The complexity of interacting automata*

Olivier Gossner[†] Penélope Hernández[‡] Ron Peretz[§]

July 24, 2015

Abstract

This paper studies the interaction of automata of size m. We characterise statistical properties satisfied by random plays generated by a correlated pair of automata with m states each. We show that in some respect the pair of automata can be identified to a more complex automaton of size comparable to $m \log m$. We investigate implications of these results on the correlated min-max value of repeated games played by automata.

1 Introduction

Automata are a central model in Game Theory when it comes to modelling agents with bounded cognitive abilities (see e.g. Aumann [1981], Abreu and Rubinstein [1988], Neyman [1997], Ben-Porath [1993]). But how complex are they? Early on, Neyman [1998] and Kalai and Stanford [1988] noted that if a repeated game strategy σ includes exactly m continuation strategies, then m is also the size of the smallest automaton that can implement σ. This gives foundation for automaton size as a measure for strategic complexity. In this paper, we take a different look at the same question by asking: how complex are phenomena that can be generated by one or several automata?

If we consider an isolated automaton of size m that generates a series of outputs in an alphabet A, then we know that this sequence is periodic of

*Olivier Gossner acknowledges financial support from Investissements d'Avenir (ANR-11-IDEX-0003/Labex Ecodec/ANR-11-LABX-0047). Penélope Hernández acknowledges financial support from the Spanish Economy and Competitiveness Ministry (ECO2013-46550-R) and Generalitat Valenciana (PROMETEOII/2014/054).

[†]École Polytechnique – CNRS and London School of Economics.

[‡]University of Valencia.

[§]London School of Economics. Email: r.perez@lse.ac.uk.

period at most m, and in fact, automata of size m can generate all such sequences. Since there are roughly speaking m^m such automata and $A^m \ll m^m$ such sequences, the automaton model is not economical when it comes to describing the behaviour of a single agent. The situation becomes more interesting when looking at interacting automata, which is the issue studied of this paper. Consider two agents interacting. Agent 1 is represented by an automaton σ^1 of size m and output space A^1 while agent 2 uses a repeated game strategy σ^2 with output space A^2. The pair (σ^1, σ^2) generates a sequence of outputs $a_1, a_2, \ldots, a_t$. An outside observer of the sequence who initially does not know (σ^1, σ^2), but knows that σ^1 is an automaton of size at most m forms beliefs at every stage on the next action of σ^1. Since σ^1 is chosen in a finite set, the outside observer is, as t grows, eventually able to predict every action of agent 1. Actually, we show a stronger result. Theorem 2.1, called the information constraint, holds: the total per stage entropy of the outside observer's prediction on agent 1's action is bounded by the logarithm of the number of strategies of agent 1.

We prove several results showing that the information constraint is tight. Since the information constraint is obtained by looking solely at the number of automata of a given size, these converse results show that automata are as rich as any other model with same number of strategies. In terms of complexity, automata generate phenomena of maximal randomness for the smallest possible number of strategies. In this sense, automata are a very economical model and do not have too many redundancies.

Our first converse result, Theorem 2.2, shows that, for any given distribution of beliefs $\mathcal{P}$ of an outside observer on the pair of actions of agents 1 and 2, there exist distributions over automata for agent 1 and strategies for agent 2 such that the average expected distribution of beliefs of an outside observer on both agents' actions between stages 1 and r is arbitrarily close to $\mathcal{P}$, as long as r is small enough so that the information constraint is satisfied. This result considers an outside observer who observes the actions of both agent 1 and 2, while the information constraint only considers agent 1's action. However, Theorem 2.2 particularized to distributions $\mathcal{P}$ in which the action of agent 2 is a function of the action of agent 1 shows that the information constraint provides a tight characterization of the beliefs an outside observer may have on the actions of an automaton that interacts with an outside strategy.

How about the complexity of two interacting automata, instead of one automaton interacting with an arbitrary strategy? Our second converse result shows that, when both agents are restricted to using automata of size at most m, the information constraint is tight up to a multiplicative constant that depends on the distribution of beliefs of the outside observer to be achieved. In particular, a pair of interacting automata of size m can generate a large subset of the set of sequences of period less than or equal to $m \log m$. We observe that $m \log m$ is much larger than the period m to which a single automaton is constrained. In terms of orders of magnitude, the logarithm of the number of pairs of automata is a constant times $m \log m$, and so is the logarithm of the number of sequences that can be generated. Therefore, this results shows that automata are not only a rich model when interacting with arbitrary strategies, but also when interactions are restricted within the model.

Therefore, it appears that while an automaton in isolation is unable to produce much complex phenomena, two interacting automata can. One may wonder what is the minimal complexity of a stream of inputs that an automaton requires in order to generate complex phenomena. Our Theorem 2.4 shows that an automaton of size m provided with a deterministic periodic stream of inputs whose period length is of order $\log m$ can generate random sequences of length of order $m \log m$.

The proof of our impossibility result relies on entropy techniques introduced by Lehrer [1988] and substantially developed by Neyman and Okada [1999, 2000, 2009] and then pursued by e.g. Gossner et al. [2006] and Peretz [2012, 2013]. The converse results rely on extensions of the construction of automata whose states are elements of De Bruijn sequences [Gossner and Hernández, 2003, 2006]. In contrast to Gossner and Hernández [2003, 2006], we are not interested here in constructing an automaton that matches a given sequence, but in the construction of a random automaton or a pair of automata that achieve desired statistical properties. We rely on a result proven independently by Ornstein [1970] and Shapira [2007] showing that the desired property is satisfied when the entropy of the generated sequence is sufficiently close to the target.

The min-max values of two-player repeated games played by automata [Ben-Porath, 1993, Neyman, 1985, 1998, Kalai, 1990, Neyman and Spencer, 2010, Neyman, 2008] or by strategies of bounded recall [Lehrer, 1988, Peretz, 2012] is relatively well understood. However, little is known about min-max values of games played with three or more finite automata, while a few results have been obtained on three players with bounded recall [see Bavly and Neyman, 2014, Peretz, 2013]. Our results on the set of random plays generated by finite automata have natural consequences on the min-max values of repeated games played by three automata, which was one of the question which originally motivated our study.

The paper is organised as follows. Section 2 presents the model and the main results, which are then proven in Section 3. Section 4 examines the min-max values of repeated games played by finite automata.

2 Model and statement of the results

Let A^1 and A^2 be finite action sets for agents 1 and 2, and $A = A^1 \times A^2$. It is assumed throughout that $|A^1|, |A^2| \geq 2$. A (pure, reduced) strategy for agent $i \in \{1, 2\}$ in the repeated interaction is a function from $\cup_{t\geq 0}(A^{-i})^t$ to A^i, and the set of strategies for agent i is denoted Σ^i. A pair of strategies $(\sigma^1, \sigma^2) \in \Sigma^1 \times \Sigma^2$ induces a play $(a_1, \ldots, a_t, \ldots) \in A^{\mathbb{N}}$, where $a_t = (a_t^1, a_t^2)$ is defined recursively by $a_t^i = \sigma^i(a_1^{-i}, \ldots, a_{t-1}^{-i})$.

A (reduced) automaton of size m for agent i is given by a state space S of cardinality m, an initial state q_0, an action function $f\colon S \to A^i$, and a transition function $h\colon S \times A^{-i} \to S$. An automaton for agent i and a sequence $(a_1^{-i}, \ldots, a_t^{-i}) \in (A^{-i})^t$ induce a sequence of states $q_0, q_1, \ldots, q_t$ given recursively by

$$\begin{aligned} &q_0, \\ &q_1 = h(q_0, a_1^{-i}), \\ &\quad\vdots \\ &q_t = h(q_{t-1}, a_t^{-i}). \end{aligned}$$

Thus, an automaton defines a strategy σ^i by

$$\sigma^i(a_1^{-i}, \ldots, a_t^{-i}) = f(q_t).$$

We let $\Sigma^i(m)$ be the set of strategies of agent i induced by automata of size m.

For a compact metric space X, the set of probability measures on X endowed with the weak-* topology is a compact metric space which is denoted $\Delta(X)$. We abbreviate $\Delta(\Delta(X))$ by $\Delta\Delta(X)$. It is noted that throughout the paper we mainly consider finitely supported probability measures, so the weak-* topology does not play a crucial role.

Shannon's entropy of a probability measure over a finite space, $Q \in \Delta(X)$, is the quantity

$$H(Q) = -\sum_{x\in X} Q(x)\log\left(Q(x)\right),$$

where $\log = \log_2$ and $0 \log 0 = 0$ by continuity. The entropy of a random variable $x \in X$, denoted $H(x)$, is the entropy of its distribution. If $x \in X$ and $y \in X$ are two finitely valued random variables $H(x, y)$ denotes the entropy to the random variable $(x, y) \in X \times Y$. The *entropy of x conditional on y* is defined by the chain rule $H(x|y) = H(x, y) - H(y)$. The *inequality of conditional entropy* asserts that $H(x) \geq H(x|y) \geq 0$, where the first inequality is equality if and only if x and y are independent. By taking π

to be a uniform random permutation on X, the inequality of conditional entropy implies that $H(x) = H(\pi(x)|\pi) \leq H(\pi(x)) = \log|X|$, with equality if and only if x has the uniform distribution. The *mutual information* of x and y is defined as $I(x;y) = H(x) - H(x|y) = I(y;x)$. By the inequality of conditional entropy, $I(x;y) \geq 0$ with equality if and only if x and y are independent.

For $\mathcal{P} \in \Delta\Delta(X)$, we denote the expected entropy of $\mathcal{P}$ by

$$\bar{H}(\mathcal{P}) = \int H(Q)\,\mathrm{d}\mathcal{P}(Q).$$

A correlated strategy is a probability distribution $\tau \in \Delta(\Sigma^1 \times \Sigma^2)$. Such τ induces a probability distribution P_τ over the set of infinite histories $A^{\mathbb{N}}$. For any finite history $(a_1, \dots a_{t-1})$ such that $P_\tau(a_1, \dots, a_{t-1}) > 0$ we let

$$\begin{aligned} p_{t,\tau}(a_1, \dots, a_{t-1}) &= P_\tau(a_t|a_1, \dots, a_{t-1}), \\ p^i_{t,\tau}(a_1, \dots, a_{t-1}) &= P_\tau(a^i_t|a_1, \dots, a_{t-1}). \end{aligned}$$

That is, $p_{t,\tau}(a_1, \dots, a_{t-1})$ and $p^i_{t,\tau}(a_1, \dots, a_{t-1})$ represent the beliefs of an outside observer on a_t and on a^i_t given $a_1, \dots, a_{t-1}$. We denote by $P_t(\tau) \in \Delta\Delta(A)$ and $P^i_t(\tau) \in \Delta\Delta(A^i)$ the laws of the random variables $p_{t,\tau}(a_1, \dots, a_{t-1})$ and of $p^i_{t,\tau}(a_1, \dots, a_{t-1})$ when $(a_1, \dots, a_{t-1})$ is drawn according to P_τ.

Finally, for $r \geq 0$, we define the r-stage *expected empirical frequency of beliefs* [Gossner and Tomala, 2006, 2007] induced by τ, $\mathcal{P}_r(\tau) \in \Delta\Delta(A)$ and $\mathcal{P}^i_r(\tau) \in \Delta\Delta(A^i)$, by

$$\begin{aligned} \mathcal{P}_r(\tau) &= \frac{1}{r}\sum_{t=1}^{r} P_t, \\ \mathcal{P}^i_r(\tau) &= \frac{1}{r}\sum_{t=1}^{r} P^i_t. \end{aligned}$$

Note that we can express $\bar{H}(\mathcal{P}_r(\tau)) = \frac{1}{r}\sum_{t=1}^{r} H(a_t|\bar{a}_{t-1})$ and $\bar{H}(\mathcal{P}^i_r(\tau)) = \frac{1}{r}\sum_{t=1}^{r} H(a^i_t|\bar{a}_{t-1})$.

Clearly, the sets of all possible $\mathcal{P}_r(\tau)$ and $\mathcal{P}^i_r(\tau)$ when $r \geq 0$ and τ is an unrestricted correlated strategy are dense in $\Delta\Delta(A)$ and $\Delta\Delta(A^i)$ (since any $\mathcal{P} \in \Delta\Delta(A)$ can be approximated by an average of Dirac measures $\mathcal{P}_r = \frac{1}{r}(\delta_{Q_1} + \cdots \delta_{Q_r})$, and such $\mathcal{P}_r$ can be implemented by correlated strategies that play Q_i at each stage $i = 1, \dots, r$ independently of the history up to stage i). Our aim is to investigate what restrictions are imposed on these sets when either or both strategies are taken among subclasses of bounded complexity.

By considering the cardinality of $\Sigma^i(m)$ we obtain the following constraint on the possible $\mathcal{P}^i_r(\tau)$ when σ^i is restricted to take values in $\Sigma^i(m)$.

Theorem 2.1 (information constraint)**.** *For every $m \in \mathbb{N}$, $\tau \in \Delta(\Sigma^1(m) \times \Sigma^2)$, and $r \in \mathbb{N}$,*

$$\frac{\bar{H}(\mathcal{P}^1_r(\tau))}{|A^2| - 1} \leq \frac{m \log m + o(m \log m)}{r}.$$

Our main effort is dedicated to obtaining converse results showing that the bounds provided by the information constraint are tight. In order to state possibility results precisely, it will be convenient to consider a metric that induces the weak-* topology on $\Delta\Delta(A)$. The Wasserstein metric $\bar{d}$ serves this purpose. Let $\langle X, d \rangle$ be a metric space, e.g., $\langle \Delta(A), \|\cdot\|_1 \rangle$. For $P, P' \in \Delta(X)$, the $\bar{d}$ distance between P and P' is defined by

$$\bar{d}(P, P') = \inf_{Q \in \Delta(X \times X), Q^1 = P, Q^2 = P'} \int d(x, y) \, \mathrm{d}Q(x, y),$$

where Q^1 and Q^2 denote the first and second marginals of Q respectively.

We are now in position to state our first converse result to Theorem 2.1.

Theorem 2.2. *For every $\mathcal{P} \in \Delta\Delta(A)$ and $\epsilon > 0$ there exists $r_0 \in \mathbb{N}$ such that for every $r \geq r_0$ there exists $m \in \mathbb{N}$ and $\tau \in \Delta(\Sigma^1(m) \times \Sigma^2)$ such that*

$$\frac{\bar{H}(\mathcal{P})}{|A^2| - 1} + \epsilon \geq \frac{m \log m}{r}$$

and

$$\bar{d}(\mathcal{P}, \mathcal{P}_r(\tau)) < \epsilon.$$

Theorem 2.2 shows that the information constraint of Theorem 2.1 is tight when particularized to distributions $\mathcal{P}$ such that $\bar{H}(\mathcal{P}) = \bar{H}(\mathcal{P}^1)$, i.e., $H(a^2|a^1) = 0$ ($\mathcal{P}$ a.s.). In terms of expected distributions of beliefs, the set of strategies $\Sigma^1(m)$ is as rich as any other set of strategies of the same cardinality can be. A natural question to ask is whether this richness still holds when the set in which σ^2 is chosen is also restricted.

In order to address this question we restrict both agents to strategies implementable through finite automata of size m. The information constraint shows that for every correlated $\tau \in \Delta(\Sigma^1(m) \times \Sigma_2(m))$,

$$\max_{i \in \{1,2\}} \left\{ \frac{\bar{H}(\mathcal{P}^{-i}_r(\tau))}{(|A^i| - 1)} \right\} \leq \frac{m \log m + o(m \log m)}{r}.$$

The next theorem asserts that the above bound is tight up to a constant multiplier that depends on $\mathcal{P}$. It shows that the information constraint is tight even when agent 2's strategy comes from the same (space) complexity class as that of agent 1.

Theorem 2.3. *For every $\mathcal{P} \in \Delta\Delta(A)$ there exists $C(\mathcal{P}) > 0$ such that for every $\epsilon > 0$ there exists $r_0 \in \mathbb{N}$ such that for every $r \geq r_0$ there exist $m \in \mathbb{N}$ and $\tau \in \Delta(\Sigma^1(m) \times \Sigma^2(m))$ such that*

$$C(\mathcal{P}) + \epsilon \geq \frac{m \log m}{r}$$

and

$$\bar{d}(\mathcal{P}, \mathcal{P}_r(\tau)) < \epsilon.$$

Note that $C(\mathcal{P})$ does not depend on ϵ. It does, however, depend on $\mathcal{P}$. Whether one could replace $C(\mathcal{P})$ by a constant that does not depend on $\mathcal{P}$ (but rather just on A) is unknown. The information constraint implies that we cannot replace $C(\mathcal{P})$ by a constant $C < \max_{i\in\{1,2\}} \left\{ \frac{\log(|A^{-i}|)}{|A^i|-1} \right\}$.

The next theorem provides a sense in which the information constraint is still tight under two further restrictions. First, the strategy of agent 2 is pure, i.e. constant on the support of τ. In particular, the strategies of 1 and 2 are now independent. Second, agent 2 uses a strategy in a much smaller space than $\Sigma^2(m)$, namely the space of periodic sequences[1] of period at most $\mathcal{O}(\log m)$. On the other hand, Theorem 2.4 considers only the distribution of predictions of agent 1, and not of both 1 and 2's actions. In this result, the sequence of actions of agent 2 can be considered as a "source" which, albeit deterministic, allows agent 1 to look more unpredictable from the point of view of an outside observer.

To fix notations, for $x > 0$ we denote by $A_p^2(x)$ the set of sequences of period at most x and identify each sequence $(a_t^2) \in A_p^2(x)$ with the strategy σ^2 given by $\sigma^2(a_1^1, \ldots, a_{t-1}^1) = a_t^2$ for every $(a_1^1, \ldots, a_{t-1}^1)$.

Theorem 2.4. *For every $\mathcal{P}^1 \in \Delta\Delta(A^1)$ there exist $C = C(\mathcal{P}^1) > 0$ such that for every $\epsilon > 0$ there exists $r_0 \in \mathbb{N}$ such that for every $r \geq r_0$ there exist $m \in \mathbb{N}$ and $\tau \in \Delta(\Sigma^1(m)) \times A_p^2(C \log m)$ such that*

$$\frac{\bar{H}(\mathcal{P}^1)}{|A^2| - 1} + \epsilon \geq \frac{m \log m}{r}$$

and

$$\bar{d}(\mathcal{P}^1, \mathcal{P}_r^1(\tau)) < \epsilon.$$

[1] Or, equivalently, in the space of oblivious automata of size at most $\mathcal{O}(\log m)$, i.e. those that ignore the actions of the other agent.

3 Proofs of the Main Results

3.1 Information constraint

Theorem 2.1 is a consequence of the following upper bound on the number of finite automata.

Lemma 3.1.

$$|\Sigma_i(m)| \leq m^{(|A^{-i}|-1)m+o(m)}.$$

Proof. All elements of $\Sigma_i(m)$ are induced by automata with state space $\{1,\ldots,m\}$, thus are described by $q_0 \in \{1,\ldots,m\}$, $f\colon \{1,\ldots,m\} \to A^i$, and $h\colon \{1,\ldots,m\} \times A^{-i} \to \{1,\ldots,m\}$. This gives $m|A^i|^m m^{|A^{-i}|m}$ different descriptions. Since strategies are invariant to permutations of states in $\{1,\ldots,m\}$, by Stirling approximation, we have:

$$|\Sigma_i(m)| \leq |A^i|^m m^{|A^{-i}|m}/(m-1)! = m^{(|A^{-i}|-1)m+o(m)}.$$

$\square$

To complete the proof of Theorem 2.1 we rely on a slight generalization of a result due to Neyman and Okada [Neyman and Okada, 1999, Section 5] showing that the strategic entropy of a strategy is at most its entropy:

Lemma 3.2. *Let $\tau \in \Delta(\Sigma'^1 \times \Sigma^2)$, where Σ'^1 is any finite set and $r \geq 1$, then,*

$$\bar{H}(\mathcal{P}^1_r(\tau)) \leq \frac{1}{r}\log\left|\Sigma'^1\right|.$$

Proof. Let $\tau \in \Delta(\Sigma'^1 \times \Sigma_2)$ be a correlated strategy. Denote by $\sigma = (\sigma^1, \sigma^2)$ a random variable with values in $\Sigma'^1 \times \Sigma^2$ with distribution τ and by $a_1,\ldots,a_r$, where $a_t = (a^1_t, a^2_t)$ the play induced by σ. We abbreviate $\bar{a}^i_t := a^i_1,\ldots,a^i_t$ $(i = 1,2)$, and $\bar{a}_t := (\bar{a}^1_t, \bar{a}^2_t)$. By the chain rule of entropy, the inequality of conditional entropy, and the fact that $\bar{a}^1_t$ is a function of $\bar{a}^2_{t-1}$ and σ^1, we have:

$$\begin{aligned}
\sum_{t=1}^{r} H(a^1_t|\bar{a}_{t-1}) + H(a^2_t|\bar{a}_{t-1}, a^1_t) &= H(\bar{a}_r) \\
&\leq H(\sigma^1, \bar{a}^2_r) = H(\sigma^1) + H(\bar{a}^2_r|\sigma^1) \\
&= H(\sigma^1) + \sum_{t=1}^{r} H(a^2_t|\bar{a}_{t-1}, a^1_t, \sigma^1).
\end{aligned}$$

And therefore:

$$\bar{H}(\mathcal{P}^1_r(\tau)) \;=\; \frac{1}{r}\sum_{t=1}^{r} H(a^1_t|\bar{a}_{t-1}) \le \frac{1}{r}H(\sigma^1) \le \frac{1}{r}\log\left|\Sigma'^1\right|.$$

□

Theorem 2.1 is an immediate consequence of Lemmata 3.1 and 3.2.

3.2 Concatenable strategy pairs

The proofs of Theorems 2.2, 2.3, 2.4 rely on the construction of adequate automata. These constructions are built in blocks, and larger automata are obtained by concatenation of smaller ones. We introduce a condition that ensures that such concatenations can be implemented by automata of size equal to the sum of the smaller ones.

We say that a pair of automata (σ^1, σ^2) is *r-concatenable* if there exists a pair of transitions (s^1, a^2) for σ^1 and, (s^2, a^1) for σ^2 such that, in the play induced by σ^1 and σ^2, the first hitting time of both (s^1, a^2) and of (s^2, a^1) is r. We denote by $(\Sigma^1(m^1) \times \Sigma^2(m^2))^r$ the subset of $\Sigma^1(m^1) \times \Sigma^2(m^2)$ consisting of all r-concatenable pairs of automata σ^1 of size m^1 and σ^2 of size m^2.

A pair given by an automaton σ^1 of agent 1 and a strategy σ^2 of agent 2 is *r-concatenable* if there exists a transition (s^1, a^2) such that, in the play induced by σ^1 and σ^2, the first hitting time of (s^1, a^2) is r. We denote by $(\Sigma^1(m^1) \times \Sigma^2)^r$ the subset of $\Sigma^1(m^1) \times \Sigma^2$ consisting of all r-concatenable pairs formed by an automaton of agent 1 of size m^1 and a strategy of agent 2.

A pair given by an automaton σ^1 of agent 1 and an r-periodic sequence of actions $\bar{a}^2 = a^2_1, \ldots, a^2_r, a^2_1 \ldots$ of agent 2 is *r-concatenable* if there exists a state s^1 of σ^1 such that, in the play induced by σ^1 and $\bar{a}^2$, the first hitting time of (s^1, a^2_r) is r. We denote by $(\Sigma^1(m^1) \times A^2_p(m^2))^r$ the subset of $\Sigma^1(m^1) \times A^2_p(m^2)$ consisting of all r-concatenable pairs formed by an automaton of agent 1 of size m^1 and a periodic sequence of actions of agent 2 whose period is at most m^2 (in particular, the period must divide r).

An r'-concatenable pair of an automaton and another strategy can be concatenated to an r''-concatenable such pair by redirecting the last transition of the former automaton to the initial state of the latter automaton. The result is an $r' + r''$-concatenable pair whose automaton size is the sum of the sizes of the two automata. This idea is formally expressed by the following three lemmata.

Lemma 3.3 (Concatenation Lemma 1)**.** *Let* $(\sigma_k^1, \sigma_k^2)_{1\le k\le n}$ *be a family of strategy pairs such that for every* k*,* $(\sigma_k^1, \sigma_k^2) \in (\Sigma^1(m_k^1) \times \Sigma^2(m_k^2))^{r_k}$*. Let* $r = \sum_k r_k$*,* $m^1 = \sum_k m_k^1$*, and* $m^2 = \sum_k m_k^2$*. Then, there exists a strategy pair* $(\sigma^1, \sigma^2) \in (\Sigma^1(m^1) \times \Sigma^2(m^2))^r$ *such that the play induced by* (σ^1, σ^2) *is the concatenation of the plays induced by* $(\sigma_k^1, \sigma_k^2)_{1\le k\le n}$*.*

Lemma 3.4 (Concatenation Lemma 2)**.** *Let* $(\sigma_k^1, \sigma_k^2)_{1\le k\le n}$ *be a family of strategy pairs such that for every* k*,* $(\sigma_k^1, \sigma_k^2) \in (\Sigma^1(m_k^1) \times \Sigma^2)^{r_k}$*. Let* $r = \sum_k r_k$*, and* $m^1 = \sum_k m_k^1$*. Then, there exists a strategy pair* $(\sigma^1, \sigma^2) \in (\Sigma^1(m^1) \times \Sigma^2)^r$ *such that the play induced by* (σ^1, σ^2) *is the concatenation of the plays induced by* $(\sigma_k^1, \sigma_k^2)_{1\le k\le n}$*.*

Lemma 3.5 (Concatenation Lemma 3)**.** *Let* $(\sigma_k^1)_{1\le k\le n}$ *be a family of strategies for agent 1 and* $\bar{a}^2$ *an* m^2*-periodic sequence for a agent 2. Suppose that for every* k*,* $(\sigma_k^1, \bar{a}^2) \in (\Sigma^1(m_k^1) \times A_p^2(m^2))^{r_k}$*. Let* $r = \sum_k r_k$*,* $m^1 = \sum_k m_k^1$*. Then, there exists a strategy* σ^1 *for agent 1 such that* $(\sigma^1, \bar{a}) \in (\Sigma^1(m^1) \times A_p^2(m^2))^r$ *and the play induced by* $(\sigma^1, \bar{a}^2)$ *is the concatenation of the plays induced by* $(\sigma_k^1, \bar{a}^2)_{1\le k\le n}$*.*

The next lemma, Lemma 3.6, utilises the Concatenation Lemmata in order to reduce Theorems 2.2, 2.3, and 2.4, to considering only r's in a sequence that does not increase too fast. Formally, Lemma 3.6 consists of three lemmata, referring to the three theorems. Since the three lemmata are very similar both in their formulation and in their proofs, we state and prove them as one result.

Lemma 3.6. *Let* $\{m_k\}$*,* $\{r_k\}_{k=1}^\infty$ *be sequences of natural numbers, and let* $\tau_k \in \Delta(\Sigma^1(m_k) \times \Sigma^2)^{r_k}$ *(respectively,* $\Delta(\Sigma^1(m_k) \times \Sigma^2(m_k))^{r_k}$*, or* $\Delta(\Sigma^1(m_k) \times A_p(f(m_k)))^{r_k}$ *for some non-decreasing* $f\colon \mathbb{N} \to \mathbb{R}_+$ *and with pure marginal* τ^2*).*

Suppose $\lim_{k\to\infty} \mathcal{P}_{r_k}(\tau_k) = \mathcal{P}$ *exists, and*

(i) $\sup_k r_{k+1} = \infty$*,*

(ii) $\sup_k \frac{r_{k+1}}{r_k} < \infty$*.*

Then, for every $\epsilon > 0$ *there exists* $r_0 \in \mathbb{N}$ *such that for every* $r > r_0$ *there exist* $m > 0$*,* $r' \le r$ *and* $\tau \in \Delta(\Sigma^1(m) \times \Sigma^2)^{r'}$ *(respectively,* $\Delta(\Sigma^1(m) \times \Sigma^2(m))^{r'}$*, or* $\Delta(\Sigma^1(m) \times A_p^2(f(m)))^{r'}$ *with* τ^2 *pure), such that*

$$\begin{aligned}
&\frac{r'}{r} > 1 - \epsilon,\\
&\frac{m \log m}{r'} < \limsup_k \frac{m_k \log m_k}{r_k} + \epsilon,\\
&\bar{d}(\mathcal{P}, \mathcal{P}_{r'}(\tau)) < \epsilon.
\end{aligned}$$

Proof. Assume w.l.o.g. that $\{r_k\}$ is increasing (otherwise consider an increasing sub-sequence). Let $1 > \epsilon > 0$ and $k_0 \in \mathbb{N}$ (later, k_0 will be assumed to be sufficiently large). Let $r_0 = \lceil \epsilon^{-1} r_{k_0} \rceil$. For $r \geq r_0$, let k be the largest integer such that $r_k \leq \epsilon r$. Let $C = \sup_k \frac{r_{k+1}}{r_k}$. Let $r' = r_k \left\lfloor \frac{r}{r_k} \right\rfloor$. It follows that

$$\frac{r'}{r} > 1 - \epsilon.$$

Let τ_k be the strategy assumed by the lemma. Set τ to be the concatenation of $\frac{r'}{r_k}$ independent copies of τ_k. It follows that

$$\mathcal{P}_{r'}(\tau) = \mathcal{P}_{r_k}(\tau_k) \xrightarrow[k_0 \to \infty]{} \mathcal{P}.$$

Let $m = \frac{r'}{r_k} m_k$. The concatenation lemma ensures that $\tau \in \Delta(\Sigma^1(m) \times \Sigma^2)^{r'}$ (respectively, $\Delta(\Sigma^1(m) \times \Sigma^2(m))^{r'}$, or $\Delta(\Sigma^1(m) \times A^2_p(f(m)))^{r'}$ with τ^2 pure).

It remains to verify that $\frac{m \log m}{r'} < \limsup_k \frac{m_k \log m_k}{r_k} + \epsilon$, for an appropriate choice of k_0. Since $C r_k \geq r_{k+1} > \epsilon r \geq \epsilon r'$, we have $\frac{r'}{r_k} < \epsilon^{-1} C$. Therefore,

$$\frac{m \log m}{r'} < \frac{m_k \log m_k}{r_k} + \frac{m_k}{r_k} \log(\epsilon^{-1} C).$$

Assuming $\limsup_k \frac{m_k \log m_k}{r_k} < \infty$, for any k large enough, we have $\frac{m_k \log m_k}{r_k} < \limsup_k \frac{m_k \log m_k}{r_k} + \frac{\epsilon}{2}$ and $\frac{m_k}{r_k} \log(\epsilon^{-1} C) < \frac{\epsilon}{2}$. □

3.3 Building blocks

We define a few building blocks that will be used in the construction of the strategies required for Theorems 2.2, 2.3, and 2.4 . For $l \in \mathbb{N}$ and $x \in (A)^l$, the *empirical frequency*, $\mathrm{emp}(x) \in \Delta(A)$, is defined by

$$\mathrm{emp}(x)(a) = \frac{1}{l} \left| \{ t \in [l] : x_t = a \} \right|,$$

where $[n] := \{1, \ldots, n\}$.

For a rational distribution Q over A, we let $T^Q(l)$ be the set of sequences of elements of A of length l with empirical frequency Q. We say that l is a common denominator of Q if $T^Q(l) \neq \emptyset$. We use the fact that if l_0 is a common denominator of Q, then [Cover and Thomas, 2006, Chapter 11]

$$\lim_{u \to \infty} \frac{\log |T^Q(u l_0)|}{u l_0} = H(Q). \tag{3.1}$$

Throughout, an *alphabet* is a set that contains at least two distinct elements. A *de Bruijn sequence of order k over a finite alphabet B* is a $|B|^k$-periodic sequence of B symbols $x_1, x_2, \ldots$, such that for every $(b_1, \ldots, b_k) \in B^k$ there exists a unique $1 \le t \le |B|^k$ such that $(b_1, \ldots, b_k) = (x_t, \ldots, x_{t+k-1})$. The existence of de Bruijn sequences was shown, e.g., by de Bruijn [1946] through the existence of Eulerian cycles in de Bruijn's graph.

The empirical frequency of de Bruijn sequences is always uniform. In order to obtain sequences with similar properties, but whose empirical frequency is arbitrary we define de Bruijn sequences over compound alphabets. For $l \ge 1$, let Y be a subset of B^l with at least two distinct element, called the set of words. A *compound de Bruijn sequence of order k over Y* is an $l|Y|^k$-periodic sequence of B elements obtained by the concatenation of the $|Y|^k$ words forming a de Bruijn sequence of order k over the alphabet Y. It is, therefore, an $l|Y|^k$-periodic sequence $x_1, x_2, \ldots$ of elements in B such that for every sequence $(y_1, \ldots, y_k) \in Y^k$, there exists a unique $1 \le j \le |Y|^k$ such that $(y_1 \cdots y_k) = (x_{jl+1}, \ldots, x_{jl+kl})$ (where uv denotes the concatenation of u and v). The item $x_t \in B$ is the *t-th element*, while we call $y_j = (x_{jl+1}, \ldots, x_{jl+l}) \in B^l$ the *j-th word*, and $(y_j, \ldots, y_{j+k-1}) \in B^{kl}$ the *j-th block*. When $Y = T^Q(l)$, the compound de Bruijn has empirical frequency Q.

The following definition provides a simple upper bound on the amount of information needed to describe a finite string of symbols given another string. We define a variant of Levenshtein's edit distance. For a set X we let $X^{<\infty}$ be the set of finite sequences of elements of X. Given a finite alphabet B, the *simple edit operations* on $B^{<\infty}$ are the following $2|B| + 2$ operations: appending one symbol at the end or the beginning of a string, and deleting the first or last element of a string. We define the *simple edit distance* between two words $w, u \in B^{<\infty}$, denoted $e(w, u)$, to be the minimal number of simple edit operations needed in order to transform w to u. As usual, the distance between a string w and a set of strings K, $e(w, K)$, is the minimal distance between w and some $u \in K$. Note that, up to a multiplier of $\log(2|B| + 2)$, the simple edit distance is an upper bound on the amount of information content of one string given another string.

The strategies required for Theorems 2.2, 2.3, and 2.4 all have a common structure. We next describe the common structure of the random plays induced by these strategies.

Definition 3.7 (Random Enumeration Scheme)**.** Let $Y \subset A^l$, for some $l \ge 1$. A *random enumeration scheme over Y* is a sequence of tuples $\langle L_k, G_k \rangle_{k=1}^{\infty}$, where for every k, $L_k = (w_1, \ldots, w_{|L_k|})$ is a finite sequence of strings in $A^{<\infty}$ and G_k is a group of permutation on $[|L_k|]$. In addition, there is a constant

C (that does not depend on k) such that

(i) $|\{s \in [|L_k|] : w_s = w_t\}| \leq C \qquad \forall t \in [|L_k|]$,

(ii) $e(w_t, Y^k) \leq C \qquad \forall t \in [|L_k|]$,

(iii) $|Y|^k / |L_k| \leq C$,

(iv) $\left(\frac{|L_k|!}{|G_k|}\right)^{\frac{1}{|L_k|}} \leq C$,

for all $k \in \mathbb{N}$.

Each element of the group G_k acts on the sequence L_k by re-ordering its elements. That is, every $\pi \in G_k$ transforms L_k into

$$(w_{\pi(1)}, \ldots, w_{\pi(|L_k|)}).$$

When $\pi \in G_k$ is random and uniformly distributed, the corresponding random sequence $\bar{a}(k) := w_{\pi(1)} \cdots w_{\pi(|L_k|)}$ is called the *induced random play.*

Informally, the idea is that L_k is approximately an enumeration of Y^k and G_k contains a substantial portion of all the permutations of L_k, so that the induced random play $\bar{a}(k)$ approximates a sequence of i.i.d. random variables drawn uniformly from Y. The two processes are similar in two ways: they produce similar sequences and they have similar entropy rates. Since for each element w_t of L_k, $e(w_t, Y^k) \leq C$, w_t is almost entirely composed of a concatenation of strings from Y, and hence so is $\bar{a}(k)$. The next lemma says that the entropy rate of $\bar{a}(k)$ is also similar to that of a uniformly distributed independent sequence of Y-valued random variables.

Lemma 3.8. *Let $\langle L_k, G_k \rangle_{k=1}^{\infty}$ be a random enumeration scheme over a compound alphabet $Y \subset A^l$. Let r_k be the total length of the random play $\bar{a}(k)$. Then,*

$$\liminf_{k \to \infty} \frac{1}{r_k} H(\bar{a}(k)) \geq \frac{\log |Y|}{l}.$$

Proof. Let π be a uniformly random permutation from G_k, and $\bar{a}(k) = a_1, \ldots, a_{r_k}$ the corresponding induced play. We first estimate $\frac{1}{r_k} H(\pi)$ and then compare it to $\frac{1}{r_k} H(\bar{a}(k))$.

Let C be as in Definition 3.7. Since the length of any word in L_k is at most $kl + C$, we have

$$r_k \leq |L_k|(kl + C).$$

Since $|G_k| \geq C^{-|L_k|}|L_k|!$ and $|L_k| \geq C^{-1}|Y|^k$,

$$\frac{1}{r_k}H(\pi) = \frac{1}{r_k}\log(|G_k|) \geq \frac{\log(|L_k|!)}{r_k} - \frac{|L_k|\log C}{r_k} \geq \frac{\log(|L_k|!)}{|L_k|(kl+C)} - \frac{\log C}{kl+C}$$
$$\geq \frac{\log(|L_k|)(1-o(1))}{kl+C} \geq \frac{k\log(|Y|)(1-o(1))}{kl+C} \geq \frac{\log|Y|}{l}(1-o(1)).$$

We now compare $\frac{1}{r_k}H(\bar{a}(k))$ to $\frac{1}{r_k}H(\pi)$ by coupling $\bar{a}(k)$ with additional random variables of low entropy, such that π can be read off from $\bar{a}(k)$ and the additional random variables. Recall that

$$\bar{a}(k) = w_{\pi(1)}w_{\pi(2)}\cdots w_{\pi(|L_k|)},$$

where $L_k = (w_1, w_2, \ldots, w_{|L_k|})$. Denote by $|w|$ the length of the word w. For each $j \in [|L_k|]$, let

$$b_j = |w_{\pi(j)}|, \text{ and}$$
$$c_j = \left|\left\{i < \pi(j) : w_i = w_{\pi(j)}\right\}\right|.$$

Note that from $\bar{a}$ and $b_1, \ldots, b_{|L_k|}$ one can read off the words $w_{\pi(1)}, \ldots, w_{\pi(|L_k|)}$, and further from $c_1, \ldots, c_{|L_k|}$ the permutation π itself. The values of each b_j range between $kl - C$ and $kl + C$, and the values of each c_j are in $[C]$. It follows that

$$\frac{1}{r_k}H(\bar{a}(k)) \geq \frac{1}{r_k}H(\pi) - \frac{1}{r_k}\left[H(b_1, \ldots, b_{|L_k|}) + H(c_1, \ldots, c_{|L_k|})\right]$$
$$\geq \frac{\log|Y|}{l}(1-o(1)) - \frac{|L_k|}{r_k}\log(2C^2) \geq \frac{\log|Y|}{l}(1-o(1)).$$

□

We now describe a framework within which one can construct schemes of automata that are naturally associated with certain random enumeration schemes.

Definition 3.9 (Random Automaton Scheme)**.** A *random automaton scheme for agent i* is a tuple $\Xi = \left\langle \mathcal{P}, l, \left\{x(k), L_k, (s_t(k), z_t(k), a_t^{-i}(k))_{t=1}^{|L_k|}, G_k\right\}_{k=1}^{\infty}\right\rangle$ where

- $\mathcal{P} = \sum_j q_j \delta_{Q_j} \in \Delta\Delta(A)$ is a finitely supported rational distribution over rational beliefs.
- $l \in \mathbb{N}$ is such that each $l_j := q_j l$ is a common denominator of Q_j (in particular, l_j is an integer).

Let $Y = \bigtimes_j T^{Q_j}(l_j)$. For every $k \in \mathbb{N}$,

- $x = x(k) = (x_1, \dots, x_{m_k})$ is a compound de Bruijn sequence of order k over Y, where $m_k = l|Y|^k$.
- $L_k = (w_1, \dots, w_{|L_k|})$ is a sequence of A-words such that $\sup_k m_k/|L_k| < \infty$.
- For $t = 1, \dots, |L_k|$,
  - $s_t = s_t(k) \in [m_k]$,
  - $z_t = z_t(k) \in [lk] \setminus [l(k-1)]$, and
  - $a_t^{-i} = a_t^{-i}(k) \in A^{-i} \setminus \{x_t^{-i}\}$,

  such that
  $$w_t = x_{s_t - z_t} \cdots x_{s_t - 1}\,(x^i_{s_t}, a_t^{-i}),$$
  and $(s_t, a_t^{-i}) \neq (s_{t'}, a_{t'}^{-i})$ for all $t \neq t'$.
- G_k is a group of permutations on $[|L_k|]$, with $\sup_k \left(\frac{|L_k|!}{|G_k|}\right)^{\frac{1}{|L_k|}} < \infty$.

Definition 3.10. Given a random automaton scheme Ξ, an automaton $\sigma^i_\Xi(k) \in \Sigma^i(m_k)$ is defined by:

- the states space $[m_k]$;
- the initial state $s_1 - z_1 \bmod m_k$;
- the action function $f(s) = x^i_s$;
- the transition function
$$h(s, a^{-i}) = \begin{cases} s + 1 \bmod m_k & \text{if } a^{-i} = x_s^{-i}, \\ s_{t+1} - z_{t+1} \bmod m_k & \text{if } s = s_t \text{ and } a^{-i} = a_t^{-i} \text{ for some } t \in [|L_k|], \\ \text{unspecified} & \text{otherwise.} \end{cases}$$

A transition of the first type is called a *+1 transition* and a transition of the second type a *jump transition*.

Note that the jump transitions are well defined, since it is assumed that the pairs (s_t, a_t^{-i}) are distinct.

The rest of Section 3.3 presents relevant properties of random enumeration schemes. For readability, the random automaton schemes are assumed to be for agent 1. Of course, similar properties hold for random enumeration schemes for agent 2.

With the notation of Definition 3.9, the concatenation of the elements of L_k is denoted

$$\bar{a}_\Xi(k) = w_1 w_2 \cdots w_{|L_k|}.$$

The next lemma ensures that the strategy $\sigma^1_\Xi(k)$ generates the desired play $\bar{a}_\Xi(k)$, that is has the necessary concanetability properties and provides the relevant upper bound on the state space cardinality.

Lemma 3.11. *Let Ξ be a random automaton scheme for agent 1, then:*

(i) the induced strategy $\sigma^1_\Xi(k)$ is consistent with the play $\bar{a}_\Xi(k)$,

(ii) the induced strategies $\sigma^1_\Xi(k)$ coupled with (any strategy consistent with) $\bar{a}_\Xi(k)$ is r_k-concatenable, and

(iii) $\limsup_{k\to\infty} \frac{m_k \log m_k}{r_k} \le \bar{H}(\mathcal{P}) \limsup_{k\to\infty} \frac{m_k}{|L_k|}$.

Proof. Part (i) follows from the definition of $\sigma^1_\Xi(k)$. In the beginning and after each jump transition, $\sigma^1_\Xi(k)$ is in state $s_t - z_t$. In the next z_t periods $\bar{a}_\Xi(k)$ prescribes the action profiles $(x_{s_t - z_t + j - 1} : j \in [z_t])$. Provided $\sigma^1_\Xi(k)$ is at state $s_t - z_t + j - 1$ it plays $x^1_{s_t - z_t + j - 1}$, as required. Provided agent 2, too, plays as required, $x^2_{s_t - z_t + j - 1}$, a “+1” transition to $s_t - z_t + j$ occurs. After z_t steps, $\sigma^1_\Xi(k)$ is at state s_t. At that time $\bar{a}_\Xi(k)$ prescribes the action profile $(x^1_{s_t}, a^2_t)$. If agent 2 plays a^2_t, then a jump transition to $s_{t+1} - z_t$ occurs, and the result follows by induction on t.

Part (ii) holds since the play ends at a jump transition and each jump transition occurs at most once by Definition 3.9.

For Part (iii) note that $r_k = \sum_{t=1}^{|L_k|} z_t + 1 \ge |L_k|(k-1)l$. Thus,

$$\frac{m_k \log m_k}{r_k} \le \frac{\log m_k}{kl} \frac{k}{k-1} \frac{m_k}{|L_k|}.$$

It remains to show that $\limsup \frac{\log m_k}{kl} \le \bar{H}(\mathcal{P})$. Recall that $m_k = l|Y|^k$. Since $\log|Y| = \sum_i \log |T^{Q_i}(l_i)| \le \sum_i l_i H(Q_i) = l\bar{H}(\mathcal{P})$,

$$\frac{\log m_k}{kl} = \frac{\log|Y|}{l} + \frac{\log l}{kl} \le \bar{H}(\mathcal{P}) + \frac{\log l}{kl} \xrightarrow[k\to\infty]{} \bar{H}(\mathcal{P}).$$

∎

A permutation $\pi \in G_k$ acts on the k-th component of Ξ by transforming $L_k = (w_1, \ldots, w_{|L_k|})$ into $(w_{\pi(1)}, \ldots, w_{\pi(|L_k|)})$, s_t into $s_{\pi(t)}$, z_t into $z_{\pi(t)}$, and a^2_t into $a^2_{\pi(t)}$ for $t = 1, \ldots, |L_k|$. This transformation defines another random automaton scheme denoted $\pi.\Xi$. The following lemma describes the relation between random automaton schemes and random enumeration schemes.

Lemma 3.12. *Let Ξ be a random automaton scheme for agent 1. Then:*

(i) $\langle L_k, G_k\rangle_{k=1}^{\infty}$ *is a random enumeration scheme over* Y.

(ii) $\sigma^1_{\pi.\Xi}$ *is consistent with the play* $w_{\pi(1)}w_{\pi(2)}\cdots w_{\pi(|L_k|)}$ *(where* $L_k = (w_1, w_2, \ldots, w_{|L_k|})$*).*

Proof. Part (ii). By the definition of $\pi.\Xi$,

$$\bar{a}_{\pi.\Xi}(k) = w_{\pi(1)}w_{\pi(2)}\cdots w_{\pi(|L_k|)}.$$

By Lemma 3.11, $\sigma^1_{\pi.\Xi}$ is consistent with $\bar{a}_{\pi.\Xi}(k)$.

Part (i). We need to verify (i)-(iv) from Definition 3.7. (iii) and (iv) follow from Definition 3.9, since $m_k = l|Y|^k$ and $\sup_k m_k/|L_k| < \infty$.

We prove (ii). Let $L_k = (w_1, \ldots, w_{|L_k|})$. Each word w_t consists of z_t consecutive elements $x_{s_t - z_t}, \ldots, x_{s_t - 1}$ of x followed by a single symbol. Since $|z_t - kl| \le l$, there is some block $B = [(u+k)l] \setminus [ul]$ such that symmetric difference between B and $\{s_t - z_t, \ldots, s_t - 1\}$ is at most $2l$; therefore the simple edit distance between w and x_B is at most $2l+1$. (ii) then holds since $x_B \in Y^k$.

We finally show (i). We must bound from above the multiplicity of elements in L_k. Fix a mapping $t \mapsto B(t)$ as described above, i.e., $B(t) = [(u+k)l] \setminus [ul]$ such that $s_t \in [(u+k)l] \setminus [(u+k-1)l]$. We must bound from above

$$\max_{t_0 \in [|L_k|]} |\{t : w_t = w_{t_0}\}|,$$

which is clearly not more than

$$\left(\max_B |\{t : B(t) = B\}|\right)\left(\max_{t_0 \in [|L_k|]} |\{B(t) : w_t = w_{t_0}\}|\right).$$

We bound each one of the above factors from above by functions of l and $|A|$ that do not depend on k. We begin with the first factor, the number of indices t mapped to any given block B. Since $s_t \in [(u+k)l] \setminus [(u+k-1)l]$, there are only l possible values for s_t. Since each (s_t, a^2_t) is unique, the multiplicity of each s_t is at most $|A^2| - 1$; therefore the number of indices mapped to any given B is at most $l(|A^2| - 1)$.

Now, fix $t_0 \in [|L_k|]$. We bound from above the cardinality of $\{B(t) : w_t = w_{t_0}\}$. Since the simple edit distance between w_t and $x_{B(t)}$ is at most $2l+1$, for any t, we have that if $w_t = w_{t_0}$ then $x_{B(t)}$ belongs to the ball of radius $4l+2$ around $x_{B(t_0)}$. Since the size of that ball is a function of l and $|A|$ (but not k) and since x_B determines B, the cardinality of $\{B(t) : w_t = w_{t_0}\}$ is bounded (by a function of l and $|A|$).

□

A random automaton scheme allows us to implement a random play with certain properties. Lemma 3.15 shows that these properties guarantee that the expected distribution of beliefs induced by the play is close to the target distribution $\mathcal{P}$.

Lemma 3.13. *Let Ξ be a random automaton scheme for agent 1 and $\mathcal{P} = \sum_{i=1}^{n} q_i \delta_{Q_i}$. Let r_k be the length of $\bar{a}_\Xi(k)$. For every $k \in \mathbb{N}$, let $\bar{a}(k) = \bar{a}_{\pi.\Xi}(k)$ be the random play obtained by taking a uniform random permutation $\pi \in G_k$. The corresponding mixture of strategies $\sigma^1_{\pi.\Xi}(k)$ coupled with any $\bar{a}_{\pi.\Xi}(k)$ consistent strategies for agent 2 defines a correlated strategy τ_k.*

(i) The play induced by τ_k is $\bar{a}(k)$.

(ii) The strategy τ_k is in $\Delta(\Sigma^1(m_k) \times \Sigma^2)^{r_k}$.

(iii) Let $Q \in \Delta(A \times [n])$ be given by $Q(a,i) = q_i Q_i(a)$. The random play $\bar{a} = \bar{a}(k)$ can be coupled with a random $[n]$-valued sequence $\bar{b} = b_1, \ldots, b_{r_k}$, such that

- $\lim_{k\to\infty} \left\|Q - \mathbb{E}\left[\mathrm{emp}(\bar{a}, \bar{b})\right]\right\| = 0,$
- $\lim_{k\to\infty} \frac{1}{r_k} H(\bar{b}) = 0.$

(iv) $\liminf_{k\to\infty} \frac{1}{r_k} H(\bar{a}(k)) \geq \bar{H}(\mathcal{P}) - f(l)$, where $f(l) \to 0$, as $l \to \infty$.

(v) $\limsup_{k\to\infty} \frac{m_k \log m_k}{r_k} \leq \bar{H}(\mathcal{P}) \limsup_{k\to\infty} \frac{m_k}{|L_k|}$.

Proof. Parts (i), (ii), and (v) immediately follow from Lemma 3.11.

Part (iv) follows from Lemma 3.11(i), Lemma 3.8 and Equation (3.1).

It remains to verify Part (iii). Associate with any state of $s \in [m_k]$ a number $i = i(s) \in [n]$, defined by

$$s \in [l_1 + \ldots + l_i] \setminus [l_1 + \ldots + l_{i-1}] \mod l.$$

Recall that the induced play is of the form $\bar{a} = w_1 w_2 \cdots w_{|L_k|}$, where $w_t = (x_{s_t^1 - z_t}, \ldots, x_{s_t - 1}, a_t)$, for some $a_t \in A$. Let $u_t = i(s_t^1 - z_t) \cdots i(s_t - 1)\, 1$, and $\bar{b} = (b_1, b_2, \ldots, b_{|L_k|})$. Since $|w_t| = |b_t| = z_t + 1 > (k-1)l$, and since the empirical distribution of any word $x_{tl+1} \ldots x_{tl+l}$ coupled with $i(tl+1) \ldots i(tl+l)$ is Q, $\|\mathrm{emp}(b_t, w_t) - Q\| = \mathcal{O}(k^{-1})$; therefore $\left\|Q - \mathbb{E}\left[\mathrm{emp}(\bar{a}, \bar{b})\right]\right\| = \mathcal{O}(k^{-1})$.

The sequence $\bar{b}$ can be read off from the random variables $\{z_t, s_t \bmod l : 1 \leq t \leq |L_k|\}$. Since these random variables take values in sets of size l,

$$\frac{1}{r_k} H(\bar{b}) \leq \frac{1}{r_k} H(z_t, s_t \bmod l : 1 \leq t \leq |L_k|) \leq \frac{|L_k| \log l}{r_k} \leq \frac{\log l}{(k-1)l} \to 0,$$

as $k \to \infty$. $\square$

3.4 Information criterion

The following lemma will allow us to determine whether a distribution of beliefs is $\bar{d}$-close to a target Dirac distribution.

Lemma 3.14. *Let A be a finite set and $H\colon \Delta(A) \to \mathbb{R}$ continuous and strictly concave. For every $\epsilon > 0$ there is $\delta > 0$ such that for every $Q \in \Delta(A)$ and $\mathcal{P} = \sum_i q_i \delta_{Q_i} \in \Delta\Delta(A)$, if*

(i) $\|Q - \sum_i q_i Q_i\| < \delta$, *and*

(ii) $\sum_i q_i H(Q_i) > H(Q) - \delta$,

then

$$\bar{d}(\delta_Q, \mathcal{P}) < \epsilon.$$

Lemma 3.14 has been shown by several authors independently, see for instance Lemma 1 in Ornstein [1970] and Lemma 22 in Shapira [2007].

A distribution over beliefs can be represented by a pair of random variables. Let x be a random variables taking values in a finite set X. Define $\mathrm{p}(x) \in \Delta\Delta(X)$ to be the Dirac measure supported on the distribution of x. For an event of positive probability E, define $\mathrm{p}(x|E) \in \Delta\Delta(X)$ to be the Dirac measure supported on the distribution of x conditional on E. Let y be another random variable taking values in a finite set Y. The distribution over beliefs $\mathrm{p}(x|y) \in \Delta\Delta(X)$ is defined by

$$\sum_{\mathbf{y} \in Y} \mathbf{P}(y = \mathbf{y}) \mathrm{p}(x|y = \mathbf{y}).$$

Lemma 3.14 says that for every $\epsilon > 0$ there is $\delta = \delta(\epsilon, |X|)$, such that

$$I(x; y) < \delta \quad \Rightarrow \quad \bar{d}(\mathrm{p}(x), \mathrm{p}(x|y)) < \epsilon.$$

Let z be a third random variable. By Markov's inequality we have

$$I(x; y|z) < \delta^2 \quad \Rightarrow \quad \bar{d}(\mathrm{p}(x|z), \mathrm{p}(x|y, z)) < \epsilon + \delta. \tag{3.2}$$

Lemma 3.15 (Information Criterion)**.** *Let A be a finite alphabet. For every $\epsilon > 0$ there is $\delta > 0$ such that for every $n \geq 1$, every $Q \in \Delta(A \times [n])$, every $r > 0$, and every random play $\bar{a} = a_1, \ldots, a_r$, if $\bar{a}$ can be coupled with an $[n]$-valued random sequence $\bar{b} = b_1, \ldots, b_r$ such that*

(i) $\left\|Q - \mathbb{E}\left[\mathrm{emp}(\bar{a}, \bar{b})\right]\right\| < \delta$,

(ii) $\frac{1}{r} H(\bar{a}|\bar{b}) > H(x|y) - \delta$, *where* $(x, y) \sim Q$,

(iii) $\frac{1}{r}I(\bar{a};\bar{b}) < \delta$,

then

$$\bar{d}(\mathrm{p}(x|y), \mathcal{P}_r(\bar{a})) < \epsilon.$$

Before we prove Lemma 3.15, note that the case of $n = 1$ is exactly Lemma 3.14.

Proof. Since the mapping that maps the distribution of (x, y) to $\mathrm{p}(x|y)$ is continuous as a function form $\Delta(A \times [n])$ to $\Delta\Delta(A)$, we may assume w.l.o.g. that $\mathbb{E}\mathrm{emp}(\bar{a}, \bar{b}) = Q$. Let $\epsilon > 0$. By (3.2) we can take $\delta > 0$ as such that

$$I(a; b|c) < \delta \quad \Rightarrow \quad \bar{d}(\mathrm{p}(a|c), \mathrm{p}(a|b, c)) < \tfrac{1}{2}\epsilon,$$

for any random variables a, b, c, where a takes values in A.

Let t be a random variable uniformly distributing in $[r]$ independently of $(\bar{a}, \bar{b})$. Let $\mathcal{H}_t = a_1, \ldots, a_t$. Note that $(a_t, b_t) \sim Q$, and $\mathcal{P}_r(\bar{a}) = \mathrm{p}(a_t|\mathcal{H}_{t-1}, t)$.

By the triangle inequality

$$\begin{aligned}
\bar{d}(\mathrm{p}(a_t|b_t), \mathrm{p}(a_t|\mathcal{H}_{t-1}, t)) \leq\ & \bar{d}(\mathrm{p}(a_t|b_t), \mathrm{p}(a_t|b_t, \mathcal{H}_{t-1}, t)) \\
& + \bar{d}(\mathrm{p}(a_t|b_t, \mathcal{H}_{t-1}, t), \mathrm{p}(a_t|\mathcal{H}_{t-1}, t)).
\end{aligned}$$

By the choice of δ, the proof will be concluded if we prove two inequalities:

$$I(a_t; \mathcal{H}_{t-1}, t|b_t) < \delta, \tag{3.3}$$

$$I(a_t; b_t|\mathcal{H}_{t-1}, t) < \delta. \tag{3.4}$$

For (3.3):

$$\begin{aligned}
I(a_t; \mathcal{H}_{t-1}, t|b_t) &= H(a_t|b_t) - H(a_t|\mathcal{H}_{t-1}, t, b_t) \\
&\leq H(a_t|b_t) - H(a_t|\mathcal{H}_{t-1}, t, \bar{b}) \\
&= H(a_t|b_t) - H(\bar{a}|\bar{b}) \\
&< \delta
\end{aligned}$$

where the last inequality is provided by condition (ii) of the lemma.

For (3.4):

$$\begin{aligned}
I(a_t; b_t|\mathcal{H}_{t-1}, t) &= H(a_t|\mathcal{H}_{t-1}, t) - H(a_t|\mathcal{H}_{t-1}, t, b_t) \\
&\leq \frac{1}{r}\left[H(\bar{a}) - H(\bar{a}|\bar{b})\right] \\
&= \frac{1}{r}I(\bar{a};\bar{b}) \\
&< \delta
\end{aligned}$$

where the last inequality is provided by condition (iii) of the lemma. $\square$

3.5 One automaton

The stage is set for proving Theorem 2.2.

Proof of Theorem 2.2. Let $\mathcal{P} \in \Delta\Delta(A)$ and $\epsilon > 0$. Since the function $\bar{H}(\cdot)$ is continuous and the set of distributions with finite support and rational coefficients is dense in $\Delta\Delta(A)$, we may assume w.l.o.g. that $\mathcal{P} = \sum_{i=1}^{n} q_i \delta_{Q_i}$ is rational. Let $\delta = \delta(\epsilon) > 0$ be given by Lemma 3.15. Let l be a common denominator of $\mathcal{P}$, such that $f(l)$ of Lemma 3.13 is less than δ.

We define a random automaton scheme for agent 1

$$\Xi = \left\langle \mathcal{P}, l, \left\{ x(k), L_k, (s_t(k), z_t(k), a_t^2(k))_{t=1}^{|L_k|}, G_k \right\}_{k=1}^{\infty} \right\rangle$$

with

- $x(k) = x_1, \ldots, x_{m_k}$ an arbitrary compound de Bruijn sequence as specified by Definition 3.9,
- $|L_k| = (|A^2| - 1)m_k$,
- $z_t(k) = kl - 1$ for every $t \in [|L_k|]$,
- $\{(s_t(k), a_t^2(k))\}$ an arbitrary enumeration of all pairs (s, a^2) such that $s \in [m_k]$ and $a^2 \neq x_s^2(k)$,
- G_k the entire symmetry group on $[|L_k|]$.

Lemma 3.13 provides the strategies $\tau_k \in \Delta(\Sigma^1(m_k) \times \Sigma^2)^{r_k}$. Also, since $|L_k| = (|A^2| - 1)m_k$, the same lemma ensures that $\limsup_{k\to\infty} \frac{m_k \log m_k}{r_k} \leq \frac{\bar{H}(\mathcal{P})}{|A^2|-1} + f(l)$. Lemma 3.13(iii) and Lemma 3.15 ensure that $\bar{d}(\mathcal{P}_{r_k}(\tau_k), \mathcal{P}) < \epsilon$, for every k large enough. Lemma 3.6 extends the construction to strategies that induce a play of any length $r > r_0$, for some $r_0 > 0$. □

3.6 One automaton, one sequence

In this section we prove Theorem 2.4. Theorem 2.4 refers only to $\mathcal{P}_r^1(\tau)$, and so we can take any coupling $\mathcal{P} \in \Delta\Delta(A)$ such that $\mathcal{P}^1 = \mathcal{P}_r^1(\tau)$. In particular, $\mathcal{P}$ can be taken such that $P^2 = \delta_{a^2}$, for some $a^2 \in A^2$ ($\mathcal{P}$ a.s.).

The strategies of agent 2 are a subset of the strategies specified in the proof of Theorem 2.2. Take the same k and l as in the proof of Theorem 2.2. Agent 2's strategy will be a $kl(|A^2| - 1)$-periodic sequence of pure actions, where at times not divisible by kl agent 2 plays a^2. At each time $t = ukl$, $u = 1, \ldots, |A^2| - 1$, agent 2 plays a different action in $A^2 \setminus \{a^2\}$. After time $kl(|A^2| - 1)$, agent 2 continues periodically.

Formally, given $\epsilon > 0$ and a rational distribution $\mathcal{P} \in \Delta\Delta(A)$ with $\mathcal{P}^2 = \delta_{\delta_{a^2}}$. Let $\delta = \delta(\epsilon) > 0$ be given by Lemma 3.15. Let l be a common denominator of $\mathcal{P}$, such that $f(l)$ of Lemma 3.13 is less than δ.

Let

$$\Xi = \left\langle \mathcal{P}, l, \left\{ x(k), L_k, (s_t(k), z_t(k), a_t^2(k))_{t=1}^{|L_k|}, G_k \right\}_{k=1}^{\infty} \right\rangle$$

be a random automaton scheme for agent 1 with

- $x(k) = x_1, \ldots, x_{m_k}$ an arbitrary compound de Bruijn sequence as specified by Definition 3.9,
- $|L_k| = (|A^2| - 1)m_k$,
- $z_t(k) = kl - 1$ for every $t \in [|L_k|]$,
- $\{(s_t(k), a_t^2(k))\}$ an enumeration of all pairs $(s, a^2) \in [m_k] \times A^2 \setminus \{\mathbf{a^2}\}$, such $a_t^2(k)$ is $(|A^2| - 1)$-periodic (in t),
- G_k is the group of all the permutations π on $[|L_k|]$ such that $\pi(t) = t \mod |A^2| - 1$, for all $t \in [|L_k|]$.

The choice of the jump transition $\{(s_t(k), a_t^2(k))\}$ was made such that the induced play $\bar{a}_\Xi^2(k) \in A_p(kl(|A^2| - 1))$. The definition of G_k is such that $\bar{a}_{\pi.\Xi}^2(k) = \bar{a}_\Xi^2(k)$, for all $\pi \in G_k$.

The rest of the specification of Ξ is the same as in the proof of Theorem 2.2, and so the length of the induced play is the same. The only thing remaining to show is that G_k is large enough, so that $\sup_k \left(\frac{|L_k|!}{|G_k|}\right)^{\frac{1}{|L_k|}} < \infty$. Indeed,

$$|G_k| = \left(\frac{|L_k|}{|A^2| - 1}!\right)^{|A^2|-1}$$

and therefore,

$$\left(\frac{|L_k|!}{|G_k|}\right)^{\frac{1}{|L_k|}} = \binom{|L_k|}{\frac{|L_k|}{|A^2|-1}, \ldots, \frac{|L_k|}{|A^2|-1}}^{\frac{1}{|L_k|}} \leq (|A^2| - 1)^{|L_k| \frac{1}{|L_k|}} = |A^2| - 1.$$

$\square$

3.7 Pairs of automata

In this section we prove Theorem 2.3. For simplicity of exposition we prove the theorem only in the case that $|A^1| = |A^2| \geq 2$ which is what we assume throughout this section.

The set of beliefs $\Delta(A)$ is divided into two regions:

$$\Delta_I(A) = \{Q \in \Delta(A) : \text{“}Q \text{ is supported on either one row or one column”}\},$$
$$\Delta_{II}(A) = \Delta(A) \setminus \Delta_I(A).$$

Every $\mathcal{P} \in \Delta\Delta(A)$ can be (uniquely) represented as $\lambda\mathcal{P}_I + (1-\lambda)\mathcal{P}_{II}$, where $\mathcal{P}_{I,II} \in \Delta\Delta_{I,II}(A)$ and $0 \le \lambda \le 1$. The constant $C(\mathcal{P})$ of Theorem 2.3 is defined as $\lambda C_I(\mathcal{P}_I) + (1-\lambda)C_{II}(\mathcal{P}_{II})$, with $C_{I,II}\colon \Delta\Delta_{I,II}(A) \to \mathbb{R}_+$ defined by

$$C_I(\mathcal{P}) = \frac{\bar{H}(\mathcal{P})}{n-1}, \qquad \text{(where } n = |A^1| = |A^2| \text{)}$$

$$C_{II}(\mathcal{P}) = \frac{\bar{H}(\mathcal{P})}{\bar{D}(\mathcal{P})}.$$

Where, $\bar{D}(\mathcal{P}) = \int D(Q)\,\mathrm{d}\mathcal{P}(Q)$ and $D\colon \Delta(A) \to [0, n-1]$ is a continuous function whose zeros are exactly $\Delta_I(A)$. The definition of D is deferred to Section 3.7.2.

The following lemma allows us to consider each one of the cases $\mathcal{P} \in \Delta\Delta_I(A)$ and $\mathcal{P} \in \Delta\Delta_{II}(A)$ separately.

Lemma 3.16. *Let $\mathcal{P}_1, \mathcal{P}_2 \in \Delta\Delta(A)$, $C_1, C_2 \ge 0$, and $0 \le \lambda \le 1$. Suppose that for every $\epsilon > 0$ and every $l \in \{1,2\}$ there exists $r_0 \in \mathbb{N}$ such that for every $r \ge r_0$ there exist $r' \le r$, $m \in \mathbb{N}$, and $\tau \in \Delta(\Sigma^1(m) \times \Sigma^2(m))^{r'}$ such that*

$$\frac{r'}{r} \ge 1 - \epsilon,$$
$$\frac{m \log m}{r'} \le C_l + \epsilon,$$

and

$$\bar{d}(\mathcal{P}_l, \mathcal{P}_{r'}(\tau)) < \epsilon.$$

Then, for every $\epsilon > 0$ there exists $r_0 \in \mathbb{N}$ such that for every $r \ge r_0$ there exist $r' \le r$, $m \in \mathbb{N}$ and $\tau \in \Delta(\Sigma^1(m) \times \Sigma^2(m))^{r'}$ such that

$$\frac{r'}{r} \ge 1 - \epsilon,$$

$$\frac{m \log m}{r'} \le \lambda C_1 + (1-\lambda)C_2 + \epsilon,$$

and

$$\bar{d}(\lambda\mathcal{P}_1 + (1-\lambda)\mathcal{P}_2, \mathcal{P}_{r'}(\tau)) < \epsilon.$$

Proof. Let $\epsilon > 0$ and $\lambda \in (0,1)$. Set $\lambda_1 = \lambda$ and $\lambda_2 = 1 - \lambda_1$, and $\mathcal{P} = \lambda_1 \mathcal{P}_1 + \lambda_2 \mathcal{P}_2$. Let $r_0 > 0$ be large enough so that for every $r \geq r_0$ and $l \in \{1,2\}$, there are $r_l \leq \lambda_l r$, $m_l \in \mathbb{N}$, and $\tau_l \in \Delta(\Sigma^1(m_l) \times \Sigma^2(m_l))^{r_l}$ such that

$$\frac{r_l}{\lambda_l r} \geq 1 - \epsilon,$$

$$\frac{m_l \log m_l}{r_l} \leq C_l + \epsilon,$$

and

$$\bar{d}(\mathcal{P}_l, \mathcal{P}_{r_l}(\tau_l)) < \epsilon.$$

Let $r' = r_1 + r_2$ and $m = m_1 + m_2$. Let τ be the concatenation of τ_1 and τ_2. We have,

$$\tau \in \Delta(\Sigma^1(m) \times \Sigma^2(m))^{r'},$$

$$\frac{r'}{r} \geq 1 - \epsilon,$$

and

$$\bar{d}(\mathcal{P}, \mathcal{P}_{r'}(\tau)) < \epsilon.$$

It remains to verify

$$\begin{aligned}
\frac{m \log m}{r'} = \sum_{l=1,2} \frac{m_l \log m}{r'} = \sum_{l=1,2} \frac{m_l \log m_l}{r'} + \sum_{l=1,2} \frac{m_l \log \frac{m}{m_l}}{r'} \\
\leq \frac{r}{r'} \sum_{l=1,2} \lambda_l \frac{m_l \log m_l}{r_l} + \frac{m}{r'} \sum_{l=1,2} \frac{m_l}{m} \log \frac{m}{m_l} \\
\leq \frac{1}{1-\epsilon} \left[\sum_{l=1,2} \lambda_l C_1 + \epsilon \right] + \frac{m}{r'} H\left(\frac{m_1}{m}, \frac{m_2}{m} \right) \\
\leq \sum_{l=1,2} \lambda_l C_1 + \frac{\epsilon}{1-\epsilon}(C_1 + C_2 + 1) + o(1).
\end{aligned}$$

□

By Lemmata 3.6 and 3.16, it is sufficient to prove the following lemma.

Lemma 3.17. *For every $\epsilon > 0$ and every $\mathcal{P} \in \Delta\Delta_I(A) \cup \Delta\Delta_{II}(A)$ there exist sequences $r_k, m_k \in \mathbb{N}$ and $\tau_k \in \Delta(\Sigma^1(m_k) \times \Sigma^2(m_k))^{r_k}$ satisfying*

$$\sup_{k\in\mathbb{N}} r_k = \infty, \tag{3.5}$$

$$\sup_{k\in\mathbb{N}} r_{k+1}/r_k < \infty, \tag{3.6}$$

$$\limsup_{k\to\infty} \frac{m_k \log m_k}{r_k} \le C(\mathcal{P}) + \epsilon, \tag{3.7}$$

$$\limsup_{k\to\infty} \bar{d}(\mathcal{P}, \mathcal{P}_r(\tau_k)) < \epsilon. \tag{3.8}$$

3.7.1 $C_I(\mathcal{P})$

In this section we prove Lemma 3.17 in the case $\mathcal{P} \in \Delta\Delta_I(A)$. We partition $\Delta_I(A)$ into finitely many regions. By virtue of Lemma 3.16, we assume w.l.o.g. that $\mathcal{P}$ is supported in one of these regions.

There are $n^2 + 2n$ regions as follows: Dirac beliefs, beliefs supported on any single row, and the beliefs supported on any single column:

$$\begin{array}{ll} \{\delta_a\} & a \in A, \\ \big\{Q \in \Delta(\big\{a^1\big\} \times A^2) : H(Q) > 0\big\} & a^1 \in A^1, \\ \big\{Q \in \Delta(A^1 \times \big\{a^2\big\}) : H(Q) > 0\big\} & a^2 \in A^2. \end{array}$$

The case $\mathcal{P} = \delta_{\delta_a}$ is simple. Let $b \in (A^1 \setminus \{a^1\}) \times (A^2 \setminus \{a^2\})$. Let σ^1_k be an automaton with k states that k-periodically outputs $k-1$ consecutive a^1s followed by a b^1. Let σ^2_k be an automaton that $k+1$-periodically outputs k consecutive a^2s followed by a b^2. A transition of the form $(\cdot, b)$ is first hit at time $k(k+1)$; therefore this pair of automata is $r_k = k(k+1)$ concatenable. Lemma 3.17 holds with $m_k = k+1$.

It remains to consider the case $\mathcal{P} \in \Delta\Delta(A^1 \times \{\mathbf{a^2}\})$ with $\bar{H}(\mathcal{P}) > 0$ (the case $\mathcal{P} \in \Delta\Delta(\{a^1\} \times A^2)$ is symmetric). The construction builds on the construction in the proof of Theorem 2.4. We consider m_k, r_k, L_k and the automata $\sigma^1(k)$ and $\sigma^2(k)$, as given in the proof of Theorem 2.4.

We construct an $(r_k + kl)$-concatenable pair of automata with $m_k + kl$ states that generate the same sequence as $(\sigma^1(k), \sigma^2)$ in the first kl steps. Since $kl \ll m_k, r_k$, the conditions of Lemma 3.17 are satisfied.

Agent 1's automaton is the concatenation of $\sigma^1(k)$ with an oblivious automaton with kl states. The oblivious automaton starts at state 1 and moves to state 2 and then 3, and so on until it reaches state kl, while always outputting some fixed action $\mathbf{a^1} \in A^1$. Note that L_k does not include a sequence where agent 1 always plays $\mathbf{a^1}$.

The $kl(n-1)$-periodic oblivious automaton σ^2 is paired with an automaton that looks for the ending sequence of kl consecutive $\mathbf{a^1}$ actions. Formally, suppose the states of σ^2, $[kl(n-1)]$, are visited in increasing order $1, 2, \ldots, kl(n-1)$. The states of the new automaton are $[kl(n-1)] \times \{0, 1\}$. The transitions are of the form $(i, b, a^1) \mapsto (i + 1 \bmod kl(n-1), x)$, where $x = 1$ if and only if either $i = kl(n-1)$, or $b = 1$ and $a^1 = \mathbf{a^1}$.

After r_k steps the automaton of agent 1 reaches the final transition of σ_k^1 and moves to the initial state of its second automaton. At the same time agent 1's automaton is at state $(1, 1)$. In the next kl steps agent 1 plays $\mathbf{a^1}$ which result in both agents reaching certain states for the first time: agent 2 – state (kl, kl), and agent 1 – the last state of its second automaton. Therefore, our construction is $r_k + kl$ concatenable.

3.7.2 $C_{II}(\mathcal{P})$

In this section we prove Lemma 3.17 in the case $\mathcal{P} \in \Delta\Delta_{II}(A)$. Recall that the definition of $C(\mathcal{P}) = C_{II}(\mathcal{P})$ depends on an appropriate definition of a function $D\colon \Delta_{II}(A) \to (0, n-1]$. We begin by defining[2] the function D.

We think of A as the complete bipartite graph with color sets A^1, A^2. A set of action profiles $J \subset A$ is called a *matching* if for every $(a^1, a^2), (b^1, b^2) \in J$, $a^1 = b^1$ if and only if $a^2 = b^2$.

Definition 3.18. We define $D\colon \Delta(A) \to [0, n-1]$ by

$$D(Q) = \max_{\text{matching } J} |J|(|J| - 1) \min_{a \in J} Q(a).$$

Note that D is continuous and its range is indeed $[0, n-1]$. The maximum, $n-1$, is attained when Q is a uniform distribution on a perfect matching.[3] [4] The minimum, 0, is attained exactly when Q is supported on either a single row or a single column.

The construction is based on the existence of a pair of random enumeration schemes, one for each agent, such that the two schemes induce exactly the same play.

Definition 3.19. A *random bi-automata scheme* is a tuple

$$\Phi = \left\langle \mathcal{P}, l, \left\{ x(k), L_k, (s_t^1(k), s_t^2(k), z_t(k))_{t=1}^{|L_k|}, G_k \right\}_{k=1}^{\infty} \right\rangle$$

[2]The function D is not the largest possible. Its concavification $cav\, D$ is also possible as discussed in Section 3.7.3.

[3]A perfect matching is a matching of size n.

[4]Section 3.7.3 suggests that D can be replaced by $cav\, D$ which implies that $n-1$ can be attained whenever the two marginals of Q, Q^1 and Q^2, are uniform.

such that for each $i \in \{1,2\}$,

$$\Phi_i = \left\langle \mathcal{P}, l, \left\{ x(k), L_k, (s^i_t(k), z_t(k), x^{-i}_{s^{-i}_t})_{t=1}^{|L_k|}, G_k \right\}_{k=1}^{\infty} \right\rangle$$

is a random automaton scheme for agent i.

Let $\mathcal{P} \in \Delta\Delta_{II}(A)$. By continuity, we may assume w.l.o.g. that $\mathcal{P} = \sum_i q_i \delta_{Q_i}$ is finitely supported rational distribution over rational believes. Let l be an arbitrary common denominator of $\mathcal{P}$. By Lemmata 3.6, 3.13, and 3.15, it is sufficient to construct a random bi-automata scheme

$$\Phi = \left\langle \mathcal{P}, l, \left\{ x(k), L_k, (s^1_t(k), s^2_t(k), z_t(k))_{t=1}^{|L_k|}, G_k \right\}_{k=1}^{\infty} \right\rangle$$

with

$$\liminf_{k\to\infty} \frac{|L_k|}{m_k} \geq \bar{D}(\mathcal{P}) - \epsilon(l),$$

where $\epsilon(l) \to 0$, as $l \to \infty$.

Let $\mathcal{P} = \sum_j q_j \delta_{Q_j}$ and l a natural number such that each $l_j = q_j l$ is a common denominator of Q_j. Let $Y = \bigtimes_j T^{Q_j}(l_j)$, and $x(k) = x_1, \ldots, x_{m_k}$ an arbitrary compound de Bruijn sequence of order k over the alphabet Y. The group G_k is the entire symmetry group on $[|L_k|]$. It remains to specify $|L_k|$, $s^1_t(k)$, $s^2_t(k)$, and $z_t(k)$.

In what follows we sometimes suppress the index k when it causes no confusion. For a state $s \in [m_k]$, let $\{lt+1, \ldots, lt+l\} \ni s$ be the unique non-overlapping interval of l consecutive states containing s. We consider the chunk of x that begins $(k-1)l$ places before that interval and ends at s. That is, let $u = u(s)$ be the unique integer such that $(k-1)l \leq u < kl$ and $s = u+1 \mod l$. We call the chunk of the de Bruijn sequence $x_{[s-u,s-1]} := (x_{s-u}, \ldots, x_{s-1})$ the *stem* of s, denoted $\mathrm{stem}(s)$. We call the state $s-u$ the *origin* of s, denoted $\mathrm{orig}(s)$.

For states $s, t \in [m_k]$ corresponding to elements of the compound de Bruijn sequence $x_s = (a^1, a^2)$ and $x_t = (b^1, b^2)$, we say that (s,t) is a *good pair of states* if

(i) $\mathrm{stem}(s) = \mathrm{stem}(t)$, and

(ii) $a^1 \neq b^1$ and $a^2 \neq b^2$.

The *played action* at (s,t) is (a_1, b_2) and we use the notations:

$$\mathrm{act}(s,t) = (a^1, b^2),$$
$$\mathrm{act}^1(s) = a^1,$$
$$\mathrm{act}^2(t) = b^2.$$

A set of good pairs of states X is *independent* if for every $(s,t),(s',t')\in X$

$$\begin{aligned} s = s' &\Rightarrow \mathrm{act}^2(t) \neq \mathrm{act}^2(t'), \qquad \text{and} \\ t = t' &\Rightarrow \mathrm{act}^1(s) \neq \mathrm{act}^1(s'). \end{aligned} \tag{3.9}$$

In graph theoretic terminology, X is an independent set of vertices in an auxiliary graph whose vertices are the good pairs of states and whose edges are given by (3.9).

From an independent set of good pairs of states of size L, $\{(s^1_t, s^2_t)\}_{t=1}^L$, we construct the remaining components of Φ by setting $|L_k| = L$, and $z_t = u(s^1_t) = u(s^2_t)$, for all $t = 1,\dots,L$. Condition (3.9) ensures that the fifth item of Definition 3.9 is satisfied.

Therefore, it remains to find an independent set of size $\geq (\bar{D}(\mathcal{P}) - \epsilon(l))m_k$ for every l and every k large enough, where $\epsilon(l)\to 0$, as $l\to\infty$.

With a tolerable abuse of notation define

$$D(J,Q) = |J|(|J|-1)\min_{a\in J} Q(a),$$

for a matching J and a distribution $Q\in\Delta(A)$. Recall that

$$D(Q) = \max_J D(J,q).$$

Let J^* be a matching such that $D(Q) = D(J^*,Q)$.

It will be convenient to consider two partitions of the states: a coarser one and a finer one. The coarser partition is made by grouping together states with the same stem. We say the two states $s,t\in[m_k]$ are equivalent if they have the same stem. Namely, $s\sim t$ if $\mathrm{stem}(s) = \mathrm{stem}(t)$. The equivalence class of $s\in[m_k]$ is denoted $\langle s\rangle$. The finer partition is given by further considering the action x_s. For $a\in A$ and $J\subset A$, we define

$$\begin{aligned} \langle s\rangle_a &= \{t : t\sim s,\ x_t = a\}, \text{ and} \\ \langle s\rangle_J &= \bigcup_{a\in J} \langle s\rangle_a. \end{aligned}$$

These partitions induce conditional distributions $q(\cdot|s)\in\Delta(A)$ defined by

$$q(a|s) = \frac{|\langle s\rangle_a|}{|\langle s\rangle|}.$$

Denote $L_j = l_1+\cdots+l_j$. Note that if $s = i \mod l$, $i\in[L_j]\setminus[L_{j-1}]$, then $q(\cdot|s)$ depends only on $x_{[s-i+L_{j-1}+1,s-1]}$. Explicitly,

$$q(a|s) = \frac{|\{y\in T^{Q_j}(l_j) : y_{[1,i-L_{j-1}-1]} = x_{[s-i+L_{j-1}+1,s-1]},\ y_i = a\}|}{|\{y\in T^{Q_j}(l_j) : y_{[1,i-L_{j-1}-1]} = x_{[s-i+L_{j-1}+1,s-1]}\}|}.$$

Explanation: the states in $\langle s\rangle$ correspond to blocks of x whose prefix is stem(s), which are exactly all the sequences in Y^k beginning with stem(s). Since the first $k-1$ Y-words of these sequences are determined by stem(s), we have

$$\begin{aligned}|\langle s\rangle| &= |\{y \in Y : y_{[1,i-1]} = x_{[s-i+1,s-1]}\}| \\ &= |\{y \in T^{Q_j}(l_j) : y_{[1,i-L_{j-1}-1]} = x_{[s-i+L_{j-1}+1,s-1]}\}| \times |\bigtimes_{j'>j} T^{Q'_j}(l_{j'})|,\end{aligned}$$

and similarly,

$$\begin{aligned}&|\langle s\rangle_a| = |\{y \in Y : y_{[1,i-1]} = x_{[s-i+1,s-1]},\ y_i = a\}| \\ &= |\{y \in T^{Q_j}(l_j) : y_{[1,i-L_{j-1}-1]} = x_{[s-i+L_{j-1}+1,s-1]},\ y_i = a\}| \times |\bigtimes_{j'>j} T^{Q'_j}(l_{j'})|.\end{aligned}$$

When s is drawn uniformly at random from $[m_k]$, $q(\cdot|s)$ becomes a random variable (a function of s). We denote this random variable q_l, emphasizing the dependence of the distribution of $q(\cdot|s)$ on l (and not on k or the choice of x). Note that the distribution of q_l is the expected distribution of beliefs of a uniform random sample from Y. By Lemma 3.15, since $\lim_{l\to\infty} \frac{\log|T^{Q_j}(l_j)|}{l_j} = H(Q_j)$, for all j, we have that $q_l \to \mathcal{P}$ in distribution (as $l \to \infty$). Since the function $D(J^*,\cdot)$ is continuous on the compact domain $\Delta(A)$, the proof will be concluded if we construct an independent set of size $m_k\mathbb{E}[D(J^*, q_l)]$.

We turn now to construct an independent set of size $m_k\mathbb{E}[D(J^*, q_l)]$. Let $s \in [m_k]$. Let $d = \min_{a\in J^*} |\langle s\rangle_a|$. It suffices to find $d|J^*|(|J^*|-1)$ independent good pairs of states from $\langle s\rangle_{J^*}$, since by doing so we obtain an independent set of pairs $I_{\langle s\rangle} \subset \langle s\rangle \times \langle s\rangle$ of size $|\langle s\rangle| D(J^*, q(\cdot|s))$. Summing over all the equivalence classes $\{\langle s\rangle : s \in [m_k]\}$ gives an independent set of the desired size,

$$\left|\bigcup_{\langle s\rangle : s\in[m_k]} I_{\langle s\rangle}\right| = \sum_{s\in[m_k]} \frac{|I_{\langle s\rangle}|}{|\langle s\rangle|} = \sum_{s\in[m_k]} D(J^*, q(\cdot|s)) = m_k\mathbb{E}[D(P^*, q_l)].$$

For every $a \in J^*$, let $S_a \subset \langle s\rangle_a$, $|S_a| = d$. For every ordered pair $(a,b) \in J^* \times J^*$, $a \neq b$, let $\varphi_{a,b} : S_a \to S_b$ be a bijection. The following set is an independent set of good pairs all coming from $\langle s\rangle_{J^*}$ whose size is $d|J^*|(|J^*|-1)$:

$$\bigcup_{(a,b)\in J^*\times J^*, a\neq b} \{(t, \varphi_{a,b}(t)) : t \in S_a\}.$$

3.7.3 Improving C_{II}

We suggest an improvement of C_{II}, by replacing the function D by its concavification $cav\, D$. Our proof that shows that $D(Q)$ is achievable could be slightly modified in order to show that $(cav\, D)(Q)$ is achievable. For the sake of simplicity and clarity of the proof, we chose to prove the slightly weaker statement, while only sketching the proof of the stronger statement.

The quantity $(cav\, D)(Q)$ is the solution of the following linear programme:

$$
\begin{aligned}
\max \sum_{\text{matching } J} x_J |J|(|J|-1) \qquad & \text{subject to} \\
\sum_{\substack{\text{matching } J: \\ a \in J}} x_J \le Q(a) \qquad & \text{for every } a \in A, \\
x_J \ge 0 \qquad & \text{for every matching } J.
\end{aligned}
$$

As before, the idea is to find an independent set of good pair of states of size $m_k \mathbb{E}\left[(cav\, D)(q(\cdot|s))\right] - o(m_k)$. Take a random state $s \in m_k$. Associate with s a bipartite multi-graph M_s whose colour sets are A^1 and A^2, and the multi-edges between each $a^1 \in A^1$ and $a^2 \in A^2$ correspond to $\langle s \rangle_{(a^1,a^2)}$. Let $\{x_J\}$ be an optimal solution for the linear programme $(cav\, D)(q(\cdot|s))$. Let $d_J = \lfloor x_J | \langle s \rangle | \rfloor$, for every matching J. Note that $d_J / | \langle s \rangle |$ is a feasible solution for the linear grogramme, and it is nearly optimal, if $| \langle s \rangle |$ is large. The multi-graph M_s contains d_J copies of each matching J. For every copy of J, take the ordered pairs of states that correspond to any ordered pair of multi-edges in J. Do the same for every J. The union of these collections of pairs of states is an independent set of good pairs of states of size $d_*(s) = \sum_J d_J |J|(|J|-1)$, with $d_*(s)/(cav\, D)(q(\cdot|s)) \to 1$, as $| \langle s \rangle | \to \infty$.

Remark 3.20. Since $cav\, D(Q)$ is a linear programme, it could be interesting to find a meaningful interpretation to the dual program, which could possibly lead to tightness results.

4 Values of repeated games

We present a few implications of our main results on the min-max values of repeated games played by finite automata.

We consider a class of three-player repeated games, parameterized by a one-shot strategic-form game, automaton size constraints, and game duration.

Formally, $G = (A = A^1 \times A^2 \times A^3, g : A \to \mathbb{R})$ is a three-player game, where g is the payoff to Player 3. The payoff function extends to $g : \Delta(A) \to \mathbb{R}$ linearly. Given such a game, we define different min-max values depending on whether the team of players 1 and 2 are restricted to pure strategies, or can randomize independently, or play correlated strategies. The correlated min-max is defined as

$$\operatorname{cor}\min\max G = \min_{\tau^{1,2}\in\Delta(A^1\times A^2)} \max_{\tau^3\in\Delta(A^3)} g(\tau^{1,2}\otimes\tau^3).$$

The pure min-max and max-min are defined as

$$\begin{aligned}\operatorname{pure}\min\max G &= \min_{a^1\in A^1, a^2\in A^2} \max_{a^3\in A^3} g(a^1,a^2,a^3),\\ \operatorname{pure}\max\min G &= \max_{a^3\in A^3} \min_{a^1\in A^1, a^2\in A^2} g(a^1,a^2,a^3).\end{aligned}$$

The r-stage repeated version of G is denoted $G_r(m_1, m_2, m_3)$, where each player i is restricted to strategies of automaton size m_i, and the payoff is the average per-stage payoff (or the limiting average, if $r = \infty$). We allow for at most one of the parameters to be infinite. In this case, either all player's strategy sets are finite, or r is finite, which guarantees that the game has a finite strategic form.

Since $G_r(m_1, m_2, m_3)$ is a finite game in strategic form, the three values, cor min max, pure min max and pure max min are well defined. For example

$$\operatorname{cor}\min\max G_r(m_1,m_2,m_3) = \\ \min_{\tau^{1,2}\in\Delta(\Sigma^1(m_1)\times\Sigma^2(m_2))} \max_{\tau^3\in\Delta(\Sigma^3(m_3))} \mathbb{E}_{\tau^{1,2,3}} \frac{1}{r}\sum_{t=1}^{r} g(a_t^1,a_t^2,a_t^3).$$

We study asymptotic properties of the min-max values of $G_r(m_1, m_2, m_3)$ and compare them to the min-max values of G.

In what follows $\{m_k\}$, $\{n_k\}$, and $\{r_k\}$ are sequences of natural numbers. If Players 1 and 2 could implement a random play of r_k independent Q-distributed actions, then they could guarantee, in the r_k-stage repeated game, the value that Q guarantees in the one-shot game. Since implementing an i.i.d. play is not always possible under automaton size constraints, we need a notion of approximation to i.i.d. play that guarantees similar strategic power.

Definition 4.1. For $Q \in (A_1 \times A_2)$, an *approximation of* r_k *independent* Q-*distributed random actions* (*approximate i.i.d. play*, for short) is a sequence of distributions $P_k \in \Delta((A^1 \times A^2)^{r_k})$ such that the expected empirical frequency of beliefs of P_k converges to δ_Q w.r.t. the $\bar{d}$ distance (as $k \to \infty$).

Theorems 2.2, 2.3, and 2.4 provide conditions under which Players 1 and 2 can implement such approximate i.i.d. plays. These conditions translate to three Propositions regarding the correlated min-max value.

Theorem 2.2 provides an implementation of an approximate i.i.d. play in situations where Player 1 is restricted and Player 2 is fully rational.

Proposition 4.2. *If* $\lim_{k\to\infty} r_k = \infty$ *and*

$$\liminf_{k\to\infty} \frac{m_k \log m_k}{r_k} \geq \frac{\log(|A^1 \times A^2|)}{(|A^2| - 1)},$$

then

$$\lim_{k\to\infty} \operatorname{cor\,min\,max} G_{r_k}(m_k, \infty, \infty) = \operatorname{cor\,min\,max} G.$$

When both players, 1 and 2, are restricted, Theorem 2.3 provides a similar result, only that in this case the ratio between r_k and $m_k \log m_k$ depends on the one-shot payoff function.

Proposition 4.3. *There exists a constant* $C > 0$*, that depends on* G*, such that if* $\lim_{k\to\infty} r_k = \infty$ *and*

$$\liminf_{k\to\infty} \frac{m_k \log m_k}{r_k} \geq C,$$

then

$$\lim_{k\to\infty} \operatorname{cor\,min\,max} G_{r_k}(m_k, m_k, \infty) = \operatorname{cor\,min\,max} G.$$

When the actions of Player 2 do not influence the payoff, the correlated min-max of the one-shot game is equal to the (uncorrelated) min-max. By Theorem 2.4, the situation is asymptotically the same w.r.t. the repeated version. Furthermore, Player 2's automaton can be very simple: pure, oblivious, and with just $\mathcal{O}(\log m)$ states.

Proposition 4.4. *If the payoff function* g *does not depend on Player 2's actions, then there is* $C > 0$ *such that if* $\lim_{k\to\infty} r_k = \infty$ *and*

$$\liminf_{k\to\infty} \frac{m_k \log m_k}{r_k} \geq \frac{\log |A^1|}{|A^2| - 1},$$

then[5]

$$\lim_{k\to\infty} \min_{\substack{\tau^1 \in \Delta(\Sigma^1(m_k)) \\ \sigma^2 \in A^2_p(\lceil C \log(m_k) \rceil)}} \max_{\sigma^3 \in \Sigma^3} g(\tau^1, \sigma^2, \sigma^3) = \operatorname{cor\,min\,max} G.$$

[5] Note that $\operatorname{cor\,min\,max} G = \min_{x^1 \in \Delta(A^1)} \max_{a^3 \in A^3} g(x^1, \cdot, a^3)$

Proofs of Propositions 4.2, 4.3, and 4.4. Let $Q \in \Delta(A^1 \times A^2)$ be a correlated min-max strategy for players 1 and 2 in the game G. By Theorem 2.3, there exist a constant $C = C(Q)$ and correlated strategies $\tau_k \in \Delta(\Sigma^1(m_k) \times \Sigma^2(m_k))$ whose induced play of length $Cm_k \log m_k$ approximates a sequence of independent Q-distributed random variables. This proves Proposition 4.3. Similarly, resorting to Theorems 2.2 and 2.4 proves Propositions 4.2 and 4.4 respectively. $\square$

Conversely to Propositions 4.2, 4.3, and 4.4, Player 3 has a pure strategy that guarantees the pure one-shot-game min-max value when the duration of the game is much more than $m \log m$.

Proposition 4.5. *If* $\frac{m_k \log m_k}{r_k} \to 0$*, as* $k \to \infty$*, then*

$$\lim_{k\to\infty} \text{pure} \max \min G_{r_k}(m_k, m_k, \infty) = \text{pure} \min \max G.$$

Proof. By Lemma 3.1, $|\Sigma^1(m_k) \times \Sigma^2(m_k)| = \mathcal{O}(m_k \log m_k)$; Theorem 1 from Neyman and Okada [2009] asserts that if one player is restricted to strategies in a set whose size is sub-exponential in the duration of the game and the other player is unrestricted, then the other player can asymptotically guarantee the pure-min-max value using pure strategies. $\square$

We turn now to settings in which Player 3's automaton size is bounded and the duration of the game r_k is not necessarily finite.

Proposition 4.6. *If* $\frac{\log n_k}{m_k \log m_k} \to 0$*, as* $k \to \infty$*, then*

$$\lim_{k\to\infty} \text{cor} \min \max G_{r_k}(m_k, m_k, n_k) = \text{cor} \min \max G.$$

Proof. Let $Q \in \Delta(A^1 \times A^2)$ be an optimal correlated strategy for Players 1 and 2 in G. By Lemmata 3.6 and 3.17, there exists $C(Q) > 0$ and $r'_k \sim C(Q)m_k \log m_k$, such that Players 1 and 2 can implement an approximation of a r'_k-periodic sequence of independent Q-distributed random actions. We assume w.l.o.g. that the induced play is stationary, because it can be made stationary by taking the average of r'_k shifts of the play. Divide the duration of the game r_k into time intervals of length r'_k with a possible remainder in the beginning. Namely, intervals $[t_{n+1}] \setminus [t_n]$, where $t_0 = 0$, $t_1 = r_k - r'_k \left\lfloor \frac{r_k}{r'_k} \right\rfloor$ (or 0 if $r_k = \infty$), and $t_n = t_1 + (n-1)r'_k$. Let σ^3 be a pure strategy for player 3 (possibly, a best response). For any $n = 0, 1, \ldots$ let $\sigma^3_{t_n}$ be Player's 3 strategy induced on the game starting at time t_n. The random variable $\sigma^3_{t_n}$ $(n > 0)$ assumes values in a set of size n_k depending on the state of Player 3's automaton at time t_n, and $\sigma^3_{t_0} = \sigma^3$ is a fixed strategy. By [Peretz, 2012, Corollary 4.3] (also, [Neyman, 2008, pp. 9, 15-16]), the expected average per-stage payoff between time t_n and time $t_{n+1}-1$ is asymptotically at least the correlated min-max value of G, as k goes to infinity. Since this is true for any n, the expectation of the payoff is asymptotically at least the correlated min-max value of G. □

Conversely to Proposition 4.6, if $m_k \log m_k$ is not large enough compared to $\log n_k$ then Player 3 can beat players 1 and 2.

Proposition 4.7. *For every $C > |A^3|(|A^1|+|A^2|)-2$, if $\log n_k \geq C m_k \log m_k$ and $r_k \geq n_k \to \infty$, as $k \to \infty$, then*

$$\lim_{k\to\infty} \operatorname{cor\,min\,max} G_{r_k}(m_k, m_k, n_k) = \operatorname{pure\,min\,max} G.$$

Proposition 4.7 strengthens Proposition 4.5 in that Player 3 needs only be exponentially smarter than Players 1 and 2 and the duration of the game is allowed to be proportional to $m_k \log m_k$. Nevertheless, Proposition 4.5 is stronger in that Player 3 can do with a pure strategy. Conversely, Players 1 and 2 have a winning pure strategy when they are much smarter than Player 3, as shown in Proposition 4.8 below.

Propositions 4.3 and 4.5 together show that around a duration proportional to $m \log m$ a phase transition occurs. The correlated min-max value of the repeated game changes from the correlated to the pure min-max values of the one-shot game.

Proof of Proposition 4.7. The proof is similar to the proof of [Neyman, 1997, Theorem 3]. Let $\gamma : A^1 \times A^2 \to A^3$ be a best-response function. That is, $g(a^1, a^2, \gamma(a^1, a^2)) = \max_{a^3 \in A^3} g(a^1, a^2, a^3)$, for every $a^1 \in A^1$, $a^2 \in A^2$. The definition of γ extends to strategies in the repeated game recursively by

$$\gamma : \Sigma^1 \times \Sigma^2 \to \left(A^3\right)^{\mathbb{N}},$$
$$\gamma_t(\sigma^1, \sigma^2) = \gamma(\sigma^{1,2}(\gamma_l(\sigma^1, \sigma^2)_{l=0}^{t-1})),$$

where $\sigma^{1,2}$ is the reduced strategy of the team of players 1 and 2 induced by σ^1 and σ^2. Let $X(m) = \{\gamma(\sigma^1, \sigma^2) : \sigma^1 \in \Sigma^1(m),\ \sigma^2 \in \Sigma^2(m)\}$. Since every pair of automata $(\sigma^1, \sigma^2) \in \Sigma^1(m) \times \Sigma^2(m)$ can be regarded as a single automaton with m^2 states, every $x \in X(m)$ can be implemented through an oblivious automaton with m^2 states.

We next construct an automaton for Player 3 with $m^3|X(m)|$ states. By Lemma 3.1, $\log(m^3|X(m)|) \leq (|A^3|(|A^1|+|A^2|)-2)m \log m + o(m \log m)$. The strategy of Player 3 is the following strategy:

(i) Choose $x \in X(m)$ uniformly at random. Play x as long as it best responds to the actions of players 1 and 2.

(ii) Repeat Step (i) $m|X(m)|$ times or until the end of the game.

(iii) Continue arbitrarily.

The implementation of x in Step (i) requires m^2 states, and so implementing $m|X(m)|$ repetitions of Step (i) requires $m^3|X(m)|$ states, as promised.

The probability of "guessing" the best response in each iteration of Step (i) is at least $|X(m)|^{-1}$, therefore the probability getting to Step (iii) is at most $(1 - |X(m)|^{-1})^{m|X(m)|} \to 0$, as $m \to \infty$.

It remains to verify the expected payoff. In each iteration of Step (i) there is at most one non-best-response stage. Since the duration of the game is much longer than the number of iterations, and the probability of getting to Step (iii) is negligible, the expected payoff is guaranteed to be asymptotically at least the pure min-max of the one-shot game. $\square$

If players 1 and 2 are much "smarter" than Player 3, they can implement a fixed play that looks like a sequence of optimal correlated mixed actions in the eyes of Player 3.

Proposition 4.8. *If* $\frac{n_k}{m_k} \to 0$, *as* $k \to \infty$, *then*

$$\lim_{k\to\infty} \text{pure} \min\max G_{r_k}(m_k, m_k, n_k) = \text{cor} \min\max G,$$

for any $r_k \geq m_k \log m_k$.

Proof. Let $Q \in \Delta(A^1 \times A^2)$ be an optimal correlated strategy for Players 1 and 2 in G. By Lemmata 3.6 and 3.17, there are $C(Q) > 0$, $r'_k \sim C(Q)m_k \log m_k$, and $\tau_k \in \Delta(\Sigma^1(m_k) \times \Sigma^2(m_k))^{r'_k}$ such that $\mathcal{P}_{r'_k}(\tau_k)$ converges to δ_Q (as $k \to \infty$) w.r.t. the $\bar{d}$ metric. We assume w.l.o.g. that r_k is a multiple of r'_k (otherwise we prove the Proposition with a $n_k \ll m'_k \ll m_k$, such that dividing r_k by r'_k leaves a negligible reminder).

Let β : *support*$(\tau_k) \to \Sigma^3(n_k)$ be any function. It suffices to show that for any such β (possibly a best-reply for Player 3) there is at least one $a \in$ *support*(τ_k) such that $g_{r_k}(a, \beta(a)) \leq \text{cor} \min\max G + o(1)$. We show a stronger statement

$$\mathbb{E}g_{r_k}(\sigma, \beta(\sigma)) \leq \text{cor} \min\max G + o(1), \tag{4.1}$$

where σ is a random strategy that distributes according to τ_k. We may assume that $r_k = r'_k$, since otherwise we can divide $[r_k]$ into intervals of length r'_k and prove the statement in each one of these intervals.

Inequality (4.1) follows from [Peretz, 2012, Corollary 4.3], since $H(\beta(\sigma)) \leq \log(|\Sigma^3(n_k)|) = o(r'_k)$. $\square$

5 Acknowledgments

We are grateful to three anonymous referees and an editor who contributed valuable suggestions that significantly improved the quality of this paper.

References

Dilip Abreu and Ariel Rubinstein. The structure of nash equilibrium in repeated games with finite automata. *Econometrica*, 56(6):1259–1281, 1988.

Robert J. Aumann. Survey of repeated games. In *Essays in Game Theory and Mathematical Economics in Honor of Oskar Morgenstern*, pages 11–42. Bibliographisches Institut, Mannheim, 1981.

Gilad Bavly and Abraham Neyman. Online concealed correlation and bounded rationality. *Games and Economic Behavior*, 88(0):71 – 89, 2014.

Elchanan Ben-Porath. Repeated games with finite automata. *Journal of Economic Theory*, 59(1):17–32, 1993.

Thomas M. Cover and Joy A. Thomas. *Elements of Information Theory.* Wiley Interscience, New York, 2nd edition, 2006.

Nicolaas Govert de Bruijn. A combinatorial problem. *Koninklijke Nederlandse Akademie v. Wetenschappen*, 49:758–764, 1946.

Olivier Gossner and Penélope Hernández. On the complexity of coordination. *Mathematics of Operations Research*, 28(1):127–140, 2003.

Olivier Gossner and Penélope Hernández. Coordination through De Bruijn sequences. *Operations Research Letters*, 34(1):17 – 21, 2006.

Olivier Gossner and Tristan Tomala. Empirical distributions of beliefs under imperfect observation. *Mathematics of Operations Research*, 31(1):13–30, 2006.

Olivier Gossner and Tristan Tomala. Secret correlation in repeated games with imperfect monitoring. *Mathematics of Operations Research*, 32(2): 413–424, 2007.

Olivier Gossner, Penelope Hernandez, and Abraham Neyman. Optimal use of communication resources. *Econometrica*, 74(6):1603–1636, 2006.

E. Kalai and W. Stanford. Finite rationality and interpersonal complexity in repeated games. *Econometrica*, 56(2):397–410, 1988.

Ehud Kalai. *Bounded rationality and strategic complexity in repeated games*. AFCET, 1990.

Ehud Lehrer. Repeated games with stationary bounded recall strategies. *Journal of Economic Theory*, 46(1):130–144, October 1988.

Abraham Neyman. Bounded complexity justifies cooperation in the finitely repeated prisoners' dilemma. *Economics Letters*, 19(3):227–229, 1985.

Abraham Neyman. Cooperation, repetition, and automata. In *Cooperation: Game Theoretic Approaches, NATO ASI Series F*, pages 233–255. Springer-Verlag, 1997.

Abraham Neyman. Finitely repeated games with finite automata. *Mathematics of Operations Research*, 23(3):513–552, 1998.

Abraham Neyman. Learning effectiveness and memory size. Discussion Paper 476, Center for the Study of Rationality, Hebrew University, Jerusalem, 2008.

Abraham Neyman and Daijiro Okada. Strategic entropy and complexity in repeated games. *Games and Economic Behavior*, 29(1–2):191 – 223, 1999.

Abraham Neyman and Daijiro Okada. Repeated games with bounded entropy. *Games and Economic Behavior*, 30(2):228–247, 2000.

Abraham Neyman and Daijiro Okada. Growth of strategy sets, entropy, and nonstationary bounded recall. *Games and Economic Behavior*, 66(1):404 – 425, 2009.

Abraham Neyman and Joel Spencer. Complexity and effective prediction. *Games and Economic Behavior*, 69(1):165–168, 2010.

Donald Ornstein. Bernoulli shifts with the same entropy are isomorphic. *Advances in Mathematics*, 4(3):337–352, 1970.

Ron Peretz. The strategic value of recall. *Games and Economic Behavior*, 74(1):332 – 351, 2012.

Ron Peretz. Correlation through bounded recall strategies. *International Journal of Game Theory*, 42(4):867–890, 2013.

Alon Shapira. *Symmetric Online Matching Pennies*. PhD thesis, Hebrew University, Jerusalem, 2007.