A Primal Condition for Approachability with Partial Monitoring

Shie Mannor — Vianney Perchet — Gilles Stoltz *

Technion — Université Paris-Diderot — CNRS / Ecole normale supérieure / HEC Paris

May 23, 2013

Abstract

In approachability with full monitoring there are two types of conditions that are known to be equivalent for convex sets: a primal and a dual condition. The primal one is of the form: a set $\mathcal{C}$ is approachable if and only all containing half-spaces are approachable in the one-shot game. The dual condition is of the form: a convex set $\mathcal{C}$ is approachable if and only if it intersects all payoff sets of a certain form. We consider approachability in games with partial monitoring. In previous works (Perchet, 2011a, Mannor et al., 2011) we provided a dual characterization of approachable convex sets and we also exhibited efficient strategies in the case where $\mathcal{C}$ is a polytope. In this paper we provide primal conditions on a convex set to be approachable with partial monitoring. They depend on a modified reward function and lead to approachability strategies based on modified payoff functions and that proceed by projections similarly to Blackwell's (1956) strategy. This is in contrast with previously studied strategies in this context that relied mostly on the signaling structure and aimed at estimating well the distributions of the signals received. Our results generalize classical results by Kohlberg (1975) (see also Mertens et al., 1994) and apply to games with arbitrary signaling structure as well as to arbitrary convex sets.

1. Introduction

Approachability theory dates back to the seminal paper of Blackwell (1956). In this paper Blackwell presented conditions under which a player can guarantee that the long-term average vector-valued reward is asymptotically close to some target set regardless of the opponent actions. If this property holds, we say that the set is approachable. In

*This work began after an interesting objection of and further discussions with Jean-François Mertens during the presentation of our earlier results (Mannor et al., 2011) on the dual characterization of approachability with partial monitoring at the conference *Games Toulouse 2011*. The material presented in this article was developed further with Sylvain Sorin back in Paris, whom we thank deeply for his advice and encouragements. Sadly, Jean-François Mertens, a close collaborator of Sylvain Sorin, passed away in the mean time. This contribution is in honor of both of these important contributors to the theory of repeated games.

the full monitoring case studied in Blackwell (1956) there are two equivalent conditions for a convex set to be approachable. The first one, known as a primal condition (or later termed the "B" condition in honor for Blackwell) states that every half-space that contains the target set is also approachable. It turns out that whether a half-space is approachable is determined by the sign of the value of some associated zero-sum game. The second characterization, known as the dual condition, states that for every mixed action of the opponent, the player can guarantee that the one-shot vector-valued reward is inside the target set.

Approachability theory has found many applications in game theory, online learning, and related fields. Both primal and dual characterizations are of interest therein. Indeed, checking if the dual condition holds is formally simple while a concrete approaching strategy naturally derives from the primal condition (it only requires solving a one-shot zero-sum game at every stage of the repeated vector-valued game).

Approachability theory has been applied to zero-sum repeated games with incomplete information and/or imperfect (or partial) monitoring. The work of Kohlberg (1975) (see also Mertens et al., 1994) uses approachability to derive strategies for games with incomplete information. The general case of repeated vector-valued games with partial monitoring was studied only recently. A dual characterization of approachable convex sets with partial monitoring was presented by Perchet (2011a). However, it is not useful for deriving approaching strategies since it essentially requires to run a calibration algorithm, which is known to be computationally hard. In a recent work (Mannor et al., 2011) we derived efficient strategies for approachability in games with partial monitoring in some cases, e.g., when the convex set to be approached is a polytope. However, these strategies are based on the dual condition, and not on any primal one: they thus do not shed light on the structure of the game.

In this paper we provide a primal condition for approachability in games with partial monitoring. It can be stated, as in Blackwell (1956), as a requirement that every half-space containing the target set is one-shot approachable. However, the reward function has to be modified in some cases for the condition to be sufficient. We also show how it leads to an efficient approachability strategy, at least in the case of approachable polytopes.

Outline. In Section 2 we define a model of partial monitoring and recall some of the basic results from approachability (both in terms of its primal and dual characterizations). In Section 3 we explain the current state-of-the-art, recall the dual condition for approachability with partial monitoring, and outline our objectives. In Section 4 we provide results for approaching half-spaces as they have the simplest characterization of approachability: we show that the signaling structure has no impact on approachability of half-spaces, only the payoff structure does. This is not the case anymore for approachability of more complicated convex sets, which is the focus of the subsequent sections. In Section 5 we discuss the case of a target set that is an orthant under additional properties on the payoff–signaling structure: we show that a natural primal condition holds. This condition, which we term the "upper-right-corner property" is the main technical

contribution of the paper. We show that basically, we can study approachability for a modified payoff function and that under some favorable conditions, a primal condition is easy to derive (which is the main conceptual contribution). As an intermezzo, we link our results to Kohlberg (1975) in Section 6 and show that repeated games with imperfect information can be analyzed using our approach for games with imperfect monitoring. In Section 7 we then analyze the case of a general signaling structure for the approachability of orthants and provide an efficient approaching strategy based on the exhibited primal conditions. Finally, we relax the shape of the target set from an orthant to a polytope in Section 8, and then to a general convex set in Section 9. Our generalizations show that the *same* primal condition holds in all cases. The generalization from orthants to polytopes is based on the observation that any polytope can be represented as an orthant in a space whose dimensionality is the number of linear inequalities describing the polytope and on a modified reward function. The generalization to general convex sets uses support functions and lifting to derive similar results; we provide some background material on support functions in the appendix.

2. Model and preliminaries

We now define the model of interest and then recall some basic results from approachability theory for repeated vector-valued games (with full monitoring).

Model and notation

We consider a vector-valued game between two players, a decision maker (or player) and Nature, with respective finite action sets $\mathcal{I}$ and $\mathcal{J}$, whose cardinalities are referred to as I and J. We denote by d the dimension of the reward vector and equip $\mathbb{R}^d$ with the ℓ^2–norm $\|\cdot\|_2$. The payoff function of the player is given by a mapping $r : \mathcal{I} \times \mathcal{J} \to \mathbb{R}^d$, which is multi-linearly extended to $\Delta(\mathcal{I}) \times \Delta(\mathcal{J})$, the set of product-distributions over $\mathcal{I} \times \mathcal{J}$.

At each round, the player and Nature simultaneously choose their actions $i_n \in \mathcal{I}$ and $j_n \in \mathcal{J}$, at random according to probability distributions denoted by $x_n \in \Delta(\mathcal{I})$ and $y_n \in \Delta(\mathcal{J})$. At the end of a round, the player does not observe Nature's action j_n nor even the payoff $r_n := r(i_n, j_n)$ he obtains: he only gets to see some signal. More precisely, there is a finite set $\mathcal{H}$ of possible signals, and the signal $s_n \in \mathcal{H}$ that is shown to the player is drawn at random according to the distribution $H(i_n, j_n)$, where the mapping $H : \mathcal{I} \times \mathcal{J} \to \Delta(\mathcal{H})$ is known by both players.

The player is said to have full monitoring if $\mathcal{H} = \mathcal{J}$ and $H(i, j) = \text{Full}(i, j) := \delta_j$, i.e., if the action of Nature is observed. We speak of a game in the dark when the signaling structure H is not informative at all, i.e., when $\mathcal{H}$ is reduced to a single signal referred to as $\emptyset$; we denote this situation by $H = \text{Dark}$.

Of major interest will be *maximal information mapping* $\boldsymbol{H} : \Delta(\mathcal{J}) \to \Delta(\mathcal{H})^{\mathcal{I}}$, which is defined as follows. The image of each $j \in \mathcal{J}$ is the vector $\boldsymbol{H}(j) = \big(H(i, j)\big)_{i \in \mathcal{I}}$, and this definition is extended linearly onto $\Delta(\mathcal{J})$. An element of the image $\mathcal{F} = \boldsymbol{H}\big(\Delta(\mathcal{J})\big)$

of $\boldsymbol{H}$ is referred to as a *flag*. The notion of "flag" is key: the player only accesses the mixed actions y of Nature through H. Indeed, as is intuitive and as is made more formal at the end of the proof of Proposition 2, he could at best access or estimate the flag $\boldsymbol{H}(y)$ but not y itself.

For every $x \in \Delta(\mathcal{I})$ and $\boldsymbol{h} \in \mathcal{F}$ the set of payoffs compatible with $\boldsymbol{h}$ is

$$\overline{m}(x, \boldsymbol{h}) = \big\{ r(x,y) : \ y \in \Delta(\mathcal{J}) \text{ such that } \boldsymbol{H}(y) = \boldsymbol{h} \big\}. \tag{1}$$

The set $\overline{m}(x, \boldsymbol{h})$ represents all the rewards that are statistically compatible with the flag $\boldsymbol{h}$ (or put differently, the set of all possible rewards we cannot distinguish from).

Note that with full monitoring, $\mathcal{H}$ reduces to $\Delta(\mathcal{J})$ and one has $\overline{m}(x,y) = \big\{ r(x,y) \big\}$ for all $y \in \Delta(\mathcal{J})$.

Finally, we denote by M a uniform ℓ_2–bound on r, that is,

$$M = \max_{i,j} \big\| r(i,j) \big\| .$$

Also, for every $n \in \mathbb{N}$ and sequence $(a_m)_{m \in \mathbb{N}}$, the average of the first n elements is referred to as $\overline{a}_n = (1/n) \sum_{m=1}^{n} a_m$. The distance to a set $\mathcal{C}$ is denoted by $d_{\mathcal{C}}$.

A behavioral strategy σ of the player is a mapping from the set of his finite histories $\cup_{n \in \mathbb{N}} (\mathcal{I} \times \mathcal{H})^n$ into $\Delta(\mathcal{I})$. Similarly, a strategy τ of nature is a mapping from $\cup_{n \in \mathbb{N}} (\mathcal{I} \times \mathcal{H} \times \mathcal{J})^n$ into $\Delta(\mathcal{J})$. As usual, we denote by $\mathbb{P}_{\sigma,\tau}$ the probability induced by the pair (σ, τ) onto $(\mathcal{I} \times \mathcal{H} \times \mathcal{J})^{\mathbb{N}}$.

Definition and some properties of approachability

A set $\mathcal{C} \subseteq \mathbb{R}^d$ is r–approachable for the signaling structure H, or, in short, is (r, H)–approachable, if, for all $\varepsilon > 0$, there exists a strategy σ_ε of the player and a natural number $N \in \mathbb{N}$ such that, for all strategies τ of Nature,

$$\mathbb{P}_{\sigma_\varepsilon, \tau} \big\{ \exists\, n \geqslant N \ \text{ s.t. } \ d_{\mathcal{C}}(\overline{r}_n) \geqslant \varepsilon \big\} \leqslant \varepsilon \, .$$

We refer to the strategy σ_ε as an ε–approachability strategy of $\mathcal{C}$. It is easy to show that the approachability of $\mathcal{C}$ implies the existence of a strategy ensuring that the sequence of the average vector-valued payoffs converges to the set $\mathcal{C}$ almost surely, uniformly with respect to the strategies of Nature. By analogy such a strategy is called a 0–approachability strategy of $\mathcal{C}$.

Conversely, a set $\mathcal{C}$ is r–excludable for the signaling structure H if, for some $\delta > 0$, the complement of the δ–neighborhood of $\mathcal{C}$ is r–approachable by Nature for the signaling structure H.

Primal characterization. We now discuss characterizations of approachability in the case of full monitoring. We will need the stronger notion of one-shot approachability (the notion of one-shot excludability is stated only for later purposes).

Definition 1. *A set $\mathcal{C}$ is one-shot r–approachable if there exists $x \in \Delta(\mathcal{I})$ such that for all $y \in \Delta(\mathcal{J})$, one has $r(x,y) \in \mathcal{C}$. A set $\mathcal{C}$ is one-shot r–excludable if for some $\delta > 0$, the complement of the δ–neighborhood of $\mathcal{C}$ is one-shot r–approachable by Nature.*

Blackwell (1956) (see also Mertens et al., 1994) provided the following *primal characterization* of approachable convex[1] sets. A set that satisfies it is called a B–set.

Theorem 1. *A convex set $\mathcal{C}$ is (r, Full)–approachable if and only if any containing half-space $\mathcal{C}_{\mathrm{hs}} \supseteq \mathcal{C}$ is one-shot r–approachable.*

This characterization also leads to an approachability strategy, which we describe with a slight modification with respect to its most classical statement. We denote by $r'_t = r(x_t, j_t)$ the expected payoff obtained at round t, which is a quantity that is observed by the player. At stage n, if $\overline{r}'_n \notin \mathcal{C}$, let $\pi_{\mathcal{C}}(\overline{r}'_n)$ denote the projection of $\overline{r}'_n$ onto $\mathcal{C}$ and consider the containing half-space $\mathcal{C}_{\mathrm{hs},\,n+1}$ whose defining hyperplane is tangent to $\mathcal{C}$ at $\pi_{\mathcal{C}}(\overline{r}'_n)$. The strategy then consists of choosing the mixed action $x_{n+1} \in \Delta(\mathcal{I})$ associated with the one-shot approachability of $\mathcal{C}_{\mathrm{hs},\,n+1}$, as illustrated in Figure 1.

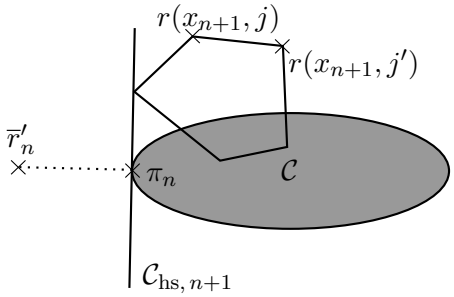


Figure 1: An illustration of Blackwell's approaching strategy. At stage $n+1$, when $\overline{r}'_n \notin \mathcal{C}$, the convex set $\mathcal{C}$ and the expected payoff r'_{n+1} are in the half-space $\mathcal{C}_{\mathrm{hs},\,n+1}$ while $\overline{r}'_n$ lies in its complement.

If $\overline{r}'_n \in \mathcal{C}$, any choice $x_{n+1} \in \Delta(\mathcal{I})$ is suitable. The above strategy ensures that for all $y \in \Delta(\mathcal{J})$,

$$\left\langle r(x_{n+1}, y) - \pi_{\mathcal{C}}(\overline{r}'_n),\ \overline{r}'_n - \pi_{\mathcal{C}}(\overline{r}'_n) \right\rangle \leqslant 0\,; \tag{2}$$

which in turn ensures the convergence to $\mathcal{C}$ of the mixed payoffs at a rate independent of d, namely

$$d_{\mathcal{C}}\left(\frac{1}{n}\sum_{t=1}^{n} r(x_t, j_t)\right) \leqslant \frac{2M}{\sqrt{n}}\,. \tag{3}$$

The uniform convergence of $\overline{r}_n$ to $\mathcal{C}$ is deduced by martingale convergence theorems (e.g., the Hoeffding–Azuma inequality) from the above uniform convergence of $\overline{r}'_n$ to $\mathcal{C}$.

[1]This primal characterization was actually stated by Blackwell (1956) in a more general way for all, non-necessarily convex, sets.

Dual characterization. In the specific case of closed convex sets, using von Neumann's min-max theorem, the primal characterization stated above can be transformed into the following *dual characterization*:

$$\mathcal{C} \subseteq \mathbb{R}^d \text{ is } (r, \text{Full})\text{–approachable} \quad \Longleftrightarrow \quad \forall\, y \in \Delta(\mathcal{J}),\ \exists\, x \in \Delta(\mathcal{I}),\ \ r(x,y) \in \mathcal{C}\,. \quad (4)$$

This characterization might be simpler to formulate and to check, yet it does not provide an explicit approachability strategy.

3. Related literature and the objective of this paper

In this section we first recall the existing results on approachability with partial monitoring and then explain in a more technical way the objectives of the paper.

Results on approachability with partial monitoring

Concerning the primal characterization. Kohlberg (1975)—see also Mertens et al., 1994—studied specific frameworks (induced by repeated games with incomplete information, see Section 6) in which approachability depends mildly on the signaling structure. A property that we define in the sequel and call the *upper-right-corner property* holds between the payoff function r and the signaling structure H. Based on this property it is rather straightforward to show that the primal characterization for the (r, H)–approachability of orthants stated in Theorem 1 still holds. Section 6 provides more details on this matter.

Concerning the dual characterization. Perchet (2011a) provided the following dual characterization of approachable closed convex sets under partial monitoring:

$$\mathcal{C} \subseteq \mathbb{R}^d \text{ is } (r,H)\text{–approachable} \quad \Longleftrightarrow \quad \forall\, \boldsymbol{h} \in \mathcal{F},\ \exists\, x \in \Delta(\mathcal{I}),\ \overline{m}(x, \boldsymbol{h}) \subseteq \mathcal{C}\,. \quad (5)$$

It indeed generalizes Blackwell's dual characterization (4) with full monitoring, as $\mathcal{F}$ can be identified with $\Delta(\mathcal{J})$ in this case.

Based on (5), Perchet constructed the first (r, H)–approachability strategy of any closed convex set $\mathcal{C}$; it was based on calibrated forecasts of the vectors of $\mathcal{F}$. Because of this, the per-stage computational complexity of this strategy increases indefinitely and rates of convergence cannot be inferred. Moreover, the construction of this strategy is unhelpful to infer a generic primal characterization.

Mannor et al. (2011) tackled the issue of complexity and devised an efficient (r, H)–approachability strategy for the case where the target set is some polytope. This strategy has a fixed and bounded per-stage computational complexity. Moreover, its rates of convergence are independent of d: they are of the order of $n^{-1/5}$, where n is the number of stages.

On the other hand, Perchet and Quincampoix (2011) unified the setups of approachability with full or partial monitoring and characterized approachable closed (convex) sets using some lifting to the Wasserstein space of probability measures on $\Delta(\mathcal{I}) \times \Delta(\mathcal{J})$.

Objectives and technical content of the paper

This paper focuses on the primal characterization of approachable closed convex sets with partial monitoring. First, note that if a closed convex set is (r, H)–approachable, then it is also (r, Full)–approachable, and therefore, by (4), any containing half-space is necessarily one-shot r–approachable. The question is: When is the latter statement a sufficient condition for (r, H)–approachability? The difficulty, as noted already by Perchet (2011a) and recalled at the beginning of Section 5, is that since the notions of approachability with full or partial monitoring do not coincide, it can be that a closed convex set is not (r, H)–approachable while every containing half-space is one-shot r–approachable.

Some situations where the usual dual characterization is indeed sufficient are formed first, by the cases when the target set is a half-space (with no condition on the game), and second, by the cases when the target set is an orthant and the structure (r, H) of the game satisfies the upper-right-corner property. This first series of results is detailed in Sections 4 and 5. Some light is then shed in Section 6 on the construction of Kohlberg (1975) for the case of repeated games with incomplete information.

The rest of the paper (Sections 7, 8, and 9) relies on no assumption on the structure (r, H) of the game. It discusses a primal condition based on one-shot approachability of half-spaces with respect to a modified payoff function $\widetilde{r}_H$ that encompasses the links between the signaling structure H and the original payoff function r. Depending on the geometry of the target closed convex set, this primal condition is stated in the original payoff space (for orthants, Section 7) or in some lifted space (for polytopes or general convex sets, see Section 8 and 9). We explain in Example 1 why such a lifting seems inevitable.

We also illustrate how the exhibited primal condition leads to a new (and efficient) strategy for (r, H)–approachability in the case of target sets given by polytopes. (Section 7 does it for orthants and the result extends to polytopes via Lemma 2.) This new strategy is based on sequences of (modified) payoffs, as in Kohlberg (1975), and is not only based on sequences of signals, as in Perchet (2011a), Mannor et al. (2011). The construction of this strategy also entails some non-linear approachability results (both in full and partial monitoring).

4. Primal approachability of half-spaces

We first focus on half-spaces, not only because they are the simplest convex sets, but because they are the cornerstones of the primal characterization of Blackwell (1956). The following proposition ties one-shot r–approachability with (r, H)–approachability of half-spaces.

Proposition 1. *For all half-spaces $\mathcal{C}_{\text{hs}}$, for all signaling structures H,*

$$\mathcal{C}_{\text{hs}} \text{ is } (r, H)\text{–approachable} \iff \mathcal{C}_{\text{hs}} \text{ is one-shot } r\text{–approachable}.$$

This result is a mere interpolation of two well-known results for the extremal cases where $H = \text{Full}$ and $H = \text{Dark}$. The former case corresponds to Blackwell's primal

characterization. In the latter case, Nature could always play the same $y \in \Delta(\mathcal{J})$ at all rounds and the player cannot infer the value of this y. So he needs to have an action $x \in \Delta(\mathcal{I})$ such that $r(x,y)$ belongs to $\mathcal{C}_{\rm hs}$, no matter y.

Stated differently, the above proposition indicates that as far as half-spaces are concerned, the approachability is independent of the signaling structure.

Proof. Only the direct implication is to be proven, as the converse implication is immediate by the above discussion about games in the dark. We thus assume that $\mathcal{C}_{\rm hs}$ is (r,H)–approachable. Using the characterization (5) of (r,H)–approachable sets, one then has that

$$\forall \boldsymbol{h} \in \mathcal{F},\ \exists\, x \in \Delta(\mathcal{I}),\ \ \overline{m}(x,\boldsymbol{h}) \subset \mathcal{C}_{\rm hs}\,,$$

which implies that

$$\forall\, y \in \Delta(\mathcal{J}),\ \exists\, x \in \Delta(\mathcal{I}),\ \ r(x,y) \in \mathcal{C}_{\rm hs}\,.$$

The implication holds because $r(x,y) \in \overline{m}(x,\boldsymbol{h})$ when $\boldsymbol{h} = \boldsymbol{H}(y)$. Now, let $a \in \mathbb{R}^d$ and $b \in \mathbb{R}$ such that $\mathcal{C}_{\rm hs} = \big\{\omega \in \mathbb{R}^d :\ \langle \omega, a\rangle \leqslant b\big\}$. The above property can be further restated as

$$\forall\, y \in \Delta(\mathcal{J}),\ \exists\, x \in \Delta(\mathcal{I}),\ \ \langle r(x,y), a\rangle \leqslant b\,,$$

or equivalently,

$$\max_{y\in\Delta(\mathcal{J})} \min_{x\in\Delta(\mathcal{I})} \langle r(x,y), a\rangle \leqslant b\,.$$

By von Neumann's min-max theorem, we then have that

$$\min_{x\in\Delta(\mathcal{I})} \max_{y\in\Delta(\mathcal{J})} \langle r(x,y), a\rangle \leqslant b\,, \quad \text{that is,} \quad \exists\, x_0 \in \Delta(\mathcal{I}),\ \forall y \in \Delta(\mathcal{J}),\ \ \langle r(x_0,y), a\rangle \leqslant b\,.$$

This is exactly the one-shot approachability of $\mathcal{C}_{\rm hs}$. □

Since the complement of any δ–neighborhood of a half-space is also a half-space, we get the following additional equivalence, in view of the respective definitions of excludability and one-shot excludability.

Corollary 1. *For all half-spaces $\mathcal{C}_{\rm hs}$, for all signaling structures H,*

$$\mathcal{C}_{\rm hs} \text{ is } (r,H)\text{–excludable} \iff \mathcal{C}_{\rm hs} \text{ is one-shot } r\text{–excludable}\,.$$

5. Primal approachability of orthants under the upper-right-corner property

This section is devoted to stating a primal characterization of (r,H)–approachable orthants, i.e., sets of the form

$$\mathcal{C}_{\rm orth}(a) = \big\{a - \omega :\ \omega \in (\mathbb{R}_+)^d\big\}$$

for some $a \in \mathbb{R}^d$. Orthants are the key for extension to polytopes, because, as we will discuss later, up to some lifting in higher dimensions, every polytope can be seen as an orthant.

We start by indicating that the primal characterization stated in the previous section in terms of the original payoff function r does not extend directly to general convex sets, not even to orthants—at least without an additional assumption. However, in this section we state such a sufficient assumption for its extension. We study the most general primal characterization in Section 7, which will involve a modified payoff function for the one-shot approachability of half-spaces.

Counter-example (adapted from Perchet, 2011a). We show that the equivalence of Proposition 1 does not hold in general if the convex set $\mathcal{C}$ at hand is not a half-space. To do so, we exhibit a game and a set $\mathcal{C}$ which is (r, Full)–approachable but not (r, Dark)–approachable. We set $\mathcal{I} = \{T, B\}$ and $\mathcal{J} = \{L, R\}$, and the payoff function r is given by the matrix

	L	R
T	$(0,0)$	$(1,-1)$
B	$(-1,1)$	$(0,0)$

We consider the set $\mathcal{C}_{\text{orth}}\big((0,0)\big) = (\mathbb{R}_-)^2$. This set is (r, Full)–approachable as indicated by the dual characterization (4): for each $\alpha \in [0,1]$,

$$r\big(\alpha T + (1-\alpha)B,\, \alpha L + (1-\alpha)R\big) = (0,0) \in \mathcal{C}_{\text{orth}}\big((0,0)\big)\,.$$

On the other hand, consider the signaling structure $H = \text{Dark}$, for which the only flag is $\emptyset$. For all actions of the player, i.e., for all $\alpha \in [0,1]$, it holds that

$$\begin{aligned}\overline{m}\big(\alpha T + (1-\alpha)B,\ \emptyset\big) &= \Big\{r\big(\alpha T + (1-\alpha)B,\ y\big) :\ y \in \Delta(\mathcal{J})\Big\} \\ &= \big\{(\lambda, -\lambda)\,;\ \lambda \in [-\alpha,\, 1-\alpha]\big\} \not\subseteq \mathcal{C}_{\text{orth}}\big((0,0)\big)\,.\end{aligned}$$

Therefore, the characterization of r–approachable closed convex sets (5) does not hold when playing in the dark.

Upper-right-corner property. We define the *upper-right corner function* $R : \Delta(\mathcal{I}) \times \mathcal{F} \to \mathbb{R}^d$ of the compatible payoff function $\overline{m}$ in a component-wise manner. We write the coordinates of R as $R = (R_1, \dots, R_d)$ and set, for all $k \in \{1, \dots, d\}$,

$$\forall\, x \in \Delta(\mathcal{I}),\ \forall\, \boldsymbol{h} \in \mathcal{F}, \qquad R_k(x, \boldsymbol{h}) = \max\Big\{\omega^k :\ \omega = \big(\omega^1, \dots, \omega^d\big) \in \overline{m}(x, \boldsymbol{h})\Big\}\,.$$

The construction of R is illustrated in Figure 2. Note that the ℓ_2–norm of R is in general bounded by $M\sqrt{d}$.

The term "upper-right corner" comes from the fact that $R(x, \boldsymbol{h})$ is the (component-wise) smallest a such that $\overline{m}(x, \boldsymbol{h}) \subseteq \mathcal{C}_{\text{orth}}(a)$. Controlling the distance of $R(x, \boldsymbol{h})$ to the orthant entails controlling the distance of the whole set $\overline{m}(x, \boldsymbol{h})$ to it. Thus, the point $R(x, \boldsymbol{h})$ is in some sense the worst-case payoff vector associated with $\overline{m}(x, \boldsymbol{h})$.

Of course, $R(x, \boldsymbol{h})$ is in general not a feasible payoff vector, i.e., $R(x, \boldsymbol{h}) \notin \overline{m}(x, \boldsymbol{h})$. We are interested in this section in the case where the upper-right corner is indeed a feasible payoff—an assumption that we call the *upper-right-corner property*.

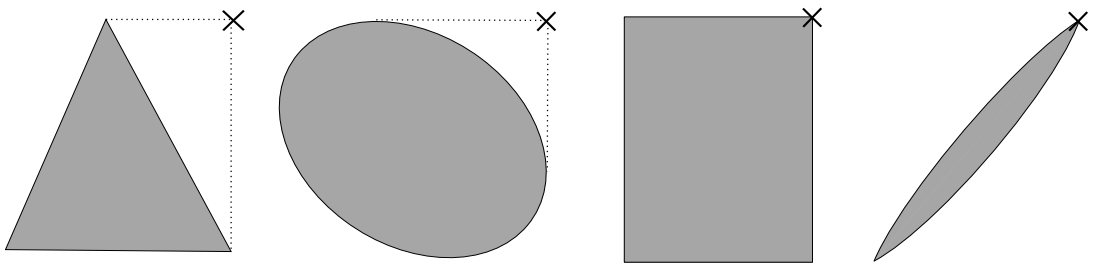


Figure 2: Four illustrations of compatible payoff sets $\overline{m}(x, \boldsymbol{h})$ and associated upper-right corners $R(x, \boldsymbol{h})$. In the two examples on the left, this upper-right corner does not belong to the set, while it does in the two on the right. (When it is so for all x and $\boldsymbol{h}$, the game is said to have the upper-right-corner property.)

Definition 2. *The game (r, H) with partial monitoring has the* upper-right-corner property *if*

$$\forall\, x \in \Delta(\mathcal{I}),\ \forall\, \boldsymbol{h} \in \mathcal{F}, \quad R(x, \boldsymbol{h}) \in \overline{m}(x, \boldsymbol{h})\,.$$

Of course, games with full monitoring have the upper-right-corner property, as for them $\mathcal{F}$ can be identified with $\Delta(\mathcal{J})$ and $\overline{m}$ can be identified with the function $\{r\}$ with values in the set of all singleton subsets of $\mathbb{R}^d$.

By definition, in a game with the upper-right-corner property, the ℓ_2–norm of R is bounded by M.

Primal characterization under the upper-right-corner property. The following proposition was implicitly used by Kohlberg (1975).

Proposition 2. *For all games (r, H) with partial monitoring that have the upper-right-corner property, for all orthants $\mathcal{C}_{\mathrm{orth}}(a)$, where $a \in \mathbb{R}^d$,*

$$\begin{aligned} & \mathcal{C}_{\mathrm{orth}}(a) \textit{ is } (r, H)\textit{–approachable} \\ \iff\quad & \textit{every half-space } \mathcal{C}_{\mathrm{hs}} \supset \mathcal{C}_{\mathrm{orth}}(a) \textit{ is one-shot } r\textit{–approachable.} \end{aligned}$$

Stated differently, using Blackwell's primal characterization of approachability (Theorem 1), an orthant $\mathcal{C}_{\mathrm{orth}}(a)$ is (r, H)–approachable in a game (r, H) satisfying the upper-right-corner property if and only if $\mathcal{C}_{\mathrm{orth}}(a)$ is r–approachable with full monitoring.

Proof. The direct implication is proved by applying Proposition 1 to any half-space $\mathcal{C}_{\mathrm{hs}} \supset \mathcal{C}_{\mathrm{orth}}(a)$, which is in particular (r, H)–approachable as soon as $\mathcal{C}_{\mathrm{orth}}(a)$ is. The interesting implication is thus the converse one.

So, we assume that every half-space $\mathcal{C}_{\mathrm{hs}} \supset \mathcal{C}_{\mathrm{orth}}(a)$ is one-shot r–approachable and, following Kohlberg's original proof and inspired by Blackwell's strategy in the case of full monitoring, we construct an (r, H)–approachability strategy of $\mathcal{C}_{\mathrm{orth}}(a)$.

Flags observed, mixed payoffs obtained. For simplicity, assume first that after stage n, the observation made by the player is not just the random signal s_n but the entire vector of probability distributions over the signals $\boldsymbol{h}_n = \boldsymbol{H}(y_n)$. (We will indicate below why this is not a restrictive assumption.) We consider the surrogate payoff vector $R_n = R(x_n, \boldsymbol{h}_n)$, which is a quantity thus observed by the player. When $\overline{R}_n$ does not already belong to $\mathcal{C}_{\mathrm{orth}}(a)$, since the latter set is convex, the half-space $\mathcal{C}_{\mathrm{hs},\,n}$ defined by

$$\mathcal{C}_{\mathrm{hs},\,n} = \Big\{\omega \in \mathbb{R}^d : \ \big\langle \omega - \pi_{\mathcal{C}_{\mathrm{orth}}(a)}\big(\overline{R}_n\big),\ \overline{R}_n - \pi_{\mathcal{C}_{\mathrm{orth}}(a)}\big(\overline{R}_n\big)\big\rangle \leqslant 0\Big\}$$

contains $\mathcal{C}_{\mathrm{orth}}(a)$, where we recall that $\pi_{\mathcal{C}_{\mathrm{orth}}(a)}$ is the orthogonal projection onto $\mathcal{C}_{\mathrm{orth}}(a)$. By assumption, $\mathcal{C}_{\mathrm{hs},\,n}$ is thus one-shot r–approachable. That is, there exists $x_{n+1} \in \Delta(\mathcal{I})$ such that

$$\forall\, y \in \Delta(\mathcal{J}), \qquad \big\langle r(x_{n+1}, y) - \pi_{\mathcal{C}_{\mathrm{orth}}(a)}\big(\overline{R}_n\big),\ \overline{R}_n - \pi_{\mathcal{C}_{\mathrm{orth}}(a)}\big(\overline{R}_n\big)\big\rangle \leqslant 0\,. \tag{6}$$

By the upper-right-corner property, $R_{n+1} \in \overline{m}(x_{n+1}, \boldsymbol{h}_{n+1})$, which entails that there exists $y'_{n+1} \in \Delta(\mathcal{J})$ such that $\boldsymbol{H}\big(y'_{n+1}\big) = \boldsymbol{h}_{n+1}$ and $R_{n+1} = r\big(x_{n+1}, y'_{n+1}\big)$. As a consequence, R_{n+1} belongs to $\mathcal{C}_{\mathrm{hs},\,n}$ and the sequence (R_n) satisfies the following condition, usually referred to as *Blackwell's condition*:

$$\big\langle R_{n+1} - \pi_{\mathcal{C}_{\mathrm{orth}}(a)}\big(\overline{R}_n\big),\ \overline{R}_n - \pi_{\mathcal{C}_{\mathrm{orth}}(a)}\big(\overline{R}_n\big)\big\rangle \leqslant 0\,.$$

This condition is trivially satisfied when $\overline{R}_n$ already belongs to $\mathcal{C}_{\mathrm{orth}}(a)$. Just as (2) leads to (3), this condition implies that $d_{\mathcal{C}_{\mathrm{orth}}(a)}\big(\overline{R}_n\big) \leqslant 2M/\sqrt{n}$.

Now, $(1/n)\sum_{t=1}^n r(x_t, y_t) \in (1/n)\sum_{t=1}^n \overline{m}(x_t, \boldsymbol{h}_t)$ and, as R is the upper-right corner function, $(1/n)\sum_{t=1}^n \overline{m}(x_t, \boldsymbol{h}_t) \subseteq \mathcal{C}_{\mathrm{orth}}\big(\overline{R}_n\big)$. That is, $(1/n)\sum_{t=1}^n r(x_t, y_t)$ is componentwise smaller than $\overline{R}_n$. Since the distance to the orthant $\mathcal{C}_{\mathrm{orth}}(a)$ equals, for all $\omega \in \mathbb{R}^d$,

$$d_{\mathcal{C}_{\mathrm{orth}}(a)}(\omega) = \sqrt{\sum_{k=1}^d \max\{\omega_k - a_k,\, 0\}^2}\,, \tag{7}$$

we get that

$$d_{\mathcal{C}_{\mathrm{orth}}(a)}\left(\frac{1}{n}\sum_{t=1}^n r(x_t, y_t)\right) \leqslant d_{\mathcal{C}_{\mathrm{orth}}(a)}\big(\overline{R}_n\big) \leqslant \frac{2M}{\sqrt{n}}\,.$$

Finally, by martingale convergence theorems (e.g., the Hoeffding–Azuma inequality), the sequence of the distances $d_{\mathcal{C}_{\mathrm{orth}}(a)}\big(\overline{r}_n\big)$ converges uniformly to 0 with respect to strategies of Nature. The above arguments are illustrated in Figure 3.

Flags not observed, only random signals observed, pure payoffs. It remains to relax the assumption that the flags $\boldsymbol{h}_n$ are observed, while only the signals s_n drawn at random according to $H(i_n, j_n)$ are. A standard trick in the literature of partial monitoring (see Kohlberg, 1975, Mertens et al., 1994, Lugosi et al., 2008) solves the issue, together with martingale convergence theorems and the fact that the upper-right corner R is a Lipschitz function for a well-chosen metric over sets (see Lemma 1 below). We briefly describe

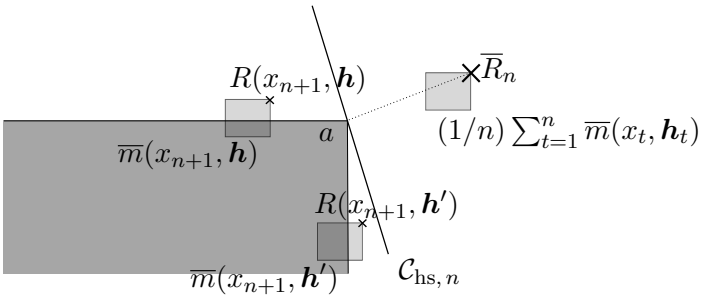


Figure 3: Illustration of the guarantee (6), in the case when the sets of compatible payoffs $\overline{m}(x, \boldsymbol{h})$ are all given by rectangles.

this trick without working out the lengthy details. Time is divided into blocks of time indexed by $b = 1, 2, \ldots$ and with respective (large and increasing) lengths L_b. Another sequence of elements $\gamma_b \in (0, 1)$ converging to 0 is needed. The same mixed distribution $x^{(b)}$ is played at all stages of block b by the player; that is, $x_{bL+t} = x^{(b)}$ for all $1 \leqslant t \leqslant L_b$. This distribution is chosen by mixing a distribution $x^{(b)}_{\text{orig.}}$ satisfying a constraint of the form of (6) with the uniform distribution. This is done with respective weights $1 - \gamma_b$ and γ_b. The distribution $x^{(b)}$ then puts a positive probability mass of at least $\gamma_b > 0$ on all actions. Doing so, an estimator of the average flag on block b can be constructed. Its accuracy, as well as the price to pay for the mixing, depend on γ_b and L_b. By such a price, we mean how farther away we are from $\mathcal{C}_{\text{orth}}(a)$ because we did not play $x^{(b)}_{\text{orig.}}$ but $x^{(b)}$ instead. Informally, each block now plays the role of a stage in the setting above when flags were observed. One can show that suitable values of L_b and γ_b lead to uniform convergence of the average payoffs, measured in terms of pure actions x_t, to $\mathcal{C}_{\text{orth}}(a)$. Also, similar martingale convergence arguments show that measuring payoffs in terms of the mixed actions x_t or pure actions i_t does not matter.

As indicated, we omit the technical proof of these facts (it already appeared in all the given references) but notice, however, that rates of convergence are adversely affected by this trick. □

Remark 1. The construction in the proof above shows that under the upper-right-corner property, it is necessary and sufficient to control the behavior of the upper-right corners R_n.

This property was used only to show that R_{n+1} is equal to some $r(x_{n+1}, y'_{n+1})$ and thus that the sequence (R_n) satisfies Blackwell's condition. When the property is not satisfied anymore, the sequence of the upper-right corners may fail to satisfy this condition. For instance, in the counter-example at the beginning of the present section, the upper-right corners equal

$$\forall\, \alpha \in [0, 1], \qquad R\big(\alpha T + (1 - \alpha)B,\ \emptyset\big) = (\alpha, 1 - \alpha)\,,$$

so that, for all strategies of the player, $\overline{R}_n = (\lambda, 1 - \lambda)$ for some $\lambda \in [0, 1]$. Thus, the distance of $\overline{R}_n$ to $\mathcal{C}_{\text{orth}}\big((0, 0)\big)$ is always larger than $1/\sqrt{2}$.

6. Intermezzo: Kohlberg's repeated games with incomplete information

We consider in this section a different, yet related framework, which is the main focus of Kohlberg (1975). We first describe a setting where d games with partial monitoring are to be played simultaneously, and then establish the formal connection with Kohlberg's results.

Simultaneous games with partial monitoring. We consider d such games, with common action sets $\mathcal{I}$ and $\mathcal{J}$ for the player and Nature and common set $\mathcal{H}$ of signals, but with possibly different payoff functions and signaling structures. We index these games by g. For each game $g \in \{1, \dots, d\}$, the player's payoff function is denoted by $r^{(g)} : \mathcal{I} \times \mathcal{J} \to \mathbb{R}$ and the signaling structure is given by $H^{(g)} : \mathcal{I} \times \mathcal{J} \to \Delta(\mathcal{H})$, with associated maximal information mapping $\boldsymbol{H}^{(g)} : \Delta(\mathcal{J}) \to \Delta(\mathcal{H})^{\mathcal{I}}$.

We put some restrictions on the strategies of the player and of Nature. The player may only choose one action $x_t \in \Delta(\mathcal{I})$ at each stage t, the same for all games g. On the other hand, Nature can choose different mixed actions $y_t^{(g)}$ in each game g, but these need to be *non-revealing*, that is, they need to induce the same flags. More formally, they need to be picked in the following set, which we assume to be non-empty:

$$\mathrm{NR} = \left\{ \big(y^{(1)}, \dots, y^{(d)}\big) \in \Delta(\mathcal{J})^d : \quad \boldsymbol{H}^{(1)}\big(y^{(1)}\big) = \cdots = \boldsymbol{H}^{(d)}\big(y^{(d)}\big) \right\}. \tag{8}$$

The above framework of simultaneous games can be embedded into an equivalent game that fits the model studied in the previous sections. Indeed, by linearity of each $\boldsymbol{H}^{(g)}$, the set NR of non-revealing actions is a polytope, thus it is the convex hull of its finite set of extremal points. We denote the cardinality of the latter by K and we write its elements as

$$\mathcal{K} = \left\{ \big(y_1^{(g)}\big)_{1 \leqslant g \leqslant d}, \ \dots, \ \big(y_K^{(g)}\big)_{1 \leqslant g \leqslant d} \right\}.$$

Each $\big(y^{(g)}\big)_{1 \leqslant g \leqslant d} \in \mathrm{NR}$ can then be represented by an element of $\Delta(\mathcal{K})$. Conversely, each $z = (z_k)_{k \leqslant K} \in \Delta(\mathcal{K})$ induces the following element of NR:

$$Y(z) = \big(Y^{(g)}(z)\big)_{1 \leqslant g \leqslant d} = \sum_{k=1}^{K} z_k \, \big(y_k^{(g)}\big)_{1 \leqslant g \leqslant d}.$$

So, with no loss of generality, we can assume that $\mathcal{K}$ is the finite set of actions of Nature and that, given $z \in \Delta(\mathcal{K})$ and $x \in \Delta(\mathcal{I})$, the payoff in the game g is $r^{(g)}\big(x, Y^{(g)}(z)\big)$.

This defines naturally an auxiliary game with linear vector-valued payoff function $r : \Delta(\mathcal{I}) \times \Delta(\mathcal{K}) \to \mathbb{R}^d$ and maximal information mapping $\boldsymbol{H} : \Delta(\mathcal{K}) \to \Delta(\mathcal{H})^{\mathcal{I}}$ defined by

$$r(x, z) = \Big(r^{(g)}\big(x, Y^{(g)}(z)\big) \Big)_{1 \leqslant g \leqslant d} \qquad \text{and} \qquad \boldsymbol{H}(z) = \boldsymbol{H}^{(g)}\big(Y^{(g)}(z)\big) \quad \text{for all } g.$$

(The definition of $\boldsymbol{H}$ is independent of g by construction, as we restricted Nature to use non-revealing strategies.) This maximal information mapping $\boldsymbol{H}$ corresponds to an underlying signaling structure which we denote by $H : \mathcal{I} \times \mathcal{K} \to \Delta(\mathcal{H})$.

The game (r, H) constructed above satisfies the upper-right-corner property. Indeed, for all $\boldsymbol{h} \in \boldsymbol{H}\big(\Delta(\mathcal{K})\big)$ and all $x \in \Delta(\mathcal{I})$,

$$\begin{aligned}\overline{m}(x, \boldsymbol{h}) &= \left\{ \Big(r^{(1)}\big(x, Y^{(1)}(z)\big), \ldots, r^{(d)}\big(x, Y^{(d)}(z)\big) \Big) : \quad z \in \Delta(\mathcal{K}) \text{ s.t. } \boldsymbol{H}(z) = \boldsymbol{h} \right\} \\ &= \left\{ \Big(r^{(1)}\big(x, y^{(1)}\big), \ldots, r^{(d)}\big(x, y^{(d)}\big) \Big) : \right. \\ &\qquad \left. \big(y^{(1)}, \ldots, y^{(d)}\big) \in \Delta(\mathcal{J})^d \text{ s.t. } \boldsymbol{H}^{(1)}\big(y^{(1)}\big) = \cdots = \boldsymbol{H}^{(d)}\big(y^{(d)}\big) = \boldsymbol{h} \right\}.\end{aligned}$$

Because of the separation of the variables in the constraint, the following set, given $\boldsymbol{h}$,

$$\left\{ \big(y^{(1)}, \ldots, y^{(d)}\big) \in \Delta(\mathcal{J})^d : \ \boldsymbol{H}^{(1)}\big(y^{(1)}\big) = \cdots = \boldsymbol{H}^{(d)}\big(y^{(d)}\big) = \boldsymbol{h} \right\}$$

is a Cartesian product of subsets of $\Delta(\mathcal{J})$. Thus, its image $\overline{m}(x, \boldsymbol{h})$ by the mapping

$$\big(y^{(1)}, \ldots, y^{(d)}\big) \in \Delta(\mathcal{J})^d \ \longmapsto \ \Big(r^{(1)}\big(x, y^{(1)}\big), \ldots, r^{(d)}\big(x, y^{(d)}\big) \Big)$$

is also a Cartesian product of closed intervals of $\mathbb{R}$. In particular, the latter set contains its upper-right corner, that is, $R(x, \boldsymbol{h}) \in \overline{m}(x, \boldsymbol{h})$, as claimed.

We assume, with no loss of generality, that in these simultaneous games, Nature maximizes the payoffs and the player minimizes them. A question that naturally arises—and whose answer will be needed below—is to determine for which vectors $a = (a_1, \ldots, a_d) \in \mathbb{R}^d$ the player can simultaneously guarantee that his average payoff will be smaller than a_g in the limit in each game $g \in \{1, \ldots, d\}$; that is, to determine which orthants $\mathcal{C}_{\text{orth}}(a)$ are (r, H)–approachable. By the exhibited upper-right-corner property, Proposition 2 shows that a necessary and sufficient condition for this is that all containing half-spaces of $\mathcal{C}_{\text{orth}}(a)$ be one-shot r–approachable. These half-spaces are parameterized by the convex distributions $q \in \Delta\big(\{1, \ldots, d\}\big)$ and are denoted by

$$\mathcal{C}_{\text{hs}}^{(q)} = \big\{ \omega \in \mathbb{R}^d : \ \langle \omega, q \rangle \leqslant \langle a, q \rangle \big\}. \tag{9}$$

Stated equivalently, the orthant $\mathcal{C}_{\text{orth}}(a)$ is (r, H)–approachable if and only if the value of the zero-sum game with payoff function $(x, z) \in \Delta(\mathcal{I}) \times \Delta(\mathcal{K}) \mapsto \langle r(x, z),\, q \rangle$ is smaller than $\langle a, q \rangle$ for all $q \in \Delta(\{1, \ldots, d\})$.

Kohlberg's model of repeated games with incomplete information. The setting of repeated games with incomplete information, introduced by Aumann and Maschler (1995), relies on the same finite family of games $\big(r^{(g)},\, H^{(g)}\big)$, where $g \in \{1, \ldots, d\}$,

as described above. They will however not be played simultaneously. Instead, a single game (state) $G \in \{1, \dots, d\}$ is drawn according to some probability distribution $p \in \Delta\big(\{1, \dots, d\}\big)$ known by both the player and Nature. Yet only Nature (and not the player) is informed of the true state G. A repeated game with partial monitoring then takes place between the player and Nature in the G–th game. Payoffs are evaluated in expectation with respect to the random draw of G according to p.

For simplicity, we assume that all mappings $\boldsymbol{H}^{(g)}$ have the same range[2] and, with no loss of generality, that p has full support. Because of these two properties, the considered setting of repeated games with incomplete information can then be embedded, from the player's viewpoint, into the above-described setting of d simultaneous games under the restriction that Nature resorts to non-revealing strategies. Indeed, from the player's viewpoint and because of the identical range of the $\boldsymbol{H}^{(g)}$, the mixed action used by Nature in the game G can be thought of as the G–th component of some vector of mixed actions in the set NR defined in (8). We use the notation defined above: as payoffs are evaluated in expectation, the payoff function is formed by the inner products $(x, z) \in \Delta(\mathcal{I}) \times \Delta(\mathcal{K}) \mapsto \langle r(x, z),\, p\rangle$.

We recall that Nature maximizes the payoff and that the player minimizes it. For each $q \in \Delta(\{1, \dots, d\})$, we denote by $u(q)$ the value of the one-shot zero-sum game $\Gamma(q)$ with payoffs $(x, z) \in \Delta(\mathcal{I}) \times \Delta(\mathcal{K}) \mapsto \langle r(x, z),\, q\rangle$. We show that, as proved in the mentioned references, the value U of this repeated game, as a function of the distribution p, may be larger than $u(p)$ and is given by $\mathrm{cav}[u](p)$, where $\mathrm{cav}[u]$ is the smallest concave function above u.

First, the so-called splitting lemma shows that U is concave. Therefore, we have $U \geqslant \mathrm{cav}[u]$. (For the splitting lemma, see Aumann and Maschler, 1995 and also Mertens et al., 1994, Section V.1 or Sorin, 2002.) The inequality of interest to us is the converse one. Using the concavity of the mapping $\mathrm{cav}[u]$, Kohlberg (1975, Corollary 2.4) proves that for all $p \in \Delta(\{1, \dots, d\})$, there exists some $a_p \in \mathbb{R}^d$ such that

$$\mathrm{cav}[u](p) = \langle a_p, p\rangle \qquad \text{and} \qquad \forall\, q \in \Delta(\{1, \dots, d\}), \quad \mathrm{cav}[u](q) \leqslant \langle a_p, q\rangle\,.$$

In particular, $u(q) \leqslant \langle a_p, q\rangle$ for all $q \in \Delta(\{1, \dots, d\})$. The equivalence stated after (9) shows that $\mathcal{C}_{\mathrm{orth}}(a_p)$ is therefore (r, H)–approachable. Hence, no matter the strategy of Nature, the payoff in state G is asymptotically smaller than the G–th component of a_p. (This is true for all realizations of G.) As a consequence, in expectation (with respect to the random choice of G), the payoff is smaller than $\langle a_p,\, p\rangle = \mathrm{cav}[u](p)$. This shows that $U(p) \leqslant \mathrm{cav}[u](p)$.

In conclusion, Kohlberg (1975) implicitly used the consequences of the upper-right-corner property detailed above when constructing an optimal strategy for the uninformed player. A close inspection reveals that Lemma 5.4 therein does not hold anymore in

[2]In full generality, when this is not the case, Nature may resort to strategies that reveal that the true state G belongs to some strict subset of $\{1, \dots, d\}$, and the player must adapt his strategy in correspondence with this knowledge, see Kohlberg (1975). But our assumption already captures the basic idea of the use of approachability in this framework and the alluded technical adaptations are beyond the scope of this paper.

the more general framework without the upper-right-corner property (in particular, one might want to read it again with Remark 1 in mind).

7. Primal approachability of orthants in the general case

We noted that the primal characterization in terms of one-shot r–approachability of containing half-spaces stated in Proposition 2 did not extend to games (r, H) without the upper-right-corner property. We show in this section that it holds true in the general case when one-shot approachability is with respect to the modified payoff function $\widetilde{r}_H : \Delta(\mathcal{I}) \times \Delta(\mathcal{J}) \to \mathbb{R}^d$ defined as follows:

$$\forall\, x \in \Delta(\mathcal{I}),\ \forall\, y \in \Delta(\mathcal{J}), \qquad \widetilde{r}_H(x, y) = R\big(x, \boldsymbol{H}(y)\big)\,.$$

The change of payoff function can be intuitively explained as follows. As noted in Section 5, when the target sets are given by orthants (and only because of this), the behavior of (averages of) sets of compatible payoffs is dictated by their upper-right corners. Now, the upper-right-corner property indicated that even when measuring payoffs with r, the worst-case payoffs were given by the upper-right corners and that it was thus necessary and sufficient to consider the latter. If this property does not hold, then evaluating actions with $\widetilde{r}_H$ enables and forces the consideration of these corners.

Of course, in the case of full monitoring, as follows from the comments after Definition 2, no modification takes place in the payoff function, that is, $\widetilde{r}_{\text{Full}} = r$.

The main result of this section is the following primal characterization. The rest of this section will then show how it leads to a new approachability strategy under partial monitoring, based on surrogate payoffs (upper-right corner payoffs) and not only on signals or on estimated flags, as previously done in the literature (e.g., in the references mentioned in the last part of the proof of Proposition 2).

Theorem 2. *For all games (r, H) with partial monitoring, for all orthants $\mathcal{C}_{\text{orth}}(a)$, where $a \in \mathbb{R}^d$,*

$$\begin{aligned} & \mathcal{C}_{\text{orth}}(a) \text{ is } (r, H)\text{–approachable} \\ \iff\ & \text{every half-space } \mathcal{C}_{\text{hs}} \supset \mathcal{C}_{\text{orth}}(a) \text{ is one-shot } \widetilde{r}_H\text{–approachable.} \end{aligned}$$

The proof of this theorem is as follows. The dual characterization (5) indicates that a necessary and sufficient condition of (r, H)–approachability for $\mathcal{C}_{\text{orth}}(a)$ is that for all $y \in \Delta(\mathcal{J})$, there exists $x \in \Delta(\mathcal{I})$ such that $\overline{m}\big(x, \boldsymbol{H}(y)\big) \subseteq \mathcal{C}_{\text{orth}}(a)$. Since, by construction of R, the smallest orthant (in the sense of inclusion) in which $\overline{m}\big(x, \boldsymbol{H}(y)\big)$ is contained is precisely $\mathcal{C}_{\text{orth}}\big(\widetilde{r}_H(x, y)\big)$, the necessary and sufficient condition can be restated as the requirement that for all $y \in \Delta(\mathcal{J})$, there exists $x \in \Delta(\mathcal{I})$ such that $\widetilde{r}_H(x, y) \in \mathcal{C}_{\text{orth}}(a)$. Now, this reformulated dual characterization of approachability in the context of orthants is seen to be equivalent to the following primal characterization, which concludes the proof of the theorem.

Proposition 3. *For all games (r, H) with partial monitoring, for all orthants $\mathcal{C}_{\mathrm{orth}}(a)$, where $a \in \mathbb{R}^d$,*

$$\begin{aligned} & \forall\, y \in \Delta(\mathcal{J}),\ \exists\, x \in \Delta(\mathcal{I}), \quad \widetilde{r}_H(x,y) \in \mathcal{C}_{\mathrm{orth}}(a) \\ \iff \quad & \textit{every half-space } \mathcal{C}_{\mathrm{hs}} \supset \mathcal{C}_{\mathrm{orth}}(a) \textit{ is one-shot } \widetilde{r}_H\textit{–approachable.} \end{aligned}$$

Before proving this proposition, we need to state some properties of the function $\widetilde{r}_H$. Given two points $a, a' \in \mathbb{R}^d$, the notation $a \preccurlyeq a'$ means that a is component-wise smaller than a'—or equivalently, that a belongs to the orthant $\mathcal{C}_{\mathrm{orth}}(a')$.

Lemma 1. *The function $\widetilde{r}_H$ is Lipschitz continuous. It is also convex in its first argument and concave in its second argument, in the sense that, for all $x, x' \in \Delta(\mathcal{I})$, all $y, y' \in \Delta(\mathcal{J})$, and all $\lambda \in [0,1]$,*

$$\begin{aligned} & \widetilde{r}_H\big(\lambda x + (1-\lambda)x',\, y\big) \preccurlyeq \lambda\, \widetilde{r}_H(x,y) + (1-\lambda)\, \widetilde{r}_H(x',y) \\ \textit{and} \qquad & \lambda\, \widetilde{r}_H(x,y) + (1-\lambda)\, \widetilde{r}_H(x,y') \preccurlyeq \widetilde{r}_H\big(x,\, \lambda y + (1-\lambda)y'\big)\,. \end{aligned}$$

Proof. Convexity and concavity follow from the concavity and the convexity of $\overline{m}$ for inclusion. Formally, it follows from the very definition (1) of $\overline{m}$ and from the linearity of r and $\boldsymbol{H}$ that, for all $x, x' \in \Delta(\mathcal{I})$, all $\boldsymbol{h}, \boldsymbol{h}' \in \mathcal{F}$, and $\lambda \in [0,1]$,

$$\begin{aligned} & \overline{m}\big(\lambda x + (1-\lambda)x',\, \boldsymbol{h}\big) \subseteq \lambda\, \overline{m}(x,\boldsymbol{h}) + (1-\lambda)\, \overline{m}(x',\boldsymbol{h}) \\ \text{and} \qquad & \lambda\, \overline{m}(x,\boldsymbol{h}) + (1-\lambda)\, \overline{m}(x,\boldsymbol{h}') \subseteq \overline{m}\big(x,\, \lambda\boldsymbol{h} + (1-\lambda)\boldsymbol{h}'\big)\,. \end{aligned}$$

The second part of the lemma follows by taking upper-right corners, which is a linear and non-decreasing operation (for the respective partial orders $\subseteq$ and $\preccurlyeq$).

As for the Lipschitz property of $\widetilde{r}_H$, it follows from a rewriting of $\overline{m}\big(x, \boldsymbol{H}(y)\big)$ as

$$\overline{m}\big(x, \boldsymbol{H}(y)\big) = \sum_{b \in \mathcal{B}} \phi_b\big(\boldsymbol{H}(y)\big)\, r\big(x,\, \boldsymbol{H}^{-1}(b)\big)\,,$$

where $\mathcal{B}$ is a finite subset of $\mathcal{F}$, the ϕ_b are Lipschitz functions $\mathcal{F} \to [0,1]$, and $\boldsymbol{H}^{-1}$ is the pre-image function of $\boldsymbol{H}$, which takes values in the set of compact subsets of $\Delta(\mathcal{J})$. This rewriting was proved in Mannor et al. (2011, Lemma 6.1 and Remark 6.1). We equip the set of compact subsets of the Euclidian ball with center $(0, \dots 0)$ and radius M, in which $\overline{m}$ takes its values, with the Hausdorff distance. For this distance, $x \mapsto r\big(x,\, \boldsymbol{H}^{-1}(b)\big)$ is M–Lipschitz for each $b \in \mathcal{B}$. All in all, given the boundedness of the ϕ_b and of r, the mapping $(x,y) \mapsto \overline{m}\big(x, \boldsymbol{H}(y)\big)$ is also Lipschitz continuous. Since taking the upper-right corner is a Lipschitz mapping as well (for the Hausdorff distance), we get, by composition, the desired Lipschitz property for $\widetilde{r}_H$. □

We are now ready to prove Proposition 3. (Note that it needs a proof and that it is not implied by the various results discussed in Section 5. Indeed, $(\widetilde{r}_H, H)$ is an auxiliary game which, by construction, has the upper-right-corner property, but $\widetilde{r}_H$ is not linear, while linearity of the payoff function was a crucial feature of the setting studied therein.)

Proof of Proposition 3. We start with the direct implication and consider some containing half-space $\mathcal{C}_{\mathrm{hs}}$. The latter is parameterized by $\alpha \in \mathbb{R}^d$ and $\beta \in \mathbb{R}$, and equals

$$\mathcal{C}_{\mathrm{hs}} = \left\{\omega \in \mathbb{R}^d : \ \langle \alpha, \omega \rangle \leqslant \beta\right\}.$$

Since $\mathcal{C}_{\mathrm{hs}}$ contains the orthant $\mathcal{C}_{\mathrm{orth}}(a)$, there are sequences (ω_n) in $\mathcal{C}_{\mathrm{hs}}$ with components tending to $-\infty$. Therefore, we necessarily have that $\alpha \succcurlyeq 0$. The convexity/concavity of $\widetilde{r}_H$ in the sense of $\preccurlyeq$ thus entails that the function $G_{\alpha,\beta} : (x, y) \longmapsto \left\langle \alpha, \widetilde{r}_H(x, y)\right\rangle - \beta$ is also convex/concave. The continuity of $G_{\alpha,\beta}$ follows from the one of $\widetilde{r}_H$. The Sion–Fan lemma applies and guarantees that

$$\min_{x\in\Delta(\mathcal{I})} \max_{y\in\Delta(\mathcal{J})} G_{\alpha,\beta}(x, y) = \max_{y\in\Delta(\mathcal{J})} \min_{x\in\Delta(\mathcal{I})} G_{\alpha,\beta}(x, y)\,,$$

(the suprema and infima are all attained and are denoted by maxima and minima). Now, by assumption, for all $y \in \Delta(\mathcal{J})$, there exists $x \in \Delta(\mathcal{I})$ such that $\widetilde{r}_H(x, y) \in \mathcal{C}_{\mathrm{orth}}(a)$. This means that the above $\max\min G_{\alpha,\beta}$ is non-positive. Putting all things together, we have proved that

$$\min_{x\in\Delta(\mathcal{I})} \max_{y\in\Delta(\mathcal{J})} G_{\alpha,\beta}(x, y) \leqslant 0\,.$$

That is, there exists $x_0 \in \Delta(\mathcal{I})$, e.g., the element attaining the above maximum, such that for all $y \in \Delta(\mathcal{J})$, one has $G_{\alpha,\beta}(x_0, y) \leqslant 0$, or, equivalently, $\widetilde{r}_H(x_0, y) \in \mathcal{C}_{\mathrm{orth}}(a)$. This property is exactly the stated one-shot $\widetilde{r}_H$–approachability of $\mathcal{C}_{\mathrm{orth}}(a)$.

Conversely, assume that there exists some $y_0 \in \Delta(\mathcal{J})$ such that, for all $x \in \Delta(\mathcal{I})$, one has $\widetilde{r}_H(x, y_0) \notin \mathcal{C}_{\mathrm{orth}}(a)$. By continuity of $\widetilde{r}_H$ and closedness of $\mathcal{C}_{\mathrm{orth}}(a)$, there exists some $\delta > 0$ such that $d_{\mathcal{C}_{\mathrm{orth}}(a)}\big(\widetilde{r}_H(x, y_0)\big) \geqslant \delta$ for all $x \in \Delta(\mathcal{I})$. Now, as indicated around (7), the distance to $\mathcal{C}_{\mathrm{orth}}(a)$ is non-decreasing for $\preccurlyeq$. In view of the convexity of $\widetilde{r}_H$ in its first argument, this shows that we also have $d_{\mathcal{C}_{\mathrm{orth}}(a)}(z) \geqslant \delta$ for all elements z of the convex hull $\mathcal{C}_{\widetilde{r}_H,y_0}$ of the set $\left\{\widetilde{r}_H(x, y_0) : \ x \in \Delta(\mathcal{I})\right\}$. That is, the closed convex sets $\mathcal{C}_{\widetilde{r}_H,y_0}$, which is compact, and $\mathcal{C}_{\mathrm{orth}}(a)$, which is closed, are disjoint and thus, by the Hahn–Banach theorem, are strictly separated by some hyperplane. One of the two half-spaces thus defined, namely, the one not containing $\mathcal{C}_{\widetilde{r}_H,y_0}$, is not one-shot $\widetilde{r}_H$–approachable. □

A new approachability strategy of an orthant under partial monitoring

Theorem 2 suggests an approachability strategy based on surrogate payoffs and not only on the information gained, i.e., based on the mapping $\widetilde{r}_H$ and not only on the signaling structure H (and the estimated flags). The first approach was already considered by Kohlberg (1975) while other works, like Perchet (2011a) and Mannor et al. (2011), resorted to the second one. The considered strategy is an adaptation of Blackwell's strategy (which was recalled after the statement of Theorem 1): such an adaptation is possible as the latter strategy only relies on the one-shot approachability of half-spaces, which is satisfied here with the surrogate payoffs $\widetilde{r}_H$.

Description and convergence analysis of the strategy. As in the proof of Proposition 2, we assume initially that flags $\boldsymbol{h}_t = \boldsymbol{H}(y_t)$ are observed at the end of each round t and that mixed payoffs are to be controlled. The player then knows his mixed payoffs $\widetilde{r}_{H,t} := \widetilde{r}_H(x_t, y_t) = R\big(x_t, \boldsymbol{h}_t\big)$ and aims at controlling his average payoffs, which we recall are denoted by $\overline{\widetilde{r}}_{H,n}$. Similarly to what was done in the proof of Proposition 2, the one-shot $\widetilde{r}_H$–approachability of the containing half-spaces of $\mathcal{C}_{\mathrm{orth}}(a)$ entails that for each round n, there exists $x_{n+1} \in \Delta(\mathcal{I})$ such that

$$\forall\, y \in \Delta(\mathcal{J}), \qquad \Big\langle \widetilde{r}_H(x_{n+1}, y) - \pi_{\mathcal{C}_{\mathrm{orth}}(a)}\big(\overline{\widetilde{r}}_{H,n}\big),\ \overline{\widetilde{r}}_{H,n} - \pi_{\mathcal{C}_{\mathrm{orth}}(a)}\big(\overline{\widetilde{r}}_{H,n}\big) \Big\rangle \leqslant 0\,.$$

The sequence $\big(\overline{\widetilde{r}}_{H,n}\big)$ thus satisfies Blackwell's condition and as a result we get

$$d_{\mathcal{C}_{\mathrm{orth}}(a)}\left(\frac{1}{n}\sum_{t=1}^{n} \widetilde{r}_H(x_t, y_t)\right) \leqslant \frac{2M\sqrt{d}}{\sqrt{n}}\,.$$

(See again the proof of Proposition 2 for this derivation and keep in mind that in the present setting where the upper-right-corner property is not satisfied, R is only bounded in ℓ_2–norm by $M\sqrt{d}$.) Since $r(x_t, y_t) \in \overline{m}\big(x_t, \boldsymbol{H}(y_t)\big) \subseteq \mathcal{C}_{\mathrm{orth}}\big(\widetilde{r}_H(x_t, y_t)\big)$, we get $r(x_t, y_t) \preccurlyeq \widetilde{r}_H(x_t, y_t)$ and, in view again of (7),

$$d_{\mathcal{C}_{\mathrm{orth}}(a)}\left(\frac{1}{n}\sum_{t=1}^{n} r(x_t, y_t)\right) \leqslant d_{\mathcal{C}_{\mathrm{orth}}(a)}\left(\frac{1}{n}\sum_{t=1}^{n} \widetilde{r}_H(x_t, y_t)\right) \leqslant \frac{2M\sqrt{d}}{\sqrt{n}}\,.$$

The same trick of playing i.i.d. in blocks as in the second part of the proof of Proposition 2, together with martingale convergence arguments, relaxes the assumptions of flags being observed and payoffs being evaluated with mixed actions, leading to the desired (r, H)–approachability strategy. (This is where we need the Lipschitzness properties stated in Lemma 1 and its proof.) A more careful study, which we omit here for the sake of brevity, shows that (r, H)–approachability takes place at a $n^{-1/5}$–rate.

What we proved in passing. We proved in a constructive way that when an orthant is (r, H)–approachable, it is also $\big(\widetilde{r}_H, \mathrm{Full}\big)$–approachable.

Conversely, assume that the equivalent conditions in Theorem 2 are not satisfied, i.e., that the orthant at hand, $\mathcal{C}_{\mathrm{orth}}(a)$, is not (r, H)–approachable. Then, (the proof of) Proposition 3 indicates that there exists some $y_0 \in \Delta(\mathcal{J})$ such that the set $\mathcal{C}_{\mathrm{orth}}(a)$ and the convex hull of $\big\{\widetilde{r}_H(x, y_0),\ x \in \Delta(\mathcal{I})\big\}$ are strictly separated. This implies in particular that $\mathcal{C}_{\mathrm{orth}}(a)$ is $\big(\widetilde{r}_H, \mathrm{Full}\big)$–excludable, and thus is not $\big(\widetilde{r}_H, \mathrm{Full}\big)$–approachable.

Putting all things together, we have proved the following equivalence:

$$\begin{aligned} & \mathcal{C}_{\mathrm{orth}}(a) \text{ is } (r, H)\text{–approachable} \\ \iff\ & \mathcal{C}_{\mathrm{orth}}(a) \text{ is } \big(\widetilde{r}_H, \mathrm{Full}\big)\text{–approachable} \\ \iff\ & \mathcal{C}_{\mathrm{orth}}(a) \text{ is not } \big(\widetilde{r}_H, \mathrm{Full}\big)\text{–excludable.} \end{aligned}$$

Note that the $\big(\widetilde{r}_H, \mathrm{Full}\big)$–approachability is a form of non-linear approachability, by which we mean that the function $\widetilde{r}_H$ is not linear and yet, approachability is possible. This result could be generalized (but we omit the description of the extension for the sake of concision).

On the computational complexity of the above-described strategy. The strategy we have exhibited reduces to solving, at each stage, a program of the form

$$\min_{x_{n+1}\in\Delta(\mathcal{I})}\left\{\max_{y\in\Delta(\mathcal{J})}\,\big\langle\widetilde{r}_H(x_{n+1},y)-\beta,\,\alpha\big\rangle\right\}$$

for some vectors $\alpha,\beta\in\mathbb{R}^d$. At first sight, it cannot be written as a finite linear program as $\widetilde{r}_H$ is not a linear function of its arguments. However, as proved in Mannor et al. (2011, Section 7.1), the function $\widetilde{r}_H$ is actually piecewise linear; that is, there exist some finite liftings of $\Delta(\mathcal{I})$ and $\Delta(\mathcal{J})$ with respect to which $\widetilde{r}_H$ is linear. (These liftings only need to be computed once, before the game starts.) Moreover, the per-step computational complexity of our strategy is constant (in fact, it is polynomial in the sizes of these liftings; see Mannor et al., 2011 for more details).

8. Primal approachability of polytopes

Recall that a convex set $\mathcal{C}_{\text{polyt}}$ is a polytope if it is the intersection of a finite number of half-spaces $\big\{\omega\in\mathbb{R}^d:\ \langle\omega,a_\ell\rangle\leqslant b_\ell\big\}$, for $a_\ell,b_\ell\in\mathbb{R}^d$ and ℓ ranging in some finite set $\mathcal{L}$. That is,

$$\mathcal{C}_{\text{polyt}}=\bigcap_{\ell\in\mathcal{L}}\Big\{\omega\in\mathbb{R}^d:\ \langle\omega,a_\ell\rangle\leqslant b_\ell\Big\}=\Big\{\omega\in\mathbb{R}^d:\ \max_{\ell\in\mathcal{L}}\langle\omega,a_\ell\rangle-b_\ell\leqslant 0\Big\}\,. \tag{10}$$

The following lemma (which is a mere exercice of rewriting) states that an approachability problem of a polytope can be transformed into an approachability problem of some negative orthant. We denote by $(0)_{\mathcal{L}}=(0,\dots,0)$ the null vector of $\mathbb{R}^{\mathcal{L}}$. The negative orthant of $\mathbb{R}^{\mathcal{L}}$ is then denoted by $\mathcal{C}_{\text{orth}}\big((0)_{\mathcal{L}}\big)$.

Lemma 2. *The convex polytope $\mathcal{C}_{\text{polyt}}$ defined in (10) is (r,H)–approachable if and only if the negative orthant $\mathcal{C}_{\text{orth}}\big((0)_{\mathcal{L}}\big)$ is (s,H)–approachable, where the vector-valued payoff function $s:\Delta(\mathcal{I})\times\Delta(\mathcal{J})\to\mathbb{R}^{\mathcal{L}}$ is defined as*

$$s(x,y)=\Big[\langle r(x,y),a_\ell\rangle-b_\ell\Big]_{\ell\in\mathcal{L}}\,. \tag{11}$$

Proof. The result follows from the equivalence (see, e.g., property 3 in Appendix A.1 of Perchet, 2011b) of the distances to $\mathcal{C}_{\text{polyt}}$ given by

$$d_{\mathcal{C}_{\text{polyt}}}\qquad\text{and}\qquad d_{\mathcal{C}_{\text{orth}}((0)_{\mathcal{L}})}\big(T(\,\cdot\,)\big)\,,$$

where $T:\mathbb{R}^d\to\mathbb{R}^{\mathcal{L}}$ is the linear transformation $\omega\mapsto\big[\langle\omega,a_\ell\rangle-b_\ell\big]_{\ell\in\mathcal{L}}$. □

Theorem 2 can then be rewritten, using Lemma 2 above, to provide the desired primal characterization of polytopes.

Corollary 2. *Consider the convex polytope $\mathcal{C}_{\text{polyt}}$ given by* (10)*, together with the payoff function s defined in* (11)*. Then,*

$$\begin{aligned} & \mathcal{C}_{\text{polyt}} \text{ is } (r,H)\text{–approachable} \qquad (12)\\ \iff \quad & \text{every containing half-space of } \mathcal{C}_{\text{orth}}\big((0)_{\mathcal{L}}\big) \text{ is one-shot } \widetilde{s}_H\text{–approachable.} \end{aligned}$$

When $\mathcal{C}_{\text{polyt}}$ is indeed (r,H)–approachable, the results of the previous section provide an approachability strategy of $\mathcal{C}_{\text{orth}}\big((0)_{\mathcal{L}}\big)$ based on the transformed payoffs $\widetilde{s}_H$. This strategy also approaches $\mathcal{C}_{\text{polyt}}$ in view of Lemma 2, however it might not be representable in the original space $\mathbb{R}^d$, as demonstrated in the following (counter-)example.

Example 1. Consider on the one hand the polytope $\mathcal{C}_{\text{polyt}} = \big\{\omega \in \mathbb{R} :\ \omega \in [-1,1]\big\}$ and the associated linear transformation T defined by $T(\omega) = (\omega - 1,\, -\omega - 1) \in \mathbb{R}^2$ for all $\omega \in \mathbb{R}$. Consider on the other hand the following game. The sets of pure actions are $\mathcal{I} = \{T, B\}$ and $\mathcal{J} = \{L, R\}$, the signaling structure is $H = \text{Dark}$ (with single signal denoted by $\emptyset$), and the payoff function r is given by the matrix

	L	R
T	-1	2
B	-2	1

We identify $\Delta(\mathcal{I})$ and $\Delta(\mathcal{J})$ with $[0,1]$.

We first discuss the dual condition (5) for (r,Dark)–approachability. For all $x \in [0,1]$, we have $\overline{m}(x,\emptyset) = [-2+x,\, 1+x]$. Thus, no mixed action x of the player is such that $\overline{m}(x,\emptyset)$ is included in $\mathcal{C}_{\text{polyt}}$, which is therefore not (r,Dark)–approachable.

We now turn to the primal condition as stated by Corollary 2. We denote by $T = (T_1, T_2)$ the components of the linear transformation T. From the linearity of T, we deduce from the above-stated expression of $\overline{m}$ (based on r) that the sets of compatible payoffs in terms of $s = T(r)$ are of the form $T\big(\overline{m}(x,\emptyset)\big)$. Taking the maxima, we thus get, for all mixed actions $x \in [0,1]$ (and all $y \in [0,1]$ as the game is played in the dark),

$$\widetilde{s}_{\text{Dark}}(x,y) = \Big(\max T_1\big([-2+x,\, 1+x]\big),\ \max T_2\big([-2+x,\, 1+x]\big)\Big) = (x, 1-x)\,.$$

Again, the necessary and sufficient condition for (r,H)–approachability of $\mathcal{C}_{\text{polyt}}$ fails, as no containing half-space of the negative orthant but two of them is one-shot $\widetilde{s}_{\text{Dark}}$–approachable. More precisely, these half-spaces are parameterized by $(p, 1-p)$ where $p \in [0,1]$ and correspond to the points $\big\{(t_1,t_2) \in \mathbb{R}^2 :\ p_1 t_1 + p_2 t_2 \leqslant 0\big\}$. Except for the case when $p = 0$ or $p = 1$, these half-spaces are strictly separated from the convex set $\big\{(x, 1-x) :\ x \in [0,1]\big\}$.

The question now is whether we could have determined this by satisfying some primal condition in the original space $\mathbb{R}$. First, consider some containing half-space of $\mathcal{C}_{\text{polyt}}$, typically, either $(-\infty, 1]$ or $[-1, +\infty)$. Their transformations by T into subsets of $\mathbb{R}^2$ are included respectively in $(-\infty, 0] \times \mathbb{R}$ or $\mathbb{R} \times (-\infty, 0]$. These are precisely the only two half-spaces that were one-shot $\widetilde{s}_{\text{Dark}}$–approachable (by resorting to one of the pure actions). Now, and more importantly, consider the containing half-space of the negative

orthant in $\mathbb{R}^2$ parameterized by $p = 1/2$, that is, $\mathcal{C}_{\mathrm{hs},\,1/2} = \big\{(t_1, t_2) \in \mathbb{R}^2 :\ t_1 + t_2 \leqslant 0\big\}$. As indicated above, it is not one-shot $\widetilde{s}_{\mathrm{Dark}}$–approachable. However, this half-space contains all the original space, in the sense that $T(\mathbb{R}) \subset \mathcal{C}_{\mathrm{hs},\,1/2}$, as follows from simple computations: $T(\omega) = (\omega - 1) + (-\omega - 1) = -2$. Therefore, there is no hope to prove, based even on general subsets of the original game with payoffs in $\mathbb{R}$, that the necessary and sufficient condition on the half-space $\mathcal{C}_{\mathrm{hs},\,1/2}$ in the transformed space $\mathbb{R}^2$ fails.

The fundamental reason why the primal characterization in the transformed space cannot be checked based on considerations in the original space is the following. In the absence of a upper-right-corner property, the range of $\widetilde{s}_{\mathrm{Dark}}$ is outside the range of s but we can only access to the latter based on the original space. The moral of this example is that we have to consider some *hidden* containing half-spaces of the polytope $\mathcal{C}_{\mathrm{polyt}}$ in order to establish some primal characterization: this is precisely what Condition (12) does.

9. Primal approachability of general convex sets

We consider in this section the primal approachability of general closed convex sets $\mathcal{C}$. In the case of polytopes, Lemma 2 was essentially indicating that only finitely many directions in $\mathbb{R}^d$ (the ones given by the a_ℓ) need be considered. In the case of general convex sets, all directions are to be studied. We do so by resorting to support functions, which we define based on the unit Euclidean sphere $\mathcal{S} = \big\{\omega \in \mathbb{R}^d :\ \|\omega\| = 1\big\}$. More formally, the support function $\phi_{\mathcal{C}} : \mathcal{S} \to \mathbb{R} \cup \{+\infty\}$ of a set $\mathcal{C} \subseteq \mathbb{R}^d$ is defined by

$$\forall s \in \mathcal{S}, \qquad \phi_{\mathcal{C}}(s) = \sup\big\{\langle c, s\rangle :\ c \in \mathcal{C}\big\}\,.$$

We now construct a lifted setting in which one-shot approaching the containing half-spaces for some payoff function will be equivalent to (r, H)–approaching the original closed convex set $\mathcal{C}$. This setting is given by some set of integrable functions on $\mathcal{S}$. We equip the latter with the (induced) Lebesgue measure, for which $\mathcal{S}$ has a finite measure. That is, we consider the set $\mathbb{L}^2(\mathcal{S})$ of Lebesgue square integrable functions $\mathcal{S} \to \mathbb{R}$, equipped with the inner product

$$(f, g) \in \mathbb{L}^2(\mathcal{S}) \times \mathbb{L}^2(\mathcal{S}) \longmapsto \int_{\mathcal{S}} f(s)\, g(s)\, \mathrm{d}s\,.$$

The orthant in $\mathbb{L}^2(\mathcal{S})$ corresponding to $\mathcal{C} \subseteq \mathbb{R}^d$ is

$$\mathcal{C}_{\mathrm{orth}}(\phi_{\mathcal{C}}) = \big\{f \in \mathbb{L}^2(\mathcal{S}) :\ f \leqslant \phi_{\mathcal{C}}\big\}\,.$$

The description of the lifted setting is concluded by stating the considered payoff function $\Phi : \Delta(\mathcal{I}) \times \Delta(\mathcal{J}) \to \mathbb{L}^2(\mathcal{S})$. It indicates, as in the previous sections, how to transform payoffs given the signaling structure H. Formally,

$$\forall\, x \in \Delta(\mathcal{I}),\ \forall\, y \in \Delta(\mathcal{J}), \qquad \Phi(x, y) = \phi_{\overline{m}(x, \boldsymbol{H}(y))}\,.$$

The square integrability of $\Phi(x,y)$ follows its boundedness, which itself stems from the boundedness of $\overline{m}\big(x,\boldsymbol{H}(y)\big)$. (See Lemma 3 in appendix, property 1, for a reminder of this well-known result and others on support functions.)

We are now ready to state the primal characterization of approachability with partial monitoring in the general case.

Theorem 3. *For all games (r,H), for all closed convex sets $\mathcal{C}\subset\mathbb{R}^d$,*

$$\begin{aligned} &\mathcal{C} \text{ is } (r,H)\text{–approachable} \\ \iff\quad &\text{every half-space } \mathcal{C}_{\rm hs}\supset\mathcal{C}_{\rm orth}(\phi_{\mathcal{C}}) \text{ is one-shot } \Phi\text{–approachable.}\end{aligned}$$

Proof. We first note that we can assume with no loss of generality that $\mathcal{C}$ is bounded thus compact. Indeed, $\mathcal{C}$ is (r,H)–approachable if and only if its intersection $\mathcal{C}\cap r\big(\Delta(\mathcal{I}\times\mathcal{J})\big)$ with the bounded convex set of feasible payoffs is approachable. This entails that $\phi_{\mathcal{C}}\in\mathbb{L}^2(\mathcal{S})$.

Now, the proof follows along the lines of the proof of Theorem 2. In particular, we exploit the dual characterization (5), that indicates that for all $y\in\Delta(\mathcal{J})$, there exists $x\in\Delta(\mathcal{I})$ such that $\overline{m}\big(x,\boldsymbol{H}(y)\big)\subseteq\mathcal{C}$. It can be restated equivalently (see Lemma 3 in appendix, property 3) as stating that for all $y\in\Delta(\mathcal{J})$, there exists $x\in\Delta(\mathcal{I})$ such that $\Phi(x,y)\leqslant\phi_{\mathcal{C}}$. We thus only need to show that the stated primal characterization is equivalent to the latter condition.

We start with the direct implication (from the dual condition to the primal condition). As recalled in the proof of Lemma 1, the function $\overline{m}$ is concave/convex, which, together with properties 3 and 4 of Lemma 3, shows that Φ is also convex/concave. Moreover, as proved at the end of the proof of Lemma 1, the function $(x,y)\mapsto\overline{m}\big(x,\boldsymbol{H}(y)\big)$ is a Lipschitz function, with Lipschitz constant denoted by $L_{\overline{m}}$, when the set of compact subsets of the Euclidean ball of $\mathbb{R}^d$ with center $(0,\dots,0)$ and radius M is equipped with the Hausdorff distance. This entails that Φ is also a Lipschitz function, with constant $L_{\overline{m}}V$, where V is the volume of $\mathcal{S}$ for the induced Lebesgue measure. This is because the Hausdorff distance δ between two sets $\mathcal{D}_1$ and $\mathcal{D}_2$ translates to a $V\delta$–Euclidean distance between $\phi_{\mathcal{D}_1}$ and $\phi_{\mathcal{D}_2}$. Indeed, we have, by definition of the Hausdorff distance, $\mathcal{D}_1\subseteq\mathcal{D}_2+B_\delta$ and $\mathcal{D}_2\subseteq\mathcal{D}_1+B_\delta$, where B_δ is the Euclidian ball of $\mathbb{R}^d$ with center $(0,\dots,0)$ and radius δ. Properties 4 and 1 of Lemma 3 respectively yield the inequalities

$$\big|\phi_{\mathcal{D}_1}-\phi_{\mathcal{D}_2}\big|=\max\big\{\phi_{\mathcal{D}_1}-\phi_{\mathcal{D}_2},\ \phi_{\mathcal{D}_2}-\phi_{\mathcal{D}_1}\big\}\leqslant\phi_{B_\delta}\leqslant\delta\,,$$

with, by integration, $\big\|\phi_{\mathcal{D}_1}-\phi_{\mathcal{D}_2}\big\|\leqslant V\delta$.

The above-stated properties of Φ imply that for all $\psi\in\mathbb{L}^2(\mathcal{S})$ with $\psi\geqslant 0$, the game $(x,y)\mapsto\langle\psi,\,\Phi(x,y)\rangle$ has a value $v(\psi)$, and that this value is achieved: there exists some $x_\psi\in\Delta(\mathcal{I})$ such that

$$\max_{y\in\Delta(\mathcal{J})}\big\langle\psi,\,\Phi(x_\psi,y)\big\rangle=v(\psi)=\max_{y\in\Delta(\mathcal{J})}\min_{x\in\Delta(\mathcal{I})}\big\langle\psi,\,\Phi(x,y)\big\rangle\,.$$

Now, consider some half-space $\mathcal{C}_{\rm hs}$ containing $\mathcal{C}_{\rm orth}\big(\phi_{\mathcal{C}}\big)$. It is of the form

$$\mathcal{C}_{\rm hs}=\big\{f\in\mathbb{L}^2(\mathcal{S}):\ \langle\psi,f\rangle\leqslant\beta\big\}\,,$$

where necessarily, as can be shown by contradiction, $\psi \geqslant 0$. The dual condition is satisfied by assumption, that is, for all $y \in \Delta(\mathcal{J})$, there exists $x \in \Delta(\mathcal{I})$ such that $\Phi(x,y) \in \mathcal{C}_{\mathrm{orth}}(\phi_{\mathcal{C}})$, and therefore, $\Phi(x,y) \in \mathcal{C}_{\mathrm{hs}}$. Thus, $v(\psi) \leqslant \beta$, as can be seen with its expression as a max/min. The mixed action x_ψ thus satisfies that $\langle \psi,\, \Phi(x_\psi, y)\rangle \leqslant \beta$ for all $y \in \Delta(\mathcal{J})$, which is exactly saying that $\Phi(x_\psi, y) \in \mathcal{C}_{\mathrm{hs}}$ for all $y \in \Delta(\mathcal{J})$. We therefore proved the desired one-shot Φ–approachability of $\mathcal{C}_{\mathrm{hs}}$.

Conversely, we assume that the dual condition is not satisfied, i.e., that there exists some $y_0 \in \Delta(\mathcal{J})$ such that for no $x \in \Delta(\mathcal{I})$ one has $\Phi(x, y_0) \in \mathcal{C}_{\mathrm{orth}}(\phi_{\mathcal{C}})$. We consider the continuous thus compact image $\Phi\big(\Delta(\mathcal{I}), y_0\big)$ of $\Delta(\mathcal{I})$ by $\Phi(\,\cdot\,, y_0)$. Its Euclidean distance to the closed set $\mathcal{C}_{\mathrm{orth}}(\phi_{\mathcal{C}})$ is thus positive, we denote it by $\delta > 0$. Now, the distance of an element $f \in \mathbb{L}^2(\mathcal{S})$ to $\mathcal{C}_{\mathrm{orth}}(\phi_{\mathcal{C}})$ is given by

$$d_{\mathcal{C}_{\mathrm{orth}}(\phi_{\mathcal{C}})}(f) = \int_{\mathcal{S}} \big(f(s) - \phi_{\mathcal{C}}(s)\big)_+ \,\mathrm{d}s\,.$$

Since in addition, Φ is convex in its first argument (as shown in the first part of this proof), we have that $d_{\mathcal{C}_{\mathrm{orth}}(\phi_{\mathcal{C}})}(f) \geqslant \delta$ not only for all $f \in \Phi\big(\Delta(\mathcal{I}), y_0\big)$ but also for all f in the convex hull of $\Phi\big(\Delta(\mathcal{I}), y_0\big)$. The latter set is pointwise bounded (by M, as follows from property 1 of Lemma 3) and is formed by equicontinuous functions (they all are M–Lipschitz continuous, as follows from property 2 of the same lemma). The Arzela–Ascoli theorem thus ensures that the closure of this set is compact for the supremum norm $\|\cdot\|_\infty$ over $\mathcal{S}$. As by integration $\|\cdot\|_\infty \geqslant \|\cdot\|/V$, the closure of the convex hull of $\Phi\big(\Delta(\mathcal{I}), y_0\big)$ and the set $\mathcal{C}_{\mathrm{orth}}(\phi_{\mathcal{C}})$ are still δ/V–separated in $\|\cdot\|_\infty$–norm, thus are disjoint. Since the former set is a convex and compact set, and the latter is a closed convex set, the Hahn–Banach theorem entails that they are strictly separated by some hyperplane. In particular, one of the two thus-defined half-spaces is not one-shot Φ–approachable. □

The above result is a generalization of the polytopial case

In Section 8 we showed that when approaching a polytope, there are only finitely many directions (i.e., finitely many elements of the sphere $\mathcal{S}$) of interest, namely, the directions corresponding to the defining hyperplanes. The results we obtained therein can in fact be obtained as a corollary of Theorem 3 when the latter is stated (and proved) with a different measure instead of the Lebesgue measure, given by the sums of the Dirac masses on the directions of the defining hyperplanes.

There are two ways to extend the primal characterization of approachability under partial monitoring from polytopes to general convex sets. The one we worked out above relies on the observation that with general convex sets, every direction might be relevant, as a general convex set is defined as the intersection of infinitely many half-spaces, one per element of $\mathcal{S}$. Based on this, we introduced for general convex sets a infinite-dimensional lifting into the space of real-valued mappings on the whole set $\mathcal{S}$. We also resorted the uniform Lebesgue measure since all directions are equally important.

The other way of generalizing the results relies on the fact that a closed convex set $\mathcal{C} \subseteq \mathbb{R}^d$ is approachable if and only if all containing polytopes are approachable. By

playing in blocks and approximating a given general convex set by a sequence of containing polytopes, one could have shown that $\mathcal{C}$ is (r, H)–approachable if and only if all containing polytopes satisfy the characterization of Corollary 2. However, while this alternative way leads to a characterization, it is less intrinsic as there is no fixed lifted space to be considered. (The finite-dimensional lifted spaces depend on the approximating polytopes.) For the sake of elegance, we thus used the infinite-dimensional lifting described above.

Acknowledgements

Shie Mannor was partially supported by the Israel Science Foundation under grant no. 920/12. Vianney Perchet acknowledges support by "Agence Nationale de la Recherche," under grant JEUDY (ANR-10-BLAN 0112).

References

R.J. Aumann and M.B. Maschler. *Repeated Games with Incomplete Information*. MIT Press, 1995.

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

E. Kohlberg. Optimal strategies in repeated games with incomplete information. *International Journal of Game Theory*, 4:7–243, 1975.

G. Lugosi, S. Mannor, and G. Stoltz. Strategies for prediction under imperfect monitoring. *Mathematics of Operations Research*, 33:513–528, 2008.

S. Mannor, V. Perchet, and G. Stoltz. Robust approachability and regret minimization in games with partial monitoring. An extended abstract was published in the *Proceedings of COLT'11*, 2011.

J.-F. Mertens, S. Sorin, and S. Zamir. Repeated games. Technical Report no. 9420, 9421, 9422, Université de Louvain-la-Neuve, 1994.

V. Perchet. Approachability of convex sets in games with partial monitoring. *Journal of Optimization Theory and Applications*, 149:665–677, 2011a.

V. Perchet. Internal regret with partial monitoring: Calibration-based optimal algorithms. *Journal of Machine Learning Research*, 12:1893–1921, 2011b.

V. Perchet and M. Quincampoix. On an unified framework for approachability in games with or without signals. 2011.

S. Sorin. *A First Course on Zero-Sum Repeated Games*. Number 37 in Mathématiques & Applications. Springer, 2002.

Appendix

A brief survey of some well-known properties of support functions

For the sake of self-completeness only we summarize in the lemma below some simple and well-known properties of support functions.

Lemma 3. *We consider a set $\mathcal{C} \subseteq \mathbb{R}^d$.*

1. *If $\mathcal{C}$ is bounded in Euclidian norm by C, then $\phi_{\mathcal{C}}$ is bounded in supremum norm by C and in Euclidean norm by VC, where V is the volume of $\mathcal{S}$ under the (induced) Lebesgue measure.*

2. *If $\mathcal{C}$ is bounded in Euclidian norm by C, then $\phi_{\mathcal{C}}$ is a Lipschitz function, with Lipschitz constant C.*

3. *For all $\mathcal{C}' \subseteq \mathbb{R}^d$, if $\mathcal{C} \subseteq \mathcal{C}'$, then $\phi_{\mathcal{C}} \leqslant \phi_{\mathcal{C}'}$. The converse implication holds if in addition $\mathcal{C}'$ is a closed convex set.*

4. *The function ϕ is linear, in the sense that for all $\gamma \geqslant 0$ and all all $\mathcal{C}' \subseteq \mathbb{R}^d$, one has $\phi_{\gamma\mathcal{C}+\mathcal{C}'} = \gamma\phi_{\mathcal{C}} + \phi_{\mathcal{C}'}$.*

Proof. Property 1 follows from the Cauchy–Schwarz inequality: for all $s \in \mathcal{S}$,

$$\left|\phi_{\mathcal{C}}(s)\right| \leqslant \sup_{c\in\mathcal{C}} \left|\langle c, s\rangle\right| \leqslant \sup_{c\in\mathcal{C}} \|c\|\,\|s\| = \sup_{c\in\mathcal{C}} \|c\|\,,$$

as the elements $s \in \mathcal{S}$ have unit norm. The bound in Euclidean norm follows by integration over $\mathcal{S}$.

For property 2, we note that $s \in \mathcal{S} \mapsto \langle c, s\rangle$ is a $\|c\|$–Lipschitz function (again, by the Cauchy–Schwarz inequality). Therefore $\phi_{\mathcal{C}}$ is the supremum of C–Lipschitz functions and as such is also a C–Lipschitz function.

The first part of property 3 holds by the definition of a supremum. To prove the converse implication, we use an argument by contradiction. We consider two sets $\mathcal{C}$ and $\mathcal{C}'$, where $\mathcal{C}'$ is closed and convex. We assume that $\mathcal{C}$ is not included in $\mathcal{C}'$ and show that the existence of a $s \in \mathcal{S}$ such that $\phi_{\mathcal{C}}(s) > \phi_{\mathcal{C}'}(s)$. The set $\mathcal{C} \setminus \mathcal{C}'$ is not empty, let x be one of its elements. The convex sets $\{x\}$, which is compact, and $\mathcal{C}'$, which is closed, are disjoint sets. The Hahn–Banach theorem ensures the existence of a strictly separating hyperplane between these convex sets, which we can write in the form $\left\{\omega \in \mathbb{R}^d :\ \langle \omega, s\rangle = \beta\right\}$ for some $s \in \mathcal{S}$ and $\beta \in \mathbb{R}$ such that

$$\phi_{\{x\}}(s) = \langle x, s\rangle > \beta \qquad \text{and} \qquad \forall c' \in \mathcal{C}', \quad \langle c', s\rangle < \beta\,.$$

This entails that $\phi_{\mathcal{C}'}(s) \leqslant \beta < \phi_{\{x\}}(s) \leqslant \phi_{\mathcal{C}}(s)$.

Finally, the last property is true because by definition

$$\gamma\mathcal{C} + \mathcal{C}' = \left\{\gamma c + c' :\ c \in \mathcal{C},\ c' \in \mathcal{C}'\right\}$$

and thus, for all $s \in \mathcal{S}$,

$$\sup_{c'' \in \gamma\mathcal{C}+\mathcal{C}'} \langle c'', s\rangle = \sup_{c\in\mathcal{C},\, c'\in\mathcal{C}'} \gamma\langle c, s\rangle + \langle c', s\rangle = \gamma \sup_{c\in\mathcal{C}} \langle c, s\rangle + \sup_{c'\in\mathcal{C}'} \langle c', s\rangle \,,$$

where we used the fact that $\gamma > 0$ in the last equality. □