{P}_i$ on $\mathcal{T}$. In this case, each player's payoff is defined relative to $\mathbb{P}_i$ (rather than relative to $\mathbb{P}$). The condition of nested information, in this case, is as follows: for each $i < n$ there is a mapping $\kappa_i : \mathcal{T}_i \to \mathcal{T}_{i+1}$ such that

$$\mathbb{P}_j\Big(t_{i+1} = \kappa_i(t_i), \quad \text{for every } 1 \leq i < n\Big) = 1, \quad \forall j = 1, \ldots, n. \tag{5}$$

Equivalently, defining $\mathbb{P} = \frac{1}{n}\sum_{i=1}^{n} \mathbb{P}_i$, it is the requirement that for each $i < n$ there is a mapping $\kappa_i : \mathcal{T}_i \to \mathcal{T}_{i+1}$ such that

$$\mathbb{P}\Big(t_{i+1} = \kappa_i(t_i), \quad \text{for every } 1 \leq i < n\Big) = 1. \tag{6}$$

To see why this extension holds, define an auxiliary Bayesian game Γ' with type space $\mathcal{T}$, common prior $\mathbb{P}$, and payoffs $(R'_i)_{i \in \mathcal{N}}$ given by $R'_i(t,a) = \frac{d\mathbb{P}_i}{d\mathbb{P}}(t) R_i(t,a)$, where $\frac{d\mathbb{P}_i}{d\mathbb{P}}$ is a fixed Borel version of the Radon-Nikodym derivative. Given a profile of strategies s, the expected payoff to Player i in Γ' is the same as the expected payoff of Player i in Γ. Since the Radon-Nikodym derivative is bounded (by n), $(R'_i)_{i \in I}$ are bounded. Theorem 1 guarantees the existence of a Bayesian 0-equilibrium in Γ', which is a Bayesian 0-equilibrium in Γ.

On the continuity of the payoff function in type. One could ask whether the proof of our result simplifies if we assume that payoffs depend continuously on types. We are not aware of any method achieving such a simplification, and doubt that such an assumption would simplify matters much at all. Indeed, consider the three-player example presented in Section 4, in which two players know everything while the third knows nothing. Even for this relatively simple case, existence of equilibrium had been an open question. However, it is known (e.g., [31, 13.11]) that given a Borel function on a Polish space, the topology on the Polish space can be refined so that it remains Polish but the function is now continuous. Hence, in this example, since the payoff functions depend on a single type (and not a product) we could have assumed w.l.o.g. that payoffs are continuous in the type.

Multi-stage Bayesian games. We showed that nested information is a sufficient condition for the existence of a Bayesian 0-equilibrium in a general class of *single-stage* Bayesian games. A natural question regards the existence of a Bayesian 0-equilibrium in *multi-stage* Bayesian games with nested information.

Specifically, a multi-stage Bayesian game with $m \geq 2$ stages is similar to a Bayesian game as in Definition 1, except that the players play for m stages and Player i's type is stage dependent. That is, Player i's type is a vector $t_i = (t_i^1, t_i^2, \ldots, t_i^m)$, the collection of types of the players are drawn at the outset, and at each stage $1 \leq k \leq m$, each player learns her own stage type t_i^k. Player i's payoff at each stage k depends on the players' stage-types $(t_i^k)_{i \in \mathcal{N}}$, and the players stage actions.

Repeating the arguments of the current paper implies existence of 0-Bayesian equilibrium in the multi-stage game whenever the nested information assumption is replaced by the following two conditions:

A1 Information is nested: For each $k \in \{1, 2, \ldots, m\}$ and each $i \in \{1, 2, \ldots, n-1\}$, t_{i+1}^k is a determined by t_i^k.

A2 Information is revealed with delay of one stage: For every $k \in \{1, 2, \ldots, m-1\}$, t_1^k is determined by t_n^{k+1}; i.e., the information of Player 1 is available to Player n with delay of one stage.

Models of multi-stage Bayesian games that satisfy these two assumptions have been studied in the literature on control under the name *information structure with one-step-delay*, see, e.g., Aicardi et al. [1], Nayyar et al. [45], and Varaiya and Walrand [60]. We conjecture that Condition **A1** (without **A2**) is not sufficient to guarantee the existence of a Bayesian 0-equilibrium in multi-stage Bayesian games.

Stopping games with asymmetric information. One class of multi-stage Bayesian games is the class of stopping games with asymmetric information. In stopping games, players choose in each round to stop or continue; the game ends when at least one player chooses to stop, and the payoff profile is given by an $\mathbb{R}^n$-valued stochastic process that depends on the set of players who chose to stop. There is also some designated payoff in case the game never terminates. Incomplete and asymmetric information can be introduced by adding uncertainty on the payoff process.

Such games have been studied both in the framework of discrete time and that of continuous time, see, e.g., Grün [18], Lempa and Matomäki [37], Gensbittel and Grün [17], Esmaeeli, Imkeller, and Nzengang [15], Gapeev and Rodosthenous [16], Pérez et al. [47], Jacobovic [27], and De Angelis et al. [11, 12]. The open problem we raised for multi-stage Bayesian games translates to the following: Does every stopping game (in discrete or continuous time) with finite horizon and information structure that satisfies **A1** (or a continuous-time analog) admit a Bayesian 0-equilibrium?

Nested information and the value of information. Various aspects of the value of information in two-player zero-sum Bayesian games have been studied, e.g., by De Meyer et al. [13], Lehrer and Rosenberg [36], and Ui [59]. As we now argue, nested information is related to the study of the value of information in multiplayer Bayesian games with symmetric information. Indeed, define the value of the information of Player i as the difference between Player i's highest expected equilibrium payoff in a multiplayer Bayesian game with symmetric information, and her highest expected equilibrium payoff in the same game, when she does not obtain any information (while the other players obtain the symmetric information).[7] As the latter game has nested information, our result ensures that it admits at least one Bayesian 0-equilibrium, and hence the measure suggested above (and other natural variants) are well-defined.

The universal belief space. The universal belief space is the space that contains all infinite belief hierarchies, see Mertens and Zamir [42]. As we have seen, when information is nested, the players' belief hierarchies are determined by the first n levels in the belief hierarchy. It will be interesting to know whether there is a canonical form to the universal belief space in this case.

In Aumann's model of incomplete information, the information of a player is given by a partition of the state space, and the information structure is nested if under some ordering of the players, Player i's partition refines Player j's partition whenever $i < j$. Call an information structure *finite* if each belief hierarchy is determined by its first k levels, for some $k \in \mathbb{N}$. Nested information structures are finite. An example of a finite information structure that is not nested is that of a "piecewise" nested information structure: the state space is divided into several common knowledge components,

[7] In case the highest expected payoff is not attained by a Bayesian 0-equilibrium, consider the supremum of Player i's expected equilibrium payoffs.

and in each component, the information is nested, possibly with a different ordering between the players. Are the piecewise nested information structures *all* finite information structures?

6 Proof of Theorem 1

The proof of Theorem 1 consists of two steps. In Section 6.1, we prove the existence of a Bayesian ε-equilibrium,[8] for every $\varepsilon > 0$. In Section 6.2 we use this result to prove the existence of a Bayesian 0-equilibrium.

6.1 Existence of a Bayesian ε-equilibrium

In this section we prove the existence of a Bayesian ε-equilibrium, for every $\varepsilon > 0$. To this end, we approximate the information structure in a way that is related to the approximation used by Shmaya and Solan [51]. In Section 6.1.1 we review the notion of δ-approximation defined over a compact set in a metric space. This notion is applied in Section 6.1.2 to approximate the information structure of the players in Γ (under the nested information assumption). We then define a new game, which is identical to Γ except that the information of the players is replaced by its approximation. In Section 6.1.3, we show that this approximated game has a Bayesian 0-equilibrium, and that this Bayesian 0-equilibrium is a Bayesian ϵ-equilibrium of the original game Γ.

6.1.1 δ-approximations

We begin by recalling definitions related to dense subsets.

Definition 4 (δ-dense subset) *Let U be a set in a metric space $(\mathcal{M}, d)$, and fix $\delta > 0$. A set $V \subseteq U$ is δ-dense in U if for every $u \in U$ there is $v \in V$ such that $d(u, v) < \delta$.*

When U is contained in a compact set, the δ-dense set V that we will consider will be implicitly assumed to be finite. In this case, there exists a measurable mapping

[8] This part of the proof does not require that the type spaces are Polish.

$v_\delta : U \to V$ such that $d(u, v_\delta(u)) < \delta$ for every $u \in U$. For every $u \in U$, the image $v_\delta(u)$ is called a *δ-approximation* of u (by V). When Ω is a measurable space and $y : \Omega \to U$ is measurable, the mapping $v_\delta(y(\cdot)) : \Omega \to V$ will be a *measurable* δ-approximation of y.

For every random vector $Z : \mathcal{T} \to \mathbb{R}^d$ with finite range $\mathcal{Z}$, and every $i \in \mathcal{N}$, denote the conditional distribution of Z given t_i by

$$\mathbb{P}(Z \mid t_i) \equiv (\mathbb{P}(Z = z \mid t_i))_{z \in \mathcal{Z}}.$$

Thus, $\mathbb{P}(Z \mid t_i)$ is a vector determined by t_i and contained in the $(|\mathcal{Z}| - 1)$-dimensional simplex. Since the $(|\mathcal{Z}| - 1)$-dimensional simplex is compact, there exists a δ-approximation $\varphi(\cdot)(t_i)$ of $\mathbb{P}(Z \mid t_i)$ which belongs to the $(|\mathcal{Z}| - 1)$-simplex and satisfies

$$\sum_{z \in \mathcal{Z}} \left| \mathbb{P}(Z = z \mid t_i) - \varphi(z)(t_i) \right| < \delta \,,\ \mathbb{P}\text{-a.s.}$$

6.1.2 A finite approximation of the information structure

In this section we present a finite approximation of the information structure, which is suited to games where information is nested.

Fix a Bayesian game with nested information Γ, and denote by M a bound on the payoff function. Since in this section we are interested in proving the existence of a Bayesian ε-equilibrium, we can assume w.l.o.g. that the range of the payoff function R is finite. Namely, there is a finite collection $\mathcal{R}$ of functions from $\mathcal{A}$ to $[-M, M]^n$ such that for *every* $t \in \mathcal{T}$, the function $R(t) = (R_i(t))_{i \in \mathcal{N}}$ is an element of $\mathcal{R}$.

For every $i \in \mathcal{N}$, every strategy $s_i \in \mathcal{S}_i$, and every action $a_i \in \mathcal{A}_i$, denote by $s_i(a_i)(t_i)$ the probability that Player i selects the action a_i under s_i when her type is t_i. For every strategy profile $s \in \mathcal{S}$ and every action profile $a \in \mathcal{A}$, the probability under s that a is selected by the players is the random variable $p_s(a)$ defined by

$$p_s(a)(t) \equiv \prod_{i=1}^{n} s_i(a_i)(t_i), \quad \forall t \in \mathcal{T}.$$

Define

$$p_{s_{-i}}(a_{-i})(t_{-i}) \equiv \prod_{j \neq i} s_j(a_j)(t_j), \quad \forall i \in \mathcal{N}, s \in \mathcal{S}, a \in \mathcal{A}, t \in \mathcal{T}.$$

We are now going to recursively define approximations of the information that the players have. Fix $\delta > 0$. Let $\psi_1 : \mathcal{T} \to \mathcal{R}$ be the random vector defined by

$$\psi_1(t) \equiv R(t), \quad \forall t \in \mathcal{T}.$$

Denote $r \equiv |\mathcal{R}|$, so that $\mathbb{P}(\psi_1 \mid t_1)$ is in the $(r-1)$-dimensional standard simplex, $\mathbb{P}$-a.s. Equip $\mathbb{R}^r$ with the L_1-norm. Let $\varphi_1 \equiv \varphi_1(\cdot)(t_1)$ be a measurable δ-approximation of $\mathbb{P}(\psi_1 \mid t_1)$, so that φ_1 has a finite range and

$$\sum_{z \in \mathcal{R}} \Big| \mathbb{P}\left(\psi_1 = z \mid t_1\right) - \varphi_1(z)(t_1) \Big| < \delta, \quad \mathbb{P}\text{-a.s.}$$

For $i \in \mathcal{N} \backslash \{1\}$, suppose we have already defined random vectors $\psi_1, \varphi_1, \ldots, \psi_{i-1}, \varphi_{i-1}$, all with finite ranges, where $\psi_j : \mathcal{T} \to \prod_{k=1}^{j-1} \mathcal{D}_k \times \mathcal{R}$ and $\varphi_j : \mathcal{T}_j \to \Delta(\prod_{k=1}^{j-1} \mathcal{D}_k \times \mathcal{R}) \subseteq \mathbb{R}^{\prod_{j=1}^{i-1} |\mathcal{D}_j| \times r}$, where for each $1 \le j \le i-1$, $\mathcal{D}_j \subseteq \Delta(\prod_{k=1}^{j-1} \mathcal{D}_k \times \mathcal{R})$ is the range of φ_j. Define

$$\psi_i(t) \equiv \big(\varphi_1(t_1), \varphi_2(t_2), \ldots, \varphi_{i-1}(t_{i-1}), R(t)\big) \in \prod_{j=1}^{i-1} \mathcal{D}_j \times \mathcal{R}, \quad \mathbb{P}\text{-a.s.}, \tag{7}$$

which is a random vector with a discrete distribution. The range of ψ_i is finite and contains at most $\prod_{j=1}^{i-1} |\mathcal{D}_j| \times r$ elements. Equip $\mathbb{R}^{\prod_{j=1}^{i-1} |\mathcal{D}_j| \times r}$ with the L_1-norm, so the $\left(\prod_{j=1}^{i-1} |\mathcal{D}_j| \times r - 1\right)$-dimensional simplex is compact. Let φ_i be a measurable δ-approximation of $\mathbb{P}\left(\psi_i \mid t_i\right)$, and hence

$$\sum_{z \in \prod_{j=1}^{i-1} \mathcal{D}_j \times \mathcal{R}} \Big| \mathbb{P}\left(\psi_i = z \mid t_i\right) - \varphi_i(z)(t_i) \Big| < \delta, \quad \mathbb{P}\text{-a.s.} \tag{8}$$

Thus, the range of φ_i is contained in a finite δ-dense subset of the $\left(\prod_{j=1}^{i-1} |\mathcal{D}_j| \times r - 1\right)$-dimensional simplex.

The random vectors $\varphi_1, \varphi_2, \ldots, \varphi_n$ have an intuitive interpretation.

- $\varphi_1(t_1)$ is an approximation of the information that Player 1 has on the payoffs when her type is t_1.

- $\varphi_2(t_2)$ is an approximation of the information that Player 2 has when her type is t_2 on the payoffs and on the approximated information that Player 1 has on the payoffs.

- $\varphi_3(t_3)$ is an approximation of the information that Player 3 has when her type is t_3 on (i) the payoffs, (ii) the approximated information that Player 1 has on the payoffs, and (iii) the approximated information that Player 2 has on the payoffs and on the approximated information that Player 1 has on the payoffs.

- Etc.

For every $i \in \mathcal{N}$ and every Borel function $f : \prod_{j=1}^{i-1} \mathcal{D}_j \times \mathcal{R} \to \mathbb{R}$, let $\mathbb{E}_{\varphi_i}[f]$ be the random variable given by

$$\mathbb{E}_{\varphi_i}[f](t_i) \equiv \sum_{z \in \prod_{j=1}^{i-1} \mathcal{D}_j \times \mathcal{R}} f(z) \cdot \varphi_i(z)(t_i), \quad \mathbb{P}\text{-a.s.}$$

The next lemma relates $\mathbb{E}_{\varphi_i}[f]$ to the conditional expectation of f given t_i.

Lemma 1 *Fix $i \in \mathcal{N}$ and let $f : \prod_{j=1}^{i-1} \mathcal{D}_j \times \mathcal{R} \to \mathbb{R}$ be a real-valued Borel function which is bounded by $M > 0$. Then,*

$$\left|\mathbb{E}\left[f(\psi_i) \mid t_i\right] - \mathbb{E}_{\varphi_i}[f](t_i)\right| < M\delta, \quad \mathbb{P}\text{-}a.s.$$

Proof: The claim holds since $\mathbb{P}$-a.s. we have

$$\begin{aligned}\left|\mathbb{E}\left[f(\psi_i) \mid t_i\right] - \mathbb{E}_{\varphi_i}[f](t_i)\right| &\leq \sum_{z \in \prod_{j=1}^{i-1} \mathcal{D}_j \times \mathcal{R}} |f(z)| \cdot \left|\mathbb{P}\left(\psi_i = z \mid t_i\right) - \varphi_i(z)(t_i)\right| \\ &\leq M \sum_{z \in \prod_{j=1}^{i-1} \mathcal{D}_j \times \mathcal{R}} \left|\mathbb{P}\left(\psi_i = z \mid t_i\right) - \varphi_i(z)(t_i)\right| < M\delta,\end{aligned}$$

where the second inequality holds by (8) and the assumption that f is bounded by M. ∎

For each $i \in \mathcal{N}$, let

$$\tau_i \equiv \tau_i(t_{i:n}) \equiv \varphi_{i:n}(t_{i:n}) = \left(\varphi_i(t_i), \varphi_{i+1}(t_{i+1}), \ldots, \varphi_n(t_n)\right).$$

Intuitively, τ_i represents the approximated information of Player i: Player i knows φ_i, and since information is nested, she also knows $\varphi_{i+1}, \ldots, \varphi_n$. The following result, which follows by the construction of $(\tau_i)_{i \in \mathcal{N}}$, details basic properties of this approximated information structure.

Lemma 2

P1. *For each $i \in \mathcal{N}$, τ_i has a finite image.*

P2. *For every $1 \leq i \leq j \leq n$, τ_j is determined by τ_i.*

P3. *For each $i \in \mathcal{N}$, $\tau_i(t_{i:n})$ is determined by t_i.*

P4. *For each $i \in \mathcal{N}$, $\varphi_i(t_i)$ is determined by $\tau_i(t_{1:n})$, $\mathbb{P}$-a.s.*

Proof: **P1**, **P2**, and **P4** follow from the construction. We show that **P3** holds as well. Fix then $i \in \mathcal{N}$, and recall that since the information is nested, $t_{i:n}$ is determined by t_i. Thus, $\tau_i(t_{1:n})$ is determined by t_i and hence **P3** follows. ∎

Remark 8 Due to **P3**, from now on, for each $i \in \mathcal{N}$, we shall write $\tau_i \equiv \tau_i(t_i) \equiv \tau_i(t_{i:n})$.

6.1.3 Proof: existence of Bayesian ε-equilibrium

We now define a Bayesian game $\widetilde{\Gamma}$, which is similar to Γ, except that the information available to each player $i \in \mathcal{N}$ is τ_i rather than t_i. Specifically, $\widetilde{\Gamma}$ is given by

- A finite set of players $\widetilde{\mathcal{N}} \equiv \mathcal{N} = \{1, 2, \ldots, n\}$.

- For every $i \in \mathcal{N}$, the set of types of Player i is the range of $\tau_i = \varphi_{i:n}$, i.e., $\mathcal{T}_i \equiv \prod_{j=i}^{n} \mathcal{D}_j$ and hence $\widetilde{\mathcal{T}} \equiv \prod_{i \in \mathcal{N}} \widetilde{\mathcal{T}}_i$.

- A common prior distribution $\widetilde{\mathbb{P}}$ on $\widetilde{\mathcal{T}}$, which is the push-forward probability measure induced by $\tau_{1:n}$ with respect to the probability measure $\mathbb{P}$:

  $$\widetilde{\mathbb{P}}(\widetilde{t}) \equiv \mathbb{P}(\tau_{1:n} = \widetilde{t}), \quad \forall \widetilde{t} \in \widetilde{\mathcal{T}}.$$

  Denote by $\widetilde{\mathbb{E}}[\cdot]$ the expectation operator that corresponds to $\widetilde{\mathbb{P}}$.

- For each $i \in \mathcal{N}$, the set of actions available to Player i is $\widetilde{\mathcal{A}}_i \equiv \mathcal{A}_i$, so that $\widetilde{\mathcal{A}} = \mathcal{A}$.

- For each $i \in \mathcal{N}$, a measurable payoff function $\widetilde{R}_i : \widetilde{\mathcal{T}} \times \mathcal{A} \to \mathbb{R}$ given by

$$\widetilde{R}_i\left(\widetilde{t}, a\right) \equiv \mathbb{E}\left[R_i(a) \mid \tau_{1:n} = \widetilde{t}\,\right],$$

  for every $a \in \mathcal{A}$ and $\widetilde{t} \in \widetilde{\mathcal{T}}$ for which $\widetilde{\mathbb{P}}(\widetilde{\mathcal{T}} = \widetilde{t}) > 0$. When $\widetilde{\mathbb{P}}(\widetilde{\mathcal{T}} = \widetilde{t}) = 0$, the definition of $\widetilde{R}_i(\widetilde{t}, a)$ is irrelevant. For each $i \in \mathcal{N}$, denote by $\widetilde{R}_i(a) : \mathcal{T} \to \mathbb{R}$ the a-section of $\widetilde{R}_i$, for each $a \in \mathcal{A}$; and by $\widetilde{R}_i(\widetilde{t}) : \mathcal{A} \to \mathbb{R}$ the $\widetilde{t}$-section of $\widetilde{R}_i$, for each $\widetilde{t} \in \mathcal{T}$. We also set $\widetilde{R} \equiv (\widetilde{R}_i)_{i\in\mathcal{N}}$.

The tower rule implies that for every $a \in \mathcal{A}$ and $i \in \mathcal{N}$,

$$\widetilde{\mathbb{E}}[\widetilde{R}_i(\widetilde{t}, a)] = \mathbb{E}[R_i(t, a)].$$

For each $i \in \mathcal{N}$, $\tau_{i:n}$ is determined by τ_i, and hence for every $i \in \mathcal{N}$, $a \in \mathcal{A}$, and $\widetilde{t} = (\widetilde{t}_i)_{i\in I} \in \widetilde{\mathcal{T}}$ for which $\widetilde{\mathbb{P}}(\widetilde{t}) > 0$, the tower rule also implies that

$$\widetilde{\mathbb{E}}\left[\widetilde{R}_i(a) \mid \widetilde{t}_i\right] = \mathbb{E}\left[R_i(a) \mid \tau_i = \widetilde{t}_i\right]. \tag{9}$$

This means that the expected payoff of Player i in $\widetilde{\Gamma}$ given her type $\widetilde{t}_i$ is the same as her expected payoff in Γ given the event $\{\tau_i = \widetilde{t}\}$.

A strategy of Player i in $\widetilde{\Gamma}$ is a function $\widetilde{s}_i : \widetilde{T}_i \to \Delta(\mathcal{A}_i)$. Such a function can be interpreted as a strategy in Γ that is determined by τ_i (which, in turn, is determined by t_i). Together with (9), this implies that a strategy profile $\widetilde{s} = (\widetilde{s}_i)_{i\in\mathcal{N}}$ is a Bayesian 0-equilibrium in $\widetilde{\Gamma}$ if and only if for every $i \in \mathcal{N}$ and every strategy $s_i \in \mathcal{S}_i$ determined by τ_i,

$$U_i(\widetilde{s})(t_i) = \mathbb{E}\left[R_i(\widetilde{s}) \mid \tau_i\right] \geq \mathbb{E}\left[R_i(s_i, \widetilde{s}_{-i}) \mid \tau_i\right] = U_i(s_i, \widetilde{s}_{-i})(t_i), \quad \mathbb{P}\text{-a.s.} \tag{10}$$

Since the sets of types in $\widetilde{\Gamma}$ are finite, this game admits a Bayesian 0-equilibrium.

The next lemma concerns the original game Γ, and states that if each Player $j \in \mathcal{N} \setminus \{i\}$ adopts a strategy s_j which is determined by τ_j (i.e., a strategy which is also feasible to her in $\widetilde{\Gamma}$), then Player i has a $(\delta M|\mathcal{A}|)$-best response which is determined by τ_i (i.e., a $\delta M|\mathcal{A}|$)-best response in Γ which is also feasible to her in $\widetilde{\Gamma}$). In view of (10), this implies that every Bayesian 0-equilibrium in $\widetilde{\Gamma}$ is a Bayesian $(\delta M|\mathcal{A}|)$-equilibrium in Γ.

Lemma 3 *Let $i \in \mathcal{N}$ be a player, and let $s_{-i} \equiv (s_j)_{j \in \mathcal{N}_{-i}} \in \mathcal{S}_{-i}$ be a strategy profile such that s_j is determined by τ_j for every $j \in \mathcal{N} \setminus \{i\}$. Then $\sup_{s_i \in \mathcal{S}_i} U_i(s_i, s_{-i} \mid t_i)$ is a random variable, and there exists $s_i^* \in \mathcal{S}_i$ which is determined by τ_i such that*

$$U_i(s_i^*, s_{-i} \mid t_i) \geq \sup_{s_i \in \mathcal{S}_i} U_i(s_i, s_{-i} \mid t_i) - \delta M \left| \mathcal{A} \right|, \quad \mathbb{P}\text{-}a.s.$$

Proof: Player i's (random) expected payoff given her information, when she selects action $a_i \in \mathcal{A}_i$ and the other players follow the strategy profile s_{-i}, is the random variable $m_i(a_i)$ determined by t_i and defined as

$$m_i(a_i)(t_i) \equiv \sum_{a_{-i} \in \mathcal{A}_{-i}} \mathbb{E}\left[R_i(a_i, a_{-i}) p_{s_{-i}}(a_{-i}) \mid t_i \right], \quad \mathbb{P}\text{-a.s.} \tag{11}$$

Since $\mathcal{A}_i$ is a finite set, the best payoff that Player i can achieve is the random variable $\widetilde{m}_i$ determined by t_i and defined by

$$\widetilde{m}_i(t_i) \equiv \max \left\{ m_i(a_i)(t_i) \ : \ a_i \in \mathcal{A}_i \right\}, \quad \forall t_i \in \mathcal{T}_i.$$

In particular, $\sup_{s_i \in \mathcal{S}_i} U_i(s_i, s_{-i} \mid t_i) = \widetilde{m}_i$ is a random variable.

Let us show that for each $a_i \in \mathcal{A}_i$, there exists a random variable $\widehat{m}_i(a_i)$ which is determined by τ_i and such that

$$\left| \widehat{m}_i(a_i) - m_i(a_i) \right| < \delta M \left| \mathcal{A} \right|, \quad \mathbb{P}\text{-a.s.} \tag{12}$$

Before proving the existence of such random variables $(\widehat{m}_i(a_i))_{a_i \in \mathcal{A}_i}$, we will show how the lemma follows from their existence. First, denote

$$\widehat{m}_i \equiv \max \left\{ \widehat{m}_i(a_i) \ : \ a_i \in \mathcal{A}_i \right\}, \tag{13}$$

which is a random variable determined by τ_i, and notice that by (12), $\left| m_i - \widehat{m}_i \right| < \delta M \left| \mathcal{A} \right|$, $\mathbb{P}$-a.s. Each of the random variables $(\widehat{m}_i(a_i))_{a_i \in \mathcal{A}_i}$ is determined by τ_i, and hence the set of maximizers $\arg\max \left\{ \widehat{m}_i(a_i); a_i \in \mathcal{A}_i \right\}$ is finite and also determined by τ_i. Therefore, there is a Borel selector[9] A_i of $\arg\max \left\{ \widehat{m}_i(a_i); a_i \in \mathcal{A}_i \right\}$, that is,

$$A_i \in \arg\max \left\{ \widehat{m}_i(a_i) \ : \ a_i \in \mathcal{A}_i \right\}, \quad \mathbb{P}\text{-a.s.},$$

[9]Let X, Y be topological spaces and $\Phi : X \rightrightarrows Y$ a correspondence (a set-valued mapping). A *Borel selector* of Φ is a Borel mapping $f : X \to Y$ such that $f(x) \in \Phi(x)$ for all $x \in X$.

which is determined by τ_i; hence A_i satisfies the requirements of the lemma.

We turn to prove the existence of $\widehat{m}_i(a_i)$. Recall that $\psi_i = (\varphi_1, \varphi_2, \ldots, \varphi_{i-1}, R)$. The strategy s_j is determined by τ_j for every $j \in \mathcal{N} \setminus \{i\}$, and by **P2** it is also determined by τ_1. Therefore, $p_{s_{-i}}(a_{-i})$ is also determined by $\tau_1 = \varphi_{1:n}$, for every $a_{-i} \in \mathcal{A}_{-i}$. By (11), there exists a measurable function $f_{a_{-i}} : \mathrm{Supp}(\psi_i, \varphi_{i:n}) \to [-M, M]$ such that

$$m_i(a_i)(t_i) = \sum_{a_{-i} \in \mathcal{A}_{-i}} \mathbb{E}\left[f_{a_{-i}}(\psi_i, \varphi_{i:n}) \mid t_i\right], \quad \mathbb{P}\text{-a.s.} \tag{14}$$

By definition, $\tau_i = \varphi_{i:n}$. Hence, by **P3**, $\varphi_{i:n}$ is also determined by t_i. Therefore,

$$m_i(a_i)(t_i) = \sum_{a_{-i} \in \mathcal{A}_{-i}} \int \mathbb{E}\left[f_{a_{-i}}(\psi_i, x) \mid t_i\right] \chi_{\varphi_{i:n}(t_i)}(\mathrm{d}x), \quad \mathbb{P}\text{-a.s.}, \tag{15}$$

where $\chi_{\varphi_{i:n}(t_i)}$ is the Dirac measure concentrated at $\{\varphi_{i:n}(t_i)\}$. Define

$$\begin{aligned} \widehat{m}_i(a_i)(t_i) &\equiv \sum_{a_{-i} \in \mathcal{A}_{-i}} \int \mathbb{E}_{\varphi_i}\left[f_{a_{-i}}(\,\cdot\,, x)\right](t_i)\, \chi_{\varphi_{i:n}(t_i)}\, \mathrm{d}x \\ &= \sum_{a_{-i} \in \mathcal{A}_{-i}} \mathbb{E}_{\varphi_i}\left\{f_{a_{-i}}\left[\,\cdot\,, \varphi_{i:n}(t_i)_{i:n}(t_i)\right]\right\}(t_i), \quad \mathbb{P}\text{-a.s.}, \end{aligned} \tag{16}$$

and notice that $\widehat{m}_i(a_i)(t_i)$ is determined by $\varphi_i(t_i)$ and $\varphi_{i:n}(t_i)$, for every realization of types profile $t \in \mathcal{T}$. Since both φ_i and $\varphi_{i:n}(t_i)$ are determined by τ_i, $\widehat{m}_i(a_i)$ is also determined by τ_i. Finally, by (15), (16), and Lemma 1,

$$\begin{aligned} &\left|\widehat{m}_i(a_i)(t_i) - m_i(a_i)(t_i)\right| \\ &< \sum_{a_{-i} \in \mathcal{A}_{-i}} \int \left|\mathbb{E}\left[f_{a_{-i}}(\psi_i, x) \mid t_i\right] - \mathbb{E}_{\varphi_i}\left[f_{a_{-i}}(\,\cdot\,, x)\right](t_i)\right| \chi_{\varphi_{i:n}(t_i)}(\mathrm{d}x) \\ &< \delta M |\mathcal{A}_{-i}|, \quad \mathbb{P}\text{-a.s.}, \end{aligned}$$

and (12) follows. ■

6.2 Existence of a Bayesian 0-equilibrium

In this section we derive Theorem 1 from the existence of Bayesian ε-equilibria for $\varepsilon > 0$, which exist by the results in Section 6.1. We will fix a sequence $(s^k)_{k \in \mathbb{N}}$ of

Bayesian $\frac{1}{k}$-equilibria, and show that this sequence has a measurable accumulation point which is a Bayesian 0-equilibrium. In Section 6.2.1 we present tools related to the existence of measurable selections and their integration. In Section 6.2.2, we represent the conditional expected payoff in a useful way. In Section 6.2.3, we define the correspondences $(\Psi_i)_{i\in\mathcal{N}}$ of the accumulation points of $(s^k)_{k\in\mathbb{N}}$ and study some of their properties, In Section 6.2.4, we characterize Bayesian 0-equilibria in terms of $(\Psi_i)_{i\in\mathcal{N}}$. In Sections 6.2.5 and 6.2.6 we show that there exist measurable selections of $(\Psi_i)_{i\in\mathcal{N}}$ satisfying the characterization of the 0-equilibrium presented in Section 6.2.4.

6.2.1 Selectors and Integration

Let X, Y be standard Borel spaces,[10] and $\Phi : X \rightrightarrows Y$ a correspondence (a set-valued mapping). We say that Φ has nonempty compact values if $\Phi(x)$ is nonempty and compact for every $x \in X$. The following classical result provides topological conditions that guarantee the existence of a Borel selector.[11]

Theorem 2 (Kuratowski and Ryll-Nardzewski [35]) *Suppose the correspondence $\Phi : X \rightrightarrows Y$ has a Borel graph and nonempty compact values. Then, Φ has a Borel selector.*

Let $\mathbf{S}_\Phi$ denote the collection of all Borel selectors of Φ. Suppose Y is a subset of a Euclidean space, and let $\mathbb{P}$ be a finite Borel measure on X. The *Aumann integral* of Φ (with respect to $\mathbb{P}$) is

$$\int_X \Phi(x)\mathbb{P}(dx) \equiv \left\{ \int_X f(x)\mathbb{P}(dx) : f \in \mathbf{S}_\Phi \right\}.$$

The following result appears in [3] and the references therein when $X = [0, 1]$ and $\mathbb{P}$ is the Lebesgue measure; the general case follows by minor modifications, or from more general results, like Theorem 4 below.

[10] A *standard Borel space* is a topological space homeomorphic to a Borel subset of a Polish space.

[11] [35] states the measurability assumption on Φ in a way that, for nonempty compact valued correspondences, is equivalent to the one we provided here; see, e.g., [25].

Theorem 3 *Suppose the correspondence* $\Phi : X \rightrightarrows \mathbb{R}^n$ *is bounded*[12] *and has a Borel graph and nonempty compact values. Then* $\int_X \Phi(x)\mathbb{P}(dx)$ *is nonempty and compact.*

We need the following slight generalization of Theorem 3.

Proposition 1 *Let* X *be a standard Borel space, and let* Y *be a compact metric space. Suppose the correspondence* $\Phi\colon X \rightrightarrows Y$ *has a Borel graph and nonempty compact values. Let* $\zeta\colon Y \to \mathbb{R}^n$ *be continuous. Then the set*

$$\left\{ \int_X \zeta \circ f(x)\mathbb{P}(dx) \colon f \in \boldsymbol{S}_\Phi \right\} \subset \mathbb{R}^n \tag{17}$$

is nonempty and compact.

Proof: Define $\Psi : X \to \mathbb{R}^n$ by $\Psi = \zeta \circ \Phi$, i.e., $\Psi(\cdot) = \zeta(\Phi(\cdot))$. Since Y is compact, ζ is continuous, and Φ has nonempty compact values, it follows that Ψ is bounded and has nonempty compact values. We contend that

$$\{\zeta \circ f \colon f \in \mathbf{S}_\Phi\} = \mathbf{S}_\Psi,$$

from which the proposition will follow due to Theorem 3. Clearly, $\{\zeta \circ f \colon f \in \mathbf{S}_\Phi\} \subseteq \mathbf{S}_\Psi$, as $\zeta \circ f \in \mathbf{S}_\Psi$ for each $f \in \mathbf{S}_\Phi$. Conversely, suppose $g \in \mathbf{S}_\Psi$. Since Y, the domain of ζ, is a nonempty compact set and ζ is continuous, the correspondence $\zeta^{-1}(\cdot)\colon \Psi(X) \to Y$ has a Borel graph and nonempty compact values. Thus, Theorem 2, applied to the correspondence $\zeta^{-1}(\cdot)$, yields a Borel mapping $\zeta'\colon \text{Image}(\zeta) \to Y$ such that $\zeta \circ \zeta' = \text{id}$. Hence, $f := \zeta' \circ g$ satisfies $g = \zeta \circ f$ and $f \in \mathbf{S}_\Phi$. ■

When $(f_k)_{k=1}^\infty$ is a sequence of mappings between two topological spaces X and Y, we denote by $\overline{\text{Lim}}((f_k)_k)\colon X \rightrightarrows Y$ the correspondence such that $\overline{\text{Lim}}((f_k)_k)(x)$ is the set of all accumulation points of $(f_k(x))_{k=1}^\infty$, for every $x \in X$. When X and Y are standard Borel spaces with Y compact, this correspondence has a Borel graph with nonempty compact values, see [40, Prop. 10.1].

[12] That is, there is a bounded $W \subseteq \mathbb{R}^n$ such that $\Phi(x) \subseteq W$ for every $x \in X$.

Lemma 4 *Let $(X, \mathbb{P})$ be a standard Borel measure space, let Y be a compact metrizable space, and for each $k \in \mathbb{N}$, let $f_k \colon X \to Y$ be measurable. Let $\zeta \colon Y \to \mathbb{R}^n$ be continuous, and suppose that*

$$\int_X \zeta(f_k(x))\mathbb{P}(dx) \xrightarrow[k\to\infty]{} z^*.$$

Then there is a Borel selector $f^* \colon X \to Y$ of $\overline{\mathrm{Lim}}((f_k)_k)$ such that

$$z^* = \int_X \zeta(f^*(x))\mathbb{P}(dx).$$

When $Y \subseteq \mathbb{R}^n$ and $\zeta = \mathrm{id}$, Lemma 4 was proven in, e.g., [25, p. 69], or [4].[13]

Proof: Denote for simplicity $\mathcal{L} = \overline{\mathrm{Lim}}((f_k)_k)$ and $\widehat{\mathcal{L}} = \overline{\mathrm{Lim}}((\zeta \circ f_k)_k)$. Applying the result in the restricted case $Y \subseteq \mathbb{R}^n$ and $\zeta = \mathrm{id}$ to the series $(\zeta \circ f_k)_k$ shows that there is Borel selector g^* of $\widehat{\mathcal{L}}$ such that

$$z^* = \int_X g^*(x)\mathbb{P}(dx).$$

We claim that[14]

$$\widehat{\mathcal{L}}(x) \subseteq \zeta(\mathcal{L}(x)), \quad \forall x \in X. \tag{18}$$

Indeed, if $y \in \widehat{\mathcal{L}}(x)$, then there are indices $(k_l)_{l\in\mathbb{N}}$ such that $\lim_{l\to\infty} \zeta(f_{k_l}(x)) = y$. Letting z be an accumulation point of $(f_{k_l}(x))_{l\in\mathbb{N}}$, which exists by compactness and metrizability of Y, we deduce by the continuity of ζ that $z \in \mathcal{L}(x)$ and $y = \zeta(z)$.

As in the proof of Proposition 1, the correspondence ζ^{-1} has a Borel graph and nonempty compact values, and hence so does $\zeta^{-1}(g^*(\cdot))$. Thus, Eq. (18) implies that for every $x \in X$, $\zeta^{-1}(g^*(x))$ is a nonempty compact subset of $\mathcal{L}(x)$. It remains to apply Theorem 2 to the correspondence $\zeta^{-1}(g^*(\cdot))$. ■

The following result is a slight generalization of the Measurable "Measurable Choice" Theorem from [40], adapted to a bounded Borel setting;[15] [40] deals with the case $W \subseteq \mathbb{R}^n$, $\zeta = \mathrm{id}$, and $B = Y \times Z$.

[13] [4] addresses the case where $\mathbb{P}$ is non-atomic; the case where $\mathbb{P}$ may have atoms follows by passing to a sequence $(f_k)_{k=1}^{\infty}$ which converges on atoms, using a diagonalization construction.

[14] In fact, there is equality in (18). However, we only need this inclusion.

[15] A similar but weaker result is proven in [2], which only gives an "almost everywhere" type of selection.

Theorem 4 *Let Y and Z be Borel spaces, and let F be a correspondence from a Borel set $B \subseteq Y \times Z$ to a compact metric space W, with nonempty compact values and a Borel graph. Let $\zeta\colon W \to \mathbb{R}^n$ be continuous, and let q be a Borel transition kernel from Y to Z,*[16] *such that $q(B_y \mid y) = 1$ for each $y \in Y$, where $B_y \equiv \{z \in Z\colon (y,z) \in B\}$ is the y-section of B. Define a correspondence $F^\diamond\colon Y \rightrightarrows \mathbb{R}^n$ by*

$$F^\diamond(y) \equiv \left\{ \int_{B_y} \zeta \circ f(y,z) q(dz \mid y)\colon f \in \boldsymbol{S}_{F(y,\cdot)} \right\}. \tag{19}$$

Then:

- *$F^\diamond$ is bounded, and has nonempty compact values and a Borel graph* $\mathrm{Gr}(F^\diamond)$.
- *There is a Borel mapping $g\colon \mathrm{Gr}(F^\diamond) \times Z \to \mathbb{R}^n$ such that for each $(y,u) \in \mathrm{Gr}(F^\diamond)$ and each $z \in B_y$, we have $g(y,u,z) \in \zeta \circ F(y,z)$ and*

$$u = \int_Z \zeta \circ g(y,u,s) q(ds \mid y).$$

Proof: As mentioned, [40] proved the case $B = Y \times Z$ (in which case $B_y = Z$ for each $y \in Y$), $W \subseteq \mathbb{R}^n$, and $\zeta = \mathrm{id}$. To prove the general case, fix $x_0 \in \mathbb{R}^n$, and define $\widetilde{F} : Y \times Z \rightrightarrows \mathbb{R}^n$ by $\widetilde{F} = \zeta \circ F$ on B, and $\widetilde{F} \equiv \{x_0\}$ outside of B. The argument from the proof of Proposition 1, together with the fact that $q(B_y \mid y) = 1$ for each $y \in Y$, shows that

$$F^\diamond(y) = \left\{ \int_Z \widetilde{f}(y,z) q(dz \mid y)\colon \widetilde{f} \in \mathbf{S}_{\widetilde{F}(y,\cdot)} \right\}.$$

Since W is compact, $F^\diamond$ is bounded, and by Proposition 1, it has nonempty compact values. By [40], $F^\diamond$ is measurable, and hence the first bullet holds.

By [40] once again, $F^\diamond$ has a Borel selector, i.e., a Borel mapping $\widetilde{g}\colon \mathrm{Gr}(F^\diamond) \times Z \to \mathbb{R}^n$ such that $g(y,u,z) \in \widetilde{F}(y,z)$ for each $(y,u) \in \mathrm{Gr}(F^\diamond)$ and $z \in Z$ (and hence $g(y,u,z) \in \zeta \circ F(y,z)$ for $z \in B_y$), and

$$u = \int_Z \widetilde{g}(y,u,s) q(ds \mid y).$$

By applying the same arguments as in the proof of Lemma 4, there is Borel mapping $\zeta'\colon \mathrm{Image}(\zeta) \to W$ such that $\zeta \circ \zeta' = \mathrm{id}$. Setting $g = \zeta' \circ \widetilde{g}$ yields the mapping indicated in the second bullet. ■

[16] A *transition kernel* is a map from Y to $\Delta(Z)$, the space of Borel probability distributions on Z, such that for each Borel $B \subseteq Y$, the mapping $z \to q(B \mid z)$ is Borel.

6.2.2 Conditional Expected Payoffs

Recall that M is a bound on the payoffs in the game. For each $i \in \mathcal{N}$, set $\mathcal{P}_{i:n} \equiv [-M, M]^{(n-i+1)\times|\mathcal{A}_{i:n}|}$, so that $R(t) \in \mathcal{P}_{1:n}$ for every $t \in \mathcal{T}$. Each vector $\rho_{i:n} \in \mathcal{P}_{i:n}$ corresponds to a vector of payoff functions for the set of players $[i : n]$, where the players' actions are $(\mathcal{A}_j)_{j=i}^n$. For such a vector, we denote by $\rho_j(a_{i:n})$ the coordinate that corresponds to Player j and to the action profile $a_{i:n}$. The multilinear extension of ρ_j is still denoted by ρ_j, so that

$$\rho_j(x_{i:n}) \equiv \sum_{a_{i:n}\in\mathcal{A}_{i:n}} \rho_j(a_{i:n}) \prod_{k=i}^{n} x_k(a_k), \quad \forall x_{i:n} \in \mathcal{X}_{i:n}.$$

Denote by $U_j(s_{1:i}, a_{i+1:n} \mid t_k)$ the expected payoff of Player j when players $[1 : i]$ follow the strategies $s_{1:i}$ and players $[i + 1 : n]$ select the actions $a_{i+1:n}$, given that Player k's type is t_k. It will be convenient to denote by

$$U_{i+1:n}(s_{1:i}, \cdot \mid t_{i+1}) \equiv \big(U_{i+1:n}(s_{1:i}, a_{i+1:n} \mid t_{i+1})\big)_{a_{i+1:n}\in\mathcal{A}_{i+1:n}} \in \mathcal{P}_{i+1:n}$$

the payoff function of players $[i + 1 : n]$ that is induced by the strategy profile $s_{1:i}$ and the type t_{i+1}. The multilinear extension of $U_{i+1:n}(s_{1:i}, \cdot \mid t_{i+1})$ is still denoted by $U_{i+1:n}(s_{1:i}, \cdot \mid t_{i+1})$, and it is a mapping from $\mathcal{X}_{i+1:n}$ to $\mathbb{R}^{n-i}$.

6.2.3 The Limit of $\frac{1}{k}$-Equilibria: The Operator Ψ_i

Since information is nested, t_i determines $t_{i+1}, t_{i+2}, \ldots, t_n$. For convenience, when $j > i$ and s_j is a strategy of Player j, we will sometimes write $s_j(t_i)$ instead of $s_j(\kappa_{j-1} \circ \kappa_{j-2} \circ \cdots \circ \kappa_i(t_i))$.

Fix a sequence $(s^k)_{k=1}^\infty$ of strategy profiles such that $s^k \equiv (s_1^k, \ldots, s_n^k)$ is a $\frac{1}{k}$-Bayesian equilibrium, for each $k \in \mathbb{N}$, which in turn is guaranteed by the result obtained in Section 6.1. We would like to prove that an appropriate limit of $(s^k)_{k=1}^\infty$ is a Bayesian 0-equilibrium. To this end, we will consider for each $i \in \mathcal{N}$ the accumulation points of the sequence $\Big(s_{i:n}^k(t_i)), U_{i:n}(s^k, \cdot \mid t_i), U_{i+1:n}(s^k, \cdot \mid \kappa_i(t_i))\Big)_{k=1}^\infty$, and show that proper Borel selectors of these correspondences (indexed by i) induce a Bayesian 0-equilibrium.

Define a correspondence $\Psi_n : \mathcal{T}_n \rightrightarrows \mathcal{X}_n \times \mathcal{P}_n$ by

$$\Psi_n(t_n) \equiv \left\{(x_n, \rho_n) \in \mathcal{X}_n \times \mathcal{P}_n : (x_n, \rho_n) \in \overline{\mathrm{Lim}}\Big(\Big(s_n^k(t_n), U_n(s^k \mid t_n)\Big)_k\Big)\right\}, \tag{20}$$

and, for every $i < n$, define the correspondence $\Psi_i : \mathcal{T}_i \times \mathcal{X}_{i+1:n} \times \mathcal{P}_{i+1:n} \rightrightarrows \mathcal{X}_{i:n} \times \mathcal{P}_{i:n}$ by

$$\begin{aligned} &\Psi_i(t_i, x_{i+1:n}, \rho_{i+1:n}) \equiv \qquad (21)\\ &\Big\{((x_i, x_{i+1:n}), \widehat{\rho}_{i:n}) \in \mathcal{X}_{i:n} \times \mathcal{P}_{i:n} \\ &\quad : ((x_i, x_{i+1:n}), \widehat{\rho}_{i:n}, \rho_{i+1:n}) \in \overline{\mathrm{Lim}}\Big(\Big(s_{i:n}^k(t_i), U_{i:n}(s_{1:i-1}^k, \cdot \mid t_i), U_{i+1:n}(s_{1:i}^k, \cdot \mid \kappa_i(t_i))\Big)_k\Big)\Big\}. \end{aligned}$$

Note that Ψ_i may have empty values. This happens, for example, when $x_{i+1:n}$ is not an accumulation point of $(s_{i+1:n}^k(t_i))_{k=1}^\infty$. The definition implies that if $(\widehat{x}_{i:n}, \widehat{\rho}_{i:n}) \in \Psi_i(t_i, x_{i+1:n}, \rho_{i+1:n})$, then $\widehat{x}_{i+1:n} = x_{i+1:n}$. The relation between $\widehat{\rho}_{i:n}$ and $\rho_{i+1:n}$ is more complex and we will not need it.

The following lemma holds from the definitions and since each $\mathcal{P}_i$ is compact.

Lemma 5 *For $i \in \mathcal{N}$, Ψ_i has a Borel graph and nonempty compact values.*

If $(x_{1:n}, \widehat{\rho}_{1:n}) \in \Psi_1(t_1, x_{2:n}, \rho_{2:n})$, then $\widehat{\rho}_{1:n}$ is a vector of payoff functions for all players. In particular, the j's coordinate $\widehat{\rho}_j$ of $\widehat{\rho}_{1:n}$ satisfies[17]

$$\widehat{\rho}_j(a) = U_j(a \mid t_1) = R_j(t, a), \tag{22}$$

for all $a \in \mathcal{A}_{1:n}$, where $t = (t_1, \kappa_1(t_1), \ldots, \kappa_{n-1} \circ \cdots \circ \kappa_2 \circ \kappa_1(t_1))$.

6.2.4 Equilibrium Characterization via $\Psi_1, \Psi_2, \ldots, \Psi_n$

In this section we will provide a characterization of Bayesian 0-equilibria in terms of the mappings $\Psi_1, \Psi_2, \ldots, \Psi_n$.

Lemma 6 *For each $i \in \mathcal{N}$, $\mathbb{P}$-a.e. $t_i \in \mathcal{T}_i$, $x_{i+1:n} \in \mathcal{X}_{i+1:n}$, and $\rho_{i+1:n} \in \mathcal{P}_{i+1:n}$, if $(x_{i:n}, \widehat{\rho}_{i:n}) \in \Psi_i(t_i, x_{i+1:n}, \rho_{i+1:n})$, then*

$$x_i \in \operatorname*{argmax}_{y_i \in \mathcal{X}_i} \widehat{\rho}_i(y_i, x_{i+1:n}). \tag{23}$$

[17] Below, $U_j(a \mid t_1) = U_j(s \mid t_1)$, where $s = (s_i)_{i \in \mathcal{N}}$, is the strategy profile in which each player selects the action a_i at all types.

The expression under the argmax on the right-hand side of (23) is the expected payoff of Player i under the payoff function $\widehat{\rho}_i$ when Players $[i:n]$ use mixed actions $y_i, x_{i+1}, \ldots, x_n$, respectively. Lemma 6, which follows by continuity arguments, is the only place in the proof where the fact that s^k is a $\frac{1}{k}$-equilibrium, for each $k \in \mathbb{N}$, is directly used.

Proof: Fix a player $i < n$, $t_i \in \mathcal{T}_i$, $x_{i+1:n} \in \mathcal{X}_{i+1:n}$, $\rho_{i+1:n} \in \mathcal{P}_{i+1:n}$, and $\widehat{\rho}_{i:n} \in \mathcal{P}_{i:n}$ such that $(x_i, \widehat{\rho}_{i:n}) \in \Psi_i(t_i, x_{i+1:n}, \rho_{i+1:n})$. By assumption, there is a sequence of indices (k_l) such that

$$\lim_{l\to\infty} \left(s^{k_l}_{i:n}(t_i), U_{i:n}(s^{k_l}_{1:i-1}, \cdot | t_i) \right) = ((x_i, x_{i+1:n}), \widehat{\rho}_{i:n}).$$

For each $l = 1, 2, \ldots$, the expected payoff of Player i with type t_i, playing mixed action $y_i \in \mathcal{X}_i$ while the others follow strategy profile s^{k_l} is,

$$U_i(s^{k_l}_{-i}, y_i \mid t_i) = \sum_{a_i \in \mathcal{A}_i} \sum_{a_{i+1:n} \in \mathcal{A}_{i+1:n}} y_i(a_i) \cdot \left(\prod_{j=i+1}^{n} s^{k_l}_j(t_i)(a_j) \right) \cdot U_i(s^{k_l}_{1:i-1}, a_{i:n} \mid t_i).$$

Since s^{k_l} is a $\frac{1}{k_l}$-equilibrium, there is a set $\Xi_l \subseteq \mathcal{T}_i$ with $\mathbb{P}(\Xi_l) = 0$ such that for every $t_i \notin \Xi_l$,

$$U_i(s^{k_l}_{-i}, y_i \mid t_i) \le U_i(s^{k_l} \mid t_i) + \frac{1}{k_l},$$

that is,

$$\sum_{a_i \in \mathcal{A}_i} \sum_{a_{i+1:n} \in \mathcal{A}_{i+1:n}} y_i(a_i) \cdot \left(\prod_{j=i+1}^{n} s^{k_l}_j(t_i)(a_j) \right) \cdot U_i(s^{k_l}_{1:i-1}, a_{i:n} \mid t_i) \tag{24}$$
$$\le \frac{1}{k_l} + \sum_{a_i \in \mathcal{A}_i} \sum_{a_{i+1:n} \in \mathcal{A}_{i+1:n}} s^{k_l}_i(a_i) \cdot \left(\prod_{j=i}^{n} s^{k_l}_j(t_i)(a_j) \right) \cdot U_i(s^{k_l}_{1:i-1}, a_{i:n} \mid t_i).$$

Taking $l \to \infty$ gives that for every $t_i \in \mathcal{T}_i \setminus \bigcup_{l \in \mathbb{N}} \Xi_l$,

$$\sum_{a_i \in \mathcal{A}_i} \sum_{a_{i+1:n} \in \mathcal{A}_{i+1:n}} y_i(a_i) \cdot \left(\prod_{j=i}^{n} x_j(a_j) \right) \cdot \widehat{\rho}_i(a_{i:n})$$
$$\le \sum_{a_i \in \mathcal{A}_i} \sum_{a_{i+1:n} \in \mathcal{A}_{i+1:n}} x_i(a_i) \cdot \left(\prod_{j=i}^{n} x_j(a_j) \right) \cdot \widehat{\rho}_i(a_{i:n}),$$

and the claim follows. The proof for $i = n$ is similar yet simpler, since in (24) the inner summations on the two sides are vacuous. ■

Corollary 1 *Let $s \in \mathcal{S}$ be a strategy profile. Suppose that for $\mathbb{P}$-almost every $t_n \in \mathcal{T}_n$*

$$(s_n(t_n), U_n(s_n \mid t_n)) \in \Psi_n(t_n),$$

and for each $i < n$ and $\mathbb{P}$-almost every $t_i \in \mathcal{T}_i$,

$$\big(s_{i:n}(t_i), U_{i:n}(s \mid t_i)\big) \in \Psi_i\big(t_i, s_{i+1:n}(t_i), U_{i+1:n}(s \mid t_{i+1})\big). \tag{25}$$

Then, s is a Bayesian 0-equilibrium.

6.2.5 Selections from $(\Psi_i)_{i\in\mathcal{N}}$

Recall that $\kappa_i : \mathcal{T}_i \to \mathcal{T}_{i+1}$ is the mapping that indicates the type of Player $i+1$ for each type of Player i. In particular, $\kappa_i^{-1}(t_{i+1}) \equiv \{t_i \in \mathcal{T}_i \mid \kappa_i(t_i) = t_{i+1}\}$ is the set of all types of Player i that are consistent with the type of Player $i+1$.

An element $(x_{i:n}, \rho_{i:n}) \in \mathcal{X}_{i:n} \times \mathcal{P}_{i:n}$ is a pair consisting of a mixed action vector for players $[i:n]$ and a payoff function for a game restricted to these players. It will prove convenient to denote, for each $j \in [i:n]$, Player j's payoff in this game by

$$\gamma_j(x_{i:n}, \rho_{i:n}) \equiv \rho_j(x_{i:n}),$$

and set

$$\gamma_{i:n}(x_{i:n}, \rho_{i:n}) \equiv (\gamma_j(x_{i:n}, \rho_{i:n}))_{j=i}^{n}.$$

The next lemma intuitively states that every point $(t_i, x_{i:n}, \rho_{i+1:n}, \widehat{\rho}_{i:n})$ in the graph of Ψ_i can be extended to a point in the graph of Ψ_{i-1}.

Lemma 7 *Fix $i = 2, 3, \dots, n$, $t_i \in \mathcal{T}_i$, $x_{i+1:n} \in \mathcal{X}_{i+1:n}$, and $\rho_{i+1:n} \in \mathcal{P}_{i+1:n}$. If $(x_{i:n}, \widehat{\rho}_{i:n}) \in \Psi_i(t_i, x_{i+1:n}, \rho_{i+1:n})$, then $\Psi_{i-1}(t_{i-1}, x_{i:n}, \widehat{\rho}_{i:n}) \neq \emptyset$, for each $t_{i-1} \in \kappa_{i-1}^{-1}(t_i)$; For $i = n$, the terms $x_{i+1:n}$ and $\rho_{i+1:n}$ are vacuous. Moreover, there exists a Borel mapping $f : \mathcal{T}_{i-1} \to \mathcal{X}_{i-1:n} \times \mathcal{P}_{i-1:n}$ such that*

$$f(t_{i-1}) \in \Psi_{i-1}(t_{i-1}, x_{i:n}, \widehat{\rho}_{i:n}), \quad \forall t_{i-1} \in \kappa_{i-1}^{-1}(t_i), \tag{26}$$

and

$$\widehat{\rho}_{i:n} = \int_{\mathcal{T}_{i-1}} \gamma_{i:n}(f(t_{i-1}))\mathbb{P}(dt_{i-1} \mid t_i). \tag{27}$$

Eq. (26) states that for any fixed $(x_{i:n}, \widehat{\rho}_{i:n})$, on $\kappa_{i-1}^{-1}(t_i)$, $f(\cdot)$ is a selector of the correspondence

$$t_{i-1} \rightrightarrows \Psi_{i-1}(t_{i-1}, x_{i:n}, \widehat{\rho}_{i:n}),$$

which, by Lemma 5, has a Borel graph and nonempty compact values.

Proof: Suppose that $(x_{i:n}, \widehat{\rho}_{i:n}) \in \Psi_i(t_i, x_{i+1:n}, \rho_{i+1:n})$. Then there is a sequence of indices $(k_l)_{l=1}^{\infty}$ such that

$$x_{i:n} = \lim_{l\to\infty} s_{i:n}^{k_l}(t_i) \quad \text{and} \quad \widehat{\rho}_{i:n} = \lim_{l\to\infty} U_{i:n}(s^{k_l}, \cdot \mid t_i). \tag{28}$$

Applying Lemma 4 to $X = \mathcal{T}_{i-1}$, $\mathbb{P}(\mathrm{d}x) = \mathbb{P}(\mathrm{d}t_{i-1} \mid t_i)$, $Y = \mathcal{X}_{i:n} \times \mathcal{P}_{i:n}$, $f_l(t_i) = (s_{i:n}^{k_l}(t_i), U_{i:n}(s_{i:n}^{k_l} \mid t_i)$ for each $l \in \mathbb{N}$, $\zeta = \gamma_{i:n}$, and $z^* = (x_{i:n}, \widehat{\rho}_{i:n})$, we conclude that there is a Borel selector f of the correspondence

$$t_{i-1} \rightrightarrows \overline{\mathrm{Lim}}\left(\left(s_{i-1:n}^{k_l}(t_{i-1}), U_{i-1:n}(s_{1:i-1}^{k_l}, \cdot \mid t_{i-1})\right)_l\right)$$

such that (27) holds. Condition (28) implies that

$$\emptyset \neq \overline{\mathrm{Lim}}\left(\left(s_{i-1:n}^{k_l}(t_{i-1}), U_{i-1:n}(s_{1:i-1}^{k_l}, \cdot \mid t_{i-1})\right)_l\right) \subseteq \Psi_{i-1}(t_{i-1}, x_{i:n}, \widehat{\rho}_{i:n}),$$

which completes the proof. ■

The next result is a measurable version of Lemma 7.

Lemma 8 *For each $i = 2, \ldots, n$, there is a mapping $f_{i-1} : \mathcal{T}_{i-1} \times \mathrm{Gr}(\Psi_i) \to \mathcal{X}_{i-1:n} \times \mathcal{P}_{i-1}$ such that for each $(t_i, x_{i+1:n}, \rho_{i+1:n}, x_{i:n}, \widehat{\rho}_{i:n}) \in \mathrm{Gr}(\Psi_i)$,*

$$f_{i-1}(t_{i-1}, t_i, x_{i:n}, \rho_{i+1:n}, \widehat{\rho}_{i:n}) \in \Psi_{i-1}(t_{i-1}, x_{i:n}, \widehat{\rho}_{i:n}) \ , \ \forall t_{i-1} \in \kappa_{i-1}^{-1}(t_i), \tag{29}$$

and

$$\widehat{\rho}_{i:n} = \int_{\mathcal{T}_{i-1}} \gamma_{i:n}\big(f_{i-1}(t_{i-1}, t_i, x_{i:n}, \rho_{i+1:n}, \widehat{\rho}_{i:n})\big)\mathbb{P}(dt_{i-1} \mid t_i). \tag{30}$$

For $i = n$, in both (29) and (30) the term $\rho_{i+1:n}$ is vacuous.

Remark 9 (Dependence of f on $\rho_{i+1:n}$) The dependence of f on $\rho_{i+1:n}$ seems superfluous, but may be indispensable. This dependence parallels the dependence of the equilibria in stochastic games on the previous state *and* on the current state, whose existence was proved by [41] using the Measurable "Measurable Choice" Theorem of [40].

Remark 10 By the properties of Ψ_{i-1}, if

$$(\widetilde{x}_{i-1:n}, \widetilde{\rho}_{i-1:n}) = f_{i-1}(t_{i-1}, t_i, x_{i:n}, \rho_{i+1:n}, \widehat{\rho}_{i:n}),$$

then $\widetilde{x}_{i:n} = x_{i:n}$. Later we will make use of this observation.

Proof of Lemma 8: Fix $i = 2, \ldots, n$, and apply Theorem 4 with the following parameters:

- $Y = \text{Gr}(\Psi_i)$.
- $Z = \mathcal{T}_{i-1}$.
- $B = \big\{(t_{i-1}, t_i, x_{i+1:n}, \rho_{i+1:n}, x_{i:n}, \widehat{\rho}_{i:n}) \in Z \times Y \mid t_i = \kappa_{i-1}(t_{i-1})\big\}$.
- $q(t_{i-1} \mid t_i, x_{i+1:n}, \rho_{i+1:n}, x_{i:n}, \widehat{\rho}_{i:n}) = \mathbb{P}(t_{i-1} \mid t_i)$.
- $W = \mathcal{X}_{i-1:n} \times \mathcal{P}_{i-1:n}$.
- ζ is the evaluation map $\gamma_{i:n}$ on $\mathcal{X}_{i-1:n} \times \mathcal{P}_{i-1:n}$.
- F is defined on B by $F(t_{i-1}, t_i, x_{i+1:n}, \rho_{i+1:n}, x_{i:n}, \widehat{\rho}_{i:n}) \equiv \Psi_{i-1}(t_{i-1}, x_{i:n}, \widehat{\rho}_{i:n})$.
- $F^{\diamond} : Y \to \mathcal{P}_{i:n}$ as defined in (19).

By Theorem 4, there is a Borel mapping $g : \text{Gr}(F^{\diamond}) \times \mathcal{T}_{i-1} \to \mathcal{X}_{i-1:n} \times \mathcal{P}_{i-1:n}$ such that for every $y = (t_i, x_{i+1:n}, \rho_{i+1:n}, x_{i:n}, \widehat{\rho}_{i:n}) \in Y = \text{Gr}(\Psi_i)$, every $u \in F^{\diamond}(y)$, and every $t_{i-1} \in \kappa_{i-1}^{-1}(t_i)$, we have

$$g(y, u, t_{i-1}) \in F(t_{i-1}, y) = \Psi_{i-1}(t_{i-1}, x_{i:n}, \widehat{\rho}_{i:n}),$$

and

$$u = \int_{\mathcal{T}_{i-1}} \gamma_{i:n}\big(g(y, u, t_{i-1})\big)\mathbb{P}(dt_{i-1} \mid t_i).$$

It follows from Lemma 7 that for any $y = (t_i, x_{i+1:n}, \rho_{i+1:n}, x_{i:n}, \widehat{\rho}_{i:n}) \in Y$, noting that $B_y = \kappa_{i-1}^{-1}(t_i)$ and hence $\mathbb{P}(B_y \mid t_i) = 1$, there is $f : \mathcal{T}_{i-1} \to \mathcal{P}_{i-1:n}$ such that $f|_{B_y}$ is a Borel selector of $t_{i-1} \to F(t_{i-1}, t_i, x_{i+1:n}, \rho_{i+1:n}, x_{i:n}, \widehat{\rho}_{i:n}) = \Psi_{i-1}(t_{i-1}, x_{i:n}, \widehat{\rho}_{i:n})$ and such that (27) holds, which means that

$$\widehat{\rho}_{i:n} \in F^{\diamond}(t_{i-1}, t_i, x_{i+1:n}, \rho_{i+1:n}, x_{i:n}, \widehat{\rho}_{i:n}).$$

Defining $f_{i-1} : \mathcal{T}_{i-1} \times \mathrm{Gr}(\Psi_i) \to \mathcal{X}_{i-1:n} \times \mathcal{P}_{i-1:n}$ by

$$f_{i-1}\big(t_{i-1}, t_i, x_{i:n}, \rho_{i+1:n}, \widehat{\rho}_{i:n}\big) \equiv g\big((t_i, x_{i+1:n}, \rho_{i+1:n}, x_{i:n}, \widehat{\rho}_{i:n}), \widehat{\rho}_{i:n}, t_{i-1}\big),$$

yields the desired result. ■

6.2.6 Construction of a Bayesian 0-Equilibrium

In this section we define a strategy profile $s_* = (s_*^1, \ldots, s_*^n)$, and prove that it is a Bayesian 0-equilibrium.

For any set A, denote by π_A the projection map to A. Let $f_{n-1}, \ldots, f_1$ be the mappings given by Lemma 8. Define mappings $(g_i)_{i=1}^n$ recursively (backwards) as follows.

- Let $g_n : \mathcal{T}_n \to \mathcal{X}_n \times \mathcal{P}_n$ be a Borel selector of Ψ_n, which exists by Theorem 2 (recall that Ψ_n depends only on the type of Player n). Define

$$s_n^* \equiv \pi_{\mathcal{X}_n} \circ g_n. \tag{31}$$

- For $i = 2, \ldots, n$, assuming that we have already defined mappings $(g_j)_{j=i}^n$, let $g_{i-1} : \mathcal{T}_{i-1} \to \mathcal{X}_{i-1:n} \times \mathcal{P}_{i-1:n}$ be defined by

$$g_{i-1}(t_{i-1}) \equiv f_{i-1}\Big(t_{i-1}, \kappa_{i-1}(t_{i-1}), s_{i:n}^*(\kappa_{i-1}(t_{i-1})), \pi_{\mathcal{P}_{i+1:n}} \circ g_{i+1}(\kappa_i(\kappa_{i-1}(t_{i-1}))), \pi_{\mathcal{P}_{i:n}} \circ g_i(\kappa_{i-1}(t_{i-1}))\Big), \tag{32}$$

  where the penultimate argument is vacuous when $i = n-1$, and set

$$s_{i-1}^* \equiv \pi_{\mathcal{X}_{i-1}} \circ g_{i-1}. \tag{33}$$

For $i = 1, \ldots, n-1$, the mapping g_i is well defined provided

$$\Big(\kappa_{i-1}(t_{i-1}), s_{i+1:n}^*(\kappa_{i-1}(t_{i-1})), \pi_{\mathcal{P}_{i+1:n}} \circ g_{i+1}(\kappa_i(\kappa_{i-1}(t_{i-1}))), s_{i:n}^*(\kappa_{i-1}(t_{i-1})), \pi_{\mathcal{P}_{i:n}} \circ g_i(\kappa_{i-1}(t_{i-1}))\Big)$$

always lies in $\mathrm{Gr}(\Psi_i)$. The next lemma states that this is indeed the case.

Lemma 9 *The mappings $(g_i)_{i=1}^{n-1}$ are well defined.*

Proof: Denote $t_i = \kappa_{i-1}(t_{i-1})$. We will prove that for each $t_{i-1} \in \mathcal{T}_{i-1}$,

$$\Big(t_i, s^*_{i+1:n}(t_i), \pi_{\mathcal{P}_{i+1:n}} \circ g_{i+1}(\kappa_i(t_i)), s^*_{i:n}(t_i), \pi_{\mathcal{P}_{i:n}} \circ g_i(t_i)\Big) \in \mathrm{Gr}(\Psi_i), \tag{34}$$

and

$$s^*_{i:n}(t_i) = \pi_{\mathcal{X}_{i:n}} \circ g_i(t_i). \tag{35}$$

Note the slight difference between (33) and (35): in the former we set $s^*_j(t_j)$ to be $\pi_{\mathcal{X}_j} \circ g_j(t_j)$, while the latter claims that $s^*_j(t_j) = \pi_{\mathcal{X}_j} \circ g_i(t_i)$ for each $j \geq i$; by construction, using Remark 10, these agree.

We prove (34) and (35) by induction, starting with $i = n$. In this case, (35) holds by (31). Moreover, the left-hand side in (34) becomes

$$\big(t_n, s^*_n(t_n), \pi_{\mathcal{P}_n} \circ g_n(t_n)\big) = \big(t_n, \pi_{\mathcal{X}_n} \circ g_n(t_n), \pi_{\mathcal{P}_n} \circ g_n(t_n)\big) = \big(t_n, g_n(t_n)\big),$$

which lies in $\mathrm{Gr}(\Psi_i)$ since g_n is a selector of Ψ_n.

Let now $i \in [1:n-1]$, and assume by induction that (34) and (35) hold for $i+1$. By (29), (32), and the induction hypothesis,

$$g_i(t_i) \in \Psi_i\big(t_i, s^*_{i+1:n}(t_{i+1}), \pi_{\mathcal{P}_{i+1:n}} \circ g_{i+1}(t_{i+1})\big), \tag{36}$$

where $t_{i+1} = \kappa_i(t^i)$. It follows from the properties of Ψ_i — or from Remark 10 and the definition of g_i in (32) — that $\pi_{X_{i+1:n}}(g_i(t_i)) = s^*_{i+1:n}(t_{i+1})$. By definition, $s^*_i(t^i) = \pi_{X_i}(g_i(t^i))$. Putting these together shows that (35) holds for i. Once we proved that (35) holds for i, we have

$$(s^*_{i:n}, \pi_{\mathcal{P}_{i:n}} \circ g_i(t_i)) = (\pi_{\mathcal{X}_{i:n}} \circ g_i(t_i), \pi_{\mathcal{P}_{i:n}} \circ g_i(t_i)) = g_i(t_i).$$

Hence, the left-hand side in (34) becomes

$$\Big(t_i, s^*_{i+1:n}(t_{i+1}), \pi_{\mathcal{P}_{i+1:n}} \circ g_{i+1}(\kappa_i(t_i)), g_i(t_i)\Big).$$

By the induction hypothesis (36), this element is in $\mathrm{Gr}(\Psi_i)$, as required. ■

The next result relates $U_{i:n}(s^*, \cdot \mid t_i)$ to $g_i(t_i)$.

Lemma 10 *For each $i \in \mathcal{N}$ and each $t_i \in \mathcal{T}_i$,*

$$U_{i:n}(s^*, \cdot \mid t_i) = \pi_{\mathcal{P}_{i:n}} \circ g_i(t_i). \tag{37}$$

Proof: We prove the claim by induction on i. We start with $i = 1$. By (32) and (29), for each $t_1 \in \mathcal{T}_1$,

$$g_1(t_1) \in \Psi_1\big(t_1, s^*_{2:n}(t_1), \pi_{\mathcal{P}_{2:n}} \circ g_2(\kappa_1(t_1))\big),$$

which, by (22), implies the result for $i = 1$.

Fix now $i > 1$, and suppose the claim holds for $i - 1$. For each $t_i \in \mathcal{T}_i$,

$$\begin{aligned}
U_{i:n}(s^*, \cdot \mid t_i) &= \int_{\mathcal{T}_{i-1}} \gamma_{i:n}\big(s^*_{i-1:n}(t_{i-1}), U_{i-1:n}(s^*_{1:i-1}, \cdot \mid t_{i-1})\big)\mathbb{P}(\mathrm{d}t_{i-1} \mid t_i) && (38)\\
&= \int_{\mathcal{T}_{i-1}} \gamma_{i:n}\big(s^*_{i-1}(t_{i-1}), \pi_{\mathcal{P}_{i-1:n}} \circ g_{i-1}(t_{i-1})\big)\mathbb{P}(\mathrm{d}t_{i-1} \mid t_i) && (39)\\
&= \int_{\mathcal{T}_{i-1}} \gamma_{i:n}\big(g_{i-1}(t_{i-1})\big)\mathbb{P}(\mathrm{d}t_{i-1} \mid t_i) && (40)\\
&= \int_{\mathcal{T}_{i-1}} \gamma_{i:n}\Big(f_{i-1}\big(t_{i-1}, t_i, s^*_{i:n}(t_i), \pi_{\mathcal{P}_{i:n}} \circ g_i(t_i), \pi_{\mathcal{P}_{i+1:n}} \circ g_{i+1}(\kappa_i(t_i))\big)\Big)\mathbb{P}(\mathrm{d}t_{i-1} \mid t_i) && (41)\\
&= \pi_{\mathcal{P}_{i:n}} \circ g_i(t_i), && (42)
\end{aligned}$$

where (38) holds since information is nested, (39) holds by the induction hypothesis, (40) holds by (35), (41) holds by (32), and (42) holds by (30). ■

We can now conclude the proof of Theorem 1.

Lemma 11 *s^* is a Bayesian 0-equilibrium.*

Proof: Fix $i \in \mathcal{N}$ and $t_i \in \mathcal{T}_i$. Then

$$\begin{aligned}
\big(s^*_i(t_i), U_{i:n}(s^* \mid t_i)\big) = g_i(t_i) &\in \Psi_i\big(t_i, s^*_{i+1:n}(t_i), \pi_{\mathcal{P}_{i+1}} \circ g_{i+1}(\kappa_i(t_i))\big)\\
&= \Psi_i\big(t_i, s^*_{i+1:n}(t_i), U_{i+1:n}(s^* \mid \kappa_i(t_i))\big),
\end{aligned}$$

where the first equality holds by (35) and Lemma 10, the inclusion holds by (32) and (29), and the second equality holds by Lemma 10. Corollary 1 now implies that s^* is a Bayesian 0-equilibrium. ■

7 Extensions

In Remark 5 we mentioned extensions of Theorem 1 to compact metric action spaces and to tree-like information structure.

In this section, we elaborate on these extensions.

7.1 Compact metric action spaces

Consider Bayesian games where for each player $i \in \mathcal{N}$, (a) the action space $\mathcal{A}_i$ is compact metric, and (b) the payoff function is continuous over $\mathcal{A}$ for each type and integrable, namely, $\mathbb{E}\big[\max_{a\in\mathcal{A}} \big|R_i(a)\big|\big] < \infty$. In this model, the existence of a Harsanyi ε-equilibrium can be established as follows. Since payoffs are integrable, there are $M > 0$ and $\mathcal{T}' \subseteq \mathcal{T}$ such that $\max_{a\in\mathcal{A}} |R_i(t, a)| \le M$ for each $t \in \mathcal{T}'$, and $\mathbb{E}\big[\mathbf{1}_{\mathcal{T}\setminus\mathcal{T}'} \max_{a\in\mathcal{A}} \big|R_i(a)\big|\big] \le \frac{\varepsilon}{2}$. Define a revised game Γ' with payoffs agreeing with R in $\mathcal{T}'$ and 0 otherwise. Then for any $\delta > 0$, a Harsanyi δ-equilibrium of Γ' is a Harsanyi $(\delta + \varepsilon)$-equilibrium of Γ.

Once this reduction is made, a Scorza-Dragoni type theorem – see Footnote 4 – shows that there is a set of types $\mathcal{T}'' \subseteq \mathcal{T}'$ such that $\mathbb{P}(\mathcal{T}'') \ge 1 - \varepsilon$ and the family of functions $(R(t, \cdot))_{t\in\mathcal{T}''}$ is uniformly equicontinuous. We can thus let Γ'' be the game with type space $\mathcal{T}''$ and with action spaces $(\mathcal{A}_i')_{i\in\mathcal{N}}$ that are finite subsets of $(\mathcal{A}_i)_{i\in\mathcal{N}}$, respectively, such that for each Player i and each $t \in \mathcal{T}''$, the family of functions $(R_i(t, a_i, \cdot))_{a_i\in\mathcal{A}_i'}$ on $\prod_{j\neq i} \mathcal{A}_i$ is ε-uniformly dense in $(R_i(t, a_i, \cdot))_{a_i\in\mathcal{A}_i}$. A Bayesian 0-equilibrium of Γ'' can then be shown to be a Harsanyi $2(M + 1)\varepsilon$-equilibrium of Γ', which is a Harsanyi $(2M + 3)\varepsilon$-equilibrium of Γ. Such a Bayesian 0-equilibrium exists by virtue of Theorem 1.

In a subsequent paper we strengthen this result and establish the existence of an ε-Bayesian equilibrium in this model.

7.2 Tree-like information

In our model, the players are ordered, and each player knows the types of all the players who follow her in that order. In some set ups, like in hierarchical organizations, the information structure is *tree-like*; that is, the players are vertices of a tree, and each player i knows the types of all her descendants, denoted $D(i)$. That is, for each $i \in \mathcal{N}$ there exists a measurable mapping $\kappa_i : \mathcal{T}_i \mapsto \prod_{j \in D(i)} \mathcal{T}_j$ such that $\mathbb{P}\left(\left((t_j)_{j \in D(i)}\right) = \kappa_i(t_i)\right) = 1$.

Denote by $\overline{D}(i)$ the *lineage set* of Player i, namely, the set that includes the ancestors of i, the descendants of i, and i herself. We here explain how to extend our result to Bayesian games with tree-like information structure, provided the payoff of each player is affected only by the types and actions of the players in $\overline{D}(i)$. Specifically, we explain how to generalize the two steps of the proof: The existence of ε-equilibrium, which was demonstrated in Section 6.1, and the existence of 0-equilibrium, which was demonstrated in Section 6.2.

For the former, note that there is at least one $i \in \mathcal{N}$ such that $D(i) = \emptyset$, because $\mathcal{N}$ is finite. In the notation of Section 6.1, for each i such that $D(i) = \emptyset$, define $\psi_i(t) \equiv (R_j(t))_{j \in \overline{D}(i)}$, and let $\varphi_i(t)$ be a δ-approximation of $\mathbb{P}(\psi_i \mid t_i)$. Recursively, for each $i \in \mathcal{N}$ for which φ_j has already been defined for every $j \in D(i)$, denote $\psi_i(t_i) \equiv \left((\varphi_j)_{j \in D(i)}, (R_j)_{j \in \overline{D}(i)}\right)$, and let φ_i be a δ-approximation of $\mathbb{P}(\psi_i \mid t_i)$.

The auxiliary game with a finite type space Γ' is defined by setting Player i's type to $\tau_i \equiv (\varphi_i, (\varphi_j)_{j \in D(i)})$, and the probability measure in that game is the push-forward measure of $(\tau_i)_{i \in \mathcal{N}}$ with respect to $\mathbb{P}$.

The proof that Γ' admits a Bayesian 0-equilibrium, that this Bayesian 0-equilibrium induces a Bayesian ε-equilibrium in Γ, are analogous to those given in Section 6 for Bayesian games with nested information.

The construction of a Bayesian 0-equilibrium from a sequence of Bayesian $\frac{1}{k}$-equilibria also follows via a similar modification to the construction carried out in Section 6.2: The construction of an equilibrium strategy profile, which had been carried recursively from the least knowledgeable player towards the most knowledgeable players in the nested-information model, is now carried out beginning with those players $i \in \mathcal{N}$ for whom $D(i) = \emptyset$, and then recursively for those players $i \in \mathcal{N}$ for whom the equilibrium strategy has already been defined for all $j \in D(i)$. The relevant construction remains well-defined by virtue of our assumption that each player's payoff depends only on action of players in their lineage.

We note that every Bayesian game Γ (without necessarily nested information) can be presented as a Bayesian game Γ' with a tree-like information structure; however, this representation cannot always be done in a way which guarantees that each players' payoff depends only on players in their lineage. Indeed, denoting by $\mathcal{N}$ the set of players in Γ, we can define Γ' with set of players $\mathcal{N} \cup \{*\}$, where player $*$ knows the types of all players in $\mathcal{N}$ and has a single action (so she is a dummy player), and each player in $\mathcal{N}$ knows only her own type. As mentioned above, there are examples of Bayesian games without Bayesian 0-equilibria and Bayesian ε-equilibria. In these examples, the payoff of each player depends on the actions of the other players in $\mathcal{N}$. Hence, without the requirement that the payoff of each player only depends on the actions of the players in her lineage, games with tree-like information structure need not admit Bayesian 0-equilibria and Bayesian ε-equilibria.

References

[1] Aicardi, M., Davoli, F., & Minciardi, R. (1987). Decentralized optimal control of Markov chains with a common past information set. *IEEE Transactions on Automatic Control*, **32**, 1028–1031.

[2] Artstein, Z. (1989). Parametrized integration of multifunctions with applications to control and optimization. *SIAM Journal on Control and Optimization*, **27**, 1369–1380.

[3] Aumann, R. J. (1965) Integrals of set-valued functions. *Journal of Mathematical Analysis and Applications*, **12**:1, 1–12.

[4] Aumann, R. J. (1976). An elementary proof that integration preserves uppersemicontinuity. *Journal of Mathematical Economics*, **3**, 15–18.

[5] Aumann, R. J., & Maschler, M., (1968) *Repeated games of incomplete information: The zero-sum extensive case*, in Report of the U.S. Arms Control and Disarmament Agency ST-143, Washington, D.C., Chapter III, pp. 37–116.

[6] Aumann, R. J., & Maschler, M., (1995) *Repeated games with incomplete information* (with the collaboration of R. E. Stearns), MIT Press, Cambridge, MA.

[7] Balbus, Ł., Dziewulski, P., Reffett, K., & Woźny, Ł. (2015). Differential information in large games with strategic complementarities. *Economic Theory*, **59**, 201–243.

[8] Barelli, P., & Duggan, J. (2015). Purification of Bayes Nash equilibrium with correlated types and interdependent payoffs. *Games and Economic Behavior*, **94**, 1–14.

[9] Carbonell-Nicolau, O., & McLean, R. P. (2018). On the existence of Nash equilibrium in Bayesian games. *Mathematics of Operations Research*, **43**, 100–129.

[10] Cardaliaguet, P., & Rainer, C. (2012). Games with incomplete information in continuous time and for continuous types. *Dynamic Games and Applications*, **2**, 206–227.

[11] De Angelis, T., Ekström, E., & Glover, K. (2022). Dynkin games with incomplete and asymmetric information, *Mathematics of Operations Research*, **47**, 560–586.

[12] De Angelis, T., Merkulov, N., & Palczewski, J. (2022). On the value of non-Markovian Dynkin games with partial and asymmetric information, *The Annals of Applied Probability*, **32**, 1774–1813.

[13] De Meyer, B., Lehrer, E., & Rosenberg, D. (2010). Evaluating information in zero-sum games with incomplete information on both sides. *Mathematics of Operations Research*, **35**, 851–863.

[14] Debo, L. G., Parlour, C., & Rajan, U. (2012). Signaling quality via queues. *Management Science*, 58(5), 876–891.

[15] Esmaeeli, N., Imkeller, P., & Nzengang, V. (2019). Dynkin game with asymmetric information, *Bulletin of the Iranian Mathematical Society*, **45**, 283–301.

[16] Gapeev, P. V., & Rodosthenous, N. (2021). Optimal stopping games in models with various information flows, *Stochastic Analysis and Applications*, **39**, 1050–1094.

[17] Gensbittel, F., & Grün, C. (2019). Zero-sum stopping games with asymmetric information, *Mathematics of Operations Research*, **44**, 277–302.

[18] Grün, C. (2013). On Dynkin games with incomplete information. *SIAM Journal on Control and Optimization*, **51**(5), 4039–4065.

[19] Hassin, R. (2016). *Rational queuing*. CRC press.

[20] Harsanyi, J. C. (1967). Games with incomplete information played by "Bayesian" players, I–III Part I. The basic model. *Management Science*, **14**(3), 159–182.

[21] Hellman, Z. (2014). A game with no Bayesian approximate equilibria. *Journal of Economic Theory*, **153**, 138–151.

[22] Hellman, Z., & Levy, Y. J. (2017). Bayesian games with a continuum of states. *Theoretical Economics*, **12**(3), 1089–1120.

[23] Hellman, Z., & Levy, Y. J. (2019). Measurable selection for purely atomic games. *Econometrica*, **87**(2), 593–629.

[24] Hellman, Z., & Levy, Y. J. (2022). Equilibria existence in Bayesian games: Climbing the countable Borel equivalence relation hierarchy. *Mathematics of Operations Research*, **47**(1), 367–383.

[25] Himmelberg, C.J. (1975) Measurable relations. *Fundamenta Mathematicae*, **78**(1), 53–72.

[26] Huber, J., Kirchler, M., & Sutter, M. (2008). Is more information always better?: Experimental financial markets with cumulative information. *Journal of Economic Behavior & Organization*, **65**(1), 86–104.

[27] Jacobovic, R. (2022). Recursive construction of a Nash equilibrium in a two-player nonzero-sum stopping game with asymmetric information. *SIAM Journal on Control and Optimization*, **60**, 3289–3301.

[28] Kamenica, E. (2019). Bayesian persuasion and information design. *Annual Review of Economics*, **11**, 249–272.

[29] Kamenica, E., & Gentzkow, M. (2011). Bayesian persuasion. *American Economic Review*, **101**(6), 2590–2615.

[30] Khan, M. A., & Sun, Y. (1999). Non-cooperative games on hyperfinite Loeb spaces. *Journal of Mathematical Economics*, **31**(4), 455–492.

[31] Kechris, A. (1995). *Classical Descriptive Set Theory.* Springer-Verlag, New York.

[32] Kim, T., & Yannelis, N. C. (1997). Existence of equilibrium in Bayesian games with infinitely many players. *Journal of Economic Theory*, **77**, 330–353.

[33] Kremer, M., & Debo, L. (2016). Inferring quality from wait time. *Management Science*, **62**, 3023–3038.

[34] Kucia, A. (1991). Scorza Dragoni type theorems. *Fundamenta Mathematicae*, **138**(3), 197–203.

[35] Kuratowski K. & Ryll-Nardzewski C. (1965) A general theorem on selectors. *Bulletin de l'Académie Polonaise des Sciences, Série des Sciences Mathématiques Astronomiques et Physique*, **13**(6), 397–411.

[36] Lehrer, E., & Rosenberg, D. (2006). What restrictions do Bayesian games impose on the value of information?. *Journal of Mathematical Economics*, **42**, 343–357.

[37] Lempa, J., & Matomäki, P. (2013) A Dynkin game with asymmetric information. *Stochastics*, **85**(5), 763–788.

[38] Levy, Y. J. (2024). Bayesian equilibrium: From local to global. *Journal of Mathematical Economics*, 103012.

[39] Taneva, I., & Mathevet, L. (2022). Organized information transmission (No. 16959). CEPR Discussion Papers.

[40] Mertens, J. F. (1987). A measurable "measurable choice" theorem. CORE Discussion Paper No. 8749. Reproduced in *Stochastic Games and Applications* (pp. 107–130). Dordrecht: Springer Netherlands, 2003.

[41] Mertens, J. F., & Parthasarathy, T. (1991). Nonzero-sum stochastic games. In: Raghavan, T.E.S., Ferguson, T.S., Parthasarathy, T., Vrieze, O.J. (eds) *Stochastic Games and Related Topics*. Theory and Decision Library, vol 7. Springer, Dordrecht.

[42] Mertens, J. F., & Zamir, S. (1985). Formulation of Bayesian analysis for games with incomplete information. *International Journal of Game Theory*, **14**, 1–29.

[43] Milgrom, P. R., & Weber, R. J. (1985). Distributional strategies for games with incomplete information. *Mathematics of Operations Research*, **10**, 619–632.

[44] Myerson, R. B., & Reny, P. J. (2020). Perfect conditional ε-equilibria of multi-stage games with infinite sets of signals and actions. *Econometrica*, **88**, 495–531.

[45] Nayyar, A., Mahajan, A., & Teneketzis, D. (2010). Optimal control strategies in delayed sharing information structures. *IEEE Transactions on Automatic Control*, **56**, 1606–1620.

[46] Olszewski, W., & Siegel, R. (2023). Equilibrium existence in games with ties. *Theoretical Economics*, **18**, 481–502.

[47] Pérez, J. L., Rodosthenous, N., & Yamazaki, K. (2024). Nonzero-sum optimal stopping game with continuous vs. periodic exercise opportunities. *Mathematics of Operations Research*, forthcoming.

[48] Radner, R., & Rosenthal, R. W. (1982). Private information and pure-strategy equilibria. *Mathematics of Operations Research*, **7**(3), 401–409.

[49] Reny, P. J. (1999). On the existence of pure and mixed strategy Nash equilibria in discontinuous games. *Econometrica*, **67**(5), 1029–1056.

[50] Reny, P. J. (2011). On the existence of monotone pure-strategy equilibria in Bayesian games. *Econometrica*, **79**(2), 499–553.

[51] Shmaya, E., & Solan, E. (2004). Two-player nonzero–sum stopping games in discrete time. *Annals of Probability,* **32**, 2733–2764.

[52] Simon, R. S. (2003). Games of incomplete information, ergodic theory, and the measurability of equilibria. *Israel Journal of Mathematics*, **138**, 73–92.

[53] Simon, R. S., & Tomkowicz, G. (2018). A Bayesian game without ε-equilibria. *Israel Journal of Mathematics*, **227**, 215–231.

[54] Simon, R., & Tomkowicz, G. (2023). A measure theoretic paradox from a continuous colouring rule. *Mathematische Annalen*, 1–22.

[55] Stinchcombe, M. B., & White, H. (1992). Some measurability results for extrema of random functions over random sets. *The Review of Economic Studies*, **59**(3), 495–514.

[56] Stinchcombe, M. B. (2011). Balance and discontinuities in infinite games with type-dependent strategies. *Journal of Economic Theory*, **146**(2), 656–671.

[57] Toth, B., Scalas, E., Huber, J., & Kirchler, M. (2007). The value of information in a multi-agent market model: The luck of the uninformed. *The European Physical Journal B*, **55**, 115–120.

[58] Ui, T. (2016). Bayesian Nash equilibrium and variational inequalities. *Journal of Mathematical Economics*, **63**, 139–146.

[59] Ui, T., & Yoshizawa, Y. (2015). Characterizing social value of information. *Journal of Economic Theory*, **158**, 507–535.

[60] Varaiya, P., & Walrand, J. (1978). On delayed sharing patterns. *IEEE Transactions on Automatic Control*, **23**, 443–445.

[61] Vives, X. (1990). Nash equilibrium with strategic complementarities. *Journal of Mathematical Economics*, **19**(3), 305–321.

[62] Wu, M., Amin, S., & Ozdaglar, A. E. (2021). Value of information in Bayesian routing games. *Operations Research*, **69**, 148–163.

[63] Yang, J. (2022). A Bayesian nonatomic game and its applicability to finite-player situations. *Journal of Mathematical Economics*, **102**, 102746.

[64] Zeng, Y. (2023). Derandomization of persuasion mechanisms. *Journal of Economic Theory*, **212**, 105690.