Complexity and Mixed Strategy Equilibria*

Tai-Wei Hu†

Penn State University

10/25/2008

Abstract

Unpredictable behavior is central for optimal play in many strategic situations because a predictable pattern leaves a player vulnerable to exploitation. A theory of unpredictable behavior is presented in the context of repeated two-person zero-sum games in which the stage games have no pure strategy equilibrium. Computational complexity considerations are introduced to restrict players' strategy sets. The use of Kolmogorov complexity allows us to obtain a sufficient condition for equilibrium existence. The resulting theory has implications for the empirical literature that tests the equilibrium hypothesis in a similar context. In particular, the failure of some tests for randomness does not justify rejection of equilibrium play.

Keywords: Kolmogorov complexity; objective probability; frequency theory of probability; mixed strategy; zero-sum game; effective randomness

*I am grateful to Professor Kalyan Chatterjee for his guidance and encouragement. All remaining errors, of course, are my own.

†E-mail: tuh129@psu.edu; Corresponding address: Department of Economics, Penn State University, 608 Kern Graduate Building, University Park, PA 16802-3306.

1 Introduction

Unpredictable behavior is central for optimal play in many strategic situations because a predictable pattern leaves a player vulnerable to exploitation—think of the direction of tennis serves or soccer penalty kicks or the pattern of bluffing in poker. In fact, Walker and Wooders [36] for tennis and Palacios-Huerta [29] for soccer find evidence that supports such unpredictable behavior. In this paper, we propose a theory of unpredictable behavior in the context of (infinitely) repeated two-person zero-sum games in which the stage games have no pure strategy equilibrium. We focus on repeated play because equilibrium makes no predictions about a single play (von Neumann and Morgenstern [23], p. 147, expresses a similar concern). Recognizing this fact, there is a literature which identifies mixed strategies with beliefs and which makes predictions about *beliefs* in one-shot games (see, for example, Harsanyi [10] and Aumann and Brandenburger [2]). In contrast, our theory, based on computational complexity considerations, has predictions about players' *actions*. For example, in a repeated matching pennies game, the sequence of plays that alternates between heads and tails is not an equilibrium play in our theory—even though it could have resulted from an i.i.d. random process (and, in fact, is no more likely or unlikely than any other sequence). Moreover, our theory has implications for empirical tests of unpredictable behavior. We find that the failure of some tests for randomness does not justify rejection of equilibrium play.

We model players' computational power by introducing a computability constraint for each player, which is the set of functions that the player can use to implement strategies. We assume that the computability constraint includes all functions that can be computed with a Turing-machine alone, and we impose conditions on the constraint so that all implementable strategies of a player can be computed in a mechanical fashion. The resulting strategy set is countable. Our framework then has three ingredients: a finite zero-sum game as the stage game, and a computability constraint for each player. For payoffs, we adopt the long-run average criterion, which we will argue is more appropriate in this context, and we show that there is no equilibrium with the discounting criterion. Our first result is a *necessary* condition for equilibrium existence. If the stage game has no pure strategy equilibrium, then, to obtain equilibrium in the repeated game with computability constraints, it is necessary that each player's constraint contains a function outside his opponent's constraint. One corollary of this result is that if equilibrium exists, the equilibrium strategy of one player is not computable by the other.

To obtain a *sufficient* condition for existence, we use Kolmogorov complexity [13] to consider the complexity of functions, which can be identified with sequences over natural numbers, in the constraints. This complexity measures the minimal description length of a finite object by using functions in a given constraint as descriptions, and we use this measure to define *complex sequences*—sequences that are hard to describe. Specifically, a sequence is complex relative to a constraint if the Kolmogorov complexities (relative to that constraint) of its initial segments are essentially the lengths of those segments. A complex sequence is itself uncomputable, but it can be thought of as the limit of finite sequences that are hard to compute. The sufficient condition we find is called *mutual complexity*: it assumes that each player can compute a complex sequence relative to the other player's computability constraint. We show that, if the constraints are mutually complex, then for any mixed strategy equilibrium of the stage game, there is a corresponding equilibrium in the repeated game with the constraints. This result cannot be obtained with the complexity notion in the machine game literature (see Ben-Porath [3] and Osborn and Rubinstein [27]), which uses finite automata to implement strategies and measures complexity with the number of states in a player's automaton. Existence result in that literature still relies on the use of mixed strategies. We overcome this difficulty by considering uncomputable sequences (which are necessary for equilibrium) relative to the other player.

Our theory has implications for the empirical literature that tests the equilibrium hypothesis by implementing statistical tests for randomness. In that empirical literature, it is assumed that equilibrium behavior is so unpredictable to pass all such tests (see, for example, Brown and Rosenthal [4], Walker and Wooders [36], and Palacios-Huerta [29]). However, O'Neill [26] doubts the relevance of all such tests to reject the equilibrium hypothesis. To understand what tests are relevant, we follow Martin-Löf [19] to define idealized statistical tests relative to a computability constraint. Such a test is based on a property with zero probability that can be detected with the functions in the computability constraint. Mutual complexity turns out to be the necessary and sufficient condition for the existence of an equilibrium strategy for each player that passes all such tests (w.r.t. an equilibrium mixed strategy of the stage game) relative to the other player. However, this does not show that all tests are relevant to the equilibrium hypothesis. We find that, under mutual complexity, there are always equilibrium strategies that will fail some tests. In matching pennies, there is an equilibrium strategy that has more heads than tails in all its initial segments. Moreover, we find a notion of unpredictability, weaker than randomness, such that any sequence of play satisfying that notion is also an equilibrium strategy under mutual complexity. This notion corresponds to a proper subset of the above tests. This result suggests that only those tests are relevant to the equilibrium hypothesis.

The rest of the paper is organized as follows: in section 2 we formulate collective games (the version of repeated games we study) and present nonexistence results; section 3 has two parts: first we formulate the notion of complex sequences using Kolmogorov complexity and give an existence result; then we discuss the statistical tests that equilibrium behavior should pass; in section 4 we give some discussions of our results and further research; the proofs of the main theorems are in section 5.

2 Repeated games with computability constraints

In this section we formulate our framework formally. We consider two alternative formulations of the repeated game, which are called the *horizontal game* (HG) and the *vertical game* (VG), respectively. In general we may refer to either of them as a *collective game*. Both the two games consist of infinite repetitions of a finite zero-sum game, but they differ in their information structures. In both games we impose computability constraints on implementable strategies. Before the formulation, we give a necessary review of computability theory, of which our main reference is Pippenger [28]. We present nonexistence results in the end.

2.1 Computability

In this section we will review some material from computability theory. We view computability as a constraint on the set of functions available to a player. Following the computability theory literature, we consider the set of all partial functions:

$$\mathcal{F} = \{f : \overline{\mathbb{N}}^k \to \overline{\mathbb{N}} \mid k > 0;\ \text{if for some } i,\ x_i = \perp,\ f(x_1, ..., x_k) = \perp\},$$

where $\overline{\mathbb{N}} = \mathbb{N} \cup \{\perp\}$, and the symbol $\perp$ means that the function is not defined there. An important subset of $\mathcal{F}$ is the set of total functions

$$\mathcal{G} = \{f : \mathbb{N}^k \to \mathbb{N} \mid k \in \mathbb{N}, k > 0\},$$

which includes functions that are well-defined everywhere. For any set $\mathcal{P} \subset \mathcal{F}$, we use $\mathcal{P}_T$ to denote the set $\mathcal{P} \cap \mathcal{G}$.

In general, computational power can be expressed as a set $\mathcal{P} \subset \mathcal{F}$, interpreted as the set of functions that can be computed by a player. We impose four axioms on the set $\mathcal{P}$, following the modern development of computability theory. We borrow the formulations from Pippenger [28].

Definition 2.1. A set $\mathcal{P} \subset \mathcal{F}$ is called a *computability constraint*[1] if it satisfies the following conditions:

FC1 $\mathcal{P}$ is closed under composition.

FC2 $\mathcal{P}$ contains the following functions:

(a) The constant zero function Z defined by $Z(x) = 0$.

(b) The successor function S defined by $S(x) = x + 1$.

(c) The projection function P_{ki} defined by $P_{ki}(x_1, ..., x_k) = x_i$, for all $k, i \in \mathbb{N}_+$.

(d) The pair function o defined by $o(x, y) = (x + 1) + (x + y + 1)(x + y)/2$.

(e) The conditional function $\mathbf{c}$ defined by $\mathbf{c}(v, w, x, y) = x$ if $v = w$ and $\mathbf{c}(v, w, x, y) = y$ otherwise.

[1] It is also called a *reflexive class* in the literature.

FC3 $\mathcal{P}$ contains a function u that satisfies the following: for each function $f \in \mathcal{P}$ with k arguments, there is a number $e \in \mathbb{N}$ such that for all $x_1, ..., x_k \in \mathbb{N}$,

$$f(x_1, ..., x_k) = u_k(e, x_1, ..., x_k),$$

where the function u_k is defined inductively: $u_0 = u$ and for all $e, x_1, ..., x_{k+1} \in \mathbb{N}$,

$$u_{k+1}(e, x_1, ..., x_{k+1}) = u_k(o(e, x_1), x_2, ..., x_{k+1}).$$

FC4 $\mathcal{P}$ contains a total function m that satisfies the following two conditions:

(a) For all $e, t \in \mathbb{N}$, $0 \leq m(e, t) \leq m(e, t+1) \leq 1$.

(b) For all $e \in \mathbb{N}$, $u(e) \in \mathbb{N}$ if and only if for some $t \in \mathbb{N}$, $m(e, t) = 1$.

The set of all computability constraints is denoted by $\mathfrak{R}$.

Conditions FC1 and FC2 assume that the player is able to compute some simple functions and is able to perform compositions. Conditions FC3 and FC4 requires that the player has a universal machine (the function u) to perform all computations. In particular, FC4 requires any computation is finished in finite steps. These conditions are meant to axiomatize computable functions by Turing-machines with a fixed oracle. As a result, if $\mathcal{P}$ is a computability constraint, it includes all Turing-computable functions, and the set of all Turing-computable functions, denoted by $\mathcal{P}^*$, is the smallest computability constraint. In our framework, a player will be endowed with a computability constraint $\mathcal{P}$. The player's computability constraint is interpreted as the set of functions available to implement strategies. Therefore, the computational constraints are put on the set of available strategies, but we give no constraints in computing optimal strategies.

2.2 Horizontal and vertical games

In this section we formulate our model, which are obtained from finite zero-sum two-person games by extending them to infinitely repeated games with computability constraints. First we define the stage games: A zero-sum two-person game g is a triple $\langle X, Y, h \rangle$, where $X = \{x_1, ..., x_m\}$ is the set of actions for player 1, $Y = \{y_1, ..., y_n\}$ is the set of actions for player 1, and $h : X \times Y \to \mathbb{Q}$ is the von Neumann-Morgenstern utility function for player 1. We use

$$\Delta(X) = \{p \in [0,1]^m : p \in \mathbb{Q}, \sum_{x \in X} p_x = 1\}$$

and

$$\Delta(Y) = \{q \in [0,1]^n : q \in \mathbb{Q}, \sum_{y \in Y} q_y = 1\}$$

to denote the set of mixed strategies (with rational probability values) for player 1 and 2, respectively. Notice that since h is rational-valued, there is always a mixed strategy equilibrium with rational probability values in g. We use $X^{<\mathbb{N}}$ ($Y^{<\mathbb{N}}$) to denote the set of finite sequences over X (Y). For any sequence $\xi \in X^{\mathbb{N}}$, we use $\xi[T]$ to denote its initial segment with length T, i.e., $\xi[T] = (\xi_0, \xi_1, ..., \xi_{T-1})$. Given the stage game g, we formulate the vertical and the horizontal games as follows:

Definition 2.2. The *vertical game* $VG(g, \mathcal{P}^1, \mathcal{P}^2)$ associated with a stage game g and computability constraints $\mathcal{P}^1$, $\mathcal{P}^2$ is a triple $\langle \mathcal{X}, \mathcal{Y}, u_h \rangle$:

(a) $\mathcal{X} = \{a : Y^{<\mathbb{N}} \to X : a \in \mathcal{P}^1_T\}$;

(b) $\mathcal{Y} = \{b : X^{<\mathbb{N}} \to Y : b \in \mathcal{P}^2_T\}$;

(c) $u_h(a, b) = \liminf_{T\to\infty} \sum_{t=0}^{T-1} \frac{h(\theta^{a,b}_t)}{T}$, where $\theta^{a,b} = (\theta^{a,b,1}, \theta^{a,b,2}) \in (X \times Y)^{\mathbb{N}}$ is the sequence of actions taken by the players under strategy profile (a, b):

$\theta^{a,b}_0 = (a(\epsilon), b(\epsilon))$;

$\theta^{a,b}_t = (a(\theta^{a,b,2}[t]), b(\theta^{a,b,1}[t]))$ for all $t \geq 1$.

Definition 2.3. The *horizontal game* $HG(g, \mathcal{P}^1, \mathcal{P}^2)$ associated with a stage game g and computability constraints $\mathcal{P}^1$, $\mathcal{P}^2$ is a triple $\langle \mathcal{X}, \mathcal{Y}, u_h \rangle$:

(a) $\mathcal{X} = \{a : \mathbb{N} \to X : a \in \mathcal{P}^1_T\}$;

(b) $\mathcal{Y} = \{b : \mathbb{N} \to Y : b \in \mathcal{P}^2_T\}$;

(c) $u_h(a, b) = \liminf_{T\to\infty} \sum_{t=0}^{T-1} \frac{h(a_t, b_t)}{T}$.

The VG resembles the standard formulation of repeated games, but we impose computability constraints on implementable strategies. In the HG, the stage games are played simultaneously. In both games, the set $\mathcal{X}$ consists of strategies available to player 1 and $\mathcal{Y}$ consists of strategies available to player 2. For the definition of VG, in (a) and (b) we implicitly use the fact that any function from $Y^{<\mathbb{N}}$ to X can be regarded as a function from natural numbers to natural numbers (because there is an effective way to assign each history in $Y^{<\mathbb{N}}$ a number).[2] Any strategy in HG can be identified with a corresponding strategy in VG which is history-independent: let a be a strategy for player 1 in HG, then we say that a', as a strategy in VG for player 1, is a *history-independent strategy based on a* if $a'(\sigma) = a(|\sigma|)$ for all $\sigma \in Y^{<\mathbb{N}}$. As a result, any equilibrium in VG that consists of history-independent strategies is also an equilibrium in HG.

The two games HG and VG differ in their information structures. In the HG, players cannot predict their opponents' actions based on previous observations, but they can in the VG. However, HG seems more appropriate to capture the unpredictable behavior behind mixed strategy equilibrium in the standard framework. An example of HG can be found in Luce and Raiffa [18], where they discuss two aerial strategists deciding the actions of their pilots in a conflict consisting of many identical aircraft fights. They use this example to illustrate the meaning of a mixed strategy, which is interpreted as the distribution of different actions assigned to the pilots. However, in many applications, VG seems a more faithful description. Thus, we consider both the vertical and the horizontal games in our discussions.

For the payoffs, we also define both the games VG and HG to be zero-sum games, and u_h is the payoff function for player 1 but the payoff to player 2 is determined by $-u_h$. This implies that player 1 uses the $\liminf$ criterion while player 2 uses the $\limsup$ criterion. However, our main results are robust to this asymmetry. An alternative payoff criterion commonly used in repeated games is the discounting criterion, which is defined as follows:

$$v_h(a,b) = (1-\delta)\sum_{t=0}^{\infty} \delta^t h(\theta_t^{a,b}), \tag{1}$$

[2] Although this assignment not unique, it does not matter which one we use.

where $\delta \in (0, 1)$ is the discounting factor. We adopt the long-run average for several reasons. First of all, the discounting criterion does not seem appropriate in HG. Second, in our framework, there is no probability involved, and so the von Neumann-Morgenstern utility function needs a foundation that uses no probability as fundamentals. In Hu [11], such a foundation is given and the long-run average criterion is used in place of expected utility criterion. The third reason, which may be the most substantial, is that both the vertical and the horizontal games thus defined may have no equilibrium at all with the discounting criterion. This result is reported in the next section.

2.3 Nonexistence results

In this section we give a preliminary analysis of the games VG and HG. First we show that, if the discounting criterion is adopted, then for games without pure strategy equilibria, the associated vertical games have no equilibrium. The proof of the following proposition is in Section 5.

Proposition 2.1. *Let $g = \langle X, Y, h \rangle$ be a two-person zero-sum game without any pure strategy equilibrium. There is no equilibrium in $VG'(g, \mathcal{P}^1, \mathcal{P}^2) = \langle \mathcal{X}, \mathcal{Y}, v_h \rangle$, with $v_h(a, b)$ being the discounted payoff defined in (1).*

This result holds for HG with the discounting criterion as well. In this proposition, there is no assumptions on $\mathcal{P}^1$ and $\mathcal{P}^2$ except for being computability constraints. This shows that, if we are interested in investigating unpredictability of the equilibrium behavior in our framework, then we cannot take the discounting criterion.

From now on, we shall consider only the long-run average criterion, that is, the HG and VG. Our main interest is to find the conditions on the computability constraints so the resulting equilibrium behavior is unpredictable in the collective games. We first consider the natural ordering on the computability constraint in terms of set-inclusion. The following proposition gives us a necessary condition for the existence of equilibrium in the collective games with the same value as their stage games. Its proof can be found in Section 5.

Proposition 2.2. *Let $g = \langle X, Y, h \rangle$ be a two-person zero-sum game without any pure strategy equilibrium for either player. Let $\mathcal{P}^1$, $\mathcal{P}^2$ be two computability constraints.*

(a) If $\mathcal{P}^1 = \mathcal{P}^2$, then there is no equilibrium in $VG(g, \mathcal{P}^1, \mathcal{P}^2)$.

(b) Suppose that $\mathcal{P}^2 \subset \mathcal{P}^1$. The value of the game $VG(g, \mathcal{P}^1, \mathcal{P}^2)$, if it exists, is

$$\min_{y \in Y} \max_{x \in X} h(x, y).$$

We know from part (a) of this proposition that there is no equilibrium in the vertical games if the two players share the same computability constraint. Part (b) shows that, if $\mathcal{P}^2 \subset \mathcal{P}^1$ and if there is no pure strategy equilibrium in the stage game, the value of the vertical game is different from that of the stage game. The case $\mathcal{P}^1 \subset \mathcal{P}^2$ is completely symmetric. Both results hold for HG as well with necessary modification. This result resembles the findings in Ben-Porath [3][3], which obtains the same value ($\min_{y \in Y} \max_{x \in X} h(x, y)$) when player 1 has a substantially stronger computational power. However, mixed strategies are allowed there and the existence problem is trivial. Nonetheless, that result and our proof suggest that this proposition does not rely on the assumption that both players' computability constraints include all Turing-computable functions. We are not able to obtain a general nonexistence result with one player's constraint being more restrictive than the other player's, but in some examples, we are able to show that there is no equilibrium in HG in this case. The following is one example.

Example 2.1. Consider the matching pennies game $g = \langle \{H, T\}, \{H, T\}, h \rangle$ with

$$h(H, H) = 1 = h(T, T) \text{ and } h(H, T) = 0 = h(T, H).$$

Suppose that $\mathcal{P}^2 \subset \mathcal{P}^1$. Then the value of the game $HG(g, \mathcal{P}^1, \mathcal{P}^2)$ is 1, if it exists. Let a^* be an equilibrium strategy of player 1 in that horizontal game. Then we have

$$\liminf_{T \to \infty} \frac{h(a^*, H)}{T} = 1 = \liminf_{T \to \infty} \frac{h(a^*, T)}{T},$$

and hence,

$$\liminf_{T \to \infty} \frac{|\{0 \le t \le T-1 : a^*_t = H\}|}{T} = 1 = \liminf_{T \to \infty} \frac{|\{0 \le t \le T-1 : a^*_t = T\}|}{T}.$$

[3]In that paper, players use finite automata to implement strategies.

But this implies that

$$1 = \liminf_{T\to\infty} \frac{|\{0 \le t \le T-1 : a_t^* = H\}| + |\{0 \le t \le T-1 : a_t^* = T\}|}{T}$$
$$\ge \liminf_{T\to\infty} \frac{|\{0 \le t \le T-1 : a_t^* = H\}|}{T} + \liminf_{T\to\infty} \frac{|\{0 \le t \le T-1 : a_t^* = T\}|}{T} = 2,$$

a contradiction. Thus, there is no equilibrium in $HG(g, \mathcal{P}^1, \mathcal{P}^2)$.

Now, Proposition 2.2 and the above example shows that, to obtain existence, it is necessary that both $\mathcal{P}^1 - \mathcal{P}^2$ and $\mathcal{P}^2 - \mathcal{P}^1$ are nonempty.[4] Moreover, this shows that if equilibrium exists in a collective game, one player's equilibrium strategy is necessarily uncomputable to the other player. The exact necessary and sufficient condition for existence seems very hard; however, in the next section, we will find a sufficient condition for existence.

3 Complexity and unpredictable behavior

In this section we give a sufficient condition, called *mutual complexity*, on the computability constraints that guarantees equilibria existence in the collective games. To formulate this condition, we will consider Kolmogorov complexity, which is introduced by Kolmogorov [13] to study the foundation of probability theory. Then, we show that, under this condition, there is an equilibrium that passes all the idealized statistical tests with respect to any mixed equilibrium of the stage game. This result in turn implies the mutual complexity condition. However, we are able to show that, under mutual complexity, there exists an equilibrium strategy that fails some tests.

3.1 Kolmogorov complexity and existence

In this section we define the notion of Kolmogorov complexity. Although our intention is to measure the complexity of functions in $\mathcal{P}$, which can be identified with infinite

[4]In fact, in Example 2.1, our simple argument for nonexistence relies on the fact that $\mathcal{P}^2 \subset \mathcal{P}^1$. However, if we assume that both players maximizes limsup of the average payoffs, then the nonexistence result remains.

sequences over $\mathbb{N}$, we begin with complexity of finite sequences over $\{0,1\}$, denoted by $\{0,1\}^{<\mathbb{N}}$. Then, we measure the complexity of an infinite sequence by considering the complexity of its initial segments, following the Algorithmic Randomness literature.[5]

The idea behind this notion is to measure the complexity of a finite object with the length of its shortest description. In our case, we consider the complexity of strings in $\{0,1\}^{<\mathbb{N}}$ and use strings in $\{0,1\}^{<\mathbb{N}}$ as descriptions. Any partial function $d : \{0,1\}^{<\mathbb{N}} \to \{0,1\}^{<\mathbb{N}}$ may be regarded as a description method, and we call elements in

$$dom(d) = \{\sigma \in \{0,1\}^{<\mathbb{N}} : d(\sigma) \neq \perp\}$$

code-words. However, to ensure the descriptions are complete, we consider only partial functions $d : \{0,1\}^{<\mathbb{N}} \to \{0,1\}^{<\mathbb{N}}$ that are *prefix-free*, [6] i.e., for any two code-words $\sigma,\ \tau \in dom(f)$, $\sigma \subset \tau$ (meaning that σ is an initial segment of τ) implies that $\sigma = \tau$. For a given computability constraint $\mathcal{P}$, we define

$$D(\mathcal{P}) = \{f \in \mathcal{P} \mid f : \{0,1\}^{<\mathbb{N}} \to \{0,1\}^{<\mathbb{N}} \text{ and } f \text{ is prefix-free}\}.$$

For any set $\mathcal{P}$, this complexity measure is asymptotically absolute in the sense that there is a function that gives the shortest descriptions among all functions in $D(\mathcal{P})$ within a constant.

To measure the complexity of infinite sequences in $\{0,1\}^{\mathbb{N}}$, we shall consider the complexity of their initial segments. In this way, we do not have a numerical measure for each sequence, but we are still able to discriminate different sequences in terms of complexity qualitatively. We first define *complex sequences.*

Definition 3.1. Let $\mathcal{P} \in \mathfrak{R}$. A sequence $\xi \in \{0,1\}^{\mathbb{N}}$ is a *complex sequence relative to* $\mathcal{P}$ if for all $f \in D(\mathcal{P})$, there is a constant b such that for all $T > 0$,

$$K_f(\xi[T]) \geq T - b,$$

where $\xi[T] = (\xi_0, ..., \xi_{T-1})$ is the initial segment of ξ with length T.

[5] See, for example, Downey *et al.* [9].

[6] For the motivation to use prefix-free functions, please see van Lambalgen [15] and references there.

Intuitively, a sequence is complex relative to $\mathcal{P}$ if it is hard to describe for all description methods in $D(\mathcal{P})$. In our definition, a sequence is complex if its initial segments can only be described by strings with almost the same lengths. As a consequence, if a sequence is complex relative to $\mathcal{P}$, then it is not in $\mathcal{P}$. A complex sequence, however, may not be maximally complex, because we know that there are sequences whose initial segments have complexity significantly higher than their lengths.[7]

Now we are ready to formulate our sufficient condition for general existence in the vertical and the horizontal games. We say that two computability constraints $\mathcal{P}^1,\ \mathcal{P}^2$ are *mutually complex* if there are there are $\xi^1 \in \mathcal{P}^1,\ \xi^2 \in \mathcal{P}^2$ such that for both $i = 1, 2$, ξ^i is a complex sequence relative to $\mathcal{P}^{-i}$. Our theorem is stated in the following, and its proof is in Section 5.

Theorem 3.1. *Let g be a finite zero-sum game with value v. Suppose that $\mathcal{P}^1,\ \mathcal{P}^2$ are mutually complex. Then there is an equilibrium for both $VG(g, \mathcal{P}^1, \mathcal{P}^2)$ and $HG(g, \mathcal{P}^1, \mathcal{P}^2)$. Moreover, the value of both HG and VG is also v.*

In the next section, we shall discuss what the statistical tests are that the equilibrium strategies should pass. Before we turn to those results, we conclude with a proposition that gives an estimation of the pervasiveness of mutually complex computability constraints. Its proof is in the appendix.

Proposition 3.1. *There are uncountably many different pairs of computability constraints that satisfy mutual complexity.*

3.2 Unpredictable behavior: randomness

In this section we consider the unpredictable behavior in the games VG and HG. In particular, we address the question of whether the equilibrium strategies should pass all statistical tests. The question corresponds to a criterion of unpredictability in the

[7] To be precise, there are infinitely many sequences ξ such that for any description method f, there is some constant b such that for all T, $K_f(\xi[T]) \geq T + K_f(T) + b$. This result is reported in Li and Vitányi [17].

literature, which is called *Martin-Löf randomness* [19]. This notion begins with a formal formulation of idealized statistical tests: a test with respect to a measure relative to a computability constraint is defined to be a property that has zero probability, i.e., a measure-zero event (w.r.t. that measure) that can be detected with functions in the constraint. Before we lay out the formal definition, we need some notations.

Let X be a finite set. We endow the product topology on the set of infinite sequences over X, which is denoted by $X^{\mathbb{N}}$. For our purpose, we consider only Borel probability measures on it. Any open set can be written as a union of basic sets, where a basic set has the form $N_\sigma = \{\zeta \in X^{\mathbb{N}} : \sigma = \zeta[|\sigma|]\}$ for some $\sigma \in X^{<\mathbb{N}}$. It is well known that any Borel measure μ is uniquely determined by its values on N_σ's, i.e., if for all $\sigma \in X^{<\mathbb{N}}$, $\mu(N_\sigma) = \nu(N_\sigma)$, then $\mu = \nu$. We give a formal definition of randomness in the following.

Definition 3.2. Let X be a finite set and let $\mathcal{P}$ be a computability constraint. Suppose that μ is a computable probability measure over $X^{\mathbb{N}}$, i.e., the mapping $\sigma \mapsto \mu(N_\sigma)$ belongs to $\mathcal{P}^*$. A sequence of open sets $\{V_t\}_{t=0}^{\infty}$ is *a μ-test relative to $\mathcal{P}$* if it satisfies the following conditions:

(1) There is a function $f : \mathbb{N} \to \mathbb{N} \times X^{<\mathbb{N}}$ in $\mathcal{P}_T$ such that for all $t \in \mathbb{N}$ and for all $\xi \in X^{\mathbb{N}}$,

$$\xi \in V_t \Leftrightarrow (\exists n)(f(n) = (t, \sigma) \wedge \sigma = \xi[|\sigma|]).$$

(2) For all $t \in \mathbb{N}$, $\mu(V_t) \leq 2^{-t}$.

A sequence $\xi \in X^{\mathbb{N}}$ is *μ-random relative to $\mathcal{P}$* if it passes all μ-tests relative to $\mathcal{P}$, i.e., for any μ-test $\{V_t\}_{t=0}^{\infty}$ relative to $\mathcal{P}$, $\xi \notin \bigcap_{t=0}^{\infty} V_t$.

Implicitly in the definition we assume that $\mu(N_\sigma)$ is always a rational number for μ to be computable. In the literature, computability of a measure is defined more generally, but this definition is sufficient for our purpose.

Each test $\{V_t\}_{t=0}^{\infty}$ is used to test a zero probability property that corresponds to the event $\bigcap_{t=0}^{\infty} V_t$. Clearly, such an event has probability zero w.r.t. the measure. Conditions (1) and (2) require that this property can be detected by functions in $\mathcal{P}$: the measure of the corresponding event is proved to be zero by a list of open sets that can be generated by functions in the computability constraint. A sequence is random if it passes all such tests. As a consequence, the set of random sequence, given a fixed measure, depends on the computability constraint in a monotonic manner. If the constraint $\mathcal{P}^1$ is a subset of another constraint $\mathcal{P}^2$, then any test relative to $\mathcal{P}^1$ is also a test relative to $\mathcal{P}^2$. Therefore, the set of random sequences relative to $\mathcal{P}^2$ is a subset of random sequences relative to $\mathcal{P}^1$. We conclude the comments with an existence theorem for random sequences. Its proof for the case with $\mathcal{P} = \mathcal{P}^*$ can be found in Martin-Löf [19] (with some minor modifications to accommodate general computable measures). See also Downey *et al.* [9].

Proposition 3.2. *Suppose that X is a finite set and μ is a computable measure over $X^{\mathbb{N}}$. Let $\mathcal{P}$ be a computability constraint. Then*

$$\mu(\{\xi \in X^{\mathbb{N}} : \xi \textit{ is } \mu\textit{-random relative to } \mathcal{P}\}) = 1.$$

Now we are ready to show that, if the collective game satisfies mutual complexity, then there is an equilibrium strategy that is random with respect to the i.i.d. measure generated by a mixed strategy equilibrium of the stage game. For any $p \in \Delta(X)$, we use μ_p to denote the measure over $X^{\mathbb{N}}$ such that for all $\sigma \in X^{<\mathbb{N}}$,

$$\mu_p(N_\sigma) = \prod_{t=0}^{|\sigma|-1} p_{\sigma_t}.$$

The measure μ_q for any $q \in \Delta(Y)$ is defined in a similar manner. Moreover, this holds for any equilibrium strategy profile of the stage game as well, for which we need to consider distribution over $X \times Y$. For any probability distribution $p \in \Delta(X)$ and $q \in \Delta(Y)$, we use $p \otimes q$ to denote the product measure of p and q over $X \times Y$ and use $\mu_{p \otimes q}$ to denote the i.i.d. Bernoulli measure generated by $p \otimes q$ over $(X \times Y)^{\mathbb{N}}$. The proof of the following theorem is in section 5.

Theorem 3.2. *Suppose that $\mathcal{P}^1, \mathcal{P}^2$ are mutually complex. Let g be a finite zero-sum game. Then, in $HG(g, \mathcal{P}^1, \mathcal{P}^2)$,*

(a) if $(p, q) \in \Delta(X) \times \Delta(Y)$ is an equilibrium of g, then there is an equilibrium (a, b) such that a (b) is a μ_p-random (μ_q-random) sequence relative to $\mathcal{P}^2$ ($\mathcal{P}^1$); moreover, any

μ_p-random (μ_q-random) sequence relative to $\mathcal{P}^2$ ($\mathcal{P}^1$) is an equilibrium strategy for player 1 (2);

(b) for such an equilibrium (a, b), the sequence $(a_t, b_t)_{t=0}^{\infty}$ is a $\mu_{p\otimes q}$-random sequence (relative to $\mathcal{P}^*$).

This theorem holds for VG as well, if we replace the strategies in HG mentioned above with the corresponding history-independent strategies in VG. Part (a) of this theorem shows that, under mutual complexity, each player has some equilibrium strategies that pass all the statistical tests relative to his opponent. This result shows that equilibrium strategies can satisfy a very strong independence condition. For each player, his opponent has an equilibrium strategy that appears random w.r.t. an i.i.d. distribution. Part (b) of this theorem is of interest because such independence may not be observable to outsiders. However, this part shows that, relative to the Turing-computability, such equilibrium strategies from both players appear to be generated by two independent random processes.

Our next concern is the necessity of mutual complexity to derive this result. We show that part (a) of Theorem 3.2 implies mutual complexity. The proof of the following theorem is in Section 5.

Theorem 3.3. *Let g be a finite zero-sum game without pure strategy equilibrium. Let $\mathcal{P}^1$, $\mathcal{P}^2$ be two computability constraints and let $(p, q) \in \Delta(X) \times \Delta(Y)$ be non-degenerate distributions.*[8] *Suppose that there is a μ_p-random sequence relative to $\mathcal{P}^2$ in $\mathcal{X}$ and there is a μ_q-random relative to $\mathcal{P}^1$ in $\mathcal{Y}$. Then $\mathcal{P}^1$, $\mathcal{P}^2$ are mutually complex.*

This result shows that if each player has a strategy that is random relative to his opponent, then their computability constraints satisfy mutual complexity. This suggests that the usual assumption that players are able to randomize independently requires high complexity in their computational powers. Moreover, this shows that mutual complexity is indispensable if we require the equilibrium strategies to pass all statistical tests. However, in the next section, we show that, under mutual complexity, some equilibrium strategies will fail some tests.

[8] A distribution p is non-degenerate if $p_{x_1} > 0$ and $p_{x_2} > 0$ for two different elements x_1, x_2.

3.3 Unpredictable behavior: stochasticity

In this section we show that there are equilibrium strategies that fail some tests in both the vertical and the horizontal games that satisfy mutual complexity. To understand what laws are violated, we introduce a weaker criterion of unpredictability proposed by von Mises [20], called *stochasticity.* Before we show the main results, we first define this notion and show that this criterion gives sufficient unpredictability to be optimal in equilibrium.

Intuitively, a sequence ξ over a finite set X is p-stochastic w.r.t. some $p \in \Delta(X)$ if there is no subsequence of ξ that has relative frequency different from p. Clearly, if there is no restriction on how one could select the subsequences, there is no p-stochastic sequence unless p is degenerate. As randomness, we shall define stochasticity relative to computability constraint $\mathcal{P}$, and consider only subsequences that can be selected by a function from $\mathcal{P}$.

Formally, a function $r : X^{<\mathbb{N}} \to \{0,1\}$ is called a *selection function.* Given a sequence $\xi \in X^{\mathbb{N}}$, we define a partial function $\pi^r : \mathbb{N} \to \mathbb{N}$ as follows:

$$
\begin{aligned}
&\text{for } t = 0,\ \pi^r(0) = \min\{T : r(\xi[T]) = 1\}; \\
&\text{for } t > 0,\ \pi^r(t) = \min\{T : T > \pi^r(t-1),\ r(\xi[T]) = 1\}.
\end{aligned} \tag{2}
$$

Then, we define the subsequence ξ^r chosen by r as follows:

$$\text{for all } t \in \mathbb{N},\ \xi^r_t = \xi_{\pi^r(t)}. \tag{3}$$

It is easy to see that $\xi^r \in X^{\mathbb{N}}$ if and only if π^r is a total function. Now, we are ready to define stochasticity.

Definition 3.3. Let $p \in \Delta(X)$ be a probability distribution and let $\mathcal{P} \in \mathfrak{R}$ be a computability constraint. We say that a sequence $\xi \in X^{\mathbb{N}}$ is *p-stochastic relative to $\mathcal{P}$* if for any selection function $r \in \mathcal{P}_T$ such that $\xi^r \in X^{\mathbb{N}}$, we have for all $x \in X$,

$$\lim_{T\to\infty} \sum_{t=0}^{T-1} \frac{c_x(\xi^r_t)}{T} = p_x,$$

where $c_x(y) = 1$ if $x = y$, and $c_x(y) = 0$ otherwise.

We now show that, if p is a mixed equilibrium strategy in the stage game, then any p-stochastic sequence is an equilibrium strategy in the collective game satisfying mutual complexity.

Theorem 3.4. *Let g be a finite zero-sum game. Suppose that $\mathcal{P}^1$, $\mathcal{P}^2$ are mutually complex. Then, for any equilibrium strategy $p^* \in \Delta(X)$ in g, if $a \in \mathcal{X}$ is p^*-stochastic relative to $\mathcal{P}^2$, then it is an equilibrium strategy of player 1 in $HG(g, \mathcal{P}^1, \mathcal{P}^2)$.*

Moreover, for such p^*, there is an equilibrium strategy a that is a p^*-stochastic sequence, but not a μ_{p^*}-random sequence.

This theorem holds as well for VG if we replace the strategies in HG mentioned above with the corresponding history independent strategies in VG. In fact, for HG, we can replace stochasticity with a weaker requirement, in which we consider only selection functions that are history independent, i.e., functions with the form $r : \mathbb{N} \to \{0, 1\}$. This theorem shows that, in equilibrium, the behavior does not have to appear as random as an i.i.d. process. Moreover, Ambos-Spies *et al.* [1] shows that each selection function (together with a deviation from the specified frequency) corresponds to a test.[9] This theorem shows that only those tests are relevant to the equilibrium hypothesis in our framework.

Although stochasticity is formally defined, we are not able to explicitly construct the tests corresponding to this criterion. However, in matching pennies, some more information is available, as the following example shows.

Example 3.1. Consider the matching pennies game g with $X = Y = \{H, T\}$. Suppose that $\mathcal{P}^1, \mathcal{P}^2$ satisfy mutual complexity. By Theorem 3.4, any $(\frac{1}{2}, \frac{1}{2})$-stochastic sequence

[9] The result in Ambos-Spies [1] shows that a sequence is stochastic if and only if there is no simple computable martingales that succeeds over it (which means that the martingale can accumulates infinite wealth by betting against the sequence). A martingale is simply a betting strategy, and it is simple if the bet is a constant proportion to the wealth. It is well-known that each martingale corresponds to a test and thus to a probability law, and so a sequence is random if and only if no approximately computable martingale succeeds over it (see Downey *et al.* [9] for a precise definition). Thus, simple martingales correspond to a proper subset of tests.

relative to $\mathcal{P}^2$ is an equilibrium strategy of player 1 in $HG(g, \mathcal{P}^1, \mathcal{P}^2)$. However, by Ville [32], there is such a stochastic sequence ξ such that for all $T \in \mathbb{N}$,

$$\frac{|\{t : 0 \leq t \leq T-1,\ \xi_t = H\}|}{T} \geq \frac{1}{2}.$$

This shows that ξ violates the Law of the Iterated Logarithm.

Such an equilibrium strategy (that violates the Law of the Iterated Logarithm in a similar manner) exists for any horizontal games with 2×2 stage games. Such a sequence is certainly not deemed to be random in any sense. The lack of randomness has been used to reject the equilibrium hypothesis in the empirical literature (see, for example, Brown and Rosenthal [4] and Walker and Wooders [36]), but we show that, in our framework, the equilibrium behavior does not have to be so unpredictable.

4 Discussions

In this section, we discuss the implications of our results to the literature, and then consider alternative formulations. In particular, we discuss the possibility of extending our framework to finitely repeated games and the difficulties of this extension.

4.1 Mixed strategies and complexity

Our results give a correspondence between the complexity of players' computability constraints and the unpredictability of equilibrium behavior. We find the necessary and sufficient condition for equilibrium behavior to pass all idealized statistical tests (relative to each other), which requires extremely high complexity in players' computational powers. This finding implies that the usual assumption that players are able to use mixed strategies actually implies that they have enough complexity in their computational abilities relative to one another and to the outside observer, which is a very strong assumption.

4.2 Extensions to finitely repeated games

There are two strong assumptions in our framework—the game is infinitely repeated and both players have Turing-computability. Although we find the relevant tests for equilibrium strategies in our framework, the applicability of these results to the real world is limited by these assumptions. In particular, even in the usual formulation, this infinitely repeated game with long-run average criterion does not have unique mixed strategy equilibrium even if the stage game has a unique equilibrium. Nonetheless, we show that some tests are not relevant because they do not correspond to profitable deviations, and hence, even though they are failed by a sequence of plays, this sequence may still be optimal. Notice that in the context of a repeated zero-sum game without pure strategy equilibrium, given a sequence of play of one player, the other player can only gain by selecting a subsequence which has a frequency different form the equilibrium mixed strategy of the stage game. This requirement is fully captured by the notion of stochasticity.

Conceptually, our framework does not depend on the infinite structures. In principle, we can put computability constraints on implementation of strategies in a finitely repeated game, and consider conditions under which unpredictable behavior is an equilibrium phenomenon. To this end, it is necessary to have both players unable to compute some Turing-computable functions. However, such a project involves many theoretical and technical difficulties. First, although the Kolmogorov complexity is well-defined for finite objects, it is independent of machines only asymptotically. Thus, the measure of complexity is machine-dependent and thus the choice of an appropriate model of computation becomes crucial to the results. In contrast, relative Turing-computability, which we use in the paper, is invariant with different formulations of machines. Second, even though in principle we can define randomness for finite sequences, there are not as many good properties as infinite random sequences. Moreover, any such definition is also machine-dependent.

5 Proofs of the main theorems

In this section we give the proofs of the main theorems. In the appendix, we review some extant results regarding Kolmogorov complexity and Martin-Löf randomness that are necessary in the proofs. We begin with the proofs of the nonexistence results, and then we present the proofs of the other theorems in the second subsection.

5.1 Proofs of the nonexistence results

Proof of Proposition 2.1: Suppose that (a^*, b^*) is an ε-equilibrium. Let

$$c = \max_{(x,y)\in X\times Y} |h(x, y)|.$$

For each $T \in \mathbb{N}$, consider the following strategy b' such that for all $\sigma \in X^{<\mathbb{N}}$, $b'(\sigma) = y_{\zeta_{|\sigma|}}$, where ζ is defined as follows:

$$\begin{aligned} &\text{for } t = 0,\ \zeta_0 = \min\{i : y_i \in \arg\min_{y\in Y} h(a^*(\epsilon), y)\}; \\ &\text{for } t = 1, ..., T,\ \zeta_t = \min\{i : y_i \in \arg\min_{y\in Y} h(a^*(y_\zeta[t]), y)\}; \\ &\text{for } t > T + 1,\ \zeta_t = 1. \end{aligned} \tag{4}$$

Since ζ_t is constant for all $t > T$, $\zeta \in \mathcal{P}^2$ and so $b' \in \mathcal{Y}$. We have

$$v(a^*, b') \leq (1 - \delta) \sum_{t=0}^{T} \delta^t v_1 + (1 - \delta) \sum_{t=T+1}^{\infty} c = v_1 + \delta^{T+1}(c - v_1).$$

Since (a^*, b^*) is an ε-equilibrium, it follows that

$$v(a^*, b^*) \leq v(a^*, b') + \varepsilon \leq v_1 + \varepsilon + \delta^{T+1}(c - v_1).$$

Since this holds for all T, we have

$$v(a^*, b^*) \leq \lim_{T\to\infty} v_1 + \varepsilon + \delta^{T+1}(c - v_1) = v_1 + \varepsilon.$$

Similarly, we can show that

$$v(a^*, b^*) \geq v_2 - \varepsilon.$$

It then follows that, for any $\varepsilon < \frac{v_2 - v_1}{2}$,

$$v_1 + \varepsilon < v_2 - \varepsilon \leq v(a^*, b^*) \leq v_1 + \varepsilon,$$

a contradiction. $\square$

Proof of Proposition 2.2: (a) We consider VG here, and the proof for HG is almost identical. We first index the actions in Y as $Y = \{y_1, ..., y_n\}$. Since there is no pure strategy equilibrium in g, it follows that

$$v_1 = \max_{x \in X} \min_{y \in Y} h(x, y) < \min_{y \in Y} \max_{x \in X} h(x, y) = v_2.$$

Suppose that, to the contrary, (a^*, b^*) is an equilibrium strategy profile in $VG(g, \mathcal{P}, \mathcal{P})$. Consider the strategy b' such that for all $\sigma \in X^{<\mathbb{N}}$, $b'(\sigma) = y_{\zeta_{|\sigma|}}$, where ζ is defined as follows:

$$\begin{aligned} t = 0,\ \zeta_0 &= \min\{i : y_i \in \arg\min_{y \in Y} h(a^*(\epsilon), y)\}; \\ t > 0,\ \zeta_t &= \min\{i : y_i \in \arg\min_{y \in Y} h(a^*(y_\zeta[t]), y)\}, \end{aligned} \tag{5}$$

where $y_\zeta[t] = (y_{\zeta_0}, ..., y_{\zeta_{t-1}})$. $\zeta \in \mathcal{P}$ because $a^* \in \mathcal{P}$ and $\mathcal{P}$ is closed under composition and primitive recursion. Thus, $b' \in \mathcal{Y}$. By construction, we have that

$$h(a^*(y_\zeta[t]), \zeta_t) \leq \max_{x \in X} \min_{y \in Y} h(x, y) = v_1.$$

Then,

$$u(a^*, b') \leq \liminf_{T \to \infty} v_1 = v_1.$$

Since b^* is a best response to a^*, it follows that

$$u(a^*, b^*) \leq u(a^*, b') \leq v_1.$$

Similarly, we can show that

$$u(a^*, b^*) \geq v_2.$$

But then

$$v_2 > v_1 \geq u(a^*, b^*) \geq v_2,$$

a contradiction.

(b) We first index the actions in X as $X = \{x_1, ..., x_m\}$. For any $b \in \mathcal{Y}$, construct the strategy a' such that for all $\sigma \in Y^{<\mathbb{N}}$, $a'(\sigma) = x_{\zeta_{|\sigma|}}$, where ζ is defined as follows:

$$\begin{aligned} &t = 0,\ \zeta_0 = \min\{i : x_i \in \arg\max_{x \in X} h(x, b(\epsilon))\}; \\ &t > 0,\ \zeta_t = \min\{i : x_i \in \arg\max_{x \in X} h(x, b(x_\zeta[t]))\}, \end{aligned} \tag{6}$$

with $x_\zeta[t] = (x_{\zeta_0}, ..., x_{\zeta_{t-1}})$. $\zeta \in \mathcal{P}^2$ because $b \in \mathcal{P}^2$ and $\mathcal{P}^2$ is closed under composition and primitive recursion. Thus, $a' \in \mathcal{X}$ since $\mathcal{P}^2 \subset \mathcal{P}^1$. By construction, we have that for all $t \in \mathbb{N}$,

$$h(\zeta_t, b(x_\zeta[t])) \geq \min_{y \in Y} \max_{x \in X} h(x, y) = v_2.$$

Then, $u_h(a', b) \geq \liminf_{T \to \infty} v_2 = v_2$. It follows that $\sup_{a \in \mathcal{X}} u_h(a, b) \geq v_2$. Now, let $y^* \in \arg\min_{y \in Y}(\max_{x \in X} h(x, y))$. Let $b \in \mathcal{P}^2$ be such that $b(\tau) = y^*$ for all $\tau \in X^{<\mathbb{N}}$. Then $\sup_{a \in \mathcal{X}} u_h(a, b) = v_2$. Thus, we have that $\min_{b \in \mathcal{Y}} \inf_{a \in \mathcal{X}} u_h(a, b) = v_2$, and hence, the value of $VG(g, \mathcal{P}^1, \mathcal{P}^2)$ is v_2. $\square$

5.2 Proofs of the existence and the unpredictable behavior results

In this section, we will begin with a theorem (Theorem 5.1) that shows that, if each player has a stochastic sequence relative to his opponent's computability constraint which has an equilibrium frequency in the stage game, then these stochastic sequences constitute an equilibrium in both HG and VG.

With this result, to show that Theorem 3.2 holds, it suffices to show that there are μ_p-random sequences relative to $\mathcal{P}^2$ for any equilibrium strategy $p \in \Delta(X)$ of g for player 1. It is completely symmetric for player 2. Moreover, Theorem 3.1 is a corollary of that theorem.

To prove Theorem 3.4, we consider a sequence of probability distributions $\mathbf{p} = \{p^t\}_{t=0}^{\infty}$ that converges to an equilibrium strategy $p \in \Delta(X)$ in g. Then, by Theorem 6.6, which can be found in the appendix, any sequence that is $\mu_{\mathbf{p}}$-random relative to $\mathcal{P}^2$ (for all $\sigma \in X^{<\mathbb{N}}$, $\mu_{\mathbf{p}}(N_\sigma) = \prod_{t=0}^{|\sigma|-1} p^t_{\sigma_t}$) is also p-stochastic relative to $\mathcal{P}^2$. By Theorem 5.1, this sequence is an equilibrium strategy. It then remains to show that this sequence is not $\mu_{p'}$-random for any $p' \in \Delta(X)$. We use the result form Vovk [33] to prove this for sequences $\mathbf{p}$ that are distant from p, which is an effective version of the result obtained in Kakutani [12].

To prove Theorem 3.3, first we show that any binary random sequence with respect to the uniform distribution is a complex sequence, and any random sequence that is generated by an non-degenerate i.i.d. measure can be used to compute a binary random sequence with respect to the uniform distribution. Of course, all these have to be relativized with respect to a computability constraint. This theorem is closely related to the principle that randomness is equivalent to extreme complexity, and interested readers may go to the survey paper Downey *et al.* [9].

First we give a lemma concerning expected values. For stochastic sequences, the expected values correspond to long-run averages.

Lemma 5.1. *Let X be a finite set. Let $\mathcal{P} \in \mathfrak{R}$ and let $p \in \Delta(X)$ be a distribution. Suppose that $h : X \to \mathbb{Q}$ is a function over X. If ξ is a p-stochastic sequence relative to $\mathcal{P}$, then, for any selection function r in $\mathcal{P}$ such that $\xi^r \in X^{\mathbb{N}}$, we have*

$$\lim_{T\to\infty} \sum_{t=0}^{T-1} \frac{h(\xi_t^r)}{T} = \sum_{x\in X} p_x h(x).$$

Proof. Let r be a selection function in $\mathcal{P}$ such that $\xi^r \in X^{\mathbb{N}}$. Then, for any $x \in X$,

$$\lim_{T\to\infty} \sum_{t=0}^{T-1} \frac{c_x(\xi_t^r)}{T} = p_x.$$

Therefore,

$$\lim_{T\to\infty} \sum_{t=0}^{T-1} \frac{h(\xi_t^r)}{T} = \lim_{T\to\infty} \sum_{x\in X} \sum_{t=0}^{T-1} \frac{c_x(\xi_t^r)h(x)}{T} = \sum_{x\in X} p_x h(x).$$

□

Now, we shall give a theorem that guarantees the existence of equilibrium in the vertical and the horizontal games that has the same value as the stage game.

Theorem 5.1. *Let g be a finite zero-sum game and let $(p^*, q^*) \in \Delta(X) \times \Delta(Y)$ be an equilibrium of g. Suppose that there are $\xi \in \mathcal{P}^1$, $\zeta \in \mathcal{P}^2$ such that ξ is p^*-stochastic relative to $\mathcal{P}^2$ and ζ is q^*-stochastic relative to $\mathcal{P}^1$. Then, in $VG(g, \mathcal{P}^1, \mathcal{P}^2)$, any strategy $a \in \mathcal{X}$ that is a history independent strategy based on a p-stochastic relative $\mathcal{P}^2$ for some equilibrium mixed strategy $p \in \Delta(X)$ is an equilibrium strategy.*

Proof. Consider the game $VG(g, \mathcal{P}^1, \mathcal{P}^2)$. First we show that

$$(\forall a \in \mathcal{X})(\limsup_{T\to\infty} \sum_{t=0}^{T-1} \frac{h(\theta_t^{a,b^*})}{T} \leq h(p^*, q^*)), \tag{7}$$

where b^* is such that $b^*(\sigma) = \zeta_{|\sigma|}$ for all $\sigma \in X^{<\mathbb{N}}$. Since (p^*, q^*) is the value of the game g, it follows that $h(x, q^*) \leq h(p^*, q^*)$ for all $x \in X$.

Suppose that $a \in \mathcal{X}$, and so $a \in \mathcal{P}^1$. For each $x \in X$, let $r^x : Y^{<\mathbb{N}} \to \{0, 1\}$ be the selection function such that $r^x(\sigma) = 1$ if $a(\sigma) = x$, and $r^x(\sigma) = 0$ otherwise. Define (see equation (3))

$$L_x(T) = |\{t \in \mathbb{N} : 0 \leq t \leq T-1,\ r^x(\zeta[t]) = 1\}| \text{ and } \zeta^x = \zeta^{r^x}.$$

It is easy to see that r^x is in $\mathcal{P}^1$ since a is. Let

$$\mathcal{E}^1 = \{x \in X : \lim_{T\to\infty} L_x(T) = \infty\} \text{ and } \mathcal{E}^2 = \{x \in X : \lim_{T\to\infty} L_x(T) < \infty\}.$$

For each $x \in \mathcal{E}^2$, let $B_x = \lim_{T\to\infty} L_x(T)$ and let $C_x = \sum_{t=0}^{B_x} h(x, \zeta_t^x)$. Then, for any $x \in \mathcal{E}^1$, by Lemma 5.1,

$$\lim_{T\to\infty} \sum_{t=0}^{T-1} \frac{h(x, \zeta_t^x)}{T} = h(x, q^*) \leq h(p^*, q^*).$$

We claim that for any $\varepsilon > 0$, there is some T' such that $T > T'$ implies that

$$\sum_{t=0}^{T-1} \frac{h(a(\zeta[t]), \zeta_t)}{T} \leq h(p^*, q^*) + \varepsilon. \tag{8}$$

Fix some $\varepsilon > 0$. Let T_1 be so large that $T > T_1$ implies that, for all $x \in \mathcal{E}^1$,

$$\sum_{t=0}^{T-1} \frac{h(x, \zeta_t^x)}{T} \leq h(p^*, q^*) + \frac{\varepsilon}{|X|}, \tag{9}$$

and, for all $x \in \mathcal{E}^2$,

$$\frac{C_x}{T} < \frac{\varepsilon}{|X|}. \tag{10}$$

Let T' be so large that, for all $x \in \mathcal{E}_1$, $L_x(T') > T_1$. If $T > T'$, then

$$\begin{aligned}
\sum_{t=0}^{T-1} \frac{h(a(\zeta[t]), \zeta_t)}{T} &= \sum_{x \in \mathcal{E}_1} \frac{L_x(T)}{T} \sum_{t=0}^{L_x(T)-1} \frac{h(x, \zeta_t^x)}{L_x(T)} + \sum_{x \in \mathcal{E}_2} \sum_{t=0}^{L_x(T)-1} \frac{h(x, \zeta_t^x)}{T} \\
&\leq \sum_{i \in \mathcal{E}_1} \frac{L_x(T)}{T} (h(p^*, q^*) + \frac{\varepsilon}{|X|}) + \sum_{x \in \mathcal{E}_2} \frac{\varepsilon}{|X|} \leq h(p^*, q^*) + \varepsilon. \qquad (11)
\end{aligned}$$

Notice that L_x is weakly increasing, and $L_x(T) \leq T$ for all T. Thus, $T > T'$ implies that $L_x(T) \geq L_x(T') > T_1$, and so $T > T_1$.

This proves the inequality (8), and it in turn implies that, for any $\varepsilon > 0$, there is some T such that

$$\alpha_T = \sup_{T' > T} \sum_{t=0}^{T'-1} \frac{h(a(\eta[t]), \eta_t)}{T'} \leq h(p^*, q^*) + \varepsilon.$$

Now, the sequence $\{\alpha_T\}_{T=0}^{\infty}$ is a decreasing sequence, and the above inequality shows that for any $\varepsilon > 0$, $\lim_{T\to\infty} \alpha_T \leq h(p^*, q^*) + \varepsilon$. Thus, we have $\lim_{T\to\infty} \alpha_T \leq h(p^*, q^*)$. This proves (7).

Now, (7) implies that

$$(\forall a \in \mathcal{X})(\liminf_{T\to\infty} \sum_{t=0}^{T-1} \frac{h(\theta_t^{a,b^*})}{T} \leq h(p^*, q^*)). \tag{12}$$

Similarly, we can show that

$$(\forall b \in \mathcal{Y})(\limsup_{T\to\infty} \sum_{t=0}^{T-1} \frac{-h(\theta_t^{a^*,b})}{T} \leq -h(p^*, q^*)), \tag{13}$$

where a^* is such that $a^*(\sigma) = \xi_{|\sigma|}$ for all $\sigma \in Y^{<\mathbb{N}}$. This implies that

$$(\forall b \in \mathcal{Y})(\liminf_{T\to\infty} \sum_{t=0}^{T-1} \frac{h(\theta_t^{a^*,b})}{T} \geq h(p^*, q^*)). \tag{14}$$

By (12), we have for all $a \in \mathcal{X}$, $u_h(a, b^*) \leq h(p^*, q^*)$. By (14), we have for all $b \in \mathcal{Y}$, $u_h(a^*, b) \geq h(p^*, q^*)$. Therefore, we have

$$u_h(a^*, b^*) \leq h(p^*, q^*) \leq u_h(a^*, b^*).$$

Therefore, (a^*, b^*) is an equilibrium. $\square$

In this theorem, we have weaker conditions than that in Theorem 3.2, but we obtain weaker results as well. In Theorem 3.2, for any equilibrium mixed strategy $p \in \Delta(X)$, there is a corresponding equilibrium strategy a that is μ_p-random relative to $\mathcal{P}^2$ in HG. In Theorem 5.1, however, we only shows that if there are stochastic sequences relative to each other's computability constraints, then these sequences are equilibrium strategies. Now we are ready to give proofs of the theorems in section 3. Notice that Theorem 3.1 is a direct implication of Theorem 3.2.

Proof of Theorem 3.2: (a) Consider the game $VG(g, \mathcal{P}^1, \mathcal{P}^2)$. Suppose that $(p, q) \in \Delta(X) \times \Delta(Y)$ is an equilibrium of g. By Theorem 6.2, there is a sequence $\zeta^1 \in X^{\mathbb{N}}$ in $\mathcal{P}^1$ that is μ_p-random relative to $\mathcal{P}^2$, and there is a sequence $\zeta^2 \in Y^{\mathbb{N}}$ in $\mathcal{P}^2$ that is μ_q-random relative to $\mathcal{P}^1$. Define a as $a(\sigma) = \zeta^1_{|\sigma|}$ for all $\sigma \in Y^{<\mathbb{N}}$ and define $b \in \mathcal{X}$ as $b(\sigma) = \zeta^2_{|\sigma|}$ for all $\sigma \in X^{<\mathbb{N}}$. By Theorem 6.6, ζ^1 is a p-stochastic random sequence relative to $\mathcal{P}^2$ and ζ^2 is a p-stochastic random sequence relative to $\mathcal{P}^1$, and, hence, by Theorem 5.1, a is an equilibrium strategy for player 1 and b is an equilibrium strategy for player 2. By Theorem 6.6, any μ_p-random sequence relative to $\mathcal{P}^2$ is also a p-stochastic sequence relative to $\mathcal{P}^2$. The result then follows directly from Theorem 5.1.

(b) By Theorem 6.3, $\zeta^1 \otimes \zeta^2$ is $\mu_{p\otimes q}$-random. Thus, the outcome $\theta^{a,b}$ is a $\mu_{p\otimes q}$-random sequence. $\square$

Proof of Theorem 3.3: Let $p \in \Delta(X)$. Since p is not degenerate, there are $x_1 \neq x_2 \in X$ such that $p_{x_1} > 0$ and $p_{x_2} > 0$. Let ξ be a strategy for player 1 in HG that is μ_p-random relative to $\mathcal{P}^2$, and let ξ' be a strategy for player 2 in HG that is μ_q-random relative to $\mathcal{P}^1$. Construct a new sequence $\zeta \in (X \times X)^{\mathbb{N}}$ as follows: $\zeta_t = (\xi_{2t}, \xi_{2t+1})$ for all $t \in \mathbb{N}$. It is easy to check that ζ is $\mu_{p\otimes p}$-random. Define $\theta : \mathbb{N} \to \mathbb{N}$ as follows:

(a) $\theta(0) = \min\{t : \zeta_t = (x_1, x_2) \vee \zeta_t = (x_2, x_1)\}$;

(b) for $t > 0$, $\theta(t) = \min\{t' : t' > \theta(t-1) \wedge (\zeta_{t'} = (x_1, x_2) \vee \zeta_{t'} = (x_2, x_1))\}$.

θ is total since ζ is $\mu_{p\otimes p}$-random and $p_{x_1} p_{x_2} > 0$. Define ξ^1 as $\xi^1_t = 0$ if $\zeta_{\theta(t)} = (x_1, x_2)$ and $\xi^1_t = 1$ if $\zeta_{\theta(t)} = (x_2, x_1)$ for all $t \in \mathbb{N}$. By Theorem 6.4, it is easy to see that ξ^1 is $\mu_{(\frac{1}{2},\frac{1}{2})}$-random.

Clearly, $\xi^1 \in \mathcal{P}^1$, and so by Theorem 6.3, ξ' is μ_q-random relative to ξ^1, and, hence, by Theorem 6.3 again, ξ^1 is $\mu_{(\frac{1}{2},\frac{1}{2})}$-random relative to $\mathcal{P}^2$. By Theorem 6.1, ξ^1 is a complex sequence relative to $\mathcal{P}^2$. The existence of ξ^2 is completely symmetric. $\square$

Proof of Theorem 3.4: Let $p \in \Delta(X)$ be an equilibrium strategy for player 1 in g. Since g has no pure strategy equilibrium, there are $x_1, x_2 \in X$ such that $p_{x_1} > 0$ and $p_{x_2} > 0$. Let $\varepsilon = \frac{\min\{p_{x_1}, p_{x_2}\}}{3}$. For any number t, let $\llcorner\sqrt{t}\lrcorner$ be the largest integer less than or equal to $\sqrt{t}$. Construct the sequence $\mathbf{p} = (p^0, p^1, ...)$ as follows:

(a) $p^t_x = p_x$ if $x \neq x_1$ and $x_1 \neq x_2$;

(b) $p^t_{x_1} = p_{x_1}$ if $\llcorner\sqrt{t}\lrcorner \leq \frac{1}{\varepsilon}$ and $p^t_{x_1} = p_{x_1} - \frac{1}{\llcorner\sqrt{t}\lrcorner}$ otherwise;

(c) $p^t_{x_2} = p_{x_2}$ if $\llcorner\sqrt{t}\lrcorner \leq \frac{1}{\varepsilon}$ and $p^t_{x_2} = p_{x_2} + \frac{1}{\llcorner\sqrt{t}\lrcorner}$ otherwise.

By construction, $p^t_x = 0$ if and only if $p^x = 0$, and $\lim_{t\to\infty} p^t = p$. Clearly, $\mathbf{p}$ is computable. By Theorem 6.2, there is a $\mu_{\mathbf{p}}$-random sequence ξ relative to $\mathcal{P}^2$ in $\mathcal{P}^1$. Now, let $X_0 = \{x \in X : p_x > 0\}$, then the sequence ξ is $\mu_{\mathbf{p}}$-random can be regarded as a sequence in $X_0^{\mathbb{N}}$. Therefore, Theorem 6.6 is applicable and so ξ is p-stochastic. Thus, by Theorem 5.1, ξ is an equilibrium strategy for player 1 in HG.

Moreover, we have that

$$\sum_{t=0}^{\infty}\sum_{x\in X}(\sqrt{p^t_x} - \sqrt{p_x})^2 \geq \sum_{t=T_0}^{\infty}(\frac{1}{4\llcorner\sqrt{t}\lrcorner^2}) \geq \sum_{t=T_0}^{\infty}\frac{1}{4t} = \infty,$$

where T_0 is the smallest number such that $\llcorner\sqrt{T_0}\lrcorner \geq \frac{1}{\varepsilon}$. By Theorem 3 in Vovk [33], ξ is not μ_p-random. Since ξ is p-stochastic, it cannot be $\mu_{p'}$-random for any $p' \in \Delta(X)$ other than p either. $\square$

6 Appendix

6.1 Recursive functions

In this section we introduce $C(\xi)$, the set of functions computable in ξ, for any $\xi \in \{0,1\}^{\mathbb{N}}$.

Definition 6.1. Let $\xi \in \{0,1\}^{\mathbb{N}}$, which can be identified as a total function in $\mathcal{F}$. Define $C(\xi)$ to be the minimal class of functions in $\mathcal{F}$ that satisfies the following conditions:

(a) $C(\xi)$ contains functions ξ, Z, S, and P_{ki} for all $k, i > 0$.

(b) $C(\xi)$ is closed under composition, primitive recursion, and minimization.

The formal definitions of operations in (b) can be found in Pippenger [28]. If $f \in C(\xi)$, we say that f is ξ-*computable*. We also say that f is *computable* if $f \in \mathcal{P}^*$. For any finite set X, we can also consider the space $X^{\mathbb{N}}$, which is called the Cantor space. Any element in $X^{\mathbb{N}}$ is called a *Turing oracle*. In the same manner, we can also consider functions that are ξ-computable for any $\xi \in X^{\mathbb{N}}$. For any computability constraint $\mathcal{P}$, there is an oracle ξ such that $\mathcal{P}$ is the set of functions that can be computed by a Turing machine with oracle ξ. For a detailed discussion and proofs, please see Pippenger [28], chapter 4. For any such a machine, we can give it a Gödel number, and we use $\varphi_e^{(k),\xi}$ to denote the function that is computed by the machine with Gödel number e and with oracle ξ, and the index k indicates that we put k numbers as input in the computation. We shall use this enumeration of functions to analyze the tests defined for randomness in section 3.

Consider a sequence $\{V_t\}_{t=0}^{\infty}$ of subsets of $X^{\mathbb{N}}$. Suppose there is a total function $f : \mathbb{N} \to \mathbb{N} \times X^{<\mathbb{N}}$ in $C(\xi)$ such that for all $t \in \mathbb{N}$ and for all $\zeta \in X^{\mathbb{N}}$,

$$\zeta \in V_t \Leftrightarrow (\exists n)(f(n) = (t, \sigma) \wedge \sigma = \zeta[|\sigma|]). \tag{15}$$

Then, we can find a total function $h \in \mathcal{P}^*$ such that[10]

$$\zeta \in V_t \Leftrightarrow \varphi_{h(t)}^{(1),\xi\oplus\zeta}(0) \downarrow, \tag{16}$$

where for any partial function f, $f(x_1, ..., x_k) \downarrow$ holds iff $f(x_1, ..., x_k) \neq \perp$.

In the following, we shall also consider functionals $\Psi : X^{\mathbb{N}} \times \mathbb{N}^2 \to \mathbb{N}$. Such a functional Ψ is ξ-*computable* if there is a number e such that for all ζ and for all x, y,

$$\Psi(\zeta, x, y) \simeq \varphi_e^{(2),\xi\oplus\zeta}(x, y).$$

[10] This follows directly from the Parametrization Theorem for relative computability. Please see Downey *et al.* [9] for a more detailed discussion.

It can also be shown that, for any sequence $\{V_t\}_{t=0}^{\infty}$, there is some $f \in C(\xi)$ such that (15) holds if and only if there is a computable functional such that

$$\zeta \in V_t \Leftrightarrow (\exists s)(\Psi(\xi, t, s) = 0). \tag{17}$$

If there is such a ξ-computable that the above relation holds, then we say that $\{V_t\}$ is of $\Sigma_1^{0,\xi}$. In case that Ψ is computable, we say that $\{V_t\}$ is of Σ_1^0.

6.2 Effective randomness

In this section we review some results from effective randomness. We include this section mainly for self-containment. We will not give all the proofs, but will refer the readers to the survey paper Downey *et al.* [9].

6.2.1 Transformations

There is a close connection between complex sequences and random sequences: any complex sequences is random with respect to the uniform distribution. The following theorem is well-known in this literature. The proof of this theorem with $\mathcal{P} = \mathcal{P}^*$ can be found in Downey *et al.* [9], and all the arguments there can be relativized and the proof is general enough to cover the general case.

Theorem 6.1. *Let $\mathcal{P} \in \mathfrak{R}$. Then, $\xi \in \{0,1\}^{\mathbb{N}}$ is a complex sequence relative to $\mathcal{P}$ if and only if ξ is a λ-random sequence relative to $\mathcal{P}$, where $\lambda(N_\sigma) = 2^{-|\sigma|}$ for all $\sigma \in \{0,1\}^{<\mathbb{N}}$.*

Our first task is to show that if $\mathcal{P}^1$ includes a complex sequence relative to $\mathcal{P}^2$, then it also includes a μ_p-random sequence relative to $\mathcal{P}^2$. As the following theorem shows, in fact, a slightly stronger result holds. The proof for the case $|X| = 2$ can be found in Zvonkin and Levin [37]. We follow a similar logic.

Theorem 6.2. *Let $\mathcal{P} \in \mathfrak{R}$. Let X be a finite set. Suppose that $\xi \in \{0,1\}^{\mathbb{N}}$ is a complex sequence relative to $\mathcal{P}$. Let $\mathbf{p} = \{p^t\}_{t=0}^{\infty}$ be a sequence over $\Delta(X)$ that is computable, i.e., is in $\mathcal{P}^*$. Then, there is a $\mu_{\mathbf{p}}$-sequence $\zeta \in X^{\mathbb{N}}$ relative to $\mathcal{P}$ belonging to $C(\xi)$.*

Proof. First we claim that there is a λ^X-random sequence $\xi' \in X^{\mathbb{N}}$ relative to $\mathcal{P}$ that is in $C(\xi)$, where $\lambda^X(N_\sigma) = |X|^{-|\sigma|}$ for all $\sigma \in X^{<\mathbb{N}}$. For a proof, see Calude [6], Theorem 7.18.

There is a natural mapping Γ between $X^{\mathbb{N}}$ and $[0,1]$:

$$\Gamma(\zeta) = \sum_{t=0}^{\infty} \iota(\zeta_t)\frac{1}{n^{t+1}},$$

where $X = \{x_1, ..., x_n\}$ and $\iota(x) = i-1$ if and only if $x = x_i$. Γ is onto but not one-to-one. However, the set $\{\zeta \in X^\omega : \Gamma(\zeta) = \Gamma(\zeta') \text{ for some } \zeta' \neq \zeta\}$ is countable, since for any such ζ, $\Gamma(\zeta)$ is a rational number.

If $\Phi : X^{\mathbb{N}} \to X^{\mathbb{N}}$ is a (Borel) measurable function, then λ^X_Φ defined as $\lambda^X_\Phi(A) = \lambda^X(\Phi^{-1}(A))$ is also a measure over $X^{\mathbb{N}}$. We will construct a computable mapping Φ such that Φ maps a λ^X-random sequence relative to $\mathcal{P}$ to a $\mu_{\mathbf{p}}$-random sequence relative to $\mathcal{P}$.

We will define Φ via a monotone function $\phi : S \to T$, where S is a and T are trees over X (a tree is a subset of $X^{<\mathbb{N}}$ that is closed under initial segments). Such a function is monotone if $\sigma \subset \tau$ implies $\phi(\sigma) \subset \phi(\tau)$. If S is a tree, $[S] = \{\zeta \in X^{\mathbb{N}} : (\forall t)\zeta[t] \in S\}$. Given such a function, let

$$D(\phi) = \{\zeta \in [S] : \lim_{t\to\infty} \phi(\zeta[t]) = \infty\}.$$

Then, define $\Phi : D(\phi) \to X^{\mathbb{N}}$ by $\Phi(\zeta) = \bigcup_{t=0}^{\infty} \phi(\zeta[t])$. We say that Φ is *obtained from ϕ via a monotone function.*

We claim that there exists a computable monotone function ϕ such that $\mu_{\mathbf{p}} = \lambda^X_\Phi$ and $\lambda^X(D(\phi)) = 1$, where $\Phi : D(\phi) \to X^{\mathbb{N}}$ is obtained from ϕ via a monotone function.

Extend Γ to $X^{<\mathbb{N}}$ as $\Gamma(\sigma) = \sum_{t=0}^{|\sigma|-1} \frac{\iota(\sigma_t)}{n^{t+1}}$. Define $g : [0,1] \to [0,1]$ as

$$g(r) = \mu_{\mathbf{p}}(\{\zeta : \Gamma(\zeta) \leq r\}),$$

i.e., the distribution function of $\mu_{\mathbf{p}}$ over $[0,1]$. Define $h = g^{-1}$. Therefore, $r \leq g(s)$ if and only if $h(r) \leq s$. Then,

$$\mu_{\mathbf{p}}(\Gamma^{-1}([0,r])) = g(r) = \lambda^X(\Gamma^{-1}([0,g(r)])) = \lambda^X(\Gamma^{-1}(h^{-1}([0,r]))).$$

Define $g^0(\epsilon) = 0$ and $g^1(\epsilon) = 1$, and, for $\tau \in X^{<\mathbb{N}} - \{\epsilon\}$, define

$$g^0(\tau) = \sum\{\mu_{\mathbf{p}}(N_\sigma) : \Gamma(\sigma) \le \Gamma(\tau) - \frac{1}{n^{|\tau|}}, |\sigma| = |\tau|\},$$

where $\sum \emptyset = 0$, and

$$g^1(\tau) = \sum\{\mu_{\mathbf{p}}(N_\sigma) : \Gamma(\sigma) \le \Gamma(\tau), |\sigma| = |\tau|\}.$$

For any $\zeta \in X^{\mathbb{N}}$,

$$\Gamma(\zeta) \le \Gamma(\tau) \Leftrightarrow \Gamma(\zeta[|\tau|]) \le \Gamma(\tau) - \frac{1}{n^{|\tau|}} \vee \Gamma(\zeta) = \Gamma(\tau),$$

and

$$\Gamma(\zeta) \le \Gamma(\tau) + \frac{1}{n^{|\tau|}} \Leftrightarrow \Gamma(\zeta[|\tau|]) \le \Gamma(\tau) \vee \Gamma(\zeta) = \Gamma(\tau) + \frac{1}{n^{|\tau|}}.$$

Since $\mu_{\mathbf{p}}$ has no atoms, we have that

$$g^0(\tau) = g(\Gamma(\tau)) \text{ and } g^1(\tau) = g(\Gamma(\tau) + \frac{1}{n^{|\tau|}}).$$

Therefore, for each $t > 0$, the class of intervals

$$\{[g^0(\tau), g^1(\tau)] : \tau \in X^{<\mathbb{N}}, |\tau| = t\} \tag{18}$$

forms a partition of $[0, 1]$.

Construct ϕ as follows: Given a string $\sigma \in X^{\mathbb{N}}$, let

$$a_\sigma = \Gamma(\sigma) \text{ and } b_\sigma = \Gamma(\sigma) + \frac{1}{n^{|\sigma|}}.$$

Let $\phi(\sigma)$ be the longest τ with $|\tau| \le |\sigma|$ such that $[a_\sigma, b_\sigma] \subset [g^0(\tau), g^1(\tau)]$. Now, for any $\sigma \in X^{<\mathbb{N}}$, $\phi(\sigma)$ is well-defined, since the intervals in (18) forms a partition and $[g^0(\epsilon), g^1(\epsilon)] = [0, 1]$. To see that ϕ is monotone, suppose that $\sigma \subset \sigma'$ and $\tau = \phi(\sigma)$, $\tau' = \phi(\sigma')$. It is easy to check that $a_\sigma \le a_{\sigma'}$ and $b_{\sigma'} \le b_\sigma$. Now, if $\Gamma(\tau') \ge \Gamma(\tau) + \frac{1}{n^{|\tau|}}$, then

$$a_{\sigma'} \ge g^0(\tau') = g(\Gamma(\tau')) \ge g(\Gamma(\tau) + \frac{1}{n^{|\tau|}}) = g^1(\tau) \ge b_\sigma \ge b_{\sigma'},$$

a contradiction. On the other hand, we have $|\tau'| \ge |\tau|$. If $\Gamma(\tau') < \Gamma(\tau)$, then $\Gamma(\tau') \le \Gamma(\tau) - \frac{1}{n^{|\tau|}}$, and hence,

$$b_\sigma \le b_{\sigma'} \le g^1(\tau') = g(\Gamma(\tau') + \frac{1}{n^{|\tau'|}}) \le g(\Gamma(\tau)) = g^0(\tau) \le a_\sigma,$$

a contradiction. Thus, we have $\tau \subset \tau'$.

Then we show that $X^{\mathbb{N}} - \{\zeta \in X^{\mathbb{N}} : h(\Gamma(\zeta)) = \frac{m}{n^t} \text{ for some } m, n, t \in \mathbb{N}\} \subseteq D(\phi)$. Suppose that $h(\Gamma(\zeta)) \neq \frac{m}{n^t}$ for any $m, n, t \in \mathbb{N}$. Let K be given. There exists some $l \in \mathbb{N}$ such that $h(\Gamma(\zeta)) \in (\frac{l}{n^K}, \frac{l+1}{n^K})$. Let

$$\varepsilon = \min\{h(\Gamma(\zeta)) - \frac{l}{n^K}, \frac{l+1}{n^K} - h(\Gamma(\zeta))\}.$$

Since h is continuous, there is some T such that $t \geq T$ implies that

$$\min\{|h(b_{\zeta[t]}) - h(\Gamma(\zeta))|, |h(\Gamma(\zeta) - h(a_{\zeta[t]})|\} \leq \frac{\varepsilon}{2} \text{ and so } [h(a_{\zeta[t]}), h(b_{\zeta[t]})] \subseteq (\frac{l}{n^K}, \frac{l+1}{n^K}).$$

Thus, if $t \geq \max\{T, K\}$, then

$$[a_{\zeta[t]}, b_{\zeta[t]}] \subset [g(\frac{l}{n^K}), g(\frac{l+1}{n^K})] = [g^0(\frac{l}{n^K}), g^1(\frac{l}{n^K})],$$

and so $|\phi(\zeta[t])| \geq K$. Therefore, $\lambda^X(D(\phi)) = 1$.

Now, we claim that if $\zeta \in D(\phi)$, then $\Gamma(\Phi(\zeta)) = h(\Gamma(\zeta))$. Let ε be given, and let K be so large that $\varepsilon < \frac{1}{n^{K-1}}$. Since $\zeta \in D(\phi)$, there exists T such that $t \geq T$ implies that $|\phi(\zeta[t])| \geq K$. Then, for all $t \geq T$,

$$h(\Gamma(\zeta)) \in [h(a_{\zeta[t]}), h(b_{\zeta[t]})] \subseteq [a_{\phi(\zeta[t])}, b_{\phi(\zeta[t])}],$$

and so

$$h(\Gamma(\zeta)) - \Gamma(\phi(\zeta[t])) \leq \frac{1}{n^K} \leq \varepsilon.$$

Thus,

$$\Gamma(\Phi(\zeta)) = \lim_{t \to \infty} \Gamma(\phi(\zeta[t])) = h(\Gamma(\zeta)).$$

Moreover, for any $r \in [0, 1]$, there is a sequence $\zeta \in X^{\mathbb{N}}$ such that $\Gamma(\Phi(\zeta)) = r$, since h is strictly increasing and is continuous. Also, we have that

$$\Gamma(\Phi(\zeta)) \geq \Gamma(\Phi(\zeta')) \Leftrightarrow \Gamma(\zeta) \geq \Gamma(\zeta').$$

Now, we show that $\lambda^X_\Phi = \mu_{\mathbf{p}}$ by demonstrating that they share the same distribution function:

$$\lambda^X_\Phi(\{\zeta : \Gamma(\zeta) \leq \Gamma(\Phi(\zeta^*))\}) = \lambda^X(\{\zeta : \Gamma(\Phi(\zeta)) \leq \Gamma(\Phi(\zeta^*))\})$$

$$= \lambda^X(\{\zeta : \Gamma(\zeta) \leq \Gamma(\zeta^*)\}) = \Gamma(\zeta^*) = g(\Gamma(\Phi(\zeta^*))).$$

Notice that if $\Gamma(\zeta) = g(\frac{m}{n^t}) \in \mathbb{Q}$, then ζ is computable. Thus, $\xi' \in D(\phi)$. Let $\zeta' = \Phi(\xi')$. Now we show that ζ' is $\mu_{\mathbf{p}}$-random relative to $\mathcal{P}$. Suppose not. Then there is a $\mu_{\mathbf{p}}$-test $\{V_t\}_{t=0}^{\infty}$ relative to $\mathcal{P}$ such that $\zeta' \in \bigcap_{t=0}^{\infty} V_t$. Let $U_t = \Phi^{-1}(V_t)$. Then, for all t, $\lambda^X(U_t) = \mu_{\mathbf{p}}(V_t) \leq \frac{1}{2^t}$. Moreover, since ϕ is computable, $\{U_t\}_{t=0}^{\infty}$ is a λ^X-test relative to $\mathcal{P}$. But $\xi' \in \bigcap_{t=0}^{\infty} U_t$ since $\zeta' \in \bigcap_{t=0}^{\infty} V_t$, a contradiction. Since ϕ is computable, $\zeta' \in C(\xi') \subset C(\xi)$. □

6.2.2 Independence

Next, we shall consider independence. This concept is crucial to prove Theorem 3.2, part (1), because in that part, we claim that not only equilibrium strategy for each player is random, but their joint behavior is random as well. First we give some notations. For any distribution $(p, q) \in \Delta(X) \times \Delta(Y)$, we define $p \otimes q$ to be the product measure of them over $X \times Y$. For any two sequences $\xi \in X^{\mathbb{N}}$ and $\zeta \in Y^{\mathbb{N}}$, we define $\xi \otimes \zeta$ as $(\xi \otimes \zeta)_t = (\xi_t, \zeta_t)$ for all $t \in \mathbb{N}$. In the axiomatic probability theory, independence of random variables is defined in terms of product measures: a random variable on X and a random variable on Y are independent in the standard theory if their joint distribution is a product distribution over $X \times Y$. Similar to Definition 3.2, we can define $\mu_{p\otimes q}$-randomness (relative to $\mathcal{P}^*$) in $(X \times Y)^{\mathbb{N}}$, where for all $x \in X$, $y \in Y$, $p \otimes q_{(x,y)} = p_x q_y$. The following theorem, essentially due to van Lambalgen [15], characterizes independence in terms of randomness with respect to product measures, which establishes a connection between our definition of independence and the measure theoretical definition. The proof for the case $|X| = |Y| = 2$ and $p = (\frac{1}{2}, \frac{1}{2}) = q$ can be found in Downey *et al.* [9], Theorem 12.12, 12.13.

Theorem 6.3. *Consider two finite sets X and Y. Suppose $\xi \in X^{\mathbb{N}}$ and $\zeta \in Y^{\mathbb{N}}$, and suppose $p \in \Delta(X)$ and $q \in \Delta(Y)$.*

(a) If $\xi \otimes \zeta$ is $\mu_{p\otimes q}$-random, then ξ is μ_p-random relative to $C(\zeta)$.

(b) If ξ is μ_p-random relative to $C(\zeta)$ and ζ is μ_q-random, then $\xi \otimes \zeta$ is $\mu_{p\otimes q}$-random.

Proof. (a) Suppose that ξ is not p-random relative to $C(\zeta)$. Then $\xi \in \bigcap_{t=0}^{\infty} U_t^{\zeta}$ for some uniformly μ_p-test $\{U_t^{\zeta}\}_{t=0}^{\infty}$ relative to $C(\zeta)$ in $X^{\mathbb{N}}$ such that $\mu_p(U_t^{\zeta}) \leq \frac{1}{2^t}$. Since it is a test, by (16), let h be a total computable function such that $\xi' \in U_t^{\zeta}$ if and only if $\varphi_{h(t)}^{(1),\zeta\oplus\xi'}(0)\downarrow$. For any $\zeta' \in Y^{\mathbb{N}}$, define

$$U_{t,s}^{\zeta'} = \{\xi' \in X^{\mathbb{N}} : \varphi_{h(t)}^{(1),(\zeta'\oplus\xi')[2s]}(0)\downarrow\},\ U_t^{\zeta'}[\frac{1}{2^t}] = \bigcup_{\mu_p(U_{t,s}^{\rho})\leq\frac{1}{2^t}} U_{t,s}^{\zeta'}. \tag{19}$$

Let

$$V_t = \{\xi' \otimes \zeta' : \xi' \in X^{\mathbb{N}},\ \zeta' \in Y^{\mathbb{N}},\ \xi' \in U_t^{\zeta'}[\frac{1}{2^t}]\}. \tag{20}$$

We claim that $\{V_t\}_{t=0}^{\infty}$ is a $\mu_{p\otimes q}$-test.

Now, by (19) and (20), $\xi' \otimes \zeta' \in V_t$ if and only if $(\exists s)(\mu_p(U_{t,s}^{\zeta'}) \leq \frac{1}{2^t} \wedge \xi' \in U_{t,s}^{\zeta'})$ if and only if $(\exists s)(\mu_p(U_{t,s}^{\zeta'}) \leq \frac{1}{2^t} \wedge \varphi_{h(t)}^{(1),\zeta'\oplus\xi'[2s]}(0)\downarrow)$. We claim that the predicates $\varphi_{h(t)}^{(1),\zeta'\oplus\xi'[2s]}(0)\downarrow$ and $\mu_p(U_{t,s}^{\zeta'}) \leq \frac{1}{2^t}$ are computable.

(a.1) The functional $(\xi' \otimes \zeta', s) \mapsto \zeta' \oplus \xi'[2s]$ is computable in $(X \times Y)^{\mathbb{N}} \times \mathbb{N}$, and so is the predicate $\varphi_{h(t)}^{(1),\sigma}(0)\downarrow$ in $\mathbb{N}^2$ (in t and σ). Thus, by generalized composition, the predicate $\varphi_{h(t)}^{(1),\zeta'\oplus\xi'[2s]}(0)\downarrow$ is computable in $(X \times Y)^{\mathbb{N}} \times \mathbb{N}^2$ (in $\xi' \otimes \zeta'$, t, and s).

(a.2) It is easy to check that

$$U_{t,s}^{\zeta'} = \bigcup\{N_\sigma : \varphi_{h(t)}^{(1),(\zeta'[s]\oplus\sigma)}(0)\downarrow, |\sigma| = s\},$$

and

$$\mu_p(U_{t,s}^{\zeta'}) = \sum_{\sigma}\{\prod_{j=0,\ldots,s-1} p_{\sigma_j} : \varphi_{h(t)}^{(1),(\zeta'[s]\oplus\sigma)}(0)\downarrow \wedge |\sigma| = s\}.$$

The functional $(\xi' \otimes \zeta', s) \mapsto \zeta'[s]$ is computable in $(X \times Y)^{\mathbb{N}} \times \mathbb{N}$ and so is the predicate $\sum_{\sigma}\{\prod_{j=0,\ldots,s-1} p_{\sigma_j} : \varphi_{h(t)}^{(1),(\sigma'\oplus\sigma)}(0)\downarrow \wedge |\sigma| = s\}$ in $\mathbb{N}^3$ (in σ', t, and s). By generalized composition, the functional $(\xi' \otimes \zeta', s) \mapsto \mu_p(U_{t,s}^{\zeta'})$ is computable, and so the predicate $\mu_p(U_{t,s}^{\zeta'}) \leq \frac{1}{2^t}$ is computable.

Thus, there is a f in $\mathcal{P}^*$ such that (15) holds. Moreover,

$$\begin{aligned}\mu_{p\otimes q}(V_t) &= \textstyle\int_{(X\times Y)^{\mathbb{N}}} \chi_{V_t}(\xi' \otimes \zeta') d\mu_{p\otimes q}(\xi' \otimes \zeta') \\ &= \textstyle\int_{Y^{\mathbb{N}}} \int_{X^{\mathbb{N}}} \chi_{U_t^{\zeta'}[\frac{1}{2^t}]}(\xi') d\mu_p(\xi') d\mu_q(\zeta') \\ &= \textstyle\int_{Y^{\mathbb{N}}} \mu_p(U_t^{\zeta'}[\frac{1}{2^t}]) d\mu_q(\zeta') \leq \frac{1}{2^t}.\end{aligned}$$

$\{V_t\}_{t=0}^\infty$ is a $\mu_{p\otimes q}$-test. But $\xi\otimes\zeta\in V_t$ for all $t\in\mathbb{N}$, and so $\xi\otimes\zeta$ is not $\mu_{p\otimes q}$-random.

(b) Suppose that $\xi\otimes\zeta\in(X\times Y)^{\mathbb{N}}$ is not $\mu_{p\otimes q}$-random. Then, $\xi\otimes\zeta\in\bigcap_{t=0}^\infty U_t$ for some $\mu_{p\otimes q}$-test $\{U_t\}$ in $(X\times Y)^{\mathbb{N}}$ such that $\mu_{p\otimes q}(U_t)\le\frac{1}{4^t}$. Suppose that $\xi'\otimes\zeta'\in U_t$ if and only if $\varphi_{h(t)}^{(1),\xi'\otimes\zeta'}(0)\downarrow$, where h is a total and computable function. Let $\zeta'\in Y^{\mathbb{N}}$, and define

$$V_t^{\zeta'}=\{\xi'\in X^{\mathbb{N}}:\xi'\otimes\zeta'\in U_t\},\ W_t=\{\zeta'\in Y^{\mathbb{N}}:\mu_p(V_t^{\zeta'})>\frac{1}{2^t}\}. \tag{21}$$

We claim that, for each $\zeta'\in Y^{\mathbb{N}}$, there is a ζ'-computable functional such that (17) holds for $\{V_t^{\zeta'}\}_{t=0}^\infty$.

(b.1) Now,

$$\xi'\in V_t^{\zeta'}\ \Leftrightarrow\ \varphi_{h(t)}^{(1),\xi'\otimes\zeta'}(0)\downarrow\Leftrightarrow(\exists s)\varphi_{h(t)}^{(1),(\xi'\otimes\zeta')[s]}(0)\downarrow.$$

Since the functional $(\xi',s)\mapsto(\xi'\otimes\zeta')[s]$ is ζ'-computable, and the predicate $\varphi_{h(t)}^{(1),\sigma}(0)\downarrow$ is computable in (t,σ), the claim is proved.

We claim that there is a computable functional that satisfies (17) for $\{W_t\}$ and $\mu_q(W_t)\le\frac{1}{2^t}$.

(b.2) $\zeta'\in W_t$ if and only if $\mu_p(V_t^{\zeta'})>\frac{1}{2^t}$. But if we define

$$V_{t,s}^{\zeta'}=\bigcup\{N_\sigma\subseteq X^{\mathbb{N}}:\sigma\in X^s\wedge\varphi_{h(t)}^{(1),\sigma\otimes(\zeta'[s])}(0)\downarrow\},$$

then $V_{t,s}^{\zeta'}\subseteq V_{t,s+1}^{\zeta'}$ for all $s\in\mathbb{N}$ and $V_t^{\zeta'}=\bigcup_{s=0}^\infty V_{t,s}^{\zeta'}$. Thus, $\mu_p(V_t^{\zeta'})>\frac{1}{2^t}$ if and only if $(\exists s)(\mu_p(V_{t,s}^{\zeta'})>\frac{1}{2^t})$. It is easy to check that

$$\mu_p(V_{t,s}^{\zeta'})=\sum_\sigma\{\prod_{j=0,\dots,s-1}p_{\sigma_j}:\varphi_{h(t)}^{(1),\sigma\otimes(\zeta'[s])}(0)\downarrow\wedge\sigma\in X^s\},$$

which is a computable functional of (ζ',s,t), and so $\{W_t\}_{t=0}^\infty$ is uniformly Σ_1^0. Now,

$$\begin{aligned}\mu_{p\otimes q}(U_t)&=\textstyle\int_{(X\times Y)^{\mathbb{N}}}\chi_{U_t}(\xi'\otimes\zeta')d\mu_{p\otimes q}(\xi'\otimes\zeta')\\&=\textstyle\int_{(X\times Y)^{\mathbb{N}}}\chi_{V_t^{\zeta'}}(\xi')d\mu_{p\otimes q}(\xi'\otimes\zeta')=\int_{Y^{\mathbb{N}}}\mu_p(V_t^{\zeta'})d\mu_q(\zeta')\\&>\textstyle\int_{Y^{\mathbb{N}}}\frac{1}{2^t}\chi_{W_t}(\zeta')d\mu_q(\zeta')=\frac{1}{2^t}\mu_q(W_t).\end{aligned}$$

Thus,

$$\mu_q(W_t)<2^t\mu_{p\otimes q}(U_t)\le\frac{1}{2^t}.$$

Since ζ is μ_q-random, by Solovay's Theorem (see Downey *et al.* Definition 3.2), there is some $L \in \mathbb{N}$ such that $\zeta \notin W_t$ for all $t \geq L$. Thus, by (21), for all $t \geq L$, $\mu_p(V_t^\zeta) \leq \frac{1}{2^t}$. By (b.1), $\{V_t^\zeta\}_{t=0}^\infty$ is a μ_p-test relative to $C(\zeta)$. But $\xi \in V_t^\zeta$ for all t, and so ξ is not p-random relative to $C(\zeta)$. $\square$

6.2.3 Decomposition

Now we shall consider conditional probability and decomposition. We use the result here only for Theorem 3.3. Let ξ be a sequence in $X^{\mathbb{N}}$. For any $A \subset X$, we define $\nu^{\xi,A} \in \{0,1\}^{\mathbb{N}}$ to be the sequence such that $\nu_t^{\xi,A} = 1$ if $\xi_t \in A$ and $\nu_t^{\xi,A} = 0$ otherwise. $\nu^{\xi,A}$ records the occurrences of the event A.

For any $\nu \in \{0,1\}^{\mathbb{N}}$, we define $\theta^\nu : \mathbb{N} \to \mathbb{N}$ as follows:

$$\theta^\nu(0) \text{ is the least } t' \text{ such that } \nu_{t'} = 1; \tag{22}$$

$$\theta^\nu(t+1) \text{ is the least } t' \text{ such that } \nu_{t'} = 1 \text{ and } t' > \theta^\nu(t). \tag{23}$$

θ^ν is then a partial ν-computable function. We can extend this for strings τ in $\{0,1\}^{\mathbb{N}}$ as well:

$$\theta^\tau(0) \text{ is the least } t' \text{ such that } \tau_{t'} = 1;$$

$$\theta^\tau(t+1) \text{ is the least } t' \text{ such that } \tau_{t'} = 1 \text{ and } t' > \theta^\tau(t).$$

In this case, θ^τ is always a partial function.

Applying the construction to $\nu^{\xi,A}$, $\theta^{\nu^{\xi,A}}$ is then ξ-computable and is total if elements in A appear in ξ infinitely often. Define $\xi_t^A = \xi_{\theta^{\nu^{\xi,A}}(t)}$, $t \in \mathbb{N}$. $\xi^A \in A^{\mathbb{N}}$ if and only if $\theta^{\nu^{\xi,A}}$ is total. The sequence ξ^A records the happenings in ξ given the event A. Intuitively, ξ^A should be a random sequence as well, and it should follow the conditional distribution. If $A = X - \{x\}$ for some $x \in X$, then $\nu^{\xi,A}$ is also denoted as $\nu^{\xi,-x}$ and ξ^A is denoted as ξ^{-x}.

On the other hand, let $\xi,\ \zeta \in X^{\mathbb{N}}$ and let $\nu \in \{0,1\}^{\mathbb{N}}$. We shall define an inverse operator that is intended to model composite random processes. We shall now define the *shuffle* of ξ and ζ using ν, denoted by $\xi \oslash_\nu \zeta$, as follows: for all $t \in \mathbb{N}$,

$$(\xi \oslash_\nu \zeta)_t = (1 - \nu_t)\xi_{\sum_{s=0}^{t-1}(1-\nu_s)} + \nu_t \zeta_{\sum_{s=0}^{t-1}\nu_s}. \tag{24}$$

The two sequences ξ and ζ can be thought of as two independent processes, and the shuffle of them using ν is the composite process taking ν as a random device. Intuitively, the shuffle is expected to be random and follow the distribution that is a convex combination of the two processes. Notice that for all $t \in \mathbb{N}$ such that $\theta^\nu(t)$ is defined, $(\xi \oslash_\nu \zeta)_{\theta^\nu(t)} = \zeta_t$. Similarly, for all $t \in \mathbb{N}$ such that $\theta^\nu(t)$ is defined, $(\xi \oslash_\nu \zeta)_{\theta^{1-\nu}(t)} = \xi_t$.

Likewise, for strings $\sigma,\ \sigma' \in X^{<\mathbb{N}}$ with and $\tau \in \{0,1\}^\mathbb{N}$ such that $|\sigma| = |\tau| = |\sigma'| = s$, $\sigma \oslash_\tau \sigma' \in X^s$ is defined as the follows: for all $t = 0, .., s-1$,

$$(\sigma \oslash_\tau \sigma')_t = (1-\tau_t)\sigma_{\sum_{u=0}^{t-1}(1-\tau_u)} + \tau_t \sigma'_{\sum_{u=0}^{t-1}\tau_u}.$$

For any $x \in X$, let $\mathbf{x}$ denote the sequence in $X^\mathbb{N}$ such that $\mathbf{x}_t = x$ for all $t \in \mathbb{N}$. Then, it is easy to check that $\mathbf{x} \oslash_{\nu^{\xi,-x}} \xi^{-x} = \xi$ for any $\xi \in X^\mathbb{N}$.

Now we shall show that all these intuitions are true. The way we prove this is to establish some measure theoretical lemmas, and then apply them to construct tests.

Lemma 6.1. *Consider any finite set X with $|X| \geq 2$.*

(a) Let $x \in X$. Consider the mapping $T : (X - \{x\})^\mathbb{N} \times \{0,1\}^\mathbb{N} \to X^\mathbb{N}$ such that $T(\xi,\nu) = \mathbf{x} \oslash_\nu \xi$, and consider the probability distribution $p \in \Delta(X)$ such that $p_x < 1$. For any measurable set $V \subseteq X^\mathbb{N}$, we have

$$\mu_p(V) = \mu_{p^{-x} \otimes (p_x, 1-p_x)}(T^{-1}(V)), \tag{25}$$

where $p^{-x} \in \Delta(X - \{x\})$ is defined as $p^{-x}_{x'} = \frac{p_{x'}}{1-p_x}$ for all $x' \neq x$.

(b) Let $p, q \in \Delta(X)$ and let $\alpha \in [0,1] \cap \mathbb{Q}$. Let $T : X^\mathbb{N} \times X^\mathbb{N} \times \{0,1\}^\mathbb{N} \to X^\mathbb{N}$ be such that $T(\xi,\zeta,\nu) = \xi \oslash_\nu \zeta$. For any measurable set $V \subseteq X^\mathbb{N}$,

$$\mu_{p\otimes q \otimes (\alpha, 1-\alpha)}(T^{-1}(V)) = \mu_{\alpha p + (1-\alpha) q}(V). \tag{26}$$

Proof. (a) Since μ_p is regular, it suffices to show that (25) holds for all open sets V in $X^\mathbb{N}$. First we remark two facts. For any $\sigma,\ \sigma' \in X^{<\mathbb{N}}$, $T^{-1}(N_\sigma \cap N'_\sigma) = T^{-1}(N_\sigma) \cap T^{-1}(N'_\sigma)$. For any collection $G \subseteq X^{<\mathbb{N}}$, $\bigcup_{\sigma\in G} T^{-1}(N_\sigma) = T^{-1}(\bigcup_{\sigma \in G} N_\sigma)$. Now, any open set $V \subseteq X^\mathbb{N}$ can be written as $V = \bigcup_{\sigma\in G} N_\sigma$ for some prefix-free set $G \subseteq X^{<\mathbb{N}}$, and so if

$$\mu_{p^{-x}\otimes(p_x,1-p_x)}(T^{-1}(N_\sigma)) = \mu_p(N_\sigma) \tag{27}$$

holds for all $\sigma \in X^{<\mathbb{N}}$, then

$$\mu_{p^{-x}\otimes(p_x,1-p_x)}(T^{-1}(V)) = \sum_{\sigma\in G} \mu_{p^{-x}\otimes(p_x,1-p_x)}(T^{-1}(N_\sigma)) = \sum_{\sigma\in G} \mu_p(N_\sigma) = \mu_p(V),$$

for any such open set V.

Now we show that (27) holds for any $\sigma \in X^{<\mathbb{N}}$. It is easy to check that $T^{-1}(N_\sigma) = N_{\sigma'} \times N_{\tau'}$, where $\tau'_t = 0$ if $\sigma_t = x$, $\tau'_t = 1$ otherwise, and $\sigma'_t = \sigma_{\theta^\tau(t)}$ for $t = 0, ..., |\sigma'| - 1$, where $|\sigma'| = |\{t : 0 \le t \le |\sigma| - 1 : \sigma_t \neq x\}|$.

Thus, $\mu_{p^{-x}\otimes(p_x,1-p_x)}(T^{-1}(N_\sigma))$

$= \mu_{p^{-x}}(N_{\sigma'}) \times \mu_{(p_x,1-p_x)}(N_{\tau'})$

$= p_x^{|\tau'|-\sum_{t=0}^{|\tau'|-1}\tau'_t}(1-p_x)^{\sum_{t=0}^{|\tau'|-1}\tau'_t}(\frac{1}{1-p_x})^{|\sigma'|}\prod_{t=0}^{|\sigma'|-1} p_{\sigma'_t}$

$= p_x^{|\tau'|-\sum_{t=0}^{|\tau'|-1}\tau'_t}\prod_{t=0}^{|\sigma'|-1} p_{\sigma'_t}$ $(|\sigma'| = \sum_{t=0}^{|\tau'|-1}\tau'_t)$

$= \prod_{t=0}^{|\sigma|-1} p_{\sigma_t} = \mu_p(N_\sigma)$.

(b) As in (a), it suffices to show that (26) holds for all basic open sets. Let $\sigma \in X^{<\mathbb{N}}$. $(\xi, \zeta, \nu) \in T^{-1}(N_\sigma)$ if and only if there are $\tau \in \{0,1\}^{<\mathbb{N}}$, $\sigma^1,\ \sigma^2 \in X^{<\mathbb{N}}$ such that

(b.1) $|\tau| = |\sigma|$;

(b.2) $|\sigma^1| = |\{s : 0 \le s \le |\tau| - 1,\ \tau_s = 0\}|$ and $|\sigma^2| = |\{s : 0 \le s \le |\tau| - 1,\ \tau_s = 1\}|$;

(b.3) for all $t = 0, ..., |\sigma^1| - 1$, $\sigma_{\theta^{1-\tau}(t)} = \sigma^1_t$, and for all $t = 0, ..., |\sigma^2| - 1$, $\sigma_{\theta^\tau(t)} = \sigma^2_t$;

(b.4) $\sigma^1 \subset \xi$, $\sigma^2 \subset \zeta$, and $\tau \subset \nu$.

Hence,

$$T^{-1}(N_\sigma) = \bigcup\{N_{\sigma^1} \times N_{\sigma^2} \times N_\tau : \tau,\ \sigma^1,\ \sigma^2 \text{satisfy (b.1-3) above}\}. \tag{28}$$

Notice that for each $\tau \in \{0,1\}^{\mathbb{N}}$, there is a unique pair (σ^1, σ^2) that satisfy (b.1-3) above.

$\mu_{p\otimes q\otimes(\alpha,1-\alpha)}(T^{-1}(N_\sigma))$

$= \sum_{|\tau|=|\sigma|}\prod_{s=0}^{|\tau|-1}(\alpha p_{\sigma_s})^{1-\tau_s}((1-\alpha)q_{\sigma_s})^{\tau_s}$

$= \prod_{s=0}^{|\sigma|-1}(\alpha p_{\sigma_s} + (1-\alpha)q_{\sigma_s}) = \mu_{\alpha p+(1-\alpha)q}(N_\sigma)$. □

The following theorem states that Martin-Löf randomness is closed under conditional probability.

Theorem 6.4. *Let A be a subset of X and let $p \in \Delta(X)$ be such that $p_A =_{def} \sum_{x \in A} p_x > 0$. Suppose that $\xi \in X^{\mathbb{N}}$ is p-random. Then,*

(a) $\nu^{\xi,A}$ is $(1-p_A, p_A)$-random relative to ξ^A;

(b) ξ^A is p^A-random relative to $\nu^{\xi,A}$, where $p^A_x = \frac{p_x}{p_A}$ for all $x \in A$ and $p^A_x = 0$ otherwise.

Proof. We first show that the theorem holds for A of the form $X - \{x\}$.

(a) (for $A = X - \{x\}$) Suppose that $\nu^{\xi,-x} \in \bigcap_{t=0}^{\infty} U_t$, where $\{U_t\}_{t=0}^{\infty}$ is a $\mu_{(p_x, 1-p_x)}$-test relative to ξ^{-x} with $\mu_{(p_x,1-p_x)}(U_t) \leq \frac{1}{2^t}$. Let $U_t = \{\nu \in \{0,1\}^{\mathbb{N}} : \varphi^{(1),\nu \oplus \xi^{-x}}_{h(t)}(0) \downarrow\}$, where h is a total computable function. For each $\zeta \in (X - \{x\})^{\mathbb{N}}$, define $U^{\zeta}_{t,s} = \{\nu \in \{0,1\}^{\mathbb{N}} : \varphi^{(1),(\nu \oplus \zeta)[2s]}_{h(t)}(0) \downarrow\}$. Let

$$U^{\zeta}_t[\frac{1}{2^t}] = \bigcup_{\mu_{(p_x,1-p_x)}(U^{\zeta}_{t,s}) \leq \frac{1}{2^t}} U^{\zeta}_{t,s}.$$

For each $\sigma \in (X - \{x\})^{<\mathbb{N}}$, we define

$$U^{\sigma}_t = \{\nu \in \{0,1\}^{\mathbb{N}} : \varphi^{(1),\nu[|\sigma|] \oplus \sigma}_{h(t)}(0) \downarrow\}.$$

By construction, we have that $\nu^{\xi,-x} \in \bigcap_{t=0}^{\infty} U^{\xi^{-x}}_t = \bigcap_{t=0}^{\infty} U^{\xi^{-x}}_t[\frac{1}{2^t}]$.

For each $t \in \mathbb{N}$, we define $V_t \subseteq X^{\mathbb{N}}$ as follows: $\xi' \in V_t$ if and only if there is some $s \in \mathbb{N}$ such that (notice that $(\xi')^{-x}[s]$ is defined if and only if $(\forall j < s)(\theta^{\nu^{\xi',-x}}(j) \downarrow)$)

(a1) $(\forall j < s)(\theta^{\nu^{\xi',-x}}(j) \downarrow)$;

(a2) $\varphi^{(1),\nu^{\xi',-x}[s] \oplus (\xi')^{-x}[s]}_{h(t)}(0) \downarrow$;

(a3) $\mu_{(p_x,1-p_x)}(U^{(\xi')^{-x}[s]}_t) \leq \frac{1}{2^t}$.

We claim that $\{V_t\}_{t=0}^{\infty}$ is is a μ_p-test with $\mu_p(V_t) \leq \frac{1}{2^t}$.

For the first half of the claim, it suffices to check that all predicates in (a1-3) are of Σ^0_1.

(a1) The functional $(\xi', j) \mapsto \theta^{\nu^{\xi',-x}}(j)$ is computable, and so the predicate in (a1) is of Σ^0_1 in $X^{\mathbb{N}} \times \mathbb{N}^2$.

(a2) The functional $(\xi', s) \mapsto (\nu^{\xi',-x}[s] \oplus (\xi')^{-x}[s])$ is computable (it is undefined if $\theta^{\nu^{\xi',-x}}(j)$ is undefined for some $j < s$). Thus, the predicate in (a2) is of Σ^0_1 in $X^{\mathbb{N}} \times \mathbb{N}^2$.

(a3) The functional $(\xi', s, t) \mapsto \mu_{(p_x,1-p_x)}(U_t^{(\xi')^{-x}[s]})$ is computable since

$$U_t^{\sigma} = \bigcup \{N_\tau : \tau \in \{0,1\}^{\mathbb{N}}, |\tau| = |\sigma|, \varphi_{h(t)}^{(1),\tau\oplus\sigma}(0) \downarrow\}$$

and so

$$\mu_{(p_x,1-p_x)}(U_t^{(\xi')^{-x}[s]}) = \sum_{\tau \in \{0,1\}^{<\mathbb{N}}} \{ \prod_{j=0,\dots,s-1} p_x^{1-\tau_j}(1-p_x)^{\tau_j} : \varphi_{h(t)}^{(1),(\tau\oplus(\xi')^{-x}[s])}(0) \downarrow \wedge |\tau| = s\}.$$

Thus, the predicate $\xi' \in V_t$ is Σ^0_1, and so $\{V_t\}_{t=0}^{\infty}$ is uniformly Σ^0_1.

Let $\Gamma = \{\xi' \in X^{\mathbb{N}} : (\forall t)(\theta^{\nu^{\xi',-x}}(t) \downarrow)\}$. Let $\Gamma_s = \{\xi' \in X^{\mathbb{N}} : \xi'_s \neq x\}$. Then $\mu_p(\Gamma_s) = 1 - p_x$, and $\{\Gamma_s\}_{s=0}^{\infty}$ is a sequence of independent events. Since $p_x < 1$,

$$\sum_{s=0}^{\infty} \mu_p(\Gamma_s) = \sum_{s=0}^{\infty} (1 - p_x) = \infty$$

and so by the second Borel-Cantelli lemma,

$$\mu_p(\Gamma) = \mu_p(\bigcap_{t=0}^{\infty} \bigcup_{s=t}^{\infty} \Gamma_s) = 1.$$

For each $t \in \mathbb{N}$, let $V_t^1 = \Gamma \cap V_t$ and let $V_t^0 = (X^{\mathbb{N}} - \Gamma) \cap V_t$. So

$$\mu_p(V_t) = \mu_p(V_t^0) + \mu_p(V_t^1) = \mu_p(V_t^1).$$

Moreover, $\xi' \in V_t^1$ if and only if

$$\xi' \in \Gamma \wedge (\exists s)(\varphi_{h(s)}^{(1),\nu^{\xi',-x}[s]\oplus(\xi')^{-x}[s]}(0) \downarrow \wedge \mu_{(p_x,1-p_x)}(U_t^{(\xi')^{-x}[s]}) \leq \frac{1}{2^t})$$

if and only if $\theta^{\nu^{\xi',-x}}$ is total and $\nu^{\xi',-x} \in U_t^{(\xi')^{-x}}[\frac{1}{2^t}]$. Thus, if we define T to be such that $T(\zeta, \nu) = \mathbf{x} \oslash_\nu \zeta$, then, by Lemma 6.1 (a),

$\mu_p(V_t^1) = \mu_{p^{-x} \otimes (p_x,1-p_x)}(T^{-1}(V_t^1))$

$\leq \int_{(X-\{x\})^{\mathbb{N}}} \int_{\{0,1\}^{\mathbb{N}}} \chi_{U_t^{\zeta}[\frac{1}{2^t}]}(\nu) d\mu_{(p_x,1-p_x)}(\nu) d\mu_{p^{-x}}(\zeta)$

$\leq \int_{(X-\{x\})^{\mathbb{N}}} \mu_{(p_x,1-p_x)}(U_t^{\zeta}[\frac{1}{2^t}]) d\mu_{p^{-x}}(\zeta) \leq \frac{1}{2^t}.$

Therefore, $\{V_t\}_{t=0}^{\infty}$ is a μ_p-test, and we have $\xi = \mathbf{x} \oslash_{\nu^{\xi,-x}} \xi^{-x} \in V_t^1 \subseteq V_t$ for all $t \in \mathbb{N}$. Hence, ξ is not μ_p-random in $X^{\mathbb{N}}$.

(b) (for $A = X - \{x\}$) Suppose that $\xi^{-x} \in \bigcap_{t=0}^{\infty} U_t^{\nu^{\xi,-x}}$, where $\{U_t^{\nu^{\xi,-x}}\}_{t=0}^{\infty}$ is a uniformly $\Sigma_1^{0,\nu^{\xi,-x}}$ sequence of sets in $(X-\{x\})^{\mathbb{N}}$ with $\mu_{p^{-x}}(U_t^{\nu^{\xi,-x}}) \le \frac{1}{2^t}$. Let

$$U_t^{\nu^{\xi,-x}} = \{\zeta \in (X-\{x\})^{\mathbb{N}} : \varphi_{h(t)}^{(1),\zeta\oplus\nu^{\xi,-x}}(0)\downarrow\},$$

where h is a total computable function. For each $\nu \in \{0,1\}^{\mathbb{N}}$, define

$$U_{t,s}^{\nu} = \{\zeta \in (X-\{x\})^{\mathbb{N}} : \varphi_{h(t)}^{(1),(\zeta\oplus\nu)[2s]}(0)\downarrow\},\ U_t^{\nu}[\frac{1}{2^t}] = \bigcup_{\mu_{p^{-x}}(U_{t,s}^{\nu})\le\frac{1}{2^t}} U_{t,s}^{\nu}.$$

For each $\tau \in \{0,1\}^{<\mathbb{N}}$, we define

$$U_t^{\tau} = \{\zeta \in (X-\{x\})^{\mathbb{N}} : \varphi_{h(t)}^{(1),\zeta[|\tau|]\oplus\tau}(0)\downarrow\}.$$

For each $t \in \mathbb{N}$, we define $V_t \subseteq X^{\mathbb{N}}$ as follows: $\xi' \in V_t$ if and only if there is some $s \in \mathbb{N}$ such that (notice that $(\xi')^{-x}[s]$ is defined if and only if $(\forall j < s)(\theta^{\nu^{\xi',-x}}(j)\downarrow)$)

(b1) $(\forall j < s)(\theta^{\nu^{\xi',-x}}(j)\downarrow)$;

(b2) $\varphi_{h(t)}^{(1),(\xi')^{-x}[s]\oplus\nu^{\xi',-x}[s]}(0)\downarrow$;

(b3) $\mu_{p^{-x}}(U_t^{\nu^{\xi',-x}[s]}) \le \frac{1}{2^t}$.

Using similar arguments as in (a), we can show that $\{V_t\}_{t=0}^{\infty}$ is uniformly Σ_1^0 and $\mu_p(V_t) \le \frac{1}{2^t}$.

Define the set Γ as in (a). We have seen that $\mu_p(\Gamma) = 1$. For each $t \in \mathbb{N}$, let $V_t^1 = \Gamma \cap V_t$ and let $V_t^0 = (X^{\mathbb{N}} - \Gamma) \cap V_t$. So $\mu_p(V_t) = \mu_p(V_t^0) + \mu_p(V_t^1) = \mu_p(V_t^1)$. Moreover, $\xi' \in V_t^1$ if and only if

$$\xi' \in \Gamma \wedge (\exists s)(\varphi_{h(t)}^{(1),(\xi')^{-x}[s]\oplus\nu^{\xi',-x}[s]}(0)\downarrow \wedge \mu_{p^{-x}}(U_t^{\nu^{\xi',-x}[s]}) \le \frac{1}{2^t})$$

if and only if $\theta^{\nu^{\xi',-x}}$ is total and $(\xi')^{-x} \in U_t^{\nu^{\xi',-x}}[\frac{1}{2^t}]$. Thus, if we define T to be such that $T(\zeta,\nu) = \mathbf{x} \oslash_{\nu} \zeta$, then, by Lemma 6.1 (a),

$\mu_p(V_t^1) = \mu_{p^{-x}\otimes(p_x,1-p_x)}(T^{-1}(V_t^1))$

$\leq \int_{\{0,1\}^{\mathbb{N}}} \int_{(X-\{x\})^{\mathbb{N}}} \chi_{U_t^{\nu}[\frac{1}{2^t}]}(\zeta) d\mu_{p^{-x}}(\zeta) d\mu_{(p_x,1-p_x)}(\nu)$

$\leq \int_{\{0,1\}^{\mathbb{N}}} \mu_{p^{-x}}(U_t^{\nu}[\frac{1}{2^t}]) d\mu_{(p_x,1-p_x)}(\nu) \leq \frac{1}{2^t}.$

Therefore, $\{V_t\}_{t=0}^{\infty}$ is an μ_p-test, and we have $\xi = \mathbf{x} \oslash_{\nu^{\xi,-x}} \xi^{-x} \in V_t$ for all $t \in \mathbb{N}$. Hence, ξ is not μ_p-random.

Now we prove the theorem for general A. Let $Y = A \cup \{y\}$, where $y \notin X$. Define $\Gamma : X^{\mathbb{N}} \to Y^{\mathbb{N}}$ as $\Gamma(\xi')_t = \xi'_t$ if $\xi'_t \in A$ and $\Gamma(\xi')_t = y$ otherwise. Γ is clearly computable. Let $q \in \Delta(Y)$ be such that $q_x = p_x$ for all $x \in A$ and $q_y = 1 - p_A$. We first show that $\Gamma(\xi) = \zeta$ is μ_q-random. Suppose not. Then $\zeta \in \bigcap_{t=0}^{\infty} U_t$ for some μ_q-test $\{U_t\}_{t=0}^{\infty}$. Define $V_t \subset X^{\mathbb{N}}$ as $V_t = \{\xi' \in X^{\mathbb{N}} : \Gamma(\xi') \in U_t\}$. We claim that for any open set $U \subset Y^{\mathbb{N}}$, $\mu_q(U) = \mu_p(\Gamma^{-1}(U))$. To see this, let $U = \bigcup_{\sigma \in G} N_\sigma$ for some prefix-free $G \subset Y^{<\mathbb{N}}$. Then $\Gamma^{-1}(U) = \bigcup_{\sigma \in G} \bigcup_{\tau \in H_\sigma} N_\tau$, where

$$H_\sigma = \{\tau \in X^{<\mathbb{N}} : |\sigma| = |\tau| \text{ and } \sigma_t = \tau_t \text{ if } \sigma_t \in A,\ \tau_t \in A \text{ otherwise}\}.$$

Notice that $H = \bigcap_{\sigma \in G} H_\sigma$ is also prefix-free. For each $\sigma \in G$,

$$\mu_q(N_\sigma) = \prod_{t=0}^{|\sigma|-1} q_{\sigma_t} = \prod_{\sigma_t \in A} p_{\sigma_t} \prod_{\sigma_t \notin A} (1 - p_A) = \sum_{\tau \in H_\sigma} \prod_{t=0}^{|\tau|-1} p_{\tau_t} = \sum_{\tau \in H_\sigma} \mu_p(N_\tau),$$

and thus

$$\mu_q(U) = \sum_{\sigma \in G} \mu_q(N_\sigma) = \sum_{\sigma \in G} \sum_{\tau \in H_\sigma} \mu_p(N_\tau) = \mu_p(\Gamma^{-1}(U)).$$

Since $V_t = \Gamma^{-1}(U_t)$, it follows that $\mu_p(V_t) = \mu_q(U_t) \leq \frac{1}{2^t}$ for all t. Hence, $\{V_t\}_{t=0}^{\infty}$ is a test. $\zeta \in \bigcap_{t=0}^{\infty} U_t$ implies that $\xi \in \bigcap_{t=0}^{\infty} V_t$ and so this is a contradiction to the fact that ξ is μ_p-random.

By construction, $\nu^{\zeta,-y} = \nu^{\xi,A}$ and $\xi^A = \zeta^{-y}$. As we have shown, $\nu^{\zeta,-y}$ is $\mu_{(q_y,1-q_y)} = \mu_{(1-p_A,p_A)}$-random relative to ζ^{-y}, and ζ^{-y} is $q^{-y} = \mu_{p^A}$-random. $\square$

6.2.4 Martingales and stochastic sequences

In this section we shall give a proof of Theorem 6.6. To this end, we need a new concept called martingales. A martingale is a formalization of a betting strategy. As we shall see, it is easy to show stochasticity using martingales instead of tests, and there is a characterization of Martin-Löf randomness using martingales. First we give the formal definition.

Definition 6.2. Let X be a finite set and let $\mathbf{p} = (p^0, p^1, ...)$ be a computable sequence of distributions over X such that $p_x^t > 0$ for all $x \in X$ and for all $t \in \mathbb{N}$. A function $M : X^{<\mathbb{N}} \to \mathbb{R}_+$ is a martingale with respect to $\mu_{\mathbf{p}}$ if for all $\sigma \in X^{<\mathbb{N}}$, $M(\sigma) = \sum_{x \in X} p_x^{|\sigma|} M(\sigma\langle x \rangle)$.

Let $\mathcal{P} \in \mathfrak{R}$. A martingale M is $\mathcal{P}$-effective if there is a sequence of martingales $\{M_t\}_{t=0}^{\infty}$ that satisfies the following properties:

(a) $M_t(\sigma) \in \mathbb{Q}_+$ for all $t \in \mathbb{N}$ and for all $\sigma \in X^{<\mathbb{N}}$;

(b) for each $t \in \mathbb{N}$, $M_t \in \mathcal{P}_T$;

(c) $\lim_{t\to\infty} M_t(\sigma) \uparrow M(\sigma)$ for all $\sigma \in X^{<\mathbb{N}}$.

In this case, we say that the sequence $\{M_t\}_{t=0}^{\infty}$ supports M.

The following theorem characterizes randomness in terms of martingales. The proof for the case $\mathcal{P} = \mathcal{P}^*$ can be found in Downey *et al.* [9], and the proof there can be easily relativized to cover the general case.

Theorem 6.5. *Let $\mathcal{P} \in \mathfrak{R}$ let $\mathbf{p}$ be a computable sequence such that $p_x^t > 0$ for all $x \in X$ and for all $t \in \mathbb{N}$. A sequence $\xi \in X^{\mathbb{N}}$ is $\mu_{\mathbf{p}}$-random relative to $\mathcal{P}$ if and only if for any $\mathcal{P}$-effective martingale M w.r.t. $\mu_{\mathbf{p}}$,*

$$\limsup_{T\to\infty} M(\xi[t]) < \infty.$$

We shall then discuss selection function and stochastic sequences. Let X be a finite set, and let $r : X^{<\mathbb{N}} \to \{0, 1\}$ be a selection function. Let $L_r^{\xi}(k) = |\{0 < t < k+1 : r(\xi[t-1]) = 1\}|$ to be the number of elements selected by r in $\xi[k]$. We have defined ξ^r before as the sequence obtained from ξ by applying r to it. With these notations, we now show that any $\mu_{\mathbf{p}}$-random sequence with $\lim_{t\to\infty} p^t = p$ is a p-stochastic sequence.

Theorem 6.6. *Let $\mathcal{P} \in \mathfrak{R}$. Let $u : X \to \mathbb{N}$ be a function. Suppose that ξ is $\mu_{\mathbf{p}}$-random relative to $\mathcal{P}$ with $p_x^t > 0$ for all $t \in \mathbb{N}$ and for all $x \in X$ and $\lim_{t\to\infty} p^t = p$, and suppose that r is a function in $\mathcal{P}_T$. If ξ^r is total, then*

$$\lim_{T\to\infty} \sum_{t=0}^{T-1} \frac{u(\xi_t^r)}{T} = \sum_{x\in X} p_x u(x).$$

Consequently, ξ is a p-stochastic sequence relative to $\mathcal{P}$.

Proof. By Theorem 6.5, for any $\mathcal{P}$-effective martingale M,

$$\limsup_{T\to\infty} M(\xi[t]) < \infty.$$

It is sufficient to show that for each $x \in X$,

$$lim_{T\to\infty} \sum_{t=0}^{T-1} \frac{\chi_x(\xi_t^r)}{T} = p_x,$$

where $\chi_x(y) = 1$ if $x = y$, $\chi_x(y) = 0$ otherwise.

Suppose that there exists some $\varepsilon > 0$ and a sequence $\{T_k\}_{k=0}^{\infty}$ such that for all $k \in \mathbb{N}$,

$$\sum_{t=0}^{T_k-1} \frac{\chi_x(\xi_t^r)}{T_k} \geq p_x + \varepsilon.$$

We shall define a martingale M as follows:

(a) $M(\epsilon) = 1$;

(b) $M(\sigma\langle x\rangle) = (1 + \kappa(1 - p_x^{|\sigma|}))M(\sigma)$ and $M(\sigma\langle y\rangle) = (1 - \kappa p_x^{|\sigma|})M(\sigma)$ for all $y \neq x$ if $r(\sigma) = 1$;

(c) $M(\sigma\langle y\rangle) = M(\sigma)$ for all $y \in X$ if $r(\sigma) = 0$.

To check that M is a martingale, note that if $r(\sigma) = 1$, then

$$\sum_{y\in X} p_y^{|\sigma|} M(\sigma\langle y\rangle) = p_x^{|\sigma|}(1 + \kappa(1 - p_x^{|\sigma|}))M(\sigma) + \sum_{y\neq x} p_y^{|\sigma|}(1 - \kappa p_x^{|\sigma|})M(\sigma)$$

$$= M(\sigma) + \kappa M(\sigma)(p_x^{|\sigma|}(1 - p_x^{|\sigma|}) - (1 - p_x^{|\sigma|})p_x^{|\sigma|}) = M(\sigma);$$

if $r(\sigma) = 0$, then $\sum_{y\in X} p_y^{|\sigma|} M(\sigma\langle y\rangle) = \sum_{y\in X} p_y^{|\sigma|} M(\sigma) = M(\sigma)$.

M is in $\mathcal{P}_T$ since r is. For $k \geq 1$, define

$$D_k = \{t \leq k-1 : r(\xi[t]) = 1, \xi_{t+1} = x\} \text{ and } E_k = \{t \leq k-1 : r(\xi[t]) = 1, \xi_{t+1} \neq x\}.$$

Then,

$$M(\xi[k]) = \prod_{t \in D_k} (1 + \kappa(1 - p_x^{t+1})) \prod_{t \in E_k} (1 - \kappa p_x^{t+1}).$$

Let $l_k = (L_r^{\xi})^{-1}(T_k)$. Since ξ^r is total, l_k is well defined for all $k \in \mathbb{N}$.

Let $\delta = \min\{p_x, 1 - p_x, \frac{\varepsilon}{2}\}$. Since $\lim_{t \to \infty} p^t = p$, let T be so large that $t \geq T$ implies that $|p_x^t - p_x| < \delta$. Let K be the first k such that $T_k > T$. Then, for all $k > K$,

$$M(\xi[l_k]) = \prod_{t \in D_{l_k}} (1 + \kappa(1 - p_x^{t+1})) \prod_{t \in E_{l_k}} (1 - \kappa p_x^{t+1})$$

$$\geq \prod_{t \in D_{l_K}} (1 + \kappa(1 - p_x^{t+1}) \prod_{t \in E_{l_K}} (1 - \kappa p_x^{t+1}))(1 + \kappa(1 - p_x - \delta))^{|D_{l_k} - D_{l_K}|}(1 - \kappa p_x - \kappa\delta)^{|E_{l_k} - E_{l_K}|}.$$

Let

$$A = \frac{\prod_{t \in D_{l_K}} (1 + \kappa(1 - p_x^{t+1}) \prod_{t \in E_{l_K}} (1 - \kappa p_x^{t+1}))}{(1 + \kappa(1 - p_x - \delta))^{L^1} (1 - \kappa p_x - \kappa\delta)^{L^2}},$$

where

$$L^1 = |D_{l_K}| \text{ and } L^2 = |E_{l_K}|.$$

Since for each k, $|D_{l_k}| \geq T_k p_x + T_k \varepsilon$,

$$M(\xi[l_k]) \geq A((1 + \kappa(1 - p_x - \delta))^{p_x + \varepsilon}(1 - \kappa p_x - \kappa\delta)^{1 - p_x - \varepsilon})^{T_k}.$$

Define

$$F(\kappa) = (1 + \kappa(1 - p_x - \delta))^{p_x + \varepsilon}(1 - \kappa p_x - \kappa\delta)^{1 - p_x - \varepsilon}.$$

We have $\ln F(0) = 1$ and

$$(\ln F)'(0) = (p_x + \varepsilon)(1 - p_x - \delta) - (1 - p_x - \varepsilon)(p_x + \delta) = \varepsilon - \delta > 0.$$

Thus, for κ small enough, $F(\kappa) > 1$, and so

$$\limsup_{T \to \infty} M(\xi[T]) = \infty,$$

a contradiction. $\square$

6.3 Proofs

Proof of Proposition 3.1: By Theorem 6.1, we know that, for any $\xi, \zeta \in \{0,1\}^{\mathbb{N}}$, if ξ is $\mu_{(\frac{1}{2},\frac{1}{2})}$-random relative to $C(\zeta)$, then ξ is complex relative to $C(\zeta)$. Moreover, we know that, by Theorem 6.3, if $\xi \otimes \zeta$ is λ^4-random, then ξ is $\mu_{(\frac{1}{2},\frac{1}{2})}$-random relative to $C(\zeta)$ and vice versa. Hence, $C(\xi)$ and $C(\zeta)$ are mutually complex. Now, by Proposition 3.2, the set

$$A = \{\xi \otimes \zeta : \xi \otimes \zeta \text{ is } \lambda^4\text{-random}\} \subset \{\xi \otimes \zeta : C(\xi) \text{ and } C(\zeta) \text{ are mutually complex}\}$$

has measure 1. Therefore, the set A is uncountable. Since for any ξ, the set of sequences ξ' such that $C(\xi) = C(\xi')$ is countable, we can then conclude that there are uncountably many different pairs of computability constraints that are mutually complex. □

References

[1] Ambos-Spies, K., E. Mayordomo, Y. Wang, and X. Zheng. (1996). " Resourcebounded dense genericity, stochasticity and weak randomness," *Proceedings of the Thirteenth Symposium on Theoretical Aspects of Computer Science (STACS96).* Springer-Verlag Lecture Notes in Computer Science 1046, 63-74.

[2] Aumman, R. J. and A. Brandenburger. (1995). "Epistemic Conditions for Nash Equilibrium," *Econometrica* vol. 63, pp. 1161-1180.

[3] Ben-Porath, E. (1993), "Repeated games with finite automata," *Journal of Economic Theory,* 59, 17-32.

[4] Brown, J. and Rosenthal, R. (1990). "Testing the Minimax Hypothesis: A Reexamination of ONeills Experiment," *Econometrica.* vol. 58, 1065-1081.

[5] Chaitin, G. J. (1975). "A theory of program size formally identical to information theory." *Journal of the ACM.* vol. 22 pp. 329-340.

[6] Calude, C. S. (2002). *Information and randomness: an algorithmic perspective.* 2nd edition, Springer.

[7] Camerer, C. (2003). *Behavioral game theory: experiments in strategic interaction.* Princeton University Press.

[8] Chiappori, P. A., Levitt, S. D. and Groseclose, T. (2002). "Testing Mixed Strategy Equilibrium When Players Are Heterogeneous: The Case of Penalty Kicks," *American Economic Review.* vol. 92(4), 1138-1151.

[9] Downey, R., D. Hirschfeldt, A. Nies, and S. Terwijn. (2006). "Calibrating randomness," *Bulletin Symbolic Logic.* volume 12, issue 3, 411-491.

[10] Harsanyi, J. C. (1973). "Games with randomly disturbed payoffs: a new rationale for mixed-strategy equilibrium points." *Int. J. Game Theory* vol. 2, pp. 123.

[11] Hu, Tai-Wei. (2008). "Expected utility theory from the frequentist perspective." *working paper.*

[12] Kakutani, S. (1948). "On equivalence of infinite product measures." *Ann. of Math.* vol. 49:214224.

[13] Kolmogorov, A. N. (1965). "Three Approaches to the Quantitative Definition of Information," in *Problems of Information Transmission (Problemy Peredachi Informatsii).* Vol. 1, p. 1-7.

[14] van Lambalgen, M. (1987) *Random Sequences.* Ph. D,. Diss. University of Amsterdam.

[15] van Lambalgen, M. (1990). "The axiomatization of randomness." *Journal of Symbolic Logic,* vol. 55, no. 3, 1143-1167.

[16] Levin, L. A. (1973). "On the notion of a random sequence." *Soviet Mathematics - Doklady.* vol. 14, pp. 1413-1416.

[17] Li, Ming and Paul Vitányi, (1997) *An introduction to Kolmogorov complexity and its applications.* Springer.

[18] Luce, R. Duncan and Howard Raiffa. (1957). *Games and Decisions: Introduction and Critical Survey.* Dover.

[19] Martin-Löf, P. (1966). "The definition of random sequences." *Information and Control*, vol. 9, 602-619.

[20] von Mises, Richard. (1939). *"Probability, Statistics, and Truth."*

[21] Muchnik, An. A., A. Semenov, and V. Uspensky. (1998). "Mathematical metaphysics of randomness." *Theoretical Computer Science*, Vol 207(2), 263-317.

[22] von Neumann, J. (1928). "Zur theorie der gesellschaftsspiele." *Math. Annalen.* v. 100, 295-320

[23] von Neumann, J. and O. Morgenstern (1944), *"Theory of Games and Economic Behavior."* Princeton University.

[24] Odifreddi, P. G. (1989). *Classical recursion theory.* (Vol. I). North-Holland Publishing Co., Amsterdam.

[25] O'Neill, B. (1987). "A Non-metric Test of the Minimax Theory of Two-person Zero-sum Games," *Proceedings of the National Academy of Science.* 84, 2106-2109.

[26] O'Neill, B. (1991). "Comments on Brown and Rosenthal's Reexamination." *Econometrica.* vol. 59, no. 2, p. 503-507.

[27] Osborne, M. J. and A. Rubinstein. (1994), *A Course in Game Theory*, MIT Express.

[28] Pippenger, Nicholas. (1997). *Theories of computability.* Cambridge University Express, Cambrigde, United Kingdom.

[29] Palacios-Huerta, I. (2003). "Professionals Play Minimax." *Review of Economic Studies.* vol. 70 (2), 395-415.

[30] Palacios-Huerta, I. and O. Volij (2008). "Experientia Docet: Professionals Play Minimax in Laboratory Experiments." *Econometrica.* vol. 76(1), 71-115.

[31] Rubinstein, A. (1991). "Comments on the interpretation of game theory." *Econometrica*, Vol. 59, No. 4. 909-924.

[32] Ville, J., (1939). *Ĕtude critique de la concept du collectif.* Gauthier-Villars.

[33] Vovk, V. G. (1987). "On a criterion for randomness." *Dokl. Akad. Nauk SSSR,* vol. 294(6). p.12981302.

[34] Wald, A. (1938). "Die widerspruchsfreiheit des kollektivsbegriffes." in *Wald: selected papers in Statistics and Probability.* MacGraw-Hill, 1955, 25-41.

[35] Wagner, E. G. (1969). "Uniformly reflexive structures: on the nature of Gödelizations and relative computability." *Trans. Amer. Math. Soc.* vol. 144. p. 1-41.

[36] Walker, M. and Wooders, J. (2001). "Minimax Play at Wimbledon." *American Economic Review.* vol. 91, 1521-1538.

[37] Zvonkin, A. K. and L. A. Levin. (1970). "The complexity of finite objects and the basing of the concepts of information and randomness on the theory of algorithms." *Uspehi Mat. Nauk.* vol. 25(6(156)). p. 85127.