A characterization of consistent assessments using power sequences of strategy profiles

Francesc Dilmé[1] ⬤

Abstract

This paper provides a new characterization of belief consistency in extensive games. We show that *all* consistent assessments are supported by sequences of strategy profiles with the property that all actions with vanishing probability are played according to power functions of the sequence index. The result makes it simpler to prove or disprove that a given assessment is consistent, facilitating the use of sequential equilibria.

1 Introduction

Sequential rationality is a key plausibility requirement in extensive games: It requires that players behave optimally at each point of the game given the other players' behavior and available information. If information is perfect, requiring equilibrium strategies to be sequentially rational is equivalent to subgame perfection (Selten 1965). If information is imperfect, requiring sequential rationality requires specifying players' beliefs about previous play at each decision node. On the path of play, Bayes' rule pins down beliefs. Off the path of play, Bayes' rule is not well defined, raising the question of which restrictions on beliefs are plausible and/or desirable. Selten (1975) addressed the question by posing the possibility that players may make mistakes with small probabilities. He defined (*trembling-hand*) *perfect equilibria* as limits of equilibria of versions of the game perturbed according to a sequence of vanishing trembles. In each perturbed game, Bayes' rule can always be applied since all instances of the game occur with positive probability on the path of play.

✉ Francesc Dilmé
  fdilme@uni-bonn.de

[1] University of Bonn, Bonn, Germany

While requiring perfection is a powerful and sensible way to refine the set of Nash equilibria, it is difficult to use in practice. The reason is that obtaining and working with the right tremble sequences and the corresponding sequences of equilibria to prove or disprove that a strategy profile is a perfect equilibrium is often cumbersome and many times not feasible. To address this problem, Kreps and Wilson (1982) proposed a weaker but easier-to-use equilibrium concept, *sequential equilibrium*, where beliefs are obtained as the limit of fully mixed sequences of strategy profiles, but sequential rationality is only required in the limit.[1] Still, while using sequential equilibria is indeed simpler than using perfect equilibria, it suffers from a similar problem: proving or disproving that a given assessment is consistent requires proving or disproving that it is not supported by any sequence of strategy profiles, which is often difficult to do. This has led to the use of other, weaker equilibrium concepts, such as different versions of perfect Bayesian equilibrium, first defined in Fudenberg and Tirole (1991).

This paper provides a simple characterization of assessment consistency. We show that an assessment (σ, μ) is consistent if and only if it is supported by a *power sequence of strategy profiles*, that is, a sequence of strategy profiles $(\sigma_n)_n$ for which, for each action a, there is a pair $(x(a), y(a)) \in \mathbb{R}_{++} \times \mathbb{R}_+$ satisfying that $\sigma_n(a) = x(a) n^{-y(a)}$ for all n when $y(a) > 0$, and $\sigma_n(a) \to x(a)$ when $y(a) = 0$.[2] We argue that since obtaining assessments from power sequences of strategy profiles is straightforward, our result greatly simplifies proving or disproving that a given assessment is consistent.

The proof of the result is involved. To understand its difficulties, take a consistent belief system (σ, μ), meaning that some sequence of fully-mixed strategy profiles $(\sigma_n)_n$ supports it. The probabilities with which actions are played along the sequence may not be power functions of the index and may differ significantly across actions. The crucial step in the proof is showing that there is a strictly increasing sequence $(\hat{n}_n \in \mathbb{N})_n$ and a K-dimensional sequence $(q_n^1, \ldots, q_n^K)_n$, for some $K \in \mathbb{N}$, with the properties that, for all $a \in A$,

$$\lim_{n \to \infty} \frac{\sigma_{\hat{n}_n}(a)}{\alpha^0(a) \prod_{k=1}^{K} (q_n^k)^{\alpha^k(a)}} = 1$$

for unique values $\alpha(a) \in \mathbb{R}^{K+1}$, and that $\lim_{n \to \infty} (q_n^k)^\gamma / q_n^{k+1} = 0$ for all k and $\gamma \in \mathbb{R}$. The sequence $(q_n^1, \ldots, q_n^K)_n$ contains the different types of convergence to 0 across different actions. The value $\alpha(a)$ can then be used to obtain a pair $(x(a), y(a))$ such that the corresponding power sequence of strategy profiles supports (σ, μ).

The class of power sequences of strategy profiles is particularly convenient to use. We show that the relative likelihoods of different histories can be straightforwardly obtained from the values that the maps x and y assign to their actions. Conversely, the consistency of an assessment is characterized by simple equations involving the

---

values that x and y take for the actions in different histories. We also show that any consistent assessment is supported by a power sequence of strategy profiles with integer exponents, further simplifying the analysis of sequential equilibria.

Literature review Since the seminal paper of Kreps and Wilson (1982), belief consistency has been seen as a natural requirement in extensive games. Still, as working with (high-dimensional) sequences of strategy profiles is not feasible in many applications, other equilibrium concepts have been favored, most notably (weak/strong) perfect Bayesian equilibrium (PBE) as defined in Fudenberg and Tirole (1991), where off-path beliefs are updated using Bayes' rule "whenever possible". PBE is often too permissive, so ad-hoc restrictions on off-path belief updating are often imposed, such as the "no signaling what you do not know" and "never dissuaded once convinced" conditions (see Osborne and Rubinstein 1994). Alternative restrictions on belief updating off the path of play have been used in Cramton (1985), Rubinstein (1985), Bagwell (1990), and Harrington (1993). The current paper takes a different direction by identifying a simple set of sequences of strategy profiles that is sufficient to characterize consistency. In separate work (Dilmé 2023a), we use the current characterization to provide a full characterization of sequential equilibria in terms of lexicographic numbers.

The rest of the paper is organized as follows. Section 2 provides the notation for extensive form games and consistent assessments. Section 3 states and proves our main result. Section 4 provides techniques to prove or disprove the consistency of an assessment and illustrates them through examples. An "Appendix" contains the omitted proofs.

2 Belief consistency in extensive-form games

We now provide the definitions and corresponding notation for an extensive-form game.

An (finite) *extensive-form game* $G \equiv \langle A, H, \mathcal{I}, N, \pi, u \rangle$ has the following components. A finite set of *actions* A. A finite set of *histories* H containing finite sequences of actions satisfying that, for all $(a_j)_{j=1}^{J} \in H$ with $J > 0$, we have $(a_j)_{j=1}^{J-1} \in H$; the set of *terminal histories* is denoted T. An *information partition* $\mathcal{I}$ of the set of non-terminal histories satisfying that there is a partition $\{A^I | I \in \mathcal{I}\}$ of A with the property that, for each $I \in \mathcal{I}$ and $h, h' \in H$, we have (1) $(h, a) \in H$ for some $a \in A^I$ if and only if $h \in I$, and (2) if $h \in I$ and $h' > h$ then $h' \notin I$.[3] A finite set of *players* $N \not\ni 0$. A *player assignment* $\iota : \mathcal{I} \to N \cup \{0\}$ assigning each information set to a player or to nature. A *strategy by nature* $\pi : \cup_{I \in \iota^{-1}(\{0\})} A^I \to (0, 1]$ satisfying $\sum_{a \in A^I} \pi(a) = 1$ for each $I \in \iota^{-1}(\{0\})$. For each player $i \in N$, a *(von Neumann–Morgenstern) payoff function* $u_i : T \to \mathbb{R}$.

A *strategy profile* is a map $\sigma : A \to [0, 1]$ such that $\sum_{a \in A^I} \sigma(a) = 1$ for all $I \in \mathcal{I}$ (i.e., it is a probability distribution for each set of actions available at each

---

[3] Note that we assume, without loss of generality, that each action only belongs to a unique information set (otherwise, one can rename actions). We use $h' > h$ to indicate that h' succedes h.

information set) and $\sigma(a) = \pi(a)$ for all a played by nature (i.e., nature plays according to π). We let Σ be the set of strategy profiles.

2.1 Consistent assessments

As explained in the introduction, requiring that a player behaves sequentially rationally requires assigning beliefs over the information she does not know at the time she makes a decision. Such beliefs are modeled using a *belief system*, which is a map $\mu : H\backslash T \to [0, 1]$ satisfying that $\sum_{h \in I} \mu(h) = 1$ for all $I \in \mathcal{I}$. Kreps and Wilson (1982) define an *assessment* as a pair (σ, μ), where σ is a strategy profile and μ is a belief system. The usual interpretation is that for each information set I, action $a \in A^I$ and history $h \in I$, $\sigma(a)$ is the probability with which player $\iota(I)$ (i.e., the player taking action at this node of the game) chooses a while the value $\mu(h)$ is the posterior belief that this player assigns to h (both probabilities are conditional on I being reached).

Kreps and Wilson (1982) pointed out that not all assessments are plausible. Fix, for instance, a strategy profile σ and an on-path information set I. If a rational player assumes that players use σ, she should use Bayes' rule to assign the probability of $h \in I$ conditional on I having been reached. If σ is a full-support strategy profile (i.e., $\sigma(a) > 0$ for all $a \in A$), Bayes' rule pins down a unique assessment, where μ is given by

$$\mu(h|\sigma) = \frac{\sigma(h)}{\sigma(I)} \tag{2.1}$$

for all information sets $I \in \mathcal{I}$ and histories $h \in I$, where abusing notation $\sigma(\cdot)$ indicates the probability with which a history or an information set occurs under σ. Instead, if there is an information set I that is not reached under σ, both the numerator and the denominator of the right side of equation (2.1) are 0, thus the Bayes' rule cannot be applied.

Selten (1975) proposed studying perturbed versions of the game with "slight mistakes" (in the form of trembles), where all information sets occur on the path of play. Following a similar approach, Kreps and Wilson (1982) introduced the concept of *consistent assessment* as one that is approximated by a sequence of fully-mixed strategy profiles. Formally, (σ, μ) is a consistent assessment if it is *supported* by some fully-mixed sequence $(\sigma_n)_n$, that is, a sequence satisfying $\sigma_n(a) > 0$ and

$$(\sigma(a), \mu(h)) = \lim_{n \to \infty} (\sigma_n(a), \mu(h|\sigma_n)) \tag{2.2}$$

for all $a \in A$ and $h \in H\backslash T$. Consistency is a crucial requirement of Kreps and Wilson 's definition of sequential equilibrium, which, additionally, requires that σ is sequentially rational given μ.[4]

---

[4] It is easy to see that for a fixed assessment, the payoff a player obtains from playing a given action conditionally on the corresponding information set being reached is uniquely pinned down. Hence, (σ, μ) is *sequentially rational* if there is *no* information set $I \in \mathcal{I}$ and actions $a, a' \in A^I$ such that $\sigma(a) > 0$ and player $\iota(I)$'s payoff from playing a at I is strictly lower than that of playing a'.

Different reasons make using sequential equilibria desirable. First, consistency provides a systematic and powerful way to discipline off-path beliefs, hence avoiding using ad-hoc requirements often made in perfect Bayesian equilibria (see the Literature Review). Second, Blume and Zame (1994) showed that the set of sequential equilibria coincides with the set of limits of Nash equilibria of perturbed games where both actions and payoffs are perturbed. Hence, they are the predictions of an external observer who cannot perfectly assess the small likelihood of trembles or the exact payoffs of the players. Third, because perfect equilibrium outcomes are sequential equilibrium outcomes (perfection requires exact equilibria along the sequence) and also sequentially stable outcomes (sequential stability requires exact stability for all trembles; see Dilmé 2023b), ruling out that an outcome is sequential permits ruling out its perfection and sequential stability. Additionally, if one can prove that there is a unique sequential outcome, then such outcome is the unique perfect outcome and the unique sequentially stable outcome (and also the unique outcome of a proper equilibrium, see Myerson 1978, and the unique quasi-perfect equilibrium, see van Damme 1984).[5]

3 Main result

3.1 Power sequences of strategy profiles

In this section, we introduce a simple and natural class of sequences of strategy profiles. These are sequences of strategy profiles where actions that are played with vanishing probability decrease according to power functions of the sequence's index.

Definition 3.1 A $(\sigma_n)_n$ is a *power sequence of strategy profiles* if there exists a pair of maps $x : A \to \mathbb{R}_{++}$ and $y : A \to \mathbb{R}_+$ such that, for every $a \in A$ and $n \in \mathbb{N}$ large enough,[6]

$$\sigma_n(a) = \begin{cases} x(a) \, n^{-y(a)} & \text{if } y(a) > 0, \\ M_n(a) \, x(a) & \text{if } y(a) = 0, \end{cases} \tag{3.1}$$

where $M_n(a)$ is a factor satisfying that $\lim_{n\to\infty} M_n(a) = 1$.

In a power sequence of strategy profiles, the probability with which an action a is played along the sequence is characterized by two parameters. The first

---

[5] In games with generic payoffs, all sequential equilibria are perfect equilibria (see Kreps and Wilson 1982), so belief consistency and sequential rationality are sufficient to prove perfection.

[6] As we are only interested in how the sequence behaves for large n, we do not impose the functional form (3.1) for small n. We argue below that this is convenient as it permits stating that "there is a power sequence of strategy profiles with (x, y) if and only if $\sum_{a \in A^I | y(a)=0} x(a) = 1$ for all information sets $I \in \mathcal{I}$." Otherwise, (x, y) would have to satisfy additional conditions for some $(\sigma_n)_n$ satisfying (3.1) to exist because if, for example, $(x(a), y(a)) = (2, 2)$ for some $a \in A$, we would have $\sigma_1(a) = 2 \cdot 1^{-2} > 1$, which is not possible.

is a "rate" $y(a) \in \mathbb{R}_+$ at which the sequence tends to 0. The action a is played with probability 0 in the limit if and only if $y(a) > 0$. Also, if $y(a) > y(a')$ then a is played with a vanishing probability *relative* to a'. The second parameter is a multiplying factor $x(a) \in \mathbb{R}_{++}$. This factor determines the relative likelihood of actions with the same rate. If $y(a) = 0$, then $x(a)$ is the asymptotic probability with which action a is played; hence, it must be that $\sum_{a \in A^I | y(a)=0} x(a) = 1$ for all I. Note that if n is large enough, then σ_n is a fully-mixed strategy profile.

Power sequences of strategy profiles are simple and often used to provide examples of consistent assessments. The reason is that the asymptotic likelihood of each action is determined by only two numbers; hence, obtaining relative likelihoods is easy. As we shall see, even though the set of power sequences of strategy profiles is small relative to the set of all sequences, it is rich enough to support all consistent assessments. Furthermore, we show in Sect. 4 that the limit relative likelihoods of histories can be straightforwardly computed from the values that x and y take on their actions. It follows that a power sequence of strategy profiles has a well-defined limit belief system; that is, it supports a consistent assessment.

3.2 Main result

Our main result establishes that any consistent assessment is supported by a power sequence of strategy profiles.
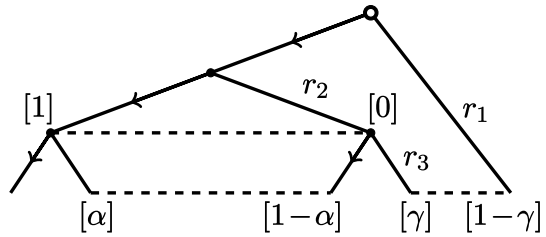
Theorem 3.1 *An assessment is consistent if and only if it is supported by a power sequence of strategy profiles.*

Theorem 3.1 is proven in Sect. 3.3. It establishes that the set of power sequences of strategy profiles is rich enough to support *all* consistent assessments. As we will see in Sect. 4, our result simplifies working with consistent assessments and hence eases the use of sequential equilibria.

We now provide an idea of the challenges that proving Theorem 3.1 poses. Since the "if" part of the statement is trivial (assessments supported by power sequences of strategy profiles are consistent), we focus on the "only if" part of the result.

Proving the "only if" part of Theorem 3.1 is involved. To see why, we fix a consistent assessment (σ, μ) and a sequence $(\hat{\sigma}_n)_n$ supporting it. Note that the rate at which $\hat{\sigma}_n(a)$ converges to $\sigma(a)$ may be very different for different actions $a \in A$, and not necessarily through a power sequence of the form (3.1). The complication of the proof lies in that each action typically belongs to different histories belonging to different information sets; hence, the choice of $(x(a), y(a))$ simultaneously affects the likelihood of all histories containing a. Guaranteeing that a power sequence of the form (3.1) supports (σ, μ) involves simultaneously satisfying a potentially large set of interconnected constraints without a clear structure, some of them in the form of inequalities. Our proof provides a method to obtain values for $\{(x(a), y(a)) | a \in A\}$ and some sequence $(\hat{n}_n)$ satisfying that, for all $h, h' \in H$,

Fig. 1 Game tree 1

$$\lim_{n\to\infty} \frac{\hat{\sigma}_{\hat{h}_n}(h)}{\hat{\sigma}_{\hat{h}_n}(h')} = \lim_{n\to\infty} \frac{\prod_{j=1}^{J} \hat{\sigma}_{\hat{h}_n}(a_j)}{\prod_{j=1}^{J'} \hat{\sigma}_{\hat{h}_n}(a'_j)} = \lim_{n\to\infty} \frac{\prod_{j=1}^{J} \sigma_n(a_j)}{\prod_{j=1}^{J'} \sigma_n(a'_j)} = \lim_{n\to\infty} \frac{\sigma_n(h)}{\sigma_n(h')},$$

where σ_n is a sequence of the form (3.1). We are then able to ensure that the asymptotic relative likelihoods of all pairs of histories under $(\hat{\sigma}_{\hat{h}_n})_n$ and $(\sigma_n)_n$ coincide.

To shed further light on the problem, consider the game in Fig. 1 (the game in the figure is assumed to continue after each depicted terminal history). Consider the depicted assessment, where the arrows indicate the actions played with positive probability. Consider a fully mixed sequence of strategy profiles $(\hat{\sigma}_n)_n$ such that

$$\hat{\sigma}_n(r_1) = \log(n+3)^{-1}, \quad \hat{\sigma}_n(r_2) = e^{-n} + e^{-n^3} \sin(n), \quad \text{and} \quad \hat{\sigma}_n(r_3) = 3\log(n+4)^{-1} e^{-n}. \tag{3.2}$$

This implies that

$$\alpha = \lim_{n\to\infty} \frac{(1 - \hat{\sigma}_n(r_1))(1 - \hat{\sigma}_n(r_2))\hat{\sigma}_n(r_3)}{(1 - \hat{\sigma}_n(r_1))(1 - \hat{\sigma}_n(r_2))\hat{\sigma}_n(r_3) + (1 - \hat{\sigma}_n(r_1))\hat{\sigma}_n(r_2)(1 - \hat{\sigma}_n(r_3))} = 1 \quad \text{and}$$

$$\gamma = \lim_{n\to\infty} \frac{\hat{\sigma}_n(r_2)\hat{\sigma}_n(r_3)}{\hat{\sigma}_n(r_2)\hat{\sigma}_n(r_3) + \hat{\sigma}_n(r_1)} = \frac{1}{4}.$$

Hence, for a power sequence of strategy profiles of the form (3.1) to exist, it must be that

$$0 < y(r_3) < y(r_2), \quad 0 < y(r_2) + y(r_3) = y(r_1), \quad \text{and} \quad x(r_1) = x(r_2)x(r_3)/3. \tag{3.3}$$

The first equation says that r_3 is played infinitely more likely than r_2 as $n \to 0$, but both probabilities tend to 0. The second equation says that the likelihood of history (r_1) vanishes at the same rate that history (r_2, r_3). The third equation says that the asymptotic relative likelihood of (r_1) relative to (r_2, r_3) is 1/3. While it is easy to find $(x, y) : A \to \mathbb{R}_{++} \times \mathbb{R}_+$ satisfying the constraints in (3.3) in this simple example, it is difficult to obtain a systematic procedure to prove the existence of solutions in general. The proof of Theorem 3.1 provides a constructive procedure that gives a power sequence of strategy profiles supporting any given consistent assessment of an arbitrary extensive-form game.

3.3 Proof of Theorem 3.1

We here present a proof of the "only if" part of the statement of Theorem 3.1, since the "if" part is trivial, as we explained before. A reader interested in the use of Theorem 3.1 to work with consistent assessments can directly move to Sect. 4.

Proof of Theorem 3.1 We begin with a useful lemma, proven in the "Appendix".

Lemma 3.1 (Representation) *Let* $\hat{A} = \{a_1, \ldots, a_J\}$ *be a finite set, and* $(\sigma_n : \hat{A} \to (0,1])_n$ *be a sequence. There are a strictly increasing sequence* $(\hat{n}_n \in \mathbb{N})_n$ *and a sequence* $((q_n^1, \ldots, q_n^K) \in \mathbb{R}_{++}^K)_n$, *for some* $K \in \{0, \ldots, J\}$, *such that*

1. $\lim_{n\to\infty} q_{\hat{n}_n}^1 = 0$ *and* $\lim_{n\to\infty} q_{\hat{n}_n}^k / (q_{\hat{n}_n}^{k-1})^\gamma = 0$ *for all* $\gamma \in \mathbb{R}$ *and* $k = 2, \ldots, K$.
2. *For each* $a \in A$ *there is a unique* $\alpha(a) \equiv (\alpha^1(a), \ldots, \alpha^K(a)) \in \mathbb{R}^K$ *such that*

$$\lim_{n\to\infty} \frac{\sigma_{\hat{n}_n}(a)}{\prod_{k=1}^K (q_{\hat{n}_n}^k)^{\alpha^k(a)}} \in \mathbb{R}_{++}.$$

3. *There are two sequences* $(\hat{j}_k)_{k=1}^K$ *with* $1 \le \hat{j}_1 < \cdots < \hat{j}_K \le J$ *and* $(\hat{k}_k)_{k=1}^K$ *with* $\hat{k}_k \ne \hat{k}_{k'}$ *for all* k, k', *such that, for all* $k \in \{1, \ldots, K\}$, *we have* (1) $\alpha^{\hat{k}_{k'}}(\hat{a}_j) = 0$ *for all* $k' > k$ *and* $j < \hat{j}_k$, *and* (2) $\alpha^{\hat{k}_k}(\hat{a}_{\hat{j}_k}) \ne 0$.

Lemma 3.1 shows that any finite collection of sequences on (0, 1] can be represented as the product of the powers of a basis of sequences. Its proof constructively provides such sequences by proceeding iteratively over the set of actions. In each iteration j, we expand the basis of sequences only if (an appropriate subsequence of) $\sigma_n(a_j)$ tends to 0 at a rate that cannot be expressed as the product of powers of the current basis.[7] The case $K = 0$ corresponds to the case where $(\hat{n}_n)_n$ is such that $\lim_{n\to\infty} \sigma_{\hat{n}_n}(a_j) > 0$ for all $j = 1, \ldots, J$.

Fix a consistent assessment (σ, μ) supported by some sequence $(\sigma_n)_n$. Let $(\hat{n}_n)_n$ and $(q_n^1, \ldots, q_n^K)_n$ be two sequences with the properties in the statement of Lemma 3.1 applied to $(\sigma_n)_n$ and A. We then have that, for each action $a \in A$, there exist unique $\alpha^0(a) \in \mathbb{R}_{++}$ and $\alpha(a) \in \mathbb{R}^K$ such that

$$\lim_{n\to\infty} \frac{\sigma_{\hat{n}_n}(a)}{\alpha^0(a) \prod_{k=1}^K (q_{\hat{n}_n}^k)^{\alpha^k(a)}} = 1.$$

---

[7] For each k, $\hat{j}_k$ indicates the step in the process where $q_n^{\hat{k}_k}$ was added to the basis. In the proof, if the current basis is $\{q_n^k | k = 1, \ldots, k'\}$ and $\sigma_n(a_j)$ cannot be expressed as the combination of the current basis, we show there are some coefficients $\{\alpha(a_j) | k = 1, \ldots, k'\}$ such that, defining $q_n = \sigma_n(a_j) / \prod_{k=1}^{k'} q_n^{a^k}$, either q_n or q_n^{-1} satisfy the conditions of the statement.

Note that $\sigma(a) > 0$ if and only if $\alpha^k(a) = 0$ for all $k = 1, \ldots, K$, in which case $\sigma(a) = \alpha^0(a)$. Similarly, it follows that, for each history $h = (a_1, \ldots, a_J) \in H$, we have

$$\lim_{n \to \infty} \frac{\sigma_{\hat{n}_n}(h)}{\left(\prod_{j=1}^{J} \alpha^0(a_j)\right) \prod_{k=1}^{K} (q_{\hat{n}_n}^k)^{\alpha^k(h)}} = 1,$$

where $\alpha(h) := \sum_{j=1}^{J} \alpha(a_j)$.

We define the reversed lexicographic order $\succ$ on $\mathbb{R}^K$ as follows. For each pair $\alpha(h), \alpha(h') \in \mathbb{R}^K$, $\alpha(h) \succ \alpha(h')$ indicates that there is some $k \in \{1, \ldots, K\}$ such that $\alpha^{k'}(h) = \alpha^{k'}(h')$ for all $k' = k+1, \ldots, K$ and $\alpha^k(h) > \alpha^k(h')$. By the first property of the statement of Lemma 3.1, we have that if $\alpha(h) \succ \alpha(h')$ then $\lim_{n \to \infty} \sigma_{\hat{n}_n}(h) / \sigma_{\hat{n}_n}(h') = 0$.[8] Our goal is to use α and α^0 to construct the maps x and y such that the power sequence of strategy profiles in Eq. (3.1) supports (σ, μ).

For each $M \in \mathbb{R}_{++}$, we define

$$(q_{M,n}^1, \ldots, q_{M,n}^K) := \left((q_n^1)^{1/M}, (q_n^2)^{1/M^2}, \ldots, (q_n^K)^{1/M^K}\right)$$

and

$$(\alpha_M^1(a), \alpha_M^2(a), \ldots, \alpha_M^K(a)) := \left(M^1 \, \alpha^1(a), M^2 \, \alpha^1(a), \ldots, M^K \, \alpha^K(a)\right),$$

and note that $\prod_{k=1}^{K} (q_{M,n}^k)^{\alpha_M^k(a)} = \prod_{k=1}^{K} (q_n^k)^{\alpha^k(a)}$. We also define the pair of maps $(x_M, y_M) : A \to \mathbb{R}_{++} \times \mathbb{R}_+$ assigning, to each action $a \in A$, the value

$$(x_M(a), y_M(a)) := \left(\alpha^0(a), \sum_{k=1}^{K} \alpha_M^k(a)\right). \tag{3.4}$$

It is then clear that, if M is large enough, then $y_M(a) > y_M(a')$ if and only if $\alpha(a) \succ \alpha(a')$, and also that for all histories $h, h' \in H$,

$$\lim_{n \to \infty} \frac{\sigma_{\hat{n}_n}(h)}{\sigma_{\hat{n}_n}(h')} = \begin{cases} x_M(h)/x_M(h') & \text{if } y_M(h) = y_M(h'), \\ 0 & \text{if } y_M(h) < y_M(h'), \\ \infty & \text{if } y_M(h) > y_M(h'). \end{cases} \tag{3.5}$$

In particular, if M is large enough, we have $\sigma(a) = 0$ if and only if $y_M(a) = 0$, and therefore $\sum_{a \in A^I | y(a) = 0} x_M(a) = 1$ (recall that $\sigma(a) = \alpha^0(a)$ when $\sigma(a) > 0$). In conclusion, if M is large enough, then (σ, μ) is supported by a power sequence of strategy profiles satisfying Eq. (3.1) with $(x, y) := (x_M, y_M)$, so the proof of Theorem 3.1 is complete. □

---

[8] See Part 4 of the proof of Lemma 3.1 for a related argument.

4 Considerations

In this section, we illustrate how Theorem 3.1 can be used to obtain consistent assessments easily. Section 4.1 introduces a result that is useful in finding consistent assessments. Section 4.2 shows that the exponents in a power sequence of strategy profiles can be chosen to be integers. Section 4.3 examines alternative classes of sequences of strategy profiles with similar properties. Section 4.4 provides some examples illustrating how our results can be used to find consistent assessments.

4.1 Constructing consistent assessments

We begin with a useful corollary of Theorem 3.1.

Corollary 4.1 *Fix a sequence of power strategy profiles* $(\sigma_n)_n$ *as in Eq.* (3.1) *for some pair* (x, y). *Then, for any history* $h \equiv (a_j)_{j=1}^{J} \in H$,

$$\lim_{n \to \infty} \overbrace{\frac{\sigma_n(h)}{(\prod_{j=1}^{J} x(a_j)) \, n^{-\sum_{j=1}^{J} y(a_j)}}}^{=: (*)} = 1. \tag{4.1}$$

Corollary 4.1 establishes that the asymptotic likelihood of a history h can be easily obtained from the values that x and y take on its actions. This is clear if h is only composed of actions with vanishing probabilities; in this case, the term $(*)$ in (4.1) is exactly 1 for all n. Otherwise, each action with a non-vanishing probability is typically not a power function of n, but the sum of polynomial functions of n. These additional terms typically make the expressions of the probability of a history h under σ_n long and complicated. Still, given that $\lim_{n \to \infty} M_n(a) = 1$ (i.e., $\sigma_n(a) \to x(a)$), the belief system can be computed "as if" the probabilities of all actions were of the form $x(a) \, n^{-y(a)}$, including those with $y(a) = 0$.

Generating consistent assessments is easy using Corollary 4.1: First, choose a pair of maps $(x, y) : A \to \mathbb{R}_{++} \times \mathbb{R}_{+}$ with the condition that $\sum_{a \in A^I \mid y(a) = 0} x(a) = 1$ for all I. For each action a define $\sigma(a) := x(a)$ if $y(a) = 0$ and $\sigma(a) := 0$ otherwise. Next, for each history $h \equiv (a_j)_{j=1}^{J}$, compute

$$x(h) := \prod_{j=1}^{J} x(a_j) \quad \text{and} \quad y(h) := \sum_{j=1}^{J} y(a_j). \tag{4.2}$$

Then, for each information set I, only histories $h \in I$ with the lowest value $y(h)$ receive a positive probability under μ, and the relative probability of two such histories $h, h' \in I$ is $x(h)/x(h')$.

Verifying that a given assessment (σ, μ) is consistent (or proving it is not) is more difficult but can often be done using Corollary 4.1 as follows. First, set $x(a) := \sigma(a)$ and $y(a) := 0$ for all $a \in A$ such that $\sigma(a) > 0$. Second, for the rest of the actions,

look for values $x(a), y(a) > 0$ so that for all histories $h, h', h'' \in H$ in the same information set with $\mu(h), \mu(h') > \mu(h'') = 0$ we have $y(h) = y(h') < y(h'')$ and $x(h)/x(h') = \mu(h)/\mu(h')$.

An important observation is that obtaining the values $\{x(a)|a \in A\}$ and $\{y(a)|a \in A\}$ can be done separately. Indeed, to obtain maps x and y supporting a given assessment, the first step consists in obtaining y, which determines the relative rates at which the probability of different histories vanish. The second step consists in finding x so that the implied relative probabilities of histories with a positive probability coincide with those implied by the belief system.[9] For example, it is easy to see that if a consistent assessment (σ, μ) is such that $\mu(h) \in \{0, 1\}$ for all $h \in H \backslash T$, then there is a power sequence of strategy profiles with $x(a) = 1$ for all actions satisfying that $\sigma(a) = 0$. The examples below illustrate this procedure.

Finally, we state that power sequences of strategy profiles can be rescaled while supporting the same assessment.

Proposition 4.1 *Fix* $(x, y) : A \to \mathbb{R}_{++} \times \mathbb{R}_+$ *such that* $\sum_{a \in A^I | y(a)=0} x(a) = 1$ *for all* $I \in \mathcal{I}$. *Let* $K_1, K_2 \in \mathbb{R}_{++}$ *be two constants. Define*

$$(\tilde{x}(a), \tilde{y}(a)) := (K_2^{y(a)} x(a), K_1 y(a)) \quad \text{for all } a \in A.$$

Then, the power sequences of strategy profiles defined by (x, y) *and* $(\tilde{x}, \tilde{y})$ *support the same assessment.*

The intuition for Proposition 4.1 is the following. First, note that two histories $h \equiv (a_j)_{j=1}^J$ and $h' \equiv (a'_{j'})_{j'=1}^{J'}$ satisfy $\tilde{y}(h) < \tilde{y}(h')$ if and only if they also satisfy $y(h) < y(h')$. Also, if $\tilde{y}(h) = \tilde{y}(h')$ then $y(h) = y(h')$, and

$$\frac{\tilde{x}(h)}{\tilde{x}(h')} = \frac{\prod_{j=1}^J \tilde{x}(a_j)}{\prod_{j'=1}^{J'} \tilde{x}(a'_{j'})} = \frac{K_2^{y(h)} \prod_{j=1}^J x(a_j)}{K_2^{y(h')} \prod_{j'=1}^{J'} x(a'_{j'})} = \frac{\prod_{j=1}^J x(a_j)}{\prod_{j'=1}^{J'} x(a'_{j'})} = \frac{x(h)}{x(h')}.$$

Hence, the parameters associated with actions played with vanishing probabilities can be rescaled while supporting the same assessment (notice that $(\tilde{x}(a), \tilde{y}(a)) = (x(a), y(a))$ when $y(a) = 0$).

4.2 Integer exponents

Our definition of power sequences of strategy profiles (Definition 3.1) allows the rate of convergence of each action's probability to be any non-negative real number. In practice, nevertheless, integer exponents are often used in examples. In this section, we show that all assessments are supported by power sequences of strategy

---

[9] Note that x and y are pinned down uniquely for actions played with positive probability in the limit. Indeed, if $\sigma(a) > 0$, then $x(a) = \sigma(a)$ and $y(a) = 0$.

profiles with integer exponents; hence, the assumption of integer exponents can be made without loss of generality.

Definition 4.1 We say that $(\sigma_n)_n$ is an *integer power sequence of strategy profiles* if it can be written as in (3.1) for some functions $x : A \to \mathbb{R}_{++}$ and $y : A \to \mathbb{Z}_+$.

The class of integer power sequences of strategy profiles is significantly smaller than the class of power sequences of strategy profiles. Still, as the following proposition states, it is large enough to support all consistent assessments.

Proposition 4.2 *An assessment is consistent if and only if some integer power sequence of strategy profiles supports it.*

Proposition 4.2 establishes that any consistent assessment is supported by an integer power sequence of strategy profiles; that is, the rates of convergence to 0 of different actions can be chosen to be non-negative integers. There are various reasons for doing so. One is pedagogical, as integers make it easier to categorize different convergence rates. Another is a potential technical advantage since looking for integer solutions to systems of equations is often numerically easier. In many cases, nevertheless, not restricting to integer exponents is easier as it gives more flexibility to find or construct consistent assessments.

Remark 4.1 Our characterization of sequential equilibria using monomials with integer powers is reminiscent of Blume and Zame (1994)'s characterization of sequential equilibria using algebraic geometry. They show that the graph of the equilibrium correspondence for sequential equilibria is a semi-algebraic set; that is, it is defined by finite systems of polynomial inequalities (with integer powers). While we would find a thorough exploration of the connection between these two approaches interesting, it lies outside the boundaries of the current paper's focus.

4.3 Other convenient sequences

Power sequences of strategy profiles are easy and convenient to use. They are determined by two maps (x and y in Definition 3.1 with the only condition that $\sum_{a \in A^I | y(a)=0} x(a) = 1$), and they permit easily computing relative likelihoods of histories (recall Corollary 4.1). In this section, we present other classes of sequences of strategy profiles that are convenient to use.

We begin by noticing three desirable properties of the class of power sequences of strategy profiles. First, a single parameter determines the relative limit likelihood of actions: $\lim_{n \to \infty} \sigma_n(a)/\sigma_n(a') = 0$ if and only if $y(a) > y(a')$. Second, another parameter determines the relative probability of the actions with the same convergence rate: $\lim_{n \to \infty} \sigma_n(a)/\sigma_n(a') = x(a)/x(a')$ whenever $y(a) = y(a')$. Finally, the procedure to obtain the rate of convergence and relative probability for histories from their actions involves simple operations (addition and multiplication, see Eq. (4.2)).

Fig. 2 Game tree 2
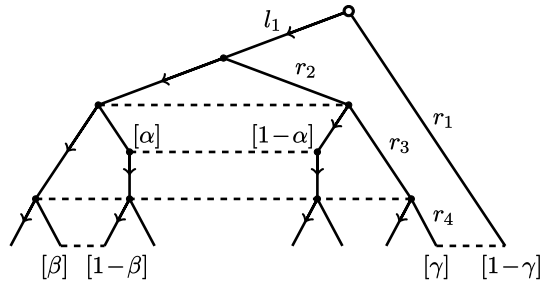
We now show that other classes of sequences of strategy profiles share similar properties. We begin noticing that, for any function $f : \mathbb{N} \to \mathbb{R}_{++}$ satisfying that $\lim_{n\to\infty} f(n) = 0$ and any consistent assessment (σ, μ), there exists a pair $(x, y) : A \to \mathbb{R}_{++} \times \mathbb{R}_+$ such that (σ, μ) is supported by

$$\sigma_n(a) := \begin{cases} x(a)f(n)^{-y(a)} & \text{if } y(a) > 0, \\ M_n(a)x(a) & \text{if } y(a) = 0, \end{cases} \tag{4.3}$$

where $M_n(a)$ is a factor satisfying that $\lim_{n\to\infty} M_n(a) = 1$. The convergence rates of histories can be computed using Eq. (4.2). So, the properties of a class of sequences of strategy profiles are the same as the properties of power sequences of strategy profiles, so they support the same assessment (for a given pair (x, y)) independently of the value of f.

A slightly different formulation is the following. It is easy to see that, for each $f : \mathbb{N} \to \mathbb{R}_+$ such that $\lim_{n\to\infty} f(n) = +\infty$ and for each consistent assessment (σ, μ), there is a pair of maps $(\hat{x}, \hat{y}) : A \to \mathbb{R}_{++} \times (0, 1]$ (note that the co-domain of $\hat{y}$ is different from before) such that (σ, μ) is supported by a sequence $(\hat{\sigma}_n)_n$ satisfying that, if n is large enough,

$$\hat{\sigma}_n(a) = \begin{cases} \hat{x}(a)\hat{y}(a)^{-f(n)} & \text{if } \hat{y}(a) > 0, \\ \hat{M}_n(a)\hat{x}(a) & \text{if } \hat{y}(a) = 1, \end{cases} \tag{4.4}$$

where again $\hat{M}_n(a)$ is a factor satisfying that $\lim_{n\to\infty} \hat{M}_n(a) = 1$. In this case, an action is played with vanishing probability if and only if $\hat{y}(a) < 1$. Now, the converge rate of a history $h \equiv (a_j)_{j=1}^J$ is given by

$$\hat{x}(h) := \prod_{j=1}^J \hat{x}(a_j) \quad \text{and} \quad \hat{y}(h) := \prod_{j=1}^J \hat{y}(a_j)$$

instead of by Eq. (4.2). In this case, it is easy to see that the assessment supported by $(\hat{\sigma}_n)_n$ for $(\hat{x}, \hat{y})$ is the same as the assessment supported by $(\sigma_n)_n$ satisfying (3.1) with $(x, y) := (\hat{x}, -\log(\hat{y}))$.

Fig. 3 Game tree 3
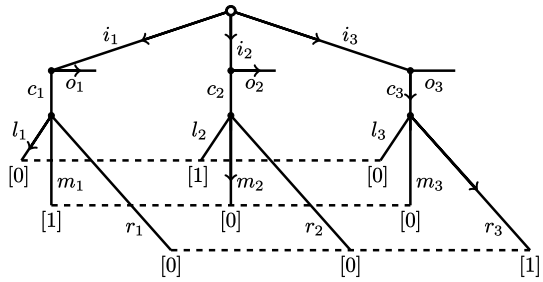
4.4 Examples

Example 4.1 We first illustrate how consistent assessments can be easily generated using power sequences of strategy profiles. Consider Fig. 2, which represents part of a tree of a game where all information sets belong to different players (the game is assumed to continue after each depicted terminal history). Assign, for example, $(x(r_1), y(r_1)) := (1, 4)$, $(x(r_2), y(r_2)) := (2, 2)$, $(x(r_3), y(r_3)) := (2, 1)$ and $(x(r_4), y(r_4)) := (1, 1)$ (note that, necessarily, all other actions get assigned $(1, 0)$). For example, note that

$$(x(r_1), y(r_1)) = (1, 4) \text{ and}$$
$$(x(l_1, r_2, r_3, r_4), y(l_1, r_2, r_3, r_4)) = (1 \cdot 2 \cdot 2 \cdot 1, 0 + 2 + 1 + 1) = (4, 4).$$

Hence, since $y(r_1) = y(l_1, r_2, r_3, r_4) = 4$, we have that $(1 - \gamma)/\gamma = x(r_1)/x(l_1, r_2, r_3, r_4) = 1/4$. It is then easy to see that this implies that $\alpha = 1$, $\beta = 1/3$, and $\gamma = 4/5$; hence, such assessment is consistent.

Example 4.2 The example illustrates how a power sequence of strategy profiles can be obtained from a given assessment. Consider Fig. 2 again. Assume that, under an assessment, all $\{r_j\}_{j=1}^4$ are played with probability 0 (hence the rest of the actions are played with probability 1) and $\alpha, \beta, \gamma \in (0, 1)$. Set $x(r_2) := 1$ and $y(r_2) := 1$ (note that, by Proposition 4.1, the values of x and y assigned to one of the actions played with probability zero can be chosen freely). That $\alpha \in (0, 1)$ implies $y(r_3) = 1$ and $x(r_3) = \alpha/(1 - \alpha)$. Similarly, that $\beta \in (0, 1)$ implies $y(r_4) = 1$ and $x(r_4) = \alpha \beta/((1 - \alpha)(1 - \beta))$. Finally, that $\gamma \in (0, 1)$ implies $y(r_1) = 3$ and $x(r_1) = \alpha^2 \beta (1 - \gamma)/((1 - \alpha)^2 (1 - \beta) \gamma)$.

Example 4.3 Fig. 3 reproduces the game depicted in Figure 235.1 in Osborne and Rubinstein (1994) (for graphical clarity, we omit the continuation game after the dashed information sets). In their Example 234.3, they argue through a lengthy argument that the assessment where only actions with an arrow are played with positive probability and with the belief system given in the picture is *not* consistent (their argument further relies on their assumption that i_1, i_2, and i_3 are played with the same probability).

We now argue that their result follows trivially using Theorem 3.1 (without the need to assume that i_1, i_2, and i_3 are played with the same probability). If the assessment in Fig. 3 would be consistent, then there would exist functions $(x, y) : A \rightarrow \mathbb{R}_{++} \times \mathbb{R}_+$ satisfying

$$y(c_2) + y(l_2) < \min\{y(c_1), y(l_3)\} \text{ and } y(c_1) + y(m_1) < \min\{y(c_2), y(m_3)\}.$$

The first equation requires that the exponent of n for the history (i_2, c_2, l_2) is strictly smaller (in absolute value) than those of the other histories in the same information set (note that $y(i_2) = 0$), while the second equation requires that the exponent of n for the history (i_1, c_1, m_1) is strictly lower than those of the other histories of the same information set. The contradiction is clear: the left equation requires $y(c_2) < y(c_1)$, while the second one requires $y(c_1) < y(c_2)$.

A Omitted proofs

Proof of Lemma 3.1

Proof We proceed by induction over the number of actions. If $\hat{A}$ has one action (i.e., $\hat{A} = \{a_1\}$) the result is clear: we let $(\hat{n}_n)_n$ be strictly increasing and such that $(\sigma_{\hat{n}_n}(a_1))_n$ converges to some $\sigma(a_1) \in [0, 1]$. If $\sigma(a_1) > 0$ then $K := 0$, and if $\sigma(a_1) = 0$ then $K := 1$ and $(q_n^1)_n := (\sigma_n(a_1))_n$. Assume then that the result is true for $\hat{A} := \{a_1, \dots, a_J\}$ with $J \geq 1$, and we prove it is true for $\hat{A} \cup \{\hat{a}\}$ (we use $\hat{a}$ to denote a_{J+1} for notational convenience). We use $(\hat{n}_n)_n$ and $(q_n^1, \dots, q_n^K)_n$ to denote the sequences provided by the theorem for $(\sigma_n : \hat{A} \rightarrow (0, 1])_n$. We divide the proof into four parts.

Part 1: The algorithm We propose the following algorithm, which follows steps $k = K, \dots, 1$ in descending order, from K to when the algorithm stops. In each step k, the algorithm takes the sequences $(\hat{n}_n^k)_n$ and $(\tilde{q}_n^k \in \mathbb{R}_{++})_n$ from the previous step, which satisfy $\lim_{n \to \infty} \tilde{q}_{\hat{n}_n^k}^k$ exists (in $\overline{\mathbb{R}}_+$). We initialize $(\hat{n}_n^K)_n$ to be a subsequence of $(\hat{n}_n)_n$ such that $(\sigma_{\hat{n}_n^K}(\hat{a}))_n$ is convergent, and $(\tilde{q}_n^K)_n := (\sigma_n(\hat{a}))_n$. The final outputs of the algorithm are $\alpha_* \in \mathbb{R}^K$, $(\hat{n}_n^*)_n$, $(\tilde{q}_n^0, \dots, \tilde{q}_n^K)_n$, and $k^* \in \{0, \dots, K\}$ (k^* indicates the step where the algorithm stopped.)

<u>Step 1:</u>

(a) If $\liminf_{n \to \infty} \tilde{q}_{\hat{n}_n^k}^k / (q_{\hat{n}_n^k}^k)^\gamma = 0$ for all $\gamma \in \mathbb{R}$, then we set $(\hat{n}_n^*)_n$ to be the subsequence of $(\hat{n}_n^k)_n$, defined as follows:

    i. $\hat{n}_1^* := \hat{n}_1^k$.

    ii. $\hat{n}_2^* := \min\{\hat{n}_n^k > \hat{n}_1^* | \tilde{q}_{\hat{n}_n^k}^k / (q_{\hat{n}_n^k}^k)^2 < 1/2\}$ (e x i s t s   s i n c e $\liminf_{n\to\infty} \tilde{q}_{\hat{n}_n^k}^k / (q_{\hat{n}_n^k}^k)^2 = 0$).

    iii.   ...

    iv. $\hat{n}_m^* := \min\{\hat{n}_n^k > \hat{n}_{m-1}^* | \tilde{q}_{\hat{n}_n^k}^k / (q_{\hat{n}_n^k}^k)^m < 1/m\}$ (e x i s t s   s i n c e $\liminf_{n\to\infty} \tilde{q}_{\hat{n}_n^k}^k / (q_{\hat{n}_n^k}^k)^m = 0$).

    v.   ...

Since $\lim_{n\to\infty} q_{\hat{n}_n^k}^k = 0$ and $\lim_{n\to\infty} \tilde{q}_{\hat{n}_n^k}^k / (q_{\hat{n}_n^k}^k)^n = 0$, it is clear that $\lim_{n\to\infty} \tilde{q}_{\hat{n}_n^k}^k / (q_{\hat{n}_n^k}^k)^\gamma = 0$ for all $\gamma \in \mathbb{R}$. We then set $k^* := k$; set $(\tilde{q}_n^{k'})_n := (\tilde{q}_n^k)_n$ and $\hat{\alpha}_*^{k'} := 0$ for all $k' \leq k$; and stop.

(b)   If the previous case fails, and $\liminf_{n\to\infty} (\tilde{q}_{\hat{n}_n^k}^k)^{-1} / (q_{\hat{n}_n^k}^k)^\gamma = 0$ for all $\gamma \in \mathbb{R}$, then set $(\hat{n}_n^*)_n$ to be a subsequence of $(\hat{n}_n^k)_n$ such that $\lim_{n\to\infty} (\tilde{q}_{\hat{n}_n^*}^k)^{-1} / (q_{\hat{n}_n^*}^k)^\gamma = 0$ for all $\gamma \in \mathbb{R}$ (which exists by the previous argument); set $k^* := k$; set $(\tilde{q}_n^{k'})_n := (\tilde{q}_n^k)_n$ and $\hat{\alpha}_*^{k'} := 0$ for all $k' \leq k$; and stop.

(c)   If the previous cases fail and $\liminf_{n\to\infty} q_{\hat{n}_n^k}^k / (\tilde{q}_{\hat{n}_n^k}^k)^\gamma = 0$ for all $\gamma \in \mathbb{R}$, then $(\hat{n}_n^{k-1})_n$ is set to be a subsequence of $(\hat{n}_n^k)_n$ such that $\lim_{n\to\infty} q_{\hat{n}_n^{k-1}}^k / (\tilde{q}_{\hat{n}_n^{k-1}}^k)^\gamma = 0$ for all $\gamma \in \mathbb{R}$ (which again exists by the previous argument), $(\tilde{q}_n^{k-1})_n := (\tilde{q}_n^k)_n$, and $\hat{\alpha}_*^k := 0$; and go to Step 2.

(d)   If the previous cases fail, we proceed as follows. Note that it must be that $\lim_{n\to\infty} \tilde{q}_{\hat{n}_n^k}^k \in \{0, +\infty\}$ (since, otherwise, we would have $\lim_{n\to\infty} q_{\hat{n}_n^k}^k / (\tilde{q}_{\hat{n}_n^k}^k)^\gamma = 0$ for all $\gamma \in \mathbb{R}$ because $q_{\hat{n}_n^k}^k$ tends to 0 as $n \to \infty$, but we assume that part (c) fails). There are then two sub-cases:

    i.   Assume first $\lim_{n\to\infty} \tilde{q}_{\hat{n}_n^k}^k = 0$. Since case (a) fails, there is some $\gamma \in \mathbb{R}_{++}$ such that

$$\liminf_{n\to\infty} \frac{\tilde{q}_{\hat{n}_n^k}^k}{(q_{\hat{n}_n^k}^k)^\gamma} > 0. \tag{A.1}$$

Let γ_1 be the infimum of the set of values γ such that the previous inequality holds. Note that $\gamma_1 \geq 0$ because both $\tilde{q}_{\hat{n}_n^k}^k$ and $q_{\hat{n}_n^k}^k$ tend to 0 as $n \to \infty$. Note that, for all $\gamma > \gamma_1$, the left-hand side of (A.1) is $+\infty$; and, if $\gamma < \gamma_1$, the left-hand side of (A.1) is 0. We then let $(\hat{n}_n^{k-1})_n$ be such that $\tilde{q}_{\hat{n}_n^{k-1}}^k / (q_{\hat{n}_n^{k-1}}^k)^{\gamma_1}$ tends to some limit in $\overline{\mathbb{R}}_+$; $\tilde{q}_n^{k-1} := \tilde{q}_n^k / (q_n^k)^{\gamma_1}$; and $\alpha_*^k := \gamma_1$; and go to Step 2. For future use, we further prove that $\gamma_1 > 0$ (recall that we already argued that $\gamma_1 \geq 0$). To see this, assume for the sake of contradiction that $\gamma_1 = 0$. Then we have that, for all $\gamma > 0$,

$$+\infty = \liminf_{n\to\infty} \frac{\tilde{q}^k_{\hat{n}^k_n}}{(q^k_{\hat{n}^k_n})^\gamma} = \left(\limsup_{n\to\infty} \frac{q^k_{\hat{n}^k_n}}{(\tilde{q}^k_{\hat{n}^k_n})^{1/\gamma}} \right)^{-\gamma}.$$

That is, replacing γ by $1/\gamma'$, we have that for all $\gamma' > 0$

$$0 = \limsup_{n\to\infty} \frac{q^k_{\hat{n}^k_n}}{(\tilde{q}^k_{\hat{n}^k_n})^{\gamma'}} \ge \liminf_{n\to\infty} \frac{q^k_{\hat{n}^k_n}}{(\tilde{q}^k_{\hat{n}^k_n})^{\gamma'}} \ge 0 \ \Rightarrow\ \liminf_{n\to\infty} \frac{q^k_{\hat{n}^k_n}}{(\tilde{q}^k_{\hat{n}^k_n})^{\gamma'}} = 0.$$

Since $\liminf_{n\to\infty} q^k_{\hat{n}^k_n}/(\tilde{q}^k_{\hat{n}^k_n})^{\gamma'} = 0$ for all $\gamma' \le 0$ (because both $q^k_{\hat{n}^k_n}$ and $\tilde{q}^k_{\hat{n}^k_n}$ tend to 0 as $n \to 0$), we reach a contradiction with the assumption that (c) does not hold.

Note finally, also for future use, that using similar arguments and that case (c) fails, we have that there exists some $\gamma_2 > 0$ such that $\liminf_{n\to\infty} q^k_{\hat{n}^k_n}/(\tilde{q}^k_{\hat{n}^k_n})^\gamma = +\infty$ for all $\gamma > \gamma_2$ and $\liminf_{n\to\infty} q^k_{\hat{n}^k_n}/(\tilde{q}^k_{\hat{n}^k_n})^\gamma = 0$ for all $\gamma < \gamma_2$. From the definition of γ_1, we have that for all $\gamma > \gamma_1$ we have that

$$+\infty = \liminf_{n\to\infty} \frac{\tilde{q}^k_{\hat{n}^k_n}}{(q^k_{\hat{n}^k_n})^\gamma} = \left(\limsup_{n\to\infty} \frac{q^k_{\hat{n}^k_n}}{(\tilde{q}^k_{\hat{n}^k_n})^{1/\gamma}} \right)^{-\gamma} \le \left(\liminf_{n\to\infty} \frac{q^k_{\hat{n}^k_n}}{(\tilde{q}^k_{\hat{n}^k_n})^{1/\gamma}} \right)^{-\gamma}.$$

Hence, we have $\gamma > \gamma_1$ implies $1/\gamma < \gamma_2$; that is, $\gamma_1 \ge 1/\gamma_2$. Similarly, if $\gamma < \gamma_1$, we have

$$0 = \liminf_{n\to\infty} \frac{\tilde{q}^k_{\hat{n}^k_n}}{(q^k_{\hat{n}^k_n})^\gamma} = \left(\liminf_{n\to\infty} \frac{q^k_{\hat{n}^k_n}}{(\tilde{q}^k_{\hat{n}^k_n})^{1/\gamma}} \right)^{-\gamma}.$$

Now we have that $\gamma < \gamma_1$ implies $1/\gamma > \gamma_2$, that is, $\gamma_1 \le 1/\gamma_2$. Overall, we have $\gamma_1 = 1/\gamma_2$, that is, $\liminf_{n\to\infty} q^k_{\hat{n}^k_n}/(\tilde{q}^k_{\hat{n}^k_n})^\gamma = +\infty$ for all $\gamma > 1/\gamma_1$ and also $\liminf_{n\to\infty} q^k_{\hat{n}^k_n}/(\tilde{q}^k_{\hat{n}^k_n})^\gamma = 0$ for all $\gamma < 1/\gamma_1$.

ii. Assume now $\lim_{n\to\infty} \tilde{q}^k_{\hat{n}^k_n} = +\infty$. We proceed analogously to case i, where now γ_1 is the *supremum* of the set of values γ such that inequality (A.1) holds. The rest holds equivalently.

Step 2: If $k > 1$, then replace k by $k - 1$ and go to Step 1. If $k = 1$, then set $(\hat{n}^*_n)_n := (\hat{n}^0_n)_n$ and $k^* := 0$, and stop.
(*End of the algorithm, the proof of Lemma 3.1 continues.*)
Part 2: Intermediate results Define $(q^*_n)_n := (\tilde{q}^0_n)_n$. We note first that

$$q^*_n = \tilde{q}^0_n = \frac{\tilde{q}^1_n}{(q^1_n)^{\alpha^1_*}} = \frac{1}{(q^1_n)^{\alpha^1_*}} \frac{\tilde{q}^2_n}{(q^2_n)^{\alpha^2_*}} = ... = \frac{\sigma_n(\hat{a})}{\prod_{k=1}^K (q^k_n)^{\alpha^k_*}}. \tag{A.2}$$

We also note that, for all $k > k^*$, it must be that

$$\lim_{n \to \infty} \frac{q^k_{\hat{n}^*_n}}{(\tilde{q}^{k-1}_{\hat{n}^*_n})^\gamma} = 0 \quad \text{for all } \gamma \in \mathbb{R} \tag{A.3}$$

(recall that $(\hat{n}^*_n)_n$ is a subsequence of $(\hat{n}^k_n)_n$). Indeed, $k > k^*$ means that both cases 1(a) and 1(b) in the algorithm do not hold for k. If case 1(c) holds for k, then it is clear that $\lim_{n \to \infty} q^k_{\hat{n}^*_n}/(\tilde{q}^{k-1}_{\hat{n}^*_n})^\gamma = 0$ for all $\gamma \in \mathbb{R}$. Assume then case 1(d) holds for k, and that $\lim_{n \to \infty} \tilde{q}^k_{\hat{n}^*_n} = 0$ (the case where $\lim_{n \to \infty} \tilde{q}^k_{\hat{n}^*_n} = +\infty$ is analogous). Assume then, for the sake of contradiction, that there is some $\gamma \in \mathbb{R}$ such that

$$0 < \limsup_{n \to \infty} \frac{q^k_{\hat{n}^*_n}}{(\tilde{q}^{k-1}_{\hat{n}^*_n})^\gamma} = \limsup_{n \to \infty} \frac{(q^k_{\hat{n}^*_n})^{1+\gamma\gamma_1}}{(\tilde{q}^k_{\hat{n}^*_n})^\gamma}, \tag{A.4}$$

where we used that $\tilde{q}^{k-1}_n = \tilde{q}^k_n/(q^k_n)^{\gamma_1}$ in case 1(d). If $1 + \gamma\gamma_1 = 0$, then it must be that $\gamma < 0$ (since $\gamma_1 > 0$), but then Eq. (A.4) does not hold because both $q^k_{\hat{n}^*_n}$ and $(\tilde{q}^{k-1}_{\hat{n}^*_n})^{-\gamma}$ tend to 0. Alternatively, if $1 + \gamma\gamma_1 > 0$, then it must be that $\gamma > 0$ for Eq. (A.4) to hold, and hence

$$0 < \limsup_{n \to \infty} \frac{(q^k_{\hat{n}^*_n})^{1+\gamma\gamma_1}}{(\tilde{q}^k_{\hat{n}^*_n})^\gamma} = \left(\liminf_{n \to \infty} \frac{\tilde{q}^k_{\hat{n}^*_n}}{(q^k_{\hat{n}^*_n})^{(1+\gamma\gamma_1)/\gamma}} \right)^{-\gamma}.$$

By the definition of γ_1, this implies that $(1 + \gamma\gamma_1)/\gamma \le \gamma_1$, which is a contradiction. Finally, if $1 + \gamma\gamma_1 < 0$, then it must be that $\gamma < 0$ (since $\gamma_1 > 0$), and hence we have

$$0 < \limsup_{n \to \infty} \frac{(q^k_{\hat{n}^*_n})^{1+\gamma\gamma_1}}{(\tilde{q}^k_{\hat{n}^*_n})^\gamma} = \left(\limsup_{n \to \infty} \frac{\tilde{q}^k_{\hat{n}^*_n}}{(q^k_{\hat{n}^*_n})^{(1+\gamma\gamma_1)/\gamma}} \right)^{-\gamma} = \left(\liminf_{n \to \infty} \frac{q^k_{\hat{n}^*_n}}{(\tilde{q}^k_{\hat{n}^*_n})^{\gamma/(1+\gamma\gamma_1)}} \right)^{1+\gamma\gamma_1}.$$

By the definition of γ_2 (in Step 1(d) of the algorithm), we have that $\gamma/(1 + \gamma\gamma_1) < \gamma_2 = 1/\gamma_1$, which is again a contradiction.

Part 3: Definition of the objects in the statement To conclude the proof of Lemma 3.1 by providing the objects in its statement. Note that there are three possibilities:

1. If $\lim_{n \to \infty} q^*_{\hat{n}^*_n} \in \mathbb{R}_{++}$ (and so $k^* = 0$) then we have that $(\hat{n}^*_n)_n$ and $(q^1_n, \dots, q^K_n)_n$ are the desired sequences. Note that in this case, the values of α, $(\hat{k}_k)^K_{k=1}$, and $(\hat{j}_k)^K_{k=1}$ are unchanged (by (A.2) we have that the second property of the statement holds for $\hat{a}$ for $\alpha(\hat{a}) = \alpha_*$).

2. If $\lim_{n \to \infty} q^*_{\hat{n}^*_n} = 0$ then the desired sequences are $(\hat{n}^*_n)_n$ and

$$(q^1_n, \dots, q^{k^*}_n, q^*_n, q^{k^*+1}_n, \dots, q^K_n)_n.$$

Indeed, in this case, the algorithm ends in one of the following two cases. In the first case, the algorithm ends because $k^* = 0$. Note that since $(q^*_n)_n = (\tilde{q}^0_n)_n$, we have $\lim_{n \to \infty} q^1_{\hat{n}^*_n}/(q^*_{\hat{n}^*_n})^\gamma = 0$ for all $\gamma \in \mathbb{R}$ (by Eq. (A.3) at $k = 1$). In the second case, the algorithm ends at Step 1(a) for some $k^* > 0$. Nevertheless, now, for all $\gamma \in \mathbb{R}$,

$$\lim_{n\to\infty} \frac{q^*_{\hat{n}_n}}{(q^{k^*}_{\hat{n}_n})^\gamma} = 0 \ \text{ and, if } k^* < K, \ \lim_{n\to\infty} \frac{q^{k^*+1}_{\hat{n}_n}}{(q^*_{\hat{n}_n})^\gamma} = 0.$$

The first equality holds from the algorithm ending at Step 1(a) for k^*, while the second equality holds because $(q^*_n)_n = (\tilde{q}^{k^*}_n)_n$ (and using Eq. (A.3) again, now at $k = k^* + 1$). Note that new α, and $(\hat{j}_k)^{K+1}_{k=1}$, and $(\hat{k}_k)^{K+1}_{k=1}$, are given by

$$\begin{cases} (\alpha^1_*(\hat{a}), \ldots, \alpha^{k^*}_*(\hat{a}), 1, \alpha^{k^*+1}_*(\hat{a}), \ldots, \alpha^K_*(\hat{a})) \text{ if } a = \hat{a}, \\ (\alpha^1_*(a), \ldots, \alpha^{k^*}_*(\hat{a}), 0, \alpha^{k^*+1}_*(a), \ldots, \alpha^K_*(a)) \text{ if } a \neq \hat{a}, \end{cases}$$

$(\hat{j}_1, \ldots, \hat{j}_K, J + 1)$, and

$$\begin{cases} \hat{k}_k & \text{if } k \leq k^*, \\ \hat{k}_k + 1 & \text{if } k \in \{k^* + 1, K\}, \\ k^* + 1 & \text{if } k = K + 1, \end{cases}$$

respectively.

3. If $\lim_{n\to\infty}(q^*_{\hat{n}_n})^{-1} = 0$ then, proceeding as in the previous case, it is easy to see that the desired sequences are $(\hat{n}^*_n)_n$ and

$$(q^1_n, \ldots, q^{k^*}_n, (q^*_n)^{-1}, q^{k^*+1}_n, \ldots, q^K_n)_n$$

(and α, and $(\hat{j}_k)^{K+1}_{k=1}$, and $(\hat{k}_k)^{K+1}_{k=1}$ coincide with the previous case except that the new value for $\alpha^{k^*+1}(\hat{a})$ is -1 instead of 1).

Part 4: Uniqueness of α We finally prove that α is unique (for some given $(q^1_n, \ldots, q^K_n)_n$ satisfying the second property of the statement of Lemma 3.1). To do so, assume that there is some $a \in A$ and $\alpha(a), \hat{\alpha}(a) \in \mathbb{R}^{K+1}$ satisfying

$$\lim_{n\to\infty} \frac{\sigma_{\hat{n}_n}(a)}{\prod^K_{k=1}(q^k_{\hat{n}_n})^{\alpha^k(a)}} \in \mathbb{R}_{++} \ \text{ and } \ \lim_{n\to\infty} \frac{\sigma_{\hat{n}_n}(a)}{\prod^K_{k=1}(q^k_{\hat{n}_n})^{\hat{\alpha}^k(a)}} \in \mathbb{R}_{++}.$$

It then follows that

$$\mathbb{R}_{++} \ni \lim_{n\to\infty} \frac{\prod^K_{k=1}(q^k_{\hat{n}_n})^{\alpha^k(a)}}{\prod^K_{k=1}(q^k_{\hat{n}_n})^{\hat{\alpha}^k(a)}} = \lim_{n\to\infty} \prod^K_{k=1}(q^k_{\hat{n}_n})^{\alpha^k(a)-\hat{\alpha}^k(a)}.$$

Let $k^\dagger \in \{1, \ldots, K\}$ be the maximum value k for which $\alpha^k(a) - \hat{\alpha}^k(a) \neq 0$ (which exists because $\alpha(a) \neq \hat{\alpha}(a)$). Since $(q^1_n, \ldots, q^K_n)_n$ satisfies the second property of the statement, we have that

$$\lim_{n\to\infty} \prod^K_{k=1}(q^k_{\hat{n}_n})^{\alpha^k(a)-\hat{\alpha}^k(a)} = \begin{cases} 0 & \text{if } \alpha^{k^\dagger}(a) - \hat{\alpha}^{k^\dagger}(a) > 0, \\ +\infty & \text{if } \alpha^{k^\dagger}(a) - \hat{\alpha}^{k^\dagger}(a) < 0. \end{cases}$$

This is a contradiction; hence, α is unique. $\qquad\square$

Proof of Corollary 4.1

Proof The result follows from Eq. (3.1) and the fact that $\lim_{n\to\infty} M_n(a) = 1$ for all a such that $y(a) = 0$. $\qquad\square$

Proof of Proposition 4.1

Proof The result follows from the arguments following the statement in the text. $\quad\square$

Proof of Proposition 4.2

Proof Like in Theorem 3.1, the "if" part of the statement is obvious. To prove the "only if" part, fix a consistent assessment (σ, μ), and let $(x, y) : A \to \mathbb{R}_{++} \times \mathbb{R}_+$ be a pair such that $(\sigma_n)_n$ defined in Eq. (3.1) supports (σ, μ) (which exists by Theorem 3.1). We will now find another pair $(\hat{x}, \hat{y}) : A \to \mathbb{R}_{++} \times \mathbb{Z}_+$ such that $\hat{x} = x$ and $\hat{y}(h) \geq \hat{y}(h')$ if and only if $y(h) \geq y(h')$, hence $(\hat{x}, \hat{y})$ will also support (σ, μ).

Fix some $\varepsilon > 0$ to be pinned down later in the proof. We define $A_0 := \{a | \sigma(a) = 0\}$. We denote the actions in A_0 as $\{a_1, \ldots, a_{J_0}\}$. We define $\hat{y}_\varepsilon : A \to \mathbb{Q}$ recursively as follows. Define $\mathcal{J}^0 := \emptyset$. Then, letting j run from 1 until J_0, proceed recursively as follows:

1. If $y(a_j)$ is a rational combination of elements in $\{y(a_{j'}) | j' \in \mathcal{J}^{j-1}\}$, given by $y(a_j) = \sum_{j' \in \mathcal{J}^{j-1}} \gamma_\varepsilon^{jj'} y(a_{j'})$ (where $\gamma_\varepsilon^{jj'} \in \mathbb{Q}$ for all $j' \in \mathcal{J}^{j-1}$), then set $\mathcal{J}^j := \mathcal{J}^{j-1}$ and $\hat{y}_\varepsilon(a_j) := \sum_{j' \in \mathcal{J}^{j-1}} \gamma_\varepsilon^{jj'} \hat{y}_\varepsilon(a_{j'}) \in \mathbb{Q}$.
2. Otherwise, if $y(a_j)$ is a not rational combination of elements in $\{y(a_{j'}) | j' \in \mathcal{J}^{j-1}\}$, set $\hat{y}_\varepsilon(a_j)$ to be a rational number in $(y(a_j) - \varepsilon, y(a_j) + \varepsilon)$ and set $\mathcal{J}^j := \mathcal{J}^{j-1} \cup \{j\}$.

Let $\mathcal{J}$ denote $\mathcal{J}^{J_0}$. Note that, by construction, $\hat{y}_\varepsilon(a) \in \mathbb{Q}$ for all $a \in A_0$. Note also that if a number $\tilde{y} \in \mathbb{R}$ is a rational combination of $\{y(a_j) | j = 1, \ldots, J_0\}$, then there exists a *unique* set of rational coefficients $\{\gamma_\varepsilon^{\tilde{y},j} | j \in \mathcal{J}\}$ such that $\tilde{y} = \sum_{j \in \mathcal{J}} \gamma_\varepsilon^{\tilde{y},j} y(a_j)$. Indeed, by construction, for each $j \in \mathcal{J}$, $y(a_j)$ cannot be written as a rational combination of $\{y(a_{j'}) | j' \in \mathcal{J} \setminus \{j\}\}$, while for all $j \notin \mathcal{J}$ we have that $y(a_j)$ can be written uniquely as a rational combination of elements in $\{y(a_j) | j \in \mathcal{J}\}$.

Note that there exists some $\bar{y} \in \mathbb{R}_{++}$ such that, for all $\varepsilon > 0$ and $\hat{y}_\varepsilon$ obtained from the previous algorithm, we have $|\hat{y}_\varepsilon(a) - y(a)| < \bar{y}\varepsilon$ for all $a \in A$. This is obviously true for all a_j with $j \in \mathcal{J}$, since $\hat{y}_\varepsilon(a_j) \in (y(a_j) - \varepsilon, y(a_j) + \varepsilon)$ by construction. The result then follows because, when a_j is such that $j \notin \mathcal{J}$, $\hat{y}_\varepsilon(a_j)$ is a linear combination of elements in $\{\hat{y}_\varepsilon(a_{j'}) | j' \in \mathcal{J}\}$ with coefficients that are independent of ε.

From the previous paragraphs, we have that if $h, h' \in H$ are such that $y(h) > y(h')$, then $\hat{y}_\varepsilon(h) > \hat{y}_\varepsilon(h')$ if ε is small enough. Finally, assume that $h, h' \in H$ are such that

$y(h) = y(h')$. Let $\{\gamma_\varepsilon^{h,j}|j \in \mathcal{J}\}$ be the unique set of rational coefficients such that $y(h) = \sum_{j \in \mathcal{J}} \gamma_\varepsilon^{h,j} y(a_j)$. Then, it is clear that

$$\hat{y}_\varepsilon(h) = \sum_{j \in \mathcal{J}} \gamma_\varepsilon^{h,j} \hat{y}_\varepsilon(a_j) = \sum_{j \in \mathcal{J}} \gamma_\varepsilon^{h',j} \hat{y}_\varepsilon(a_j) = \hat{y}_\varepsilon(h').$$

Note finally that $\hat{y}_\varepsilon(a) = 0$ if and only if $y(a) = 0$, hence $\sum_{a \in A^I} \mathbb{I}_{\hat{y}_\varepsilon(a)=0} x(a) = 1$ for all information sets $I \in \mathcal{I}$. Therefore, a strategy defined as in (3.1) with $(x, \hat{y}_\varepsilon)$ exists and supports (σ, μ), so the proof is concluded.1 □

References

Bagwell K (1990) Informational product differentiation as a barrier to entry. Int J Ind Organ 8:207–223

Blume LE, Zame WR (1994) The algebraic geometry of perfect and sequential equilibrium. Econometrica J Econom Soc 62:783–794

Cramton PC (1985) Sequential bargaining mechanisms. Cambridge University Press, Cambridge

Dilmé F (2023a) Lexicographic numbers in extensive form games. Mimeo

Dilmé F (2023b) Sequentially stable outcomes. Mimeo

Fudenberg D, Tirole J (1991) Perfect Bayesian equilibrium and sequential equilibrium. J Econ Theory 53:236–260

Harrington JE (1993) The impact of reelection pressures on the fulfillment of campaign promises. Games Econ Behav 5:71–97

Kreps DM, Wilson R (1982) Sequential equilibria. Econometrica J Econom Soc 50:863–894

Myerson RB (1978) Refinements of the Nash equilibrium concept. Int J Game Theory 7:73–80

Osborne MJ, Rubinstein A (1994) A course in game theory. MIT Press, Cambridge

Rubinstein A (1985) A bargaining model with incomplete information about time preferences. Econometrica J Econom Soc 53:1151–1172

Selten R (1965) Spieltheoretische Behandlung eines Oligopolmodells mit Nachfrageträgheit: Teil I: Bestimmung des dynamischen Preisgleichgewichts. Zeitschrift für die gesamte Staatswissenschaft/J Inst Theor Econ 121:301–324

Selten R (1975) Reexamination of the perfectness concept for equilibrium points in extensive games. Int J Game Theory 4:25–55

van Damme E (1984) A relation between perfect equilibria in extensive form games and proper equilibria in normal form games. Int J Game Theory 13:1–13

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.