LEARNING AND STABILITY IN BIG UNCERTAIN GAMES

EHUD KALAI AND ERAN SHMAYA

Abstract. In a big uncertain game a stage game is played repeatedly by a large anonymous population. Players' privately known types are correlated through an unknown state of fundamentals and the game is played with imperfect monitoring. Under simple behavioral assumptions, the game admits Markov-perfect equilibria. We show that with time, equilibrium play in these games becomes highly predictable and stable, if unceraínty that is not explained by fundamentals is small.

1. Overview of big uncertain games

Strategic interaction in large populations is a subject of interest in economics, political science, computer science, biology, and more. Indeed, the last three decades have produced a substantial literature that deals with large strategic games in a variety of applications, such as markets [31], bargaining [25], auctions [32, 33], voting [10, 29], electronic commerce [14], and market design [3, 6].

This paper focuses on games that are *big* in two senses. First, they are played by a *large* anonymous population of players, as in many of the current models of large

Date: May 4, 2016.

Key words and phrases. Anonymous games, Nash equilibrium, Repeated games, Large games, Bayesian equilibrium, Markov equilibrium, Price taking, Rational expectations.

Acknowledgment: Early versions of the paper benefitted from presentations at Tel Aviv University, The Hebrew University, Northwestern University, University of California San Diego, Microsoft Research New England, Harvard University, SUNY at Stony Brook, Duke University, Stanford University, University of Pennsilvania, California Institute of Technology, University of California Berkeley, Oxford University and more. Parts of the paper were presented in the 2015 Nancy L Schwartz Memorial Lecture at Northwestern University.

games. But in addition, they are played *repeatedly* with no end in sight. Moreover, these big games are subject to fundamental uncertainty for three reasons: (1) There is an unknown state of fundamentals of the game. (2) There are privately known player types with probability distributions that depend on this unknown state; thus, the player types are statistically correlated. (3) Period outcomes are determined randomly, based on the fundamental state and the players' actions.

Such big games are common. For example, they include repeated production and consumption games, in which period outcomes are prices; repeated rush-hour-commute games, in which outcomes are driving times; and network-communication games, in which outcomes are population distributions of communication devices. But the complexity, due to repetition and uncertainty, makes Bayesian analysis of these games difficult.

The main body of the paper provides a formal explanation of a phenomenon often observed in a big games that are subject only to low levels of unexplained uncertainty: except for a limited proportion of *learning periods,* equilibrium play becomes highly *predictable* and *stable.*

To motivate the notions of predictability and stability, consider repeated *rush-hour-commute* games, in a transportation system that has gone through subtantial changes, e.g., roads or train lines were added or updated. The players (and even the designers of the system) may be unsure about the fundamentals of the new systems (e.g., capacity and safety of the roads at various levels of congestion), and about the individual player information and preferences over the use of roads.

To say that a morning rush-hour commute is *highly predictable* means that there is near certainty that all the commuters' assessments of the driving times, made prior to the start of the rush hour, turn out to be quite accurate, say, within a few minutes of the actually realized rush-hour driving times. *High stability* in such a period is meant in a *hindsight* sense, sometimes referred to as a *no regret*, or *ex-post* Nash. It means that there is near certainty that once everybody starts driving and becomes informed of the realized morning driving times, perhaps from a radio report, no player may recognize a gain of more than a few minutes by switching to a route different from the one she had chosen. Notice that lack of stability implies (potential) *chaos*: some players may choose to deviate from their optimally chosen routes, their deviations may lead other drivers to deviate from their chosen routes, and so forth.[1] As one may anticipate, at equilibrium a high level of predictability implies a high level of stability.

In price formation games, where period prices are determined by period consumption and production decisions, hindsight stability means that the prices are competitive and reflect rational expectations. In the concluding section of the paper, we study in detail a big repeated Cournot game. There, the state of fundamentals is the difficulty of producing a new product and the types describe privately known production costs of the individual producers, which are correlated through the unknown state of fundamentals.

In these Cournot games, the general findings presented in this paper make it is easy to construct and play the following unique equilibria: Every period belongs to one of three possible regimes that depend on the ease of production perceived prior to the start of the period: (1) "perceived-easy," (2) "perceived-difficult," or (3) "not

[1]This paper does not model (1) the behavior of the players within a chaotic period, or (2) the possible effect of the chaotic behavior on the play in future periods. In other words, in their future play the players will take into consideration the initially reported driving times and not the realized chaotic driving times that followed later in the morning.

determined." In the periods in regime 1, all the inefficient producers stay out and the efficeint producers produce according to a mixed strategy. In the periods in regime 2, all the efficient producers produce at full capacity and the inefficient ones produce according to a mixed strategy. In the periods in regime 3, both efficient and inefficient producers follow mixed strategies. Roughly speaking, in any state of fundamentals, in regimes 1 and 2 the outcomes are highly predictable and stable, while regime 3 consists of learning periods.

In the examples above and in other big games, period outcomes may be unpredictable due to randomness that cannot be explained by the fundamentals of the game. For example, an unanticipated traffic accident in a morning rush hour may throw off rational predictions of the morning driving times. Similarly, the presence of unanticipated noise traders on a given day may cause significant deviation from the anticipated day prices. Such randomness interferes with the players' ability to predict the outcomes of periods, even in the absence of fundamental uncertainty.

The main findings of the paper include bounds on (1) the finite number of unpredictable chaotic learning periods, K, during which the players learn all they can about the uncertain fundamentals; and (2) the level of stability possible in the nonlearning periods, and its dependence on the level of unexplained uncertainty in period outcomes.

To argue for the existence of the finite number of learning periods, K, we extend results from the rational learning literature: in n-person Bayesian repeated games, players learn to *forecast probabilities* of period outcomes in all but K periods. To argue for predictability in the remaining nonlearning periods, we extend results from the literature on large one-shot games: when the number of players is large, forcasting probabilities of period outcomes can be replaced by learning to *predict* the realized outcomes and not just their probabilities. By generalizing and combining the results above, under the behavioral assumptions of this paper, we obtain highly accurate predictions (up to the level permitted by the unexplained randomness of the outcomes) in all but the K learning periods.

It is important to note that our players are "rational learners," in that their objective is to maximize their expected payoff in the play of Bayesian equilibrium. "Irrational learners," such as followers of naive best-reply dynamics, may never converge to stability. Consider, for example, a repeated rush-hour game with two identical parallel routes, A and B, both from starting point S to destination T, and drivers who wish to travel on the least congested of the two routes. If on day one most drivers use A, then the best-reply dynamics would have them all alternate between $B, A, B, A, ...$ from the second day on, and in no period is the equilibrium driving pattern stable.

The Bayesian analysis in this paper is greatly simplified under the behavioral assumption of *imagined-continuum* reasoning, which means that the players replace random variables of the population by their expected values. For example, in the two-route commute game above, if a player believes that each of the commuters independently and randomly chooses A or B with probabilities $2/3$ and $1/3$, respectively, then the player would assume that *with certainty* $2/3$ of the population ends up choosing A and $1/3$ choosing B. This assumption is made throughout the game whenever random choices are made independently, even when conditioned on past events. Moreover, in doing such “compression” to (conditional) expected values, players ignore their own effect on population proportions. For example, a player who believes that the population is divided 2/3 to 1/3 between A and B would hold this belief no matter which route she herself chooses.[2]

The imagined-continuum view is natural when players wish to simplify highly complex computations. In our big games, many of the Bayesian calculations required of rational players are eliminated. But since the continuum is only imagined, as game theorists we are careful to compute probabilities in the actual process, the one in which n continuum-imagining players use best-response strategies to generate events in the repeated game. After all, whether a player would want to deviate from a route that she has chosen depends on the actual observed driving times determined by the n drivers on the road and not by the hypothetical continuum that she or the other drivers imagine.

The discrepancy between the continuum game and the n- player imagined continuum game is important when we study repeated games, due to a *discontinuity in probability computations.* For example, in repeated games there are events that have probability one in the continuum game, but probability zero in the n-player imagined-continuum games, regardless of the number of players n (see our companion paper [22]).

An important consequence of the imagined-continuum assumption is the existence of myopic Markov-perfect Nash equilibria. Focusing on these equilibria in this paper simplifies the presentation of the results without much loss of generality, since (1) these equilibria seem to be focal in the minds of rational players, and (2) most of

[2]See, for example, Mcafee [27] for an earlier use of this idea. In his paper a sellers offer competing mechanisms to buyers. While the analysis is performed on a finite set of sellers, these sellers neglect their impact on the utility of buyers who don't participate in their mechanism.

the observations discussed in this paper are true for Nash equilibria, without these restrictions (see our companion paper [22]).

The findings of the current paper, presented in Part 1, may be used as an initial building block toward a more general theory of big games, in which the fundamentals change during the play of the game. This is left for future research, but we present preliminary observations, that point to the emergence of *stability cycles*, in Part 2. Part 3 presents the proofs used for the observations discussed in Part 1.

2. Relationship to Earlier literature

Kalai [19] and following papers [8, 3, 9, 13] demonstrate that (hindsight) stability is obtained in large one-shot Bayesian games with independent player types. These papers discuss examples from economics, politics, and computer science in which stability is highly desirable. In market games in particular, stability implies that Nash equilibrium prices are competitive.[3]

The interdependency of player types is a fundamental difficulty when dealing with stability, as illustrated by following example from Kalai [19]: There is an unknown state of fundamentals, $s = A$ or $s = B$, about which every player i is given a private 70 percent accurate signal g_i ($\Pr(g_i = s) = 0.7$ and $\Pr(g_i = s^c) = 0.3$), independtently of the signals of the others. Trying to match the state, every player chooses an action, $a_i = A$ or $a_i = B$, and is paid 1 if $a_i = s$, but 0 if $a_i = s^c$. Assuming that every player follows her dominant strategy ($a_i = g_i$) and that the

[3]Kalai [19] also demonstrates that hindsight stability implies a stronger stability property, referred to as "structural robustness," and Deb and Kalai [9] allow a continuum of actions and heterogeneous types of players. The current paper does not deal with either of these features. See also Azrieli and Shmaya [4], and Babichenko [5], both of which follow Kalai's paper, but take it in different directions.

number of players is large, in hindsight, approximately 30 percent of the players would like to change their actions .

The example above explains why Kalai [19] had to be restricted to independent types. This restriction, which disallows correlation of types through unknown fundamentals, is a severe limitation in economic applicability of the paper.

By extending the model to repeated games, the current paper expands the applicability in two ways. First is the obvious point that many large games are played repeatedly. But more subtle is the observation made through the main findings of the paper: with time, players "learn to be independent" and predictability is obtained.

One-shot large games with complete information were introduced by Schmeidler [35], who showed that with a continuum of players such games admit pure-strategy equilibria.[4] The more recent literature of mean field games (see the survey in Lions' lecture notes [16]), may be viewed as a continuous-time dynamic extension of Schmeidler's model that allows for mixed strategies and randomly changing states. The continuum of players allows the modelers to cancel aggregate uncertainties, in a manner similar to the imagined-continuum model of this paper.[5] But the important differences with the current paper exist: (1) In the imagined-continuum model there are n players who follow the continuum reasoning, whereas in the mean field model, as in Schmeidler's paper, there is a continuum of players. The discontinuity in the computation of probabilities, discussed earlier, limits the relevance of mean field

[4] See [23] for citations of articles that follow up on Schmeidler's paper [35]. See Sorin and Wan [37] for a hybrid model that allows atomic and non-atomic players.

[5] See also papers by Judd [18] and Sun [38] for exact law of large numbers, that formalizes the idea that there is no aggregate uncertainty when a continuum of players randomize independently.

games to repeated games with a large *but finite* number of players. For elaboration, see our companion paper [22].

(2) In mean field games the state of fundamentals changes with time. But this changing state is *publicly observed*, as compared with this paper in which the fixed state of fundamentals is *unknown* and has to be learned from the play. With additionly complex notations, this paper can also accomodate *publicly observed* changing states.[6]

(3) More generally, the mean field approach deals with uncertainties that are "eliminated in the average." This is not the case when we deal with an unknown state of fundamentals. The issues of Bayesian updating and learning of the fundamentals, which are central to this paper, have no analogues in the current literature on mean field games.

A pioneering paper on large repeated games is Green [15], which studies large repeated strategic interaction under restrictions to complete information and to pure strategies. Green and Sabourian [34] derive conditions under which the Nash correspondence is continuous, in the sense that the equilibrium of the continuum game is a limit of the equilibria of the standard n-player games with increasing n (as opposed to the n-player imagined continuum games). In addition to Green's paper, the myopic property of large repeated games has been studied in Al-Najjar and Smorodinsky [1].

In regard to incomplete information, the "learning to predict" theorem presented in this paper relies on earlier results from the literature on rational learning in repeated games, such as Fudenberg and Levine [12], Kalai and Lehrer [21, 20], and Sorin [36].

[6]The concluding part of this paper discusses preliminary results on big games with *unobserved* changing fundamentals.

An early reference to Markov-perfect equilibria is Maskin and Tirole's paper [26]. The body of the current paper elaborates on some properties of the imagined-continuum Markov equilibrium, enough for us to present the results on predictions and hindsight stability. Our companion paper [22], henceforth KS, studies this type of equilibrium in depth and presents results on how the n-player imagined-continuum Markov equilibrium offers a good asymptotic approximation to the Nash equilibrium of the standard n-player game. The current paper uses the same model as in KS, described in Section 3.

As mentioned above, compressing computations to expected values is used in a variety of current models in economics. See, for example, McCaffee [27], Angeletos et al. [2], and Jehiel and Koessler [17], all of whom study a dynamic global game with fundamental uncertainty. Lykouris at el. [24] introduced stability cycles, in a much different environment than the one discussed in our concluding section.

Part 1. Big uncertain games with fixed unknown fundamentals

3. The model

A stage game is played repeatedly in an environment with an unknown fixed state of fundamentals, s (also referred to as a state of nature), by a population of n players whose fixed, privately known types t^i are statistically correlated through the state s. The environment and the game are symmetric and anonymous.

We consider first the *game skeleton* that consists of all the primitives of the game other than the number of players. Augmenting the game skeleton with a number of players n results in a fully specified Bayesian repeated game. This organization eases the presentation of asymptotic analysis, as one can keep all the primitives of the game fixed while varying only the number of players n.

Definition 1. [Game] A *game skeleton* is given by $\Gamma = (S, \theta_0, T, \tau, A, X, \chi, u)$ with the following interpretation:

- S is a finite set of possible *states of nature*; $\theta_0 \in \Delta(S)$ is an initial *prior* probability distribution over S.[7]
- T is a finite set of possible player *types*. The function $\tau : S \to \Delta(T)$ is a stochastic *type-generating function* used initially to establish types. Conditional on the state s, $\tau_s(t)$ is the probability that a player is of type t, and it is (conditionally) independent of the types of the opponents. The selected types, like the initially drawn state of fundamentals, remain fixed throughout the repeated game.
- A is a finite set of possible player's *actions*, available to a player in every period.
- X is a countable set of *outcomes*, and for every $s \in S$ and every $e \in \Delta(T \times A)$, $\chi_{s,e}$ is a probability density function over X. In every period, $e(t, a)$ is the empirical proportion of players in the population who are of type t and who choose the action a. $\chi_{s,e}(x)$ is the probability of the outcome x being realized and announced at the end of the period. We assume that the function $e \mapsto \chi_{s,e}(x)$ is continuous for every s and x.
- $u : T \times A \times X \to [0, 1]$ is a function that describes the player's *payoff*: $u(t, a, x)$ is the period payoff of a player of type t who plays a when the announced period outcome is x.

[7] As usual, $\Delta(B)$ is the set of all probability distributions over the set B.

Remark 1. All the definitions and results hold also when X is a subset of Euclidean space, and $\chi_{s,e}$ is a density function. In this case we have to assume that the function $e \mapsto \chi_{s,e}$ is continuous when the range is equipped with the L^1-norm and that the payoff function u is a Borel function. See an example in Section 6. To simplify notations we assume that X is countable but our proofs do not rely on this assumption. We also identify density functions with probability distribution, so for density function χ over X we let $\chi(B)$ be the probability of $B \subseteq X$. This abuse of notations is common in the case of discrete X but not in the case of continuous distributions.

Example 1 (Repeated computer-choice game with correlated types)**.** As in the example of the one-shot computer-choice game from Kalai [19], let $S = T = A = \{\mathcal{PC}, \mathcal{M}\}$ denote two possible *states* of nature, two possible *types* of players, and two possible *actions* to select. But now these selections are made repeatedly in discrete time periods $k = 0, 1, 2, ...$

Initially, an unknown state s is chosen randomly with equal probabilities, $\theta_0(s = \mathcal{PC}) = \theta_0(s = \mathcal{M}) = 1/2$; and conditional on the realized state s, the fixed types of the n players are drawn by an independent identical distribution: $\tau_s(t^i = s) = 0.7$ and $\tau_s(t^i = s^c) = 0.3$, where s^c is the unrealized state. Each player is privately informed of her type t^i. Both s and the vector of t^is remain fixed throughout the repeated game.

Based on player i's information at the beginning of each period $k = 0, 1, ...,$ she selects one of the two computers, $a^i_k = \mathcal{PC}$ or $a^i_k = \mathcal{M}$. These selections determine the empirical distribution of type-action pairs, e_k, where $e_k(t, a)$ is the proportion of players who are of type t and who choose the computer a in the kth period.

At the end of each period, a random sample (with replacement) of J players is selected, and the sample proportion of $\mathcal{PC}$ users $x = x_k(\mathcal{PC})$ is publicly announced ($x_k(\mathcal{M}) \equiv 1 - x$). Thus, the probability of the outcome $x = y/J$ being selected (when the state is s and the period's empirical distribution is e) is determined by a binomial probability of having y successes in J tries, with a probability of success $e_k(\mathcal{PC}, \mathcal{PC}) + e_k(\mathcal{M}, \mathcal{PC})$.

Player i's payoff in period k is given by

$$u_k^i(t^i, a_k^i, x_k) = (x_k[a_k^i])^{1/3} + 0.2\delta_{a_k^i = t^i},$$

where $\delta_{a_k^i = t^i}$ is 1 if she choose her type and 0 otherwise. The game is infinitely repeated, and a player's overall payoff is the discounted sum of her period payoffs.

3.1. **Bayesian Markov strategies.** We study a symmetric equilibrium in which all the players use the same strategy κ. Normally, a player's strategy in the repeated game specifies a probability distribution by which the player selects an action in every period, as a function of (i) her type, (ii) the observed history of past publicly announced outcomes, and (iii) her own past actions. However, we are interested only in a certain class of strategies: "Bayesian Markov strategies" (or Markov strategies, for short). When playing a Markov strategy, the player does not condition her selection on her own past actions. Moreover, her selection of an action depends on the past publicly announced outcomes only through a Markovian state, which is the posterior public beliefs over the state of nature.

Definition 2. A *(Bayesian) Markov strategy* is a function $\kappa : \Delta(S) \times T \to \Delta(A)$.

The interpretation is that $\kappa_{\theta,t}(a)$ is the probability that a player of type $t \in T$ selects the action $a \in A$, in periods in which the 'public belief' about the state of nature is θ. The term 'public belief' is in quotes because these beliefs are not derived from correct Bayesian reasoning. Instead, they are derived by the imagined-continuum reasoning described in Section 3.2.

Notice that as defined, a Markov strategy κ may be used by any player regardless of the number of opponents and the repetition-payoff structure.

3.2. **Beliefs in the imagined-continuum model.** By the 'public belief' at the beginning of period k, we mean the belief over the state of nature held by an outside observer who (i) knows the players' strategy κ, and (ii) has observed the past publicly announced outcomes of the game. A main feature of our definition of Markov strategies and equilibrium is that these beliefs differ from the correct posterior conditional distributions over the state of nature: They are updated during the game using what we call "imagined-continuum reasoning." Under imagined-continuum reasoning, all uncertainty about players' types and actions conditioned on the state of nature is compressed to its conditional expectations, resulting in known, deterministic, conditional distributions. Specifically, the *public beliefs* are defined recursively by the following process:

- The initial public belief is that the probability of every state s is $\theta_0(s)$.
- In every period that starts with a public belief θ, the *imagined empirical proportion* of a type-action pair (t, a) in the population is

$$d_\theta(t, a) \equiv \tau_s(t) \cdot \kappa_{\theta,t}(a), \tag{3.1}$$

and the posterior public belief assigned to every state s is computed by Bayes rule to be

$$\beta_{\theta,x}(s) \equiv \frac{\theta(s) \cdot \chi_{s,d_\theta}(x)}{\sum_{s' \in S} \theta(s') \cdot \chi_{s',d_\theta}(x)}. \tag{3.2}$$

Like the public observer each player ignores the impact of her type and actions on period outcomes. In addition, she has additional information for assessing probabilities of states of nature, namely, the realization of her own type. Under imagined-continuum reasoning, her type and the public outcome are conditionally independent of each other for any given state of nature. This implies that we can use Bayes formula to compute her *private belief* about the state of nature from the public belief.

Formally, in every period that starts with the public belief θ, for any player of type t the private belief probability assigned to the state of nature s is

$$\theta^{(t)}(s) \equiv \frac{\theta(s) \cdot \tau_s(t)}{\sum_{s' \in S} \theta(s') \cdot \tau_{s'}(t)}. \tag{3.3}$$

3.3. **Markov-perfect equilibrium.** We are now in a position to define the equilibrium concept used in this paper. Under the imagined-continuum view, the players ignore the impact of their own action on the outcome, and a player of type t believes the outcome is drawn from the distribution $\phi(\theta^{(t)}, \theta)$ where $\theta^{(t)}$ is given by (3.3) and $\phi : \Delta(S) \times \Delta(S) \to \Delta(X)$ is given by

$$\phi(\mu, \theta) = \sum_{s \in S} \mu(s) \chi_{s,d_\theta}, \tag{3.4}$$

where d_θ is given by (3.1). Thus, $\phi(\mu, \theta)$ is the *forecast* about the period outcome of an observer whose belief about the state of nature is μ when the public belief about the state of nature is θ.

Definition 3. A *(symmetric, imagined-continuum) Markov (perfect) equilibrium* is given by a Markov strategy $\kappa : \Delta(S) \times T \to \Delta(A)$, such that

$$[\kappa(\theta, t)] \subseteq \mathrm{argmax}_a \sum_{x \in X} \phi(\theta^{(t)}, \theta)(x) u(t, a, x)$$

for every public belief $\theta \in \Delta(S)$ about the state of nature and every type $t \in T$, where $[\kappa(\theta, t)]$ is the support of $\kappa(\theta, t)$, the private belief $\theta^{(t)}$ is given by (3.3), and ϕ is given by (3.4).

According to the imagined-continuum equilibria, each player of type t treats the public outcome as a random variable with distribution $\phi(\theta^{(t)}, \theta)$, ignoring her impact on the outcome. This is a generalization of the economic 'price-taking' property in Green [15] to a *stochastic* setting and to applications other than market games. For this reason our players may be viewed as *stochastic outcome takers*. Note that imagined-continuum equilibria are, by definition, myopic: at every period the players play an imagined-continuum equilibrium in the one-shot Bayesian game for that period.

Remark 2. In our companion paper [22] we define the notion of imagined-continuum equilibrium more generally (without assuming the Markov property and myopia) and prove that (i) every imagined-continuum equilibrium is myopic, (ii) probabilities of certain outcomes computed in the imagined game approximate the real probabilities computed in the standard finite large versions of the game, and (iii) best-responses (and Nash equilibria) in the imagined game are uniformly ϵ best-responses (and ϵ Nash equilibria) for all sufficiently large finite standard versions of the game.

Notice also that under myopia, the equilibrium that we study and the main results that follow are applicable for a variety of repetition and payoff structures. For example, the game may be repeated for a finite number of periods with overall payoffs assessed by the average period payoff, or the game may be infinitely repeated with payoffs discounted by different discount parameters by different players.

Another consequence of myopia is that the set of players may change and include combinations of long-lived players, short-lived players and overlapping generations, provided that (i) the death and birth process keeps the size of the population large, (ii) that process does not alter the state and the players' type distribution, and (iii) that players of a new generation are informed of the latest public belief about the unknown state.[8]

3.4. **The induced play path.** To compute the probability of actual events in the game, we need to describe the actual probability distribution induced over play paths when players follow a Markov strategy κ (as opposed to the beliefs that are derived from the imagined-continuum reasoning).

We use boldface letters to denote random variables that assume values from corresponding sets. For example, $\mathbf{S}$ is the random variable that describes a randomly-selected state from the set of possible states S. We use superscripts to denote players' names, superscripts in parenthesis to denote players' types and subscripts to denote periods' numbers.

[8]Games in which the number of players is large and unknown and follows a Poisson distribution have been studied in Myerson [28]. By restricting ourselves to games of proportions, lack of knowledge of the number of players becomes a trivial issue in the current paper.

The definition below is applicable to a game with a set N of n players with any repetition-payoff specification. As already stated, all the players use the same strategy κ.

Definition 4. Let κ be a Markov strategy of the finite game with n players. The *random κ-play-path* is a collection $\left(\mathbf{S}, \mathbf{T}^i, \mathbf{A}^i_k, \mathbf{X}_k\right)_{i \in N, k=0,1,\ldots}$ of random variables[9], representing the state of nature, types, actions, and outcomes, such that:

- The state of nature $\mathbf{S}$ is distributed according to θ_0.
- Conditional on $\mathbf{S}$, the players types $\mathbf{T}^i$ are independent and identically distributed with the distribution $\tau_{\mathbf{S}}$.
- Conditional on the history of periods $0, \ldots, k-1$, players choose period k actions $\mathbf{A}^i_k$ independently of each other. Every player $i \in N$ uses the distribution $\kappa_{\mathbf{T}^i,\ \mathbf{\Theta}_k}$, where $\mathbf{\Theta}_k$ is the public belief at the beginning of period k, given by

$$\begin{aligned} &\mathbf{\Theta}_0 = \theta_0, \text{ and} \\ &\mathbf{\Theta}_{k+1} = \beta_{\mathbf{\Theta}_k, \mathbf{X}_k}, \text{ for } k \geq 0, \end{aligned} \tag{3.5}$$

  and β is defined in (3.2).
- The outcome $\mathbf{X}_k$ of period k is distributed according to $\chi_{\mathbf{S}, \mathbf{e}_k}$, where

$$\mathbf{e}_k(t, a) = \#\{i \in N | \mathbf{T}^i = t, \mathbf{A}^i_k = a\}/n \tag{3.6}$$

  is the (random) *empirical type-action distribution* in period k.

[9]We do not specify the probability space or the domain of these variables, but only the probability distribution over their values. The play path is unique in distribution.

In equations:

$$
\begin{aligned}
&\mathbb{P}\left(\mathbf{S}=s, \mathbf{T}^i=t^i \ i\in N\right) = \theta_0(s)\cdot\prod_{i\in N}\tau_s(t^i);\\
&\mathbb{P}\left(\mathbf{A}^i_k=a^i \ i\in N \,\middle|\, \mathbf{S},\mathbf{T}^i,\mathbf{A}^i_l,\mathbf{X}_l \ l<k, i\in N\right) = \prod_{i\in N}\kappa_{\mathbf{\Theta}_k,\mathbf{T}^i}(a^i);\\
&\mathbb{P}\left(\mathbf{X}_k=x \,\middle|\, \mathbf{S},\mathbf{T}^i,\mathbf{A}^i_l \ l\le k, i\in N, \mathbf{X}_0,\dots,\mathbf{X}_{k-1}\right) = \chi_{\mathbf{S},\mathbf{e}_k}(x),
\end{aligned}
\tag{3.7}
$$

where $\mathbf{e}_k$ is given by (3.6), and $\mathbf{\Theta}_k$ is given by (3.5).

Note that the imagined-continuum reasoning enters our definition only through (3.5), which reflects the way that the outside observer and the players process information. The assumption of imagined-continuum reasoning lies behind the simple form of the public beliefs process $\mathbf{\Theta}_0, \mathbf{\Theta}_1, \dots$. Two important properties are a consequence of this definition: (i) $\mathbf{\Theta}_k$ admits a recursive formula (i.e., the outside observer and the players need keep track of only their current belief about state of nature and not their beliefs about players' types and actions), and (ii) this formula does not depend on the number of players. Both these properties do not hold for the beliefs $\mathbb{P}(\mathbf{S}\in\cdot\,|\mathbf{X}_0,\dots,\mathbf{X}_{k-1})$ of the game theorists who do the correct Bayesian reasoning.

4. Correct predictions

Consider a game skeleton Γ played repeatedly by n players. Let κ be a Markov strategy and consider the random κ play path $\left(\mathbf{S},\mathbf{T}^i,\mathbf{A}^i_k,\mathbf{X}_k\right)_{i\in N, k=0,1,\dots}$.

Recall that we denote by $\mathbf{\Theta}_k$ the public belief about the state of nature at the beginning of period k, given by (3.5). For every type $t\in T$, let

$$
\mathbf{\Theta}^{(t)}_k(s) = \frac{\mathbf{\Theta}(s)\cdot\tau_s(t)}{\sum_{s'\in S}\mathbf{\Theta}(s')\cdot\tau_{s'}(t)} \tag{4.1}
$$

be the belief of a player of type t about the state of nature, computed under the imagined-continuum reasoning, as in (3.3). Also let $\boldsymbol{\phi}_k^{(t)} = \phi(\boldsymbol{\Theta}_k^{(t)}, \boldsymbol{\Theta}_k)$ be the probability distribution of the $\Delta(X)$-valued random variable that represents the forecast of a player of type t about the outcome of period k, where the forecast function ϕ is given by (3.4). In this section we give conditions under which these probabilistic forecasts can be said to predict the outcome.

We assume hereafter that the space X of outcomes is equipped with a (complete, separable) metric η. The event that *players make* (r, ϵ)*-correct predictions in period* k is given by

$$R(k, r, \epsilon) = \left\{ \boldsymbol{\phi}_k^{(t)} \left(B(\mathbf{X}_k, r) \right) > 1 - \epsilon \text{ for every } t \in T \right\} \tag{4.2}$$

where $B(\mathbf{X}_k, r) = \{x \in X | \eta(x, \mathbf{X}_k) \leq r\}$ is the closed ball of radius r around $\mathbf{X}_k$. Thus, players make (r, ϵ) correct predictions at period k if each player assigns a probability at least $1 - \epsilon$ to a ball of radius r around the realized outcome $\mathbf{X}_k$, before she observes its realized value.

Definition 5. Let Γ be a game skeleton and let κ be a Markov strategy. We say that *players make asymptotically* (r, ϵ, ρ)*-correct predictions under* κ *in period* k if there exists some n_0 such that

$$\mathbb{P}(R(k, r, \epsilon)) > 1 - \rho$$

in every n-player game with $n > n_0$.

We proceed to provide conditions on the game skeleton under which players make asymptotically correct predictions. For every probability distribution function ν over

X, let $Q_\nu : [0,\infty) \to [0,1]$ be the *concentration function* of ν given by

$$Q_\nu(r) = 1 - \sup_D \nu(D), \tag{4.3}$$

where the supremum ranges over all closed subsets B of X with diameter $\operatorname{diam}(D) \leq r$ (where $\operatorname{diam}(D) = \sup_{x,x' \in B} \eta(x,x')$ where η is the metric on X). When ν is the distribution of a random variable $\mathbf{X}$, we also denote $Q_{\mathbf{X}} = Q_\nu$. The following are examples of concentration functions:

Example 2. [Concentration functions]

- If for some a, $a \leq \mathbf{X} \leq a + .01$, then $Q_{\mathbf{X}}(0.01) = 0$.
- If X is a finite set and $\mathbb{P}(\mathbf{X} = x_0) = 1 - \epsilon$ for some $x_0 \in X$ and small $\epsilon > 0$, then $Q_{\mathbf{X}}(0) = \epsilon$.
- If $X = \mathbb{R}$ and $\mathbf{X}$ is a random variable with variance σ^2, then from Chebyshev's Inequality it follows that $Q_{\mathbf{X}}(r) \leq 4\sigma^2/r^2$.
- If $X = \mathbb{R}$ and $\mathbf{X}$ is a random variable with a Normal distribution with standard deviation σ, then $Q_{\mathbf{X}}(r) = 2(1 - \Phi(r/2\sigma)) \leq 2\exp(-r^2/2\sigma^2)$.

For every game skeleton Γ, we let $Q_\Gamma : [0,\infty) \to [0,1]$ be given by $Q_\Gamma(r) = \sup_{s,e} Q_{\chi_{s,e}}(r)$. For example $Q_\Gamma(0.01) = 0$ in the *round-off case*, where outcomes are empirical distributions randomly rounded off to integral percentages.

Theorem 1 (Correct predictions)**.** *Fix a game skeleton Γ. For every $\epsilon, \rho > 0$ there exists an integer K such that under every Markov strategy κ and every $r > 0$, in all but at most K periods players make $[r, Q_\Gamma(r) + \epsilon, Q_\Gamma(r) + \rho]$-asymptotically correct predictions.*

The appearance of $Q_\Gamma(r)$ in Theorem 1 is intuitively clear: increasing the concentration of the random outcome (e.g., by taking a larger sample size J in Example 1) improves the level of predictability and stability. But if the variance is large (as in Example 1 with a small sample size), predictability and stability are not to be expected.

5. Stability

Consider a game skeleton Γ played repeatedly by n players. Let κ be a Markov strategy and consider the random κ play path $\left(\mathbf{S}, \mathbf{T}^i, \mathbf{A}^i_k, \mathbf{X}_k\right)_{i\in N, k=0,1,\ldots}$. The event that period k is ϵ *-hindsight stable* is given by

$$H(k,\epsilon) = \left\{u(\mathbf{T}^i, \mathbf{A}^i_k, \mathbf{X}^i) + \epsilon \geq u(\mathbf{T}^i, a, \mathbf{X}^i) \text{ for every player } i \text{ and action } a \in A\right\}.$$

This is the event that after observing the realized outcome of period k no player can improve her payoff by more than ϵ through a unilateral revision of her period-k action. The probability of this event is, $P_{\text{HS}}(k, \epsilon; \kappa) = \mathbb{P}\big(u(\mathbf{T}^i, \mathbf{A}^i_k, \mathbf{X}^i) + \epsilon \geq u(\mathbf{T}^i, a, \mathbf{X}^i)$ for every player i and action $a\big)$.

Definition 6. Let Γ be a game skeleton and let κ be a Markov strategy. We say that period k is *asymptotically* (ϵ, ρ)*-stable under* κ if there exists some n_0 such that

$$\mathbb{P}(H(k,\epsilon)) > 1 - \rho$$

in every n-player game with $n > n_0$.

We proceed to provide bounds on the level of hindsight stability in natural classes of large games. For this purpose, in addition to the earlier assumptions about the game skeleton, we now make an assumption about the modulus of continuity of the payoff function. Let $\omega : [0, \infty) \to [0, \infty)$ be a continuous, monotone increasing function with $w(0) = 0$. We say that the utility function u admits ω as a *modulus of continuity* if $|u(t, a, x) - u(t, a, x')| \leq \omega(\eta(x, x'))$ for all $t \in T, a \in A$, and $x, x' \in X$, where η is the metric on X. The special case of a *Lipschitz payoff function with constant* L is described by the function $\omega(d) = Ld$.

The following lemma says that correct predictions imply hindsight stability.

Lemma 1. *Fix $r, \epsilon > 0$. Let κ be a Markov strategy of a game skeleton Γ in which the payoff function u has a modulus of continuity ω, and consider the random κ-play-path. For every period k, the event $R(k, r, \epsilon)$ that players make (r, ϵ) -correct predictions in period k is contained in the event $H(k, 2\omega(r) + \epsilon/(1-\epsilon))$ that period k is $(2\omega(r) + \epsilon/(1-\epsilon))$-stable.*

The following theorem follows from Theorem 1 and Lemma 1:

Theorem 2. *[Hindsight Stability] Fix a game skeleton Γ in which the payoff function u has a modulus of continuity ω . Then for every $\epsilon, \rho > 0$, there exists an integer K such that in every Markov equilibrium κ and for every $d > 0$, all but at most K periods are $[2d + 2Q_\Gamma(\omega^{-1}(d)) + 2\epsilon, Q_\Gamma(\omega^{-1}(d)) + \rho]$ asymptotically-stable.*[10]

Under the theorems above we deduce the following examples.

Example 3 (Rounded-off empirical distribution)**.** Consider a game Γ in which the reported outcome x is the realized empirical distribution of the population e, randomly rounded off, up or down, to the nearest percentage. Then $Q_\Gamma(0.01) = 0$. Let

[10]If the function ω is not invertible, then $r = \omega^{-1}(d)$ is defined in any way such that $\omega(r) \leq d$.

r, ϵ, ρ be arbitrarily small positive numbers; then there is a finite number of periods K such that:

(1) In all but the Kperiods, under any strategy κ the players make correct predictions up to $[r, \epsilon, \rho]$.
(2) If the payoff function is Lipschitz with constant $L = 1$ and κ is a Markov equilibrium, then all but K periods are $(0.02 + 2\epsilon, \rho)$-stable.

6. Cournot example: Price stability

In this example of an n-person Cournot production game, the state of nature determines whether it is easy or difficult to produce a certain good, and producers are of two types: efficient and inefficient. At the beginning of every period, each one of the producers chooses whether or not to produce a single unit of the good. The total production determines the period's price through a random inverse demand function.

Let $S = \{\text{easy}, \text{difficult}\}$ denote the set of possible states equipped with a uniform prior $\theta_0(\text{easy}) = \theta_0(\text{difficult}) = 1/2$, and let $T = \{\text{efficient}, \text{inefficient}\}$ denote the set of players' types. Let the type-generating function τ be given by

$$\tau(\text{efficient}|\text{easy}) = \tau(\text{inefficient}|\text{difficult}) = 3/4, \text{ and}$$

$$\tau(\text{inefficient}|\text{easy}) = \tau(\text{efficient}|\text{ difficult}) = 1/4.$$

A player's period production levels are described by the set of actions $A = \{0, 1\}$, and a price $x \in R$ is the outcome of every period. The period price depends entirely on the period's total production, and not on the state of nature and the types. Formally, for every $s \in S$ and empirical distribution of type-action pairs $e \in \Delta(T \times$

A), let $\chi_{s,e} = \text{Normal}(1/2 - r, \sigma^2)$, where $r = e(\text{easy}, 1) + e(\text{ difficult}, 1)$ is the proportion of players who produce the good. One interpretation in the n-player game is that there are n buyers whose demand at price x is given by $1/2 - r + \epsilon$, where $\epsilon \sim$ Normal$(0, \sigma^2)$ is the same for all buyers. Another interpretation is that ϵ represents noisy traders who may either buy or sell the good.

The payoff function is given by $u(t, 0, x) = 0$ for every $t \in T$ and $x \in X$ and $u(t, 1, x) = x - (1/8)\delta_{t=\text{inefficient}}$. These means that not producing results in zero payoff and that per unit production cost is zero for an efficient producer and 1/8 for an inefficient one.

The repeated game admits the following unique imagined-continuum Markov equilibrium: Let θ_k be the public belief about the state of nature at the beginning of period k, computed (according to the imagined-continuum reasoning) by an outsider who observes the prices but not the players' types and actions. We identify θ_k with the probability assigned to $s = \text{easy}$, so $\theta_k \in [0, 1]$. Note that if the public belief is θ_k, then the belief of every efficient player is

$$\theta_k^{(\text{efficient})} = \frac{3/4 \cdot \theta_k}{3/4 \cdot \theta_k + 1/4(1 - \theta_k)} = \frac{3\theta_k}{1 + 2\theta_k},$$

and the belief of every inefficient players is

$$\theta_k^{(\text{inefficient})} = \frac{\theta_k}{3 - 2\theta_k}.$$

The equilibrium strategies in the repeated game are defined by the following:

(1) When $\theta_k \geq (7 + \sqrt{33})/16 = 0.796..$, each efficient player produces with prob $p = \frac{4\theta_k + 2}{8\theta_k + 1}$ (thus, under imagined-continuum reasoning, $\frac{4\theta_k + 2}{8\theta_k + 1}$ of them produce)

and the inefficient players are idle. Here p is the solution to the equation:

$$\theta_k^{(\text{efficient})} 3/4 \cdot p + (1 - \theta_k^{(\text{ efficient})}) 1/4 \cdot p = 1/2,$$

so that the efficient players expect a selling price of zero and zero profit. In particular, when $\theta_k = 1$, a proportion $p = 2/3$ of the efficient players produce and the inefficient players are idle.

(2) When $(35 - \sqrt{649})/64 < \theta_k < (7 + \sqrt{33})/16$, each efficient player produces with probability p and each inefficient player produces with probability q, where $0 < p, q < 1$ are the unique solution to the following equations

$$\theta_k^{(\text{efficient})}(3/4 \cdot p + 1/4 \cdot q) + (1 - \theta_k^{(\text{efficient})})(1/4 \cdot p + 3/4 \cdot q) = 1/2$$
$$\theta_k^{(\text{inefficient})}(3/4 \cdot p + 1/4 \cdot q) + (1 - \theta_k^{(\text{inefficient})})(1/4 \cdot p + 3/4 \cdot q) = 3/8,$$

so that the efficient players expect price 0 and the inefficient players expect price 1/8. For example, when $\theta = 1/2$, the strategies are $p = 11/16$ and $q = 3/16$.

(3) When $\theta \leq (35 - \sqrt{649})/64 = 0.148...$, the efficient players all produced and the inefficient players produce with probability $q = (3 - 6\theta)/(18 - 16\theta)$. Here q is the solution to the following equation:

$$\theta_k^{(\text{inefficient})} \cdot (3/4 + 1/4 \cdot q) + (1 - \theta_k^{(\text{efficient})}) \cdot (1/4 + 3/4 \cdot q) = 3/8,$$

so that the inefficient player expects price 1/8 and zero profit. In this case the efficient players expect a positive profit.

After each period the players update their beliefs using Bayes' formula:

$$\theta_{k+1} = \frac{\theta_k \cdot \exp\left(-(x_k - (3/4p_k + 1/4q_k))^2/2\right)}{\theta_k \cdot \exp\left(-(x_k - (3/4p_k + 1/4q_k))^2/2\right) + (1-\theta_k) \cdot \exp\left(-(x_k - (1/4p_k + 3/4q_k))^2/2\right)}$$

where p_k and q_k are the equilibrium strategy under θ_k, and x_k is the outcome of period k.

By Theorem 2 it follows that for every $\epsilon, \rho > 0$ and every $d > 0$, every period except for a finite number is asymptotically hindsight-stable at level $(2d + 2Q_\Gamma(d) + 2\epsilon, Q_\Gamma(d) + \rho)$. Assume, for example, that $\sigma = 0.01$. Choosing $d = 0.05$, we get (by item 2 in Example 2) $Q_G(d) = 0.012$. Therefore, every period except for a finite number is asymptotically $(0.11 + 2\epsilon, 0.012 + \rho)$-stable.

Remark 3. Why is the equilibrium unique? Let θ be the outsider belief about the state of nature at the beginning of some period. Let p be the proportion of efficient players who produce at that period, and q the proportion of inefficient players who produce.

Under this profile the supplied quantity that the efficient players expect is

$$\theta^{(\text{efficient})}(3/4 \cdot p + 1/4 \cdot q) + (1 - \theta^{(\text{ efficient})})(1/4 \cdot p + 3/4 \cdot q),$$

and the supplied quantity that the inefficient players expect is

$$\theta^{(\text{inefficient})}(3/4 \cdot p + 1/4 \cdot q) + (1 - \theta^{(\text{ inefficient})})(1/4 \cdot p + 3/4 \cdot q).$$

Assume now by contradiction that (p, q) and (p', q') are two equilibrium profiles and that $q > q'$. The equilibrium condition implies that the supplied quantity that the inefficient players expect under (p, q) is weakly smaller than what they expect under (p', q'). Because $q > q'$, this implies that $p < p'$, so that, again by the equilibrium condition, the supplied quantity that the efficient players expect under (p, q) is weakly larger than under (p', q'). This is a contradiction since the difference between the expected supplied quantities of the efficient and inefficient players is monotone-increasing in p and monotone-decreasing in q.

Part 2. Big games with unobserved changing fundamentals

We next discuss preliminary observations about equilibrium models of big games with unobserved changing fundamentals. There are several key questions about such games, for example: (1) What is the process that governs the changes in fundamentals, the time of changes and the selection of new states? (2) What is the information that the players receive about the time of changes? (3) what is the information that the players receive about the new states of fundamentals?

Assuming that the changes of fundamentals occure at random discrete times, $0 = C_0 < C_1 < C_2, \ldots$, the play of the big game consists of *segments* of periods $[C_i, C_{i+1})$ during which the state of fundamentals is fixed, and in each such segment we have a repeated big game with unknown state of fundamentals. We now proceed to discuss how the Markov equilibria for such segments, from the earlier part of this paper, may be used to construct equilibria for the multi-segment game. This construction is limited to big *repeated segment games* in which the changes in fundamentals and information is very simple. Nevertheless, it may offer an approach for research dealing with more general cases.

In a repeated segment game the changes in fundamentals are due to exogenous iid random shocks that are independent of the play, and when a change takes place the players are informed about it, but not about the new realized state of fundamentals. More precisely, a repeated segment game is described by a pair (Γ, ξ) in which $\Gamma = (S, \theta_0, T, \tau, A, X, \chi, u)$ is a fixed multi-period stage game with fixed unknown fundamentals as discussed earlier, and $1 > \xi > 0$ describes the probability of change at the end of every period of play.

The extensive form of the game is described recursively as follows: (i) Independent of any past events, nature draws a random state of nature s according to the distribution θ_0; it also independently draws player types by the distribution τ_s, and privately informs the players of their realized types. (ii) The segment game Γ is played repeatedly in successive periods as described earlier, but subject to the following modification: At the end of every period, independent of the history, and with probability ξ, the game *restarts* and with probability $1-\xi$ it *continues.* Following the continue event, the play continues to the next period of the segment. But following a restart event, the players are informed that the game has been restarted, and a new segment starts with step (i) above.

Let κ be a Bayesian Markov strategy of the stage game Γ , according to Definition 3. Then κ induces the following natural multi segment strategy: Initially and after every change in fundamentals, the public belief is set to be θ_0 and players use their current type when applying κ throughout the segment. Even without giving a formal definition of an equilibrium in the multi segment game, it is intuitively clear that if κ is an equilibrium in the stage game, then the induced strategy is an equilibrium in the multi segment game. Our results from the previous part imply that in every segment there is a bounded number K of periods which are not hindsight-stable. If the probability of transition ξ is sufficeintly small then we get a low bound of $\xi \times K$ on the frequency of periods in the multi segment game which are not hindsight stable.

Notice that the construction above with small ξ suggests that in big games we may anticipate the emergence of stability cycles. Every new segment of play consists of learning unstable periods, followed by many predictable stable periods played until the next change of fundamentals.

The construction above may be extended to a broader class of multi segment games, provided that the transitions to new segments are governed entirely by the Markovian state of the segment and its equilibrium currently played.

Moving to more general models, we may view big games as large imperfectly observed Bayesian stochastic games, in which the transition probability and the new state depend on the empirical distribution of players, actions and on the current state. This broader view gives rise to many questions for future research. A special challenge are games in which players are not informed when fundamental changes occure, they can only infere this by observing statistical changes in period outcomes. We leave these issues to future research.

Part 3. Proofs

The main result of the paper – that asymptotic hindsight stability holds in all but finitely many chaotic learning periods – is proven in two steps.

Step 1 argues that the result holds in the imagined processes that describe the beliefs of the players. Building on the result of Step 1, Step 2 shows that the result holds in the real process.

In Step 1, the intuition may be broken into two parts. First, relying on the literature on merging (see Fudenberg and Levin [12], Sorin [36], and Kalai and Lehrer [20, 21]) we argue that in an equilibrium of our model there is only a finite number of learning periods in which the forecasted probability of the period outcome is significantly different from its real probability. In other words, the players' belief about the fundamental state s leads to approximately the same probability distribution over the future outcomes as under the realized state s. One issue we need to address in applying these results is that, in our multi player setup, players with different types have different beliefs, and so may make mistakes in forecasts in different periods. But we need to bound the number of periods in which some player makes a forecasting mistake. To do that, we extend the previous learning result to a multi player setup.

The second part of Step 1 is based on the following reasoning. If we assume that the uncertainty in the determination of period outcomes is low, in the imagined process the period outcomes under every state s are essentially deterministic. This implies that in every non learning period the players learn to predict the outcomes (and not just forecast their probabilities). When the predicted period outcome (on which a player bases her optimal choice of an action) is correct, she has no reason to revise her choice. Thus, in the imagined processes we have hindsight stability in these non learning periods.

In Step 2, to argue that hindsight stability holds in all the non learning periods in the real process, we rely on arguments developed in our companion paper [22].

These arguments show that the probability of outcomes in the real process are approximately the same as their counterparts in the imagined process. Thus, the high level of hindsight stability obtained in Step 1 also applies to the real process.

Section 7 gives a formal definition of the imagined play path: the incorrect one computed in the minds of the players (recall Section 3.2). Section 8 presents the result from our companion paper: when the number of players is large, the probabilities of outcomes computed in the imagined process are close to the probabilities obtained from the real process (recall Section 3.4). Section 9 presents a uniform merging (learning) result: in an environment with many player types, each starting with a different initial signal, there is a bound on the number of periods in which some types update their beliefs. Section 10 connects the dots.

7. Imagined-continuum view

In this section we describe the imagined play path that justifies the imagined-continuum reasoning of Section 3.2. This is needed because the Bayesian updating performed by the imagined-continuum players is done relative to this imagined play path, and not relative to the real play path described in Section 3.4.

In order to distinguish between corresponding entities in the actual play path and in the imagined play path, we denote the random variables that represent the outcomes in the imagined play by $\tilde{\mathbf{X}}_0, \tilde{\mathbf{X}}_1, \dots$, and the random variables that represent public beliefs by $\tilde{\mathbf{\Theta}}_0, \tilde{\mathbf{\Theta}}_1, \dots$.

Let κ be a Markov strategy. An *imagined random κ-play-path* is a collection $(\mathbf{S}, \tilde{\mathbf{T}}, \tilde{\mathbf{X}}_0, \tilde{\mathbf{X}}_1, \dots)$ of random variables, representing the state of nature, type of a representative player and outcomes, such that: the state of nature $\mathbf{S}$ is distributed according to θ_0 and conditional on the history of periods $0, \ldots, k-1$, the outcome $\tilde{\mathbf{X}}_k$ is drawn randomly according to the probability density function $\chi_{\mathbf{S}, d_{\mathbf{S}, \tilde{\boldsymbol{\Theta}}_k}}$, where the imagined public beliefs $\tilde{\boldsymbol{\Theta}}_k$ are given by

$$\begin{aligned} &\tilde{\boldsymbol{\Theta}}_0 = \theta_0, \text{ and} \\ &\tilde{\boldsymbol{\Theta}}_{k+1} = \beta_{\tilde{\boldsymbol{\Theta}}_k, \tilde{\mathbf{X}}_k}, \text{ for } k \geq 0; \end{aligned} \tag{7.1}$$

β is defined in (3.2); and $d_{s,\theta}$ for every state s and belief θ is defined in (3.1).

In equations:

$$\begin{aligned} &\mathbb{P}(\mathbf{S} = \cdot) = \theta_0; \\ &\mathbb{P}\left(\tilde{\mathbf{T}} = \cdot | \mathbf{S}\right) = \tau_{\mathbf{S}}; \\ &\mathbb{P}\left(\tilde{\mathbf{X}}_k = x \,\middle|\, \mathbf{S}, \tilde{\mathbf{T}}, \tilde{\mathbf{X}}_0, \ldots, \tilde{\mathbf{X}}_{k-1}\right) = \chi_{\mathbf{S}, d_{\mathbf{S}, \tilde{\boldsymbol{\Theta}}_k}}(x), \end{aligned} \tag{7.2}$$

The difference between Equations (7.2) and the Equations (3.7) that defined the actual random play path is that in the latter, the outcome is generated from the random empirical types-actions distribution $\mathbf{e}_k$ of n players, whereas in the former the outcome is generated from $d_{\mathbf{S},\ \tilde{\boldsymbol{\Theta}}_k}$. For this reason the beliefs $\tilde{\boldsymbol{\Theta}}_k$ are the conditional probabilities over the state of nature of an observer who views the outcome process $\tilde{\mathbf{X}}_0, \tilde{\mathbf{X}}_1, \ldots$, and $\tilde{\boldsymbol{\Theta}}_k^{(t)}$ are the correct conditional probabilities over the state of nature of a player of type $\mathbf{T}$, i.e.,

$$\begin{aligned} \tilde{\boldsymbol{\Theta}}_k &= \mathbb{P}\left(\mathbf{S} = \cdot \,\middle|\, \tilde{\mathbf{X}}_0, \ldots, \tilde{\mathbf{X}}_{k-1}\right), \text{ and} \\ \tilde{\boldsymbol{\Theta}}_k^{(t)} &= \mathbb{P}\left(\mathbf{S} = \cdot \,\middle|\, \mathbf{T} = t, \tilde{\mathbf{X}}_0, \ldots, \tilde{\mathbf{X}}_{k-1}\right) \end{aligned} \tag{7.3}$$

where for every $t \in T$, $\tilde{\mathbf{\Theta}}_k^{(t)}$ is given by

$$\tilde{\mathbf{\Theta}}_k^{(t)}(s) = \frac{\tilde{\mathbf{\Theta}}(s) \cdot \tau_s(t)}{\sum_{s' \in S} \tilde{\mathbf{\Theta}}(s') \cdot \tau_{s'}(t)}$$

as in (4.1). Similarly, the forecasts of the public observer and the players about the next period outcome are correct in the imagined process:

$$\begin{aligned} \phi(\tilde{\mathbf{\Theta}}_k, \tilde{\mathbf{\Theta}}_k) &= \mathbb{P}\left(\tilde{\mathbf{X}}_k = \cdot \,\middle|\, \tilde{\mathbf{X}}_0, \dots, \tilde{\mathbf{X}}_{k-1}\right), \text{ and} \\ \phi(\tilde{\mathbf{\Theta}}_k^{(t)}, \tilde{\mathbf{\Theta}}_k) &= \mathbb{P}\left(\tilde{\mathbf{X}}_k = \cdot \,\middle|\, \tilde{\mathbf{T}} = t, \tilde{\mathbf{X}}_0, \dots, \tilde{\mathbf{X}}_{k-1}\right). \end{aligned} \tag{7.4}$$

From (7.4) it follows that if κ is an imagined-continuum equilibrium, then at every round the players choose the optimal actions for the imagined beliefs:

$$[\kappa(t, \tilde{\Theta}_k)] \subseteq \arg\max_a \mathbb{E}\left(u(t, a, \tilde{\mathbf{X}}_k) \,\middle|\, \tilde{\mathbf{T}} = t, \tilde{\mathbf{X}}_0, \dots, \tilde{\mathbf{X}}_{k-1}\right), \tag{7.5}$$

for every period k and every player's type $t \in T$. As mentioned in Remark 2, in our companion paper we define the concept of imagined equilibrium (not necessarily Markovian) using this property and prove that every equilibrium of this kind is myopic.

8. Validation of the imagined view

We prove Theorem 1 using a proposition that couples a play path in the actual game with an imagined play path that reflects the players imagined-continuum reasoning. By *coupling* we mean that both processes are defined on the same probability space. The coupling presented in Proposition 1 is such that when the number of players is large, the realization of the processes is with high probability the same. In particular, the forecasts about the outcome sequence made by the imagined-continuum reasoning are not far from the correct forecasts made by an observer who performs the correct Bayesian calculation[11]. We prove Proposition 1 in our companion paper[12]. See also Carmona and Podczeck [7, 8] and the reference therein for results of similar spirit in a static (single period) game.

Proposition 1. *Fix a game skeleton and a Markov strategy κ. There exist random variables $\mathbf{S}, \mathbf{T}^i, \mathbf{A}^i_k, \mathbf{X}_k, \tilde{\mathbf{X}}_k$ for $i \in N$ and $k = 0, 1, \ldots$ such that*

- $(\mathbf{S}, \mathbf{T}^i, \mathbf{A}^i_k, \mathbf{X}_0,\ \mathbf{X}_1, \ldots)$ *is a random κ-play path of the repeated game.*
- *The outcome sequence $\mathbf{S}, \tilde{\mathbf{X}}_0,\ \tilde{\mathbf{X}}_1, \ldots$ is an imagined random κ-play path.*
- *For every k,*

$$\mathbb{P}\left(\mathbf{X}_0 = \tilde{\mathbf{X}}_0, \ldots, \mathbf{X}_k = \tilde{\mathbf{X}}_k\right) > 1 - C \cdot k\sqrt{\frac{\log n}{n}}, \tag{8.1}$$

where C is a constant that depends only on the game skeleton.

9. A uniform merging theorem

The proof of Theorem 1 relies on the notion of merging. Roughly speaking, the merging literature shows that if forecasts over a sequence of future outcomes are made by a Bayesian agent whose belief contains some 'grain of truth' (in a sense defined by this literature) then the agent's forecasts cannot be wrong in many periods. We follow Sorin's paper [36]. For our purpose there is a set of agents, each of whom holds a belief that has some grain of truth, and we want to bound the number of

[11]In a sense, the proposition states that the incorrect imagined-continuum computations are validated by the observed outcomes. This idea is a similar to self-confirming equilibrium [11] with two exceptions: (i) Here the players may have an incorrect understanding of the game, not just of the opponents' strategies. (ii) our validation is probabilistic.

[12]See Lemma 1 in that paper. The version of the lemma in that paper is more general than here in that it does not assume that all players play the same Markov strategy and also allows arbitrary (non Markovian) deviations of a player.

periods in which at least one of these agents makes a wrong forecast. In principle, different beliefs with a grain of truth may induce wrong forecasts in different periods. Nevertheless, in this section we use Sorin's result to show that we can still bound the number of periods in which some agents make a wrong forecast. Proposition 2 essentially appears in Sorin's paper. Proposition 3 is our generalization for a multi-agent setup.

Let X be a Borel space of outcomes and P a probability distribution over $X^{\mathbb{N}}$. For every $x = (x_0, x_1, \dots) \in X^{\mathbb{N}}$, we denote by $P_k(x_0, \dots, x_{k-1})$ the forecast made by P over the next period outcome conditioned on $x_0, \dots, x_{k-1}$. Let S be a finite set of states and for every $s \in S$, let $P^s \in \Delta(X^{\mathbb{N}})$. For a belief $\theta \in \Delta(S)$ let $P^\theta = \sum_{s \in S} \theta(s) P^s \in \Delta(X^{\mathbb{N}})$. Thus, P^θ is the belief over outcomes of a player with a prior θ over the states of nature.

Consider a probability space equipped with random variables $\mathbf{S}, \mathbf{X}_0, \mathbf{X}_1, \dots$ such that

$$\mathbb{P}(\mathbf{S} = s) = \theta_0 \text{ for every } s \in S, \text{ and}$$

$$\mathbb{P}(\mathbf{X}_0 = \cdot, \mathbf{X}_1 = \cdot, \dots \mid \mathbf{S} = s) = P^s$$

for every $s \in S$.

For $P, Q \in \Delta(X^{\mathbb{N}})$ and $\epsilon > 0$ we denote by $D_{k,\epsilon}(P, Q)$ the event that the forecast about $\mathbf{X}_K$ made by P and Q are differ by more than ϵ, i.e.,

$$D_{k,\epsilon}(P, Q) = \{\|P_k(\mathbf{X}_0, \dots, \mathbf{X}_{k-1}) - Q_k(\mathbf{X}_0, \dots, \mathbf{X}_{k-1})\| > \epsilon\}.$$

Here and later the norm is the total variation norm of signed measure. In terms of density function this norm is the L^1-norm.

The following proposition is included in results that appear in Sorin [36] and Neyman [30]. In order to keep the paper self-contained we include a proof following Neyman's argument.

Proposition 2. *For every $\epsilon, \rho > 0$ there exists $K = K(\epsilon, \rho)$ such that for every prior belief $\theta_0 \in \Delta(S)$ and every collection of distributions $\{P^s \in \Delta(X^{\mathbb{N}}) | s \in S\}$, in every period k except at most K it holds that:*

$$\mathbb{P}\left(D_{k,\epsilon}\left(P^{\mathbf{S}}, P^{\theta_0}\right)\right) > 1 - \rho.$$

The meaning of the assertion in Proposition 2 is that if the state of nature is randomized according to θ_0, then at period k , with high probability an agent who does not observe the state of nature make correct forecast, as if she knew the realized state of nature.

Proof. One has

$$\begin{aligned}
&\sum_{k=0}^{\infty} \mathbb{E}\left\|P_k^{\mathbf{S}}(\mathbf{X}_0, \ldots, \mathbf{X}_{x-1}) - P_k^{\theta_0}(\mathbf{X}_0, \ldots, \mathbf{X}_{x-1})\right\|^2 = \\
&\sum_{k=0}^{\infty} \mathbb{E}\left\|\mathbb{P}(\mathbf{X}_k \in \cdot \,|\mathbf{S}, \mathbf{X}_0, \ldots, \mathbf{X}_{k-1}) - \mathbb{P}(\mathbf{X}_k \in \cdot \,|\mathbf{X}_0, \ldots, \mathbf{X}_{k-1})\right\|^2 \leq \\
&2\sum_{k=0}^{\infty} \mathbb{E}\ D_{KL}\left(\mathbb{P}(\mathbf{X}_k \in \cdot \,|\mathbf{S}, \mathbf{X}_0, \ldots, \mathbf{X}_{k-1}), \mathbb{P}(\mathbf{X}_k \in \cdot \,|\mathbf{X}_0, \ldots, \mathbf{X}_{k-1})\right) = \\
&2\sum_{k=0}^{\infty} \mathbb{E}\ I\left(S; \mathbf{X}_k \,|\mathbf{X}_0, \ldots, \mathbf{X}_{k-1}\right) = \\
&2\sum_{k=0}^{\infty} \left(H(S|\mathbf{X}_0, \ldots, \mathbf{X}_{k-1}) - H(S|\mathbf{X}_0, \ldots, \mathbf{X}_k)\right) = \\
&2 \cdot \left(H(S) - H(S|\mathbf{X}_0, \mathbf{X}_1 \ldots)\right) \leq 2H(S) \leq 2|S| \log |S|.
\end{aligned} \tag{9.1}$$

Here D_{KL} denotes the Kullback-Leibler divergence between distributions, the first inequality follows from Pinsker's Inequality, I is the mutual information, and H is the entropy. The assertion follows by choosing $K = 2|S|\log|S|/(\epsilon^2\rho)$. □

We now turn to the main contribution of this section, which is a generalization of Proposition 2 to a multi-agent setup, where the agents may have different beliefs over state of nature S because they receive different private signals. A *stochastic signal* is given by a function $\zeta : S \to [0,1]$, with the interpretation that an agent observes *the signal* with probability $\zeta(s)$ if the state of nature is s. In our setup, an agent of type t receives the stochastic signal ζ that is given by $\zeta(s) = \tau_s(t)$. An agent who has some prior θ about S and receives a signal ζ updates her belief to

$$\theta^{(\zeta)}(s) = \frac{\theta(s)\cdot\zeta(s)}{\sum_{s'\in S}\theta(s')\zeta(s')}. \tag{9.2}$$

This is the same formula as (3.3) except that we use the abstract notation of a stochastic signal. The posterior belief is underined when $\zeta(s) = 0$. Finally, for $\delta > 0$, let $Z_\delta = \{\zeta : S \to [\delta, 1]\}$ be the set of stochastic signals with probability at least δ in every state. The following proposition is the main result of this section:

Proposition 3. *For every $\delta, \epsilon, \rho > 0$, there exists $K = K(\delta, \epsilon, \rho)$ such that for every prior belief $\theta_0 \in \Delta(S)$ and every collection of distributions $\{P^s \in \Delta(X^{\mathbb{N}})|s \in S\}$, in every period k except at most K one has:*

$$\mathbb{P}\left(\bigcap_{\zeta\in Z_\delta} D_{k,\epsilon}\left(P^{\mathbf{S}}, P^{\theta_0^{(\zeta)}}\right)\right) > 1-\rho.$$

The meaning of the assertion in Proposition 3 is that if the state of nature is randomized according to θ_0, then at period k , with high probability all agents who receive signals in Z_δ about the state of nature make simultaneously correct forecasts, as if they knew the realized state of nature.

To prove Proposition 3, we will use the following claim, which is a generalization of Cauchy-Schwarz Inequality $|\mathrm{Cov}(\mathbf{X},\mathbf{Y})| \le \sqrt{\mathrm{Var}(X)\ \mathrm{Var}(Y)}$ to random variables that assume values in a separable Banach space. For such variables the expectation is the Bochner integral. The case we are interested in is when the random variable $\boldsymbol{\phi}$ is an agent's forecast, which is a probability distribution over X. The set of all forecasts is a subset of the Banach space of finite signed measure over X. [13]

Claim 1. *Let $\boldsymbol{\phi}$ be a random variable which assumes values in some Banach space V, and let $\boldsymbol{\zeta}$ be a real-valued random variable, both bounded. Define $Cov(\boldsymbol{\zeta},\ \boldsymbol{\phi}) = \mathbb{E}\boldsymbol{\zeta}\boldsymbol{\phi} - \mathbb{E}\boldsymbol{\zeta}\mathbb{E}\boldsymbol{\phi} \in V$. Then $\|Cov(\boldsymbol{\zeta},\ \boldsymbol{\phi})\| \le \sqrt{\mathrm{Var}(\boldsymbol{\zeta}) \cdot \mathbb{E}\|\boldsymbol{\phi} - \mathbb{E}\boldsymbol{\phi}\|^2}$.*

Proof. From the linearity of the expectation, it follows that

$$\mathrm{Cov}(\boldsymbol{\zeta},\boldsymbol{\phi}) = \mathbb{E}\left(\boldsymbol{\zeta} - \mathbb{E}\boldsymbol{\zeta}\right)\left(\boldsymbol{\phi} - \mathbb{E}\boldsymbol{\phi}\right)$$

Therefore, the following holds:

$$\|\mathrm{Cov}(\boldsymbol{\zeta},\boldsymbol{\phi})\| \le \mathbb{E}\|\left(\boldsymbol{\zeta} - \mathbb{E}\boldsymbol{\zeta}\right)\left(\boldsymbol{\phi} - \mathbb{E}\boldsymbol{\phi}\right)\| = \\ \mathbb{E}\left(|\boldsymbol{\zeta} - \mathbb{E}\boldsymbol{\zeta}| \cdot \|\boldsymbol{\phi} - \mathbb{E}\boldsymbol{\phi}\|\right) \le \sqrt{\ \mathrm{Var}(\boldsymbol{\zeta}) \cdot \mathbb{E}\|\boldsymbol{\phi} - \mathbb{E}\boldsymbol{\phi}\|^2},$$

[13] Recall that X is either countable or a subset of an Euclidean space, in which case we assume also that the forecasts admit density, and view $\Delta(X)$ as a subset of the Banach space of finite signed measures which are absolutely continuous w.r.t. Lebesgue's measure. In both case the norm is the total variation norm, which equal the L^1-norm over the corresponding densities (and, therefore, separable).

where the first inequality follows from Jensen's inequality the and convexity of the norm, the equality follows from properties of the norm, and the last inequality follows from the Cauchy-Schwarz inequality. □

Proof of Proposition 3. Let $\gamma = \epsilon^2\rho\delta/8$ and $K = 2|S|\log|S|/\gamma = 16|S|\log|S|/(\epsilon^2\rho\delta)$. From (9.1) it follows that for every period k except at most K of them, it holds that

$$\mathbb{E}\left\|P_k^{\mathbf{S}}(\mathbf{X}_0,\dots,\mathbf{X}_{x-1}) - P_k^{\theta_0}(\mathbf{X}_0,\dots,\mathbf{X}_{x-1})\right\|^2 < \gamma. \tag{9.3}$$

Let $\mathcal{F}_k$ be the sigma-algebra that is generated by $\mathbf{X}_0,\dots,\mathbf{X}_{k-1}$ and let

$$\boldsymbol{\phi}_k = P_k^{\mathbf{S}}(\mathbf{X}_0,\dots,\mathbf{X}_{x-1})$$

be the $\Delta(X)$-valued random variable that represents the forecast about $\mathbf{X}_k$ of an agent who knows the state of nature and has observed previous outcomes. Note that the forecast about $\mathbf{X}_k$ of an agent who observes previous outcomes but does not observe the state of nature is given by

$$P_k^{\theta_0}(\mathbf{X}_0,\dots,\mathbf{X}_{x-1}) = \mathbb{E}(\boldsymbol{\phi}_k|\mathcal{F}_k).$$

More generally, let $\zeta : S \to [0,1]$ and let $\boldsymbol{\zeta} = \zeta(\mathbf{S})$. Then the forecast about $\mathbf{X}_k$ of an agent who receives the signal ζ is given by:

$$\boldsymbol{\phi}_k^{(\zeta)} = \mathbb{E}(\boldsymbol{\zeta}\boldsymbol{\phi}_k|\mathcal{F}_k)/\mathbb{E}(\boldsymbol{\zeta}|\mathcal{F}_k).\text{[14]} \tag{9.4}$$

From (9.3)

$$\mathbb{E}\left\|\boldsymbol{\phi}_k - \mathbb{E}(\boldsymbol{\phi}_k|\mathcal{F}_k)\right\|^2 < \gamma.$$

[14] As in (9.2), the forecast is undefined on the event that $\mathbb{E}(\boldsymbol{\zeta}|\mathcal{F}_k)$. This event has probability 0 from the perspective of an agent who get the signal ζ

We call periods k in which this inequality holds *good periods*. It follows that in a good period k there exists an $\mathcal{F}_k$ -measurable event G_k and an event H_k (not necessarily $\mathcal{F}_k$-measurable) such that $\mathbb{P}(G_k) > 1-\rho/2, \mathbb{P}(H_k) > 1-\rho/2$ and

$$\mathbb{E}\left(\|\boldsymbol{\phi}_k - \mathbb{E}(\boldsymbol{\phi}_k|\mathcal{F}_k)\|^2 \,\middle|\mathcal{F}_k\right) < 2\gamma/\rho \text{ on } G_k, \text{ and} \tag{9.5}$$

$$\|\boldsymbol{\phi}_k - \mathbb{E}(\boldsymbol{\phi}_k|\mathcal{F}_k)\|^2 < 2\gamma/\rho \text{ on } H_k. \tag{9.6}$$

From the concavity of the square root function, Jensen's inequality and (9.5), it follows that

$$\mathbb{E}\left(\|\boldsymbol{\phi}_k - \mathbb{E}(\boldsymbol{\phi}_k|\mathcal{F}_k)\|\middle|\mathcal{F}_k\right) < \sqrt{2\gamma/\rho}, \tag{9.7}$$

on G_k. From (9.4) it follows that

$$\boldsymbol{\phi}_k^{(\zeta)} - \mathbb{E}(\boldsymbol{\phi}_k|\mathcal{F}_k) = \frac{\text{Cov}\left(\boldsymbol{\zeta}, \boldsymbol{\phi}_k\,|\mathcal{F}_k\right)}{\mathbb{E}(\boldsymbol{\zeta}|\mathcal{F}_k)}$$

Assume now that $\zeta \in Z_\delta$, so that $\delta \le \boldsymbol{\zeta} \le 1$. Then

$$\frac{\sqrt{\text{Var}(\boldsymbol{\zeta}|\mathcal{F}_k)}}{|\mathbb{E}(\boldsymbol{\zeta}|\mathcal{F}_k)|} \le \frac{\sqrt{\mathbb{E}(\boldsymbol{\zeta}^2|\mathcal{F}_k)}}{|\mathbb{E}(\boldsymbol{\zeta}|\mathcal{F}_k)|} \le \sqrt{\frac{1}{\delta}}, \tag{9.8}$$

where the second inequality follows from Claim 2 below (conditioned on $\mathcal{F}_k$). Therefore, by Claim 1 (conditioned on $\mathcal{F}_k$), that

$$\|\boldsymbol{\phi}_k^{(\zeta)} - \mathbb{E}(\boldsymbol{\phi}_k|\mathcal{F}_k)\| = \frac{\|\text{Cov}\left(\zeta, \boldsymbol{\phi}_k\,|\mathcal{F}_k\right)\|}{|\mathbb{E}(\;\boldsymbol{\zeta}|\mathcal{F}_k)|} \le \frac{\sqrt{\text{Var}(\boldsymbol{\zeta}|\mathcal{F}_k)}}{|\mathbb{E}(\boldsymbol{\zeta}|\mathcal{F}_k)|}\sqrt{\mathbb{E}\left(\|\boldsymbol{\phi}_k - \mathbb{E}(\boldsymbol{\phi}_k|\mathcal{F}_k)\|^2\middle|\mathcal{F}_k\right)} < \sqrt{\frac{2\gamma}{\delta\rho}}$$

on G_k, where the last inequality follows fram (9.8) and (9.5). From the last equation and (9.6) it follows that

$$\|\boldsymbol{\phi}_k^{(\zeta)} - \boldsymbol{\phi}_k\| < \sqrt{\frac{2\gamma}{\rho}}(1 + 1/\sqrt{\delta}) < \epsilon$$

on $G_k \cap H_k$. Therefore,

$$\mathbb{P}\left(\cap_{\zeta \in Z_\delta} D_{k,\epsilon}\left(P^{\mathbf{S}}, P^{\theta_0^{(\zeta)}}\right)\right) = \mathbb{P}\left(\cap_{\zeta \in Z_\delta}\{\|\boldsymbol{\phi}_k^{(\zeta)} - \boldsymbol{\phi}_k\| < \epsilon\}\right) \geq \mathbb{P}(G_k \cap H_k) > 1 - \rho.$$

□

Claim 2. *If $\boldsymbol{\zeta}$ is a random variable such that $\delta \leq \boldsymbol{\zeta} \leq 1$ then $\mathbb{E}\boldsymbol{\zeta}^2 \leq \frac{1}{\delta}(\mathbb{E}\boldsymbol{\zeta})^2$.*

Proof.

$$\mathbb{E}\boldsymbol{\zeta}^2 \leq \mathbb{E}\boldsymbol{\zeta} \leq \frac{1}{\delta}\left(\mathbb{E}\boldsymbol{\zeta}\right)^2$$

where the first inequality follows from $\boldsymbol{\zeta} \leq 1$ and the third from $\delta \leq \mathbb{E}\boldsymbol{\zeta}$. □

10. Proof of Theorem 1

Claim 3. *Let $\mathbf{X}$ be an X-valued random variable and let $\nu \in \Delta(X)$ be the distribution of $\mathbf{X}$. Then for every $r > 0$,*

$$\mathbb{P}\Big(\nu(B(\mathbf{X}, r)) \geq 1 - Q_\nu(r)\Big) \geq 1 - Q_\nu(r),$$

where $Q_\nu(r)$ is the concentration function of ν given by (4.3), and $B(\mathbf{X}, r)$ is the ball of radius r around $\mathbf{X}$.

Proof. Let $D_1, D_2, \ldots$ be a subset of X such that diameter$(D_n) \leq r$ and $\nu(D_n) > 1 - Q_\nu(r) - 1/n$. Then the event $\mathbf{X} \in D_n$ implies the event $D_n \subseteq B(\mathbf{X}, r)$, which implies the event $\nu(B(\mathbf{X},r)) \geq 1 - Q_\nu(r) - 1/n$. Therefore,

$$\mathbb{P}\Big(\nu(B(\mathbf{X},r)) \geq 1 - Q_\nu(r) - 1/n\Big) \geq \mathbb{P}(\mathbf{X} \in D_n) = \nu(D_n) > 1 - Q_\nu(r) - 1/n.$$

The assertion follows by taking the limit when $n \to \infty$. □

Proof of Theorem 1. Consider the coupling $\mathbf{S}, \mathbf{T}^i, \mathbf{A}^i_k, \mathbf{X}_k, \tilde{\mathbf{X}}_k$ of the real and imagined κ-play paths given in Proposition 1. We prove that players make asymptotically correct perdictions in the imagined play path and then use (8.1) to deduce that they make correct predictions in the real play path.

We first prove that in the imagined game, players make correct forecasts, as if they knew the state of nature. We use Proposition 3 where P^s is the joint distribution of $\tilde{\mathbf{X}}_0, \tilde{\mathbf{X}}_1, \ldots$ conditioned on $\mathbf{S} = s$ for every state s. Let $K = K(\delta, 2\epsilon, \rho) = 4|S| \log |S| / (\epsilon^2 \rho \delta)$ as in Proposition 3, where $\delta = \min \tau_s(t)$ and the minimum ranges over all states s and all types t such that $\tau_s(t) > 0$. Then it follows from Proposition 3 that

$$\mathbb{P}\left(\left\|\mathbb{P}(\tilde{\mathbf{X}}_k = \cdot | \tilde{T} = t, \tilde{\mathbf{X}}_0, \ldots, \tilde{\mathbf{X}}_{k-1}) - \mathbb{P}(\tilde{\mathbf{X}}_k = \cdot | \tilde{\mathbf{S}}, \tilde{\mathbf{X}}_0, \ldots, \tilde{\mathbf{X}}_{k-1})\right\| < 2\epsilon \text{ for every } t \in T\right) > 1-\rho$$

for all periods except at most K of them. From the last equation, (7.4), and (7.2), we deduce that in all good periods

$$\mathbb{P}\left(\left\|\tilde{\boldsymbol{\phi}}^{(t)}_k - \chi_{\,\mathbf{S}, d_{\mathbf{S}, \tilde{\boldsymbol{\Theta}}_k}}\right\| < 2\epsilon \text{ for every } t \in T\right) > 1 - \rho, \tag{10.1}$$

where $\tilde{\boldsymbol{\phi}}^{(t)}_k = \phi(\tilde{\boldsymbol{\Theta}}^{(t)}_k, \tilde{\boldsymbol{\Theta}}_k)$ is the $\Delta(X)$-valued random variable that represents the forecast of a player of type t about the outcome of period k, computed in the imagined play path.

From Claim 3 conditioned on $\mathbf{S}, \tilde{\mathbf{X}}_0, \ldots, \tilde{\mathbf{X}}_{k-1}$ and from (7.2), it follows that

$$\mathbb{P}\left(\chi_{\mathbf{S}, d_{\mathbf{S}, \tilde{\mathbf{\Theta}}_k}}(B(\tilde{\mathbf{X}}_k, r)) \geq 1 - Q_\Gamma(r)\right) \geq 1 - Q_\Gamma(r).$$

From the last inequality and (10.1), we deduce that

$$\mathbb{P}\Big(\tilde{\boldsymbol{\phi}}_k^{(t)}(B(\tilde{\mathbf{X}}_k, r)) > 1 - Q_\Gamma(r) - \epsilon \quad \text{for all } t \in T\Big) > 1 - Q_\Gamma(r) - \rho.^{15}$$

All this was for the imagined play path. From Proposition 1 it now follows that

$$\mathbb{P}\left(\boldsymbol{\phi}_k^{(t)}(B(\mathbf{X}_k, r)) > 1 - Q_\Gamma(r) - \epsilon\right) > 1 - Q_\Gamma(r) - \rho$$

for sufficiently large n, as desired. □

11. Proof of Theorem 2

Proof of Lemma 1. Let $\boldsymbol{\phi}_k^{(t)}$ be the $\Delta(X)$-valued random variable that represents the forecast of a player of type t about the outcome of period k computed under the imagined reasoning. On $R(k, r, \epsilon)$

(11.1)

$$(1-\epsilon)\,(u(t,a,\mathbf{X}_k) - \omega(r)) \leq \sum_x u(t,a,x)\ \boldsymbol{\phi}_k^{(t)}(x) \leq (1-\epsilon)\,(u(t,a,\mathbf{X}_k) + \omega(r)) + \epsilon$$

[15] We use the fact that for two distributions P, Q over X it holds that $|P(B) - Q(B)| \leq \|P - Q\|/2$ for every event B in X. Recall that the norm here is the total variation norm, which is the L^1 norm for densities.

for every type t and action a. Therefore, on $R(k,r,\epsilon)$

$$u(t,b,\mathbf{X}_k) \leq \frac{1}{1-\epsilon}\sum_x u(t,b,x) \;\; \boldsymbol{\phi}_k^{(t)}(x) + \omega(r) \leq$$
$$\frac{1}{1-\epsilon}\sum_x u(t,a,x) \; \boldsymbol{\phi}_k^{(t)}(x) + \omega(r) \leq u(t,a,\mathbf{X}_k) + 2\omega(r) + \epsilon/(1-\epsilon)$$

for every $a \in [\kappa(t,\boldsymbol{\Theta}_k)]$ and $b \in A$, where the first inequality follows from (11.1), the second from the equilibrium condition $\sum_x u(t,b,x) \; \boldsymbol{\phi}_k^{(t)}(x) \leq \sum_x u(t,a,x) \; \boldsymbol{\phi}_k^{(t)}(x)$ and the third from (11.1). □

Proof of Theorem 2. Let $r = \omega^{-1}(d)$ so that $\omega(r) \leq d$. By Theorem 1 there exists an integer K such that, under every Markov strategy κ and every $r > 0$, in all but at most K periods it holds that

$$\mathbb{P}(R(k,r,Q_\Gamma(r)+\epsilon)) > 1-(Q_\Gamma(r)+\rho).$$

By Lemma 1

$$\{R(k,r,Q_\Gamma(r)+\epsilon))\} \subseteq \{H(k,2\omega(r)+(Q_\Gamma(r)-\epsilon)/(1-Q_\Gamma(r)+\epsilon)\} \subseteq \{H(k,2\omega(r)+2Q_\Gamma(r)+2\epsilon).$$

It follows that

$$\mathbb{P}\big(H(k,2\omega(r)+2Q_\Gamma(r)+2\epsilon)\big) > 1-(Q_\Gamma(r)+\rho),$$

as desired. □

References

[1] Nabil I Al-Najjar and Rann Smorodinsky. Pivotal players and the characterization of influence. *Journal of Economic Theory*, 92(2):318–342, 2000.

[2] George-Marios Angeletos, Christian Hellwig, and Alessandro Pavan. Dynamic global games of regime change: Learning, multiplicity, and the timing of attacks. *Econometrica*, 75(3):711–756, 2007.

[3] Eduardo Azevedo and Eric Budish. Strategyproofness in the large. Technical report, Working paper, 2012.

[4] Yaron Azrieli and Eran Shmaya. Lipschitz games. *Mathematics of Operations Research*, 38(2):350–357, 2013.

[5] Yakov Babichenko. Best-reply dynamics in large binary-choice anonymous games. *Games and Economic Behavior*, 81:130–144, 2013.

[6] Aaron Bodoh-Creed. Efficiency and information aggregation in large uniform-price auctions. *Journal of Economic Theory*, 148(6):2436–2466, 2013.

[7] Guilherme Carmona and Konrad Podczeck. Approximation and characterization of nash equilibria of large games. *University of Cambridge and Universitat Wien*, 2011.

[8] Guilherme Carmona and Konrad Podczeck. Ex-post stability of bayes–nash equilibria of large games. *Games and Economic Behavior*, 74(1):418–430, 2012.

[9] Joyee Deb and Ehud Kalai. Stability in large bayesian games with heterogeneous players. *Journal of Economic Theory*, 157:1041–1055, 2015.

[10] Timothy Feddersen and Wolfgang Pesendorfer. Voting behavior and information aggregation in elections with private information. *Econometrica: Journal of the Econometric Society*, pages 1029–1058, 1997.

[11] Drew Fudenberg and David K. Levine. Self-confirming equilibrium. *Econometrica: Journal of the Econometric Society*, pages 523–545, 1993.

[12] Drew Fudenberg and David K. Levine. *The theory of learning in games*, volume 2. MIT press, 1998.

[13] Ronen Gradwohl and Omer Reingold. Partial exposure in large games. *Games and Economic Behavior*, 68(2):602–613, 2010.

[14] Ronen Gradwohl and Omer Reingold. Fault tolerance in large games. *Games and Economic Behavior*, 2013.

[15] Edward J Green. Continuum and finite-player noncooperative models of competition. *Econometrica: Journal of the Econometric Society*, pages 975–993, 1984.

[16] Olivier Guéant, Jean-Michel Lasry, and Pierre-Louis Lions. Mean field games and applications. In *Paris-Princeton lectures on mathematical finance 2010*, pages 205–266. Springer, 2011.

[17] Philippe Jehiel and Frederic Koessler. Revisiting games of incomplete information with analogy-based expectations. *Games and Economic Behavior*, 62(2):533–557, 2008.

[18] Kenneth L Judd. The law of large numbers with a continuum of iid random variables. *Journal of Economic theory*, 35(1):19–25, 1985.

[19] Ehud Kalai. Large robust games. *Econometrica*, pages 1631–1665, 2004.

[20] Ehud Kalai and Ehud Lehrer. Rational learning leads to nash equilibrium. *Econometrica: Journal of the Econometric Society*, pages 1019–1045, 1993.

[21] Ehud Kalai and Ehud Lehrer. Subjective games and equilibria. *Games and Economic Behavior*, 8(1):123–163, 1995.

[22] Ehud Kalai and Eran Shmaya. A behavioral model for big games with uncertain fundamentals. https://www.researchgate.net/publication/280312972_A_behavioral_model_for_big_games_with_uncertain_fundamentals, 2015.

[23] M Ali Khan and Yeneng Sun. Non-cooperative games with many players. *Handbook of game theory with economic applications*, 3:1761–1808, 2002.

[24] Thodoris Lykouris, Vasilis Syrgkanis, and Eva Tardos. Learning and efficiency in games with dynamic population. *arXiv preprint arXiv:1505.00391*, 2015.

[25] George J. Mailath and Andrew Postlewaite. Asymmetric information bargaining problems with many agents. *The Review of Economic Studies*, 57(3):351–367, 1990.

[26] Eric Maskin and Jean Tirole. Markov perfect equilibrium: I. observable actions. *Journal of Economic Theory*, 100(2):191–219, 2001.

[27] R Preston McAfee. Mechanism design by competing sellers. *Econometrica: Journal of the econometric society*, pages 1281–1312, 1993.

[28] Roger B. Myerson. Population uncertainty and Poisson games. *Internat. J. Game Theory*, 27(3):375–392, 1998.

[29] Roger B. Myerson. Large Poisson games. *J. Econom. Theory*, 94(1):7–45, 2000.

[30] Abraham Neyman. The maximal variation of martingales of probabilities and repeated games with incomplete information. *Journal of Theoretical Probability*, 26(2):557–567, 2013.

[31] William Novshek and Hugo Sonnenschein. Cournot and Walras equilibrium. *J. Econom. Theory*, 19(2):223–266, 1978.

[32] Wojciech Olszewski and Ron Siegel. Large contests. Econometrica, forthcoming, 2013.

[33] A. Rustichini, M.A. Satterthwaite, and S.R. Williams. Convergence to efficiency in a simple market with incomplete information. *Econometrica: Journal of the Econometric Society*, pages 1041–1063, 1994.

[34] Hamid Sabourian. Anonymous repeated games with a large number of players and random outcomes. *Journal of Economic Theory*, 51(1):92–110, 1990.

[35] David Schmeidler. Equilibrium points of nonatomic games. *Journal of Statistical Physics*, 7(4):295–300, 1973.

[36] Sylvain Sorin. Merging, reputation, and repeated games with incomplete information. *Games and Economic Behavior*, 29(1):274–308, 1999.

[37] Sylvain Sorin and Cheng Wan Finite composite games: Equilibria and dynamics. *Journal of Dynamics and Games*, 3(1):101–120, 2016.

[38] Yeneng Sun. The exact law of large numbers via fubini extension and characterization of insurable risks. *Journal of Economic Theory*, 126(1):31–69, 2006.

Kellogg School of Management, Northwestern University, 2001 Sheridan Road, Evanston IL 60208

E-mail address: kalai@kellogg.northwestern.edu

School of Mathematical Sciences, Tel Aviv University and Kellogg School of Management, Northwestern University

E-mail address: e-shmaya@kellogg.northwestern.edu