INNOCENZO GASPARINI INSTITUTE FOR ECONOMIC RESEARCH

Institutional Members: CEPR, NBER and Università Bocconi

WORKING PAPER SERIES

Sophisticated Reasoning, Learning, and Equilibrium in Repeated Games with Imperfect Feedback

Pierpaolo Battigalli, Davide Bordoli

Working Paper n. 702

This Version: July, 2023

IGIER – Università Bocconi, Via Guglielmo Röntgen 1, 20136 Milano –Italy
http://www.igier.unibocconi.it

Sophisticated Reasoning, Learning, and Equilibrium in Repeated Games with Imperfect Feedback*

Pierpaolo Battigalli
Bocconi University and IGIER, pierpaolo.battigalli@unibocconi.it

Davide Bordoli
Yale University, davide.bordoli@yale.edu

July 31, 2023

Abstract

We analyze the infinite repetition with imperfect feedback of a simultaneous or sequential game, assuming that players are strategically sophisticated—but possibly impatient—expected-utility maximizers. Sophisticated strategic reasoning in the repeated game is combined with belief updating to provide a foundation for a refinement of self-confirming equilibrium. In particular, we model strategic sophistication as *rationality and common strong belief in rationality.* Then, we combine belief updating and sophisticated reasoning to provide sufficient conditions for a kind of learning—that is, the ability, in the limit, to exactly forecast the sequence of future observations—thus showing that impatient agents end up playing a sequence of *self-confirming equilibria in strongly rationalizable conjectures* of the one-period game. Irrespective of whether individuals value the future, if they are able to learn then they will play in the limit a self-confirming equilibrium in strongly rationalizable conjectures of the continuation (infinitely repeated) game.

*We thank Edoardo Bonaglia, Nicodemo De Vito, Ryota Iijima, Paolo Leonetti, Shuige Liu, Anna Merotto, and Ferdinand Pieroth for useful comments.

1. Introduction

In this paper we analyze the limits of learning dynamics in the infinite repetition with imperfect monitoring of a one-period game played by strategically sophisticated agents. The one-period game may be sequential or with simultaneous moves. Focusing on the case of impatient agents who maximize their subjective expected one-period payoff, we relate such limits to solutions of the one-period game, that is, self-confirming equilibrium and rationalizability.

In a **self-confirming equilibrium** (**SCE**), players best respond to confirmed conjectures (first-order beliefs) about co-players' behavior, where "confirmed" means that each player, given her conjecture, correctly predicts what she observes about the play. The SCE concept characterizes the rest-point limits of learning dynamics in games played recurrently given the possibly imperfect feedback about play obtained by each player at the end of each period (e.g., Fudenberg & Kreps 1995 and Gilli 1999).[1] Since the SCE concept is not meant to capture strategic reasoning, in such rest points players' conjectures may be incompatible with strategic reasoning based on what is commonly known about the game. Indeed, in an environment with possibly incomplete information and private values, the SCE set at any given state of nature is independent of players' interactive knowledge of the profile of payoff functions. It is then natural to ask how one can characterize the limits of learning dynamics when beliefs are shaped by sophisticated strategic reasoning, which we take to mean some form of *common belief in rationality*.

The literature offers two kinds of answers that directly focus on refinements of SCE, neglecting an explicit analysis of learning dynamics. The simplest one can be found in the works that first put forward a version of the SCE concept (Battigalli 1987, and Battigalli & Guaitoli 1988): SCE should be refined by requiring that players' conjectures about co-players' behavior assign probability 1 to co-players' rationalizable strategies, a condition that follows from common belief in rationality. Yet, such **SCE in rationalizable conjectures** allows for the possibility that confirmation of conjectures is not commonly believed, which may be thought to

[1]Note that the SCE term was coined by Fudenberg & Levine (1993), but the concept was also previously or simultaneously called "conjectural equilibrium" (Battigalli 1987, Battigalli & Guaitoli 1988, Rubinstein & Wolinsky 1994) and "subjective equilibrium" (Kalai & Lehrer 1993, 1995). Here we stick to the more explicative SCE terminology (see the discussion in Battigalli et al. 2015).

jeopardize the stability of the equilibrium. Intuitively, if confirmed conjectures is a pre-requisite to play again the same strategies, why should a sophisticated player who is unsure whether her co-players' conjectures are confirmed expect them to behave in the future as in the current period? And if they don't, why should she? Motivated by such informal considerations, Rubinstein & Wolinsky (1994) proposed an even more refined notion of SCE: while an SCE in rationalizable conjectures obtains if—on top of rationality—there is common belief in rationality and conjectures are confirmed, in a **rationalizable SCE**[2] players conjectures about behavior are compatible with common belief of *both* rationality and confirmation of conjectures. Rationalizable SCE is elegant and intuitive, but—unlike the mere SCE concept, to the best of our knowledge—there is no formal result relating it to learning in recurrent interactions. Instead, here we obtain a kind of learning foundation for SCE in rationalizable conjectures.

To formally represent rationality and strategic sophistication, we adopt the approach of epistemic game theory[3] extended to infinitely repeated games as in Battigalli & Tebaldi (2019). To ease notation, we assume *complete information*: the rules of the game and players' expected-utility preferences over streams of stochastic outcomes are commonly known. Since the one-period game being repeated may have a sequential (multistage) structure, we need to distinguish between strategies of the one-period game and strategies of the repeated game; we call the latter **superstrategies**. Players are endowed with **conditional probability systems (CPSs)**, which specify subjective beliefs about the behavior and beliefs of co-players *in the infinitely repeated game* conditional on every personal history (roughly, information set) so as to satisfy the *chain rule*. We assume that players are **rational**, that is, they carry out (super)strategies that maximize their subjective expected utility conditional on every personal history, including those that they did not expect to observe according to earlier conjectures specified by their CPSs. Of course, assumptions about intertemporal preferences are crucial. We mostly focus on the extreme case of **impatient** players who do not value future payoffs, as in much of the literature on learning in games, but we also consider the case of a positive discount factor. To model strategic sophistication, we assume *common strong belief in rationality* (Battigalli & Siniscalchi 2002): each player

[2]the words of Rubinstein & Wolinsky, "rationalizable conjectural equilibrium."

[3]See, e.g., the survey of Dekel & Siniscalchi (2015).

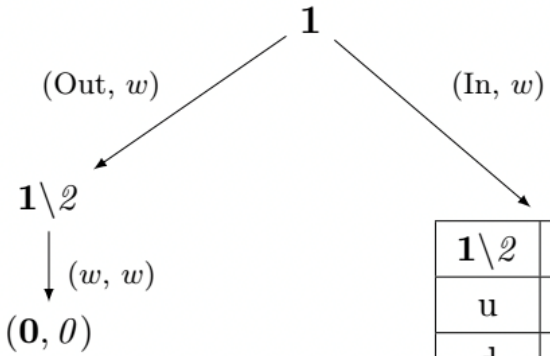
strongly believes in the co-players' rationality, i.e., she assigns probability 1 to it conditional on every personal history that does not contradict it; furthermore, she strongly believes that, on top of being rational, her co-players also strongly believe in the rationality of others; analogous assumptions hold for higher and higher levels of beliefs about beliefs. With this, in every period impatient agents play (strongly) rationalizable strategies, and *assign probability 1 to the (strongly) rationalizable strategies of others* even if they are surprised.[4] The reason is that, on a rationalizable path, unexpected observations cannot be due to deviations from rationalizability; therefore, common strong belief in rationality implies that even surprised players keep believing in rationalizability. To obtain convergence to SCE play, we assume that the profile of superstrategies and CPSs satisfy an **"observational grain of truth"** condition (cf. Kalai & Lehrer 1993, 1995): after some time T, each player assigns positive probability to what she is actually going to observe in the continuation (infinitely repeated) game.[5] This implies that, in the long-run limit, players assign probability 1 to what they observe, i.e., their *conjectures are confirmed.* Since impatient players maximize their one-period subjective expected utility, there must be *convergence to playing an SCE in rationalizable conjectures in each period.* However, *the SCE played in the limit may change from period to period*, because convergence of conjectures about superstrategies in the infinite repetition does not imply convergence of marginal one-period conjectures. We also show a converse: for every sequence of one-period SCEs in rationalizable conjectures there is a profile of superstrategies and CPSs satisfying the aforementioned conditions that yields such sequence in the limit. Finally, we extend our results to allow for a positive discount factor: under rationality, common strong belief in rationality and observational grain of truth, players' *behavior and conjectures converge to an SCE in rationalizable conjectures of the repeated game*, that is, in the long-run limit players best respond to confirmed conjectures assigning probability 1 to co-players' rationalizable superstrategies of the continuation game.

[4]Strong rationalizability is akin to the notion of rationalizability for sequential games put forward by Pearce (1984), often called "extensive-form rationalizability" (see Battigalli 2003). Thus, it coincides with the usual rationalizability concept in games with simultaneous moves. Since it is the only version of the rationalizability idea considered here, we sometimes simplify our language and omit the adjective "strong."

[5]Absent randomization, which we exclude because expected utility maximizers have no need to randomize, our assumption is a generalization of the "grain of truth" condition of Kalai & Lehrer (1993).

1.1 Heuristic Example

To give guidance and provide intuition we informally analyze an example we will repeatedly refer to in the rest of the paper.



1*2*	ℓ	r
u	**1**, *0*	**-1**, *1*
d	**-2**, *0*	**-1**, *-1*

Figure 1: Entry game

Consider the two-person, two-stage, one-period game depicted in Figure 1. In the first stage, only player **1** (**she**, player label and payoffs in **bold)** is active and can go Out or In, while player *2* (*he*, player label and payoffs in *Italics*) can only "wait". Action Out effectively terminates the game. But we find it convenient in the analysis of the repetition of such one-period games to have a fixed number L of stages, in this case $L = 2$. Thus, we have player **1** and *2* "wait" for one stage after Out before they get their payoffs. A key ingredient of the analysis is what players are able to observe, i.e., their feedback, at the end of each stage, including the last one. End-game information is crucial to determine the information structure of the repeated game and, relatedly, the set of SCEs. *In this example, we assume that players observe* (i) first-stage actions, so that action In leads to a proper subgame, and (ii) *only their realized payoff at the end of the second stage.* Of course, we also assume that players remember what they did and their earlier information. Thus, if player *2* chooses ℓeft after In, he observes that his payoff is 0, but cannot ascertain whether **1** chose up or down. With this, the game has two SCE outcomes:

1. Player **1** goes Out.[6]

[6]Out is also the unique subgame perfect equilibrium outcome.

(a) Indeed, she is afraid that *2* would chose *r*ight in the subgame. Her conjecture is trivially confirmed, because she cannot observe how the co-player would have reacted.[7]

(b) Player *2*'s plan for the subgame is immaterial. He is initially certain of Out; hence, his conjecture is confirmed. If he—counterfactually—observed In, he would revise his conjecture about **1** and choose a best reply.

2. Player **1** goes In and then up in the subgame, player *2* goes ℓeft in the subgame.

   (a) Indeed, player **1** is certain that *2* would go ℓeft in the subgame and her conjecture is confirmed.

   (b) Player *2* is certain of In and deems In.d more likely than In.u. He only observes that his realized payoff is 0, as he—necessarily—expected, and cannot identify **1**'s subgame action; hence, his conjecture is confirmed.

Both one-period equilibrium outcomes can occur infinitely often in the limit play of the repeated game if players are impatient, rational (i.e., they always best respond to their beliefs), and their conjectures about the co-player's superstrategy converge, as must be the case if the observational-grain-of-truth condition holds. However, only the first outcome is consistent with sophisticated strategic thinking based on common knowledge of the game (complete information): Strategy In.d of player **1** is strictly dominated by Out. If player *2* strongly believes in **1**'s rationality, upon observing action In—even if surprised—he would infer that **1** is going up in the subgame; thus, he would react by going *r*ight. If player **1** is certain that *2* is rational and strongly believes in her rationality, then she expects r in the subgame and goes Out.

Since in this simple example there is only one strongly rationalizable outcome of the one-period game, under impatience, strong rationalizablity in the repeted game yields the infinite repetition of this outcome. In less simple examples with

[7]Given such conjecture, if she plans by folding back, she plans to minimize her loss in the subgame by choosing up. But, whatever her continuation plan for the subgame, she believes she is better off going Out at the root.

multiple one-period strongly rationalizable outcomes, players may assign subjective probability 0 to some of them and be surprised by observations concerning the previous periods. But strong rationalizability in the repeated game implies that players, even if surprised, hold on to the belief that co-players are rational and strategically sophisticated.[8]

The rest of the paper is organized as follows. Section 2 contains some mathematical preliminaries. Section 3 describes one-period multistage games with imperfectly observable actions and their infinite repetition. Section 4 analyzes rationality for the one-period game and its repetition, and characterizes the behavioral and first-order belief implications of rationality and common strong belief in rationality. Section 5 analyzes convergence of conjectures. Section 6 contains the main results of the paper. Section 7 discusses in detail the related literature and some possible extensions of our work.

2. Preliminaries

We provide some mathematical and notational preliminaries.

2.1 Mathematical notation

We let $[n] = \{1, \ldots, n\}$ denote the set of the first n natural numbers. Given a finite set X, we let $X^{[n]}$ denote the set of functions from $[n]$ to X (i.e., the sequences of length n of elements of X) $X^{\mathbb{N}}$ denote the set of infinite sequences of elements of X, $X^{[0]} = \{\varnothing\}$ denote the singleton containing the empty sequence $\varnothing$, $X^{<\mathbb{N}_0} = \cup_{n\in\mathbb{N}_0} X^{[n]}$ denote the set of finite sequences of elements of X, and $X^{\leq\mathbb{N}_0} = X^{<\mathbb{N}_0} \cup X^{\mathbb{N}}$ denote the set of finite and infinite sequences of elements of X.[9] We write $x^{[n]} = (x_k)_{k=1}^n \in X^{\leq\mathbb{N}_0}$ for any $n \in \mathbb{N} \cup \{\infty\}$.

We endow every finite set X with the discrete topology and any Cartesian product of sets with the product topology, and we consider the corresponding Borel

[8] The following variation of the example has multiple rationalizable outcomes: Player **1** at the root has also a third action zs that leads to a zero-sum subgame with zero-maxmin value. In this variation, both actions Out and zs (followed by appropriate continuations in the subgame) are strongly rationalizable.

[9] That is, we regard such sequences as functions with domain $[n]$ or $\mathbb{N}$ and codomain X.

σ-algebras. Given a sequence of finite sets $(X_n)_{n\in\mathbb{N}}$, the σ-algebra $\mathcal{B}(X)$ on their product $X = \prod_{n\in\mathbb{N}} X_n$ is the one generated by all the cylinders of the form $\{x_1\} \times \ldots \times \{x_n\} \times X_{n+1} \times \ldots$, with $n \in \mathbb{N}$. Given any topological space Y endowed with its Borel σ-algebra $\mathcal{B}(Y)$, we let $\Delta(Y)$ denote the space of probability measures defined on $(Y, \mathcal{B}(Y))$, which we endow with the topology of weak convergence.

We let $\preceq$ denote the canonical "prefix-of" partial order over sequences.[10] Given a sequence $x^{[n]} \in X^{[n]}$, we define its length as $\ell(x^{[n]}) = n$.

2.2 Beliefs representation and properties

In this subsection we introduce conditional probability systems, used to represent players' beliefs in sequential games, and we give the definition of strong belief.

Definition 1. *Let Y be a Polish space and $\mathcal{C} \subseteq \mathcal{B}(Y)$ be a countable collection of clopen*[11] *subsets of Y. A* ***conditional probability system (CPS)*** *on $(Y, \mathcal{C})$ is an array of probability measures $\mu = (\mu(\cdot|C))_{C\in\mathcal{C}} \in [\Delta(Y)]^{\mathcal{C}}$ such that:*

(i) for all $C \in \mathcal{C}$, $\mu(C|C) = 1$;

(ii) for all $E \in \mathcal{B}(Y)$ and $C, D \in \mathcal{C}$ such that $E \subseteq D \subseteq C$,

$$\mu(E|C) = \mu(E|D)\mu(D|C).$$

We let $\Delta^{\mathcal{C}}(Y)$ denote the set of all CPSs on $(Y, \mathcal{C})$. CPSs will be used to represent the beliefs of a player, compactly modeling the way in which, upon observing some personal history of actions and messages (from which a conditioning event can be inferred), a player updates or revise her beliefs. In particular, (i) upon observing the event corresponding to any personal history, a player is certain of it; (ii) beliefs comply with the chain rule of conditional probabilities, hence, standard updating holds whenever an observed event was previously deemed possible: $\mu(D|C) > 0$ implies $\mu(E|D) = \mu(E|C)/\mu(D|C)$.

[10] That is, $x^{[m]} \prec y^{[n]}$ if $m < n$ and $y^{[n]} = (x^{[m]}, ...)$; $x^{[m]} \preceq y^{[n]}$ if either $x^{[m]} \prec y^{[n]}$, or $m = n$ and $x^{[m]} = y^{[n]}$.

[11] Simultaneously closed and open. See the discussion and motivation in Battigalli & Tebaldi (2019).

Definition 2. *Fix an event $E \in \mathcal{B}(Y)$ and a CPS $\mu \in \Delta^{\mathcal{C}}(Y)$. We say that μ* ***strongly believes*** *E if, for every $C \in \mathcal{C}$,*

$$E \cap C \neq \emptyset \Rightarrow \mu(E|C) = 1.$$

In words, μ strongly believes E if an agent with CPS μ is certain of E whenever possible, i.e., upon observing any evidence that does not contradict E.

3. Games

We give a formal description of the one-period game and of its repetition.

3.1 One-period game

A finite multistage game with feedback is a game that may last for more than one stage, where at each stage every player chooses an action and then observes a message about the play. We represent the information accruing to agents as the play unfolds with a formalism that is similar to the one used to represent information (monitoring) in repeated games.[12] Stages are indexed by natural numbers: stage k starts after the end of stage $k-1$ and ends with the realization of the profiles of actions played and messages received by players. To ease notation, we adopt the convention that the set of available actions (after some non-terminal play) of an **inactive** player is the singleton $\{w\}$, where w is interpreted as the action "wait." A finite multistage game has necessarily a finite horizon, that is, a maximum number of stages $L \in \mathbb{N}$ after which the game ends. In order to simplify the formal representation of the infinite repetition of the game, we adopt the convention that, *each time the one-period game is played, the play lasts L stages.* If at some history shorter than L the game ends, then players are assumed to play the action "wait" for all the following stages, until stage L. The rules of interaction are represented by the primitive elements

$$\langle I, (A_i, M_i, \mathcal{A}_i, f_i)_{i\in I}\rangle,$$

where:

[12] See Battigalli & Generoso (2021).

- I is the finite set of players;
- A_i is the finite set of all actions player i may ever take at any point in the game;
- M_i is the finite set of all messages player i may observe at any point in the game, including the initial message m_i^0, saying that the game is about to start and informing i of her feasible actions;
- $\mathcal{A}_i = (\mathcal{A}_i^k : M_i \rightrightarrows A_i)_{k=1}^L$ is a *sequence of constraint correspondences*, where for every k and possible message m_i^{k-1}, $\mathcal{A}_i^k\left(m_i^{k-1}\right)$ specifies the set of i's feasible actions at stage k; thus, i is informed of her feasible actions, independently of her mnemonic abilities;
- $f_i = (f_i^k : A^{[k]} \to M_i)_{k=0}^L$ is the *incremental feedback structure*, where for every k and every conceivable sequence of action profiles $a^{[k]}$, $f_i^k\left(a^{[k]}\right)$ is the message that would be observed by player i at the end of stage k if the sequence of actions $a^{[k]}$ were played.

The sequence of own actions (which are automatically observed as soon as they are irreversibly chosen) and received messages up to the current stage, or personal history, determines the information potentially available to a player. Since the initial message m_i^0 is fixed, we ignore it in the following notation and it does not affect the length of histories.

From these primitives we derive the sets of one-period histories and personal histories, that is, the *objective* and the *subjective trees* generated by the one-period game form:[13]

- $\bar{H}$ is the set of **histories,** that is, partial, or complete plays made of *feasible* sequences of action profiles including the empty sequence $a^{[0]} = \varnothing$ (root).
- $Z = \left\{z \in \bar{H} : \ell(z) = L\right\}$ is the set of **terminal histories**;
- $H = \bar{H} \setminus Z$ is the set of **non-terminal histories**;

[13]See the complete formalism in Appendix 8.1.

- $\bar{H}_i$ is the set of **personal histories** is, that is, feasible sequences of own actions and messages, and it is partitioned into Z_i and H_i (terminal and non-terminal personal histories).

Personal histories represent what a player is able to observe at each stage given that a certain history has occurred. We informally assume that the *rules of interaction* represented by the foregoing formal structure are *common knowledge.* Whether at a certain stage the player is able to use the sequence of own actions and messages she has observed as information depends on her memory. In this framework, the assumption of *perfect memory* consists in saying that, at every stage k, each player i remembers her personal history. Under perfect memory, for each player, personal histories of actions and messages yield an information partition that satisfies the standard perfect recall assumption (see Battigalli & Generoso 2021). In particular, let

$$\begin{array}{rlcl} P_i: & \bar{H}_i & \rightrightarrows & \bar{H}, \\ & (a_i, m_i)^{[k]} & \mapsto & \left\{ h \in \bar{H} \ : \ a_i^{[k]} = \mathrm{proj}_{A_i^{\leq \mathbb{N}}} h \ , \ f_i^{[k]}(h) = m_i^{[k]} \right\}, \end{array}$$

denote the correspondence from each personal history to the objective histories consistent with it. As anticipated, P_i is necessarily partitional and satisfies perfect recall.[14] Then, let

$$\begin{array}{rlcl} \bar{\mathrm{h}}_i: & \bar{H} & \to & \bar{H}_i, \\ & a^{[k]} & \mapsto & P_i^{-1}\left(a^{[k]}\right) \end{array}$$

denote the map that associates each history (play) $a^{[k]} \in \bar{H}$ with the corresponding personal history. Notice that $\bar{\mathrm{h}}_i := P_i^{-1}$ is a well-defined function.

We now define strategies, that is, descriptions of information-dependent behavior. Since rational players behave as planned, we also interpret strategies as plans, consistently with the meaning of the term in the natural language.[15] A **strategy** of player i is a function $s_i = (s_i(h_i))_{h_i \in H_i}$ such that, for each $h_i \in H_i$, $s_i(h_i) \in \mathcal{A}_i^{\ell(h_i)+1}(\mathrm{m}_i\,(h_i))$, where $\mathrm{m}_i\,(h_i)$ denotes the last message in personal history h_i. The set of i's strategies is denoted by S_i, whereas $S = \times_{i \in I} S_i$ and

[14] Note that $P_i\,(h_i)$ does not depend on the initial message m_i^0, because it is fixed.

[15] We do not offer a theory of irrationality, or of how players think about the irrationality of co-players. Furthermore, we assume it to be transparent that observed behavior is intentional. Therefore, here we need not consider the relationship between plans and behavior of irrational players (see Section 6 of Battigalli & De Vito 2021).

$S_{-i} = \times_{j\neq i} S_j$ denote the sets of strategy profiles and co-players' strategy profiles. From these elements we can derive the path function $\zeta : S \to Z$ mapping strategies to terminal histories.[16]

It is also useful to define, for each player $i \in I$, the sets of profiles that induce, and strategies that allow, some personal history $h_i \in \bar{H}_i$:

- $S(h_i) = \{s \in S : \exists x \in P_i(h_i), x \preceq \zeta(s)\}$, the set of strategy profiles inducing h_i;
- $S_i(h_i) = \{s_i \in S_i : \exists s_{-i} \in S_{-i}, \exists x \in P_i(h_i), x \preceq \zeta(s_i, s_{-i})\} = \mathrm{proj}_{S_i} S(h_i)$, the set of strategies of i that allow h_i;
- $S_{-i}(h_i) = \{s_{-i} \in S_{-i} : \exists s_i \in S_i, \exists x \in P_i(h_i), x \preceq \zeta(s_i, s_{-i})\} = \mathrm{proj}_{S_{-i}} S(h_i)$, the set of co-players' strategy profiles that allow h_i.

Intuitively, $S_{-i}(h_i)$ represents the information that h_i reveals to i about the co-players' behavior, that is, the strategies that the co-players are carrying out. One can verify that, for every $h_i \in H_i$, $S(h_i)$ can be factorized as $S(h_i) = S_i(h_i) \times S_{-i}(h_i)$. Hence, for every i,

$$\mathcal{C}_i = \{S_{-i}(h_i) : h_i \in H_i\}$$

is the relevant collection of observable events about co-players' behavior that will be used to define the set of CPSs of i. A CPS of i for the one-period game is an element of $\Delta^{\mathcal{C}_i}(S_{-i})$, that is, it is a system of conjectures that are connected, whenever possible, by the chain rule.

Positing a set of outcomes (consequences) Y, an outcome function $g : Z \to Y$ and von Neumann-Morgenstern utility functions $(v_i : Y \to \mathbb{R})_{i\in I}$, we construct payoff functions $(u_i = v_i \circ g : Z \to \mathbb{R})_{i\in I}$ on terminal histories. With this, we can also conveniently define strategic-form utility functions $U_i = u_i \circ \zeta : S \to \mathbb{R}$. The continuation value of strategy s_i at personal history h_i, given one-period CPS $\gamma^i \in \Delta^{\mathcal{C}_i}(S_{-i})$, is

$$V^{\gamma^i}_{i,h_i}(s_i) = \sum_{s_{-i}\in S_{-i}(h_i)} U_i((s_i|h_i, s_{-i}))\gamma^i(s_{-i}|S_{-i}(h_i)),$$

[16] For every $s \in S = \times_{i\in I} S_i$, $\zeta(s) = (a^k)_{k=1}^{L} \in Z$, where $a^1 = (s_i(m_i^0))_{i\in I}$ and, for every $\ell \geq 2$, $a^\ell = ((s_i(\bar{\mathrm{h}}_i(a^{[\ell-1]})))_{i\in I})$.

where $s_i|h_i$ is the modified strategy allowing h_i and playing like s_i at each personal history that does not strictly precede h_i.

Then, a **one-period multistage game** Γ with feedback structure $f = (f_i)_{i\in I}$ is defined as

$$\Gamma = \langle I, (A_i, M_i, \mathcal{A}_i, f_i, u_i)_{i\in I}\rangle.$$

To conclude the subsection, we define the notion of observationally equivalent strategy profiles of co-players, which plays a fundamental role in our analysis.

Definition 3. *Fix $i \in I$ and $s_i \in S_i$. Two profiles of co-players' strategies $\bar{s}_{-i}, s_{-i} \in S_{-i}$ are **observationally equivalent**, given s_i, if $\bar{\mathrm{h}}_i(\zeta(s_i, \bar{s}_{-i})) = \bar{\mathrm{h}}_i(\zeta(s_i, s_{-i}))$.*

In words, s_{-i} is observationally equivalent to $\bar{s}_{-i}$ given s_i if these two profiles, when played along with s_i, induce the same sequence of messages observed by player i, who thus is unable to distinguish between the two profiles.

We refer to the example of the Introduction to illustrate the formalism.

Example 1. Go back to the two-stage ($L = 2$) one-period game depicted in Figure 1 of the Introduction. Note that, formally, players choose simultaneously in each sage, but inactive players can only choose the waiting action. The first-stage feedback is perfect, that is, each function f_i^1 is injective. With this, $\mathcal{A}_{\mathbf{1}}^1(m_{\mathbf{1}}^0) = \{\text{In}, \text{Out}\}$, $\mathcal{A}_{\mathit{2}}^1(m_{\mathit{2}}^0) = \{w\}$, $\mathcal{A}_i^2(\mathrm{m}_i(\text{Out}, \text{w})) = \{w\}$ ($i \in \{\mathbf{1}, \mathit{2}\}$), $\mathcal{A}_{\mathbf{1}}^2(\mathrm{m}_{\mathbf{1}}(\text{Out}, \text{w})) = \{\text{u}, \text{d}\}$, and $\mathcal{A}_{\mathit{2}}^2(\mathrm{m}_{\mathit{2}}(\text{Out}, \text{w})) = \{\ell, r\}$. The second-stage feedback of each player i coincides with her payoff function, that is, $f_i^2\left(a^{[2]}\right) = f_i^2\left(\bar{a}^{[2]}\right)$ if and only if $u_i\left(a^{[2]}\right) = u_i\left(\bar{a}^{[2]}\right)$. Player **1** has 4 strategies, $S_{\mathbf{1}} = \{\text{Out.u}, \text{Out.d}, \text{In.u}, \text{In.d}\}$, two of which correspond to the reduced strategy Out. Player *2* has 2 strategies, which can be identified with her actions in the subgame, $S_{\mathit{2}} = \{\ell, r\}$. ▲

3.2 Infinitely repeated game

The infinite repetition of the game is itself a multistage game, whose elements are characterized by the primitives of the one-period game. For example, the set of feasible actions for player i after personal history $(a_i, m_i)^{[k]} \in (A_i \times M_i)^{[k]}$ is $\mathcal{A}_i^{k-\tau(\mathbf{h}_i)+1}(m_i^k)$, where $\tau(\mathbf{h}_i) := \lfloor \ell(\mathbf{h}_i)/L \rfloor$ denotes number of periods that have elapsed in $\mathbf{h}_i$. For the sake of brevity, we abuse notation letting $\mathcal{A}_i^{k+1}(\mathbf{h}_i)$ denote the set of feasible actions, given the last message in $\mathbf{h}_i$, a personal history of length k. Similarly, the message observed by i after history $a^{[k]}$ is $f_i^{k-\tau(a^{[k]})}\left((a^j)_{j=\tau(a^{[k]})+1}^k\right)$, where again $\tau(a^{[k]}) := \lfloor k/L \rfloor$. We let $\mathbf{f}_i^k$ denote the feedback function mapping from sequences of action profiles of length k to the corresponding message observed by i. More generally, we use symbols in **bold** to denote mathematical objects related to the repeated game, such as the various sets of histories analogous to the one-period game ones. Terminal histories of the repeated game are infinite sequences of feasible action profiles, that is, elements of $Z^{\mathbb{N}}$.

We similarly define the informational correspondences

$$\mathbf{P}_i : \quad \bar{\mathbf{H}}_i \quad \rightrightarrows \quad \bar{\mathbf{H}}$$

such that

$$\mathbf{P}_i((a_i, m_i)^{[k]}) = \left\{ a^{[k]} : a_i^{[k]} = \operatorname{proj}_{A_i^{\leq \mathbb{N}}} a^{[k]}, \left(\mathbf{f}_i^\ell \left(a^{[\ell]}\right)\right)_{\ell=1}^k = m_i^{[k]} \right\}.$$

The information content of infinite personal histories is similarly defined, replacing $[k]$ with $\mathbb{N}$. The inverses $\bar{\mathbf{h}}_i := \mathbf{P}_i^{-1}$ $(i \in I)$ are well-defined functions.

In this context, we still call "strategy" the description of the information-dependent behavior of a player *in a single period*. We instead call "**superstrategy**" the description of the information-dependent behavior of a player in the repeated game; we let $\mathbf{s}_i \in \mathbf{S}_i = \times_{\mathbf{h}_i \in \mathbf{H}_i} \mathcal{A}_i^{\ell(\mathbf{h}_i)+1}(\mathbf{h}_i)$ denote the superstrategies of player i.

As for the one-period game, one can define the path function over superstrategies $\boldsymbol{\zeta} : \mathbf{S} \to Z^{\mathbb{N}}$, and the sets of (profiles of co-players') superstrategies $\mathbf{S}_i(\mathbf{h}_i)$, $\mathbf{S}_{-i}(\mathbf{h}_i)$ allowing any personal history $\mathbf{h}_i \in \mathbf{H}_i$ of the repeated game.

Under the assumption of perfect memory, the relevant collection of conditioning events for the CPSs of player i is

$$\mathfrak{C}_i = \{\mathbf{S}_{-i}(\mathbf{h}_i) : \mathbf{h}_i \in \mathbf{H}_i\}.$$

In particular, systems of conjectures of player i are represented by elements of $\Delta^{\mathfrak{C}_i}(\mathbf{S}_{-i})$.

Definition 4. *Fix $i \in I$ and $\mathbf{s}_i \in \mathbf{S}_i$. Two profiles of superstrategies $\bar{\mathbf{s}}_{-i} \in \mathbf{S}_{-i}$ and $\mathbf{s}_{-i} \in \mathbf{S}_{-i}$ are* ***observationally equivalent*** *given $\mathbf{s}_i$ if $\bar{\mathbf{h}}_i(\boldsymbol{\zeta}(\mathbf{s}_i, \bar{\mathbf{s}}_{-i})) =$*

$\bar{\mathbf{h}}_i\left(\boldsymbol{\zeta}(\mathbf{s}_i, \mathbf{s}_{-i})\right)$.

At the end of every period an outcome of the one-period game Γ is generated. We endow every player i with a discount factor $\delta_i \in [0, 1)$ representing his intertemporal preferences over streams of outcomes.

When players are **impatient**, meaning that they have zero discount factor, no payoff can be meaningfully attached to infinite terminal histories. To cope with this issue and generalize the analysis over any possible $\delta_i \in [0, 1)$, we directly rely on a form of **sequential rationality** based on **continuation values**, that is, we require players to take, at every personal history, choices that maximize the discounted expected utility computed *at that point in time*. If the discount factor is strictly positive, this implies *ex ante* expected utility maximization.

For each infinite history $\mathbf{z} \in Z^{\mathbb{N}}$, let $\mathrm{z}_t(\mathbf{z})$ denote its t-th coordinate projection, that is, the one-period terminal history played in period t. The continuation value of superstrategy $\mathbf{s}_i \in \mathbf{S}_i$ at personal history $\mathbf{h}_i$, given $\mu^i \in \Delta^{\mathfrak{C}_i}(\mathbf{S}_{-i})$, is

$$V^{\mu^i}_{i,\mathbf{h}_i}(\mathbf{s}_i) = \sum_{t=\tau(\mathbf{h}_i)+1}^{\infty} \delta_i^{t-\tau(\mathbf{h}_i)-1} \cdot \int_{\mathbf{S}_{-i}(\mathbf{h}_i)} u_i(\mathrm{z}_t(\boldsymbol{\zeta}(\mathbf{s}_i|\mathbf{h}_i, \mathbf{s}_{-i})))\mu^i(\mathrm{d}\mathbf{s}_{-i}|\mathbf{S}_{-i}(\mathbf{h}_i)),$$

where $\mathbf{s}_i|\mathbf{h}_i$ is the modified superstrategy allowing $\mathbf{h}_i$ and playing like $\mathbf{s}_i$ at each personal history that does not strictly precede $\mathbf{h}_i$.

We provided all the elements to define and analyze the infinite repetition of the multistage game Γ: letting $\delta = (\delta_i)_{i\in I}$,

$$\Upsilon(\Gamma, \delta) = \langle I, (A_i, M_i, \mathcal{A}_i, u_i, f_i, \delta_i)_{i\in I}\rangle.$$

4. Rationality and strategic reasoning

We begin this section with our definition of rationality, based on the previously introduced continuation values. We then connect rationality and one-period rationality, and we characterize and connect the behavioral and first-order belief implications of (one-period) rationality and common strong belief in (one-period) rationality, i.e., RCSBR (one-period RCSBR).

4.1 Rational planning

Player i is rational if she plays a strategy that satisfies one-step optimality given her CPS. This definition of rationality can be seen as a generalization of folding-back optimality to the infinite-horizon case.

Definition 5. *A superstrategy* $\mathbf{s}_i$ *is one-step optimal in the repeated game given* $\mu^i \in \Delta^{\mathfrak{C}_i}(\mathbf{S}_{-i})$*—written* $\mathbf{s}_i \in \mathcal{BR}_i(\mu^i)$*—if, for all* $\mathbf{h}_i \in \mathbf{H}_i$,

$$\mathbf{s}_i(\mathbf{h}_i) \in \arg \max_{a_i \in \mathcal{A}_i^{\ell(\mathbf{h}_i)+1}(\mathbf{h}_i)} V_{i,\mathbf{h}_i}^{\mu^i}(\mathbf{s}_i|_{\mathbf{h}_i} a_i),$$

where $\mathbf{s}_i|_{\mathbf{h}_i} a_i$ *is the superstrategy that allows* $\mathbf{h}_i$*, plays* a_i *at* $\mathbf{h}_i$ *and behaves like* $\mathbf{s}_i$ *at any other personal history that does not precede* $\mathbf{h}_i$*. Similarly, a strategy* s_i *is one-step optimal in the one-period game given a one-period CPS* $\gamma^i \in \Delta^{\mathcal{C}_i}(S_{-i})$*—written* $s_i \in \mathcal{BRO}_i(\gamma^i)$*—if, for all* $h_i \in H_i$,

$$s_i(h_i) \in \arg \max_{a_i \in \mathcal{A}_i^{\ell(h_i)+1}(h_i)} V_{i,h_i}^{\gamma^i}(s_i|_{h_i} a_i).$$

We point out that we operate under the *one-shot deviation principle*: since discounting implies continuity at infinity, one-step optimality is equivalent to sequential optimality, that is, maximization with respect to continuation (super)strategies conditional on every personal history.[17]

We are interested in studying the implications for the one-period game of assumptions *on the infinite interaction* for impatient players. Therefore, we need to identify one-period objects induced by the infinite repetition. First notice that from superstrategy $\mathbf{s}_i$, for every period t and every history $z^{[t-1]}$ that describes the play up to, but excluding, period t, one can derive the **strategy induced by** $\mathbf{s}_i$ **at** $z^{[t-1]}$: it starts by playing like $\mathbf{s}_i$ at $\bar{\mathbf{h}}_i(z^{[t-1]})$, and for every one-period personal history h_i it plays the action prescribed by $\mathbf{s}_i$ at the concatenation $(\bar{\mathbf{h}}_i(z^{[t-1]}), h_i)$.

Similarly, a CPS over the infinite repetition specifies, given $\bar{\mathbf{h}}_i(z^{[t-1]})$, a system of conjectures for every personal history of the one-period game following

[17] Furthermore, sequential optimality (hence, one-step optimality) is realization equivalent to the requirement that the continuations of of the given strategy maximize expected utility starting from histories consistent with it. Rationalizability for sequential games is often equivalently defined by means of this weaker version of sequential optimality. See Appendix 8.4 for details.

$\bar{h}_i(z^{[t-1]})$. These conjectures are probability measures over superstrategies, but induce conjectures over co-players' strategies for the relevant one-period game: the probability assigned to a profile of strategies s_{-i} is just the probability assigned to the (closed) set of profiles of superstrategies inducing s_{-i}. Remark 1 states that this system of conjectures is indeed a CPS of the one-period game, which we thus refer to as **induced one-period CPS**.

Remark 1. *Fix $\mu^i \in \Delta^{\mathfrak{C}_i}(\mathbf{S}_{-i})$, a period $t \in \mathbb{N}$, and a path $z^{[t-1]} \in Z^{[t-1]}$. Then the system of marginal probabilities induced by μ^i on the co-players' strategies played after $\bar{\mathbf{h}}_i(z^{[t-1]})$ is a CPS of the one-period game.*

If rational players only care about the present, then in every one-period game they should act so as to maximize their current one-period expected utility. Similarly, if they behave in such way then they can be seen as rational impatient players. Proposition 1 formalizes this fact.

Proposition 1. *When player i is impatient, a superstrategy $\bar{\mathbf{s}}_i$ is one-step optimal given $\mu^i \in \Delta^{\mathfrak{C}_i}(\mathbf{S}_{-i})$ if and only if, for every period t and path $z^{[t-1]}$, the strategy induced in the corresponding one-period game is one-step optimal given the induced one-period CPS.*

Remark 2. *For every player and every CPS (one-period CPS), there always exists a one-step optimal superstrategy (strategy).*[18]

4.2 Strategic thinking and strong rationalizability

Battigalli & Tebaldi (2019) extend the analysis of rationality and common strong belief in rationality of Battigalli & Siniscalchi (2002) to a class of infinite sequential games, which includes the infinite repetition of finite one-period games (simultaneous or sequential). To provide perspective for our results, it is useful to relate to their work. Events about behavior and interactive strategic thinking can be defined within the canonical type structure $\left(\beta_i : \mathbf{T}_i \to \Delta^{\mathfrak{C}_i}\left(\mathbf{S}_{-i} \times \mathbf{T}_{-i}\right)\right)_{i \in I}$ based on the given multistage game—in our case, the infinitely repeated game: $\mathbf{T}_i$ is the space of epistemic types (ways of thinking) of player i, that is, infinite hierarchies of conditional probability systems based on the countable collection

[18]See Appendix 8.4.

$\{\mathbf{S}_{-i}(\mathbf{h}_i) \times \mathbf{T}_{-i}\}_{\mathbf{h}_i \in \mathbf{H}_i}$ of conditioning events corresponding to personal histories; $\beta_i(\mathbf{t}_i) = (\beta_{i,\mathbf{h}_i}(t_i))_{\mathbf{h}_i \in \mathbf{H}_i}$ (with $\beta_{i,\mathbf{h}_i}(t_i) \in \Delta(\mathbf{S}_{-i}(\mathbf{h}_i) \times \mathbf{T}_{-i})$ for each $\mathbf{h}_i \in \mathbf{H}_i$) is the CPS over superstrategies and types of the co-players associated with type (infinite hierarchy) $\mathbf{t}_i$, and belief map β_i is a homeomorphism.[19] With this,

- an **event** about player i is a measurable subset of $\mathbf{S}_i \times \mathbf{T}_i$;
- R_i is event "i is **rational**," that is, the set of i-states $(\mathbf{s}_i, \mathbf{t}_i)$ such that $\mathbf{s}_i$ is one-step (hence, sequentially) optimal given the first-order CPS in hierarchy $\mathbf{t}_i$, which is obtained from the marginal on $\mathbf{S}_{-i}$ of each conditional conjecture $\beta_{i,\mathbf{h}_i}(\mathbf{t}_i)$;
- $\mathrm{SB}_i(E_{-i})$ is the event that i **strongly believes** E_{-i}, that is, CPS $\beta_i(\mathbf{t}_i)$ assigns probability 1 to E_{-i} whenever $E_{-i} \cap (\mathbf{S}_{-i}(\mathbf{h}_i) \times \mathbf{T}_{-i}) \neq \emptyset$;
- $R_i^{m+1} = R_i^m \cap \mathrm{SB}_i\left(R_{-i}^{m-1}\right)$, with $R_i^1 = R_i$; for example, R_i^2 is the event that i is rational and strongly believes in the co-players' rationality;
- **rationality and common strong belief in rationality** (RCSBR) is event $\times_{i\in I} R_i^\infty = \times_{i\in I} \cap_{m=1}^\infty R_i^m$;
- finally note that $\cap_{m=1}^\infty \mathrm{SB}_i\left(R_{-i}^{m-1}\right) = \mathrm{SB}_i\left(R_{-i}^\infty\right)$; thus, $(\mathbf{s}_i, \mathbf{t}_i) \in R_i^\infty$ satisfies the **best rationalization principle**: for every $i \in I$ and $m \in \mathbb{N} \cup \{\infty\}$, $\beta_{i,\mathbf{h}_i}(\mathbf{t}_i)$ *assigns probability* 1 *to* R_{-i}^m *whenever* $R_{-i}^m \cap (\mathbf{S}_{-i}(\mathbf{h}_i) \times \mathbf{T}_{-i}) \neq \emptyset$, in particular, $\beta_{i,\mathbf{h}_i}(\mathbf{t}_i)\left(R_{-i}^\infty\right) = 1$ whenever $R_{-i}^\infty \cap (\mathbf{S}_{-i}(\mathbf{h}_i) \times \mathbf{T}_{-i}) \neq \emptyset$.

Of course, a similar analysis applies to all finite games (see Battigalli & Siniscalchi 2002), including the one-period games considered here. We are interested in the *implications of RCSBR for strategic behavior and conjectures* about co-players' behavior (first-order beliefs). Building on Battigalli & Tebaldi (2019) and adapting their results, one can show that such implications are *characterized by the strong rationalizability solution concept* defined below.

Definition 6. *For every player $i \in I$, let*

$$\mathbf{\Sigma}_i^1 = \{(\mathbf{s}_i, \mu^i) \in \mathbf{S}_i \times \Delta^{\mathfrak{C}_i}(\mathbf{S}_{-i}) : \mathbf{s}_i \in \mathcal{BR}_i(\mu^i)\},$$

[19] What really matters is that the type structure à la Battigalli & Siniscalchi features continuous and onto belief maps.

$$\Sigma_i^1 = \{(s_i, \gamma^i) \in S_i \times \Delta^{\mathcal{C}_i}(S_{-i}) : s_i \in \mathcal{BRO}_i(\gamma^i)\},$$

and recursively define, for each $k \in \mathbb{N}$,

$$\begin{aligned}\mathbf{\Sigma}_i^{k+1} &= \{(\mathbf{s}_i, \mu^i) \in \mathbf{S}_i \times \Delta^{\mathfrak{C}_i}(\mathbf{S}_{-i}) : \mathbf{s}_i \in \mathcal{BR}_i(\mu^i),\ \forall m \leq k, \forall \mathbf{h}_i \in \mathbf{H}_i, \\ &\quad \mathrm{proj}_{\mathbf{S}_{-i}} \mathbf{\Sigma}_{-i}^m \cap \mathbf{S}_{-i}(\mathbf{h}_i) \neq \emptyset \Rightarrow \mu^i(\mathrm{proj}_{\mathbf{S}_{-i}} \mathbf{\Sigma}_{-i}^m | \mathbf{S}_{-i}(\mathbf{h}_i)) = 1\},\end{aligned}$$

and

$$\begin{aligned}\Sigma_i^{k+1} &= \{(s_i, \gamma^i) \in S_i \times \Delta^{\mathcal{C}_i}(S_{-i}) : s_i \in \mathcal{BRO}_i(\gamma^i), \forall m \leq k, \forall h_i \in H_i, \\ &\quad \mathrm{proj}_{S_{-i}} \Sigma_{-i}^m \cap S_{-i}(h_i) \neq \emptyset \Rightarrow \gamma^i(\mathrm{proj}_{S_{-i}} \Sigma_{-i}^m | S_{-i}(h_i)) = 1\},\end{aligned}$$

where $\mathbf{\Sigma}_{-i}^m = \prod_{j \neq i} \mathbf{\Sigma}_j^m$ and $\Sigma_{-i}^m = \prod_{j \neq i} \Sigma_j^m$. Then let

$$\mathbf{\Sigma}_i^\infty = \cap_{k \in \mathbb{N}} \mathbf{\Sigma}_i^k \quad \text{and} \quad \Sigma_i^\infty = \cap_{k \in \mathbb{N}} \Sigma_i^k.$$

Any profile $(\mathbf{s}_i, \mu^i)_{i \in I} \in \prod_{i \in I} \mathbf{\Sigma}_i^\infty$ (one-period profile $(s_i, \mu^i)_{i \in I} \in \prod_{i \in I} \Sigma_i^\infty$) is called ***strongly rationalizable****; for each player $i \in I$, superstrategy $\mathbf{s}_i$ (strategy s_i) is strongly rationalizable if $\mathbf{s}_i \in \mathrm{proj}_{\mathbf{S}_i} \mathbf{\Sigma}_i^\infty$ ($s_i \in \mathrm{proj}_{S_i} \Sigma_i^\infty$), and CPS μ^i (one-period CPS γ^i) is strongly rationalizable if $\mu^i \in \mathrm{proj}_{\Delta^{\mathfrak{C}_i}(\mathbf{S}_{-i})} \mathbf{\Sigma}_i^\infty$ ($\gamma^i \in \mathrm{proj}_{\Delta^{\mathcal{C}_i}(S_{-i})} \Sigma_i^\infty$).*

Example 2. Consider again the one-period game depicted in Figure 1 of the Introduction. Formally, one-period strong rationalizability works as follows:

$$\begin{aligned}\Sigma_{\mathbf{1}}^1 &= \{\mathrm{Out.u}\} \times \left\{\gamma^{\mathbf{1}} : \gamma^{\mathbf{1}}(r|\{\ell, r\}) \geq \frac{1}{2}\right\} \cup \{\mathrm{In.u}\} \times \left\{\gamma^{\mathbf{1}} : \gamma^{\mathbf{1}}(\ell|\{\ell, r\}) \geq \frac{1}{2}\right\}, \\ \Sigma_{\mathit{2}}^1 &= \{\ell\} \times \left\{\gamma^{\mathit{2}} : \gamma^{\mathit{2}}(\mathrm{In.d}|\{\mathrm{In.u}, \mathrm{In.d}\}) \geq \frac{1}{2}\right\} \cup \{r\} \times \left\{\gamma^{\mathit{2}} : \gamma^{\mathit{2}}(\mathrm{In.u}|\{\mathrm{In.u}, \mathrm{In.d}\}) \geq \frac{1}{2}\right\},\end{aligned}$$

$$\begin{aligned}\Sigma_{\mathbf{1}}^2 &= \Sigma_{\mathbf{1}}^1, \\ \Sigma_{\mathit{2}}^2 &= \{r\} \times \left\{\gamma^{\mathit{2}} : \gamma^{\mathit{2}}(\mathrm{In.u}|\{\mathrm{In.u}, \mathrm{In.d}\}) = 1\right\},\end{aligned}$$

$$\begin{aligned}\Sigma_{\mathbf{1}}^3 &= \{\mathrm{Out.u}\} \times \left\{\gamma^{\mathbf{1}} : \gamma^{\mathbf{1}}(r|\{\ell, r\}) = 1\right\}, \\ \Sigma_{\mathit{2}}^3 &= \Sigma_{\mathit{2}}^2,\end{aligned}$$

and, for all $k > 3$,

$$\begin{aligned}
\Sigma_{\mathbf{1}}^{k} &= \Sigma_{\mathbf{1}}^{3}, \\
\Sigma_{\mathit{2}}^{k} &= \{r\} \times \left\{\gamma^{\mathit{2}} : \gamma^{\mathit{2}}\left(\text{Out.u}|S_{\mathbf{1}}\right) = 1, \gamma^{\mathit{2}}\left(\text{In.u}|\left\{\text{In.u}, \text{In.d}\right\}\right) = 1\right\},
\end{aligned}$$

The last equality rests on the best rationalization principle: even if, by strategic reasoning, player *2* is initially certain of Out, if upon observing In (a non-rationalizable choice of player **1**), he would still believe that player **1** is—at least—rational, and that she is going up in the subgame. ▲

We can use the aforementioned characterization result to study the implications of RCSBR in the case of impatient players. The intuition is that the behavior of sophisticated impatient players is consistent with one-period strong rationalizability at every history (by Proposition 1 and an inductive argument); thus, each player at the beginning of the game expects with probability 1 such co-players' behavior. As long as players carry out strongly rationalizable strategies—by the *best rationalization principle* embedded in RCSBR—common strong belief in rationality implies that players keep assigning probability 1 to the co-players' strongly rationalizable strategies, even if they observe personal histories to which their earlier conjectures assigned probability 0. The same argument applies to the second period and all subsequent periods. The following result formalizes this intuition. Since we are ultimately interested in players' behavior and their beliefs about co-players' behavior, in our formal analysis we only consider "first-order states" $(\mathbf{s}_i, \mu^i)_{i \in I} \in \prod_{i \in I} \mathbf{S}_i \times \Delta^{\mathfrak{C}_i}(\mathbf{S}_{-i})$, the induced path of play $\boldsymbol{\zeta}(\mathbf{s})$, and players' beliefs held upon observing personal histories that realize as path $\boldsymbol{\zeta}(\mathbf{s})$ unfolds. In particular, we say that players **satisfy strong rationalizability** when we consider a strongly rationalizable state $(\mathbf{s}_i, \mu^i)_{i \in I} \in \prod_{i \in I} \boldsymbol{\Sigma}_i^{\infty}$.

Theorem 1. *Let players be impatient and satisfy strong rationalizability. Then, in every period, on the actual path of play, players carry out one-period strongly rationalizable strategies and hold strongly rationalizable one-period CPSs.*

Of course, one-period strong rationalizability is not implied after deviations from strong rationalizability (hence, off path), such deviations are rationalized ascribing to co-players lower levels of sophisticated reasoning.

5. Learning

Let us fix a profile $((\mathbf{s}_i, \mu^i))_{i\in I} \in \prod_{i\in I} \mathbf{S}_i \times \Delta^{\mathfrak{C}_i}(\mathbf{S}_{-i})$, which we interpret as the *true state*, that is, the profile of superstrategies actually played and CPSs actually held by players.[20] Players' superstrategies yield a path of play $\boldsymbol{\zeta}(\mathbf{s})$ in the repeated game and corresponding personal histories observed by each player. In particular, we focus on the sequence of personal histories observed at the beginning of each period (or end of the previous period) and the corresponding sequence of conjectures about co-players' superstrategies and one-period strategies implied by each player's CPS at the given state. For every $t \in \mathbb{N}$, we let $\mathbf{h}_i^t$ denote the personal history of i at the beginning of period t induced by the true state, that is, $\mathbf{h}_i^t \prec \bar{\mathbf{h}}_i(\boldsymbol{\zeta}(\mathbf{s}))$ with $\ell(\mathbf{h}_i^t) = L(t-1)$. We refer to the belief of player i upon observing $\mathbf{h}_i^t$ as his conjecture at time t. For pedagogical purposes, we first consider convergence of conjectures in finite time and then the more general case of asymptotic convergence.

Definition 7. *The conjectures of player i **have converged from time** T if, for every $t \geq T$,*

$$\mu^i(\cdot|\mathbf{S}_{-i}(\mathbf{h}_i^t)) = \mu^i(\cdot|\mathbf{S}_{-i}(\mathbf{h}_i^T)).$$

Let $\hat{t} > t$. For every $E_{-i} \subseteq \mathbf{S}_{-i}(\mathbf{h}_i^{\hat{t}})$, the chain rule implies

$$\mu^i(E_{-i}|\mathbf{S}_{-i}(\mathbf{h}_i^t)) = \mu^i(E_{-i}|\mathbf{S}_{-i}(\mathbf{h}_i^{\hat{t}})) \cdot \mu^i(\mathbf{S}_{-i}(\mathbf{h}_i^{\hat{t}})|\mathbf{S}_{-i}(\mathbf{h}_i^t)).$$

Hence, convergence requires that $\mu^i(\mathbf{S}_{-i}(\mathbf{h}_i^{\hat{t}})|\mathbf{S}_{-i}(\mathbf{h}_i^t)) = 1$. Since this must hold for all $\hat{t}, t \geq T$, we obtain the following characterization.

Remark 3. *Player i's conjectures have converged from time T if and only if the conjecture conditional on the observed personal history at every $t \geq T$, $\mu^i(\cdot|\mathbf{S}_{-i}(\mathbf{h}_i^t))$, assigns probability 1 to the set $\mathbf{S}_{-i}\left(\bar{\mathbf{h}}_i(\boldsymbol{\zeta}(\mathbf{s}_i, \mathbf{s}_{-i}))\right)$ of co-players' superstrategies that are observationally equivalent, given i's own superstrategy, to the true ones.*

Hence, convergence of conjectures is equivalent to a form of learning, that is, eventually acquiring the ability to perfectly forecast the future messages one will observe. Intuitively, a player that is certain, and correct, about the message

[20]From an epistemic perspective, we are considering the behavior and first-order beliefs at some state of the world $(\mathbf{s}_i, \mathbf{t}_i)_{i\in I}$ in the canonical type structure.

she will observe in every period, has no reason to change her conjectures about others' behavior. We express convergence of conjectures after a finite number of periods with the phrase "learning in finite time", whereas asymptotic convergence (formally defined below) is expressed as "asymptotic learning", which is weaker than the former. Note that we do not mean that players "learn the truth", we only mean that they stop changing (from some T or in the limit) their beliefs about co-players' behavior.

Since conjectures are updated according to the chain rule, a sufficient condition for asymptotic learning is the well-known requirement that conjectures assign positive probability to the relevant part of the "true state of the world", at least from some time onward (at some finite history on path). In our setting, this relevant part is the set of co-players' superstrategies observationally equivalent to the true ones—given the player's feedback and her own superstrategy. We call this property "observational grain of truth". Our results are similar to the ones of Kalai & Lehrer (1993), extending them to the case of imperfect monitoring and multistage one-period games.[21] In the "medium run", for every $\varepsilon > 0$, there exists a time starting from which conjectures are "ε-close" to the objective distribution of observations. In our case, "ε-closeness" means that the conjecture assigns probability at least $1 - \varepsilon$ to the set of superstrategies observationally equivalent to the true ones.

Definition 8. *We say that **observational grain of truth** holds for player i if there exists a time $T \in \mathbb{N}$ such that*[22]

$$\mu^i(\mathbf{S}_{-i}(\bar{\mathbf{h}}_i(\boldsymbol{\zeta}(\mathbf{s}_i, \mathbf{s}_{-i})))|\mathbf{S}_{-i}(\mathbf{h}_i^T)) > 0.$$

Proposition 2. *If observational grain of truth holds for player i, then, for every $\varepsilon > 0$, there exists a time T such that, for all $t \geq T$*

$$\mu^i(\mathbf{S}_{-i}(\bar{\mathbf{h}}_i(\boldsymbol{\zeta}(\mathbf{s}_i, \mathbf{s}_{-i})))|\mathbf{S}_{-i}(\mathbf{h}_i^t)) \geq 1 - \varepsilon.$$

[21] Yet, we simplify the analysis of learning by focusing on pure (super)strategies.

[22] While in our definitions we refer to conjectures held at the beginning of periods, both Definition 7 and Definition 8 can be given equivalently in terms of personal histories of general length k.

Therefore,

$$lim_{t\to\infty}\mu^i(\mathbf{S}_{-i}(\bar{\mathbf{h}}_i(\boldsymbol{\zeta}(\mathbf{s}_i,\mathbf{s}_{-i})))|\mathbf{S}_{-i}(\mathbf{h}_i^t)) = 1.$$

Of course, learning in finite times implies observational grain of truth. More generally, we obtain the following corollary.

Corollary 1. *Asymptotic learning ($lim_{t\to\infty}\mu^i(\mathbf{S}_{-i}(\bar{\mathbf{h}}_i(\boldsymbol{\zeta}(\mathbf{s}_i,\mathbf{s}_{-i})))|\mathbf{S}_{-i}(\mathbf{h}_i^t)) = 1$) and observational grain of truth are equivalent.*

6. Strong rationalizability, learning, and equilibrium

In this section we provide the main results of the paper. We begin by formally defining SCE and SCE in strongly rationalizable conjectures. In Section 6.1 we discuss the implications of learning in finite time. Section 6.2 illustrates the concepts with some examples. Section 6.3 analyzes asymptotic learning.

Definition 9. *State $((\mathbf{s}_i,\mu^i))_{i\in I}$ is a **self-confirming equilibrium (SCE)** if, for every i:*

(i) (confirmation of conjectures) $\mu^i\left(\mathbf{S}_{-i}(\bar{\mathbf{h}}_i(\boldsymbol{\zeta}(\mathbf{s})))|\mathbf{S}_{-i}\right) = 1$*;*

(ii) (rationality) $\mathbf{s}_i \in \mathcal{BR}_i(\mu^i)$.

*The state $((\mathbf{s}_i,\mu^i))_{i\in I}$ is an **SCE in strongly rationalizable conjectures** if it is an SCE of the infinitely repeated game $\Upsilon(\Gamma,\delta)$ such that, for every $i\in I$, $(\mathbf{s}_i,\mu^i)\in\mathbf{\Sigma}_i^\infty$.*

The one-period counterparts are analogously defined. If condition (i) of Definition 9 is replaced by "(i') there exists $T\in\mathbb{N}$ such that $\mu^i\left(\mathbf{S}_{-i}(\bar{\mathbf{h}}_i(\boldsymbol{\zeta}(\mathbf{s})))|\mathbf{S}_{-i}(\mathbf{h}_i^T)\right) = 1$", we call the state an **eventual SCE** (in strongly rationalizable conjectures).

6.1 Implications of learning in finite time and RCSBR

Here we study the implications on play and beliefs when conjectures have converged in a finite number of periods, that is, under learning in finite time. Theorem 2 below states the implications of strong rationalizability for impatient players whose conjectures converge in a finite number of periods.

Theorem 2. *Let players be impatient. Assume strong rationalizability is satisfied and there is a time $T > 0$ starting from which all players' conjectures have converged. Then the state induces, starting from T, a sequence of one-period self-confirming equilibria in strongly rationalizable conjectures.*

We use the first part of this subsection to prove the theorem through a series of other results, that we find interesting on their own.

We first take a further step in connecting confirmation of conjectures (about the repeated play) and one-period conjectures

Remark 4. *For every player $i \in I$, for every $\varepsilon \geq 0$ and $T \in \mathbb{N}$ such that, for all $t \geq T$,*

$$\mu^i(\mathbf{S}_{-i}(\bar{\mathbf{h}}_i(\boldsymbol{\zeta}(\mathbf{s}_i, \mathbf{s}_{-i})))|\mathbf{S}_{-i}(\mathbf{h}_i^t)) \geq 1 - \varepsilon,$$

all one-period CPSs $(\mu_t^i)_{t \geq T}$, induced by μ^i at every period $t \geq T$ starting at personal history $\mathbf{h}_i^t$, satisfy

$$\mu_t^i(S_{-i}\left(\bar{\mathrm{h}}_i(\zeta(s^t))\right)|S_{-i}) \geq 1 - \varepsilon.$$

Remark 4 tells us that ε-confirmed beliefs about the infinite interaction induce one-period ε-confirmed beliefs. This is also true for $\varepsilon = 0$, which is what we use in the current theorem. However, we state a more general result because it will be useful later. The next result follows from this observation and Theorem 1.

Corollary 2. *If players are impatient, every SCE (in strongly rationalizable conjectures) of the infinite repetition induces a sequence of one-period SCEs (in strongly rationalizable conjectures).*

Hence, it is possible to translate results on SCE in results on sequences of one-period SCEs. Definition 9 and Theorem 1 also yield the following fact.

Remark 5. *If strong rationalizability is satisfied and conjectures converge in finite time, then the true state of the game must feature an eventual SCE in strongly rationalizable conjectures.*

Corollary 2 and Remark 5 are all we need to prove Theorem 2. If players are impatient, their conjectures converge from time T, and strong rationalizability is satisfied, then the true state of the game must induce, from time T onward, a sequence of one-period SCEs in strongly rationalizable conjectures.

In the rest of this subsection we show a kind of converse result: every sequence of one-period SCEs in strongly rationalizable conjectures can be induced under the assumptions stated in Theorem 2.

Theorem 3. *Let players be impatient. Every sequence of one-period SCEs in strongly rationalizable conjectures starting from some time T can be induced by a strongly rationalizable state that features learning in finite time.*

The key ingredient to obtain this result is some kind of converse of Theorem 1: when (one-period) strong rationalizability is satisfied and players are impatient, we can find a state consistent with their play in which strong rationalizability for the infinite repetition is satisfied.

Proposition 3. *Let players be impatient. If an infinite (i.e., terminal) history of the repeated game consists of a sequence of one-period terminal histories consistent with one-period strong rationalizability, then there exists a profile of strongly rationalizable superstrategies that induce it.*

Since a sequence of (one-period) strongly rationalizable strategies induces a sequence of one-period terminal histories consistent with one-period strong rationalizability, Proposition 3 allows us to conclude that such sequence of strategies can be induced on path by a profile of strongly rationalizable superstrategies. Then one can find strongly rationalizable CPSs that justify these superstrategies as sequential best replies and are consistent with the sequence of strongly rationalizable one-period CPSs, which is turn are part of the sequence of one-period equilibria posited in Theorem 3. Confirmation of conjectures for such one-period CPSs implies confirmation for the CPSs on the infinite repetition.

A result similar to Theorem 3 also holds when players have positive discount factors.

6.2 Examples

We illustrate some of the concepts and results presented so far with examples.

Example 3. Let us go back to the entry-game example of Figure 1. As explained above, the strongly rationalizable strategies are: $s_{\mathbf{1}}$ =Out.u and s_{2} =r. Hence, the only outcome of SCE in rationalizable conjectures is Out. Strategy pair (In.u, ℓ) is part of an SCE (supported by a non-rationalizable conjecture of player *2*): player **1** holds a correct conjecture, player *2* assigns probability 1 to In (observed), and at least 1/2 to In.d. This SCE can occur in the limit if we remove common strong belief in rationality, and maintain rationality and observational grain of truth. ▲

The following example illustrates the difference between rationalizable self-confirming equilibrium and SCE in rationalizable conjectures. Intuitively, the former concept follows from rationality, confirmation of conjectures, and *common belief in both.*[23]

Example 4. Consider a discrete location game with two players and simultaneous moves. Each player can choose among the same 4 locations, $\{1, 2, 3, 4\}$. Players only observe their realized payoffs, which are negatively proportional to the distance between their location choice and the co-player's location: $u_i(a_i, a_j) = -|a_i - a_{-i}|$. With this, all actions are rationalizable. The case in which one player chooses location 2 and the other chooses location 3 is part of an SCE in rationalizable conjectures: It can be justified by conjectures assigning (sufficiently close to) uniform probability to the other being in a location at distance 1. These conjectures assign probability (close to) 1/2 to 1—for the player in location 2—and to 4—for the player in location 3. Locations 1 and 4 are inconsistent with rationality and confirmation of conjectures upon observing distance 1. Therefore, this is not a rationalizable self-confirming equilibrium: belief in rationality and confirmation of conjectures is violated. ▲

The last example of this section illustrates the difference between convergence of beliefs over superstrategies and convergence of marginal one-period beliefs over strategies.

Example 5 (Battle of the Sexes)**.** Let the one-period game be the Battle of the Sexes (BoS):

[23] We consider a simple discretized version of Example (a) in Rubinstein & Wolinsky (1994), who first put forward the rationalizable SCE concept for games with simultaneous moves.

1 \ 2	B	S
B	2, 1	0, 0
S	0, 0	1, 2

Suppose there are observable actions, or—equivalently—that $f_i = u_i$ for each $i \in \{1, 2\}$. Thus, SCE, SCE in rationalizable conjectures and (pure) Nash equilibrium coincide. There exist strongly rationalizable superstrategies $\mathbf{s}_1$ and $\mathbf{s}_2$ inducing the alternated path $\zeta(\mathbf{s}_1, \mathbf{s}_2) = ((B, B), (S, S), (B, B), (S, S), \ldots)$ justified by CPSs that strongly believe in the superstrategy of the co-player. Such CPSs are confirmed, which means that we have convergence of conjectures from period $T = 1$. Thus, the above path is consistent with learning and strong rationalizability. Next, consider a different situation. Suppose that player 1 believes that player 2 wants to "cooperate" and play the alternated sequence of equilibria. In particular, 1 is sure that 2 will start playing B. Upon observing S, she would be sure that 2 decided to start with S, but would now play B. Upon observing S again, 1 would think that 2 was hoping that 1 would have come along, but now has understood 1's intention and would play B. Upon observing S for the fourth time, she would make the same identical reasoning. If 1 goes on with this kind of thinking, then her CPS is such that, at any time along a sequence of (B, S)'s, 1's one-period conjecture assigns probability 1 to 2 playing B. One can check that this is consistent with the definition of CPS. As a result, 1's one-period conjectures repeat steadily and unchanged in every period. However, they are not confirmed. Observe that, if 2 follows an analogous reasoning, with S in place of B, then the actual play is exactly an infinite sequence of (B, S)'s, motivated by conjectures over superstrategies that have not converged, but one-period conjectures that have converged. ▲

6.3 Medium-run and long-run implications of asymptotic learning and RCSBR

We now offer some results on the medium and long-run implications of asymptotic learning, or equivalently, observational grain of truth (Corollary 1) at a strongly rationalizable state. As shown in Section 5, under observational grain of truth conjectures (and CPSs) become ε-confirmed in finite time, that is, in the medium run. By strong rationalizability, impatient players carry out (on path) one-period sequential best reply to strongly rationalizable one-period CPSs. We can show that, after a sufficiently long time, by finiteness of the one-period game, these strategies are also sequential best replies to rationalizable "fully confirmed modifications" of these CPSs. Thus, the play is eventually consistent with one-period SCEs in strongly rationalizable conjectures. A similar result holds with positive discount factors.

Definition 10. *Fix a profile* $((\mathbf{s}_i, \mu^i))_{i\in I} \in \prod_{i\in I} \mathbf{S}_i \times \Delta^{\mathfrak{C}_i}(\mathbf{S}_{-i})$ *that satisfies observational grain of truth. For each* $i \in I$*, let* $\hat{\mathbf{h}}_i \prec \bar{\mathbf{h}}_i(\zeta(\mathbf{s}))$ *be the first on-path personal history such that such that* $\mu^i\left(\mathbf{S}_{-i}\left(\bar{\mathbf{h}}_i(\zeta(\mathbf{s}))\right)|\hat{\mathbf{h}}_i\right) > 0$*; the* ***fully confirmed modification of*** μ^i *is the CPS* $\hat{\mu}^i$ *such that*

(i) *for every* $\mathbf{h}_i \in \mathbf{H}_i$ *with* $\hat{\mathbf{h}}_i \preceq \mathbf{h}_i \prec \bar{\mathbf{h}}_i(\zeta(\mathbf{s}))$ *and every (measurable)* $E_{-i} \subseteq \mathbf{S}_{-i}$*,*

$$\hat{\mu}^i(E_{-i}|\mathbf{S}_{-i}(\mathbf{h}_i)) = \frac{\mu^i\left(E_{-i} \cap \mathbf{S}_{-i}\left(\bar{\mathbf{h}}_i(\zeta(\mathbf{s}))\right)|\mathbf{S}_{-i}(\mathbf{h}_i)\right)}{\mu^i\left(\mathbf{S}_{-i}\left(\bar{\mathbf{h}}_i(\zeta(\mathbf{s}))\right)|\mathbf{S}_{-i}(\mathbf{h}_i)\right)};$$

(ii) *for every other* $\mathbf{h}_i \in \mathbf{H}_i$*,* $\hat{\mu}^i\left(\mathbf{S}_{-i}\left(\bar{\mathbf{h}}_i(\zeta(\mathbf{s}))\right)|\mathbf{S}_{-i}(\mathbf{h}_i)\right) = \mu^i\left(\mathbf{S}_{-i}\left(\bar{\mathbf{h}}_i(\zeta(\mathbf{s}))\right)|\mathbf{S}_{-i}(\mathbf{h}_i)\right)$.

In words, $\hat{\mu}^i$ is an on-path rescaling of μ^i that assigns probability 1 to $\mathbf{S}_{-i}\left(\bar{\mathbf{h}}_i(\zeta(\mathbf{s}_i, \mathbf{s}_{-i}))\right)$ (the set of co-players' superstrategies observationally equivalent to the true ones, given i's superstrategy $\mathbf{s}_i$) when μ^i assigns strictly positive probability to $\mathbf{S}_{-i}\left(\bar{\mathbf{h}}_i(\zeta(\mathbf{s}_i, \mathbf{s}_{-i}))\right)$, preserving probability ratios within $\mathbf{S}_{-i}\left(\bar{\mathbf{h}}_i(\zeta(\mathbf{s}_i, \mathbf{s}_{-i}))\right)$. One can show that $\hat{\mu}^i$ is indeed a CPS. In particular, the chain rule is trivially satisfied as $0 = 0$ when we relate an "on-path conjecture" $\hat{\mu}^i(\cdot|\mathbf{S}_{-i}(\mathbf{h}_i))$, that is, with $\hat{\mathbf{h}}_i \preceq \mathbf{h}_i \prec \bar{\mathbf{h}}_i(\zeta(\mathbf{s}))$, to an "off-path conjecture" $\hat{\mu}^i(\cdot|\mathbf{S}_{-i}(\mathbf{h}'_i))$ with $\mathbf{h}_i \prec \mathbf{h}'_i \not\prec \bar{\mathbf{h}}_i(\zeta(\mathbf{s}))$. As we look at the one-period CPSs μ^i_t induced on path by μ^i, we obtain from $\hat{\mu}^i$ corresponding fully confirmend modifications $\hat{\mu}^i_t$. When the on path conjectures become "ε-confirmed", i.e., $\mu^i\left(\mathbf{S}_{-i}\left(\bar{\mathbf{h}}_i(\zeta(\mathbf{s}))\right)|\mathbf{S}_{-i}(\mathbf{h}^t_i)\right) \geq 1-\varepsilon$, the induced fully confirmed modifications are close to these on-path conjectures.

Definition 11. *A one-period* ***ε-self-confirming equilibrium in strongly rationalizable conjectures*** *is a profile* $((s_i, \gamma^i))_{i\in I} \in \prod_{i\in I} S_i \times \Delta^{\mathcal{C}_i}(S_{-i})$ *such that, for every* $i \in I$*,*

(i) $\gamma^i(S_{-i}(\bar{\mathrm{h}}_i(\zeta(s)))|S_{-i}) \geq 1-\varepsilon$*;*

(ii) $(s_i, \gamma^i) \in \Sigma_i^\infty$.

With this, we can state our main result.

Theorem 4. *Let players be impatient. If observational grain of truth and strong rationalizability hold for every player, then for every $\varepsilon > 0$ there exists a time T starting from which the state induces a sequence of one-period ε-self-confirming equilibria in strongly rationalizable conjectures. Furthermore, there exists some $\hat{T}$ starting from which, in each period, the induced profile of strategies is part of a one-period self-confirming equilibrium in strongly rationalizable conjectures, in which each player i's CPS is the fully confirmed modification of i's one-period CPS μ_t^i.*

The intuition is quite simple: after a sufficiently long time, players' one-period conjectures are almost confirmed. Since the one-period game is finite and strategies are sequential best-replies to one-period almost confirmed CPSs, they are also sequential best-replies to their fully-confirmed modifications. Since the on-path one-period CPSs are strongly rationalizable, it turns out that also their fully confirmed modifications are strongly rationalizable. The result follows.

We now provide a result with a similar flavor for players with positive discount factor. Once again, conjectures (about the infinite repetition) become ε-confirmed in finite time. As an ancillary result, by continuity at infinity, the (continuation) superstrategies are close to exact sequential best replies to the fully confirmed modifications. However, without full convergence—that only needs to happen asymptotically—(continuation) superstrategies need not be exact sequential best replies to the fully confirmed modifications.

We relegate to Appendix 8.3 the additional formalism of this subsection. In short, at every history $\mathbf{h}$ one can define the concepts of continuation superstrategies ($\mathbf{s}_i^{\succeq \bar{\mathbf{h}}_i(\mathbf{h})} \in \mathbf{S}_i^{\succeq \bar{\mathbf{h}}_i(\mathbf{h})}$), conditioning events ($\mathfrak{C}_i^{\succeq \bar{\mathbf{h}}_i(\mathbf{h})}$), CPSs ($\Delta^{\mathfrak{C}_i^{\succeq \bar{\mathbf{h}}_i(\mathbf{h})}}(\mathbf{S}_{-i})$), and continuation values for every player i starting from personal history $\bar{\mathbf{h}}_i(\mathbf{h})$, and define SCE of the continuation game with root $\mathbf{h}$.

Definition 12. *Fix a personal history $\mathbf{h}_i$. We say that a continuation superstrategy $\mathbf{s}_i^{\succeq \mathbf{h}_i}$ is a sequential ε-best reply to a (continuation) CPS ν^i, where $\varepsilon > 0$, if, for all $\bar{\mathbf{s}}_i^{\succeq \mathbf{h}_i}$ and $\mathbf{g}_i \succeq \mathbf{h}_i$,*

$$V_{i,\mathbf{g}_i}^{\nu^i}(\mathbf{s}_i^{\succeq \mathbf{h}_i}) \geq V_{i,\mathbf{g}_i}^{\nu^i}(\bar{\mathbf{s}}_i^{\succeq \mathbf{h}_i}) - \varepsilon.$$

Definition 13. *State* $(\mathbf{s}_i, \mu^i)_{i\in I}$ *is an* ***eventual*** $\boldsymbol{\varepsilon}$***-self-confirming equilibrium in strongly rationalizable conjectures*** *if there exists a time* T *such that, for every* $i \in I$,

(i) *for every personal history* $\mathbf{h}_i$ *with* $\mathbf{h}_i^T \preceq \mathbf{h}_i \prec \bar{\mathbf{h}}_i(\boldsymbol{\zeta}(\mathbf{s}))$,

$$\mu^i(\mathbf{S}_{-i}(\bar{\mathbf{h}}_i(\boldsymbol{\zeta}(\mathbf{s})))|\mathbf{S}_{-i}(\mathbf{h}_i)) \geq 1 - \varepsilon;$$

(ii) $(\mathbf{s}_i, \mu^i) \in \mathbf{\Sigma}_i^\infty$.

Theorem 5. *If strong rationalizability and observational grain of truth are satisfied, then for every* $\varepsilon > 0$ *the state is an eventual* ε*-self-confirming equilibrium in strongly rationalizable conjectures. Furthermore, for all* $\varepsilon' > 0$ *and* $i \in I$*, there exists* $\mathbf{h}_i$ *such that* $\mathbf{s}_i^{\succeq \mathbf{h}_i}$ *is a sequential* ε'*-best reply to the fully confirmed modification of* μ^i.

7. Literature review and discussion

In this paper we analyze the the medium and long-run behavior of strategically sophisticated rational players in infinitely repeated games with imperfect feedback. We model sophisticated strategic thinking by assuming common strong belief in rationality and prove that, under an "observational grain of truth" assumption, players' behavior and conjectures, i.e., first-order beliefs, converge to an SCE with strongly rationalizable conjectures of the repeated game. If players are impatient, in the long run they play SCEs with strongly rationalizable conjectures of the one-period game, but the one-period equilibrium may change over time. We also show that our assumptions are tight. We are now in a position to discuss the related literature in detail. While doing this, we consider the limitations and possible extensions of our work.

Drawing on Battigalli (1987), Battigalli & Guaitoli (1988) use the notion of SCE in (strongly) rationalizable conjectures to analyze economic policy in a macroeconomic game with incomplete information. This equilibrium concept is adapted and used by Schipper (2021) to analyze discovery and equilibrium in games with unawareness (lack of conception of some features of the game). Here we provide both an epistemic and a learning foundation to the equilibrium concept. Although we assume complete information, we can easily extend our results to environments with incomplete information about payoff functions, as in the epistemic analysis of Battigalli & Siniscalchi (2002) and Battigalli & Tebaldi (2019). We conjecture that our approach can be extended to analyze processes of learning and discovery as (impatient) agents repeatedly play a game with unawareness, but this is well beyond the scope of this paper.[24]

Fudenberg & Levine (1993) coined the term "self-confirming equilibrium." They put forward a notion of randomized SCE motivated by a population-game scenario whereby agents are drawn from large populations and randomly matched in every period to play a sequential game, so that randomized strategies of the one-period game are interpreted as stable statistical distributions of pure strategies within populations. In this case, conjecture-confirmation means that each agent assigns probability 1 to the set of co-players' randomized strategies inducing the actual frequency distribution of observations given her (pure) strategy. The large-population scenario also justifies one-period expected payoff maximization despite a positive discount factor, as agents understand that they cannot affect the behavior of future co-players, who are almost certainly different from their current co-players, and—in the long run—they also have no incentive to experiment. We do not consider a population-game scenario for two reasons. First, many recurrent interactions feature a fixed set of players. Second, the analysis would be technically more difficult. We relate to one-period game equilibria by assuming impatient players, while with patient players we obtain convergence to repeated-game SCE (cf. Kalai & Lehrer 1993, 1995). We conjecture that we could cover the case of large but finite populations allowing for chance moves and analyzing the population game as a grand game with (finitely) many agents partitioned according to their role. Another difference with Fudenberg & Levine (1993) is that, unlike us, they assume perfect feedback about chosen actions (terminal history/node) at the end of the one-period game. When the latter is a sequential game, co-players' one-period strategies are nonetheless imperfectly observable, which is what makes their SCE concept different from Nash equilibrium. Note, however, that under perfect feedback pure SCEs in two-person games are realization-equivalent to Nash equilibria.[25] Fudenberg &

[24]See the discussion in Schipper (2021), pages 3-4.

[25]The latter may be partially randomized off path. Cf. Battigalli (1987) and Fudenberg &

Kamada (2015, 2018) remove the perfect feedback assumption, positing a terminal information partition for each player.[26]

We explained in the Introduction the main conceptual difference between SCE with rationalizable conjectures and the rationalizable SCE concept of Rubinstein & Wolinsky (1994): unlike the former, the latter postulates common certainty of the confirmation of conjectures. This is argued informally in their paper, and it is formally proved in the epistemic analysis of Esponda (2013), who focuses on games with incomplete information. Another important difference between our work and these papers on rationalizable SCE is that they consider simultaneous-move games. While the SCE concept, which does not presume strategic sophistication, can be meaningfully applied to the strategic form of a sequential game,[27] notions of SCE with strategically sophisticated players must be adapted to take sequential moves into account, because their application to the strategic form of a sequential game with feedback would allow for non-credible threats.[28] Dekel et al. (1999) analyze a version of rationalizable SCE for sequential games with perfect feedback. As mentioned above, Fudenberg & Kamada (2015, 2018) allow for imperfect feedback. These papers on rationalizable SCE in sequential games feature a weak notion of strategic sophistication, as they assume that there is common certainty of rationality and conjecture confirmation at the beginning of the game, but not if players are surprised by moves that are compatible with such assumptions. We instead assume common strong belief in rationality. Yet, we do not assume common strong belief in confirmation and we do not allow for randomization; thus, the two concepts are not nested.[29]

The learning aspect of our paper is related to Kalai & Lehrer (1993) who analyze repeated games with perfect monitoring where each player knows her payoff function, and Kalai & Lehrer (1995) on repeated games with imperfect monitoring

Levine (1993).

[26]For this reason, they call the equilibrium "partition confirmed." Instead, we keep the same terminology independently of the information/feedback structure.

[27]Provided that also feedback, besides the payoff functions, is accurately represented in strategic form. See the discussion in Battigalli et al. (2019), who point out that this claim is not true when players are ambiguity averse.

[28]Rubinstein & Wolinsky (1994) write that their analysis concerns "normal-form games." They do not clarify whether they mean that the analysis can be meaningfully applied to the normal/strategic form of the given game with feedback. But it is obvious that this is not the case.

[29]Except, of course, within the class of simultaneous-move games, where rationalizable SCE refines SCE in rationalizable conjectures. See Example 4.

and imperfect knowledge of one's own payoff function. As in their work, we obtain convergence of conjectures about superstrategies from a kind of "grain of truth" condition. As in Kalai & Lehrer (1995), our condition concerns the personal observations made by each player, rather than the path of play. Furthermore, since we model beliefs as conditional probability systems, we can state this condition as something that holds eventually, that is, we allow for finitely many surprises. The most important difference between our work and these papers is that they do not assume sophisticated strategic thinking, which is the reason why only knowledge of one's own payoff function matters, rather than interactive knowledge about the game.

Finally, we do not model the information structure of the one-period game and of the repeated game by means of information partitions. We represent the flow of information accruing to players between stages and periods by means of feedback functions and thereby comply with the following "separation principle" of Battigalli & Generoso (2021): the description of the rules of the game is independent of players' personal features, such as their mnemonic abilities.[30] Besides this conceptual advantage, our representation allows to seamlessly blend information flows within each one-period game with repeated-game monitoring. To simplify the exposition, we assume a multistage structure (cf. Myerson 1986), but our analysis and results can be extended to more general sequential games represented as in Battigalli & Generoso (2021).

[30] As explained in Section 3, we informally assume that players have perfect memory. This assumption about players' mnemonic abilities is expressed formally in Battigalli & Generoso (2021).

8. Appendix

8.1 Additional formalism and remarks on game structure

We provide the formalism for all game objects defined implicitly, or somewhat informally, in the main text, and we provide related remarks to be used in the proofs of our results.[31] Let us define the different sets of histories for the one-period game:

- $\bar{H}$ is the set of **histories,** that is, the feasible sequences of action profiles including the empty sequence $a^{[0]} = \varnothing$ (root).

$$\bar{H} = \{\varnothing\} \cup \left\{ (a^k)_{k=1}^{\ell} : \ell \leq L, \forall k \in [\ell], a^k \in \prod_{i \in I} \mathcal{A}_i^k(f_i^{k-1}(a^{[k-1]})) \right\};$$

- $Z = \left\{ z \in \bar{H} : \ell(z) = L \right\}$ is the set of **terminal histories**;
- $H = \bar{H} \setminus Z$ is the set of **non-terminal histories**;
- The set of **personal histories** is

$$\begin{aligned} \bar{H}_i =& \{(a_i, m_i)^{[k]} \in (A_i \times M_i)^{[k]} : k \leq L, \exists a_{-i}^{[k]} \in \bar{H}_{a_i^{[k]}}, f_i^{[k]}(a_i^{[k]}, a_{-i}^{[k]}) = m_i^{[k]}\} \\ & \cup \{(f_i^0(\varnothing))\}, \end{aligned}$$

  and it is partitioned into Z_i and H_i (terminal and non-terminal personal histories).[32]

Remark 6. *If we endow $\bar{H}$ with the natural partial order on sequences $\preceq$, we obtain the objective game with root $\varnothing$, $(\bar{H}, \preceq)$. Endowing $\bar{H}_i$ with the partial order*

$$h_i \preceq g_i \quad \Leftrightarrow \quad a_i^{[\ell(h_i)]}(h_i) \preceq a_i^{[\ell(g_i)]}(g_i) \ \wedge \ m_i^{[\ell(h_i)]}(h_i) \preceq m_i^{[\ell(g_i)]}(g_i),$$

where $h_i, g_i \in \bar{H}_i$ and $\ell(h_i) \leq \ell(g_i)$, we obtain player i's subjective game with root $\varnothing$, $(\bar{H}_i, \preceq)$.

[31] The proofs of these quite intuitive remarks are available upon request.

[32] $\bar{H}_{a_i^{[k]}}$ is the section of $\bar{H}$ at $a_i^{[k]}$.

Recalling that $\zeta : S \to Z$ is the function associating each strategy profile with the induced terminal history, define the sets of strategies and strategy profiles consistent with "objective" histories $h \in \bar{H}$ as follows:

- $S(h) = \{s \in S : h \preceq \zeta(s)\}$;
- $S_i(h) = \{s_i \in S_i : \exists s_{-i} \in S_{-i}, h \preceq \zeta(s_{-i}, s_i)\} = \mathrm{proj}_{S_i} S(h)$;
- $S_{-i}(h) = \{s_{-i} \in S_{-i} : \exists s_i \in S_i, h \preceq \zeta(s_{-i}, s_i)\} = \mathrm{proj}_{S_{-i}} S(h)$.

We can define analogous sets for personal histories h_i as follows (cf. Section 3):

- $S(h_i) = \{s \in S : h_i \preceq \bar{\mathrm{h}}_i(\zeta(s))\}$;
- $S_i(h_i) = \{s_i \in S_i : \exists s_{-i} \in S_{-i}, h_i \preceq \bar{\mathrm{h}}_i(\zeta(s_i, s_{-i}))\} = \mathrm{proj}_{S_i} S(h_i)$;
- $S_{-i}(h_i) = \{s_{-i} \in S_{-i} : \exists s_i \in S_i, h_i \preceq \bar{\mathrm{h}}_i(\zeta(s_i, s_{-i}))\} = \mathrm{proj}_{S_{-i}} S(h_i)$.

Remark 7. *For all $h_i \in H_i$ and $i \in I$, $S(h_i) = S_i(h_i) \times S_{-i}(h_i)$. For all $g_i, h_i \in \bar{H}_i$,*

$$g_i \preceq h_i \quad \Rightarrow \quad S(h_i) \subseteq S(g_i) \quad \Leftrightarrow \quad S_i(h_i) \subseteq S_i(g_i) \wedge S_{-i}(h_i) \subseteq S_{-i}(g_i).$$

By inspection of the definition observe that, for every $h_i \in \bar{H}_i$, $S(h_i) = \cup_{h \in P_i(h_i)} S(h)$. Thus, by the above remark, $S_{-i}(h_i) = \cup_{h \in P_i(h_i)} S_{-i}(h)$.

Define the objects for the infinite repetition in an analogous way as just done for the one-period game. Then the following remarks also hold.

Remark 8. *$(\mathbf{H}, \preceq)$ is an objective tree, $(\mathbf{H}_i, \preceq)$ is a subjective tree, where $\preceq$ is the "prefix of" relation, inherited, respectively, from $A^{\mathbb{N}}$ and $(A_i \times M_i)^{\mathbb{N}}$.*

Remark 9. *$\mathbf{S}(\mathbf{h}_i) = \mathbf{S}_i(\mathbf{h}_i) \times \mathbf{S}_{-i}(\mathbf{h}_i)$. Moreover, for every $\mathbf{g}_i, \mathbf{h}_i \in \bar{\mathbf{H}}_i$,*

$$\mathbf{g}_i \preceq \mathbf{h}_i \Rightarrow \mathbf{S}(\mathbf{h}_i) \subseteq \mathbf{S}(\mathbf{g}_i) \Leftrightarrow \mathbf{S}_i(\mathbf{h}_i) \subseteq \mathbf{S}_i(\mathbf{g}_i) \wedge \mathbf{S}_{-i}(\mathbf{h}_i) \subseteq \mathbf{S}_{-i}(\mathbf{g}_i).$$

Remark 10. *$\mathbf{Z} = Z^{\mathbb{N}}$, $\mathbf{Z}_i = Z_i^{\mathbb{N}}$, $\mathbf{H} = \cup_{n \geq 0} \left(Z^{[n]} \times H\right)$ and $\mathbf{H}_i = \cup_{n \geq 0} \left(Z_i^{[n]} \times H_i\right)$.*

The last two equalities follow from the observation that a finite history of the repeated game is the concatenation between a (possibly empty) finite sequence of terminal histories of the one-period game with a (possibly empty) non-terminal history of the one-period game, and similarly for personal histories.

8.2 Proofs

We first state two preliminary and quite standard continuity results.[33]

Lemma 1. *The infinite repetition $\Upsilon(\Gamma, \delta)$ of the multistage game Γ with discount factors $\delta = (\delta_i)_{i\in I}$ satisfies continuity at infinity for continuation values, i.e.,*

$$\forall i \in I, \forall \mathbf{h}_i \in \mathbf{H}_i, \quad lim_{t\to\infty}[\sup\{|V^{\mu^i}_{i,\mathbf{h}_i}(\mathbf{s}_i) - V^{\mu^i}_{i,\mathbf{h}_i}(\bar{\mathbf{s}}_i)| \ : \ \mu^i \in \Delta^{\mathfrak{C}_i}(\mathbf{S}_{-i}), \ \mathbf{s}_i, \bar{\mathbf{s}}_i \in \mathbf{S}_i,$$

$$\forall \mathbf{g}_i \in \mathbf{H}_i, \ell(\mathbf{g}_i) < t, \ \mathbf{s}_i(\mathbf{g}_i) = \bar{\mathbf{s}}_i(\mathbf{g}_i)\}] = 0.$$

Lemma 2. *For all $i \in I$ and $\mathbf{h}_i \in \mathbf{H}_i$, $V_{i,\mathbf{h}_i} : \mathbf{S}_i \times \Delta^{\mathfrak{C}_i}(\mathbf{S}_{-i}) \to \mathbb{R}$ is jointly continuous.*

Of course, an analogous result holds for the one-period game.

8.2.1 Proofs for Section 4

Proof of Remark 1: Let

$$\mu^i_t(\cdot|S_{-i}(\text{proj}_{H_i}\mathbf{h}_i)) := marg_{\prod_{\mathbf{h}_i\in\left\{\bar{\mathbf{h}}_i(z^{[t-1]})\right\}\times H_i}\mathcal{A}^{\ell(\mathbf{h}_i)+1}_{-i}(\mathbf{h}_i)}\mu^i(\cdot|\mathbf{S}_{-i}(\mathbf{h}_i))$$

for every $\mathbf{h}_i \in \left\{\bar{\mathbf{h}}_i(z^{[t-1]})\right\} \times H_i$. Then, $\left\{\bar{\mathbf{h}}_i(z^{[t-1]})\right\} \times H_i \cong H_i$. We want to show that the map $h_i \mapsto \mu^i_t(\cdot|S_{-i}(h_i))$ satisfies the chain rule. Note that

$$\mu^i_t\left(S_{-i}(h_i)|S_{-i}(h_i)\right) = \mu^i\left(\mathbf{S}_{-i}\left((\bar{\mathbf{h}}_i(z^{[t-1]}), h_i)\right)|\mathbf{S}_{-i}\left((\bar{\mathbf{h}}_i(z^{[t-1]}), h_i)\right)\right) = 1$$

for every $h_i \in H_i$. For any $E_{-i} \subseteq S_{-i}$, let

$$\mathbf{S}^{E_{-i}}_{-i}\left(\bar{\mathbf{h}}_i(z^{[t-1]})\right) = \left\{\mathbf{s}_{-i} \in \mathbf{S}_{-i}\left(\bar{\mathbf{h}}_i(z^{[t-1]})\right) : \mathbf{s}_{-i}|_{\left\{\bar{\mathbf{h}}_i(z^{[t-1]})\right\}\times H_i} \in E_{-i}\right\}.$$

denote the set of co-players' superstrategy profiles inducing one-period strategy profiles in E_{-i} in the t-th period after any history that i cannot distinguish from $z^{[t-1]}$. With this, for all E_{-i} and $g_i, h_i \in H_i$ such that $h_i \preceq g_i$,

$$\mu^i_t\left(E_{-i} \cap S_{-i}(g_i)|S_{-i}(h_i)\right) =$$

[33] The proofs are available upon request.

$$\mu^{i}\left(\mathbf{S}_{-i}^{E_{-i}}\left(\bar{\mathbf{h}}_{i}(z^{[t-1]})\right)\cap\mathbf{S}_{-i}\left((\bar{\mathbf{h}}_{i}(z^{[t-1]}),g_{i})\right)|\mathbf{S}_{-i}\left((\bar{\mathbf{h}}_{i}(z^{[t-1]}),h_{i})\right)\right)=$$

$$\mu^{i}\left(\mathbf{S}_{-i}^{E_{-i}}\left(\bar{\mathbf{h}}_{i}(z^{[t-1]})\right)|\mathbf{S}_{-i}\left((\bar{\mathbf{h}}_{i}(z^{[t-1]}),g_{i})\right)\right)\cdot$$

$$\mu^{i}\left(\mathbf{S}_{-i}\left((\bar{\mathbf{h}}_{i}(z^{[t-1]}),g_{i})\right)|\mathbf{S}_{-i}\left((\bar{\mathbf{h}}_{i}(z^{[t-1]}),h_{i})\right)\right)=$$

$$=\mu_{t}^{i}\left(E_{-i}|S_{-i}(g_{i})\right)\mu_{t}^{i}\left(S_{-i}(g_{i})|S_{-i}(h_{i})\right).$$

■

Proof of Proposition 1: Take a time t and an objective history $z^{[t-1]}$. Take any personal history $\mathbf{g}_i=\left(\bar{\mathbf{h}}_i(z^{[t-1]}),g_i\right)\in\left\{\bar{\mathbf{h}}_i(z^{[t-1]})\right\}\times H_i$, i.e., any personal history obtained by concatenating the $(t-1)$-period personal history induced by $z^{[t-1]}$ with one-period personal history g_i. We look at the restriction of superstrategies to particular subsets of $\mathbf{H}_i$, using the standard notation for the restriction of functions: $\mathbf{s}_i|\mathbf{F}_i$, for any $\mathbf{F}_i\subseteq\mathbf{H}_i$. We write $\mathbf{s}_i|\mathbf{h}_i$ to denote the superstrategy allowing $\mathbf{h}_i$ and playing like $\mathbf{s}_i$ at each personal history that does not strictly precede $\mathbf{h}_i$. For fixed $\mathbf{s}_i$, $\mathbf{s}_{-i}$, and $z^{[t-1]}$, let $s_{i,t}(\mathbf{s}_i)$ denote the (one-period) strategy played by $\mathbf{s}_i$ in period t following history $z^{[t-1]}$, and similarly define $s_{-i,t}(\mathbf{s}_{-i})$ the profile of strategies played by $\mathbf{s}_{-i}$. Observe that, for every $\mathbf{s}_i\in\mathbf{S}_i$, since players are impatient, we have

$$V_{i,\mathbf{g}_i}^{\mu^i}(\mathbf{s}_i)=\int_{\mathbf{S}_{-i}(\mathbf{g}_i)}u_i\left(\mathrm{z}_t\left(\boldsymbol{\zeta}(\mathbf{s}_i|\mathbf{g}_i,\mathbf{s}_{-i})\right)\right)\mu^i(\mathrm{d}\mathbf{s}_{-i}|\mathbf{S}_{-i}(\mathbf{g}_i))=$$

$$\int_{\mathbf{S}_{-i}(\mathbf{g}_i)}u_i\left(\zeta\left(s_{i,t}(\mathbf{s}_i|\mathbf{g}_i),\ s_{-i,t}(\mathbf{s}_{-i})\right)\right)\mu^i(\mathrm{d}\mathbf{s}_{-i}|\mathbf{S}_{-i}(\mathbf{g}_i)),$$

where, again, $\mathrm{z}_t(\mathbf{z})$ denotes the one-period terminal history played by $\mathbf{z}$ in period t, that is, the t-th coordinate of the sequence $\mathbf{z}\in\mathbf{Z}=Z^{\mathbb{N}}$. Moreover, notice that

$$s_{i,t}(\mathbf{s}_i|\mathbf{g}_i)=s_{-i,t}(\mathbf{s}_i)|g_i.$$

Let

$$\mathbf{S}_{-i}^{s_{-i}}(\mathbf{g}_i)=\left\{\mathbf{s}_{-i}\in\mathbf{S}_{-i}(\mathbf{g}_i):\ s_{-i,t}(\mathbf{s}_{-i})=s_{-i}\right\}.$$

Since players are impatient, their continuation value at the beginning of period t is fully determined by the expected payoff from the subsequent one-period game.

Recall from the proof of Remark 1 the definition of the induced one-period CPS μ_t^i. Then one can see that

$$V_{i,\mathbf{g}_i}^{\mu^i}(\mathbf{s}_i) = \sum_{s_{-i}\in S_{-i}(g_i)} u_i\left(\zeta\left(s_{i,t}(\mathbf{s}_i)|g_i,\ s_{-i}\right)\right)\mu_t^i(s_{-i}|g_i)$$

$$= V_{i,g_i}^{\mu_t^i(\cdot|g_i)}\left(s_{i,t}(\mathbf{s}_i)|g_i\right).$$

Since all $z^{[t-1]} \in Z^{[t-1]}$ and $g_i \in H_i$ yield the repeated-game personal history $\mathbf{g}_i = (\bar{\mathbf{h}}_i(z^{[t-1]}), g_i)$, and for every $\mathbf{g}_i \in \mathbf{H}_i$ there are $g_i \in H_i$ and $z^{[t-1]} \in Z^{[t-1]}$ such that $\mathbf{g}_i = (\bar{\mathbf{h}}_i(z^{[t-1]}), g_i)$, we conclude that, for all $g_i \in H_i$ and $a_i \in \mathcal{A}_i^{\ell(\mathbf{g}_i)+1}(\mathbf{g}_i) = \mathcal{A}_i^{\ell(g_i)+1}(g_i)$,

$$V_{i,g_i}^{\mu_t^i(\cdot|g_i)}(\bar{s}_i^t) = V_{i,\mathbf{g}_i}^{\mu^i}(\bar{\mathbf{s}}_i) \geq V_{i,\mathbf{g}_i}^{\mu^i}(\bar{\mathbf{s}}_i|_{\mathbf{g}_i}a_i) = V_{i,g_i}^{\mu_t^i(\cdot|g_i)}(\bar{s}_i^t|_{g_i}a_i),$$

and that, for all $\mathbf{g}_i \in \mathbf{H}_i$ and $a_i \in \mathcal{A}_i^{\ell(\mathbf{g}_i)+1}(\mathbf{g}_i) = \mathcal{A}_i^{\ell(g_i)+1}(g_i)$,

$$V_{i,\mathbf{g}_i}^{\mu^i}(\bar{\mathbf{s}}_i) = V_{i,g_i}^{\mu_t^i(\cdot|g_i)}(\bar{s}_i^t) \geq V_{i,g_i}^{\mu_t^i(\cdot|g_i)}(\bar{s}_i^t|_{g_i}a_i) = V_{i,\mathbf{g}_i}^{\mu^i}(\bar{\mathbf{s}}_i|_{\mathbf{g}_i}a_i).$$

■

Proof of Theorem 1: Fix $((\mathbf{s}_i, \mu^i))_{i\in I} \in \prod_{i\in I} \mathbf{\Sigma}_i^\infty$. We want to show that, for all $t \in \mathbb{N}$, in the one-period game starting after $z^{[t-1]} = (\mathrm{z}_1(\boldsymbol{\zeta}(\mathbf{s})), ..., \mathrm{z}_{t-1}(\boldsymbol{\zeta}(\mathbf{s})))$, the induced strategy $s_{i,t}(\mathbf{s}_i)$ and the induced one-period CPS μ_t^i are strongly rationalizable, i.e., $(s_{i,t}(\mathbf{s}_i), \mu_t^i) \in \Sigma_i^\infty$.

We are going to prove by induction the following claim.

Claim 1. *For all $k \in \mathbb{N}$, if $((\mathbf{s}_i, \mu^i))_{i\in I} \in \prod_{i\in I} \mathbf{\Sigma}_i^k$ then, for all $t \in \mathbb{N}$, in the one-period game starting at $z^{[t-1]}$,*

$$S_{-i}(h_i) \cap \mathrm{proj}_{S_{-i}}\Sigma_{-i}^m \neq \emptyset \Rightarrow \mu_t^i(\mathrm{proj}_{S_{-i}}\Sigma_{-i}^m|S_{-i}(h_i)) = 1,$$

for all $m < k$ and $h_i \in H_i$.

The claim implies that, for every k, $\mu_t^i \in \mathrm{proj}_{\Delta^{\mathcal{C}_i}(S_{-i})}\Sigma_i^k$. Since $s_{i,t}(\mathbf{s}_i) \in \mathcal{BRO}_i(\mu_t^i)$ (by Proposition 1), we have $(s_{i,t}(\mathbf{s}_i), \mu_t^i) \in \Sigma_i^k$. To prove the claim, we prove that, for all $k \in \mathbb{N}$, $i \in I$, $t \in \mathbb{N}$, $h_i \in H_i$, and $s'_{-i} \in S_{-i}(h_i) \cap \mathrm{proj}_{S_{-i}}\Sigma_{-i}^{k-1}$

(provided this last intersection is not empty), there exists $\mathbf{s}'_{-i} \in \mathrm{proj}_{\mathbf{S}_{-i}} \mathbf{\Sigma}^{k-1}_{-i} \cap \mathbf{S}_{-i}(\bar{\mathbf{h}}_i(z^{[t-1]}))$ such that $s_{-i,t}(\mathbf{s}'_{-i}) = s'_{-i}$.

Proof: *Basis step:* We start with $k = 2$. By Proposition 1, for all $i \in I$ and $t \in \mathbb{N}$, if $(\mathbf{s}_i, \mu^i) \in \mathbf{\Sigma}^1_i$ then $s_{i,t}(\mathbf{s}_i) \in \mathcal{BRO}_i(\mu^i_t)$, which implies that $s_{i,t}(\mathbf{s}_i) \in \Sigma^1_i$. Now suppose that $((\mathbf{s}_i, \mu^i))_{i \in I} \in \prod_{i \in I} \mathbf{\Sigma}^2_i$. Take all $h_i \in H_i$ such that $S_{-i}(h_i) \cap \mathrm{proj}_{S_{-i}} \Sigma^1_{-i} \neq \emptyset$. We want to show that, for any profile of strategies $s'_{-i} = (s'_j)_{j \neq i} \in S_{-i}(h_i) \cap \mathrm{proj}_{S_i} \Sigma^1_{-i}$, there exists a profile of superstrategies $\mathbf{s}'_{-i} \in \mathrm{proj}_{\mathbf{S}_{-i}} \mathbf{\Sigma}_{-i}$ such that $s_{-i,t}(\mathbf{s}'_{-i}) = s'_{-i}$. To see this, consider the case $t = 1$. For each $j \neq i$, there is some $\gamma^j \in \Delta^{\mathcal{C}_j}(S_{-j})$ such that $s'_j \in \mathcal{BRO}_j(\gamma^j)$. Then, we can find some $\nu^i \in \Delta^{\mathfrak{C}_i}(\mathbf{S}_{-i})$ such that, for all $h_i \in H_i$ and $E_{-i} \subseteq S_{-i}$,

$$\nu^i(\mathbf{S}^{E_{-i}}_{-i} | \mathbf{S}_{-i}(h_i)) = \gamma^i(E_{-i} | S_{-i}(h_i)),$$

where $\mathbf{S}^{E_{-i}}_{-i} := \{\mathbf{s}_{-i} \in \mathbf{S}_{-i} : \ s_{-i,t}(\mathbf{s}_{-i}) \in E_{-i}\}$. Indeed, consider $\nu^j \in \Delta^{\mathfrak{C}_j}(\mathbf{S}_{-j})$ such that, for all $h_j \in H_j$ and $s_{-j} \in \mathrm{supp}\gamma^j(\cdot | S_{-j}(h_j)) \setminus \cup_{g_j \prec h_j} \mathrm{supp}\mu^j(\cdot | g_j)$,

$$\nu^j(s^{\mathbb{N}}_{-j} | \mathbf{S}_{-j}(h_j)) = \gamma^j(s_{-j} | S_{-j}(h_j)),$$

where $s^{\mathbb{N}}_{-j}$ is the superstrategy playing like s_{-j} in every period. Indeed, the above condition is consistent with ν^j being a CPS. By Proposition 1, there exists $\mathbf{s}'_j \in \mathcal{BR}_j(\nu^j)$ such that $s^1_j(\mathbf{s}'_j) = s_j$, and hence $\mathbf{s}'_{-i} \in \mathrm{proj}_{\mathbf{S}_{-i}} \mathbf{\Sigma}^1_{-i} \cap \mathbf{S}_{-i}(h_i)$.

Suppose now that $t \geq 2$. There exists $\mathbf{s}'_{-i} \in \mathbf{S}_{-i}(\bar{\mathbf{h}}_i(z^{[t-1]}))$ such that $\mathbf{s}'_{-i} \in \mathrm{proj}_{\mathbf{S}_{-i}} \mathbf{\Sigma}^1_{-i}$. For every $h_i \in H_i$ for which it is possible, take some $s'_{-i} \in S_{-i}(h_i) \cap \mathrm{proj}_{S_{-i}} \Sigma^1_{-i} \neq \emptyset$. We want to show that there exists $\mathbf{s}'_{-i} \in \mathbf{S}_{-i}(\mathbf{h}_i) \cap \mathrm{proj}_{\mathbf{S}_{-i}} \mathbf{\Sigma}^1_{-i}$, where $\mathbf{h}_j = (\bar{\mathbf{h}}_j(z^{[t-1]}), h_j)$. For every $j \neq i$, take any $\gamma^j \in \Delta^{\mathcal{C}_j}(S_{-j})$ such that $s'_j \in \mathcal{BRO}_j(\gamma^j)$. Then, let $\nu^j \in \Delta^{\mathfrak{C}_j}(\mathbf{S}_{-j})$ be such that:

(i) $\mathbf{s}'_j \in \mathcal{BR}_j(\nu^j)$;

(ii) for all $h_j \in H_j$ and $\bar{s}_{-j} \in \mathrm{supp}\gamma^j(\cdot | S_{-j}(h_j)) \setminus \cup_{g_j \prec h_j} \mathrm{supp}\gamma^j(\cdot | g_j)$,

$$\nu^j \left(\mathbf{S}^{\bar{s}^{\mathbb{N}}_{-j}}_{-j} \left((\bar{\mathbf{h}}_j(z^{[t-1]}), h_j) \right) | \mathbf{S}_{-j} \left((\bar{\mathbf{h}}_j(z^{[t-1]}), h_j) \right) \right) = \gamma^j(\bar{s}_{-j} | S_{-j}(h_j)),$$

where $\mathbf{S}^{\bar{s}^{\mathbb{N}}_{-j}}_{-j} \left((\bar{\mathbf{h}}_j(z^{[t-1]}), h_j) \right)$ is the set of j's co-players' superstrategies that allow $(\bar{\mathbf{h}}_j(z^{[t-1]}), h_j)$ and play like $\bar{s}_{-j}$ in the rest of the period and in every subsequent period on and off path;

(iii) $\mathbf{s}'_{-i} \in \mathbf{S}_{-i}(\bar{\mathbf{h}}_i(z^{[t-1]}))$.

Conditions (i) and (iii) are compatible. Condition (ii), as before, does not contradict the fact that ν^j is a CPS, nor can it prevent the superstrategy from satisfying one-step optimality in the previous periods, as that only depends on the past induced one-period CPSs, not modified by this requirement. Indeed, the second condition only affects conjectures about continuation strategies from $\bar{\mathbf{h}}_j(z^{[t-1]})$ onward. Thus, it allows $s_{j,t}(\mathbf{s}'_j) = s'_j$. Consequently, there exists $\mathbf{s}'_{-i} \in \mathbf{S}_{-i}((\bar{\mathbf{h}}_i(z^{[t-1]}), h_i)) \cap \mathrm{proj}_{\mathbf{S}_{-i}} \mathbf{\Sigma}^1_{-i}$. Hence, for all $i \in I$, $t \in \mathbb{N}$, and $h_i \in H_i$

$$S_{-i}(h_i) \cap \mathrm{proj}_{S_{-i}} \Sigma^1_{-i} \neq \emptyset \Rightarrow \mathbf{S}_{-i}((\bar{\mathbf{h}}_i(z^{[t-1]}), h_i)) \cap \mathrm{proj}_{\mathbf{S}_{-i}} \mathbf{\Sigma}^1_{-i} \neq \emptyset \Rightarrow$$

$$\Rightarrow \mu^i(\mathrm{proj}_{\mathbf{S}_{-i}} \mathbf{\Sigma}^1_{-i} | \mathbf{S}_{-i}((\bar{\mathbf{h}}_i(z^{[t-1]}), h_i))) = 1 \Rightarrow \mu^i_t(\mathrm{proj}_{S_{-i}} \Sigma^1_{-i} | S_{-i}(h_i)) = 1$$

where the last implication follows from the fact that, for all $j \in I$ and $t \in \mathbb{N}$,

$$\mathbf{s}'_j \in \mathrm{proj}_{\mathbf{S}_j} \mathbf{\Sigma}^1_j \Rightarrow s_{j,t}(\mathbf{s}'_j) \in \mathrm{proj}_{S_j} \Sigma^1_j.$$

Hence, $\mu^i_t \in \Sigma^2_i$. Then again,

$$\mathbf{s}_i \in \mathcal{BR}_i(\mu^i) \Rightarrow s_{i,t}(\mathbf{s}_i) \in \mathcal{BRO}_i(\mu^i_t)$$

for every $i \in I$, which implies that $((s_{i,t}(\mathbf{s}_i), \mu^i_t))_{i \in I} \in \Sigma^2$.

Inductive step: Suppose that, for some $k \in \mathbb{N}$, for all $m \geq k$ and $t \in \mathbb{N}$, if $((\mathbf{s}_i, \mu^i))_{i \in I} \in \prod_{i \in I} \mathbf{\Sigma}^m_i$, then $((s_{i,t}(\mathbf{s}_i), \mu^i_t))_{i \in I} \in \prod_{i \in I} \Sigma^m$. Suppose also that there exists $\mathbf{s}'_{-i} \in \mathrm{proj}_{\mathbf{S}_{-i}} \mathbf{\Sigma}^\ell_{-i} \cap \mathbf{S}_{-i}(\bar{\mathbf{h}}_i(z^{t-1}))$ such that $s_{-i,t}(\mathbf{s}'_{-i}) = s'_{-i}$, for all $\ell < k$, $h_i \in H_i$ such that $S_{-i}(h_i) \cap \mathrm{proj}_{S_{-i}} \Sigma^\ell_{-i} \neq \emptyset$, and $s'_{-i} \in S_{-i}(h_i) \cap \mathrm{proj}_{S_{-i}} \Sigma^\ell_{-i}$.

First, we want to show that, given any suitable h_i, for every $s'_{-i} \in S_{-i}(h_i) \cap \mathrm{proj}_{S_{-i}} \Sigma^k_{-i}$ there exists $\mathbf{s}'_{-i} \in \mathbf{S}_{-i}(\bar{\mathbf{h}}_i(z^{[t-1]})) \cap \mathrm{proj}_{\mathbf{S}_{-i}} \mathbf{\Sigma}^k_{-i}$ such that $s_{-i,t}(\mathbf{s}'_{-i}) = s'_{-i}$. Condition $\mathbf{S}_{-i}(\bar{\mathbf{h}}_i(z^{[t-1]})) \cap \mathrm{proj}_{\mathbf{S}_{-i}} \mathbf{\Sigma}^k_{-i} \neq \emptyset$ holds because the LHS contains $\mathbf{s}_{-i}$. Conditions $\mathbf{s}'_{-i} \in \mathbf{S}_{-i}(\bar{\mathbf{h}}_i(z^{[t-1]}))$ and $s_{-i,t}(\mathbf{s}'_{-i}) = s_{-i}$ do not contrast one another. Most importantly, neither are $\mathbf{s}'_{-i} \in \mathrm{proj}_{\mathbf{S}_{-i}} \mathbf{\Sigma}^k_{-i}$ and $s_{-i,t}(\mathbf{s}'_{-i}) = s_{-i}$. Indeed, there exist $\mathbf{s}'_i$ such that $\mathbf{s}' = (\mathbf{s}'_i, \mathbf{s}'_{-i}) \in \mathrm{proj}_{\mathbf{S}} \prod_{i \in I} \mathbf{\Sigma}_i^k \cap \mathbf{S}(\bar{\mathbf{h}}_i(z^{[t-1]}))$, which implies that $s_{-i,t}(\mathbf{s}'_{-i}) \in \mathrm{proj}_{S_{-i}} \Sigma_{-i}^k$. In particular, we want to build $\mathbf{s}'_{-i}$ such that $s_{-i,t}(\mathbf{s}'_{-i}) = s'_{-i}$.

Take any $j \neq i$. Let $\gamma^j \in \Delta^{\mathcal{C}_j}(S_{-j})$ be a one-period CPS justifying $s'_j = \mathrm{proj}_{S_j} s'_{-i}$, that is, γ^j strongly believes $\mathrm{proj}_{S_{-j}} \Sigma_{-j}^{k-1}$ and $s'_j \in \mathcal{BRO}_j(\gamma^j)$. Hence, by the inductive hypothesis, for all h_j such that $S_{-j}(h_j) \cap \mathrm{proj}_{S_{-j}} \Sigma_{-j}^{k-1}$, and $s'_{-j} \in \mathrm{supp}\gamma^j(\cdot|S_{-j}(h_j))$, there exists $\mathbf{s}''_{-j} \in \mathrm{proj}_{\mathbf{S}_{-j}} \mathbf{\Sigma}_{-j}^{k-1} \cap \mathbf{S}_{-j}(\bar{h}_j(z^{[t-1]}))$ such that $s^t_{-j}(\mathbf{s}''_{-j}) = s'_{-j}$. By defining $\nu^j \in \Delta^{\mathfrak{C}_j}(\mathbf{S}_{-j})$ as a CPS that strongly believes $\mathbf{s}''_{-j}$, and such that $\nu^j_t = \gamma^j$, it immediately follows that there exists $\mathbf{s}'_j \in \mathcal{BR}_j(\nu^j)$ such that $\mathbf{s}'_j \in \mathrm{proj}_{\mathbf{S}_j} \mathbf{\Sigma}_j^k \cap \mathbf{S}_j(\bar{h}_j(z^{[t-1]}))$ and $s_{j,t}(\mathbf{s}'_j) = s'_j$. Letting $\mathbf{s}'_j \in \mathbf{S}_j(\bar{h}_j(z^{[t-1]}))$ is possible because $\bar{h}_j(z^{[t-1]})$ is consistent with strong belief of level k in rationality, and thus any superstrategy that is a sequential best reply to a CPS assigning probability one to $\bar{h}_j(z^{[t-1]})$ can allow it without loss of generality, and independently of the subsequent choices (since the player is impatient and the personal history is terminal for period $t-1$).

This holds for every $j \neq i$. Observe that for every $\mathbf{h} \in \mathbf{H}$ it holds that

$$\mathbf{S}_{-i}(\bar{\mathbf{h}}_i(\mathbf{h})) \supseteq \mathbf{S}_{-i}(\mathbf{h}) = \prod_{j \neq i} \mathbf{S}_j(\bar{\mathbf{h}}_j(\mathbf{h})).$$

Hence, we have shown the existence of $\mathbf{s}'_{-i} \in \mathbf{S}_{-i}(\bar{\mathbf{h}}_i(z^{[t-1]})) \cap \mathrm{proj}_{\mathbf{S}_{-i}} \mathbf{\Sigma}_{-i}^k$ such that $s_{-i,t}(\mathbf{s}'_{-i}) = s'_{-i}$, and thus $\mathbf{S}_{-i}\left((\bar{\mathbf{h}}_i(z^{[t-1]}), h_i)\right) \cap \mathrm{proj}_{\mathbf{S}_{-i}} \mathbf{\Sigma}_{-i}^k \neq \emptyset$. Assume now that $((\mathbf{s}_i, \mu^i))_{i \in I} \in \mathbf{\Sigma}^{k+1}$. Then, for all $i \in I$, $t \in \mathbb{N}$, $m \neq k$, and $h_i \in H_i$,

$$S_{-i}(h_i) \cap \mathrm{proj}_{S_i} \Sigma_{-i}^m \neq \emptyset \Rightarrow \mathbf{S}_{-i}((\bar{\mathbf{h}}_i(z^{[t-1]}), h_i)) \cap \mathrm{proj}_{\mathbf{S}_{-i}} \mathbf{\Sigma}_{-i}^m \neq \emptyset \Rightarrow$$

$$\Rightarrow \mu^i(\mathrm{proj}_{\mathbf{S}_{-i}} \mathbf{\Sigma}_{-i}^m | \mathbf{S}_{-i}(\bar{\mathbf{h}}_i(z^{[t-1]}), h_i)) = 1 \Rightarrow \mu^i_t(\mathrm{proj}_{S_{-i}} \Sigma_{-i}^m | S_{-i}(h_i)) = 1,$$

where the last implication follows from the inductive hypothesis. Hence, $\mu^i_t \in \mathrm{proj}_{\Delta^{\mathcal{C}_i}(S_{-i})} \Sigma_i^{k+1}$. Then again,

$$\mathbf{s}_i \in \mathcal{BR}_i(marg_{\mathbf{S}_{-i}} \mu^i) \Rightarrow s_{i,t}(\mathbf{s}_i) \in \mathcal{BRO}_i(\mu^i_t)$$

for every $i \in I$, that is, $((s_{i,t}(\mathbf{s}_i), \mu^i_t))_{i \in I} \in \Sigma^{k+1}$.

If $((\mathbf{s}_i,\mu^i))_{i\in I}\in\mathbf{\Sigma}^\infty$, then, for all $t\in\mathbb{N}$, $i\in I$, $h_i\in H_i$, and $k\in\mathbb{N}$,

$$S_{-i}(h_i)\cap\mathrm{proj}_{S_{-i}}\Sigma^\infty_{-i}\neq\emptyset\Rightarrow\mu^i_t(\mathrm{proj}_{S_{-i}}\Sigma^k_{-i}|S_{-i}(h_i))=1\Rightarrow\mu^i_t(\mathrm{proj}_{S_{-i}}\Sigma^\infty_{-i}|S_{-i}(h_i))=1.$$

Hence, $((s_{i,t}(\mathbf{s}_i),\mu^i_t))_{i\in I}\in\prod_{i\in I}\Sigma^\infty$. ■

8.2.2 Proofs for Section 5

Proof of Remark 3: Observe that $\mu^i(\cdot|\mathbf{S}_{-i}(\mathbf{h}^t_i))=\mu^i(\cdot|\mathbf{S}_{-i}(\mathbf{h}^k_i))$ if and only if, for all $k\geq t\geq T$,

$$\mu^i(\mathbf{S}_{-i}(\mathbf{h}^k_i)|\mathbf{S}_{-i}(\mathbf{h}^t_i))=1.$$

This requires that, for every $t\geq T$,

$$\mu^i(\mathbf{S}_{-i}(\mathbf{h}^t_i)|\mathbf{S}_{-i}(\mathbf{h}^T_i))=1,$$

which happens if and only if

$$\mu^i(\cap_{t\geq T}\mathbf{S}_{-i}(\mathbf{h}^t_i)|\mathbf{S}_{-i}(\mathbf{h}^T_i))=1,$$

i.e.,

$$\mu^i(\mathbf{S}_{-i}(\bar{\mathbf{h}}_i(\boldsymbol{\zeta}(\mathbf{s}_i,\mathbf{s}_{-i})))|\mathbf{S}_{-i}(\mathbf{h}^T_i))=1.$$

If this holds for T, it must also hold for every $t\geq T$. In words, if the conjectures of player i over co-players superstrategies has converged starting from T, then at every $t\geq T$ player i believes with certainty in co-players' superstrategy profiles that are observationally equivalent, given i's own superstrategy, to the true ones. Of course, also the reverse implications hold. ■

Proof of Proposition 2: Observe that the sequence $(\mathbf{S}_{-i}(\mathbf{h}^t_i))_{t\in\mathbb{N}}$ is decreasing, and such that $\mathbf{S}_{-i}(\mathbf{h}^t_i)\downarrow\mathbf{S}_{-i}(\bar{\mathbf{h}}_i(\boldsymbol{\zeta}(\mathbf{s})))=\cap_{t\in\mathbb{N}}\mathbf{S}_{-i}(\mathbf{h}^t_i)$. Hence, by continuity of measures, for every $k\in\mathbb{N}$,

$$lim_{t\to\infty}\mu^i(\mathbf{S}_{-i}(\mathbf{h}^t_i)|\mathbf{S}_{-i}(\mathbf{h}^k_i))=\mu^i(\mathbf{S}_{-i}(\bar{\mathbf{h}}_i(\boldsymbol{\zeta}(\mathbf{s}_i,\mathbf{s}_{-i})))|\mathbf{S}_{-i}(\mathbf{h}^k_i)).$$

If there exists some $T\in\mathbb{N}$ such that $\mu^i(\mathbf{S}_{-i}(\bar{\mathbf{h}}_i(\boldsymbol{\zeta}(\mathbf{s}_i,\mathbf{s}_{-i})))|\mathbf{S}_{-i}(\mathbf{h}^T_i))>0$, then, for all $\ell \geq t \geq T$,

$$\mu^i(\mathbf{S}_{-i}(\bar{\mathbf{h}}_i(\boldsymbol{\zeta}(\mathbf{s}_i, \mathbf{s}_{-i})))|\mathbf{S}_{-i}(\mathbf{h}_i^t)) > 0 \quad \wedge \quad \mu^i(\mathbf{S}_{-i}(\mathbf{h}_i^\ell)|\mathbf{S}_{-i}(\mathbf{h}_i^t)) > 0,$$

by the chain rule. Hence, for all $\ell \geq t \geq T$, again applying the chain rule, it holds that

$$\mu^i(\mathbf{S}_{-i}(\bar{\mathbf{h}}_i(\boldsymbol{\zeta}(\mathbf{s}_i, \mathbf{s}_{-i})))|\mathbf{S}_{-i}(\mathbf{h}_i^\ell)) = \frac{\mu^i(\mathbf{S}_{-i}(\bar{\mathbf{h}}_i(\boldsymbol{\zeta}(\mathbf{s}_i, \mathbf{s}_{-i})))|\mathbf{S}_{-i}(\mathbf{h}_i^t))}{\mu^i(\mathbf{S}_{-i}(\mathbf{h}_i^\ell)|\mathbf{S}_{-i}(\mathbf{h}_i^t))}.$$

Taking the limit as $\ell \to \infty$, we obtain

$$lim_{\ell \to \infty} \mu^i(\mathbf{S}_{-i}(\bar{\mathbf{h}}_i(\boldsymbol{\zeta}(\mathbf{s}_i, \mathbf{s}_{-i})))|\mathbf{S}_{-i}(\mathbf{h}_i^\ell)) = \frac{\mu^i(\mathbf{S}_{-i}(\bar{\mathbf{h}}_i(\boldsymbol{\zeta}(\mathbf{s}_i, \mathbf{s}_{-i})))|\mathbf{S}_{-i}(\mathbf{h}_i^t))}{\mu^i(\mathbf{S}_{-i}(\bar{\mathbf{h}}_i(\boldsymbol{\zeta}(\mathbf{s}_i, \mathbf{s}_{-i})))|\mathbf{S}_{-i}(\mathbf{h}_i^t))} = 1,$$

proving the first claim. The second claim follows immediately from the definition of limit. ■

8.2.3 Proofs for Section 6

Proof of Remark 4: For every $t \in \mathbb{N}$,

$$\mu_t^i(S_{-i}\left(\bar{\mathrm{h}}_i(\zeta(s^t))\right)|S_{-i}) = \mu^i(\mathbf{S}_{-i}\left(\left(\mathbf{h}_i^t, \bar{\mathrm{h}}_i(\zeta(s^t))\right)\right)|\mathbf{S}_{-i}(\mathbf{h}_i^t)) \geq$$

$$\geq \mu^i(\mathbf{S}_{-i}\left(\bar{\mathbf{h}}_i(\boldsymbol{\zeta}(\mathbf{s}))\right)|\mathbf{S}_{-i}(\mathbf{h}_i^t)),$$

where the first inequality is by definition of probability measure and the second by the fact that $\left(\mathbf{h}_i^t, \bar{\mathrm{h}}_i(\zeta(s^t))\right) \preceq \bar{\mathbf{h}}_i(\boldsymbol{\zeta}(\mathbf{s}))$. Hence, for all $t \in \mathbb{N}$ and $\varepsilon \geq 0$,

$$\mu^i(\mathbf{S}_{-i}\left(\bar{\mathbf{h}}_i(\boldsymbol{\zeta}(\mathbf{s}))\right)|\mathbf{S}_{-i}(\mathbf{h}_i^t)) \geq 1-\varepsilon \Rightarrow \mu_t^i(S_{-i}\left(\bar{\mathrm{h}}_i(\zeta(s^t))\right)|S_{-i}) \geq 1-\varepsilon.$$

■

Proof of Remark 5: We have seen in Section 5 that conjectures converging in finite time is equivalent to the existence of a time T starting from which

$$\forall i, \quad \mu^i(\mathbf{S}_{-i}(\bar{\mathbf{h}}_i(\boldsymbol{\zeta}(\mathbf{s})))|\mathbf{S}_{-i}(\mathbf{h}_i^T)) = 1,$$

and thus condition (i') of Definition 9 is satisfied. Then strong rationalizability takes care of the rest. ■

Proof of Proposition 3: We have shown in the proof of Theorem 1 that, when players are impatient, for all $k \in \mathbb{N}$, $t \in \mathbb{N}$, $i \in I$, $z^{[t]} \in Z^{[t]}$ such that $\mathbf{S}(z^{[t]}) \cap \mathrm{proj}_{\mathbf{S}}\mathbf{\Sigma}^k \neq \emptyset$, and $s'_i \in \mathrm{proj}_{S_i}\Sigma^k$, there exists $\mathbf{s}'_i \in \mathrm{proj}_{\mathbf{S}_i}\mathbf{\Sigma}^k_i \cap \mathbf{S}_i(z^{[t]})$ such that $s_{i,t}(\mathbf{s}'_i) = s'_i$.

Since $\mathbf{S}(\mathbf{h}) = \prod_{i\in I}\mathbf{S}_i(\mathbf{h})$ for every $\mathbf{h} \in \bar{\mathbf{H}}$, it follows that, for all $t \in \mathbb{N}$, $\mathbf{h} \in Z^{[t]}$ such that $\mathbf{S}(\mathbf{h}) \cap \mathrm{proj}_{\mathbf{S}}\mathbf{\Sigma}^\infty \neq \emptyset$, $z \in Z$ with $\mathrm{proj}_S\Sigma^\infty \cap S(z) \neq \emptyset$, and $k \in \mathbb{N}$, there exists $\mathbf{s}_k \in \mathrm{proj}_{\mathbf{S}}\mathbf{\Sigma}^k \cap \mathbf{S}((\mathbf{h}, z))$, i.e., $\mathrm{proj}_{\mathbf{S}}\mathbf{\Sigma}^k \cap \mathbf{S}((\mathbf{h}, z)) \neq \emptyset$. Since $\mathrm{proj}_{\mathbf{S}}\mathbf{\Sigma}^k$ and $\mathbf{S}((\mathbf{h}, z))$ are closed, $\mathrm{proj}_{\mathbf{S}}\mathbf{\Sigma}^k \cap \mathbf{S}((\mathbf{h}, z))$ is closed, and thus compact. By finite intersection property of compact sets, and because for every $m < k$ we know that $\mathrm{proj}_{\mathbf{S}}\mathbf{\Sigma}^m \cap \mathrm{proj}_{\mathbf{S}}\mathbf{\Sigma}^k = \mathrm{proj}_{\mathbf{S}}\mathbf{\Sigma}^k$, it holds that $\cap_{k\in\mathbb{N}}\left(\mathbf{S}((\mathbf{h}, z)) \cap \mathrm{proj}_{\mathbf{S}}\mathbf{\Sigma}^k\right) = \mathbf{S}((\mathbf{h}, z)) \cap \mathrm{proj}_{\mathbf{S}}\mathbf{\Sigma}^\infty \neq \emptyset$. In other words, there exists $\mathbf{s}' \in \mathbf{S}(\mathbf{h}) \cap \mathrm{proj}_{\mathbf{S}}\mathbf{\Sigma}^\infty$ such that $\zeta(s^t(\mathbf{s}')) = z$.

By the above, for every $z \in Z$ consistent with one-period strong rationalizability, there exists $\mathbf{s}' \in \mathrm{proj}_{\mathbf{S}}\mathbf{\Sigma}^\infty$ such that $\mathrm{z}_1(\boldsymbol{\zeta}(\mathbf{s}')) = z$. Now proceed by induction: as inductive hypothesis, suppose that, for a fixed $t \in \mathbb{N}$, for every $\mathbf{h} \in Z^{[t]}$ with $\mathrm{z}_k(\mathbf{h})$ consistent with one-period strong rationalizability for every $k \leq t$, there exists $\mathbf{s}' \in \mathrm{proj}_{\mathbf{S}}\mathbf{\Sigma}^\infty$ such that $\mathbf{s}' \in \mathbf{S}(\mathbf{h})$. Then, for every $z \in Z$ consistent with one-period strong rationalizability there exists $\mathbf{s}'' \in \mathbf{S}((\mathbf{h}, z)) \cap \mathrm{proj}_{\mathbf{S}}\mathbf{\Sigma}^\infty$.

Hence, if $\mathbf{z} \in Z^\infty$ is such that, for every $t \in \mathbb{N}$, $\mathrm{z}_t(\mathbf{z})$ is consistent with one-period strong rationalizability, then, for every t, $\mathbf{S}((\mathrm{z}_1(\mathbf{z}), \dots, \mathrm{z}_t(\mathbf{z}))) \cap \mathrm{proj}_{\mathbf{S}}\mathbf{\Sigma}^\infty \neq \emptyset$. As above, $\mathbf{S}(z^{[t]}(\mathbf{z})) \cap \mathrm{proj}_{\mathbf{S}}\mathbf{\Sigma}^\infty$ is compact. Moreover, $\mathbf{S}((\mathrm{z}_1(\mathbf{z}), \dots, \mathrm{z}_t(\mathbf{z}))) \cap \mathbf{S}(z^{[\ell]}(\mathbf{z})) = \mathbf{S}((\mathrm{z}_1(\mathbf{z}), \dots, \mathrm{z}_t(\mathbf{z})))$ whenever $\ell \leq t$. Thus, by the finite intersection property of compact sets, $\cap_{t\in\mathbb{N}}\left(\mathbf{S}((\mathrm{z}_1(\mathbf{z}), \dots, \mathrm{z}_t(\mathbf{z}))) \cap \mathrm{proj}_{\mathbf{S}}\mathbf{\Sigma}^\infty\right) = \mathbf{S}(\mathbf{z}) \cap \mathrm{proj}_{\mathbf{S}}\mathbf{\Sigma}^\infty \neq \emptyset$. ■

Proof of Theorem 3: Let $\mathbf{z} \in Z^\infty$ be the terminal history induced by the sequence of one-period SCEs in strongly rationalizable conjectures $(((s^t_i, \gamma^i_t))_{i\in I})_{t\in\mathbb{N}}$. For every $i \in I$, take $\nu^i \in \Delta^{\mathfrak{C}_i}(\mathbf{S}_{-i})$ so that, for all $t \in \mathbb{N}$, $h_i \in H_i$, and $s_{-i} \in S_{-i}$,

$$\nu^i\left(\mathbf{S}^{s_{-i}}_{-i}(\bar{\mathbf{h}}_i((\mathrm{z}_1(\mathbf{z}), \dots, \mathrm{z}_{t-1}(\mathbf{z})))) \cap \mathrm{proj}_{\mathbf{S}_{-i}}\mathbf{\Sigma}^\ell_{-i}|\mathbf{S}_{-i}((\bar{\mathbf{h}}_i((\mathrm{z}_1(\mathbf{z}), \dots, \mathrm{z}_{t-1}(\mathbf{z}))), h_i))\right)$$

$$= \gamma^i_t(s_{-i}|S_{-i}(h_i)),$$

where once again

$$\mathbf{S}_{-i}^{s_{-i}}(\bar{\mathbf{h}}_i((\mathrm{z}_1(\mathbf{z}),\ldots,\mathrm{z}_{t-1}(\mathbf{z})))=\{\mathbf{s}_{-i}\in\mathbf{S}_{-i}(\bar{\mathbf{h}}_i((\mathrm{z}_1(\mathbf{z}),\ldots,\mathrm{z}_{t-1}(\mathbf{z})))\ :\ \ s_{-i,t}(\mathbf{s})=s_{-i}\}.$$

Since γ_t^i strongly believes $(\mathrm{proj}_{S_{-i}}\Sigma^k)_{k=1}^{\infty}$, then, by definition and by Proposition 3, there exists $\nu^i\left(\cdot|\mathbf{S}_{-i}((\bar{\mathbf{h}}_i((\mathrm{z}_1(\mathbf{z}),\ldots,\mathrm{z}_{t-1}(\mathbf{z}))),h_i))\right)$ that strongly believes $(\mathrm{proj}_{S_{-i}}\mathbf{\Sigma}^k)_{k=1}^{\infty}$. Observe that, by definition, ν^i assigns initial probability 1 to the collection of sets $(\mathbf{S}_{-i}(\bar{\mathbf{h}}_i(\mathrm{z}_1(\mathbf{z}),\ldots,\mathrm{z}_{t-1}(\mathbf{z}))))_{t\in\mathbb{N}}$, and thus is confirmed on path. Furthermore, let ν^i strongly believe in $(\mathrm{proj}_{\mathbf{S}_{-i}}\mathbf{\Sigma}_{-i}^k)_{k=0}^{\infty}$ (which consists in imposing constraints on ν^i at personal histories outside $\cup_{t\in\mathbb{N}}\left(\{\bar{\mathbf{h}}_i(\mathrm{z}_1(\mathbf{z}),\ldots,\mathrm{z}_{t-1}(\mathbf{z}))\}\times H_i\right)$). Then, there exists $\mathbf{s}_i'\in\mathbf{S}_i$ such that $\mathbf{s}_i\in\mathcal{BR}_i(\nu^i)$, and $s_{i,t}(\mathbf{s}_i')=s_i^t$. Therefore, $((\mathbf{s}_i',\nu^i))_{i\in I}\in\mathbf{\Sigma}^{\infty}$ is such that $\boldsymbol{\zeta}(\mathbf{s}')=\mathbf{z}$, and ν^i has converged for every player. ■

Proof of Theorem 4:

Claim 2. *There exists $\bar{\varepsilon}>0$ such that, for all $i\in I$, $\gamma^i\in\Delta^{\mathcal{C}_i}(S_{-i})$, $s_i\in\mathcal{BRO}_i(\gamma^i)$, and $z_i\in Z_i$, if*

$$\gamma^i(S_{-i}(z_i)|S_{-i})\geq 1-\bar{\varepsilon}$$

then $s_i\in\mathcal{BRO}_i(\nu^i)$, where ν^i is the fully confirmed modification on γ^i.

Proof: Let $s_i'\in\mathcal{BRO}_i(\nu^i)$. Let

$$m:=\min\{u_i(z)-u_i(z')>0:\ i\in I,\ z,z'\in Z\},$$

where $m>0$ by finiteness of Z. Let

$$M:=\max\{u_i(z)-u_i(z')>0:\ i\in I,\ z,z'\in Z\},$$

where $M<\infty$, again, by finiteness of Z. Notice that there must exist $h_i\in H_i$ such that $V_{i,h_i}^{\nu^i}(s_i')-V_{i,h_i}^{\nu^i}(s_i)>0$ (and thus in particular $\geq m$), otherwise $s_i\in\mathcal{BRO}_i(\nu^i)$ and we are done. Notice that

$$V_{i,h_i}^{\gamma^i}(s_i)-V_{i,h_i}^{\gamma^i}(s_i')\leq\gamma^i(S_{-i}(z_i)|S_{-i})\cdot(-m)+\left(1-\gamma^i(S_{-i}(z_i)|S_{-i})\right)\cdot M.$$

Then $\bar{\varepsilon}=\frac{m}{m+M}$ does the job. □

When a one-period ε-SCE in strongly rationalizable conjectures is played with

$\varepsilon \le \bar{\varepsilon}$ above, then for all i $s_i \in \mathcal{BRO}_i(\nu^i)$, where ν^i is a confirmed conjectured (given s_{-i}). Since $(s_i, \gamma^i) \in \Sigma_i^\infty$ and supp$\nu^i \subseteq$ suppγ^i, then $(s_i, \nu^i) \in \Sigma_i^\infty$. This concludes the proof. ■

Proof of Theorem 5: The first part needs no further proof. The auxiliary result follows from the following claim.

Claim 3. *Fix $i \in I$, $\mu^i \in \Delta^{\mathfrak{C}_i}(\mathbf{S}_{-i})$, $\mathbf{s}_i \in \mathcal{BR}_i(\mu^i)$, and $\mathbf{z}_i \in \mathbf{Z}_i$. If for every $\bar{\varepsilon} > 0$ there exists $\mathbf{h}_i^{\bar{\varepsilon}} \prec \mathbf{z}_i$ such that*

$$\mu^i(\mathbf{S}_{-i}(\mathbf{z}_i)|\mathbf{S}_{-i}(\mathbf{h}_i^{\bar{\varepsilon}})) \ge 1 - \bar{\varepsilon},$$

then, for every $\varepsilon' > 0$, there exists $\mathbf{h}_i^{\varepsilon}$ and $\nu^i \in \Delta^{\mathfrak{C}_i^{\mathbf{h}_i^{\varepsilon}}}(\mathbf{S}_{-i})$ such that

$$\nu^i(\mathbf{S}_{-i}(\mathbf{z}_i)|\mathbf{S}_{-i}(\mathbf{h}_i)) = 1 \quad \forall \mathbf{h}_i^{\varepsilon} \preceq \mathbf{h}_i \preceq \mathbf{z}_i,$$

and $\mathrm{proj}_{\mathbf{S}_i^{\succeq \mathbf{h}_i^{\varepsilon}}} \mathbf{s}_i \in \mathcal{BR}_i^{\varepsilon'}(\nu^i)$.

Proof: Fix $i \in I$. Let $z_i \in Z_i$ and $\gamma^i \in \Delta^{\mathcal{C}_i}(S_{-i})$. Assume that, for some $\varepsilon > 0$ and for every $h_i \preceq z_i$, there exists ε_{h_i} such that

$$\gamma^i\left(S_{-i}(z_i)|S_{-i}(h_i)\right) = 1 - \varepsilon_{h_i} \ge 1 - \varepsilon.$$

Define $\nu_\varepsilon^i \in [\Delta(S_{-i})]^{\mathcal{C}_i}$ in the following way:

$$\forall\ h_i \preceq z_i,\ \forall E_{-i} \subseteq S_{-i} \quad \nu_\varepsilon^i\left(E_{-i}|S_{-i}(h_i)\right) = \frac{\gamma^i\left(E_{-i} \cap S_{-i}(z_i)|S_{-i}(h_i)\right)}{1 - \varepsilon_{h_i}}$$

$$\forall\ h_i \npreceq z_i,\ \forall E_{-i} \subseteq S_{-i}, \quad \nu_\varepsilon^i\left(E_{-i}|S_{-i}(h_i)\right) = \gamma^i\left(E_{-i}|S_{-i}(h_i)\right).$$

It can be checked that ν_ε^i is a CPS, i.e., $\nu_\varepsilon^i \in \Delta^{\mathcal{C}_i}(S_{-i})$. Moreover, for all $h_i \in H_i$, $\nu_{\varepsilon,h_i}^i := \nu_\varepsilon^i(\cdot|S_{-i}(h_i))$ is absolutely continuous with respect to $\gamma_{h_i}^i := \mu^i(\cdot|S_{-i}(h_i))$. The Radon-Nikodym derivative is

$$\frac{\mathrm{d}\nu_{\varepsilon,h_i}^i}{\mathrm{d}\gamma_{h_i}^i} = \frac{1}{1 - \varepsilon_{h_i}} \mathbb{1}_{S_{-i}(z_i)}$$

whenever $h_i \preceq z_i$, and simply 1 otherwise. Take any measurable function u, then

$$\int_{S_{-i}} u \mathrm{d}\nu^i_{\varepsilon,h_i} = \int_{S_{-i}} u \frac{\mathrm{d}\nu^i_{\varepsilon,h_i}}{\mathrm{d}\gamma^i_{h_i}} \mathrm{d}\gamma^i_{h_i}$$

becomes, when $h_i \preceq z_i$,

$$\int_{S_{-i}(z_i)} u \mathrm{d}\gamma^i_{h_i} = (1 - \varepsilon_{h_i}) \int_{S_{-i}} u \mathrm{d}\nu^i_{\varepsilon,h_i}.$$

Hence, for all h_i and $\bar{s}_i \in S_i$,

$$\mathbb{E}_{\gamma^i_{h_i}}[U_i(\bar{s}_i, \cdot)|h_i] = \sum_{s_{-i} \in S_{-i}(h_i)} U_i(\bar{s}_i|h_i, s_{-i}) \cdot \gamma^i_{h_i}(s_{-i})$$

$$= (1-\varepsilon_{h_i}) \sum_{s_{-i} \in S_{-i}(z_i)} U_i(\bar{s}_i|h_i, s_{-i}) \cdot \nu^i_{\varepsilon,h_i}(s_{-i}) + \sum_{s_{-i} \in S_{-i}(h_i) \setminus S_{-i}(z_i)} U_i(\bar{s}_i|h_i, s_{-i}) \cdot \gamma^i_{h_i}(s_{-i})$$

$$= (1 - \varepsilon_{h_i}) \mathbb{E}_{\nu^i_{\varepsilon,h_i}}[U_i(\bar{s}_i, \cdot)|h_i] + \sum_{s_{-i} \in S_{-i}(h_i) \setminus S_{-i}(z_i)} U_i(\bar{s}_i|h_i, s_{-i}) \cdot \gamma^i_{h_i}(s_{-i}).$$

Let

$$n_i = min_{s \in S} U_i(s), \quad N_i = max_{s \in S} U_i(s), \quad \kappa_i = min_{w,v \in U_i(S), w \neq v} |w - v|.$$

Suppose by contradiction that, for every $\varepsilon' > 0$ and some $h'_i \preceq z_i$, there exists $\widehat{s}_i$ such that $\mathbb{E}_{\nu^i_{\varepsilon',h'_i}}[U_{i,\widehat{s}_i}] > \mathbb{E}_{\nu^i_{\varepsilon',h'_i}}[U_{i,s_i}]$. Then

$$0 \geq \mathbb{E}_{\gamma^i_{h'_i}}[U_{i,\widehat{s}_i}] - \mathbb{E}_{\gamma^i_{h'_i}}[U_{i,s_i}] = (1 - \varepsilon_{h_i}) \left[\mathbb{E}_{\nu^i_{\varepsilon,h'_i}}[U_{i,\widehat{s}_i}] - \mathbb{E}_{\nu^i_{\varepsilon',h'_i}}[U_{i,s_i}] \right] +$$

$$+ \sum_{s_{-i} \in S_{-i}(h'_i) \setminus S_{-i}(z_i)} [U_i(\widehat{s}_i|h'_i, s_{-i}) - U_i(s_i|h'_i, s_{-i})] \cdot \gamma^i_{h'_i}(s_{-i})$$

$$\geq (1 - \varepsilon_{h_i})\kappa_i - \varepsilon_{h_i}(M_i - n_i).$$

Thus the inequality is satisfied only if

$$\varepsilon' \geq \varepsilon_{h_i} \geq \frac{\kappa_i}{\kappa_i + M_i - N_i} \in (0, 1).$$

Then, there exists $\bar{\varepsilon} < \frac{\kappa_i}{\kappa_i + M_i - N_i}$ such that a contradiction obtains. Since, for every $\bar{s}_i \in S_i$,

$$\mathbb{E}_{\nu^i_{\varepsilon,h_i}}[U_{i,s_i}] = \mathbb{E}_{\gamma^i_{h_i}}[U_{i,s_i}] \geq \mathbb{E}_{\gamma^i_{h_i}}[U_{i,\bar{s}_i}] = \mathbb{E}_{\nu^i_{\varepsilon,h_i}}[U_{i,\bar{s}_i}]$$

when $h_i \not\preceq z_i$, then the statement is satisfied. ■

8.3 Formalism of Section 6.3

We define formally i's "continuation objects" for the infinite repetition, given a certain personal history $\mathbf{h}_i$. Let

$$\mathfrak{C}_i^{\succeq \mathbf{h}_i} = \{\mathbf{S}_{-i}(\mathbf{g}_i) \subseteq \mathbf{S}_{-i} : \mathbf{g}_i \succeq \mathbf{h}_i\}$$

be the set of conditional events that are induced by the sub-tree of $\bar{\mathbf{H}}_i$ with root $\mathbf{h}_i$, and let $\Delta^{\mathfrak{C}_i^{\succeq \mathbf{h}_i}}(\mathbf{S}_{-i})$ denote the corresponding set of CPSs. After $\mathbf{h}_i$, "optimality from there on" depends only on these CPSs. In other words, optimality starting at $\mathbf{h}_i$ should be intended as optimality in the "subjective continuation game with root $\mathbf{h}_i$". Given a continuation superstrategy $\mathbf{s}_i^{\succeq \mathbf{h}_i} \in \mathbf{S}_i^{\succeq \mathbf{h}_i} := \times_{\mathbf{g}_i \succeq \mathbf{h}_i} \mathcal{A}^{\ell(\mathbf{g}_i)+1}(\mathbf{g}_i)$ and a CPS $\nu^i \in \Delta^{\mathfrak{C}_i^{\succeq \mathbf{h}_i}}(\mathbf{S}_{-i})$, the continuation value for the continuation game is clear: for every $\mathbf{g}_i \succeq \mathbf{h}_i$, for all $\mathbf{s}_i \in \mathbf{S}_i$ and $\mu^i \in \Delta^{\mathfrak{C}_i}(\mathbf{S}_{-i})$ such that $\mathrm{proj}_{\mathbf{S}_i^{\succeq \mathbf{h}_i}} \mathbf{s}_i = \mathbf{s}_i^{\succeq \mathbf{h}_i}$ and $\mathrm{proj}_{[\Delta(\mathbf{S}_{-i})]^{\mathfrak{C}_i^{\mathbf{h}_i}}} \mu^i = \nu^i$,

$$V^{\nu^i}_{i,\mathbf{g}_i}(\mathbf{s}_i^{\succeq \mathbf{h}_i}) = \sum_{t=\tau(\mathbf{g}_i)+1}^{\infty} \delta_i^{t-\tau(\mathbf{g}_i)-1} \int_{\mathbf{S}_{-i}(\mathbf{g}_i)} u_i(\mathrm{z}_t(\boldsymbol{\zeta}(\mathbf{s}_i^{\succeq \mathbf{h}_i} | \mathbf{g}_i, \mathbf{s}_{-i}))) \nu^i(\mathrm{d}\mathbf{s}_{-i} | \mathbf{S}_{-i}(\mathbf{g}_i)) =$$

$$= \sum_{t=\tau(\mathbf{g}_i)+1}^{\infty} \delta_i^{t-\tau(\mathbf{g}_i)-1} \int_{\mathbf{S}_{-i}(\mathbf{g}_i)} u_i(\mathrm{z}_t(\boldsymbol{\zeta}(\mathbf{s}_i | \mathbf{g}_i, \mathbf{s}_{-i}))) \mu^i(\mathrm{d}\mathbf{s}_{-i} | \mathbf{S}_{-i}(\mathbf{g}_i)) = V^{\mu^i}_{i,\mathbf{g}_i}(\mathbf{s}_i),$$

where $\boldsymbol{\zeta}(\mathbf{s}_i^{\succeq \mathbf{h}_i} | \mathbf{g}_i, \mathbf{s}_{-i})$ is the terminal history induced by playing continuation superstrategy $\mathbf{s}_i^{\succeq \mathbf{h}_i}$ after $\mathbf{g}_i$ and superstrategies $\mathbf{s}_{-i}$. Let $\mu^i_{\succeq \mathbf{h}_i}$ denote the projection of CPS μ^i over the set of CPSs $\Delta^{\mathfrak{C}_i^{\succeq \mathbf{h}_i}}(\mathbf{S}_{-i})$. Then, $\mathbf{s}_i$ is optimal starting at $\mathbf{h}_i$, with respect to μ^i, if $\mathbf{s}_i^{\succeq \mathbf{h}_i}$ is one-step optimal given $\mu^i_{\succeq \mathbf{h}_i}$, that is, for every $\mathbf{g}_i \succeq \mathbf{h}_i$,

$$\mathbf{s}_i^{\succeq \mathbf{h}_i}(\mathbf{g}_i) \in \arg \max_{a_i \in \mathcal{A}_i^{\ell(\mathbf{g}_i)+1}} V^{\mu^i_{\succeq \mathbf{h}_i}}_{i,\mathbf{g}_i}(\mathbf{s}_i^{\succeq \mathbf{h}_i}|_{\mathbf{g}_i} a_i),$$

where $V^{\mu^i_{\succeq \mathbf{h}_i}}_{i,\mathbf{g}_i}(\mathbf{s}_i^{\succeq \mathbf{h}_i}|_{\mathbf{g}_i} a_i)$ denotes, in the usual way, the continuation value at $\mathbf{g}_i$ of playing continuation like superstrategy $\mathbf{s}_i^{\succeq \mathbf{h}_i}$ after $\mathbf{g}_i$ and a_i. With an abuse of notation, we denote such optimality property as

$$\mathbf{s}_i^{\succeq \mathbf{h}_i} \in \mathcal{BR}_i(\mu^i_{\succeq \mathbf{h}_i}).$$

Remark 11. *If a superstrategy is one-step optimal given a CPS, then the continuation superstrategy is optimal starting at any personal history given the induced continuation CPS.*

Formally, Remark 11 says that, given $\mathbf{s}_i$, $\mathbf{s}_i^{\succeq \mathbf{h}_i}$, μ^i, and $\mu^i_{\succeq \mathbf{h}_i}$ as above, for every $\mathbf{h}_i$,

$$\mathbf{s}_i \in \mathcal{BR}_i(\mu^i) \quad \Rightarrow \quad \mathbf{s}_i^{\succeq \mathbf{h}_i} \in \mathcal{BR}_i(\mu^i_{\succeq \mathbf{h}_i}).$$

8.4 Optimality and existence

In this section we define sequential optimality and weak sequential optimality and compare them with one-step optimality. The use of continuation values allows us to extend optimality conditions "computed under this ex-ante perspective" to the case of impatient intertemporal preferences in a multi-period game. Subsequently, we present a version of the *One-Shot Deviation Principle*, which states the equivalence between one-step optimality and sequential optimality. Known arguments can be adapted to prove the existence of sequentially optimal and weakly sequentially optimal superstrategies (and strategies), and to prove that a superstrategy (strategy) is weakly sequentially optimal if and only if there exists a behaviorally equivalent sequentially optimal superstrategy (strategy). This implies that our definition of strong rationalizability is behaviorally equivalent to the one based on weak sequential optimality used by Pearce (1984) and many following papers for finite games, and by Battigalli & Tebaldi (2019) for infinite-horizon games (see also the discussion in Battigalli & De Vito 2021).

Definition 14. *A superstrategy* $\mathbf{s}_i^*$ *is sequentially optimal given* $\mu^i \in \Delta^{\mathfrak{C}_i}(\mathbf{S}_{-i})$ *if, for every* $\mathbf{h}_i \in \mathbf{H}_i$,

$$\mathbf{s}_i^* \in \arg\max_{\mathbf{s}_i \in \mathbf{S}_i} V^{\mu^i}_{i,\mathbf{h}_i}(\mathbf{s}_i).$$

Similarly, a strategy s_i^* *is sequentially optimal given a one-period CPS* $\gamma^i \in \Delta^{\mathcal{C}_i}(S_{-i})$

if, for every $h_i \in H_i$,

$$s_i^* \in \arg\max_{s_i \in S_i(h_i)} V_{i,h_i}^{\gamma^i}(s_i).$$

Observe that, for all $i \in I$, $\mathbf{h}_i \in \mathbf{H}_i$, and $\mu^i \in \Delta^{\mathfrak{C}_i}(\mathbf{S}_{-i})$, by compactness of $\mathbf{S}_i$ and continuity of $V_{i,\mathbf{h}_i}^{\mu^i}(\cdot)$, $V_{i,\mathbf{h}_i}^{\mu^i}(\cdot)$ admits a maximizer. Of course, the same holds with respect to the one-period continuation values. For any $\mathbf{s}_i \in \mathbf{S}_i$, let

$$\mathbf{H}_i(\mathbf{s}_i) = \left\{\mathbf{h}_i \in \mathbf{H}_i : \exists \mathbf{s}_{-i} \in \mathbf{S}_{-i}, \mathbf{h}_i \prec \bar{\mathbf{h}}_i\left(\zeta\left(\mathbf{s}_i, \mathbf{s}_{-i}\right)\right)\right\}$$

denote the set of personal histories allowed by $\mathbf{s}_i$, that is, those that can occur if $\mathbf{s}_i$ is played. The set $H_i\left(s_i\right)$ of histories of the one-period game allowed by strategy s_i is similarly defined.

Definition 15. *A superstrategy* $\mathbf{s}_i^*$ *is weakly sequentially optimal given* $\mu^i \in \Delta^{\mathfrak{C}_i}(\mathbf{S}_{-i})$ *if, for every* $\mathbf{h}_i \in \mathbf{H}_i(\mathbf{s}_i^*)$,

$$\mathbf{s}_i^* \in \arg\max_{\mathbf{s}_i \in \mathbf{S}_i} V_{i,\mathbf{h}_i}^{\mu^i}(\mathbf{s}_i).$$

Similarly, a strategy s_i^* *is weakly sequentially optimal given a one-period CPS* $\gamma^i \in \Delta^{\mathcal{C}_i}(S_{-i})$ *if, for every* $h_i \in H_i(s_i^*)$,

$$s_i^* \in \arg\max_{s_i \in S_i} V_{i,h_i}^{\gamma^i}(s_i).$$

The following two propositions establish that our definition of rationality is equivalent to sequential optimality, and it is hence behaviorally equivalent to the rationality definition of Battigalli & Tebaldi (2019), provided that players are not impatient.

Proposition 4 (One-Shot Deviation Principle)**.** *Fix a player* i*, a superstrategy* $\mathbf{s}_i$ *and a CPS* μ^i *over co-players' superstrategy profiles. Then,* $\mathbf{s}_i$ *is one-step optimal given* μ^i *if and only if* $\mathbf{s}_i$ *is sequentially optimal given* μ^i.

Proposition 5. *Fix a player* i *and a CPS* μ^i *over co-players' superstrategy profiles. Then, there always exists at least one sequentially optimal superstrategy and one weakly sequentially optimal superstrategy. Furthermore, every superstrategy behaviorally equivalent to a sequentially optimal superstrategy is weakly sequentially optimal. The same results hold for the one-period game.*

Corollary 3. *Fix $i \in I$, superstrategy $\mathbf{s}_i$, and CPS μ^i over superstrategies. Then $\mathbf{s}_i$ is weakly sequentially optimal given μ^i if and only if there exists a behaviorally equivalent strategy $\bar{\mathbf{s}}_i$ which is one-step optimal given μ^i.*

In conclusion, our representation of rationality is equivalent to sequential optimality, and behaviorally equivalent to weak sequential optimality. The latter—whenever players are not impatient—coincides with the representation of rationality in Battigalli & Tebaldi (2019).

References

[1] Battigalli, P. (1987): "Comportamento Razionale ed Equilibrio nei Giochi e nelle Situazioni Sociali," unpublished thesis, Università Bocconi.

[2] Battigalli, P., E. Catonini, G. Lanzani, and M. Marinacci (2019): "Ambiguity attitudes and self-confirming equilibrium in sequential games," *Games and Economic Behavior*, 115, 1-29.

[3] Battigalli, P., S. C. Cerreia-Vioglio, F. Maccheroni, and M. Marinacci (2015): "Self-Confirming Equilibrium and Model Uncertainty," *American Economic Review*, 105, 646-677.

[4] Battigalli, P., and N. De Vito (2021): "Beliefs, plans, and perceived intentions in dynamic games," *Journal of Economic Theory*, 195, 1052-83.

[5] Battigalli, P., and N. Generoso (2021): "Information Flows and Memory in Games", IGIER Working Paper, no. 678.

[6] Battigalli, P., and D. Guaitoli (1988): "Conjectural Equilibria and Rationalizability in a Game with Incomplete Information," Quaderni di Ricerca, Università Bocconi (published in *Decisions, Games and Markets*, Kluwer, Dordrecht, 97-124, 1997).

[7] Battigalli, P., and M. Siniscalchi (2002): "Strong Belief and Forward Induction Reasoning," *Journal of Economic Theory*, 106, 356-391.

[8] Battigalli, P., and P. Tebaldi (2019): "Interactive Epistemology in Simple Dynamic Games with a Continuum of Strategies," *Economic Theory*, 68, 737-763.

[9] Dekel, E., D. Fudenberg, and D.K. Levine (1999): "Payoff Information and Self-confirming Equilibrium," *Journal of Economic Theory*, 89, 165-185.

[10] Dekel, E., and M. Siniscalchi (2015): "Epistemic Game Theory," in *Handbook of Game Theory with Economic Applications, Volume 4*, ed. by P. Young and S. Zamir. Amsterdam: North-Holland, 619-702.

[11] Esponda, I. (2013): "Rationalizable Conjectural Equilibrium: A Framework for Robust Predictions," *Theoretical Economics*, 8, 467-501.

[12] Fudenberg, D., and Y. Kamada (2015): "Rationalizable Partition-Confirmed Equilibrium," *Theoretical Economics*, 10, 775-806.

[13] Fudenberg, D., and Y. Kamada (2018): "Rationalizable Partition-Confirmed Equilibrium with Heterogeneous Conjectures," *Games and Economic Behavior*, 109, 364-381.

[14] Fudenberg, D., and D.M. Kreps (1995): "Learning in Extensive-Form Games I. Self-Confirming Equilibria," *Games and Economic Behaviour*, 8, 20-55.

[15] Fudenberg, D., and D.K. Levine (1993): "Self-Confirming Equilibrium," *Econometrica*, 61, 523-545.

[16] Gilli, M. (1999): "Adaptive Learning in Imperfect Monitoring Games," *Review of Economic Dynamics*, 2, 472-485.

[17] Kalai, E., and E. Lehrer (1993): "Rational Learning leads to Nash Equilibrium," *Econometrica*, 61, 1019-1045.

[18] Kalai, E., and E. Lehrer (1995): "Subjective Games and Equilibria," *Games and Economic Behaviour*, 8, 123-163.

[19] Myerson, E. (1986): "Multistage Games with Communication," *Econometrica*, 54, 323-58.

[20] Pearce, D.G. (1984): "Rationalizable Strategic Behavior and the Problem of Perfection," *Econometrica*, 52, 1029-50.

[21] Rubinstein, A., and A. Wolinsky (1994): "Rationalizable Conjectural Equilibrium: Between Nash and Rationalizability," *Games and Economic Behavior*, 6, 299-311.

[22] Schipper, B. (2021): "Discovery and Equilibrium in Games with Unawareness," *Journal of Economic Theory*, 198, 105365.