Games and Economic Behavior 119 (2020) 383–391

Contents lists available at ScienceDirect

Games and Economic Behavior

www.elsevier.com/locate/geb

Jointly controlled lotteries with biased coins

Eilon Solan [a,*], Omri N. Solan [a], Ron Solan [b]

[a] The School of Mathematical Sciences, Tel Aviv University, Tel Aviv 6997800, Israel
[b] The Faculty of Life Sciences, Tel Aviv University, Tel Aviv 6997800, Israel

ARTICLE INFO

Article history:
Received 10 April 2018
Available online 20 November 2018

JEL classification:
C65
C72
C73

Keywords:
Jointly controlled lotteries
Biased coin
Quitting games
Undiscounted equilibrium
Uniform equilibrium

ABSTRACT

We study the implementation of a jointly controlled lottery when the coins that are used by the players are exogenously given. We apply this result to show that every quitting game in which at least two players have at least two continue actions has an undiscounted ε-equilibrium, for every $\varepsilon > 0$.

1. Introduction

Jointly controlled lotteries are a technique that allows group members to reach a random joint decision in such a way that no member can affect the distribution of the final decision (see, e.g., Aumann and Maschler, 1995; Forges, 1995; Lehrer, 1996; Lehrer and Sorin, 1997; Forges and Koessler, 2008, or Kalai et al., 2010). For example, if two people want to jointly selects a color, blue or green, each with probability $\frac{1}{2}$, in such a way that none of them can manipulate the choice, then they can each toss a fair coin and determine the selected color according to whether the outcomes of the two coins are the same (blue) or different (green). Similarly, if the two players want to jointly select the color with unequal probabilities, say, blue with probability $\frac{1}{3}$ and green with probability $\frac{2}{3}$, they can each toss a fair dice, and select blue if the outcome of both dice is in the set $\{1, 2\}$, or if the outcome of both dice is in the set $\{3, 4\}$, or if the outcome of both dice is in the set $\{5, 6\}$, and select green otherwise. In both examples provided above, the probability measure of the randomization devices of the participants are not exogenously given, but are rather devised by the designer of the choice process to fit the desired probability measure that is to be implemented.

In this paper we study the problem, when the distribution of the randomization devices of the group members are exogenously given, and we allow the selection process to last more than one stage. Formally, for each $i \in \{1, 2, \dots, k\}$ we are given a randomization device, which selects at every stage a letter from a given finite alphabet A_i according to a given probability measure $\mathbf{P}_i$ over A_i. A mechanism is characterized by a stopping time τ and a deterministic rule f, that dictates which letter in some finite alphabet J is selected based on the letters that were selected by the randomization devices up

* Corresponding author.
E-mail addresses: eilons@post.tau.ac.il (E. Solan), omrisola@post.tau.ac.il (O.N. Solan), ron_solan@walla.com (R. Solan).

https://doi.org/10.1016/j.geb.2018.11.008
0899-8256/

to time τ. Given a probability measure ν over J, the goal is to devise a mechanism that is immune to deviations of some of the participants: even if $k-1$ of the randomization devices become faulty and produce letters according to some arbitrary law that may depend on past choices of all devices, the distribution of the outcome of the mechanism is still ν.

We will provide two mechanisms for selecting an element of J, both of which depend on a parameter $\varepsilon > 0$. One mechanism has a bounded length and selects an element in J with a distribution ε-close to ν: as long as at least one randomization device is not faulty, the probability that each element $j \in J$ is selected is ε-close to $\nu(j)$. The second mechanism may never terminate yet as long as no randomization device is faulty it is finite a.s. and selects each element $j \in J$ with probability exactly $\nu(j)$. Moreover, as long as at least one randomization device is not faulty, the probability that each element $j \in J$ is selected does not exceed $\nu(j)$, and, in case the mechanism never terminates, the identity of the faulty devices is revealed by the information they produce. Thus, even if some (but not all) devices are faulty, no element in J is chosen with probability greater than the intended probability that it is chosen.

To demonstrate the usefulness of the result, we apply it to the study of undiscounted equilibria in stochastic games. Whether every multiplayer stochastic game admits an undiscounted ε-equilibrium for every $\varepsilon > 0$ is one of the main open problems in game theory to date; see Flesch et al. (1997), Solan (1999, 2000), Vieille (2000a, 2000b), Solan and Vieille (2001), Flesch et al. (2007), Flesch et al. (2008, 2009), Simon (2012, 2016), and Solan and Solan (2017) for partial results, and Jaśkiewicz and Nowak (2017) for a recent survey. The class of games that we study in this paper is the class of positive recursive general quitting games. Those are quitting games in which (a) each player has a single quitting action and may have several continue actions, (b) the payoff if no absorption ever occurs is 0, and (c) the absorbing payoffs are nonnegative. This class of games was studied by Solan and Solan (2018), who showed that those games admit a sunspot ε-equilibrium for every $\varepsilon > 0$; that is, an ε-equilibrium in an extended game in which at every stage the players observe the outcome of a uniformly distributed random variable on $[0, 1]$, which is independent of past signals and past play. Using jointly controlled lotteries with biased coins we will show that if at least two players have at least two continue actions, an undiscounted ε-equilibrium exists.

To date it is not known whether quitting games in which each player has a single quitting action and a single continue action admit undiscounted ε-equilibria. Our result shows that when players have enough flexibility in coordinating their play, an undiscounted ε-equilibrium does exist.

Biased coins are not prevalent in game theory, since usually it is assumed that players have all randomization means that they need. One exception is Gossner and Vieille (2002), who studied two-player zero-sum repeated games in which the randomization device of one of the players is a biased coin that is tossed once at the beginning of every stage. Gossner and Vieille (2002) showed that the player can do better than using at every stage the outcome of the toss performed at the beginning of that stage, and characterized the value of the game as a function of the distribution of the coin. In their model, the player need not use the information provided by the coin at the stage in which it is obtained, but rather may use this information in subsequent stages. In our model, in contrast, aggregating the random information is impossible, since the letter chosen by a faulty device may depend on past choices of the unfaulty devices.

Though jointly controlled lotteries with biased coins remind one of mediated talk (see, e.g., Lehrer, 1996 and Lehrer and Sorin, 1997) and cheap talk (see, e.g., Farrell and Rabin, 1996 and Aumann and Hart, 2003), there are some significant differences among the models. Indeed, while in mediated talk and cheap talk the players are free to select the messages they send out and the goal is to choose an action for each player, in our model, when unfaulty, the randomization devices choose messages according to a known stationary probability measure and the goal is to choose one outcome.

The paper is arranged as follows. In Section 2 we discuss jointly controlled lotteries with biased coins, and in Section 3 we apply the mechanism of jointly controlled lotteries to positive recursive general quitting games.

2. Jointly controlled lotteries with biased coins

In the Introduction we presented the problem of jointly controlled lotteries with biased coins for any number of randomization devices. To simplify the presentation we will present the model and results for the case of two randomization devices; the definitions and constructions for any number of randomization devices follow the same lines.

2.1. On mechanisms

Let A_1 and A_2 be two finite sets, each containing at least two elements. The set of *finite histories*[1] is $H := \cup_{t=0}^{\infty}(A_1 \times A_2)^t$, and the set of *infinite histories* is $H^\infty := (A_1 \times A_2)^{\mathbf{N}}$. An infinite history is denoted $h = (a_i^{t'})_{i=1,2,t'\ge 0}$. For every infinite history $h = (a_i^{t'})_{i=1,2,t'\ge 0}$ and $t \ge 0$ we denote by $h^t = (a_i^{t'})_{i=1,2,0\le t'\le t}$ the prefix of h for the first t stages. The set H^∞ is a measurable space when equipped with the product sigma-algebra. For every $t \ge 0$ denote by $\mathcal{F}^t$ the sigma-algebra over H^∞ defined by all t-stage histories; it is the sigma-algebra spanned by the sets $C(h^t) := \{(h^t, h) \in H^\infty : h \in H^\infty\}$ for every $h^t \in (A_1 \times A_2)^t$, where (h^t, h) is the concatenation of the finite history h^t and the infinite history h.

[1] By conventions, the set $(A_1 \times A_2)^0$ contains only the empty history.

A *stopping time* (for the filtration $(\mathcal{F}^t)_{t\geq 0}$) is a function $\tau : H^\infty \to \mathbf{N} \cup \{\infty\}$ such that for every $t \in \mathbf{N}$ the set $\{h' \in H^\infty : \tau(h') = t\}$ is in $\mathcal{F}^t$. When τ is a stopping time, the sigma-algebra $\mathcal{F}^\tau$ is the σ-algebra of all events that are "known at time τ"; that it, the sigma-algebra generated by the collections of sets $\{A \in \mathcal{F}^t : \tau(h) \geq t \quad \forall h \in A\}$, for $t \in \mathbf{N}$.

The basic concept that we need is that of a mechanism, which describes how to generate an element from a set J given an infinite history.

Definition 2.1. A *mechanism* is a triplet $M := (\tau, J, f)$ where

- τ is a stopping time w.r.t. the filtration $(\mathcal{F}^t)_{t\geq 0}$.
- J is a finite set.
- $f : H^\infty \to J$ is a function that is measurable w.r.t. the sigma-algebra $\mathcal{F}^\tau$.

Remark 2.2. To better digest the concept of mechanism, consider the following game theoretic interpretation. Each device is a player, the players are engaged in a repeated game, and would like to collectively select some equilibrium play out of $|J|$ possible plays. The only means of communication between the players is through their actions. With this interpretation, A_i is the set of actions of player i, τ is the stage in which the selection process ends, and the function f dictates which equilibrium play is selected as a function of the players' actions.

Let $i \in \{1, 2\}$. A *(behavior) strategy* for the i's device is a function $\sigma_i : H \to \Delta(A_i)$ that assigns a distribution over A_i to each finite history. The set of all strategies for the i'th device is denoted Σ_i. A strategy is *stationary* if $\sigma_i(h^t)$ is independent of $h^t \in H$. When σ_i is a stationary strategy, we denote by $\sigma_i(a_i)$ the per-stage probability that strategy σ_i selects the element a_i, for each $a_i \in A_i$.

Every pair of strategies (σ_1, σ_2) defines a probability measure $\mathbf{P}_{\sigma_1,\sigma_2}$ over H^∞. We denote by $\mathbf{E}_{\sigma_1,\sigma_2}$ the expectation operator that corresponds to the probability measure $\mathbf{P}_{\sigma_1,\sigma_2}$.

2.2. Strong secure implementability

We now present three properties of mechanisms: having finite length, being able to implement a given probability measure, and being able to implement the probability measure in a secure way.

Definition 2.3. Let $T \in \mathbf{N}$. A mechanism $M = (\tau, J, f)$ *has length at most* T if $\mathbf{P}_{\sigma_1,\sigma_2}(\tau \leq T) = 1$ for every pair of strategies (σ_1, σ_2). A mechanism *has finite length* if it has length at most T for some $T \in \mathbf{N}$.

Definition 2.4. Let $\varepsilon \geq 0$ and let ν be a probability measure over some finite set J. The mechanism $M = (\tau, J, f)$ and the pair of strategies $(\sigma_1, \sigma_2) \in \Sigma_1 \times \Sigma_2$ *ε-implement the distribution ν* if

$$|\mathbf{P}_{\sigma_1,\sigma_2}(f = j) - \nu(j)| \leq \varepsilon, \quad \forall j \in J. \tag{1}$$

The mechanism $M = (\tau, J, f)$ and the pair of strategies $(\sigma_1, \sigma_2) \in \Sigma_1 \times \Sigma_2$ *ε-implement the distribution ν in a strong secure way* if for every $i \in \{1, 2\}$ and every strategy $\sigma_i' \in \Sigma_i$, the mechanism M and the pair of strategies $(\sigma_i', \sigma_{3-i})$ ε-implement the distribution ν.

Remark 2.5. Continuing the game theoretic interpretation of a mechanism presented in Remark 2.2, suppose that to satisfy some incentive constraints, the probabilities by which the various equilibrium plays are chosen should be close to some given probability measure ν. Suppose also that until the selection of an equilibrium play is made, due to incentive constraints, each player i should play some given stationary strategy σ_i. The mechanism has finite length if the selection process always ends, whatever the players play. The mechanism and the pair of stationary strategies (σ_1, σ_2) ε-implement the distribution ν in a strong secure way if by deviating no player can significantly affect the probability by which each equilibrium play is chosen.

Our first result concerns the possibility of ε-implementing any distribution in a strong secure way given any pair of stationary strategies.

Theorem 2.6. *Let A_1, A_2, and J be three finite sets, each of which contains at least two elements. Let σ_1 (resp. σ_2) be a stationary strategy that selects all elements in A_1 (resp. A_2) with positive probability, and let ν be any distribution on J. For every $\varepsilon > 0$ there is a mechanism $M = (\tau, J, f)$ that has finite length and, together with the pair of stationary strategies (σ_1, σ_2), ε-implements ν in a strong secure way.*

The intuition for the result is as follows. Suppose that $A_1 = A_2 = \{\alpha, \beta\}$, and consider the two-player zero-sum strategic-form game that appears in Fig. 1, where each player has two actions, α and β. The value of this game is 0, an optimal

		Player II	
		α	β
Player I	α	$-\sigma_1(\beta)\sigma_2(\beta)$	$\sigma_1(\beta)\sigma_2(\alpha)$
	β	$\sigma_1(\alpha)\sigma_2(\beta)$	$-\sigma_1(\alpha)\sigma_2(\alpha)$

Fig. 1. A strategic-form game with value 0

strategy for Player I is $(\sigma_1(\alpha), \sigma_1(\beta))$, and an optimal strategy for Player II is $(\sigma_2(\alpha), \sigma_2(\beta))$. Consider now the repeated game that is based on this strategic-form game, and denote by Y^t the payoff at stage t, for each $t \geq 0$. Since the value of the stage game is 0, as soon as Player I (resp. Player II) follows the stationary strategy σ_1 (resp. σ_2) for T stages, the sum $\sum_{t=1}^{T} Y^t$ is approximately normally distributed around 0, provided T is sufficiently large.

The variance d of the distribution of $\sum_{t=1}^{T} Y^t$ depends on the players' strategies, and, since players only know their own strategy and the play, this variance is not known by the players. However, the Martingale Central Limit Theorem (McLeish, 1974) allows us to approximate d by the outcomes, and therefore, a proper normalization of $\sum_{t=1}^{T} Y^t$ has approximately the standard normal distribution as soon as one of the players follows σ_1 or σ_2. The desired mechanism suggests itself: we divide the real line $\mathbf{R}$ into $|J|$ disjoint intervals such that the probability under the standard normal distribution of the j'th interval is $\nu(j)$, and set $f(h) = j$ if and only if the proper normalization of $\sum_{t=1}^{T} Y^t(h)$ lies in the j'th interval. We now turn to the formal proof.

Proof. Assume w.l.o.g.[2] that $|A_1| = |A_2| = 2$, and denote $A_i = \{\alpha, \beta\}$ for $i \in \{1, 2\}$. For every $t \in \mathbf{N}$ define a random variable Y^t over H^∞ as follows (see Fig. 1):

$$Y^t := \begin{cases} -\sigma_1(\beta)\sigma_2(\beta) & \text{if } a^t = (\alpha, \alpha), \\ \sigma_1(\beta)\sigma_2(\alpha) & \text{if } a^t = (\alpha, \beta), \\ \sigma_1(\alpha)\sigma_2(\beta) & \text{if } a^t = (\beta, \alpha), \\ -\sigma_1(\alpha)\sigma_2(\alpha) & \text{if } a^t = (\beta, \beta). \end{cases}$$

We observe that $\mathbf{E}_{\sigma_1, \sigma_2'}[Y^t] = \mathbf{E}_{\sigma_1', \sigma_2}[Y^t] = 0$, for every $t \geq 0$ and every pair of strategies $(\sigma_1', \sigma_2') \in \Sigma_1 \times \Sigma_2$.

For every real number $C > 0$ let τ_C be the stopping time

$$\tau_C := \min \left\{ t \in \mathbf{N} \colon \sum_{k=1}^{t} (Y^k)^2 \geq C \right\}. \tag{2}$$

Denoting by

$$c_0 := \min\{\sigma_1(\alpha), \sigma_1(\beta), \sigma_2(\alpha), \sigma_2(\beta)\} > 0, \tag{3}$$

we obtain that the stopping time τ_C is bounded by $\frac{C}{(c_0)^4}$. Denote $Z_C := \frac{\sum_{t=1}^{\tau_C} Y^t}{\sqrt{C}}$. The Martingale Central Limit Theorem (see, e.g., McLeish, 1974), implies that for each player i and each strategy $\sigma_i' \in \Sigma_i$ of player i, under the pair of strategies $(\sigma_i', \sigma_{3-i})$ the distribution of Z_C converges weakly to the standard normal distribution as C goes to infinity. Moreover, the rate of convergence is independent of σ_i'.

Let now $I_1, I_2, \cdots, I_J$ be a partition of the real line $\mathbf{R}$ into J disjoint intervals such that the probability of the interval I_j under the standard normal distribution is $\nu(j)$, for each $j \in J$. Let C be sufficiently large such that, for all strategies σ_1' and σ_2' of the two devices,

$$|\mathbf{P}_{\sigma_1, \sigma_2'}(Z_C \in I_j) - \nu(j)| < \varepsilon,$$
$$|\mathbf{P}_{\sigma_1', \sigma_2}(Z_C \in I_j) - \nu(j)| < \varepsilon.$$

For every infinite history $h \in H^\infty$ define $f(h)$ to be the unique $j \in J$ such that $Z_C(h) \in I_j$. It follows that the mechanism $M = (\tau_C, J, f)$ together with the pair of stationary strategies (σ_1, σ_2) ε-implement ν in a strong secure way. □

2.3. Weak secure implementability

We now weaken the security requirement of the mechanism. The weaker condition does not require that the mechanism stops in finite time, but rather that it stops in finite time when the two randomization devices are not faulty, that as long as the mechanism stops, no element in J is selected with probability higher than it was intended, and that if the mechanism does not stop (that is, on the set $\tau = \infty$), then the outputs of the faulty device reveal that it is faulty.

[2] If the set A_i contains more than two elements, divide it arbitrarily into two subsets, and treat all elements that lie in the same subset as equivalent.

Definition 2.7. Let $\varepsilon \geq 0$ and let ν be a probability measure over J. The mechanism $M = (\tau, J, f)$ and the pair of strategies $(\sigma_1, \sigma_2) \in \Sigma_1 \times \Sigma_2$ *ε-implement the distribution ν in a weak secure way* if the following conditions hold:

(W.1) $\mathbf{P}_{\sigma_1,\sigma_2}(\tau < \infty) = 1$ and M and (σ_1, σ_2) ε-implement the distribution ν.
(W.2) For every strategy $\sigma_1' \in \Sigma_1$ we have $\mathbf{P}_{\sigma_1',\sigma_2}(\tau < \infty \text{ and } f = j) \leq \nu(j)$, for every $j \in J$.
(W.3) For every strategy $\sigma_2' \in \Sigma_2$ we have $\mathbf{P}_{\sigma_1,\sigma_2'}(\tau < \infty \text{ and } f = j) \leq \nu(j)$, for every $j \in J$.
(W.4) There are two events D_1 and D_2 such that
- $D_1 \cup D_2 \subseteq \{\tau = \infty\}$.
- For every strategy $\sigma_2' \in \Sigma_2$ we have $\mathbf{P}_{\sigma_1,\sigma_2'}(D_1) = 0$.
- For every strategy $\sigma_1' \in \Sigma_1$ we have $\mathbf{P}_{\sigma_1',\sigma_2}(D_2) = 0$.
- For every strategy $\sigma_1' \in \Sigma_1$ we have $\mathbf{P}_{\sigma_1',\sigma_2}(D_1) + \mathbf{P}_{\sigma_1',\sigma_2}(\{\tau < \infty\}) = 1$.
- For every strategy $\sigma_2' \in \Sigma_2$ we have $\mathbf{P}_{\sigma_1,\sigma_2'}(D_2) + \mathbf{P}_{\sigma_1,\sigma_2'}(\{\tau < \infty\}) = 1$.

The events D_1 and D_2 in Definition 2.7 are used to reveal the identity of the faulty device: on the event D_i device i is faulty, for $i = 1, 2$; indeed, this event occurs with probability 0 if device i is unfaulty, and it occurs whenever the mechanism does not stop and device i is faulty. Note that whereas strong security requires the stopping time τ to be uniformly bounded, weak security has no such restriction.

Remark 2.8. Continuing Remarks 2.2 and 2.5 about the game theoretic interpretation of mechanisms, we now allow the players to deviate in a way that makes the selection process infinite, yet we require that even if a player deviates, he cannot increase the probability of each outcome j above $\nu(j)$. This property is important for the following reason. If each outcome j leads to some utility for every player, then the expected utility of a player may increase when the player deviates. If the utility from being detected as a deviator is not higher than the utility from each outcome in J (for example, if a detected deviation leads to indefinite punishment), then as soon as no element j is ever chosen with probability greater than $\nu(j)$, by deviating no player can increase his expected utility.

Theorem 2.9. *Let A_1, A_2, and J be three finite sets, each of which contains at least two elements. Let σ_1 (resp. σ_2) be a stationary strategy that selects all elements in A_1 (resp. A_2) with positive probability, and let ν be any distribution on J. There is a mechanism $M = (\tau, J, f)$ that together with the pair of stationary strategies (σ_1, σ_2) 0-implement ν in a weak secure way.*

Proof. Assume w.l.o.g. that $|A_1| = |A_2| = 2$, and denote $A_i = \{\alpha, \beta\}$ for $i = 1, 2$. Let $(Z^t)_{t \in \mathbf{N}}$ be a stochastic process with values in $\Delta(J)$, adapted to the filtration $(\mathcal{F}^t)_{t \geq 0}$, which satisfies the following properties:

(C.1) $Z^0 = \nu$.
(C.2) Z^{t+1} depends deterministically on Z^t, a_1^t, and a_2^t, for every $t \geq 0$. Consequently, Z^{t+1} is a function of the t-stage history h^t. We will therefore sometimes write $Z^{t+1}(h^t)$ rather than $Z^{t+1}(h)$.
(C.3) $\mathbf{E}_{\sigma_1,\sigma_2}[Z^{t+1} \mid Z^t, a_2^t = \alpha] = \mathbf{E}_{\sigma_1,\sigma_2}[Z^{t+1} \mid Z^t, a_2^t = \beta] = Z^t$.
(C.4) $\mathbf{E}_{\sigma_1,\sigma_2}[Z^{t+1} \mid Z^t, a_1^t = \alpha] = \mathbf{E}_{\sigma_1,\sigma_2}[Z^{t+1} \mid Z^t, a_1^t = \beta] = Z^t$.
(C.5) For every $(t-1)$-stage history $h^{t-1} \in H$ such that the support of $Z^t(h^{t-1})$ contains more than one element, the support of at least one of the distributions $Z^{t+1}(h^{t-1}, \alpha, \alpha)$, $Z^{t+1}(h^{t-1}, \alpha, \beta)$, $Z^{t+1}(h^{t-1}, \beta, \alpha)$, and $Z^{t+1}(h^{t-1}, \beta, \beta)$ is a strict subset of the support of $Z^t(h^{t-1})$.

Condition (C.3) states that whatever be the choice of Device 2 at stage t, as long as Device 1 follows σ_1, the expected value of Z^{t+1} is equal to the value of Z^t. Condition (C.4) is analogous. This implies in particular that the support of $Z^{t+1}(h^t)$ cannot contain elements that are not in the support of $Z^t(h^{t-1})$.

To show that such a process exists, let λ be a possible value of Z^t. Conditions (C.3) and (C.4) determine three equalities that the four variables $Z^{t+1}(h^{t-1}, \alpha, \alpha)$, $Z^{t+1}(h^{t-1}, \alpha, \beta)$, $Z^{t+1}(h^{t-1}, \beta, \alpha)$, and $Z^{t+1}(h^{t-1}, \beta, \beta)$ should satisfy. One solution of these equalities is $Z^{t+1}(h^{t-1}, a_1, a_2) = \lambda$ for every $a_1, a_2 \in \{\alpha, \beta\}$. Since the number of conditions is smaller by one than the number of variables, the set of solutions is a line, hence there is a solution on the boundary of the set $(\Delta(J))^4$, and therefore one can define Z^{t+1} in such a way that Conditions (C.2) and (C.5) hold.

Conditions (C.3) and (C.4) imply that the process $(Z^t)_{t \in \mathbf{N}}$ is a martingale under (σ_1, σ_2), hence it converges $\mathbf{P}_{\sigma_1,\sigma_2}$-a.s. to a random variable Z^∞. By Condition (C.5), under the stationary strategy pair (σ_1, σ_2), for every $t \in \mathbf{N}$, the probability that the support of Z^{t+1} is strictly contained in the support of Z^t (whenever the support of Z^t contains at least two elements) is at least $(c_0)^2$, where c_0 is defined as in Eq. (3). It follows that Z^∞ is a Dirac measure $\mathbf{P}_{\sigma_1,\sigma_2}$-a.s. Since the process $(Z^t)_{t \in \mathbf{N}}$ is a martingale, it follows that for every $j \in J$ we have $\mathbf{P}_{\sigma_1,\sigma_2}(Z^\infty = j) = Z^0(j) = \nu(j)$. Setting $\tau = \infty$ and $M = (\tau, J, Z^\infty)$ we obtain that M 0-implements the distribution ν, and Condition (W.1) holds.

Condition (C.3) implies that the process $(Z^t)_{t \geq 0}$ is a martingale under (σ_1, σ_2') for every strategy $\sigma_2' \in \Sigma_2$, which implies that Condition (W.3) holds. Analogously, Condition (W.2) holds as well.

We complete the proof by proving that Condition (W.4) holds. Denote by $(\widehat{a}_1^t, \widehat{a}_2^t) \in A_1 \times A_2$ an action pair such that the support of $Z^{t+1}(h^{t-1}, \widehat{a}_1^t, \widehat{a}_2^t)$ is strictly contained in the support of $Z^t(h^{t-1})$. We note that under strategy σ_i we have

$$\mathbf{P}_{\sigma_i, \sigma'_{3-i}}(a_i^t = \widehat{a}_i^t \text{ infinitely often}) = 1, \quad \forall \sigma'_{3-i} \in \Sigma_{3-i}.$$

For $i \in \{1, 2\}$ define an event D_i by

$$D_i := \{\tau = \infty \text{ and } a_i^t = \widehat{a}_i^t \text{ finitely many times}\}.$$

The event D_i contains all histories in which the mechanism does not stop and device i refrains from choosing infinitely many times the action that leads to a decrease in the support of Z^t. The reader can verify that Condition (W.4) in Definition 2.7 holds, and therefore the mechanism M 0-implements the distribution ν in a weak secure way. □

2.4. Discussion and open problems

We now discuss a variant of the mechanism described in the proof of Theorem 2.9 that has finite length. Fix $T \in \mathbf{N}$, and consider the mechanism described in the proof of Theorem 2.9 with two exceptions. First, the mechanism terminates after T periods. Second, the output of the mechanism may be undefined. Specifically, if the support of $Z^{T+1}(h^T)$ contains a single element of J, then this is the output $f(h^T)$ of the mechanism; otherwise we keep $f(h^T)$ undefined. If T is sufficiently large, then when no device is faulty, with high probability the support of Z^{T+1} is a singleton. Hence the mechanism together with the pair of stationary strategies (σ_1, σ_2) ε-implement ν for some small $\varepsilon = \varepsilon(T)$. To define the events D_1 and D_2, denote by $k_i(\alpha)$ the number of times that device i selects the letter α up to stage T. The set D_1 contains all T-stage histories $h^T \in H$ such that $Z^{T+1}(h^T)$ is not a singleton and $\left|\frac{k_1(\alpha)}{T} - \sigma_1(\alpha)\right| > \left|\frac{k_2(\alpha)}{T} - \sigma_2(\alpha)\right|$. The set D_2 contains all T-stage histories for which $Z^{T+1}(h^T)$ is not a singleton and $\left|\frac{k_1(\alpha)}{T} - \sigma_1(\alpha)\right| \le \left|\frac{k_2(\alpha)}{T} - \sigma_2(\alpha)\right|$. If T is sufficiently large and $Z^{T+1}(h^T)$ is not a singleton, then in the presence of one faulty device, with high probability the identity of the faulty device is correctly revealed.

The definitions of implementability make various assumptions: (a) the sets A_1, A_2, and J are finite, (b) the devices' outputs are chosen according to given stationary laws, (c) the devices observe each other's choice in every period, and (d) there is no external source of randomization. One can ask whether the results hold when one weakens any of these assumptions.

If the set J is infinite, then Theorem 2.6 still applies, because one can make J into a finite set by dropping all elements that have low probability and normalizing the distribution ν. That is, one defines the set J' to be a smallest subset of J satisfying $\sum_{j \in J'} \nu(j) \ge 1 - \frac{\varepsilon}{2}$ and the probability measure $\nu' \in \Delta(J')$ by $\nu'(j) := \frac{\nu(j)}{\nu(J')}$. A mechanism $M = (\tau, J', f)$ and a pair of strategies $(\sigma_1, \sigma_2) \in \Sigma_1 \times \Sigma_2$ that $\frac{\varepsilon}{2}$-implement the distribution ν' in a strong secure way, also ε-implement the distribution ν in a strong secure way.

If one changes the definition of implementability to require that the *ratio* $\frac{\mathbf{P}_{\sigma_1, \sigma_2}(f = j)}{\nu(j)}$ is between $1 - \varepsilon$ and $1 + \varepsilon$ for every $j \in J$ (instead of requiring that the difference between these two quantities is small, see Eq. (1)), then every distribution in $\Delta(J)$ is ε-implementable in a strong secure way, but the mechanism need not have finite length.

If the laws of the outputs of the devices are not stationary but arbitrary, then the results hold as soon as the laws "contain enough joint randomness". One sufficient condition is that, loosely speaking, with high probability there are sufficiently many stages in which at least two devices select each of at least two letters with probability uniformly bounded away from 0.

An interesting problem is the characterization of the monitoring structures that allow for a jointly controlled lottery in the presence of partial monitoring.

3. Undiscounted ε-equilibrium in general quitting games

In this section we provide an application of jointly controlled lotteries with biased coins to the area of stochastic games. As mentioned in the introduction, whether every stochastic game admits an undiscounted equilibrium payoff is one of the most challenging open problems in game theory to date. We will use the tools developed in Section 2 to prove the existence of an undiscounted ε-equilibrium in a class of stochastic games that was termed *positive recursive general quitting game* in Solan and Solan (2018). A *positive recursive general quitting game* is a vector $\Gamma = (I, (A_i^c)_{i \in I}, u)$ where

- $I = \{1, 2, \ldots, |I|\}$ is a nonempty finite set of players.
- A_i^c is a nonempty finite set of *continue actions*, for each player $i \in I$. The set of all *actions* of player i is $A_i := A_i^c \cup \{Q_i\}$, where Q_i is interpreted as a *quitting action*. The set of all action profiles is $A := \times_{i \in I} A_i$, and the set of all action profiles in which at least one player plays a quitting action is $A^* := A \setminus \left(\times_{i \in I} A_i^c\right)$.
- $u : A^* \to [0, 1]^I$ is a payoff function.

The game proceeds as follows. At every stage $t \in \mathbf{N}$, each player $i \in I$ chooses an action $a_i^t \in A_i$. Let $a^t = (a_i^t)_{i\in I}$ be the action profile chosen at stage t. Denote by t_* the first stage in which at least one player selects his quitting action; that is, the first stage t such that $a^t \in A^*$. If no player ever selects his quitting action, then $t_* = \infty$.

A *(behavior) strategy* of player i is a function $\sigma_i\colon \left(\cup_{t=0}^{\infty} A^t\right) \to \Delta(A_i)$. A *strategy profile* is a vector of strategies $\sigma = (\sigma_i)_{i\in I}$, one for each player. Every strategy profile σ induces a probability measure over the set of infinite histories A^∞. Denote by $\mathbf{E}_\sigma$ the corresponding expectation operator and by

$$\gamma(\sigma) := \mathbf{E}_\sigma\left[\mathbf{1}_{\{t_*<\infty\}} u(a^{t_*})\right]$$

the (expected undiscounted) *payoff* under strategy profile σ. Note that the way a strategy is defined after the termination stage t_* does not affect the payoff.

A mixed action profile $x \in \times_{i\in I}\Delta(A_i)$ is *nonabsorbing* if under x all players play continue actions with probability 1, and it is *absorbing* otherwise.

Let $\varepsilon \geq 0$. A strategy profile $\sigma = (\sigma_i)_{i\in I}$ is an ε*-equilibrium*[3] if for every player $i \in I$ and every strategy σ_i' of player i,

$$\gamma_i(\sigma) \geq \gamma_i(\sigma_i', \sigma_{-i}) - \varepsilon.$$

A *sunspot* ε*-equilibrium* is an ε-equilibrium in an extended game Γ^E that contains a correlation device, which sends a public signal s^t at the beginning of each stage $t \in \mathbf{N}$. The random variable s^t is uniformly distributed in $[0,1]$, and independent of $s^1, \cdots, s^{t-1}$ and of the past actions played by the players. In particular, a *(behavior) strategy* for player i in the extended game Γ^E is a measurable function $\xi_i\colon \left(\cup_{t=0}^{\infty}([0,1]\times A)^t\right) \times [0,1] \to \Delta(A_i)$. The payoff induced by a strategy profile $\xi = (\xi_i)_{i\in I}$ is

$$\gamma^E(\xi) := \mathbf{E}_\xi\left[\mathbf{1}_{\{t_*<\infty\}} u(a^{t_*})\right],$$

where $\mathbf{E}_\xi$ is the expectation w.r.t. the probability measure induced by ξ over the space of infinite histories $([0,1]\times A)^\infty$ in the game Γ^E. The strategy profile ξ is a *sunspot* ε*-equilibrium* in the game Γ if $\gamma_i^E(\xi) \geq \gamma_i^E(\xi_i', \xi_{-i}) - \varepsilon$, for every player $i \in I$ and every strategy ξ_i' of player i.

Solan and Solan (2018) proved that every positive recursive general quitting game admits a sunspot ε-equilibrium, for every $\varepsilon > 0$. Our main result in this section is that when at least two players have at least two continue actions, the game admits an ε-equilibrium, for every $\varepsilon > 0$.

Theorem 3.1. *Let $\Gamma = (I, (A_i^c)_{i\in I}, u)$ be a positive recursive general quitting game that satisfies $|I| \geq 2$, $|A_1^c| \geq 2$, and $|A_2^c| \geq 2$. Then for every $\varepsilon > 0$ the game admits an ε-equilibrium.*

To prove Theorem 3.1 we describe the structure of the sunspot ε-equilibrium constructed in Solan and Solan (2018). In that paper, it was proven that for every positive recursive general quitting game $\Gamma = (I, (A_i^c)_{i\in I}, u)$ there exists a mixed action profile $x = (x_i)_{i\in I} \in \times_{i\in I}\Delta(A_i)$ such that one of the following two alternatives holds for every $\varepsilon > 0$:

(A.1) The mixed action profile x is absorbing and a stationary 0-equilibrium.
(A.2) The mixed action profile x is nonabsorbing, and the game admits a sunspot ε-equilibrium ξ in which at every stage t the players play the mixed action profile x, except of possibly one player i^t, whose identity is determined by the correlation device, who plays the mixed action $(1-\eta^t)x_{i^t} + \eta^t Q_{i^t}$, where the random variable η^t has values in $(0, \varepsilon)$ and depends on the history before stage t and on i^t. Moreover, under ξ the play terminates with probability 1.

Thus, if Alternative (A.2) holds, then the players play mainly the stationary strategy profile x, and take turns in stopping the game: in each stage t the correlation device may designate one player i^t as the possible quitter, and that player stops the game with a history-dependent probability η^t. If the correlation device did not designate any player as the possible quitter, then all players follow x. The order in which the players are selected by the correlation device is random, depends on the device's past choices, and is crafted so as to keep incentive constraints.

In Alternative (A.2) statistical tests are conducted to ensure that the players do not deviate from the prescribed strategy profile: the players verify that the distribution of continue actions played by each player i is close to x_i. In Alternative (A.1) no statistical tests are required.

As mentioned above, in Alternative (A.2), at stage t the players play either the mixed action x or the mixed action $((1-\eta^t)x_{i^t} + \eta^t Q_{i^t}, x_{-i^t})$. Since payoffs are bounded by 1 and $\eta^t < \varepsilon$, if some player, say player j, quits, his payoff is ε-close to $u_j(Q_j, x_{-j})$, where $u_j(Q_j, x_{-j})$ is the multilinear extension of u_j. Since ξ is a sunspot ε-equilibrium, no player can gain more than ε by quitting. In particular, we can add a finite number of stages in which the players play the mixed action profile x without affecting the ε-equilibrium property of the strategy profile. In our construction, before *every* stage of the sunspot ε-equilibrium we add a large number of stages in which the players play the mixed action profile x, and these stages are used to implement a jointly controlled lottery by Players 1 and 2, which replaces the public correlation.

[3] The concept that we define is that of undiscounted ε-equilibrium. Theorem 3.1 below holds for the stronger notion of uniform ε-equilibrium as well.

Proof of Theorem 3.1. To prove the result we need to consider Alternative (A.2) only. Fix then $\varepsilon > 0$ and let ξ be a sunspot ε-equilibrium in the extended game Γ^E in which the players play mainly some nonabsorbing mixed action profile x. We will divide the play into blocks of random size; block t will correspond to stage t of the implementation of ξ. All stages of the block except the last one will be used to perform a jointly controlled lottery by Players 1 and 2, which will mimic the correlation device; that is, in this lottery Players 1 and 2 will select an element $i^t \in I \cup \{0\}$ according to a probability measure that is close to that indicated by ξ for stage t, where 0 will mean that no player is designated to quit. In the last stage of the block the players will play as ξ plays in stage t, given the outcome of the jointly controlled lottery conducted in that block.

Formally, for each $t \in \mathbf{N}$ denote by k^t the stage of the game in which block t starts, by $\widehat{a}^t$ the action profile that the players play in the last stage of block t (stage $k^{t+1}-1$), and by i^t the player who is selected by Players 1 and 2 in block t using the jointly controlled lottery mechanism of Theorem 2.6 (which will be described shortly in the context of the positive recursive general quitting game). Let $T_1 \in \mathbf{N}$ be sufficiently large such that

$$\mathbf{P}_\xi \left(\prod_{t=1}^{T_1} (1-\eta^t) > \varepsilon \right) < \varepsilon : \tag{4}$$

under the sunspot ε-equilibrium ξ, with probability at least $1-\varepsilon$, the play terminates before stage T_1 with high probability.

Assume first that both x_1 and x_2 are not pure. Let A_1 (resp. A_2) be the support of x_1 (resp. x_2). Let T_2 be sufficiently large such that for every probability measure ν over $J := I \cup \{0\}$, the jointly controlled lottery described in the proof of Theorem 2.6 for the $\frac{\varepsilon}{T_1}$-implementation of ν with the alphabets A_1 and A_2 and the probability measures $\mathbf{P}_1 = x_1$ and $\mathbf{P}_2 = x_2$ has length less than T_2.

Let σ be the following strategy profile in the positive recursive general quitting game Γ that plays in blocks. For every $t \in \mathbf{N}$, in block t the strategy profile is defined as follows.

(B.1) Consider the mechanism $M = (J, \tau^t, f)$ described in the proof of Theorem 2.6 that $\frac{\varepsilon}{T_1}$-implements the probability measure $\xi^t(\widehat{a}^1, \cdots, \widehat{a}^{t-1}, i^1, \cdots, i^{t-1})$, where $A_1 = \mathrm{supp}(x_1)$, $A_2 = \mathrm{supp}(x_2)$, $J = I \cup \{0\}$, $\mathbf{P}_1 = x_1$, and $\mathbf{P}_2 = x_2$.
In block t the players play the mixed action profile x until the game terminates (if some player quits) or until stage τ^t of the block (stage $k^t + \tau^t$ of the game). Note that the length of this phase is uniformly bounded by T_2, even if one player deviates from the play described herein. Denote by $\widehat{i}^t := f(a_1^{k^t}, a_2^{k^t}, \cdots, a_1^{k^t+\tau^t}, a_2^{k^t+\tau^t}) \in I \cup \{0\}$ the outcome of the jointly controlled lottery described in Theorem 2.6.

(B.2) In the last stage of the block, stage $k^t + \tau^t + 1$, the players follow the mixed action prescribed by the sunspot ε-equilibrium ξ at stage t, given the history $(\widehat{a}^1, \cdots, \widehat{a}^{t-1}, \widehat{i}^1, \cdots, \widehat{i}^t)$. Set $k^{t+1} := k^t + \tau^t + 1$.

We thus defined a strategy profile σ in the general quitting game Γ. By Eq. (4) and since the difference between the distribution of the jointly controlled lottery at each block t and $\xi(h^t)$ is at most $\frac{\varepsilon}{T_1}$, a standard coupling argument shows that $\|\gamma(\sigma) - \gamma^E(\xi)\|_\infty \le 3\varepsilon$; that is, the expected payoff under σ is 3ε-close to the expected payoff under ξ.

We argue that no player can profit more than 7ε by deviating to a pure strategy. Fix then a player $i \in I$ and a pure strategy σ_i' of that player. Using the strategy σ_i' we will define a strategy ξ_i' in the game with correlation device and show that $\gamma_i(\sigma_i', \sigma_{-i}) \le \gamma_i^E(\xi_i', \xi_{-i}) + 3\varepsilon$. Since ξ is a sunspot ε-equilibrium, it will follow that

$$\gamma_i(\sigma_i', \sigma_{-i}) \le \gamma_i^E(\xi_i', \xi_{-i}) + 3\varepsilon \le \gamma_i^E(\xi) + 4\varepsilon \le \gamma_i(\sigma) + 7\varepsilon,$$

as claimed.

Our goal now is to construct a strategy ξ_i' in the game with correlation device and prove that $\gamma_i(\sigma_i', \sigma_{-i}) \le \gamma_i^E(\xi_i', \xi_{-i}) + 3\varepsilon$. As described above, the strategy profile (σ_i', σ_{-i}) defines a partition of the stages $\mathbf{N}$ into blocks.[4] For each block t, the play defines an element $\widehat{i}^t \in I \cup \{0\}$ that indicates if some player has to quit with low probability, and if so, his identity, and an action profile $\widehat{a}^t \in A$, which is composed of the actions played at the last stage of block t by each player. Let ρ^t be the conditional probability that under (σ_i', σ_{-i}) player i quits during the first τ^t stages of block t, given $\widehat{i}^1, \cdots, \widehat{i}^{t-1}, \widehat{a}^1, \cdots, \widehat{a}^{t-1}$. For every action $a_i \in A_i$, let $\mu^t(a_i)$ be the conditional probability that under (σ_i', σ_{-i}) we have $\widehat{a}_i^t = a_i$, given $\widehat{i}^1, \cdots, \widehat{i}^{t-1}, \widehat{i}^t, \widehat{a}^1, \cdots, \widehat{a}^{t-1}$. Let ξ_i' be the strategy of player i that plays as follows at stage t:

- The quitting action Q_i is played with probability ρ^t.
- For each $a_i \in A_i$, the action a_i is played with probability $(1-\rho^t)\mu^t(a_i)$.

Since under ξ_{-i} and σ_{-i} the designated player quits with probability at most ε at every stage, it follows that $\|\gamma(\sigma_i', \sigma_{-i}) - \gamma^E(\xi_i', \xi_{-i})\| \le 3\varepsilon$, as claimed.

[4] In fact, the partition concerns only the stages up to the termination stage.

It is left to take care of the situation that one (or both) of the mixed actions x_1 or x_2 is pure. If the mixed action x_1 is pure, then, since $|A_1^c| \geq 2$, we can find a mixed action $x_1' \in \Delta(A_1^c)$ that is not pure and ε-close to x_1 in the l_∞-norm. A similar statement holds for x_2. In Step (B.1) we then change x_i by x_i' for each player $i \in \{1, 2\}$ whose mixed action x_i is pure. The only effect that this change has is that if a player quits, then his payoff changes by at most 2ε. Consequently the strategy profile described above is an 11ε-equilibrium. □

There are various ways in which one can strive to extend the equilibrium existence result.

- Our method shows that jointly controlled lotteries with biased coins enable one to transform sunspot ε-equilibria into undiscounted ε-equilibria, in various settings of stochastic games. Can one extend the existence result to other classes of stochastic games that include more than one nonabsorbing state? To the characterization of the set of equilibrium payoffs in discounted repeated games with imperfect monitoring?
- One property of the class of general quitting games is that some players have two actions that induce the same transitions, for every given action profile of the other players. Is it true that an undiscounted ε-equilibrium exists in any stochastic game in which for every state s, every player i and every action a_i of player i, there is an action $a_i' \neq a_i$ that yields the same transition as a_i at state s (for every action profile a_{-i} of the other players)?

4. Acknowledgment

The authors thank Johannes Hörner, Ehud Lehrer, Asaf Nachmias, Tristan Tomala, and Nicolas Vieille for useful discussions, and two anonymous referees whose comments improved the presentation. E. Solan acknowledges the support of the Israel Science Foundation, grant #217/17.

References

Aumann, R.J., Hart, S., 2003. Long cheap talk. Econometrica 71 (6), 1619–1660.
Aumann, R.J., Maschler, M.B., 1995. Repeated Games with Incomplete Information. MIT Press.
Farrell, J., Rabin, M., 1996. Cheap talk. J. Econ. Perspect. 10 (3), 103–118.
Flesch, J., Schoenmakers, G., Vrieze, K., 2008. Stochastic games on a product state space. Math. Oper. Res. 33, 403–420.
Flesch, J., Schoenmakers, G., Vrieze, K., 2009. Stochastic games on a product state space: the periodic case. Int. J. Game Theory 38, 263–289.
Flesch, J., Thuijsman, F., Vrieze, O.J., 1997. Cyclic Markov equilibrium in stochastic games. Int. J. Game Theory 26, 303–314.
Flesch, J., Thuijsman, F., Vrieze, O.J., 2007. Stochastic games with additive transitions. Eur. J. Oper. Res. 179, 483–497.
Forges, F., 1995. Non-zero sum repeated games and information transmission. In: Megiddo, Nimrod (Ed.), Essays in Game Theory: In Honor of Michael Maschler. Springer Verlag, pp. 65–95.
Forges, F., Koessler, F., 2008. Long persuasion games. J. Econ. Theory 143 (1), 1–35.
Gossner, O., Vieille, N., 2002. How to play with a biased coin. Games Econ. Behav. 41, 206–226.
Jaśkiewicz, A., Nowak, A.S., 2017. Non-zero-sum stochastic games. In: Handbook of Dynamic Games, vol. I.
Kalai, A.T., Kalai, E., Lehrer, E., Samet, D., 2010. A commitment folk theorem. Games Econ. Behav. 69 (1), 127–137.
Lehrer, E., 1996. Mediated talk. Int. J. Game Theory 25, 177–188.
Lehrer, E., Sorin, S., 1997. One-shot public mediated talk. Games Econ. Behav. 20, 131–148.
McLeish, D.L., 1974. Dependent central limit theorems and invariance principles. Ann. Probab. 2 (4), 620–628.
Simon, R.S., 2012. A topological approach to quitting games. Math. Oper. Res. 37, 180–195.
Simon, R.S., 2016. The challenge of non-zero-sum stochastic games. Int. J. Game Theory 45, 191–204.
Solan, E., 1999. Three-player absorbing games. Math. Oper. Res. 24, 669–698.
Solan, E., 2000. Absorbing team games. Games Econ. Behav. 31, 245–261.
Solan, E., Solan, O.N., 2017. Quitting games and linear complementarity problems. arXiv:1707.02598.
Solan, E., Solan, O.N., 2018. Sunspot equilibrium in general quitting games. arXiv:1803.00878.
Solan, E., Vieille, N., 2001. Quitting games. Math. Oper. Res. 26, 265–285.
Vieille, N., 2000a. Equilibrium in 2-person stochastic games I: a reduction. Isr. J. Math. 119, 55–91.
Vieille, N., 2000b. Equilibrium in 2-person stochastic games II: the case of recursive games. Isr. J. Math. 119, 93–126.