Approachability with Constraints*

Gaëtan Fournier[†], Eden Kuperwasser[‡], Orin Munk[§], Eilon Solan[¶] and Avishay Weinbaum[‖]

October 31, 2020

Abstract

We study approachability theory in the presence of constraints. Given a repeated game with vector payoffs, we study the pairs of sets (A, D) in the payoff space such that Player 1 can guarantee that the long-run average payoff converges to the set A, while the average payoff always remains in D. We provide a full characterization of these pairs when D is convex and open, and a sufficient condition when D is not convex.

Keywords: Game Theory, Approachability, Optimization

1 Introduction

Approachability theory, which was first introduced in Blackwell 1956, is an extension of the theory of zero-sum strategic-form games to the situation where the outcome is multidimensional. In a two-player repeated game in which the outcome is an n-dimensional vector, a target set A in $\mathbb{R}^n$ is *approachable* by Player 1 if she has a strategy that ensures that the long-run average payoff converges to the set, whatever strategy Player 2 uses. Blackwell 1956 provided a geometric condition that ensures that a set is approachable by Player 1. Hou 1971 and Spinat 2002 completed the characterization of approachable sets, by showing that a set is approachable only if it contains a set that satisfies Blackwell's geometric condition.

Approachability theory was used to study no regret with partial monitoring (see Perchet 2009, Perchet and Quincampoix 2014 and Lehrer and Solan 2016). In fact, approachability is equivalent to regret minimization and calibration (see Cesa-Bianchi

*The authors acknowledge the support of the COST Action 16228, the European Network of Game Theory. Munk, Solan, and Weinbaum acknowledge the support of the Israel Science Foundation, Grants #323/13 and #217/17. Fournier acknowledges the support through the ANR Labex IAST. This work was supported by the the French National Research Agency Grant 2017-EUR-20.

†Aix-Marseille Univ, CNRS, AMSE, France. 5-9 Boulevard Maurice Bourdet, 13001 Marseille, France. e-mail: gaetan.fournier@univ-amu.fr (Corresponding author)

‡The School of Mathematical Sciences, Tel Aviv University, Tel Aviv 6997800, Israel. e-mail: kuperwasser@mail.tau.ac.il

§The School of Mathematical Sciences, Tel Aviv University, Tel Aviv 6997800, Israel. e-mail: orin25@gmail.com

¶The School of Mathematical Sciences, Tel Aviv University, Tel Aviv 6997800, Israel. e-mail: eilons@post.tau.ac.il

‖The School of Mathematical Sciences, Tel Aviv University, Tel Aviv 6997800, Israel. e-mail: avishay.weinbaum@gmail.com

and Lugosi 2006 and Perchet 2014). The theory was also used to study continuous-time network flows with capacity and unknown demand (Bauso, Blanchini, et al. 2010), production-inventory problems in both discrete and continuous-time (Khmelnitsky and Tzur 2004), and to construct normal numbers (Lehrer 2004). The geometric principle that lies behind approachability theory has been studied by Lehrer 2002, the rate of convergence to the target set was studied in Mannor and Perchet 2013, and variants of the basic notion of approachability have been studied by, e.g., Vieille 1992, Lehrer and Solan 2009, Shani and Solan 2014, Mannor, Stoltz, et al. 2014, and Bauso, Lehrer, et al. 2015.

In various situations, in addition to having a target set, the player has constraints that have to be satisfied. For example, an investment firm makes daily investment decisions and may have various goals, like maximizing the value of its portfolio, keeping the value of its portfolio higher than the value of other investment firms, and attracting investors. The firm may also have various constraints, like keeping its Sharpe ratio above a certain level, or its conditional value at risk below a certain threshold, see, e.g., Uziel and El-Yaniv 2018. A second example concerns fellowships obtained by students in various universities. In every quarter or semester the student has to dedicate her time to several activities, like studying, extracurricular classes, and spending time with friends and family. In addition, to keep receiving the fellowship, the student must keep her average grade above a certain level. A third example concerns decision problems like expert selection and regret minimization, see Chapter 13 in Maschler et al. 2013. In these applications, constraints arise naturally. For example, in expert selection, the decision maker may be required to select experts from each gender with some minimal frequency, and regret minimization with constraints has been studied by, e.g., Mannor, Tsitsiklis, et al. 2009, Mahdavi et al. 2012, Sadeghi and Fazel 2020, Liakopoulos et al. 2019, and Du and Lehrer 2020.

The goal of this paper is to study approachability in the presence of constraints. Specifically, we consider a two-player repeated game with vector payoffs, and are given two sets A and D in the payoff space. The set A is the target set that Player 1 would like to approach, and the set D represents the set of allowable average payoffs: after every history that occurs with positive probability, the average payoff must be in D. We call this problem *approaching A while remaining in D*.

If the set A is approachable by Player 1 while remaining in the set D, then as we prove below the set $\overline{A} \cap \overline{D}$ is approachable by Player 1, where $\overline{A}$ and $\overline{D}$ are the closures of A and D, respectively. Since the outcome after the first stage must be in D, a second necessary condition is that Player 1 has a safe action; that is, an action s such that when Player 1 plays s the outcome is in D, whatever Player 2 plays. We will show that the first condition, together with a stronger version of the second condition, is sufficient for approachability with constraints: when D is convex, if the set $\overline{A} \cap \overline{D}$ is approachable by Player 1, and if there is an action s of Player 1 that ensures that the outcome is in the interior of the closure of D, then the set A is approachable by Player 1 while remaining in D. When the set D is open, the necessary conditions coincide with the sufficient conditions, and we obtain a characterization of the pairs of sets (A, D) such that A is approachable by Player 1 while remaining in D.

As we will prove, to approach the set A while remaining in D, Player 1 can follow a strategy that approaches the set $\overline{A} \cap \overline{D}$, and, whenever the average outcome becomes close to the boundary of D, she should play a safe action, which ensures that the average outcome gets farther away from the boundary of D. We will show that under this strategy, the number of stages up to stage t in which the average outcome is close to the boundary of D is of the order $\sqrt{t}$, hence this strategy indeed approaches A while remaining in D. We moreover show that the rate of convergence of the average payoff to the set A in our setup is the same as the rate of convergence given by Blackwell 1956 or Mannor and Perchet 2013, so that the presence of constraints does not slow down the rate of convergence to A.

We then study the case where the set D is not convex, provide two sufficient conditions that guarantee that the set A is approachable while remaining in D, and provide an example that shows that the conditions are not necessary. This example exhibits the difficulty in providing a general characterization of the pairs of sets (A, D) such that A is approachable by Player 1 while remaining in D.

In Section 2 we define the model and the concept of approachability while remaining in a set and state our main result. In Section 3 we prove Proposition 3. This Proposition and Proposition 2 directly imply our main result, Theorem 4. Section 4 is devoted to the case in which the set D is not convex. In Section 5 we discuss the effect of constraints on variants of approachability, like weak approachability and strong approachability, and present open problems.

2 The Model and the Main Result

A two-player *repeated game with vector payoffs* is a triplet (I, J, U), where I and J are finite sets of actions for the two players, and $U = I \times J \to \mathbb{R}^n$ is a vector payoff matrix. We assume w.l.o.g. that payoffs are nonnegative and bounded by 1, that is $0 \leq U_k(i, j) \leq 1$ for every $i \in I$, every $j \in J$, and every $1 \leq k \leq n$. To eliminate trivial cases we assume that both players have at least two actions: $|I| \geq 2$ and $|J| \geq 2$. The bilinear extension of U to $\Delta(I) \times \Delta(J)$ is still denoted by U, where $\Delta(A)$ is the space of probability distributions over $A = I, J$.

At every stage $t \geq 1$, Player 1 (resp. Player 2) chooses an action $i_t \in I$ (resp. $j_t \in J$), which is observed by the other player.[1] A *history* of length t is a sequence $h_t = (i_1, j_1, i_2, j_2, \cdots, i_t, j_t) \in (I \times J)^t$ for some $t \geq 0$. The empty history is denoted $\emptyset$. We denote by $H := \cup_{t=0}^{\infty} (I \times J)^t$ the set of all histories. When $h_t = (i_1, j_1, i_2, j_2, \cdots, i_t, j_t) \in H$ and $t' \leq t$ we denote by $h_{t'} = (i_1, j_1, i_2, j_2, \cdots, i_{t'}, j_{t'})$ the prefix of h_t of length t'.

We assume perfect recall, and consequently by Kuhn's Theorem we can restrict attention to behavior strategies. A (behavior) *strategy* for Player 1 (resp. Player 2) is a function $\sigma : H \to \Delta(I)$ (resp. $\tau : H \to \Delta(J)$). We denote by $\mathcal{S}$ and $\mathcal{T}$ the strategy spaces of Players 1 and 2, respectively.

The set F of *feasible payoffs* is the convex hull of possible one stage payoffs, that is,

$$F := conv\{U(i, j) \colon (i, j) \in I \times J\} \subset \mathbb{R}^n.$$

The set of plays is $H^\infty := (I \times J)^\infty$. This set, when supplemented with the sigma-algebra generated by all finite cylinders, is a measurable space. Every pair of strategies $(\sigma, \tau) \in \mathcal{S} \times \mathcal{T}$ defines a probability distribution $P_{\sigma,\tau}$ over H^∞. We denote by $E_{\sigma,\tau}$ the expectation with respect to this probability distribution.

The average vector payoff up to stage $t \geq 1$ is

$$g_t := \frac{1}{t} \sum_{l=1}^{t} U(i_l, j_l).$$

Note that for every $t \geq 1$, the average vector payoff g_t is a random variable with values in $\mathbb{R}^n$, whose distribution is determined by the strategies of both players. When we

[1] As in Blackwell 1956, for our results it is sufficient to assume that the players observe the outcome $U(i_t, j_t)$ at every stage t.

wish to calculate the average payoff up to stage t along a history $h_t \in H$ we use the notation $g_t(h_t)$.

Let $d(x,y) := ||x-y||_2$ denote the Euclidean distance between the points x and y in $\mathbb{R}^n$.

Blackwell 1956 defined the concept of approachable sets in repeated games with vector payoffs. A subset $A \in \mathbb{R}^n$ is *approachable* by Player 1 if there exists a strategy $\sigma \in \mathcal{S}$ such that for every $\epsilon > 0$ there exists an integer $T \geq 1$ such that for every strategy $\tau \in \mathcal{T}$ of Player 2 we have

$$P_{\sigma,\tau}\Big[\forall t \geq T, d(g_t, A) < \epsilon\Big] > 1-\epsilon. \tag{1}$$

We say that the strategy σ *approaches* the set A. This paper concerns the concept of approachability with constraints, which is defined as follows.

Definition 1. *Let A and D be two subsets of $\mathbb{R}^n$. The set A is* approachable by Player 1 while remaining in the set D *if there exists a strategy $\sigma \in \mathcal{S}$ such that for every $\epsilon > 0$ there exists an integer $T \geq 1$ such that for every strategy $\tau \in \mathcal{T}$ of Player 2, we have*

$$P_{\sigma,\tau}\Big[\forall t \geq T,\ d(g_t, A) < \epsilon\Big] \quad > \quad 1-\epsilon, \tag{2}$$

$$P_{\sigma,\tau}\Big[\forall t \geq 1,\ g_t \in D\Big] \quad = \quad 1. \tag{3}$$

Condition (4) ensures that the strategy σ approaches the set A. Condition (3) is concerned with the constraints: when playing σ, Player 1 guarantees that the sequence of realized average payoffs always remains in the set D. Our main goal is the characterization of the pairs of sets (A, D) such that A is approachable by Player 1 while remaining in D.

For every mixed action $p \in \Delta(I)$ define

$$R_1(p) := \{U(p,q) \colon q \in \Delta(J)\} = \mathrm{conv}\{U(p,j) \colon j \in J\}.$$

This is the set of all possible expected outcomes when Player 1 plays the mixed action p. For every set X in a Euclidean space we denote by $\overline{X}$ the closure of X.

The following proposition lists two necessary conditions to approaching A while remaining in D.

Proposition 2. *Let A and D be two subsets of $\mathbb{R}^n$. If the set A is approachable by Player 1 while remaining in the set D, then the following two conditions hold.*

(C1) The set $\overline{A} \cap \overline{D}$ is approachable by Player 1.

(C2) There exists an action $s \in I$ such that for every action $j \in J$ we have $U(s,j) \in D$.

An action $s \in I$ that satisfies Condition (C2) is termed a *safe action*, since it ensures that the stage payoff is in D.

Proof. To see that Condition (C1) is necessary, we recall the following equivalent definition of approachability, see Hou (1971): A set A is approachable by Player 1 if, and only if, there is a strategy σ for Player 1 such that

$$P_{\sigma,\tau}[\lim_{t\to\infty} d(g_t, \overline{A}) = 0] = 1 \tag{4}$$

for every strategy τ for Player 2. To see that the two definitions are indeed equivalent, note that if Eq. (4) holds then for every $\epsilon > 0$ there is a $P_{\sigma,\tau}$-a.s. finite stopping time

τ_ϵ such that $P_{\sigma,\tau}[\forall t \geq \tau_\epsilon, d(g_t, \overline{A}) \leq \epsilon] = 1$, hence Eq. (1) holds with T such that $P_{\sigma,\tau}[\tau_\epsilon \leq T] \geq 1 - \epsilon$. Conversely, if Eq. (4) does not hold, then there exists $\delta > 0$ such that $P_{\sigma,\tau}[\limsup_{t\to\infty} d(g_t, \overline{A}) \geq \delta] \geq \delta$, and then Eq. (1) does not hold for $\epsilon < \delta$.

Suppose that the set A is approachable by Player 1 while remaining in the set D, and let σ be a strategy for Player 1 that guarantees that the average payoff converges to A while remaining in D. In particular, the strategy σ approaches the set A, hence by the equivalent definition of approachability, for every strategy τ for Player 2 we have $P_{\sigma,\tau}[\lim_{t\to\infty} d(g_t, \overline{A}) = 0] = 1$. Moreover, $P_{\sigma,\tau}[\forall t \geq 1, g_t \in D \subseteq \overline{D}] = 1$, hence $P_{\sigma,\tau}[\lim_{t\to\infty} d(g_t, \overline{A} \cap \overline{D}) = 0] = 1$. By the equivalent definition of approachability once again, this implies that the set $\overline{A} \cap \overline{D}$ is approachable by Player 1.

We now argue that Condition (C2) is necessary. Denote by s any action that the strategy σ plays with positive probability at the first stage. Let τ be any strategy for Player 2 that plays all actions in J with positive probability. Under the strategy pair (σ, τ) the probability that the action pair (s, j) is played is positive, and therefore Eq. (3) implies that $U(s, j) \in D$, for every $j \in J$. □

The next proposition asserts that Condition (C1) together with a variation of Condition (C2) are sufficient for approachability with constraints, provided the set D is convex.

Proposition 3. *Let A be a subset of $\mathbb{R}^n$ and let D be a convex subset of $\mathbb{R}^n$. If Condition (C1) and the following Condition (C2') hold, then Player 1 can approach A while remaining in D:*

(C2') There exists an action $s \in I$ such that for every action $j \in J$ the vector $U(s, j)$ is in the interior of the closure of D.

We note that if Condition (C2') holds, then, provided D is convex, we have $R_1(s) \subseteq D$ and, moreover, $\delta := d(F \setminus D, R_1(s)) > 0$. Section 3 is devoted to the proof of Proposition 3.

As a conclusion from Propositions 2 and 3 we obtain our main result, which is a characterization of the pairs of sets (A, D) such that the set A is approachable while remaining in D, and it is valid whenever the set D is open and convex.

Convexity is a natural assumption, as often constraints have the form of linear inequalities. The requirement that the set D is open means that these inequalities should be strict.

Theorem 4. *Let A be a subset of $\mathbb{R}^n$ and let D be an open and convex subset of $\mathbb{R}^n$. Player 1 can approach A while remaining in D if and only if Conditions (C1) and (C2) hold.*

The following example shows that when D is not open, Conditions (C1) and (C2) are not sufficient to imply that A is approachable while remaining in D.

Example 5. *Consider the game that is depicted in Figure 1 with $A = \{(0, 0)\}$ and $D = \{(x, y) \in \mathbb{R}^2 : x \geq 0\}$. For convenience we consider, in this example and in the following ones, payoffs that do not necessarily belong to the interval $[0, 1]$.*

The set $A = \overline{A} \cap \overline{D}$ is approachable by Player 1, for example, by the stationary strategy $[\frac{1}{2}(B_1), \frac{1}{2}(B_2)]$, and therefore Condition (C1) holds. The action T is a safe action for Player 1, and therefore Condition (C2) holds, yet since $(0, 1)$ is on the boundary of D, Condition (C2') does not hold.

We argue that the set A is not approachable by Player 1 while remaining in D. Indeed, to approach A Player 1 has to play one of the action B_1 and B_2 at least once (in fact, with probability 1 she should play these actions infinitely often). Suppose that

	L	R
T	$(0,1)$	$(0,1)$
B_1	$(1,0)$	$(-1,0)$
B_2	$(-1,0)$	$(1,0)$

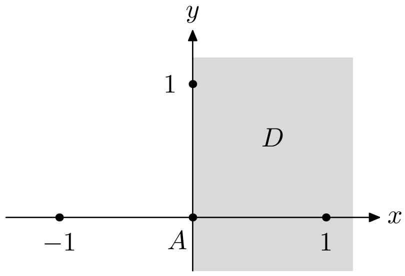


Figure 1: The payoff matrix in Example 5 and the payoff space.

Player 2 plays the stationary strategy $[\frac{1}{2}(L), \frac{1}{2}(R)]$*. Then in the first stage in which Player 1 plays* B_1 *or* B_2*, there is a probability* $\frac{1}{2}$ *that the outcome is* $(-1,0)$*, in which case the first coordinate of the average payoff is negative, so that the average payoff is not in* D*.*

Example 12 in Section 4 shows that when D is not convex, Conditions (C1) and (C2) are not sufficient to imply that A is approachable while remaining in D.

3 The Proof of Proposition 3

This section is devoted to the proof of Proposition 3. In particular, the set D is assumed to be convex.

3.1 B-Sets and Blackwell's Characterization of Approachable Sets

Blackwell 1956 provided a geometric characterization for approachable sets (without constraints). The basic concept that Blackwell used was that of B-sets. In this subsection we review the definition of B-sets and Blackwell's characterization.

A *hyperplane* in $\mathbb{R}^n$ is any set of the form $\mathcal{H} := \{x \in \mathbb{R}^n \colon \sum_{k=1}^n \alpha^k x^k = \beta\}$, where $\alpha^1, \cdots, \alpha^k, \beta \in \mathbb{R}$. For every hyperplane $\mathcal{H}$ we denote by $\mathcal{H}^+ := \{x \in \mathbb{R}^n \colon \sum_{k=1}^n \alpha^k x^k \geq \beta\}$ and $\mathcal{H}^- := \{x \in \mathbb{R}^n \colon \sum_{k=1}^n \alpha^k x^k \leq \beta\}$. These are the two half-spaces defined by $\mathcal{H}$.

Definition 6. *A set* $A \subset \mathbb{R}^n$ *is a* B-set *for Player 1 if for every point* $x \in F \setminus A$ *there exists a point* $y \in \overline{A}$ *and a mixed action* $p \in \Delta(I)$ *such that (a)* y *minimizes the distance to* x *among the points in* $\overline{A}$*, and (b) there is a hyperplane* $\mathcal{H}$ *that is (i) perpendicular to the line that connects* x *to* y*, (ii) passes through* y*, and (iii) separates* x *from* $R_1(p)$*, that is,* $x \in \mathcal{H}^-$ *and* $R_1(p) \subseteq \mathcal{H}^+$*, or* $x \in \mathcal{H}^+$ *and* $R_1(p) \subseteq \mathcal{H}^-$*.*

Blackwell 1956 proved that every B-set is approachable. Hou 1971 and Spinat 2002 proved that every approachable set contains a B-set.

3.2 A Definition of a Strategy σ^*.

We now define a strategy σ^* for Player 1 that, as will be shown later, approaches A while remaining in D. This strategy is based on two components: a strategy $\widehat{\sigma}$ that approaches the set $\overline{A} \cap \overline{D}$, and a set H_{unsafe} of histories that we now define. Roughly, the strategy σ^* follows the strategy $\widehat{\sigma}$ that approaches the set $\overline{A} \cap \overline{D}$, but whenever the average payoff gets close to the boundary of D, it plays the safe action, ensuring that the average payoff gets farther away from the boundary.

Let $H_{\text{unsafe}} \subseteq H$ be the set of all histories $h_t \in H$ that satisfy $d(g_t(h_t), F \setminus D) \leq \frac{\sqrt{n}}{t}$. We also add the empty history $\emptyset$ to H_{unsafe}. Such histories are called *unsafe histories*. The complement $H_{\text{safe}} := H \setminus H_{\text{unsafe}}$ contains all histories $h_t \in H$ that satisfy $d(g_t(h_t), F \setminus D) > \frac{\sqrt{n}}{t}$. Such histories are called *safe histories*. As we will show in the proof of Lemma 7 below, because payoffs are bounded by 1, whatever Player 1 plays at stage $t+1$ after a history $h_t \in H_{\text{safe}}$, the average payoff up to stage $t+1$ will be in D.

For every history $h_t \in H$ denote by $\varphi(h_t)$ the history where we only keep actions played at stages $t' \leq t$ that follow a safe history (i.e., $h_{t'-1} \in H_{\text{safe}}$); actions at stages $t' \leq t$ that follow an unsafe history (i.e., $h_{t'-1} \in H_{\text{unsafe}}$) are deleted. Formally, we define φ recursively as follows.

$$\varphi(\emptyset) := \emptyset, \tag{5}$$

$$\varphi(h_t, i_{t+1}, j_{t+1}) := \begin{cases} \varphi(h_t) & \text{if } h_t \in H_{\text{unsafe}}, \\ \varphi(h_t) \circ (i_{t+1}, j_{t+1}) & \text{if } h_t \in H_{\text{safe}}, \end{cases} \tag{6}$$

where $\varphi(h_t) \circ (i_{t+1}, j_{t+1})$ is the concatenation of the history $\varphi(h_t)$ with the action pair (i_{t+1}, j_{t+1}).

We turn to the definition of the strategy σ^*. Let σ^* be the strategy of Player 1 that plays the safe action s whenever $h_t \in H_{\text{unsafe}}$, and plays the strategy $\widehat{\sigma}$ that approaches $\overline{A} \cap \overline{D}$ whenever $h_t \in H_{\text{safe}}$: after a history $h_t \in H_{\text{safe}}$, when determining which mixed action she should play, Player 1 ignores past stages that followed unsafe histories, and follows $\widehat{\sigma}$ as if past play is $\varphi(h_t)$. Formally,

$$\sigma^*(h_t) := \begin{cases} s & \text{if } h_t \in H_{\text{unsafe}}, \\ \widehat{\sigma}(\varphi(h_t)) & \text{if } h_t \in H_{\text{safe}}. \end{cases}$$

3.3 The strategy σ^* approaches A while remaining in D

The proof that the strategy σ^* approaches A while remaining in D is done in four steps.

1. In Lemma 7 we prove that when Player 1 plays σ^*, the average payoff always remains in D, whatever Player 2 plays.
2. In Lemma 8 we prove that when Player 1 plays σ^*, the frequency of stages at which the realized history is unsafe goes to 0.
3. In Lemma 9 we prove a geometric inequality used in the proof of Lemma 8.
4. In Lemma 10 we prove that the strategy σ^* approaches the set A.

Lemma 7. *For every strategy τ of Player 2, we have*

$$P_{\sigma^*,\tau}\left[\exists t \geq 1,\ g_t \notin D\right] = 0.$$

The properties that are needed to prove Lemma 7 are that (a) after histories in H_{unsafe} the strategy σ^* plays a safe action s, and (b) the set D is convex.

Proof. Fix a strategy τ of Player 2. For every history $h_t \in H$, denote by $P_{\sigma^*,\tau}[h_t]$ the probability that, under the strategy pair (σ^*, τ), the realized history of length t is h_t. We will prove by induction on t that for every history $h_t \in H$ that satisfies $P_{\sigma^*,\tau}[h_t] > 0$, we have $g_t(h_t) \in D$.

Since at the first stage the strategy σ^* plays a safe action, the claim holds for $t = 1$. Assume then by induction that the claim holds for $t-1$, and let h_t be a history of length t with $P_{\sigma^*,\tau}[h_t] > 0$. In particular, h_{t-1}, the prefix of h_t of length $t-1$, satisfies $P_{\sigma^*,\tau}[h_{t-1}] > 0$, and consequently, by the induction hypothesis, we have $g_{t-1}(h_{t-1}) \in D$.

If $h_{t-1} \in H_{\text{unsafe}}$ then at stage t the strategy σ^* plays a safe action. Since the set D is convex, since $g_{t-1}(h_{t-1}) \in D$, and since $R_1(s) \subseteq D$, it follows that $g_t(h_t) \in D$.

If $h_{t-1} \in H_{\text{safe}}$ then $d(g_{t-1}(h_{t-1}), F \setminus D) > \frac{\sqrt{n}}{t-1}$. Because payoffs are in $[0,1]^n$, we have that $d(g_{t-1}(h_{t-1}), g_t(h_t)) \le \frac{\sqrt{n}}{t}$. This implies that $d(g_t(h_t), F \setminus D) > 0$, so that $g_t(h_t) \in D$ as well. $\square$

Denote by $f(h_t)$ the number of times along the history h_t in which the subhistory $h_{t'}$ is unsafe for $t' < t$, that is,

$$f(h_t) := \#\{0 \le t' < t \colon h_{t'} \in H_{\text{unsafe}}\}.$$

For every history $h_t \in H$, we denote the average payoff during the stages in which the subhistory is unsafe by

$$\alpha_t := \frac{1}{f(h_t)} \sum_{0 \le l < t, h_l \in H_{\text{unsafe}}} U(i_{l+1}, j_{l+1}) = \frac{1}{f(h_t)} \sum_{0 \le l < t, h_l \in H_{\text{unsafe}}} U(s, j_{l+1}),$$

and the average payoff up to stage t during the stages in which the history is safe by

$$\beta_t := \frac{1}{t - f(h_t)} \sum_{0 \le l < t, h_l \in H_{\text{safe}}} U(i_{l+1}, j_{l+1}).$$

Note that $\alpha_t \in R_1(s)$, so in particular $d(\alpha_t, F \setminus D) \ge \delta > 0$, where δ was defined to be $d(F \setminus D, R_1(s))$. Note also that

$$g_t = \frac{1}{t} \sum_{l=1}^{t} U(i_l, j_l) = \frac{f(h_t)}{t} \alpha_t + \frac{t - f(h_t)}{t} \beta_t. \tag{7}$$

The following result, which provides a uniform upper bound on $f(h_t)$, implies that with high probability, the frequency of stages at which the realized history is unsafe goes to 0. Recall that δ was defined to be $d(F \setminus D, R_1(s))$

Lemma 8. *For every $\epsilon > 0$ there exists a constant $c' > 0$ such that for every strategy τ of Player 2 we have*

$$P_{\sigma^*,\tau}\left[\forall t \ge 1,\ f(h_t) \le \tfrac{c'\sqrt{t}}{\delta}\right] \ge 1 - \epsilon.$$

To prove Lemma 8 we need the following technical result.

Lemma 9. *For every $x \in D$, every $y \in \mathbb{R}^n$, and every $\lambda \in [0,1]$, we have*

$$d(\lambda x + (1-\lambda) y, F \setminus D) \ge \lambda d(x, F \setminus D) - (1-\lambda) d(y, D)$$

Proof. **Step 1:** Definitions.

Define two continuous functions $g_1, g_2 : [0,1] \to \mathbb{R}$ by

$$g_1(\lambda) := d(\lambda x + (1-\lambda) y, F \setminus D), \quad \forall \lambda \in [0,1],$$

and

$$g_2(\lambda) := \lambda d(x, F \setminus D) - (1-\lambda) d(y, D).$$

Since the set D is open, $g_1(1) = g_2(1) = d(x, F \setminus D) > 0$. Define

$$\lambda_0 := \inf\{\lambda \in [0,1] \colon g_1(\lambda) > 0\}.$$

The intuition of the proof of the lemma is as follows. The point $\lambda x + (1-\lambda)y$ is in D whenever $\lambda \in [0, \lambda_0)$, and outside D whenever $\lambda \in (\lambda_0, 1]$. By definition, $g_1(\lambda) = 0$ for every $\lambda \in [0, \lambda_0]$. The convexity of D will imply that the function g_1 is concave on $(\lambda_0, 1]$ (Step 2) and that $g_2(\lambda) \le 0$ for every $\lambda \in [0, \lambda_0)$ (Step 3). This will imply that $g_1(\lambda) = 0 \ge g_2(\lambda)$ for every $\lambda \in [0, \lambda_0]$. Since $g_1(1) = g_2(1)$ while $g_1(\lambda_0) \ge g_2(\lambda_0)$, it will then follow that $g_1(\lambda) \ge g_2(\lambda)$ for every $\lambda \in [\lambda_0, 1]$.

Step 2: The function g_1 is concave on $(\lambda_0, 1]$.

The claim holds since the set D is convex. Indeed, suppose that $\lambda, \lambda' \in (\lambda_0, 1]$, $g_1(\lambda) = c > 0$, and $g_1(\lambda') = c' > 0$. This implies that $d(\lambda x + (1-\lambda)y, F \setminus D) = c$ and $d(\lambda' x + (1-\lambda')y, F \setminus D) = c'$. Consequently, $B(\lambda x + (1-\lambda)y, c) \subseteq D$ and $B(\lambda' x + (1-\lambda')y, c') \subseteq D$, where $B(z, r)$ is the open ball around z with radius r, for every $z \in \mathbb{R}^n$ and every $r \ge 0$. It follows that $B(\lambda'' x + (1-\lambda'')y, c'') \subseteq D$, where $\lambda'' := \frac{1}{2}\lambda + \frac{1}{2}\lambda'$ and $c'' := \frac{1}{2}c + \frac{1}{2}c'$. Therefore $g_1(\lambda'') = d(\lambda'' x + (1-\lambda'')y, F \setminus D) \ge c''$. The function g_1 is therefore mid-point concave and continuous, hence concave.

Step 3: $g_2(\lambda) \le 0$ for every $\lambda \in [0, \lambda_0)$.

The claim holds trivially whenever $\lambda_0 = 0$. We therefore assume that $\lambda_0 > 0$, so in particular y is not in D. We will show that $g_2(\lambda_0) \le 0$. The result for every $\lambda \in [0, \lambda_0)$ will follow since the function g_2 is monotone increasing.

Set $q := \lambda_0 x + (1-\lambda_0)y$ (see Figure 2). Then q lies on the boundary of the set D. Since the set D is convex, so is its closure $\overline{D}$. Hence there is a unique closest point to y in $\overline{D}$, denoted z. Denote by θ the angle between the line segment $[y, z]$ and the line segment $[y, x]$. If $\theta > 0$, denote by w the intersection point of the half line $[z, q]$, and the half line that starts at x, lies on the plane defined by x, y, and z, and has angle θ relative to the line segment $[x, y]$. Since D is convex, since the points q and z are on the boundary of D, and since q lies on the line segment $[w, z]$, we have $w \in \overline{F \setminus D}$. The triangles (x, w, q) and (y, z, q) are similar, hence $(1-\lambda_0)d(y, z) = \lambda_0 d(x, w)$. If $\theta = 0$, define $w := z$, and then $(1-\lambda_0)d(y, z) = \lambda_0 d(x, w)$ holds as well. We conclude that

$$(1-\lambda_0)d(y, D) = (1-\lambda_0)d(y, z) = \lambda_0 d(x, w) \ge \lambda_0 d(x, F \setminus D),$$

where the last inequality follows from the fact that w is in $\overline{F \setminus D}$. Consequently, $g_2(\lambda_0) \le 0$ as desired.

Step 4: $g_1(\lambda) \ge g_2(\lambda)$ for every $\lambda \in [0, 1]$.

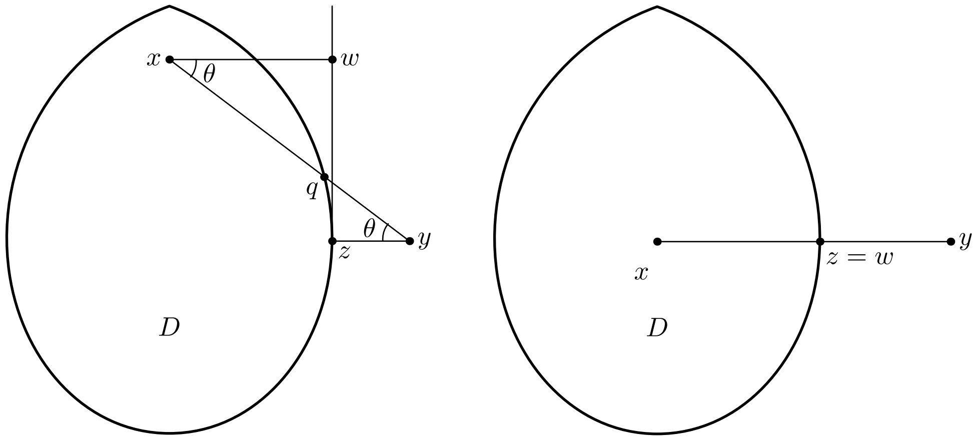
For $\lambda \in [0, \lambda_0)$ we have by Step 3

$$g_1(\lambda) = 0 \ge g_2(\lambda).$$

By the continuity of g_1 and g_2 this inequality extends to $\lambda = \lambda_0$. By Step 2 the function g_1 is concave on $(\lambda_0, 1]$ and by its definition the function g_2 is linear on this interval. Since $g_1(1) = g_2(1)$ while $g_1(\lambda_0) \ge g_2(\lambda_0)$, it follows that $g_1(\lambda) \ge g_2(\lambda)$ for every $\lambda \in [\lambda_0, 1]$.

□

Proof of Lemma 8. It is sufficient to prove the claim for histories $h_t \in H$ that satisfy $f(h_t) \ge 2$. Fix $\epsilon > 0$, a strategy τ of Player 2, and a history $h_t \in H$ that satisfies $f(h_t) \ge 2$. Denote by $m = m(h_t) \in \{1, 2, \cdots, t\}$ the last stage along h_t such that $h_m \in H_{\text{unsafe}}$. Note that if $m < t$ then $f(h_t) = f(h_m) + 1$, while if $m = t$ then



The case $\theta > 0$ The case $\theta = 0$

Figure 2: The construction in the proof of Lemma 9.

$f(h_t) = f(h_m)$. Since $h_m \in H_{\text{unsafe}}$ we have $d(g_m, F \setminus D) \leq \frac{\sqrt{n}}{m}$. By Eq. (7) and Lemma 9 we deduce that

$$\begin{aligned} \frac{\sqrt{n}}{m} &\geq d(g_m, F \setminus D) \\ &= d\left(\frac{f(h_m)}{m}\alpha_m + \frac{m-f(h_m)}{m}\beta_m, F \setminus D\right) \\ &\geq \frac{f(h_m)}{m} d(\alpha_m, F \setminus D) - \frac{m-f(h_m)}{m} d(\beta_m, D). \end{aligned} \tag{8}$$

Since $\alpha_m \in R_1(s)$ we have $d(\alpha_m, F \setminus D) \geq \delta$. Since β_m is the average payoff when playing the approachability strategy for $m - f(h_m)$ stages, there is a constant $c \geq \sqrt{n}$, such that[2]

$$P_{\sigma^*,\tau}\left[\forall t \geq 1,\ d(\beta_m, D) \leq \frac{c}{\sqrt{m-f(h_m)}}\right] \geq 1 - \epsilon. \tag{9}$$

Together with Eq. (8) this implies that on the event $\left\{\forall t \geq 1,\ d(\beta_m, D) \leq \frac{c}{\sqrt{m-f(h_m)}}\right\}$ we have

$$\begin{aligned} \frac{c}{m} \geq \frac{\sqrt{n}}{m} &\geq \frac{f(h_m)}{m}\delta - \frac{m-f(h_m)}{m} \cdot \frac{c}{\sqrt{m-f(h_m)}} \\ &= \frac{f(h_m)}{m}\delta - \frac{c\sqrt{m-f(h_m)}}{m} \qquad (10) \\ &\geq \frac{f(h_m)}{m}\delta - \frac{c\sqrt{m}}{m}, \qquad (11) \end{aligned}$$

which solves to $f(h_m) \leq \frac{c(1+\sqrt{m})}{\delta}$. Consequently, on this event

$$f(h_t) \leq f(h_m) + 1 \leq \frac{c(1+\sqrt{m})}{\delta} + 1 \leq \frac{c(2+\sqrt{t})}{\delta},$$

where the first inequality holds by the choice of m and the last inequality holds since $m \leq t$. Therefore, on an event of probability larger than $1-\epsilon$, we have that $f(h_t) \leq \frac{c'\sqrt{t}}{\delta}$ where $c' = 3c$, and the result follows. □

[2]To properly interpret Eq. (9) and the event $\{\forall t \geq 1,\ d(\beta_m, D) \leq \frac{c}{\sqrt{m-f(h_m)}}\}$, recall that m depends on t.

We now complete the proof of Theorem 4 by showing that the strategy σ^* approaches the set $\overline{A} \cap \overline{D}$, and, in particular, approaches the set A. This result holds since in most stages Player 1 plays a strategy that approaches the set $\overline{A} \cap \overline{D}$.

Lemma 10. *The strategy σ^* approaches the set $\overline{A} \cap \overline{D}$.*

Proof. Fix $\epsilon > 0$. Since the strategy $\widehat{\sigma}$ approaches the set $\overline{A} \cap \overline{D}$, there is $T_0 \in \mathbb{N}$ such that for every strategy τ of Player 2,

$$P_{\widehat{\sigma},\tau}\Big[\forall t \geq T_0,\ d(g_t, \overline{A} \cap \overline{D}) < \epsilon\Big] > 1 - \epsilon. \tag{12}$$

Consider an outside observer who observes the play only at stages t such that $h_{t-1} \in H_{\text{safe}}$. This observer is not aware of the stages t such that $h_{t-1} \in H_{\text{unsafe}}$ and from her point of view Player 1 follows the strategy $\widehat{\sigma}$. Let c' be the constant of Lemma 8, and let $T_1 \geq 1$ be sufficiently large so that $T_1 - \frac{c'\sqrt{T_1}}{\delta} \geq T_0$. By Lemma 8, with high probability the number of stages which the observer missed up to stage T_1 is at most $\frac{c'\sqrt{T_1}}{\delta}$. Hence, up to stage T_1 of the actual game, with high probability the observer observed at least T_0 stages.

Recall that β_t is the average payoff up to stage t during the stages m in which the partial history up to stage m is in H_{safe}; that is, this is the average payoff as observed by the observer. Let Ω' be the event that, for the observer, $d(\beta_t, \overline{A} \cap \overline{D}) \leq \epsilon$ for every $t \geq T_0$:

$$\Omega' := \{d(\beta_t, \overline{A} \cap \overline{D}) \leq \epsilon, \forall t \text{ such that } h_t \in H_{\text{safe}}, t - f(h_t) \geq T_0\}.$$

By Eq. (12) we have $P_{\sigma^*,\tau}(\Omega') > 1 - \epsilon$.

Denote by Ω'' the event

$$\Omega'' := \Omega' \cap \left\{ f(h_t) \leq \tfrac{c'\sqrt{t}}{\delta}, \forall t \geq 1 \right\}.$$

By Lemma 8 we have $P_{\sigma^*,\tau}[\Omega''] \geq 1 - 2\epsilon$.

From now on we restrict our attention to the event Ω'' and we fix $t \geq T_1$. By definition we have $d(\beta_t, \overline{A} \cap \overline{D}) \leq \epsilon$. Since payoffs are between 0 and 1, we have $d(\alpha_t, \overline{A} \cap \overline{D}) \leq 1$, which implies that

$$d(g_t, \overline{A} \cap \overline{D}) \leq \tfrac{f(h_t)}{t} + \tfrac{t - f(h_t)}{t} d(\beta_t, \overline{A} \cap \overline{D}) \leq \epsilon + \tfrac{c'}{\delta\sqrt{t}}. \tag{13}$$

Taking $T_2 := \max\{T_1, \frac{c'^2}{\epsilon^2\delta^2}\}$ we obtain that on Ω''

$$d(g_t, \overline{A} \cap \overline{D}) \leq 2\epsilon, \quad \forall t \geq T_2,$$

and the desired result follows. □

3.4 Rate of Convergence

In addition to proving that every B-set is approachable, Blackwell identified a strategy for Player 1 that guarantees that the average payoff converges to A at a rate of $O(\frac{1}{\sqrt{t}})$, where t is the number of stages played so far; that is, there is a strategy σ for Player 1 such that for every $\epsilon > 0$ there is a constant $c > 0$ (which depends only on the payoff function U) such that for every strategy τ of Player 2 and every $t \geq 1$,

$$P_{\sigma,\tau}[\forall t \in \mathbb{N}, d(g_t, A) < \tfrac{c}{\sqrt{t}}] \geq 1 - \epsilon, \tag{14}$$

see, e.g., Corollary 14.16 in Maschler et al. 2013.

We now show that the rate of convergence to the target set A is not harmed by the introduction of constraints. In particular, remaining in D as part of the approachability strategy does not incur additional penalties on the rate of approachability from an asymptotic perspective.

Theorem 11. *The rate at which the strategy σ^* approaches $\overline{A} \cap \overline{D}$ is $O(\frac{1}{\delta\sqrt{t}})$; that is, for every $\epsilon > 0$ there exists a constant $c'' > 0$ (which depends only on the payoff function U) such that for every strategy τ of Player 2 we have*

$$P_{\sigma^\star,\tau}\left[\forall t \geq 1,\ d(g_t, \overline{A} \cap \overline{D}) < \frac{c''}{\delta\sqrt{t}}\right] \geq 1 - \epsilon.$$

Proof. We use the notations of Lemma 10. Fix $\epsilon > 0$. Recall that β_t is the average payoff up to stage t in safe stages . In these stages the strategy σ^* follows the strategy $\widehat{\sigma}$ that approaches $\overline{A} \cap \overline{D}$. Hence, there is $c > 0$ (which depends only on the payoff function U)such that with probability at least $1-\epsilon$ we have $d(\beta_t, \overline{A} \cap \overline{D}) \leq \frac{c}{\sqrt{t-f(h_t)}}$ for every $t \geq 1$ and every strategy τ of Player 2 (see Eq. (14)). By Eq. (13) and Lemma 8, with probability at least $1 - \epsilon$ we have

$$\begin{aligned} d(g_t, \overline{A} \cap \overline{D}) &\leq \tfrac{t-f(h_t)}{t} d(\beta_t, \overline{A} \cap \overline{D}) + \tfrac{f(h_t)}{t} \\ &\leq \tfrac{c\sqrt{t-f(h_t)}}{t} + \tfrac{c'\sqrt{t}}{\delta t} \qquad (15) \\ &\leq \tfrac{c}{\sqrt{t}} + \tfrac{c'}{\delta\sqrt{t}} = \left(c + \frac{c'}{\delta}\right)\frac{1}{\sqrt{t}}, \qquad (16) \end{aligned}$$

and the claim follows. □

4 The Case that D is not Convex

The proof of Theorem 4 hinges on the assumption that the set D is convex. In this section we study approachability with constraints when the set D is not convex. We start with an example, which shows that in the absence of convexity, Conditions (C1) and (C2) are not sufficient to guarantee that Player 1 can approach A while remaining in D. This example will lead us to a weaker concept of approachability with constraints that we will examine.

Example 12. *Consider the game that appears in Figure 3, where Player 1 has four actions, T_1, T_2, B_1, and B_2, Player 2 has two actions, L and R, and the payoffs are two-dimensional.*

	L	R
T_1	$(1,2)$	$(2,2)$
T_2	$(2,2)$	$(1,2)$
B_1	$(1,1)$	$(2,1)$
B_2	$(2,1)$	$(1,1)$

Figure 3: The payoff matrix in Example 12.

Let $0 < \alpha' < \alpha < \frac{1}{2}$ and define $A := B((\frac{3}{2},1),\alpha')$ the ball with center $(\frac{3}{2},1)$ and radius α', and $D := ([1-\alpha, 2+\alpha] \times [2-\alpha, 2+\alpha]) \cup ([\frac{3}{2}-\alpha, \frac{3}{2}+\alpha] \times [1-\alpha, 2+\alpha])$, see Figure 4.

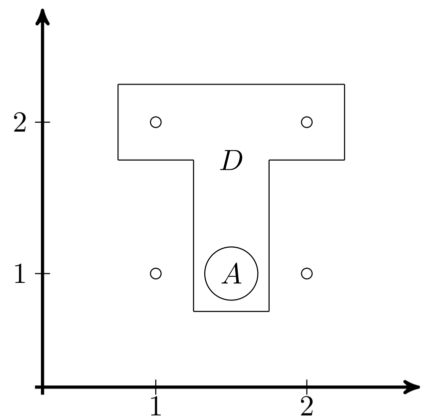


Figure 4: The sets A and D in Example 12.

Conditions (C1) and (C2) are satisfied; indeed, the actions T_1 and T_2 are safe actions, and the strategy $\frac{1}{2}B_1 + \frac{1}{2}B_2$ approaches the set A, which is a strict subset of D, and therefore it also approaches the set $\overline{A} \cap \overline{D}$.

Player 1 cannot approach A while remaining in D. Indeed, assume to the contrary that Player 1 has a strategy σ that approaches A while remaining in D, and let τ be the stationary strategy of Player 2 that plays L and R with equal probability at every stage. Fix $\epsilon > 0$ sufficiently small and suppose that the average payoff at stage t_0 is in the set $B(A, \epsilon)$. It might happen with positive probability, albeit small, that in the next $\frac{t_0}{\epsilon}$ stages the first coordinate of the outcome is 1. In this case there will be $t > t_0$ such that the average payoff at stage t is not in D.

Nevertheless, as we now argue, Player 1 can approach A while remaining in D with high probability. To do so, Player 1 *plays the mixed action $\frac{1}{2}T_1 + \frac{1}{2}T_2$ for K stages, where $K \geq 1$ is sufficiently large, and afterwards she plays the mixed action $\frac{1}{2}B_1 + \frac{1}{2}B_2$. During the first K stages the average payoff is in the convex hull of $(1,2)$ and $(2,2)$, and in particular it remains in D. Moreover, by the strong law of large numbers, provided K is sufficiently large, with high probability the first coordinate of the average payoff at stage K is between $\frac{3}{2} - \frac{\alpha}{2}$ and $\frac{3}{2} + \frac{\alpha}{2}$. By the strong law of large numbers once again, and provided K is sufficiently large, with high probability the first coordinate of the average payoff at every stage $t \geq K$ is between $\frac{3}{2} - \alpha$ and $\frac{3}{2} + \alpha$, in which case the average payoff remains in D. It follows that Player 1 can indeed approach A while remaining in D with high probability.*

Example 12 leads us to the study of probabilistic approachability with constraints, which requires that the constraints are satisfied with high probability.

Definition 13. *Let A and D be two subsets of $\mathbb{R}^n$. Given $\epsilon > 0$, we say that Player 1* can approach A while remaining in D with probability at least $1 - \epsilon$ *if there exist a strategy σ_ϵ and an integer $T_\epsilon \geq 1$ such that for every strategy τ of Player 2 we have*

$$P\Big[\forall t \geq T_\epsilon,\ d(g_t, A) < \epsilon\Big] > 1 - \epsilon, \tag{17}$$

$$P\Big[\forall t \geq 1,\ g_t \in D\Big] > 1 - \epsilon. \tag{18}$$

We say that Player 1 can approach A while remaining in D with high probability *if, for every $\epsilon > 0$, Player 1 can approach A while remaining in D with probability at least $1 - \epsilon$.*

To prove Theorem 4, which studied the case in which the set D is convex, we constructed a strategy that "directly" approaches $\overline{A} \cap \overline{D}$: the strategy attempted to approach the set $\overline{A} \cap \overline{D}$, and played a safe action only to ensure that the average payoff does not leave D. In the next example, we illustrate a more complex strategy that handles the nonconvexity of the set D by setting intermediate goals to Player 1.

Example 14. *Consider the game that is displayed in Figure 5, where Player 1 has four actions, x_0, x_1, x_2, and x_3, Player 2 has two actions, L and R, and the payoff is two-dimensional.*

	L	R
x_0	$(1,1)$	$(1,1)$
x_1	$(4,1)$	$(4,1)$
x_2	$(2,3)$	$(4,3)$
x_3	$(4,3)$	$(2,3)$

Figure 5: The payoff matrix in Example 14.

Let $0 < \alpha' < \alpha < \frac{1}{2}$ and define $A := B((3,3),\alpha')$ and $D := ([1-\alpha, 3+\alpha] \times [1-\alpha, 1+\alpha]) \cup ([3-\alpha, 3+\alpha] \times [1-\alpha, 3+\alpha])$, see Figure 6.

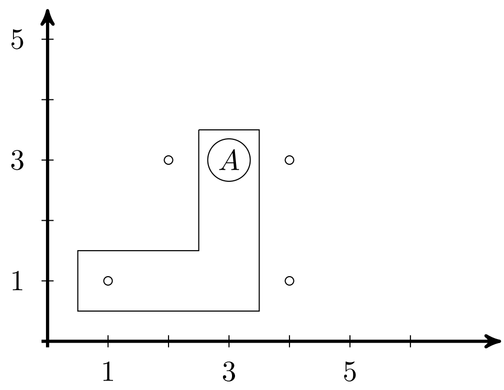


Figure 6: The sets A and D in Example 14.

To approach the set A while remaining in D with high probability, Player 1 can use the following strategy.

- *Play the action x_0 during T_0 stages, where $T_0 \geq 1$ is sufficiently large. Regardless of the play of Player 2, the average payoff is $g_t = (1,1)$ for every $t \in \{1, 2, \ldots, T_0\}$.*
- *Between stages T_0+1 and T_0+T_1, play the action x_1, where $T_1 = 3T_0$. Regardless of the play of Player 2, we have $g_{T_0+T_1} = (3,1)$, and for every $t \in \{T_0+1, T_0+2, \ldots, T_0+T_1\}$ the average payoffs g_t is in the convex hull of $(1,1)$ and $(3,1)$, hence in D.*
- *Then play forever the mixed action $[\frac{1}{2}(x_2), \frac{1}{2}(x_3)]$, which approaches the set A.*

By the strong law of large numbers, the probability that the average payoff always remains in D goes to 1 as T_0 goes to infinity. Indeed, the average payoff will leave the set D only if in the last phase of the play, the percentage of the number of stages in which the outcome is $(2,3)$ is far from 0.5, an event that occurs with a small probability when T_0 is sufficiently large.

While the strategy used in the proof of Theorem 4 plays either the safe action or a strategy that approaches the set $\overline{A} \cap \overline{D}$, the strategy used in Example 14 starts by playing a sequence of actions that leads the average payoff towards the intermediate target $(3, 1)$, and then a sequence of actions that leads the average payoff towards $\overline{A} \cap \overline{D}$. There are two main differences between the two strategies.

- First, since in Theorem 4 the set D is convex, it follows that the convex hull of $R_1(s)$ and any point $g_t \in D$ is a subset of D. Hence, Player 1 can switch from playing the safe action s to a strategy that approaches $\overline{A} \cap \overline{D}$, back and forth, to maintain the average payoff in D. When the set D is not convex, one needs to "lead" the average payoff from $R_1(s)$ to some point x that satisfies that the convex hull of x and a subset of $\overline{A} \cap \overline{D}$ remains in D, and once the average payoff gets close to x, switch to the strategy that approaches $\overline{A} \cap \overline{D}$. Thus, the play before switching to the strategy that approaches $\overline{A} \cap \overline{D}$ is more involved when the set D is not convex.
- Second, the convexity of the set D in Theorem 4 ensures that whenever the average payoff gets close to $F \setminus D$ and the strategy plays the safe action again, the average payoff, which is the average of two vectors in D, is in D. When the set D is not convex, this property does not hold, and it may be impossible to play the safe action, in a way that ensures that the average payoff remains in D. Hence in this case we have to work with the weaker concept of Definition 13.

Example 14 shows that it is unlikely that there is an elegant characterization to approachability with constraints when the set D is not convex, because the structure of the payoff function may allow Player 1 to make the average payoff follow a complicated path in D that starts in some set $R_1(s)$ and ends at A. The next theorem presents one sufficient condition for a pair of sets (A, D) to be such that Player 1 can approach A while remaining in D with high probability. This sufficient condition generalizes the insight of Example 14. A mixed action of Player 1 is called *safe* if each action in the support of the mixed action is safe.

Theorem 15. *Let A and D be two subsets of $\mathbb{R}^n$, the latter being open. Suppose that*

(D1) The set $A \cap D$ is approachable.

(D2) There exist $\delta > 0$, a safe mixed-action x_0, a positive integer m, m mixed actions $x_1, \ldots, x_m \in \Delta(I)$, and m open subsets $A_1, \ldots, A_m$ of $\mathbb{R}^n$ such that $A_m \subseteq A \cap D$ and for every $0 \leq \ell \leq m-1$ the following hold:

$$\mathrm{conv}[\overline{A}_\ell \cup B(R_1(x_{\ell+1}), \delta)] \setminus A_{\ell+1} \text{ is not path-connected, and} \tag{19}$$
$$\mathrm{conv}[A_\ell \cup B(A_{\ell+1}, \delta)] \subseteq D, \tag{20}$$

where $A_0 := R_1(x_0)$.

Then Player 1 can approach A while remaining in D with probability $1 - \epsilon$, for every $\epsilon > 0$.

Remark 16. *Theorem 15 claims that we can weaken the convexity assumption. To approach A we can specify objective sets $A_1, A_2, \ldots, A_m$ and have the average payoff reach the sets successively. The role of Condition (D2) is to ensure that there exists actions that makes the move from A_ℓ and $A_{\ell+1}$ possible, while remaining in D. The role of Condition (D1) is to ensure that once the average payoff reaches the set A_m, it can remain in $A \cap D$ with high probability.*

Remark 17. *The game described in Example 14 satisfies Conditions (D1) and (D2). Condition (D1) is satisfied because $A = A \cap D$ is approachable. Condition (D2) is satisfied because we can define $A_0 := R_1(x_1)$ and $A_2 = B((3,1),\epsilon)$ for a small $\epsilon > 0$. The average payoff reaches A_1 from A_0 by playing x_1 and it reaches $A_2 = A$ from A_1 by playing $\frac{1}{2}x_2 + \frac{1}{2}x_3$.*

Proof. The proof is quite technical, yet it poses no conceptual difficulties. We therefore only present the main steps of the proof.

Eq. (19) implies that every path in $\mathrm{conv}[A_\ell \cup R_1(x_{\ell+1})]$ that links A_ℓ to $R_1(x_{\ell+1})$ intersects $A_{\ell+1}$. Moreover, because $A_{\ell+1}$ is open while $\overline{A}_\ell$ and $R_1(x_{\ell+1})$ are closed, the length of this intersection is bounded away from 0. Denote by $\Lambda_\ell > 0$ a lower bound on the length of these intersections, and define

$$\Lambda := \min_{0 \le \ell \le m-1} \Lambda_\ell > 0.$$

Fix $\epsilon > 0$ and let T such that $T \gg \frac{1}{\Lambda}$. Define a collection $\tau_0, \tau_1, \cdots, \tau_{m-1}$ of stopping times as follows:

$$\tau_0 \quad := \quad T, \tag{21}$$

$$\tau_\ell \quad := \quad \min\{t > \tau_{\ell-1} \colon g_t \in A_\ell\}, \quad \ell \in \{1, 2, \cdots, m-1\}. \tag{22}$$

For $1 \le \ell \le m$, the stopping time τ_ℓ is the first stage after stage $\tau_{\ell-1}$ in which the average payoff reaches the set A_ℓ. Define a strategy $\sigma(T)$ as follows:

- Until stage τ_0 play the safe mixed action x_0.
- Between stages $\tau_{\ell-1}$ and $\tau_\ell - 1$ play the mixed action x_ℓ, for each $\ell \in \{1, 2, \cdots, m-1\}$.
- From stage τ_{m-1} and onward play a strategy $\widehat{\sigma}$ that approaches the set $A \cap D$.

Let T_0 be sufficiently large. We argue that if Player 1 plays the strategy $\sigma(T_0)$, then with high probability

- The stopping times $\tau_1, \cdots, \tau_{m-1}$ are bounded, regardless of the strategy played by Player 2.
- The average payoff remains in D.

Since the mixed action x_0 is safe, $g_t \in D$ for every $t \le \tau_0$.

Assume by induction that, for a given $\ell \in \{1, 2, \cdots, m-1\}$, there is an integer $T_{\ell-1}$ such that $P_{\sigma(T),\tau}[\tau_{\ell-1} \le T_{\ell-1}] > 1 - \ell\epsilon$. In particular, $g_{\tau_{\ell-1}} \in A_{\ell-1}$ on the event $\{\tau_{\ell-1} \le T_{\ell-1}\}$. At stage $\tau_{\ell-1}$ Player 1 starts playing the mixed action x_ℓ, so the expected stage payoff is in $R_1(x_\ell)$. The sequence of the average payoffs $(g_t)_{t=\tau_{\ell-1}}^{\tau_\ell}$ starts at $A_{\ell-1}$ and moves towards $R_1(x_\ell)$. Since $\tau_{\ell-1} \ge \tau_0 = T_0$, we have $\|g_t - g_{t-1}\| < \frac{2}{T_0} < \Lambda$. By the strong law of large numbers, provided T_0 is much larger than $\frac{1}{\Lambda}$, there is $T_\ell \ge T_{\ell-1}$ such that $g_t \in A_\ell$ for some $t \le T_\ell$, and therefore τ_ℓ is finite. Moreover, Eq. (20) implies that provided T_0 is sufficiently large, $P_{\sigma(T_0),\tau}[\tau_\ell \le T_\ell] > 1 - (\ell+1)\epsilon$, for every strategy τ of Player 2.

Since at stage τ_{m-1} Player 1 starts following a strategy that approaches the set $A \cap D$, there is an integer T_m such that $P_{\sigma(T_0),\tau}[d(g_t, A \cap D) \le \epsilon, \quad \forall t \ge \tau_m] > 1 - (m+2)\epsilon$. Eq. (20), applied to $\ell = m$, implies that, provided T_0 is sufficiently large, with high probability $g_t \in D$ for every $t \ge \tau_m$. The result follows. ☐

The next example shows that the sufficient condition provided by Theorem 15 is not necessary.

Example 18. *Consider the game that appears in Figure 7, where Player 1 has three actions, T_1, T_2, and B, Player 2 has two actions, L and R, and the payoff is two-dimensional.*

	L	R
T_1	$(1, 2)$	$(1, 2)$
T_2	$(2, 1)$	$(2, 1)$
B	$(2, 3)$	$(3, 2)$

Figure 7: The payoff matrix in Example 18.

Let $0 < \alpha' < \alpha < \frac{1}{2}$ and define $A := \{(2, 2)\}$ and $D := ([2 - \alpha, 2 + \alpha] \times [2 - \alpha, 3 + \alpha]) \cup ([2 - \alpha, 3 + \alpha] \times [2 - \alpha, 2 + \alpha])$, see Figure 8.

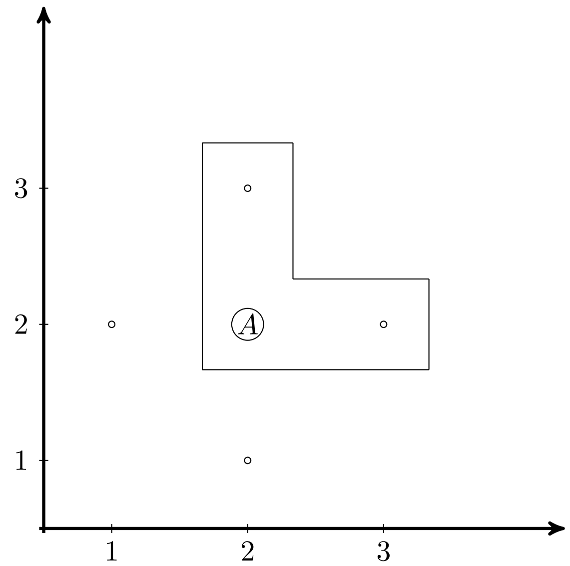


Figure 8: The sets A and D in Example 18.

In this game, B is the only safe action of Player 1, yet $R_1(B) \not\subset D$, and therefore Eq. (20) is not satisfied for $\ell = 0$. Nevertheless, Player 1 can approach A while remaining in D. To do that, Player 1 plays in blocks of size 2. In the first stage of each block Player 1 plays the action B. If Player 2 played the action L (resp. R) at the first stage of the block, then in the second stage of the block Player 1 plays the action T_2 (resp. T_1). The average payoff in each block is $(2, 2)$. The reader can verify that the strategy described above ensures that the average payoff converges to $(2, 2)$ while remaining in D.

5 Discussion and Open Problems

In this paper we presented two variants of the concept of approachability with constraints, characterized approachability with constraints when the set of constraints is convex (Theorem 4), showed that the rate of convergence in this case is the same as when constraints are absent (Theorem 11), and provided a sufficient condition for approachability with constraint with high probability when the set of constraints is not convex (Theorem 15). Example 18 shows that the sufficient condition in the latter case is not necessary. This example highlights the complexity that the presence of constraints introduces to the game: Player 1 can ensure that the average payoff satisfies the constraints by balancing her behavior in different stages. The sufficient condition of Theorem 15 is useful only in cases that balancing is achieved by means of the strong law of large numbers: Player 1 ensures that the average payoff is close to the expected average payoff. In Example 18 Player 1 balances the average payoff in a different way: in every two consecutive stages the average payoff is a given vector. In such a case, the sufficient condition of Theorem 15 is not useful.

The difficulty posed by the presence of constraints is not unique to the concept of approachability. Two other notions of approachability have been studied in the literature. Vieille 1992 studied the concept of weak approachability: a set A is weakly approachable if for every $t \geq 1$ sufficiently large Player 1 has a strategy $\sigma = \sigma(t)$ that guarantees that g_t, the average payoff up to stage t, is close to A, whatever Player 2 plays. Shani and Solan 2014 studied the concept of strong approachability: a set A is strongly approachable if there is $T_0 \geq 1$ such that Player 1 can ensure that the average payoff g_t is in A for every $t \geq T_0$. Introducing constraints to the model of weak approachability or strong approachability exhibits the same difficulties presented in this paper. This is because the difficulties arise due to the presence of constraints and not due to the variant of approachability that is studied: Player 1 has various ways to ensure that the average payoff satisfies the constraints, and to date we are not aware of an elegant characterization of these ways. The difficulties are also present if instead of studying the average payoff g_t one considers the average of the expected stage payoff, namely, $\frac{1}{t}\sum_{l=1}^{t} U(\sigma(h_{l-1}), \tau(h_{l-1}))$.

A related model that was studied in the literature is that of approachability in stochastic games with vector payoffs, see, e.g., Shimkin and Shwartz 1993, Milman 2006, Flesch et al. 2018. The main difference between the models is that while in stochastic games transitions depend on the state variable and on the players' actions, in repeated games with constraints the failure of the constraint depends on the average payoff, and therefore to model a repeated games with constraints as a stochastic game one needs an infinite state space.

We end the paper by introducing some open problems. The identification of a necessary and sufficient condition for approachability with constraints in the nonconvex case is an important open problem that is left for future research. Characterizing pairs of sets (A, D) such that A is weakly approachable or strongly approachable when D is convex or not convex is another interesting question.

Once the characterization of approachability with constraints in the general case will be completed, it will be important to study the convergence rate to the set $\overline{A} \cap \overline{D}$. By Theorem 11, the rate of convergence when D is convex is $O(\frac{1}{\sqrt{t}})$. We conjecture that this will be the convergence rate also in the general case, since the two forces that play role in approachability with constraints are Blackwell's approachability strategy and the strong law of large numbers, both of which have a convergence rate of $O(\frac{1}{\sqrt{t}})$.

Another issue, which already arises in the setup of approachability without constraints, is what happens when Player 2 (nature) cannot play any mixed action, but is restricted to a subset of mixed actions. This happens, e.g., when Player 2 is composed of two players who cannot correlate their actions. A characterization of the collection of approachable sets in this case, as well as understanding how this collection changes as the set of mixed action available to Player 2 varies, are two questions that call for further study.

References

Bauso, D., Blanchini, F., & Pesenti, R. (2010). Optimization of long-run average-flow cost in networks with time-varying unknown demand. *IEEE Transactions on Automatic Control*, *55*(1), 20–31.

Bauso, D., Lehrer, E., Solan, E., & Venel, X. (2015). Attainability in repeated games with vector payoffs. *Mathematics of Operations Research*, *40*, 739–755.

Blackwell, D. (1956). An analog of the minimax theorem for vector payoffs. *Pacific Journal of Mathematics*, *6*(1), 1–8.

Cesa-Bianchi, N., & Lugosi, G. (2006). *Prediction, learning, and games.* Cambridge university press.

Du, Y., & Lehrer, E. (2020). Constrained no-regret learning. *Journal of Mathematical Economics.*

Flesch, J., Laraki, R., & Perchet, V. (2018). Approachability of convex sets in generalized quitting games. *Games and Economic Behavior*, *108*, 411–431.

Hou, T.-F. (1971). Approachability in a two-person game. *Annals of Mathematical Statistics*, *42*(2), 735–744.

Khmelnitsky, E., & Tzur, M. (2004). Parallelism of continuous-and discrete-time production planning problems. *IIE Transactions*, *36*(7), 611–628.

Lehrer, E. (2002). Approachability in infinitely dimensional spaces. *International Journal of Game Theory*, *31*, 253–268.

Lehrer, E. (2004). The game of normal numbers. *Mathematics of Operations Research*, *29*, 259–265.

Lehrer, E., & Solan, E. (2009). Approachability with bounded memory. *Games and Economic Behavior*, *66*, 995–1004.

Lehrer, E., & Solan, E. (2016). A general internal regret-free strategy. *Dynamic Games and Applications*, *6*(1), 112–138.

Liakopoulos, N., Paschos, G., & Spyropoulos, T. (2019). No regret in cloud resources reservation with violation guarantees. *IEEE INFOCOM 2019-IEEE Conference on Computer Communications*, 1747–1755.

Mahdavi, M., Jin, R., & Yang, T. (2012). Trading regret for efficiency: Online convex optimization with long term constraints. *Journal of Machine Learning Research*, *13*(Sep), 2503–2528.

Mannor, S., & Perchet, V. (2013). Approachability, fast and slow. *Journal of Machine Learning Research: Workshop and Conference Proceedings*, *30*, 1–16.

Mannor, S., Stoltz, G., & Perchet, V. (2014). Set-valued approachability and online learning under partial monitoring. *Journal of Machine Learning Research*, *1*, 3247–3295.

Mannor, S., Tsitsiklis, J. N., & Yu, J. Y. (2009). Online learning with sample path constraints. *Journal of Machine Learning Research*, *10*(Mar), 569–590.

Maschler, M., Solan, E., & Zamir, S. (2013). Game theory (translated from the hebrew by ziv hellman and edited by mike borns). *Cambridge University Press, Cambridge, pp. xxvi*, *979*, 4.

Milman, E. (2006). Approachable sets of vector payoffs in stochastic games. *Games and Economic Behavior*, *56*(1), 135–147.

Perchet, V. (2009). Calibration and internal no-regret with random signals. *International Conference on Algorithmic Learning Theory*, 68–82.

Perchet, V. (2014). Approachability, regret and calibration. implications and equivalences. *Journal of Dynamics and Games*, *1*(2), 181–254.

Perchet, V., & Quincampoix, M. (2014). On a unified framework for approachability with full or partial monitoring. *Mathematics of Operations Research*, *40*(3), 596–610.

Sadeghi, O., & Fazel, M. (2020). Online continuous dr-submodular maximization with long-term budget constraints. *International Conference on Artificial Intelligence and Statistics*, 4410–4419.

Shani, B., & Solan, E. (2014). Strong approachability. *Journal of Dynamic Games*, *1*, 507–535.

Shimkin, N., & Shwartz, A. (1993). Guaranteed performance regions in markovian systems with competing decision makers. *IEEE Transactions on Automatic Control*, *38*(1), 84–95.

Spinat, X. (2002). A necessary and sufficient condition for approachability. *Mathematics of Operations Research*, *27*(3), 31–44.

Uziel, G., & El-Yaniv, R. (2018). Growth-optimal portfolio selection under cvar constraints. *International Conference on Artificial Intelligence and Statistics*, 48–57.

Vieille, N. (1992). Weak approachability. *Mathematics of Operations Research*, *17*(4), 781–791.