

## IDENTIFYING THE DEVIATOR

BY NOGA ALON<sup>1,a</sup>, BENJAMIN GUNBY<sup>2,c</sup>, XIAOYU HE<sup>1,b</sup>, ERAN SHMAYA<sup>3,d</sup> AND  
 EILON SOLAN<sup>4,e</sup>

<sup>1</sup>Department of Mathematics, Princeton University, <sup>a</sup>[anlon@math.princeton.edu](mailto:anlon@math.princeton.edu), <sup>b</sup>[xiaoyuh@princeton.edu](mailto:xiaoyuh@princeton.edu)

<sup>2</sup>Department of Mathematics, Rutgers University, <sup>c</sup>[bg570@connect.rutgers.edu](mailto:bg570@connect.rutgers.edu)

<sup>3</sup>Department of Economics, Stony Brook University, <sup>d</sup>[eran.shmaya@stonybrook.edu](mailto:eran.shmaya@stonybrook.edu)

<sup>4</sup>The School of Mathematical Sciences, Tel Aviv University, <sup>e</sup>[eilons@tauex.tau.ac.il](mailto:eilons@tauex.tau.ac.il)

A group of players are supposed to follow a prescribed profile of strategies. If they follow this profile, they will reach a given target. We show that if the target is not reached because some player deviates, then an outside observer can identify the deviator. We also construct identification methods in two nontrivial cases.

**1. Introduction.** Alice and Bob alternately report outcomes (heads or tails), which each of them is supposed to generate by tossing a fair coin. If both of them follow through, then the realized sequence of outcomes is random and with probability 1 will pass known statistical tests. Suppose that the sequence of outcomes does not pass a given test: for example, the long-run frequency of heads does not converge to  $1/2$ . Can an outside observer who observes only the sequence of outcomes identify who among Alice and Bob did not generate the outcomes by tossing a fair coin? The answer is positive: if the outcome sequence fails the test, it is easy to identify who is responsible by checking separately the long-run frequency of heads in the sequences produced by each player.

Consider now a different test: Alice's and Bob's outcomes control a one-dimensional random walk that moves to the right when a head is reported, and to the left when a tail is reported. Alice controls the odd periods and Bob controls the even periods, and they pass the test if the realized walk crosses the origin infinitely often. We assume that the coin flips are generated sequentially, and that each player observes the previous flips of the other player before announcing her or his next outcome. A version of this test that is also interesting is the one in which Alice and Bob pass the test if the walk visits the origin at least once after the initial step.

Here, too, if Alice and Bob generate the outcomes by tossing fair coins, the realized sequence passes the test almost surely. Suppose this does not happen. Observing only the reported outcomes, can one identify who among the two is responsible for the test's failure? The reader may want to stop at this point and think whether this is possible.

We study a more general form of this question, in which each of several players is supposed to generate outcomes according to some probabilistic rule in every period. We are also given a *target set*, which is a set of infinite sequences of outcomes, and assume that if the players follow their probabilistic rules, then with probability 1 the generated sequence is within the set. A *blame function* is a function from the complement of the target set to the set of players: if the generated sequence happens to be outside the target set, the blame function identifies a player who is proclaimed the deviator. In our opening example, the target set is the set of realizations with long-run frequency  $1/2$  of heads, and in the second example the target set is the set of all realizations where the induced random walk crosses the origin infinitely often.

Received July 2022; revised September 2023.

*MSC2020 subject classifications.* Primary 91A35; secondary 62C05, 62F07.

*Key words and phrases.* multiplayer games, detecting deviations, random walk, zero-sum games.

Given the players' probabilistic rules and the target set, we seek a blame function with the property that, if only one player deviates from her prescribed rule, then the probability that the realization is outside the target set and an innocent player is blamed is small.

Our motivation comes from game theory. The most studied solution concept in game theory is the Nash equilibrium ([14]), which is a profile of strategies, one for each player, that is immune to *unilateral deviations*, that is, such that no player can profit by deviating from her strategy assuming her opponents stick to their equilibrium strategies. A common way to construct a Nash equilibrium in dynamic games is to find a profile of strategies such that if played together these strategies yield a high payoff to all players, and punish a player who is caught deviating; see, for example, [4]. To implement such a construction, the players must be able to identify the deviator once a deviation is detected, assuming there is only one deviator. The goal of our paper is to study the extent to which a deviator can be correctly identified.

Despite its application to game theory, it appears that in its full generality the question has not been addressed before in the literature. The reason is perhaps that most papers about Nash equilibrium in dynamic games either study two-player games, where identifying the deviator does not play any role (because if a player did not deviate, he knows the other must have deviated), or environments in which the detection of the deviator is easily done. The latter include (a) environments where the profile of strategies the players should follow is deterministic, so that a deviation is immediately detected, as in [4], and (b) environments where the payoff functions are additive over periods (such as stochastic games with the long-run average payoff), and the profile of strategies the players should follow has a stationary flavor. In such an environment, the underlying statistical tests are based on long-run frequency of actions, as in our opening paragraph, and detecting deviations is done using a law of large numbers; see, for example, [16, 17].

Dynamic games where the payoff functions are arbitrary functions of the players' actions have been studied in the past only in the context of two-player zero-sum games; see, for example, [12, 13]. In the last few years, researchers turned to study Nash equilibria in multiplayer dynamic games with general payoff functions, which require detecting deviations from more complicated strategies; see, for example, [2, 3]. Our result turns out to be useful in this area. Indeed, it can be used to provide an alternative proof for the existence of an  $\varepsilon$ -equilibrium in repeated games with tail-measurable payoffs (see [9]), and to prove that in every multiplayer stochastic game with finite state and action spaces and with bounded, Borel measurable payoff functions, for every  $\varepsilon > 0$  there is a subgame in which an  $\varepsilon$ -equilibrium exists; see [10].

The paper is also related to statistical decision theory. Recall that in a statistical decision problem (see, e.g., [1], Chapter 7), a statistician observes a realization from a distribution that depends on an unknown parameter, and then makes a decision. The statistician's loss is a function of the unknown parameter and her decision. In our problem, a realization is a finite or infinite sequence of outcomes, the parameter space is all distributions induced by possible deviations of a single player, the statistician's decision is a player to blame, and the statistician loses if an innocent player is blamed. Our blame function is, in the terminology of statistical decision theory, a decision rule. The twist is that in our problem the statistician makes a decision only if the realization is outside the target set. We show that in our environment the statistician has a decision rule with zero risk (or with a small risk, in case the realization is finite yet sufficiently long).

The paper is structured as follows. In Section 2 we formally describe the model, the concept of a blame function, and the main results. In Section 3 we construct blame functions for two nontrivial examples, including the one-dimensional random walk. In Section 4 we present a nonconstructive proof for the general case, and we conclude in Section 5.

**2. Model and main results.** Throughout the paper, we fix a finite set  $I$  of players, and for each player  $i \in I$  we fix a finite set of actions  $A_i$ . Denote by  $A = \prod_{i \in I} A_i$  the set of action profiles.

The set of *finite realizations* is the set  $A^{<\mathbb{N}}$  of finite sequences of action profiles. A *pure strategy* for player  $i$  is a function  $\sigma_i : A^{<\mathbb{N}} \rightarrow A_i$ , and a *behavior strategy* for player  $i$  is a function  $\sigma_i : A^{<\mathbb{N}} \rightarrow \Delta(A_i)$ , where  $\Delta(A_i)$  is the set of probability distributions over  $A_i$ . Denote by  $z^n = (z_i^n)_{i \in I}$  the action profile selected by the players in period  $n$ , and by  $\sigma_i(z_i^n | z^1, z^2, \dots, z^{n-1})$  the probability that  $\sigma_i$  selects the action  $z_i^n$  in period  $n$ , provided the action profiles selected in the first  $n - 1$  periods are  $z^1, z^2, \dots, z^{n-1}$ . Denote by  $\Sigma_i$  the set of behavior strategies of player  $i$ .

We endow the space  $A^{\mathbb{N}}$  of *realizations* with the product topology and the induced Borel  $\sigma$ -algebra. Every behavior strategy profile  $\sigma = (\sigma_i)_{i \in I} \in \prod_i \Sigma_i$  induces a probability distribution  $\mathbf{P}_\sigma$  over realizations. Abusing notation, for every finite realization  $z \in A^{<\mathbb{N}}$  we denote by  $\mathbf{P}_\sigma(z)$  the probability that the sequence  $z$  will be generated under  $\sigma$ . Given a strategy profile  $\sigma = (\sigma_i)_{i \in I}$ , denote by  $\sigma_{-i} = (\sigma_j)_{j \neq i}$  the strategies of all players except player  $i$ .

Even though we use game theoretic terminology (players, actions, strategies), we emphasize that we do not define a game between the players, as there are no payoff functions.

### 2.1. Testability.

**DEFINITION 2.1 (Goal).** A *goal* is a pair  $(\sigma^*, D)$  where  $\sigma^* \in \prod_i \Sigma_i$  is a behavior strategy profile and  $D \subseteq A^{\mathbb{N}}$  is a Borel set of realizations, which is termed the *target set*.

The strategy profile  $\sigma^*$  is a prescribed way for the players to play. The target set  $D$  is a set of realizations that they are supposed to reach if they follow through their prescribed strategy. We are interested in cases in which the probability  $\mathbf{P}_{\sigma^*}(D)$  that the prescribed strategy profile attains the target set is 1 or close to 1.

**REMARK 2.2 (Alternate play).** In our model, players make their choices simultaneously. Yet, the model can accommodate alternate play. Indeed, if  $I = \{0, 1, \dots, |I| - 1\}$ , and if for each player  $i$ , the strategy  $\sigma_i^*$  randomizes only in periods  $n$  such that  $n \bmod |I| = i$ , then in effect the players play alternately.

**DEFINITION 2.3 (Blame function).** A *blame function* is a Borel function  $f : D^c \rightarrow I$ .

The interpretation of a blame function is that if the players generate a realization  $s \in A^{\mathbb{N}}$  that misses the target set, player  $f(s)$  is blamed as the player who did not follow her part of  $\sigma^*$ .

**DEFINITION 2.4 ( $\delta$ -testability).** Fix  $\delta \geq 0$ . The goal  $(\sigma^*, D)$  is  $\delta$ -*testable* if there exists a blame function  $f$  such that for every player  $i \in I$  and every strategy  $\sigma_i$  for player  $i$  we have  $\mathbf{P}_{\sigma_i, \sigma_{-i}^*}(D^c \text{ and } \{f \neq i\}) \leq \delta$ .

The interpretation of testability is that if some player  $i$  deviates, then the probability that a different player  $j$  is blamed is at most  $\delta$ . Thus, with probability  $\mathbf{P}_{\sigma_i, \sigma_{-i}^*}(D)$  no player is blamed, and with probability at least  $\mathbf{P}_{\sigma_i, \sigma_{-i}^*}(D^c) - \delta$  the blame function identifies the player who deviated from  $\sigma^*$ . We note that if  $(\sigma^*, D)$  is  $\delta$ -testable then

$$\begin{aligned} \mathbf{P}_{\sigma^*}(D) &= 1 - \mathbf{P}_{\sigma^*}(D^c) = 1 - \sum_{i \in I} \mathbf{P}_{\sigma^*}(D^c \text{ and } \{f = i\}) \\ &= 1 - \frac{1}{|I| - 1} \sum_{i \in I} \mathbf{P}_{\sigma^*}(D^c \text{ and } \{f \neq i\}) \geq 1 - \frac{|I|}{|I| - 1} \delta, \end{aligned}$$

where the inequality follows from the fact that, by Definition 2.4 with  $\sigma_i = \sigma_i^*$ , we have  $\mathbf{P}_{\sigma^*}(D^c \text{ and } \{f \neq i\}) \leq \delta$  for every  $i$ .

**REMARK 2.5** (Deviations by more than one player). Our definition of testability involves deviations of a single player. If more than one player deviates, then identifying all the deviators may not be possible. For simplicity, consider the following example, where each of Alice and Bob have two actions, denoted  $a$  and  $b$ , and the players play for a single period: Alice's strategy  $\sigma_A^*$  selects each action with probability  $\frac{1}{2}$ , while Bob's strategy  $\sigma_B^*$  selects  $a$  with probability 1. Let the target set  $D$  be the set that contains the two realizations in which Bob selects  $a$ . While detecting a deviation of Bob is trivial, detecting a deviation of Alice is not possible. Thus, if both Alice and Bob deviate, we will not be able to identify Alice's deviation.

**2.2. Main results.** Our main results are the following.

**THEOREM 2.6.** *Every goal  $(\sigma^*, D)$  is  $2\sqrt{(|I| - 1)\varepsilon}$ -testable, as soon as  $\mathbf{P}_{\sigma^*}(D) > 1 - \varepsilon$ .*

**REMARK 2.7** (Tightness of the bound in Theorem 2.6). The bound in Theorem 2.6 is tight up to a constant factor; see Section 3.1.

**REMARK 2.8** (Random blame function). A *random* blame function is a function  $\varphi : D^c \rightarrow \Delta(I)$ , with the interpretation that if the realization  $s$  is not in  $D$ , then each player  $i$  is blamed with probability  $\varphi_i(s)$ . If in the definition of  $\delta$ -testability we allow for random blame functions, then we could get rid of the constant 2 in Theorem 2.6, as can be seen in the proof of the theorem.

As the next result states, in the case that  $\mathbf{P}_{\sigma^*}(D) = 1$ , the goal  $(\sigma^*, D)$  is in fact 0-testable.

**THEOREM 2.9.** *Every goal  $(\sigma^*, D)$  such that  $\mathbf{P}_{\sigma^*}(D) = 1$  is 0-testable.*

As we now observe, Theorem 2.9 is a consequence of Theorem 2.6.

**PROOF OF THEOREM 2.9 USING THEOREM 2.6.** Fix a goal  $(D, \sigma^*)$  with  $\mathbf{P}_{\sigma^*}(D) = 1$ . Let  $(\delta_k)_{k=1}^\infty$  be a sequence of positive reals such that  $\sum_{k=1}^\infty \delta_k < \infty$ . By Theorem 2.6, the goal  $(D, \sigma^*)$  is  $\delta_k$ -testable for every  $k \in \mathbb{N}$ . Let  $f_k : D^c \rightarrow I$  be a blame function such that  $\mathbf{P}_{\sigma_i, \sigma_{-i}^*}(f_k \neq i) < \delta_k$ , for every player  $i \in I$  and every strategy  $\sigma_i \in \Sigma_i^B$ . By the Borel–Cantelli lemma,  $\mathbf{P}_{\sigma_i, \sigma_{-i}^*}(f_k \neq i \text{ for infinitely many } k) = 0$ .

Let  $f : D^c \rightarrow I$  be such that  $f(s) = i$  only if  $f_k(s) = i$  for infinitely many  $k$ 's. Then  $f$  satisfies that  $\mathbf{P}_{\sigma_i, \sigma_{-i}^*}(f \neq i) = 0$ , for every player  $i \in I$  and every strategy  $\sigma_i \in \Sigma_i^B$ , as desired.  $\square$

**REMARK 2.10** (Finite horizon). Though our main results, Theorems 2.6 and 2.9, are phrased for the problem with infinite horizon, both apply in particular to the case of finite horizon. That is, suppose there is  $n \in \mathbb{N}$  such that  $D \subseteq A^{\mathbb{N}}$  depends only on the first  $n$  action profiles: two realizations  $s$  and  $s'$  coincide in their first  $n$  coordinates, then either both are in  $D$  or both are outside  $D$ .

According to, for example, Theorem 2.6, if  $\mathbf{P}_{\sigma^*}(D) > 1 - \varepsilon$ , and if the choices of the players in the first  $n$  stages  $(z^1, z^2, \dots, z^n)$  satisfy that all extensions of this sequence are not in  $D$ , then after stage  $n$ , with high probability, one can correctly determine which player deviated from  $\sigma^*$ .

REMARK 2.11 (Construction of blame functions for specific goals). Theorems 2.6 and 2.9 are theoretical results that assert the existence of a blame function that identifies the deviator. A natural question is whether there is an efficient procedure that constructs such a blame function for every given problem. Unfortunately, we do not have such a procedure. In Section 3 we analyze two specific goals and illustrate the complexity of the required blame functions. Since our proof of Theorem 2.6 uses the minmax theorem, when considering a set up with finite horizon, as in Remark 2.10, a blame function that correctly identifies the deviator can be calculated using a linear program, yet the size of the linear program is double exponential in the length of the problem.

**3. Examples.** In this section we present three examples; the first shows the tightness of Theorem 2.6, and in the other two we describe explicit blame functions. The second example is rather simple, while the third is more sophisticated and is analyzed both in an infinite-horizon and in a finite-horizon framework. In all examples, there are two players, denoted A and B. To simplify notation, in the second and third examples we will assume that Player A is active *only* in odd periods, and Player B is active *only* in even periods.

3.1. *Tightness of Theorem 2.6.* We here consider a problem with finite horizon (of length 1). Each of the two players chooses a single random bit, where he is supposed to choose 1 with probability  $\mu$  and 0 with probability  $1 - \mu$ . The target set is  $D = \{00, 01, 10\}$ . Here the probability of  $D^c = \{11\}$  is  $\mu^2$  and each player can deviate by choosing 1, ensuring that if the other player is honest, then the resulting pair of choices lies in  $D^c$  with probability  $\mu$ . In this scenario, any blame function must err with probability at least  $\mu$  under some deviation. Indeed, suppose that the blame function  $f$  satisfies  $f(11) = A$ . Then when Player B deviates and selects 1 with probability 1, with probability  $\mu$  Player A is erroneously blamed.

3.2. *Adjacent ones.* Two players generate an infinite sequence in  $\{0, 1\}^{\mathbb{N}}$ , where Player A chooses the odd bits and Player B chooses the even ones. In period  $n$ , the player whose turn it is to choose the bit is supposed to choose 1 with probability  $\mu/n$ , and 0 with probability  $1 - \mu/n$ , where  $\mu > 0$  is a small real. Formally, the strategy pair  $\sigma^* = (\sigma_A^*, \sigma_B^*)$  is given by  $\sigma_A^*(1|z) = \mu/n$  for every  $z \in \{0, 1\}^{n-1}$  such that  $n$  is odd, and  $\sigma_B^*(1|z) = \mu/n$  for every  $z \in \{0, 1\}^{n-1}$  such that  $n$  is even.

Call a sequence in  $\{0, 1\}^{\mathbb{N}}$  *bad* if there is an odd number  $n \geq 1$  such that the bits in both periods  $n$  and  $n + 1$  are 1. That is, Player B chooses 1 exactly one period after Player A chooses 1. A sequence is *good* if it is not bad, and the target set  $D$  is the set of all good infinite sequences. The probability of falling into the target set under the process above is  $1 - \varepsilon$ , where  $\varepsilon$  is given by

$$\varepsilon = \sum_{k=0}^{\infty} \left( \prod_{0 \leq i < k} \left( 1 - \frac{\mu^2}{(2i+1)(2i+2)} \right) \right) \cdot \frac{\mu^2}{(2k+1)(2k+2)} < \mu^2.$$

By Theorem 2.6, the goal  $(\sigma^*, D)$  described above is  $O(\mu)$ -testable. Intuitively, if Player A selected 1 too often, she is proclaimed the deviator, and otherwise it is Player B. Formally, let  $s = (s_1, s_2, s_3, \dots)$  be a sequence in  $D^c$ , that is, a sequence containing two consecutive ones in positions  $2k + 1, 2k + 2$  for some  $k \geq 0$ . If

$$\sum_{k \geq 0, s_{2k+1}=1} \frac{\mu}{2k+2} > \mu,$$

then  $f(s) = A$ . Otherwise,  $f(s) = B$ .

Suppose Player A is honest. If so, then the expectation of the sum

$$\sum_{k \geq 0, s_{2k+1}=1} \frac{\mu}{2k+2}$$

is

$$\sum_{k \geq 0} \frac{\mu}{2k+1} \cdot \frac{\mu}{2k+2} < \mu^2.$$

Therefore, by Markov's inequality, the probability this sum exceeds  $\mu$  is smaller than  $\mu$ , showing that in this case the probability that the blame function blames Player A is less than  $\mu$ . This conclusion holds for every strategy of Player B, as this step is independent of the bits selected by B.

Suppose next that Player B is honest. We show that no strategy of Player A causes Player B to be blamed with probability exceeding  $\mu$  while keeping the condition

$$\sum_{k \geq 0, s_{2k+1}=1} \frac{\mu}{2k+2} \leq \mu.$$

To do so, for every  $k \in \mathbb{N}$  define a random variable  $\delta_k$  by  $\delta_k = \frac{\mu}{2k+2}$  when  $s_{2k+1} = 1$ , and  $\delta_k = 0$  when  $s_{2k+1} = 0$ . Let  $y$  be a uniform random number in  $[0, 1]$  that is selected at the outset, and suppose Player B is blamed in step  $k$  if and only if  $y$  lies in  $(\sum_{i < k} \delta_i, \sum_{i=1}^k \delta_i]$ . Since Player B is honest, for every given history in which Player B has not been already blamed, the conditional probability Player B is blamed in step  $k$  is  $\frac{\delta_k}{1 - \sum_{i < k} \delta_i} \geq \delta_k$ . As a result, the probability that Player B is blamed under this construction,  $\sum_{k=1}^{\infty} \delta_k$ , is at least the probability she is blamed in the original problem. This sum is at most  $\mu$ , implying the desired result.

This shows that the goal considered here is  $\mu$ -testable by the explicit blame function  $f$  described above.

**3.3. Avoiding the origin in a random walk.** In this subsection we return to the random walk example in the [Introduction](#). Two players generate an infinite sequence in  $\{-1, 1\}^{\mathbb{N}}$ , thought of as the moves in an infinite walk on  $\mathbb{Z}$  that originates at 1, where Player A chooses the odd terms and Player B chooses the even terms. In each period, the player is supposed to choose 1 with probability  $1/2$ . We start by studying the problem with infinite horizon, and then discuss the problem with finite horizon.

Suppose that the target set  $D$  is the set of all infinite random walks  $w \in \{-1, 1\}^{\mathbb{N}}$  such that there exists  $n$  for which  $\sum_{i \leq n} w_i = -1$ , that is, the walk reaches  $-1$  at least once. The random walk is recurrent with probability 1, so  $\mathbf{P}_{\sigma^*}(D) = 1$ .

Let  $a_1, a_2, \dots \in \{-1, 1\}$  and  $b_1, b_2, \dots \in \{-1, 1\}$  be the moves that player A plays and player B plays, in order, that is,  $w = (a_1, b_1, a_2, b_2, \dots)$ . Let  $A_n = \sum_{i=1}^n a_i$  and  $B_n = \sum_{i=1}^n b_i$ . Let  $s_n$  be the position of the random walk at time  $n$ , so that  $s_0 = 1$ , and for  $n \geq 1$ ,  $s_{2n} = 1 + A_n + B_n$  and  $s_{2n-1} = 1 + A_n + B_{n-1}$ . Call the player that plays randomly Honest, and the other player Deviator. We claim that the following blame function  $f : D^c \rightarrow \{A, B\}$  allows the statistician to detect Deviator with probability 1. Given  $w \in D^c$ , the following steps are performed in order to determine the value of  $f$ .

1. If  $\limsup_{n \rightarrow \infty} (\frac{A_n}{\sqrt{n \log \log n}}) > 0$ , choose player A. Otherwise, if  $\limsup_{n \rightarrow \infty} (\frac{B_n}{\sqrt{n \log \log n}}) > 0$ , choose player B.
2. If  $\sum_{n=2}^{\infty} \frac{a_n}{\sqrt{n}(\log n)^{3/4}}$  diverges, choose player A. Otherwise, if  $\sum_{n=2}^{\infty} \frac{b_n}{\sqrt{n}(\log n)^{3/4}}$  diverges, choose player B.

3. If  $\sum_{n=1}^{\infty} \frac{s_{2n+1}^2 - s_{2n}^2}{2n \log(2n)} = \infty$ , choose player B. Otherwise, if  $\sum_{n=2}^{\infty} \frac{s_{2n}^2 - s_{2n-1}^2}{(2n-1) \log(2n-1)} = \infty$ , choose player A.
4. Otherwise, choose player A.

**THEOREM 3.1.** *The blame function  $f$  above correctly identifies the Deviator with probability 1, regardless of Deviator's strategy.*

**REMARK 3.2** (Visiting the origin infinitely often). In the second example in the [Introduction](#), the target set  $D$  contains all realizations that visit the origin infinitely often. The algorithm above can be adapted to this case. Indeed, supposing that the realization  $s$  is such that the origin is visited finitely many times, then applying the above algorithm to the suffix of the realization after the last visit to the origin identifies Deviator with probability 1.

The idea of the proof is that since Deviator must move to the right during periods where Honest moves substantially to the left (to avoid going below zero), Deviator must thus move to the left when Honest moves to the right to avoid being clearly right-biased (Steps 1 and 2 detect right-biased behavior). Thus Deviator must keep the walk fairly close to 0. Since  $s_n^2$  increases by 1 in expectation on Honest's moves, it should decrease on Deviator's moves to keep the walk close to 0, and this discrepancy is what is detected in Step 3.

The success of the algorithm can be deduced from the following lemmas.

**LEMMA 3.3.** *Honest is chosen on Step 1 with probability 0.*

**LEMMA 3.4.** *Honest is chosen on Step 2 with probability 0.*

**LEMMA 3.5.** *If no player is chosen on Steps 1 and 2, and  $s_n$  is nonnegative for all  $n$ , then  $\sum_{n=2}^{\infty} \frac{s_n^2}{n^2 \log n}$  converges.*

**LEMMA 3.6.** *If A is Honest, then the probability that no player is chosen on Step 1 and  $\sum_{n=1}^{\infty} \frac{s_{2n+1}^2 - s_{2n}^2}{2n \log(2n)}$  does not diverge to  $\infty$  is 0, regardless of Deviator's strategy.*

*Similarly, if B is Honest, the probability that no player is chosen on Step 1 and  $\sum_{n=2}^{\infty} \frac{s_{2n}^2 - s_{2n-1}^2}{(2n-1) \log(2n-1)}$  does not diverge to  $\infty$  is 0, regardless of Deviator's strategy.*

Lemma 3.5 is the only place where the fact that the walk must be nonnegative is used. Indeed, in a purely random walk, the analogous statements to the other three lemmas are true, but since  $s_n^2$  is generally of order  $n$  in a purely random walk,  $\sum s_n^2 / (n^2 \log n)$  would be on the order of the divergent sum  $\sum 1 / (n \log n)$ .

We now deduce Theorem 3.1 from these lemmas.

**PROOF OF THEOREM 3.1.** We show that the probability that the algorithm chooses the wrong player at any given step is 0.

First, Lemma 3.3 and Lemma 3.4 show that the algorithm chooses the wrong player on Steps 1 and 2 with probability 0.

By Lemma 3.6, the probability that no player is chosen in Steps 1, 2, and 3 combined is 0. Thus the algorithm chooses the wrong player on Step 4 with probability 0.

It remains to bound the probability that the algorithm chooses the wrong player on Step 3. By Lemma 3.6, if A is Honest, then the algorithm fails on Step 3 with probability 0.



Suppose B is Honest. Then by Lemma 3.5, if we reach Step 3, then  $\sum_{n=2}^{\infty} \frac{s_n^2}{n^2 \log n}$  converges. Notice that  $\frac{1}{x \log x}$  has derivative  $(-1 + o(1)) \frac{1}{x^2 \log x}$ , so

$$\frac{1}{(n-1) \log(n-1)} - \frac{1}{n \log n} = (1 + o(1)) \frac{1}{n^2 \log n}.$$

Since the series  $\sum s_n^2 / (n^2 \log n)$  converges, we may substitute (dropping the  $n = 2$  term) to obtain that

$$(1) \quad \sum_{n=3}^{\infty} \left( \frac{1}{(n-1) \log(n-1)} - \frac{1}{n \log n} \right) s_n^2$$

converges. Regrouping terms, we obtain that

$$(2) \quad \sum_{n=2}^{\infty} \frac{s_{n+1}^2 - s_n^2}{n \log n}$$

converges, as its partial sums differ from those of expression (1) by  $\frac{s_2^2}{2 \log 2} - \frac{s_{n+1}^2}{(n+1) \log(n+1)}$ , which is bounded because  $s_n = O(\sqrt{n} \log \log n)$  (otherwise a player would have been chosen on Step 1).

Almost surely either a player was chosen on Step 1 or 2 or the sum of just the odd-numbered terms (given by B's moves) of expression (2) diverges to  $\infty$ , by Lemma 3.6. In the latter case, the sum of the even-numbered terms must diverge to  $-\infty$  (as the sum of all terms is convergent), and therefore cannot diverge to  $\infty$ . Thus player B is chosen on this step with probability 0, finishing the proof.  $\square$

We now prove the various lemmas.

**PROOF OF LEMMA 3.3.** As Honest's partial sums form a truly random walk, this follows from the law of the iterated logarithm.  $\square$

**PROOF OF LEMMA 3.4.** Suppose without loss of generality that A is Honest. Let  $X_n = \frac{a_n}{\sqrt{n}(\log n)^{3/4}}$  for  $n \geq 2$ . Since

$$\sum_{n=3}^{\infty} \mathbf{E}(X_n^2) = \sum_{n=3}^{\infty} \frac{1}{n(\log n)^{3/2}} < \infty.$$

Kolmogorov's two-series theorem (e.g., [6], Theorem 2.5.6) implies that  $\sum_{n=2}^{\infty} X_n$  converges almost surely. Thus A is chosen incorrectly on Step 2 with probability 0.  $\square$

**PROOF OF LEMMA 3.5.** Since no player was chosen on Step 1,  $s_{2n-1} = 1 + A_n + B_{n-1}$  and  $s_{2n} = 1 + A_n + B_n$  are both  $o(\sqrt{n}(\log n)^{1/4})$ , so  $s_n = o(\sqrt{n}(\log n)^{1/4})$ .

Since no player was chosen on Step 2,

$$\sum_{n=2}^{\infty} \frac{a_n + b_n}{\sqrt{n}(\log n)^{3/4}}$$

converges. We may rewrite this sum using the fact that  $a_n + b_n = s_{2n} - s_{2n-2}$  to obtain

$$\sum_{n=2}^{\infty} \frac{s_{2n} - s_{2n-2}}{\sqrt{n}(\log n)^{3/4}} = O(1) + \sum_{n=2}^{\infty} \left( \frac{1}{\sqrt{n}(\log n)^{3/4}} - \frac{1}{\sqrt{n+1}(\log(n+1))^{3/4}} \right) s_{2n}.$$



This last regrouping is possible because the difference in the partial sums of the two sides is  $\frac{s_{2n}}{\sqrt{n+1}(\log(n+1))^{3/4}}$ , which approaches 0 as  $s_n = O(\sqrt{n}(\log n)^{1/4})$ .

Let  $c_n = \frac{1}{\sqrt{n}(\log n)^{3/4}} - \frac{1}{\sqrt{n+1}(\log(n+1))^{3/4}}$ . Since the derivative of the function  $\frac{1}{\sqrt{x}(\log x)^{3/4}}$  is

$$\left(-\frac{1}{2} + o(1)\right) \frac{1}{x^{3/2}(\log x)^{3/4}}$$

as  $x \rightarrow \infty$ , we have  $c_n = \Theta(\frac{1}{n^{3/2}(\log n)^{3/4}})$ . Since  $\sum c_n s_{2n}$  converges and  $s_n \geq 0$  for all  $n$ ,

$$\sum_{n=2}^{\infty} \frac{s_{2n}}{n^{3/2}(\log n)^{3/4}}$$

must converge as well. Since  $|s_{2n-1} - s_{2n}| \leq 1$  and  $\sum n^{-3/2}(\log n)^{-3/4}$  converges,

$$(3) \quad \sum_{n=2}^{\infty} \frac{s_{2n-1} + s_{2n}}{n^{3/2}(\log n)^{3/4}}$$

converges. Now, the coefficient of  $s_n$  in expression (3) is  $(2\sqrt{2} + o(1))n^{-3/2}(\log n)^{-3/4}$  for  $n \geq 3$ , so finally the sum

$$(4) \quad \sum_{n=2}^{\infty} \frac{s_n}{n^{3/2}(\log n)^{3/4}}$$

must converge.

Since  $s_n = o(\sqrt{n}(\log n)^{1/4})$ , we may multiply each term in expression (4) by  $\frac{s_n}{\sqrt{n}(\log n)^{1/4}}$  and retain convergence, using the nonnegativity of  $s_n$ . Thus  $\sum_{n=1}^{\infty} \frac{s_n^2}{n^2 \log n}$  converges.  $\square$

**REMARK 3.7** (The positivity of  $(s_n)$ ). The last step of the proof of Lemma 3.5 is the only location where the positivity of  $s_n$  is used. Indeed,  $\sum s_n / (n^{3/2}(\log n)^{3/4})$  would converge with high probability for a random walk as well. It is only because all  $s_n$  are positive that this is surprising.

**PROOF OF LEMMA 3.6.** We prove the lemma under the assumption that A is Honest; the proof is analogous when B is Honest.

Notice that

$$\sum_{n=1}^{\infty} \frac{s_{2n+1}^2 - s_{2n}^2}{2n \log(2n)} = \sum_{n=1}^{\infty} \frac{1 + 2a_{n+1}s_{2n}}{2n \log(2n)}.$$

Now,  $\sum 1/(2n \log(2n))$  diverges to  $\infty$  (at rate on the order of  $\log \log n$ ). It thus suffices to show that the probability that no player is chosen on Step 1 and

$$(5) \quad \sum_{n=1}^{\infty} \frac{a_{n+1}s_{2n}}{n \log(2n)}$$

diverges is 0.

Let  $s'_n = \min(s_n, \sqrt{n} \log \log n)$ . Notice that if no player is chosen on Step 1,  $s'_n = s_n$  for all sufficiently large  $n$ . Thus the probability that no player is chosen on Step 1 and expression (5) diverges is at most the probability that  $\sum_{n=1}^{\infty} \frac{a_{n+1}s'_{2n}}{n \log(2n)}$  diverges.

Let  $D_n = \frac{a_{n+1}s'_{2n}}{n \log(2n)}$ . Since  $a_{n+1}$  is chosen at random after  $s'_{2n}$  is already fixed, it follows that  $D_n$  is a sequence of martingale differences. Since  $s'_n \leq \sqrt{n} \log \log n$ , it follows that

$$\sum_{n=1}^{\infty} \mathbf{E}(D_n^2) \leq \sum_{n=1}^{\infty} \frac{1}{n(\log n)^{3/2}} + O(1) < \infty.$$

Therefore, the partial sums of the infinite series  $\sum_{n=1}^{\infty} D_n$  form a martingale bounded in  $L^2$ , and therefore the series converges almost surely.  $\square$

We turn to handle the example when the horizon is finite. If we only have a finite number of samples, say  $N$ , we can guarantee a probability of failure  $\epsilon \rightarrow 0$  as  $N \rightarrow \infty$  by modifying the test as follows.

1. If  $\frac{A_n}{\sqrt{n} \log \log n} > \epsilon$  for any  $n$  with  $\log \log N < n \leq N$ , choose player A. Otherwise, if  $\frac{B_n}{\sqrt{n} \log \log n} > \epsilon$  for any  $n$  with  $\log \log N < n \leq N$ , choose player B.
2. Take  $C = \sum_{n=1}^{\infty} \frac{1}{n(\log n)^{3/2}}$ . If  $\sum_{n=2}^N \frac{a_n}{\sqrt{n}(\log n)^{3/4}} > \sqrt{3C/\epsilon}$ , choose player A. Otherwise, if  $\sum_{n=2}^N \frac{b_n}{\sqrt{n}(\log n)^{3/4}} > \sqrt{3C/\epsilon}$ , choose player B.
3. If  $\sum_{n=1}^{\lfloor \frac{N-1}{2} \rfloor} \frac{s_{2n+1}^2 - s_{2n}^2}{2n \log(2n)} > (\log \log N)/3$ , choose player B. Otherwise, if  $\sum_{n=2}^{\lfloor \frac{N}{2} \rfloor} \frac{s_{2n}^2 - s_{2n-1}^2}{(2n-1) \log(2n-1)} > (\log \log N)/3$ , choose player A.
4. Otherwise, choose player A.

The next result provides the analog of Theorem 3.1 to the finite horizon set up.

**PROPOSITION 3.8.** *If  $N$  is sufficiently large in terms of  $\epsilon$ , the algorithm above identifies the deviator with failure probability at most  $\epsilon$ .*

**PROOF.** The proof is almost identical to the proof of Theorem 3.1, so we briefly outline the changes. For sufficiently large  $N$ , the probability of failure at Step 1 is at most  $\epsilon/3$ . Similarly, the probability of failure at Step 2 is also at most  $\epsilon/3$  (using a second moment bound and Markov's inequality). If both Steps 1 and 2 are passed, then

$$\sum_{n=\log \log N}^N \frac{s_n^2}{n^2 \log n}$$

is bounded by an absolute constant by a similar argument to earlier, so

$$\sum_{n=1}^N \frac{s_n^2}{n^2 \log n} = o(\log \log N).$$

For the same reason as in the original argument, a player will be chosen in Step 3 with probability approaching 1 as  $N \rightarrow \infty$  (independently of  $\epsilon$ ), and adding the sums yields (up to a constant)  $\sum_{n=1}^N \frac{s_n^2 - s_{n-1}^2}{n \log n} \approx \sum_{n=1}^N \frac{s_n^2}{n^2 \log n} = o(\log \log N)$ , so only one player is chosen during this step. If, for example, A is incorrectly chosen as the deviator, a similar computation to earlier yields that the partial sum of the series (5) up to  $N$  is  $\Omega(\log \log N)$ , which occurs with a probability approaching 0 as  $N \rightarrow \infty$ . Hence, choosing  $N$  sufficiently large, the deviator is incorrectly identified with probability at most  $\epsilon/3$  in this step. In conclusion, the total probability of failure is at most  $\epsilon$ .  $\square$

**4. Proof of Theorem 2.6.** In this section we prove Theorem 2.6 via a game-theoretic approach. Roughly speaking, we consider a zero-sum game between an adversary and a statistician, in which the adversary chooses a deviation and the statistician, after observing the realization  $s$ , has to guess the deviator if  $s \notin D$ . A strategy for the statistician in this game is a blame function. We use the minimax theorem to establish that the statistician has a strategy that guarantees high payoff. However, for the minimax theorem to apply, we need to make some modifications to the game.<sup>1</sup>

First, we may assume without loss of generality that  $D$  is closed. Indeed, every probability distribution over  $A^{\mathbb{N}}$  is regular, so every Borel set  $D$  with  $\mathbf{P}_{\sigma^*}(D) > 1 - \varepsilon$  contains a closed subset  $F$  with  $\mathbf{P}_{\sigma^*}(F) > 1 - \varepsilon$ . Hence, if  $D$  is not closed, we can replace it by  $F$ .

For every  $z = (z^1, z^2, \dots, z^n) \in A^{<\mathbb{N}}$  denote by  $[T_z] = \{s \in A^{\mathbb{N}} : s^k = z^k \text{ for } 1 \leq k \leq n\}$  the cylinder set of all realizations with initial segment  $z$ . A set in  $A^{\mathbb{N}}$  is open if and only if it is a union of cylinder sets.

Since the set  $D$  is closed, its complement  $D^c$  is open, and therefore there is a set  $Z \subseteq A^{<\mathbb{N}}$  that satisfies the following properties:

- For every two elements in  $Z$ , none is the prefix of the other.
- $D^c = \bigcup_{z \in Z} [T_z]$ .

Since  $A^{<\mathbb{N}}$  is countable, so is  $Z$ .

We now consider the following auxiliary zero-sum game  $\Gamma(D)$  between an adversary and a statistician:

- The adversary selects an element of  $i \in I$  (a player in the original problem) and a behavior strategy  $\sigma_i \in \Sigma_i$  for that player.
- Nature chooses a realization  $s \in A^{\mathbb{N}}$  according to  $\mathbf{P}_{\sigma_i, \sigma_{-i}^*}$ .
- If  $s \notin D$ , the statistician is told the element  $z \in Z$  such that  $s \in [T_z]$ . The statistician has then to select an element  $j \in I$ .
- The statistician wins if  $s \in D$  or  $i = j$ .
- The adversary wins otherwise, that is, if  $s \notin D$  and  $i \neq j$ .

The interpretation of the game is as follows. The statistician has to detect which player deviated, and the adversary tries to cause the statistician to blame an innocent player. Thus, the adversary's strategy is to select the identity of the deviator  $i \in I$  and a strategy for that deviator. Then Nature chooses a realization  $s$  according to the strategy that player  $i$  deviated to and the prescribed strategies of the other players. If  $s \in D$ , then the statistician wins. If  $s \notin D$ , then the statistician learns the minimal prefix  $z$  of the realization all of whose extensions are not in  $D$ , and she wins only if she correctly guesses the identity of the deviator based on this information.

**LEMMA 4.1.** *Let  $D$  be a closed set such that  $\mathbf{P}_{\sigma^*}(D) > 1 - \varepsilon$ . Then for every strategy of the adversary, the statistician has a response that wins in  $\Gamma(D)$  with probability at least  $1 - \sqrt{(|I| - 1)\varepsilon}$ .*

**PROOF.** Fix a strategy  $(q, \sigma) \in \Delta(I) \times \prod_i \Sigma_i$  of the adversary in  $\Gamma(D)$ .

Recall that  $D^c = \bigcup_{z \in Z} [T_z]$  for some countable set  $Z$  of finite realizations with the property that no element of  $Z$  is a prefix of a different element of  $Z$ .

For a finite realization  $z = (z^1, z^2, \dots, z^m) \in Z$  let

$$\ell_i(z) := \prod_{n=1}^m \frac{\sigma_i(z_i^n | z^1, z^2, \dots, z^{n-1})}{\sigma_i^*(z_i^n | z^1, z^2, \dots, z^{n-1})} \quad \forall i \in I,$$

<sup>1</sup>Blackwell [5] gives an example of a statistical game without a value.

where  $\frac{0}{0} = 1$  and  $\frac{c}{0} = \infty$  for  $c > 0$ . Then  $\ell_i(z)$  is the likelihood ratio of  $\sigma_i$  (the deviation strategy of player  $i$ ) over  $\sigma_i^*$  (the goal strategy of player  $i$ ) under the realized sequence  $z$ . We recall that in the general model, each player chooses an outcome in all periods, and therefore  $\sigma_i(z_i^n | z^1, z^2, \dots, z^{n-1})$  is defined for every  $n \in \mathbb{N}$ .

Note that the probability that a finite realization  $z$  will be realized under  $(\sigma_i, \sigma_{-i}^*)$  is  $\mathbf{P}_{\sigma_i, \sigma_{-i}^*}(z) = \ell_i(z) \mathbf{P}_{\sigma^*}(z)$ , provided  $\ell_i(z) < \infty$ . Similarly, the probability that  $z$  will be realized under  $(\sigma_i, \sigma_j, \sigma_{-i,j}^*)$  is  $\mathbf{P}_{\sigma_i, \sigma_j, \sigma_{-i,j}^*}(z) = \ell_i(z) \ell_j(z) \mathbf{P}_{\sigma^*}(z)$ , provided  $\ell_i(z), \ell_j(z) < \infty$ , where  $\sigma_{-i,j}^* = (\sigma_k^*)_{k \in I \setminus \{i,j\}}$ .

Consider a pure strategy of the statistician that, after observing a finite realization  $z \in Z$ , blames a player  $j$  whose likelihood ratio is maximal. For each  $j \in I$ , denote by  $E_j$  the set of all sequences in  $Z$  where the statistician blames  $j$ . Then

$$E_j \subseteq \{z \in Z : \ell_j(z) \geq \ell_i(z) \text{ for every } i \neq j\},$$

where the inclusion may be strict when for some  $z \in Z$  the maximum of  $\{\ell_i(z), i \in I\}$  is attained at  $j$  together with some other index.

Observe that

$$\begin{aligned} (\mathbf{P}_{\sigma_i, \sigma_{-i}^*}(E_j))^2 &= \left( \sum_{z \in E_j} \ell_i(z) \mathbf{P}_{\sigma^*}(z) \right)^2 \\ (6) \quad &\leq \left( \sum_{z \in E_j} \ell_i(z)^2 \mathbf{P}_{\sigma^*}(z) \right) \cdot \left( \sum_{z \in E_j} \mathbf{P}_{\sigma^*}(z) \right) \\ (7) \quad &\leq \left( \sum_{z \in E_j} \ell_i(z) \ell_j(z) \mathbf{P}_{\sigma^*}(z) \right) \cdot \left( \sum_{z \in E_j} \mathbf{P}_{\sigma^*}(z) \right) \\ (8) \quad &= \mathbf{P}_{\sigma_i, \sigma_j, \sigma_{-i,j}^*}(E_j) \cdot \mathbf{P}_{\sigma^*}(E_j) \leq \mathbf{P}_{\sigma^*}(E_j), \end{aligned}$$

where equation (6) holds by the Cauchy–Schwarz inequality, equation (7) holds since  $\ell_i(z) \leq \ell_j(z)$  on  $E_j$ , and equation (8) follows from the definitions. By the Cauchy–Schwarz inequality once again, it follows that

$$\begin{aligned} \left( \sum_{j \neq i} \mathbf{P}_{\sigma_i, \sigma_{-i}^*}(E_j) \right)^2 &\leq (|I| - 1) \cdot \left( \sum_{j \neq i} \mathbf{P}_{\sigma_i, \sigma_{-i}^*}(E_j)^2 \right) \\ &\leq (|I| - 1) \cdot \left( \sum_{j \neq i} \mathbf{P}_{\sigma^*}(E_j) \right) \\ &= (|I| - 1) \cdot \mathbf{P}_{\sigma^*}(Z) \leq (|I| - 1) \cdot \varepsilon, \end{aligned}$$

and the claim follows.  $\square$

We now conclude the proof of Theorem 2.6. For every  $n \in \mathbb{N}$  let  $Z_n = \{z \in Z : \text{length of } z < n\}$ , and let  $D_n \subseteq A$  be the set whose complement is given by

$$D_n^c = \bigcup \{[T_z] : z \in Z_n\}.$$

The sequence  $(D_n)_{n \in \mathbb{N}}$  is a decreasing sequence of closed sets that contain  $D$ , and because  $D$  is closed,  $D = \bigcap_{n \in \mathbb{N}} D_n$ . In particular,  $\mathbf{P}_{\sigma^*}(D_n) > 1 - \varepsilon$  for every  $n$ . The set of pure strategies of the statistician in the game  $\Gamma(D_n)$  is finite, as the game ends after  $n$  periods. By a standard minimax theorem, the game has a value in mixed strategies, and the statistician has an optimal strategy,  $\xi_n : Z_n \rightarrow \Delta(I)$ .

By Lemma 4.1, For every  $n \in \mathbb{N}$ , the value of the game  $\Gamma(D_n)$  is at least  $1 - \sqrt{(|I| - 1)\varepsilon}$ . Let  $f_n : Z_n \rightarrow I$  be a blame function such that  $f_n(z) \in \operatorname{argmax}_{i \in I} \xi_n[z]$ . It follows that if  $f_n(z) \neq i$  then  $(\xi_n(z))(I \setminus \{i\}) \geq \frac{1}{2}$ , and hence

$$\mathbf{P}_{\sigma_i, \sigma_{-i}^*}(D_n^c \text{ and } \{f_n \neq i\}) \leq 2\mathbf{P}_{\sigma_i, \sigma_{-i}^*, \xi_n}(D_n^c \text{ and } \{j \neq i\}) \leq 2\sqrt{(|I| - 1)\varepsilon},$$

for every  $i \in I$  and every  $\sigma_i \in \Sigma_i^B$ . Abusing notation, we view  $f_n$  as a function from  $D_n^c$  to  $I$ , such that for every  $z \in Z_n$  and every  $s \in [T_z]$ , we set  $f_n(s) = f_n(z)$ . It follows that  $f_n$  is a blame function that guarantees to the statistician at least  $1 - 2\sqrt{(|I| - 1)\varepsilon}$  in  $\Gamma(D_n)$ .

For every  $n \in \mathbb{N}$ , the domain of  $f_n$  is the finite set  $Z_n$ . By a diagonal argument, there is a function  $f : D^c \rightarrow I$  that is an accumulation point of the sequence  $(f_n)_{n \in \mathbb{N}}$ : there is a subsequence  $(n_k)_{k \in \mathbb{N}}$  such that for every  $z \in Z$  and every  $s \in T_{[z]}$ ,  $f(s)$  is equal to  $f_{n_k}(s)$ , for all sufficiently large  $k \in \mathbb{N}$ .

We argue that  $f$  guarantees to the statistician at least  $1 - 2\sqrt{(|I| - 1)\varepsilon}$  in  $\Gamma(D)$ . Indeed, let  $i \in I$  and  $\sigma_i \in \Sigma_i^B$  be arbitrary. Since  $D^c = \bigcup_{k \in \mathbb{N}} D_{n_k}^c$ ,

$$\mathbf{P}_{\sigma_i, \sigma_{-i}^*}(D^c \text{ and } \{f \neq i\}) = \lim_{k \rightarrow \infty} \mathbf{P}_{\sigma_i, \sigma_{-i}^*}(D_{n_k}^c \text{ and } \{f \neq i\}) \leq 2\sqrt{(|I| - 1)\varepsilon},$$

and the result follows.

**REMARK 4.2** (The value of the infinite-horizon game). Instead of studying the truncated games  $\Gamma(D_n)$ , we could have proved that the game  $\Gamma(D)$  has a value by showing that the statistician's payoff function is upper-semi-continuous and her strategy space is compact, and use a general minimax theorem, like [11], Theorem 4. We chose the path above, as it uses the simpler version of von Neumann's minimax theorem. Moreover, Theorems 2.6 and 2.9 are valid also when the set of actions  $A$  is countably infinite.

**5. Concluding remarks and open problems.** The constant  $\sqrt{|I| - 1}$  in Theorem 2.6 can be improved to  $\lceil \ln_2(I) \rceil$ , by iteratively dividing the set of players who are suspected as possible deviators into two groups, and identifying whether a player in one of the groups deviated.

The example in Section 3.1 shows that the bound in Theorem 2.6 is tight up to a constant factor. As we now argue, the same holds in the example described in Section 3.2. Indeed, the set  $D$  in this example has probability  $1 - \varepsilon = 1 - \Theta(\mu^2)$ , and the blame function shows that the corresponding goal is  $\mu = O(\sqrt{\varepsilon})$ -testable. It is in fact not difficult to see that for this example the goal is not  $\delta$ -testable for any  $\delta$  smaller than some  $\Theta(\sqrt{\varepsilon}) = \Theta(\mu)$ , showing that the quantitative estimate given in Theorem 2.6 is tight up to a constant factor. Indeed, consider the following two scenarios.

1. Player A chooses  $s_1 = 1$  and later plays honestly according to the rules. Player B plays honestly.
2. Player A plays honestly. Player B chooses  $s_2 = 1$  and later plays honestly.

In both scenarios, the probability that  $s_1 = s_2 = 1$  is  $\Theta(\mu)$ . Moreover, in both scenarios, if indeed  $s_1 = s_2 = 1$ , then the conditional distribution of  $s$  is identical. Therefore, on the subset of  $D^c$  consisting of all sequences  $s$  above with  $s_1 = s_2 = 1$ , the two scenarios are indistinguishable and any blame function chooses one of the players with probability at least  $1/2$ . It thus follows that the probability that  $s$  lies in this subset and the blame function blames the honest player is  $\Omega(\mu)$ .

Theorem 2.6 is not constructive. In Section 3 we described two cases where a blame function could be identified, and in these cases, especially in the second one, the blame function

is quite involved. There are other interesting cases where identifying an explicit blame function looks challenging. For example, consider the two-dimensional analog of the example in Section 3.3: two players control a two-dimensional random walk, and the set  $D$  is the set of all realizations that visit the origin infinitely often (or at least once after the initial position). The same question can be considered for a random walk on any recurrent graph.

It is easy and well known that the probability that a one-dimensional honest random walk starting at the origin never returns to the origin for  $N$  steps is  $\Theta(1/\sqrt{N})$ . This implies, by Theorem 2.6, that for the version of the game considered in Section 3.3, where the set  $D$  consists of all walks that visit the origin at least once during the first  $N$  steps, there is a blame function that errs with probability at most  $O(1/N^{1/4})$ . The quantitative estimate that can be derived from the explicit proof described in Section 3.3 is far weaker. Indeed, consider item 1 in the description following the proof of Lemma 3.6. If  $A$  is an honest player, then the probability that for some fixed  $n$ ,  $\frac{A_n}{\sqrt{n} \log \log n} > 1(> \varepsilon)$  is roughly  $e^{-(\log \log n)^2/2}$ . Therefore for, say,  $n = \log N$  this probability is roughly  $e^{-(\log \log \log N)^2/2}$  which is (much) larger than  $1/\log N$ . It follows that the probability that the honest player  $A$  is blamed in this algorithm is larger than  $1/\log N \gg 1/N^{1/4}$ . It may be interesting to find an explicit blame function with a better quantitative performance. The corresponding question for an  $N$ -steps two-dimensional random walk is even more challenging. It is well known ([7], see also [8, 15]) that the probability that a standard two-dimensional random walk does not return to the origin for  $N$  steps is  $\Theta(1/\log N)$ . Theorem 2.6 thus shows that the corresponding goal here is  $O(1/(\log N)^{1/2})$ -testable. It would be very interesting to find an explicit description of a blame function demonstrating this bound.

Theorem 2.6 states that if there is an agreed upon strategy profile  $\sigma^*$  that reaches some desired target set  $D$  with high-probability, and if a single player deviates, then with high probability, the identity of the deviator can be found by all players when the target set is not reached. Such a result calls for applications in the construction of equilibria in game theory. As mentioned in the Introduction, Theorem 2.6 can be used to provide an alternative proof for the existence of an  $\varepsilon$ -equilibrium in repeated games with finitely many players each having finitely many actions, and tail-measurable payoffs; see [9]. Theorem 2.6 has also been used to prove that in every multiplayer stochastic game with finite state and action spaces, and with bounded, Borel measurable payoff functions, for every  $\varepsilon > 0$  there is a subgame in which an  $\varepsilon$ -equilibrium exists; see [10].

**Acknowledgments.** We thank Orin Munk, Asaf Nachmias, Bhargav Narayanan, Ron Solan, and David Steinberg for useful discussions.

**Funding.** Alon's research was supported in part by the National Science Foundation Grant DMS-2154082.

He acknowledges the support of the National Science Foundation Grant DMS-2103154.

Shmaya acknowledges the support of the National Science Foundation Grant 89371-1164105-1, and the Binational Science Foundation, Grant 2022178.

Solan acknowledges the support of the Israel Science Foundation Grants #217/17 and #211/22, and the Binational Science Foundation, Grant 2022178.

## REFERENCES

- [1] ABRAMOVICH, F. and RITOV, Y. (2013). *Statistical Theory: A Concise Introduction. Texts in Statistical Science Series*. CRC Press, Boca Raton, FL. [MR3114871](#)
- [2] ASHKENAZI-GOLAN, G., FLESCH, J., PREDTETCHINSKI, A. and SOLAN, E. (2022). Existence of equilibria in repeated games with long-run payoffs. *Proc. Natl. Acad. Sci. USA* **119** Paper No. e2105867119, 8. [MR4417583](#)

- [3] ASHKENAZI-GOLAN, G., FLESCHE, J., PREDTETCHINSKI, A. and SOLAN, E. (2023). Regularity of the minmax value and equilibria in multiplayer Blackwell games. *Israel J. Math.* forthcoming.
- [4] AUMANN, R. J. and SHAPLEY, L. S. (1994). Long-term competition—a game-theoretic analysis. In *Essays in Game Theory* (Stony Brook, NY, 1992) 1–15. Springer, New York. [MR1300862](#)
- [5] BLACKWELL, D. (1983). A hypothesis-testing game without a value. In *A Festschrift for Erich L. Lehmann. Wadsworth Statist./Probab. Ser.* 79–82. Wadsworth, Belmont, CA. [MR0689739](#)
- [6] DURRETT, R. (2019). *Probability—Theory and Examples*. 2th ed. *Cambridge Series in Statistical and Probabilistic Mathematics* **49**. Cambridge Univ. Press, Cambridge. [MR3930614](#) <https://doi.org/10.1017/9781108591034>
- [7] DVORETZKY, A. and ERDŐS, P. (1951). Some problems on random walk in space. In *Proceedings of the Second Berkeley Symposium on Mathematical Statistics and Probability*, 1950 353–367. Univ. California Press, Berkeley. [MR0047272](#)
- [8] ERDŐS, P. and TAYLOR, S. J. (1963). Some problems concerning the structure of random walk paths. *Acta Math. Acad. Sci. Hung.* **11** 137–162. [MR0121870](#) <https://doi.org/10.1007/BF02020631>
- [9] FLESCHE, J. and SOLAN, E. (2023). Repeated games with tail-measurable payoffs. *David Gale: Mathematical economist: Essays in appreciation on His 100th Birthday*. Monographs in Mathematical Economics. (Springer), forthcoming.
- [10] FLESCHE, J. and SOLAN, E. (2023). Stochastic games with general payoff functions. *Math. Oper. Res.* forthcoming.
- [11] HA, C. W. (1980). Minimax and fixed point theorems. *Math. Ann.* **248** 73–77. [MR0569411](#) <https://doi.org/10.1007/BF01349255>
- [12] MAITRA, A. and SUDDERTH, W. (1998). Finitely additive stochastic games with Borel measurable payoffs. *Internat. J. Game Theory* **27** 257–267. [MR1645524](#) <https://doi.org/10.1007/s001820050071>
- [13] MARTIN, D. A. (1998). The determinacy of Blackwell games. *J. Symbolic Logic* **63** 1565–1581. [MR1665779](#) <https://doi.org/10.2307/2586667>
- [14] NASH, J. F. JR. (1950). Equilibrium points in  $n$ -person games. *Proc. Natl. Acad. Sci. USA* **36** 48–49. [MR0031701](#) <https://doi.org/10.1073/pnas.36.1.48>
- [15] RÉVÉSZ, P. (2013). *Random Walk in Random and Non-random Environments*, 3rd ed. World Scientific, Hackensack, NJ. [MR3060348](#) <https://doi.org/10.1142/8678>
- [16] SOLAN, E. (1999). Three-player absorbing games. *Math. Oper. Res.* **24** 669–698. [MR1854248](#) <https://doi.org/10.1287/moor.24.3.669>
- [17] VIEILLE, N. (2000). Solvable states in  $n$ -player stochastic games. *SIAM J. Control Optim.* **38** 1794–1804. [MR1776656](#) <https://doi.org/10.1137/S0363012998345937>