When Can Limited Randomness Be Used in Repeated Games?

Pavel Hubáček[*]   Moni Naor[†]   Jonathan Ullman[‡]

December 1, 2015

Abstract

The central result of classical game theory states that every finite normal form game has a Nash equilibrium, provided that players are allowed to use randomized (mixed) strategies. However, in practice, humans are known to be bad at generating random-like sequences, and true random bits may be unavailable. Even if the players have access to enough random bits for a single instance of the game their randomness might be insufficient if the game is played many times.

In this work, we ask whether randomness is necessary for equilibria to exist in finitely repeated games. We show that for a large class of games containing arbitrary two-player zero-sum games, approximate Nash equilibria of the n-stage repeated version of the game exist if and only if both players have $\Omega(n)$ random bits. In contrast, we show that there exists a class of games for which no equilibrium exists in pure strategies, yet the n-stage repeated version of the game has an exact Nash equilibrium in which each player uses only a constant number of random bits.

When the players are assumed to be computationally bounded, if cryptographic pseudo-random generators (or, equivalently, one-way functions) exist, then the players can base their strategies on "random-like" sequences derived from only a small number of truly random bits. We show that, in contrast, in repeated two-player zero-sum games, if pseudorandom generators *do not* exist, then $\Omega(n)$ random bits remain necessary for equilibria to exist.

Contents

1 Introduction

The signature result of classical game theory states that a Nash equilibrium exists in every finite normal form game, provided that players are allowed to play randomized (mixed) strategies. It is easy to see in some games (e.g. Rock-Paper-Scissors) that randomization is necessary for the existence of Nash equilibrium. However, the assumption that players are able to randomize their strategies in an arbitrary manner is quite strong, as sources of true randomness may be unavailable and humans are known to be bad at generating random-like sequences.

Motivated by these considerations, Budinich and Fortnow [BF11] investigated the question of whether Nash equilibria exist when players only have access to *limited randomness*. Specifically, they looked at the "repeated matching pennies." Matching pennies is a very simple, two-player, two-action, zero-sum game in which the unique equilibrium is for each player to flip a fair coin and play an action uniformly at random. If the game is repeated for n stages, then the unique Nash equilibrium is for each player to play an independent, uniformly random action in each of the n stages. Budinich and Fortnow considered the case where the players only have access to $\ll n$ bits of randomness, which are insufficient to play the unique equilibrium of the game, and showed that there does not even exist an *approximate* equilibrium (where the approximation depends on the deficiency in randomness). That is, if the players cannot choose independent, uniformly random actions in each of the n stages, then no approximate equilibrium exists.

In this work, we further investigate the need for randomness in repeated games by asking whether the same results hold for *arbitrary* games. That is, we start with an arbitrary multi-player game such that Nash equilibria only exist if players can use β bits of randomness. Then we consider the n-stage repetition of that game. Do equilibria exist in the n-stage game if players only have access to $\ll \beta n$ bits of randomness? First, we show that the answer is essentially *no* for arbitrary zero-sum games, significantly generalizing the results of Budinich and Fortnow. On the other hand, we show that the answer is *yes* for a large class of general games.

These results hold when both players are assumed to be computationally unbounded. As noted by Budinich and Fortnow, if we assume that the players are required to run in polynomial time, and cryptographic pseudorandom generators (or, equivalently, one-way functions) exist, then a player equipped with only $\ll n$ truly random bits can generate n *pseudorandom bits* that appear truly random to a polynomial time adversary. Thus, in the computationally bounded regime, if pseudorandom generators exist, then linear randomness is not necessary. We show that, in contrast, in arbitrary repeated two-player zero-sum games, if pseudorandom generators *do not* exist, then linear randomness remains necessary.

1.1 Our Results

Suppose we have an arbitrary finite strategic game among k players. We consider the n-stage repetition of this game in which in each of the n consecutive stages, each of the k players simultaneously chooses an action (which may depend on the history of the previous stages). We assume that in the 1-stage game $\beta > 0$ bits of randomness for each player are necessary and sufficient for an equilibrium to exist. We ask whether or not the existence of approximate equilibria in the n-stage game requires a linear amount of randomness ($\Omega(n)$ bits) per player.

The case of computationally unbounded players. Our first set of results concerns players who are computationally unbounded, which is the standard model in classical game theory. In

this setting, our first result shows that linear randomness is necessary for a large class of games including every two-player zero-sum game.

Theorem 1 (informal). *For any k-player strategic game in which every Nash equilibrium achieves the minmax payoff profile, in any Nash equilibrium of its repeated version the players' strategies use randomness at least linear in the number of stages.*

An important subset of strategic games where any Nash equilibrium achieves the minmax payoff profile is the class of two-player zero-sum games where, as implied by the von Neumann's minmax theorem, the concept of Nash equilibrium collapses to the minmax solution. Hence, to play a Nash equilibrium in any finitely repeated two-player zero-sum game the players must use randomness at least linear in the number of stages.

Second, we show that the above results cannot be extended to arbitrary games. That is, there exists a class of strategic games that, in their repeated version, admit "randomness efficient" Nash equilibria:

Theorem 2 (informal). *For any k-player strategic game in which for every player there exists a Nash equilibrium that achieves strictly higher expectation than the minmax strategy, there exists a Nash equilibrium of its repeated version where the players use total randomness independent of the number of stages.*

As we shall see, this result is related to the "finite horizon Nash folk theorem," which roughly states that in finitely repeated games every payoff profile in the stage game that dominates the minmax payoff profile can be achieved as a payoff profile of some Nash equilibrium of the repeated game.

The case of computationally efficient players. For repeated two-player zero-sum games we study the existence of Nash equilibria with limited randomness when the players are computationally bounded. Under the assumption that one-way functions do not exist (see the above discussion), we show that in every repeated two-player zero-sum game it is possible to *efficiently exploit* any opponent that uses low randomness (i.e., gain a non-negligible advantage over the value of the stage game). Hence, in repeated two-player zero-sum games there are no computational Nash equilibria in which one of the players uses randomness sub-linear in the number of the stages.

Theorem 4 (informal). *In any repeated two-player zero-sum game, if one-way functions do not exist, then for any strategy of the column player using sub-linear randomness, there is a computationally efficient strategy for the row player that achieves an average payoff non-negligibly higher than his minmax payoff in the stage game.*

The proof of this result employs the algorithm of Naor and Rothblum [NR06] for learning adaptively changing distributions. The main idea is to adaptively reconstruct the small randomness used by the opponent in order to render his strategy effectively deterministic and then improve the expectation by playing the best response.

Strong exploitation of low-randomness players. In the classical setting, i.e., without restrictions on the computational power of the players, it was shown by Neyman and Okada [NO00] that in every repeated two-player zero-sum game it is possible to extract utility proportional to the randomness deficiency of the opponent. On the other hand, our result in the setting with computationally efficient players guarantees only a non-negligible advantage in the presence of a

low-randomness opponent. This leaves open an intriguing question of how much utility can one *efficiently* extract from an opponent that uses low randomness in a repeated two-player zero-sum game (see Section 5 for additional discussion).

The case of matching pennies. As noticed by Budinich and Fortnow [BF11], the repeated game of matching pennies exhibits clear tradeoffs between the randomness available to players and existence of ε-Nash equilibria. Our work generalizes their results already in the context of repeated matching pennies, since they assumed that the players randomize their strategies by flipping limited number of coins, whereas we only assume that the players' strategies are of low entropy. Our results for the game of matching pennies are provided in Appendix B.

1.2 Other Related Work

In one of the first works to consider the relation between the randomness available to players and the existence of equilibria Halpern and Pass [HP14] introduced a computational framework of machine games that explicitly incorporates the cost of computation into the utility functions of the players and specifically the possibility of randomness being expensive. They demonstrated this approach on the game of Rock-Paper-Scissors, and showed that in machine games where randomization is costly then Nash equilibria do not necessarily exist. However, in machine games where randomization is free then Nash equilibria always exist.

Based on derandomization techniques, Kalyanaraman and Umans [KU07] proposed randomness efficient algorithms both for finding equilibria and for playing strategic games. In the context of finitely repeated two-player zero-sum games where one of the players (referred to as the learner) is uninformed of the payoff matrix, they gave an adaptive on-line algorithm for the learner that can reuse randomness over the stages of the repeated game.

Halprin and Naor [HN10] suggested the possibility of using randomness generated by human players in repeated games for generation of pseudorandom sequences. The strategic game they proposed for this purpose is a two-player zero-sum game. As shown by our results, their choice improves the likelihood of extracting truly random bits from the gameplay, since the players must use linear randomness in the number of stages in equilibria of any repeated two-player zero-sum game.

2 Notation and Background

2.1 Game Theoretic Background

Here we provide the concepts from game theory that we use in this work (for an in-depth study see the classical textbook by Osborne and Rubinstein [OR94]).

Definition 1 (strategic game). A *strategic game* $G = \langle N, (A_i), (u_i) \rangle$ is a tuple consisting of

- a finite set of players N

- for each player $i \in N$ a nonempty set of actions A_i

- for each player $i \in N$ a utility function $u_i : A \to \mathbb{R}$ assigning each action profile $a \in A = \times_{j \in N} A_j$ a real-valued payoff $u_i(a)$.

In the special case when G is a two-player zero-sum game we use the notation $\langle (A_1, A_2), u \rangle$ instead of $\langle \{1, 2\}, (A_1, A_2), (u_1, u_2) \rangle$, since there are only two players and $u_1(a) = -u_2(a)$ for all $a \in A_1 \times A_2$. We refer to player 1 as the row player (also known as Rowena) and to player 2 as the column player (also known as Colin).[1]

We denote by S_i the set of mixed strategies of player i, i.e., the set $\Delta(A_i)$ of all probability distributions on the action space of player i. For a strategy profile $\sigma \in S = \times_{j \in N} S_j$ we use σ_i to denote the strategy of player i in σ and σ_{-i} to denote the profile of strategies of all the players in N except for player i in σ, and we write σ equivalently as (σ_i, σ_{-i}).

Definition 2 (Nash equilibrium in strategic game). A *Nash equilibrium* of a strategic game $\langle N, (A_i), (u_i) \rangle$ is a profile σ of strategies with the property that for every player $i \in N$ we have

$$\mathbf{E}[u(\sigma_i, \sigma_{-i})] \geq \mathbf{E}[(\sigma_i', \sigma_{-i})] \text{ for all } \sigma_i' \in S_i .$$

Definition 3 (minmax payoff). The *minmax payoff of player i* in strategic game $\langle N, (A_i), (u_i) \rangle$, denoted v_i, is the lowest payoff that the other players *can force upon player i*, i.e.,

$$v_i = \min_{\sigma_{-i} \in S_{-i}} \max_{\sigma_i \in S_i} \mathbf{E}[u_i(\sigma_i, \sigma_{-i})] .$$

A *minmax strategy of player i* in G is a strategy $\hat{\sigma}_i \in S_i$ such that $\mathbf{E}[u_i(\hat{\sigma}_i, \sigma_{-i})] \geq v_i$ for all $\sigma_{-i} \in S_{-i}$.

Definition 4 (feasible and individually rational payoff profile). An *individually rational payoff profile* of G is a vector $p \in \mathbb{R}^{|N|}$ that is in each coordinate at least as large as the minmax payoff of the corresponding player, i.e., a vector for which $p_i \geq v_i$ for all $i \in N$. A vector $p \in \mathbb{R}^{|N|}$ is a *feasible payoff profile of G* if there exists a collection $\{\alpha_a\}_{a \in A}$ of nonnegative rational numbers such that $\sum_{a \in A} \alpha_a = 1$ and $p_i = \sum_{a \in A} \alpha_a u_i(a)$ for all $i \in N$.

Note that since in every finite strategic game a Nash equilibrium always exists, there also always exists an individually rational payoff profile (the payoff profile of the Nash equilibrium). However, the Nash equilibrium payoff profile is not necessarily feasible in the above sense.

Definition 5 (n-stage repeated game). Let $G = \langle N, (A_i), (u_i) \rangle$ be a strategic game. An *n-stage repeated game of G* is an extensive form game with perfect information and simultaneous moves $G^n = \langle N, H, P, (u_i^*) \rangle$ in which:

- $H = \{\emptyset\} \cup \{\bigcup_{t=1}^n A^t\}$, where $\emptyset$ is the initial history and A^t is the set of sequences of action profiles in G of length t

- $P(h) = N$ for each non-terminal history $h \in H$

- $u_i^*(a^1, \ldots, a^n) = \frac{1}{n} \sum_{t=1}^n u_i(a^t)$ for every terminal history $(a^1, \ldots, a^n) \in A^n$.

A *behavioral strategy of player i* is a collection $(\sigma_i(h))_{h \in H \setminus A^n}$ of independent probability measures (one for each non-terminal history), where each $\sigma_i(h)$ is a probability measure over A_i.

Definition 6 (Nash equilibrium in n-stage repeated game). A *Nash equilibrium* of an n-stage repeated game of $G = \langle N, (A_i), (u_i) \rangle$ is a profile σ of behavioral strategies with the property that for every player $i \in N$ and every behavioral strategy σ_i', we have

$$\mathbf{E}[u^*(\sigma_i, \sigma_{-i})] \geq \mathbf{E}[u^*(\sigma_i', \sigma_{-i})] .$$

---

[1]We have adopted Colin and Rowena from Aumann and Hart [AH03].

2.2 Cryptographic Background

Pseudorandom generators and one-way functions. The notion of cryptographic pseudorandom generators was introduced by Blum and Micali [BM84], who defined them as algorithms that produce sequences of bits unpredictable in polynomial time, i.e., no efficient next-bit-test is able to predict the next output of the pseudorandom generator given the sequence of bits generated so far. As Yao [Yao82] showed, this is equivalent to a generator whose output is indistinguishable from a truly random string to any polynomial time observer. One of the central questions in cryptography is to understand the assumptions that are sufficient and necessary for implementing a particular cryptographic task. Impagliazzo and Luby [IL89] (see also Impagliazzo [Imp92]) showed that one-way functions are essential for many cryptographic primitives (e.g., private-key encryption, secure authentication, coin-flipping over telephone). Håstad, Impagliazzo, Levin and Luby [HILL99] showed that pseudorandom generators exist if and only if one-way functions exist. Therefore the existence of one-way functions is the major open problem of cryptography. For an in depth discussion see Goldreich [Gol01].

Standard notation. A function $\mu : \mathbb{N} \to \mathbb{R}^+$ is *negligible* if for all $c \in \mathbb{N}$ there exists $n_c \in \mathbb{N}$ such that for all $n \geq n_c$, $\mu(n) \leq n^{-c}$. A function $\mu : \mathbb{N} \to \mathbb{R}^+$ is *noticeable* if there exists $c \in \mathbb{N}$ and $n_c \in \mathbb{N}$ such that for all $n \geq n_c$, $\mu(n) \geq n^{-c}$.

Definition 7 (statistical distance). The statistical distance between two distributions X and Y over $\{0,1\}^\ell$, denoted by $\mathrm{SD}(X,Y)$, is defined as:

$$\mathrm{SD}(X,Y) = \frac{1}{2} \sum_{\alpha \in \{0,1\}^\ell} |\Pr[X = \alpha] - \Pr[Y = \alpha]| \ .$$

The most fundamental notion for measuring randomness is the Shannon entropy:

Definition 8 (Shannon entropy). Given a probability distribution $\rho \in \Delta(A)$, the *Shannon entropy of* ρ is defined as

$$H(\rho) := \mathop{\mathbf{E}}_{a \leftarrow \rho} \left(\log_2 \left(\frac{1}{\Pr(\rho = a)} \right) \right) .$$

As mentioned above, if we have a one-way function then many cryptographic primitives are possible and in particular we can stretch a short seed into a long seemingly random one. Hence, we will be interested in the case that such functions do not exist.

Definition 9 (almost one-way function). A function $f : \{0,1\}^{k(n)} \to \{0,1\}^{\ell(n)}$ is an *almost one-way function* if it is computable in polynomial time, and for infinitely many input lengths, for any PPTM $\mathcal{M}$, the probability that $\mathcal{M}$ inverts f on a random input is negligible. Namely, for any polynomial p, there exist infinitely many choices of $n \in \mathbb{N}$ such that

$$\Pr_{x \leftarrow U_{k(n)}, \mathcal{M}}[\mathcal{M}(f(x)) \in f^{-1}(x)] < \frac{1}{p(n)} \ .$$

3   Low-Entropy Nash Equilibria of Finitely Repeated Games

In this section we show that, in the setting with players that have unbounded computational power, there are two classes of k-player strategic games at the opposite sides of the spectrum with respect to the amount of randomness necessary for the existence of equilibria of their repeated versions.

To measure the randomness of a player's strategy we consider the maximal total Shannon entropy of his strategies used along any terminal history.

Definition 10 (Shannon entropy of a strategy in repeated game). Let $G = \langle N, (A_i), (u_i) \rangle$ be a finite strategic game and let σ_i be a strategy of player i in the n-stage repeated game of G. For any terminal history $a = (a^1, \ldots, a^n) \in A^n$, let $(\sigma_i(\emptyset), \sigma_i(a^1), \sigma_i(a^1, a^2), \ldots, \sigma_i(a^1, \ldots, a^{n-1}))$ be the n-tuple of strategies of player i in σ_i at all the non-terminal subhistories of a. We define the *Shannon entropy of σ_i*, denoted as $H(\sigma_i)$, as

$$H(\sigma_i) := \max_{a \in A^n} \left\{ H(\sigma_i(\emptyset)) + \sum_{j=1}^{n-1} H(\sigma_i(a^1, \ldots, a^j)) \right\} .$$

This is a worst case notion, in that it measures the entropy of the strategy of player i irrespective of the strategies of the other players. For some of our results we consider its alternative variant of *effective Shannon entropy of a strategy σ_i in a strategy profile σ*, defined as the maximal total entropy of σ_i along terminal histories that are sampled in σ with non-zero probability:[2]

$$H(\sigma_i, \sigma) := \max_{\substack{a \in A^n \\ \sigma(a) \neq 0}} \left\{ H(\sigma_i(\emptyset)) + \sum_{j=1}^{n-1} H(\sigma_i(a^1, \ldots, a^j)) \right\} .$$

For the class of games in which any Nash equilibrium payoff profile is exactly the minmax payoff profile (e.g. any two-player zero-sum game), the following proposition relates the Nash equilibria of the strategic game to the structure of Nash equilibria in its n-stage repeated version.[3]

Proposition 1. *Let $G = \langle N, (A_i), (u_i) \rangle$ be a strategic game such that any Nash equilibrium payoff profile is equal to the minmax payoff profile. For all $n \in \mathbb{N}$, if σ is a Nash equilibrium of $G^n = \langle N, H, P, (u_i^*) \rangle$, the n-stage repeated game of G, then for every non-terminal history $h \in H$ sampled with non-zero probability by σ the strategy profile $\sigma(h)$ is a Nash equilibrium of G.*

Proof. Assume to the contrary that there exists a Nash equilibrium σ of G^n such that for some non-terminal history $h \in H$, sampled with non-zero probability by σ, the strategy profile $\sigma(h)$ is not a Nash equilibrium of G. Let h be without loss of generality the longest history such that $\sigma(h)$ is not a Nash equilibrium of G. There exists a player i with a profitable deviation σ_i^* in the stage game to his strategy in the strategy profile $\sigma(h)$. Consider the strategy σ_i' of player i in G^n defined in the following way: $\sigma_i'(h') = \sigma_i(h')$ for any history $h' \in H$ that does not contain h as a subhistory, $\sigma_i'(h) = \sigma_i^*$ for the history h, and $\sigma_i'(h'')$ is the minmax strategy $\hat{\sigma}_i$ of player i in G for any history $h'' \neq h$ that contains h as a subhistory.

---

[2]Alternatively, we could define an average-case notion that weights the entropy by the probability of reaching the terminal history. We note that the worst-case notion makes our result (i.e., Theorem 2) only stronger.

[3]A variant of Proposition 1 with respect to pure equilibria is given in Osborne and Rubinstein [OR94] as Proposition 155.1.

Note that for any history $h' \in H$ that does not contain h as a subhistory, $\mathbf{E}[u_i((\sigma'_i, \sigma_{-i})(h'))] = \mathbf{E}[u_i((\sigma_i, \sigma_{-i})(h'))]$ by the construction of σ'_i. Since the minmax strategy $\hat{\sigma}_i$ of player i guarantees at least the minmax payoff v_i (equal to any Nash equilibrium payoff of player i in G), $\mathbf{E}[u_i((\sigma'_i, \sigma_{-i})(h''))] \geq \mathbf{E}[u_i((\sigma_i, \sigma_{-i})(h''))]$ for any history $h'' \neq h$ that contains h as a subhistory. Finally, $\mathbf{E}[u_i((\sigma'_i, \sigma_{-i})(h))] > \mathbf{E}[u_i((\sigma_i, \sigma_{-i})(h))]$ because σ_i^* is a profitable deviation for player i in G given the strategy profile $\sigma(h)$.

Recall that the history h is sampled in σ with non-zero probability, and hence $\mathbf{E}[u_i^*(\sigma'_i, \sigma_{-i})] > \mathbf{E}[u_i^*(\sigma_i, \sigma_{-i})]$, i.e., the alternative strategy σ'_i increases the expectation of player i in G^n given that the other players follow σ_{-i}, a contradiction to σ being a Nash equilibrium of G^n. □

For strategic games that satisfy the requirement of Proposition 1, we immediately get a linear lower bound on entropy needed to play Nash equilibria in their repeated games.

Theorem 1. *Let G be a strategic game such that any Nash equilibrium payoff profile is equal to the minmax payoff profile. For all $n \in \mathbb{N}$ and every player $i \in N$, if in any Nash equilibrium of G the strategy of player i is of entropy at least β_i then in any Nash equilibrium of G^n, the n-stage repeated game of G, the strategy of player i is of entropy at least $n\beta_i$.*

Proof. Assume to the contrary that there exists a Nash equilibrium σ of the n-stage repeated game of G with strategy of entropy strictly smaller than $n \cdot \beta_i$ for player i. By Proposition 1, $\sigma(h)$ is a Nash equilibrium of G for all h sampled by σ with non-zero probability. Hence, there must exist a history $h^* \in H$ sampled with non-zero probability in σ such that $\sigma(h^*)$ is a Nash equilibrium of G and the entropy $H(\sigma_i(h^*))$ of $\sigma_i(h^*)$ is strictly smaller than β_i, a contradiction. □

	Left (L)	Heads (H)	Tails (T)	Right (R)
Up (U)	$0, -1$	$0, -1$	$0, -1$	$0, \ \ 0$
Heads (H)	$0, -1$	$1, -1$	$-1, \ \ 1$	$-1, \ \ 0$
Tails (T)	$0, -1$	$-1, \ \ 1$	$1, -1$	$-1, \ \ 0$
Down (D)	$0, \ \ 0$	$-1, \ \ 1$	$-1, \ \ 1$	$1, \ \ 0$

Figure 1: The payoff matrix of an extended game of matching pennies.

Repeated non-zero-sum game requiring a lot of randomness. Theorem 1 applies not only to two-player zero-sum games but also to some non-zero-sum games. The game G given by the payoff matrix in Figure 1 is a variant of the game of matching pennies where the players have two additional options. There are three mixed Nash equilibria in G: $(\frac{1}{2}H + \frac{1}{2}T, \frac{1}{2}H + \frac{1}{2}T)$, $(\frac{1}{2}U + \frac{1}{2}D, \frac{1}{2}H + \frac{1}{2}R)$, and $(\frac{1}{2}U + \frac{1}{2}D, \frac{1}{2}T + \frac{1}{2}R)$; all the three Nash equilibria achieve the same payoff profile $(0, 0)$ and require each player to use one random bit. Notice that the row player can get utility 0 irrespective of the strategy of the column player by selecting his action "Up", and similarly the column player can ensure utility 0 by playing "Right". Hence, the minmax payoff profile is $(0, 0)$. Since none of the three Nash equilibria of G improves over the minmax payoff profile, we get by Theorem 1 that each player must use strategy of entropy at least n in any Nash equilibrium of the n-stage repeated game of G.

7

	Cooperate (C)	Heads (H)	Tails (T)	Punish (P)
Cooperate (C)	3, 3	-3, 6	-3, 6	$-3, -3$
Heads (H)	$6, -3$	$1, -1$	-1, 1	$-3, -3$
Tails (T)	$6, -3$	-1, 1	$1, -1$	$-3, -3$
Punish (P)	$-3, -3$	$-3, -3$	$-3, -3$	$-4, -4$

Figure 2: The payoff matrix of an extended game of matching pennies.

Repeated non-zero-sum game requiring low randomness. On the other hand, there are strategic games for which Theorem 1 does not apply, and the players may use in the n-stage repeated game equilibrium strategies of entropy proportional only to the entropy needed in the single-shot game.

Consider for example the strategic game G given by the payoff matrix in Figure 2. The strategy profile $\sigma = (\frac{1}{2}H + \frac{1}{2}T, \frac{1}{2}H + \frac{1}{2}T)$ is the unique Nash equilibrium of G that achieves payoff profile $(0,0)$. The minmax payoff profile is $(-3, -3)$, since any player can get utility at least -3 by playing C. We show that the n-stage repeated game of G admits a Nash equilibrium that requires only a single random coin, i.e., the same amount of randomness as the Nash equilibrium σ of the stage game G. Consider the strategy profile in which both players play C in the first $n-1$ rounds and in the last round each player plays H and T with equal probability, and if any player deviates from playing C in one of the first $n-1$ rounds then the opponent plays P throughout all the remaining stages. To see that this strategy profile is a Nash equilibrium of the n-stage repeated game of G note that any deviation from playing C in the first $n-1$ rounds can increase the utility of any player by at most 3 (by playing either H or T instead of C), however the subsequent punishment induces a loss of at least -3 which renders any deviation unprofitable.

The randomness efficient Nash equilibrium from the above example resembles the structure of Nash equilibria constructed in the proof of the *Nash folk theorem for finitely repeated games*. This theorem characterizes the payoff profiles that can be achieved by Nash equilibria of the repeated game. In particular, it shows that in strategic games G such that for very player i there exists a Nash equilibrium σ_i strictly improving over his minmax payoff any feasible payoff profile (i.e., any convex combination of payoff profiles in G with rational coefficients) that is individually rational (i.e., achieves at least the minmax level for every player) can be approximated by a Nash equilibrium of sufficiently long finitely repeated game of G (cf. Osborne and Rubinstein [OR94] for a survey of known folk theorems).

The main idea behind the proof of the folk theorem is that for every player i the gap between the payoff in the Nash equilibrium σ_i and the minmax payoff v_i can be used to punish the player in case he deviates from the strategy that approximates any feasible and individually rational payoff profile. In particular, in any such Nash equilibrium the players use a fixed number of rounds (independent of the number of stages n) before the last round in which they play according to some (possibly mixed) Nash equilibria of the stage game and in the preceding rounds they play pure strategies so that the overall payoff approximates the feasible payoff profile. Hence, the amount of randomness on all the equilibrium paths is independent of the number of stages in any such Nash equilibrium of the repeated game.

Theorem 2. *Let G be a strategic game such that for every player i there exists a Nash equilibrium σ_i of G in which the payoff of player i exceeds his minmax payoff v_i and there exists a feasible and individually rational payoff profile in G. Let β_i be such that in any Nash equilibrium of G the strategy of player i is of entropy at most β_i. There exists $c \in \mathbb{N}$ such that for all sufficiently large $n \in \mathbb{N}$ and every player $i \in N$ there exists a Nash equilibrium of G^n, the n-stage repeated game of G, in which the strategy of player i is of* effective entropy *at most $c \cdot \beta_i$.*

Proof. Let $p \in \mathbb{R}^{|N|}$ be the feasible and individually rational payoff profile of G. There exist coefficients $\{\alpha_a\}_{a \in A} \subset \mathbb{Q}$ such that $\sum_{a \in A} \alpha_a = 1$ and for all $i \in N$, $p_i = \sum_{a \in A} \alpha_a u_i(a)$. Let K be the smallest integer such that each α_a can be written as α'_a / K for $\alpha'_a \in \mathbb{N}$. For some $\ell \in \mathbb{N}$, we divide the stages in G^n into two parts of length $\ell \cdot K$ and $m = n - \ell \cdot K$. Let s be a strategy profile in G^n that schedules the first $\ell \cdot K$ stages such that each action profile a for which $\alpha_a \neq 0$ is played by the players in exactly $\ell \cdot \alpha'_a$ number of stages. In the remaining m stages the players cycle between the Nash equilibria $\{\sigma_i\}_{i \in N}$, i.e., for all $j \in \{0, \ldots, m-1\}$ at the stage $n - m + 1 + j$ the players play the Nash equilibrium $\sigma_{j'}$, where $j' = 1 + (j \mod |N|)$. In case any player i deviates from s in one of the first $\ell \cdot K$ rounds, the remaining players play the strategy that forces the minmax level v_i on player i.

Note that if the number m of the last stages is such that for all action profiles $a \in A$ with $\alpha_a \neq 0$ and for every player i: $\frac{m}{|N|} \left(\sum_{j \in N} \mathbf{E}[u_i(\sigma_j)] - |N| v_i \right) \geq \max_{a'_i \in A_i} u_i(a'_i, a_{-i}) - u_i(a)$, then no player has a profitable deviation and σ is a Nash equilibrium of G^n. The number m of last stages can be bounded by some constant c selected independently of n. Since the number of stages in which the players play according to some Nash equilibrium of G is at most c (the players take pure actions in all the first $n - c$ stages), for any player i the effective entropy of s_i in s is at most $c \cdot \beta_i$. $\qquad\square$

Randomness in Subgame Perfect Equilibria of Finitely Repeated Games. An unavoidable shortcoming of the solution concept of Nash equilibrium in the context of repeated (and in general extensive form) games is that it is possible for equilibria to be established based on non-credible threats. This issue can be circumvented by the stronger requirement of *subgame perfection* that demands the players' strategies to be best response at every history (even off the equilibrium path), and hence implicitly eliminates all empty threats.

Since any subgame perfect equilibrium is a Nash equilibrium, the linear lower bound on the amount of entropy applies to subgame perfect equilibria when the minmax payoff profile cannot be improved upon by any Nash equilibrium in the stage game. On the other hand, it is possible to construct a randomness efficient subgame perfect equilibrium in the n-stage repeated game if in the underlying game there are two Nash equilibria with different payoffs for each player. Such subgame perfect equilibrium is constructed in the proof of perfect finite horizon Folk theorem of Benoît and Krishna [BK85].

Characterization of games with randomness efficient equilibria. The condition on the structure of the stage game in Theorem 2 (i.e., that for every player there exists a Nash equilibrium of the stage game that strictly improves over his minmax payoff) is the same as in the Nash Folk theorem of Benoît and Krishna [BK87]. We leave it as an open problem whether ideas from a proof of a more general finite horizon Nash folk theorem (e.g. the one given by González-Díaz [Gon06]) could help to extend (or to characterize) the class of games that admit randomness efficient equilibria in their repeated versions.

4 Low-Entropy Computational Nash Equilibria of Finitely Repeated Two-Player Zero-Sum Games

In this section we study randomness in equilibria of repeated two-player zero-sum games with computationally efficient players. The solution concept we consider in this setting is *computational Nash equilibrium* (introduced in the work of Dodis, Halevi and Rabin [DHR00]) that assumes that the players are restricted to computationally efficient strategies and indifferent to negligible improvements in their utilities, i.e., a computational Nash equilibrium is analogous to the concept of ε-Nash equilibrium with a negligible ε, where the player's strategies, as well as any deviations, must be computationally efficient.

To capture the requirement of computational efficiency, the players' strategies must be implemented by families of polynomial-size circuits. For a two-player zero-sum game G, we denote by *repeated game of G* the infinite collection $\{G^n\}_{n \in \mathbb{N}}$ of all the n-stage repeated games of G. A family of polynomial size circuits $\{C_n\}_{n \in \mathbb{N}}$ implements the strategy of the row player in the repeated game of G as follows. In G^n, the n-stage repeated game of G, the circuit C_n takes as input a string corresponding to a non-terminal history h in G^n and $s(n)$ random bits; it outputs an action to be taken at history h. If the strategy of player $i \in \{1, 2\}$ is implemented by family $\{C_n^i\}_{n \in \mathbb{N}}$ then the gameplay in the n-stage repeated game of G is defined in the following way: player i samples a random string $r_i \in \{0,1\}^{s_i(n)}$ and at each stage of G^n takes the action $a = C_n^i(h, r_i) \in A_i$, given that the history of play up to the current stage is h. The utility function u_n^* is for all n defined as in the standard n-stage repeated game of G (i.e., it is the average utility achieved in the stage game over the n stages).

Definition 11 (computational Nash equilibrium of repeated game)**.** For a two-player zero-sum game $G = \langle (A_1, A_2), u \rangle$, *a computational Nash equilibrium of the repeated game of G* is a strategy profile $(\{C_n^1\}_{n \in \mathbb{N}}, \{C_n^2\}_{n \in \mathbb{N}})$ given by polynomial-size circuit families such that for every player $i \in \{1, 2\}$ and every strategy $\{\tilde{C}_n^i\}_{n \in \mathbb{N}}$ given by a polynomial-size circuit family it holds for all large enough $n \in \mathbb{N}$ that

$$\mathbf{E}[u_n^*(C_n^i, C_n^{-i})] \geq \mathbf{E}[u_n^*(\tilde{C}_n^i, C_n^{-i})] + \varepsilon(n) \ ,$$

where ε is a negligible function.

We show that if one-way functions do not exist, then in repeated two-player zero-sum games there are no computational Nash equilibria in which the players' strategies use random strings of length sub-linear in the number of the stages.

Our result follows by showing that finding efficiently a best response to the opponent's strategy that uses limited randomness can be seen as a special case of the problem of *learning an adaptively changing distribution* (introduced by Naor and Rothblum [NR06]). The goal in their framework is for a learner to recover a secret state used to sample a publicly observable distribution, in order to be able to predict the next sample. In particular, this would allow the learner to be competitive to someone who knows the secret state (Naor and Rothblum [NR06] considered this problem in the context of an adversary trying to impersonate someone in an authentication protocol). In the setting of repeated games, the random string used by the opponent's strategy can be thought of as the secret state. Note that learning it at any non-terminal history would give rise to efficient profitable deviation, since the player could just compute the next move of his opponent and play the best response to it.

Learning adaptively changing distributions. An adaptively changing distribution is given by a pair of algorithms $\mathcal{G}$ and $\mathcal{D}$ for generating an initial state and sampling. The algorithm $\mathcal{G}$ is a randomized function $\mathcal{G} : R \to S_p \times S_{init}$ that outputs an initial public state p_0 and a secret state s_0. The sampling algorithm $\mathcal{D}$ is a randomized function $\mathcal{D} : S_p \times S_s \times R \to S_p \times S_s$ that at each stage takes the current public and secret states, updates its secret sate and outputs a new public state. A learning algorithm $\mathcal{L}$ for $(\mathcal{G}, \mathcal{D})$ is given the initial public state p_0 ($\mathcal{L}$ does not get the initial secret state s_0) and at each round i: i) $\mathcal{L}$ either outputs prediction of the conditional distribution $D_{i+1}^{s_0}(p_0, \ldots, p_i)$ of the public output of $\mathcal{D}$ given the initial secret s_0 and the observed public states $p_0, \ldots, p_i$, or ii) $\mathcal{L}$ proceeds to round $i+1$ after observing a new public state $p_{i+1} \leftarrow D_{i+1}^{s_0}(p_0, \ldots, p_i)$. The goal of the learning algorithm is to output a hypothesis (in a form of a distribution) that is with high probability close in statistical distance to $D_{i+1}^{s_0}(p_0, \ldots, p_i)$. In other words, $\mathcal{L}$ is trying to be competitive to somebody who knows the initial secret state s_0. In the setting where $\mathcal{G}, \mathcal{D}$ are efficiently constructible Naor and Rothblum [NR06] gave an algorithm $\mathcal{L}$ that learns s_0 in probabilistic polynomial time provided that one-way functions do not exist. Moreover, their algorithm outputs a hypothesis after seeing a number of samples proportional to the entropy of the initial secret state.

Theorem 3 (Naor and Rothblum [NR06])**.** *Almost one-way functions exist if and only if there exists an adaptively changing distribution $(\mathcal{G}, \mathcal{D})$ and polynomials $\varepsilon(n), \delta(\epsilon)$ such that it is hard to $(\delta(n), \epsilon(n))$-learn the adaptively changing distribution $(\mathcal{G}, \mathcal{D})$ with $O\left(\delta^{-2}(n) \cdot \varepsilon^{-4}(n) \cdot \log |S_{init}|\right)$ samples.*

The strategy of the column player (Colin) with limited randomness gives rise to a natural adaptively changing distribution and we show that the algorithm of Naor and Rothblum [NR06] can be used to construct a computationally efficient strategy for the row player (Rowena) that achieves utility noticeably larger than the value of the stage game. Hence, if one-way functions do not exist, then in repeated two-player strategic games there are no computational Nash equilibria with strategies that use sub-linear randomness in the number of the stages.

Theorem 4. *Let $G = \langle (A_1, A_2), u \rangle$ be a two-player zero-sum strategic game with no weakly dominant pure strategies and with value v. If almost one-way functions do not exist then for any strategy $\{C_n\}_{n \in \mathbb{N}}$ of Colin in the repeated game of G that uses $o(n)$ random bits, there exists a polynomial time strategy of Rowena with expected average utility $v + \delta(n)$ against $\{C_n\}_{n \in \mathbb{N}}$ for some noticeable function δ.*

Proof. Let $\{C_n\}_{n \in \mathbb{N}}$ be an arbitrary strategy of Colin that takes $s(n) \in o(n)$ random bits. Let μ be the minmax strategy of Rowena in G. We define the following adaptively changing distribution $(\mathcal{G}, \mathcal{D})$. The generating algorithm $\mathcal{G}$ on input 1^n outputs a random string of length $s(n)$ as the initial secret state s_0 and the initial history $\emptyset$ of the n-stage repeated game of G as the initial public state p_0. The sampling algorithm $\mathcal{D}$ outputs the new secret state s_{i+1} identical to the secret state s_i that it received as an input (i.e., the secret state remains fixed as the $s(n)$ random coins s_0) and updates the input public state p_i in the following way. The sampling algorithm parses p_i as a history of length i in the n-stage repeated game of G and computes Colin's action $c_i = C_n(p_i, s_i)$ at p_i using randomness s_i. $\mathcal{D}$ additionally samples Rowena's action $r_i \leftarrow \mu$ according to her minmax strategy and then outputs the history $(p_i, (r_i, c_i))$ of length $i + 1$ as the new public state p_{i+1}. Note that after sampling the initial secret state s_0 the only randomness used by $\mathcal{D}$ is to sample the minmax strategy of Rowena.

11

It follows from Theorem 3 that there exists an efficient learning algorithm $\mathcal{L}$ that after at most $k = k(n) \in O(s(n) \cdot \delta^{-2}(n)\epsilon^{-4}(n))$ samples from $\mathcal{D}$ outputs a hypothesis h such that $\Pr[\text{SD}(D_{k+1}^{s_0}, D_{k+1}^h) \leq \epsilon(n)] \geq 1 - \delta(n)$. Consider the strategy of Rowena that uses $\mathcal{L}$ in order to learn Colin's random coins. In particular, a strategy that at each stage i runs $\mathcal{L}$ on the current history p_{i-1} and if $\mathcal{L}$ outputs some hypothesis h then the strategy plays the best response to Colin's action at stage i sampled according to D_{i+1}^h; and otherwise it plays according to Rowena's minmax strategy μ. This strategy can be efficiently implemented and it achieves expectation at least v in the $n-1$ stages in which Rowena plays according to her minmax strategy.[4] It remains to show that Rowena has a noticeable advantage over the value of the game at the stage in which $\mathcal{L}$ outputs the hypothesis h about s_0 and Rowena selects her strategy as the best response to Colin's action sampled according to D_{k+1}^h.

First, note that since G has no weakly dominant strategies, the best response to any pure action a_2 of Colin achieves a positive advantage over the value of the game. This observation follows from the fact that Rowena's minmax strategy achieves expectation at least v against any action of Colin and from the fact that the minmax strategy must be mixed (as there are no weakly dominant strategies). By moving all the probability in the minmax strategy to the action with highest payoff given that Colin plays a_2, Rowena achieves a value strictly larger than v. Hence, there exists some constant e (depending only on G) such that if D_{k+1}^h is e-close in statistical distance to $D_{k+1}^{s_0}$ then the expectation of the best response against D_{k+1}^h achieves expectation at least $v + c$ for some constant $c > 0$. Moreover, it is good enough if $\mathcal{L}$ outputs such h with probability at least $1 - \delta$ for some constant $\delta > 0$. Since ϵ and δ can be constant, for all large enough n the learning algorithm $\mathcal{L}$ outputs the hypothesis after receiving at most $k < n$ samples which allows Rowena to get expectation at least $v + \frac{1}{n}c$. □

It follows from Theorem 4 that if one-way functions do not exist, then there is no computational Nash equilibrium of repeated two-player zero-sum games where one of the players uses random strings of length sub-linear in the number of stages.

Corollary 1. *Let $G = \langle (A_1, A_2), u \rangle$ be a two-player zero-sum strategic game with no weakly dominant pure strategies and with value v. If almost one-way functions do not exist then there is no computational Nash equilibrium of the repeated game of G in which strategy of one of the players uses $o(n)$ random bits.*

Proof. Assume that there exists a computational Nash equilibrium $(\{C_n^1\}_{n \in \mathbb{N}}, \{C_n^2\}_{n \in \mathbb{N}})$ of $\{G^n\}_{n \in \mathbb{N}}$, the repeated game of G, in which the strategy of one of the players uses random strings of length $o(n)$. Without loss of generality, let Colin be the player with strategy that uses sub-linear randomness in the number of stages.

Denote by $w(n)$ the expectation of Rowena in this computational Nash equilibrium, i.e., for all $n \in \mathbb{N}$, $w(n) = \mathbf{E}[u_n^*(C_n^1, C_n^2)]$. First, consider the case when $w(n) \leq v + \eta(n)$ for some negligible function η. By Theorem 4 there exists a polynomial-time strategy of Rowena that achieves expectation $v + \delta(n)$ against $\{C_n^2\}_{n \in \mathbb{N}}$ for some noticeable function δ. Thus, this strategy constitutes Rowena's computationally efficient deviation to the above strategy profile that is profitable by some non-negligible amount. Second, consider the case when $w(n) = v + \delta(n)$ for some noticeable function δ. Colin can efficiently approximate the strategy that at each stage achieves his minmax payoff

---

[4]Note that if $\mathcal{L}$ does not output a hypothesis at the current stage, then Rowena chooses her action according to the same distribution as in $\mathcal{D}$, her minmax strategy, and her expectation is v.

profile in the stage game to achieve expected payoff in the repeated game at least $-v - \eta(n)$, where η is a negligible function. Such strategy constitutes Colin's computationally efficient deviation that achieves non-negligible advantage over the above utility profile. In both cases, $(\{C_n^1\}_{n \in \mathbb{N}}, \{C_n^2\}_{n \in \mathbb{N}})$ is not a computational Nash equilibrium of the repeated game of G. □

5   Strong Exploitation of Low-Entropy Opponents

We showed in the previous sections that equilibrium strategies in repeated two-player zero-sum games (both with or without restrictions on the computational power of the players) require entropy at least linear in the number of stages. A natural approach for enabling equilibria that require lower amount of randomness might be to relax the solution concept and consider ε-Nash equilibria, i.e., to ask what is the amount of randomness necessary for equilibrium strategies when the players are indifferent to improvements in utility smaller than ε.

As can be seen from the following argument, an equivalent question is how much can a player exploit an opponent that uses a strategy of low-entropy. Let α be an entropy level such that Rowena can exploit any Colin's strategy of entropy below α by more than ϵ (i.e., she can achieve expected utility in the repeated game improving by at least ε over the value of the stage game). Then in any ε-Nash equilibrium of the repeated game the strategy of the column player must be of entropy at least α.

5.1   Computationally Unbounded Players

The performance of strategies with bounded entropy in repeated two-player zero-sum games was previously studied in the standard setting with players that do not face any computational limitations. Towards this direction, Neyman and Okada [NO99] introduced a notion of *strategic entropy* in the context of repeated two-player zero-sum games in order to analyze repeated games played by bounded automata or players with bounded recall. Subsequently, [NO00] gave an asymptotic characterization of the value of repeated two-player zero-sum games when one of the players is restricted to strategies of bounded strategic entropy. In particular, they showed that if the row player can use strategies of strategic entropy at most γn, then in the n-stage game she can guarantee expected average utility at most $(\text{cav } U)(\gamma)$; where $U(\gamma)$ is the maximal expected utility the row player can guarantee in the stage game by a strategy of entropy at most γ, and cav U is the concavification of U (i.e., the smallest concave function larger or equal to U for all $\gamma \geq 0$).

Repeated matching pennies.   For the special case of the repeated game of matching pennies (given in Figure 3), Budinich and Fortnow [BF11] noticed a smooth tradeoff between the amount of entropy available to players and the necessary relaxation of the Nash equilibrium solution concept. In particular, they showed that in any ε-Nash equilibrium of the n-stage repeated game of matching pennies the players must use strategies of entropy at least $(1 - \varepsilon)n$ (for all $0 \leq \varepsilon \leq 1$). Their result follows by observing that in the n-stage game of matching pennies for all $0 \leq \varepsilon \leq 1$, the best response of the column player to any strategy of the row player of entropy at most $(1 - \epsilon)n$ achieves expected utility at least ϵ. This observation can be derived from the result of Neyman and Okada [NO00] by noticing that in the one-shot game of matching pennies $(\text{cav } U)(1 - \varepsilon) = -\varepsilon$. Hence, in the n-stage game of matching pennies the row player can guarantee for herself average expected utility at most

$(\operatorname{cav} U)(1 - \varepsilon) = -\varepsilon$ by a strategy of entropy at most $(1 - \varepsilon)n$, and equivalently the column player can achieve expectation at least ε.

In fact, the result of Neyman and Okada [NO00] implies that the relation between ε-Nash equilibria and the entropy of the players' strategies can be extended to all repeated two-player zero-sum games.

Theorem 5. *Let $G = \langle (A_1, A_2), u \rangle$ be a two-player zero-sum strategic game of value v and let $\beta > 0$ denote the minimal entropy of a minmax strategy for the column player in G. For any $0 < \varepsilon \leq 1$, there exists $c > 0$ such that if σ is a strategy of the column player of entropy $(1 - \varepsilon)\beta n$ in the n-stage repeated game of G then the row player has a deterministic strategy that achieves average payoff of at least $v + c$ against σ.*

For completeness we provide the proof of Theorem 5 in Appendix A.

Limits on exploiting a low-entropy opponent in non-zero-sum games.   In repeated non-zero-sum games, unlike in repeated two-player zero-sum games, it is in general not possible for a player to always achieve utility strictly above his minmax level given that his opponent uses low-entropy strategy. We illustrate this phenomenon on the game G given by the payoff matrix in Figure 1 that we discussed in Section 3. Note that if Colin plays his pure action "left" then Rowena gets utility 0, her minmax payoff, irrespective of her strategy. Even though Colin needs at least one random bit to play his equilibrium strategy in G, Rowena cannot benefit from the imperfect play of her opponent at all. Note that this limitation occurs even if any strategy of Colin in a Nash equilibrium of the repeated game of G must use randomness linear in the number of stages.

5.2   Computationally Efficient Players

Our results from Section 4 (i.e., Theorem 4) show that if one-way functions do not exist, then it is possible to efficiently gain a noticeable advantage over an opponent that uses randomness sub-linear in the number of the stages. We find it as an intriguing open problem to show a stronger version of Theorem 4 analogous to know results in the setting with computationally unbounded players (i.e., Theorem 5). In particular, to show that it is possible to efficiently gain *a constant advantage* over an opponent that uses randomness sub-linear in the number of the stages (even for the special case of the repeated game of matching pennies).

References

[AH03]   Robert J. Aumann and Sergiu Hart. Long cheap talk. *Econometrica*, 71(6):1619–1660, 2003.

[BF11]   Michele Budinich and Lance Fortnow. Repeated matching pennies with limited randomness. In *Proceedings 12th ACM Conference on Electronic Commerce (EC-2011), San Jose, CA, USA, June 5-9, 2011*, pages 111–118, 2011.

[BK85]   Jean-Pierre Benoît and Vijay Krishna. Finitely repeated games. *Econometrica*, 53(4):905–922, 1985.

[BK87]   Jean-Pierre Benoît and Vijay Krishna. Nash equilibria of finitely repeated games. *International Journal of Game Theory*, 16(3):197–204, 1987.

[BM84]   Manuel Blum and Silvio Micali. How to generate cryptographically strong sequences of pseudo-random bits. *SIAM J. Comput.*, 13(4):850–864, 1984.

[DHR00]   Yevgeniy Dodis, Shai Halevi, and Tal Rabin. A cryptographic solution to a game theoretic problem. In *Advances in Cryptology - CRYPTO 2000, 20th Annual International Cryptology Conference, Santa Barbara, California, USA, August 20-24, 2000, Proceedings*, pages 112–130, 2000.

[Gol01]   Oded Goldreich. *The Foundations of Cryptography - Volume 1, Basic Techniques*. Cambridge University Press, 2001.

[Gon06]   Julio González-Díaz. Finitely repeated games: A generalized nash folk theorem. *Games and Economic Behavior*, 55(1):100–111, 2006.

[HILL99] Johan Håstad, Russell Impagliazzo, Leonid A. Levin, and Michael Luby. A pseudorandom generator from any one-way function. *SIAM J. Comput.*, 28(4):1364–1396, 1999.

[HN10]   Ran Halprin and Moni Naor. Games for extracting randomness. *ACM Crossroads*, 17(2):44–48, 2010.

[HNU15]   Pavel Hubáček, Moni Naor, and Jonathan Ullman. When can limited randomness be used in repeated games? In *Algorithmic Game Theory - 8th International Symposium, SAGT 2015, Saarbrücken, Germany, September 28-30, 2015, Proceedings*, pages 259–271, 2015.

[HP14]   Joseph Y. Halpern and Rafael Pass. Algorithmic rationality: Game theory with costly computation. *Journal of Economic Theory*, 2014.

[IL89]   Russell Impagliazzo and Michael Luby. One-way functions are essential for complexity based cryptography (extended abstract). In *30th Annual Symposium on Foundations of Computer Science, Research Triangle Park, North Carolina, USA, 30 October - 1 November 1989*, pages 230–235, 1989.

[Imp92]   Russell Impagliazzo. *Pseudo-random generators for cryptography and for randomized algorithms*. PhD thesis, PhD thesis, University of California, Berkeley, 1992.

[KU07]   Shankar Kalyanaraman and Christopher Umans. Algorithms for playing games with limited randomness. In *Algorithms–ESA 2007*, pages 323–334. Springer, 2007.

[NO99]   Abraham Neyman and Daijiro Okada. Strategic entropy and complexity in repeated games. *Games and Economic Behavior*, 29(1):191–223, 1999.

[NO00]   Abraham Neyman and Daijiro Okada. Repeated games with bounded entropy. *Games and Economic Behavior*, 30(2):228–247, 2000.

[NR06]   Moni Naor and Guy N. Rothblum. Learning to impersonate. In *Machine Learning, Proceedings of the Twenty-Third International Conference (ICML 2006), Pittsburgh, Pennsylvania, USA, June 25-29, 2006*, pages 649–656, 2006.

[OR94]   Martin J. Osborne and Ariel Rubinstein. *A course in game theory*. MIT press, 1994.

[Yao82]   Andrew Chi-Chih Yao. Theory and applications of trapdoor functions (extended abstract). In *23rd Annual Symposium on Foundations of Computer Science, Chicago, Illinois, USA, 3-5 November 1982*, pages 80–91, 1982.

A Exploiting Low Entropy in Two-Player Zero-Sum Games

In this appendix we provide the proof of Theorem 5 that establishes that if one player uses a constant fraction less randomness in the repeated two-player zero-sum game, then the other player can obtain an average payoff that is larger than the value of the stage game by a constant.

We use the following lemma about performance of low-entropy strategies in two-player zero-sum games in the proof of Theorem 5.

Lemma 1. *Let $G = \langle (A_1, A_2), u \rangle$ be a two-player zero-sum strategic game of value v and let $\beta > 0$ denote the minimal entropy of a minmax strategy for the column player in G. For every $\varepsilon > 0$, there exists $c_\varepsilon > 0$ such that if σ is a strategy of the column player of entropy $(1 - \varepsilon)\beta$ then the row player has a strategy that achieves utility at least $v + c_\varepsilon$ against σ.*

Proof. Let σ be an arbitrary strategy of Colin in G of entropy $(1 - \epsilon) \cdot \beta$ for some $\varepsilon > 0$, and let ρ_σ denote the best response strategy of Rowena to σ. First, we show that Rowena's expected utility $\mathbf{E}[u(\rho_\sigma, \sigma)]$ is at least $v + c$ for some $c > 0$. Suppose to the contrary that Rowena's best response to σ achieves expectation at most v. Let $\hat{\rho}$ be the minmax strategy of Rowena in G, the profile $(\hat{\rho}, \sigma)$ is a Nash equilibrium of G: Rowena's minmax strategy guarantees at least the value of the game v. On the other hand, by the hypothesis her best response to σ achieves at most v, so Rowena's expectation in $(\hat{\rho}, \sigma)$ is equal to v. There are no profitable deviations for Colin, since he cannot decrease Rowena's expectation below v given that she plays according to her minmax strategy. The strategy σ of Colin is of entropy $(1 - \varepsilon) \cdot \beta < \beta$, and the strategy profile $(\hat{\rho}, \sigma)$ is a Nash equilibrium of G contradicting that β is the minimal entropy of Colin's strategy in any Nash equilibrium of G. Hence, the best response to σ must increase Rowena's expectation by a non-zero amount over v. The statement of the lemma follows by setting c_ε to be the infimum of the set of all c achieved against Colin's strategies of entropy $(1 - \varepsilon) \cdot \beta$. $\square$

Theorem 5. *Let $G = \langle (A_1, A_2), u \rangle$ be a two-player zero-sum strategic game of value v and let $\beta > 0$ denote the minimal entropy of a minmax strategy for the column player in G. For any $0 < \varepsilon \leq 1$, there exists $c > 0$ such that if σ is a strategy of the column player of entropy $(1 - \varepsilon)\beta n$ in the n-stage repeated game of G then the row player has a deterministic strategy that achieves average payoff of at least $v + c$ against σ.*

Proof. Let σ be an arbitrary strategy of the column player (Colin) of Shannon entropy $n \cdot \beta(1 - \varepsilon)$ for some $\varepsilon \in [0, 1]$. Let ρ_σ be the strategy of the row player (Rowena) that at each non-terminal history a plays the best response in G to Colin's strategy $\sigma(a)$. Rowena's expectation $\mathbf{E}_{a \leftarrow (\rho_\sigma, \sigma)}[u^*(a)]$ is

$$\frac{1}{n} \left(\mathop{\mathbf{E}}_{a \leftarrow (\rho_\sigma, \sigma)}[u(a^1)] + \mathop{\mathbf{E}}_{a \leftarrow (\rho_\sigma, \sigma)}[u(a^2)|a^1] + \cdots + \mathop{\mathbf{E}}_{a \leftarrow (\rho_\sigma, \sigma)}[u(a^n)|(a^1, \ldots, a^{n-1})] \right) .$$

By the definition of conditional expectation, we rewrite her expectation as a summation over all terminal histories, i.e.,

$$\frac{1}{n} \left(\sum_{b \in A^n} (\rho_\sigma, \sigma)(b) \cdot \left(\mathop{\mathbf{E}}_{a \leftarrow (\rho_\sigma, \sigma)}[u(a^1)] + \mathop{\mathbf{E}}_{a \leftarrow (\rho_\sigma, \sigma)}[u(a^2)|a^1 = b^1] + \cdots \right.\right.$$

$$\left.\left. + \mathop{\mathbf{E}}_{a \leftarrow (\rho_\sigma, \sigma)}[u(a^n)|(a^1, \ldots, a^{n-1}) = (b^1, \ldots, b^{n-1})] \right) \right) . \quad (1)$$

16

Note that for every terminal history $b \in A^n$ the summands correspond to the expectation of Rowena at the non-terminal subhistories of b. For any terminal history $b \in A^n$, the total sum of entropy used in σ at the subhistories of b is at most $(1 - \varepsilon)\beta n$, which implies that there are at least $n' = n\left(1 - (1 - \varepsilon)/(1 - \frac{\varepsilon}{2})\right)$ subhistories of b where the Colin's strategy has entropy at most $(1 - \frac{\varepsilon}{2})\beta$. To see this assume that there exists a terminal history b with less than n' subhistories where σ uses entropy at most $(1 - \frac{\varepsilon}{2})\beta$. Then the total entropy of σ on all subhistories of b is strictly larger than

$$(n - n')\left(1 - \frac{\varepsilon}{2}\right)\beta = \left(n - n\left(1 - \frac{(1 - \varepsilon)}{(1 - \frac{\varepsilon}{2})}\right)\right)\left(1 - \frac{\varepsilon}{2}\right)\beta = (1 - \varepsilon)\beta n \ ,$$

a contradiction. As shown in Lemma 1, for each subhistory of b where Colin uses strategy of entropy at most $(1 - \frac{\varepsilon}{2})\beta$, Rowena's best response achieves at least $v + c$, where $c = c_{\varepsilon/2} > 0$ is a value determined by the game G (and a function of epsilon). On all other subhistories of b (with Colin's strategy of entropy larger than $(1 - \frac{\varepsilon}{2})\beta$) the value of Rowena is at least v. Therefore the total utility (the sum of the expectations over all subhistories of b) is at least $nv + c \cdot n' = n(v + c')$, where $c' = c\left(1 - (1 - \varepsilon)/(1 - \frac{\varepsilon}{2})\right) > 0$.

Since this holds for every terminal history of G^n, it follows from (1) that the strategy ρ_σ of Rowena achieves average expected utility at least $v + c'$ against σ in G^n. □

Note that the constant c by which the row player can exploit strategy of the column player of entropy $(1 - \varepsilon)\beta n$ is related to the possible gain of the row player in the stage game, given that the column player plays strategy of entropy $(1 - \varepsilon)\beta$. To make the connection explicit, we use the following notation from Neyman and Okada [NO00]. Let $G = \langle (A_1, A_2), u \rangle$ be the stage game and for $\gamma \geq 0$ define

$$U(\gamma) = \max_{\substack{\sigma \in \Delta(A_1) \\ H(\sigma) \leq \gamma}} \min_{a_2 \in A_2} \mathbf{E}[u(\sigma, a_2)].$$

Hence, $U(\gamma)$ is the maximal expected utility the row player can guarantee with a strategy of entropy at most γ; or equivalently, $-U(\gamma)$ is the minimal expected utility that the column player can achieve by a best response to any strategy of the row player of entropy at most γ. Note that $U(0)$ is equal to the row player's minmax level in pure strategies, and for all $\gamma \geq 0$, $U(\gamma)$ is at most the value of the game. Using this notation the statement of Theorem 5 can be restated as:

Theorem 5 (restated). *Let $G = \langle (A_1, A_2), u \rangle$ be a two-player zero-sum strategic game of value v and let $\beta > 0$ denote the minimal entropy of a minmax strategy for the row player in G. For any $0 < \varepsilon \leq 1$, if σ is a strategy of the row player of entropy $(1 - \varepsilon)\beta n$ in the n-stage repeated game of G then the column player has a deterministic strategy that achieves average payoff of at least $-v - \left(1 - \frac{(1 - \varepsilon)}{(1 - \frac{\varepsilon}{2})}\right) U((1 - \frac{\varepsilon}{2})\beta)$ against σ.*

We remark that an improved bound on the expectation can be obtained using the technique of Neyman and Okada and the column player can in fact achieve average expected utility at least $-v - (\text{cav}\, U)((1 - \varepsilon)\beta)$, where $\text{cav}\, U$ is the smallest concave function larger or equal than U.

Theorem 6 below can be seen as a "converse" of Theorem 5. Specifically, we show that even if the players are restricted to strategies of entropy $(1 - \varepsilon)\beta n$ then there exists an ε'-Nash equilibrium of G^n for some ε' proportional to ε.

Theorem 6. *Let G be a two-player zero-sum strategic game such that the minimal entropy of a minmax strategy is $\beta > 0$ for both players. There exists $c > 0$ such that for all $0 < \varepsilon \leq 1$ and for all n, there exists a $\left(c \cdot \frac{\lceil n\varepsilon \rceil + 1}{n}\right)$-Nash equilibrium of the n-stage repeated game of G in which the players' strategies are of entropy at most $(1 - \varepsilon)\beta n$.*

Proof. Let σ be the strategy profile in the n-stage repeated game of G in which the players play in the first $\lfloor n(1-\varepsilon) \rfloor$ stages according to their minmax strategies of minimal entropy (i.e., entropy β), and in the remaining $\lceil n\varepsilon \rceil$ stages the players alternate between playing the (pure) action profiles $a^* \in A_1 \times A_2$ and $a^\dagger \in A_1 \times A_2$, such that $p^* = u(a^*)$ is the maximum payoff of Rowena in G and $p^\dagger = u(a^\dagger)$ is the minimal payoff of Rowena in G. Note that by construction of σ, the players use strategies of entropy at most $n \cdot \beta(1 - \varepsilon)$.

Assume $\lceil n\varepsilon \rceil$ is odd (the argument for $\lceil n\varepsilon \rceil$ even is analogous). The expected utility of Rowena in σ in the n-stage repeated game of G is

$$\mathbf{E}[u^*(\sigma)] = \frac{1}{n}\left(\lfloor n(1 - \varepsilon) \rfloor \cdot v + \frac{1}{2}(\lceil n\varepsilon \rceil - 1)(p^* + p^\dagger) + p^*\right),$$

where v is the value of G. The expectation of every deviating strategy σ'_2 of Colin is

$$-\mathbf{E}[u^*(\sigma_1, \sigma'_2)] \leq \frac{1}{n}\left(-\lfloor n(1 - \varepsilon) \rfloor \cdot v + \frac{1}{2}(\lceil n\varepsilon \rceil - 1)(-p^\dagger - p^\dagger) - p^\dagger\right),$$

hence Colin can increase his utility by at most $\frac{1}{2n}(p^* - p^\dagger)(\lceil n\varepsilon \rceil + 1)$. Similarly, the increase in expectation from any deviating strategy of Rowena can be upper bounded by $\frac{1}{2n}(p^* - p^\dagger)(\lceil n\varepsilon \rceil - 1)$. Therefore, σ is a $\left(c \cdot \frac{\lceil n\varepsilon \rceil + 1}{n}\right)$-Nash equilibrium of the n-stage repeated game of G for $c = \frac{1}{2}(p^* - p^\dagger)$, and the statement of the proposition follows since $\frac{1}{2}(p^* - p^\dagger)$ is a constant independent of ε and n. $\square$

B Matching Pennies

The game of *matching pennies* is a two-player zero-sum strategic game given by the payoff matrix in Figure 3. Both players can either play Heads (H) or Tails (T). The only Nash equilibrium is the strategy profile $(\frac{1}{2}H + \frac{1}{2}T, \frac{1}{2}H + \frac{1}{2}T)$ in which both players randomize uniformly over H and T. By Theorem 1, in the equilibrium for the n-stage repeated game of matching pennies both

	Heads (H)	Tails (T)
Heads (H)	$1, -1$	$-1, \ 1$
Tails (T)	$-1, \ 1$	$1, -1$

Figure 3: The payoff matrix of the game of matching pennies.

players randomize uniformly between playing Heads and Tails at each stage, and the entropy of the equilibrium strategy of each player is exactly n.

We now give a generalization of Lemma 3.1 from Budinich and Fortnow [BF11].

Theorem 7. *For any $\varepsilon \in [0, 1]$, let σ be a strategy of the column player of entropy $n(1 - \varepsilon)$ in the n-stage repeated game of matching pennies. The row player has a deterministic strategy that achieves payoff of at least ε against σ.*

Proof. Let σ be an arbitrary strategy of Colin of Shannon entropy $n(1 - \varepsilon)$ for some $\varepsilon \in [0, 1]$. Let ρ_σ be the strategy of Rowena that at each non-terminal history a plays the best response to Colin's strategy $\sigma(a)$. We can express Rowena's expectation $\mathbf{E}_{a \leftarrow (\rho_\sigma, \sigma)}[u^*(a)]$ as

$$\frac{1}{n} \left(\underset{a \leftarrow (\rho_\sigma, \sigma)}{\mathbf{E}} [u(a^1)] + \underset{a \leftarrow (\rho_\sigma, \sigma)}{\mathbf{E}} [u(a^2)|a^1] + \cdots + \underset{a \leftarrow (\rho_\sigma, \sigma)}{\mathbf{E}} [u(a^n)|(a^1, \ldots, a^{n-1})] \right) ,$$

which can be rewritten due to the definition of conditional expectation as a summation over terminal histories

$$\frac{1}{n} \left(\sum_{b \in A^n} (\rho_\sigma, \sigma)(b) \cdot \left(\underset{a \leftarrow (\rho_\sigma, \sigma)}{\mathbf{E}} [u(a^1)] + \underset{a \leftarrow (\rho_\sigma, \sigma)}{\mathbf{E}} [u(a^2)|a^1 = b^1] + \cdots \right. \right.$$
$$\left. \left. + \underset{a \leftarrow (\rho_\sigma, \sigma)}{\mathbf{E}} [u(a^n)|(a^1, \ldots, a^{n-1}) = (b^1, \ldots, b^{n-1})] \right) \right) .$$

For every terminal history $b = (b_1, \ldots, b_n)$, the total entropy of σ over the non-terminal subhistories of b is bounded by $n(1 - \varepsilon)$, i.e.,

$$H(\sigma(\emptyset)) + \sum_{i=1}^{n-1} H(\sigma(b_1, \ldots, b_i)) \le n(1 - \varepsilon) . \tag{2}$$

We define $\varepsilon_0 = 1 - H(\sigma(\emptyset))$ and for every $i \in \{1, \ldots, n-1\}$ we define $\varepsilon_i = (1 - H(\sigma(b_1, \ldots, b_i)))$. Note that $0 \le \varepsilon_i \le 1$ for every $i \in \{0, \ldots, n-1\}$ and from inequality (2) we get that $\varepsilon \le \frac{1}{n} \sum_{i=0}^{n-1} \varepsilon_i$. In order to conclude that Rowena's expected utility in the strategy profile (ρ_σ, σ) is at least ϵ, it is sufficient to show that for every subhistory b' of b the expectation $\mathbf{E}[u(\rho_\sigma(b'), \sigma(b'))]$ is least $1 - H(\sigma(b'))$.

For an arbitrary non-terminal history h, consider Rowena's expectation in G given the strategy profile $(\rho_\sigma(h), \sigma(h))$. Since $\rho_\sigma(h)$ is the best response to $\sigma(h)$, Rowena's expectation is $2p - 1$, where p is the probability of Colin's most probable action at history h. We need to show that for all $p \in [1/2, 1]$

$$2p - 1 \ge 1 - H(\sigma(h)) = 1 + p \log_2(p) + (1 - p) \log_2(1 - p) .$$

For p equal $1/2$ or 1, the left side and the right side of the inequality are equal. Since $2p - 1$ is a linear function and $1 + p \log_2(p) + (1 - p) \log_2(1 - p)$ is a convex function on $[1/2, 1]$, the inequality holds. This concludes the proof. $\square$

It follows form Theorem 7 that if the players can use only strategies of entropy $(1 - \epsilon)n$ (i.e., lower than n-times the entropy of an equilibrium of the single-shot matching pennies) then Nash equilibria in the n-stage repeated game of matching pennies do not exist.

Proposition 2. *Let G^n be the n-stage repeated game of matching pennies.*

1. *For all $0 \le \varepsilon \le 1$, if σ is an ε-Nash equilibrium of G^n then the players' strategies in σ are of entropy at least $n(1 - \varepsilon)$.*

2. *For all $0 \le \varepsilon \le 1$, there exists an $(\varepsilon + \frac{2}{n})$-Nash equilibrium of G^n in which the players' strategies are of entropy at most $(1 - \varepsilon)n$.*

19

Proof. First, we show that any ε-Nash equilibrium σ in the n-stage repeated game of matching pennies comprises of strategies of entropy at least $(1 - \varepsilon)n$. Assume that there is an ε-Nash equilibrium in which both players use a strategy of strictly smaller entropy than $(1 - \varepsilon)n$, i.e., of entropy $(1 - \varepsilon')n$ for some $\varepsilon' > \varepsilon$. By Theorem 7, each player i has a strategy σ_i' that achieves at least ε' against σ_{-i}. Since σ is an ε-Nash equilibrium then for any player i

$$\mathbf{E}[u_i^*(\sigma)] \geq \mathbf{E}[u_i^*(\sigma_i', \sigma_{-i})] - \varepsilon \geq \varepsilon' - \varepsilon > 0 .$$

This implies that for both players $\mathbf{E}[u_i^*(\sigma)] > 0$, however it cannot be the case that the expectation of both players is strictly larger than zero, since matching pennies is a zero-sum game.

Second, we show that if the players can use strategies of entropy $(1 - \varepsilon)n$ then there exists an $(\varepsilon + \frac{2}{n})$-Nash equilibrium of the n-stage repeated game of matching pennies. To see this, consider a strategy profile in which the players play uniformly at random H and T in the first $\lfloor (1-\varepsilon)n \rfloor$ stages and in the remaining $\lceil \varepsilon n \rceil$ stages Rowena plays always H and Colin alternates between T and H (i.e., the outcome at stage $\lfloor (1-\varepsilon)n \rfloor + 1$ is (H, T)). If $\lceil \varepsilon n \rceil$ is odd then Rowena's expectation is $-\frac{1}{n}$ and otherwise it is 0. Both Colin and Rowena can improve their expectation only in the last $\lceil \varepsilon n \rceil$ stages by matching/countering the opponent, but any such deviation can achieve utility at most

$$\frac{\lceil \varepsilon n \rceil}{n} \leq \frac{\varepsilon n + 1}{n} \leq \varepsilon + \frac{1}{n} .$$

Hence, both players can improve the utility by at most $\varepsilon + \frac{2}{n}$ by deviating from the prescribed strategy profile, and it constitutes an $(\varepsilon + \frac{2}{n})$-Nash equilibrium. $\qquad \square$

B.1   Matching Pennies with Computationally Efficient Players

In this section we prove the statement of Theorem 4 for the special case of the game of matching pennies without relying on the framework of adaptively changing distributions of Naor and Rothblum [NR06], but using the classical results on pseudorandomness discussed in Section 2.2. In particular, that if one-way functions do not exist, then the players cannot efficiently generate unpredictable sequences of bits using only a few truly random bits. Hence, in the repeated game of matching pennies any player can at some stage efficiently predict and exploit the next move of an opponent that uses amount of random bits sub-linear in the number of stages.

Theorem 8. *If one-way functions do not exist then for any polynomial-size circuit family $\{C_n\}_{n \in \mathbb{N}}$ implementing a strategy of Colin in the repeated game of matching pennies using at most $n - 1$ random bits, there exists a polynomial time strategy of Rowena with expected utility $\delta(n)$ against C_n for some noticeable function δ.*

Proof. Let $\{X_n\}_{n \in \mathbb{N}}$ be a probability ensemble defined for all n as the random variable over $2n$-bit strings corresponding to the terminal histories in the n-stage repeated matching pennies (where H corresponds to 0 and T to 1) when Rowena plays uniformly at random and Colin plays according to C_n. Note that X_n is of length $2n$ and it can be generated in polynomial time given at most $2n - 1$ random bits, since Colin's strategy uses random strings of length at most $n - 1$.

Since one-way functions do not exist, the ensemble $\{X_n\}_{n \in \mathbb{N}}$ cannot be pseudorandom. In particular, it cannot be unpredictable in polynomial time in the following sense. There exists a polynomial time predictor algorithm A that reads $x \leftarrow X_n$ bit by bit and succeeds in predicting the next value with probability noticeably larger than one half. Formally, let $\mathsf{next}_A(x)$ be a function

that returns the i-th bit of x if on input $(1^{|x|}, x)$ algorithm A reads only the first $i - 1 < |x|$ bits of x, and returns a uniformly chosen bit in case A reads the entire string x. There exists a predictor algorithm A and some positive polynomial p, such that

$$\Pr[A(1^{|X_n|}, X_n) = \mathsf{next}_A(X_n)] \geq \frac{1}{2} + \frac{1}{p(n)} \ ,$$

where the probability is taken over the randomness of A.

We show that Rowena can guarantee for herself at least noticeable expected utility by emulating A on the transcript of the repeated game. Consider the strategy R_A of Rowena that at each stage i samples a uniformly random bit r_i, and if $A(b_i)$ outputs any prediction c_i^* of Colin's action then Rowena plays c_i^* (to match Colin) and otherwise it plays r_i and uses the action played by Colin at stage i as the next input to A. After the stage in which A outputs a prediction R_A plays uniformly at random. The expectation of Rowena can be lower bounded in the following way:

$$\mathbf{E}[u^*(R_A, C)] \geq \frac{1}{n} \left(\Pr[A \text{ outputs } c_i^*] \cdot \left((n-1) \cdot 0 + 2 \left(\frac{1}{2} + \frac{1}{p(n)} \right) - 1 \right) \right.$$

$$\left. + (1 - \Pr[A \text{ outputs } c_i^*]) \cdot 0 \right) \ .$$

Recall that the actions of Rowena are chosen uniformly at random and the predictor A has to guess a uniformly random bit if it reads the whole terminal history $x \leftarrow X_n$. Hence, in order to gain noticeable advantage over one half, A must output its prediction to one of the actions of Colin with at least noticeable probability, i.e., $\Pr[A \text{ outputs } c_i^*]$ is at least $\delta'(n)$ for some noticeable function δ'. Thus, the strategy R_A achieves expectation at least $\delta(n) = \delta'(n) \cdot (n \cdot p(n))^{-1}$, which is a noticeable function of n. $\qquad\square$