HAL
open science

Approachability of convex sets in generalized quitting games

János Flesch, Rida Laraki, Vianney Perchet

To cite this version:

János Flesch, Rida Laraki, Vianney Perchet. Approachability of convex sets in generalized quitting games. Conference On Learning Theory, Jun 2016, New York, United States. hal-01373174

HAL Id: hal-01373174

https://hal.science/hal-01373174v1

Submitted on 28 Sep 2016

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Approachability of convex sets in generalized quitting games*

János Flesch[†], Rida Laraki[‡] and Vianney Perchet[§]

September 28, 2016

Abstract

We consider Blackwell approachability, a very powerful and geometric tool in game theory, used for example to design strategies of the uninformed player in repeated games with incomplete information. We extend this theory to "generalized quitting games", a class of repeated stochastic games in which each player may have quitting actions, such as the Big-Match. We provide three simple geometric and strongly related conditions for the weak approachability of a convex target set. The first is sufficient: it guarantees that, for any fixed horizon, a player has a strategy ensuring that the expected time-average payoff vector converges to the target set as horizon goes to infinity. The third is necessary: if it is not satisfied, the opponent can weakly exclude the target set.

In the special case where only the approaching player can quit the game (Big-Match of type I), the three conditions are equivalent and coincide with Blackwell's condition. Consequently, we obtain a full characterization and prove that the game is weakly determined - every convex set is either weakly approachable or weakly excludable.

In games where only the opponent can quit (Big-Match of type II), none of our conditions is both sufficient and necessary for weak approachability. We provide a continuous time sufficient condition using techniques coming from differential games, and show its usefulness in practice, in the spirit of Vieille's seminal work for weak approachability.

*An extended version of the abstract has been published at the 29th edition of the annual Conference On Learning Theory, 2016.

[†]Department of Quantitative Economics, Maastricht University, The Netherlands. Email: j.flesch@maastrichtuniversity.nl

[‡]Université Paris-Dauphine, PSL Research University, CNRS, Lamsade, 75016 Paris, France. Also affiliated with Department of Economics, Ecole Polytechnique, France. Laraki's work was supported by grants administered by the French National Research Agency as part of the Investissements d'Avenir program (Idex [Grant Agreement No. ANR-11-IDEX-0003-02/Labex ECODEC No. ANR11- LABEX-0047] and ANR-14-CE24-0007-01 CoCoRICo-CoDecreceived

[§]Centre de Mathématiques et de Leurs Applications, ENS Cachan, France & Criteo Labs, Paris, France. Email vianney.perchet@normalesup.org. V. Perchet is partially funded by the ANR grant ANR-13-JS01-0004-01 and he benefited from the support of the « FMJH Program Gaspard Monge in optimization and operation research» and from EDF.

Finally, we study uniform approachability where the strategy should not depend on the horizon and demonstrate that, in contrast with classical Blackwell approachability for convex sets, weak approachability does not imply uniform approachability.

Keywords: Blackwell Approachability, Stochastic games, Absorbing Games, Big-Match, Calibration, Regret Learning, Determinacy.

1 Introduction

We study a class of 2-player stochastic games with vector payoffs, building upon the classical models proposed by Shapley [32] and Blackwell [5].

A finite zero-sum stochastic game is a repeated game with perfect monitoring where the action spaces are finite and the stage payoff $g_t \in \mathbb{R}$ depends on a state-parameter that can have finitely many different values and whose evolution is controlled by both players. Shapley [32] proved that such a game with the λ-discounted evaluation $\sum_t \lambda(1-\lambda)^{t-1} g_t$ has a value v_λ. This existence result extends easily to more general evaluations, such as $\sum_t \theta_t g_t$ where θ_t is the weight of stage t, see Laraki and Sorin [19]. For example, in the classical T-stage game, where $\theta_t = \frac{1}{T}$ for $1 \leq t \leq T$, the corresponding value is denoted v_T. Bewley and Kohlberg [4] proved that every stochastic game has an asymptotic value v, i.e. v_λ and v_T both converge to the same limit v, as $\lambda \to 0$ and $T \to \infty$. Mertens and Neyman [22] showed that the players can guarantee v uniformly in the sense that for every $\epsilon > 0$, each player has a strategy that guarantees v up to ε simultaneously in every λ-discounted game for sufficiently small λ and in every T-stage game for sufficiently large T. Such a result was earlier obtained by Blackwell and Ferguson [7] for the game Big-Match (introduced by Gillette [12]) and for the class of all absorbing games by Kohlberg [16]. In fact, absorbing games are one of the very few classes where the asymptotic value has an explicit formula in terms of the one-shot game, see Laraki [18]. We recall here that the asymptotic value may even fail to exist if we drop any of the following assumptions: finite action spaces, see Vigeral [37], finite state space, or perfect monitoring, see Ziliotto [38]. Ergodic stochastic games, where all states are visited infinitely often almost surely regardless the actions chosen by the players, are another example.

A Blackwell approachability problem is a 2-player repeated game with perfect monitoring and stage vector-payoffs $g_t \in \mathbb{R}^d$ in which player 1's objective is to enforce the convergence of $\frac{1}{T}\sum_{t=1}^{T} g_t$ to some target set $\mathcal{C} \subset \mathbf{R}^d$. On the other hand, player 2 aims at preventing and his ultimate objective is to exclude this target set, i.e., to approach the complement of some δ-neighborhood of it. Blackwell [5] proved that the game is uniformly determined for any convex target set: either player 1 can uniformly approach $\mathcal{C}$ or player 2 can uniformly exclude $\mathcal{C}$. More importantly, Blackwell provided a simple geometric characterization of approachable sets from which one can easily build an optimal strategy. However, if $\mathcal{C}$ is not convex, uniform determinacy fails. This led Blackwell to define a weaker version of determinacy by allowing the strategy to depend on the horizon T. Several years later, Vieille [36] solved the problem and proved that

a set is weakly approachable if and only if the value of an auxiliary differential game is zero. Weak determinacy follows from the existence of the value in differential games.

Combining the models of Shapley and Blackwell: It is natural to consider stochastic games with vector payoffs and try to characterize approachable target sets in these games, notably to develop new tools for stochastic games with incomplete information. This challenging problem has already been tackled but, so far, only few results have been achieved. The most relevant work, by Milman [23], only apply to ergodic stochastic games and no geometric characterization of approachable sets has been provided. On a different matter, it has been remarked that uniform determinacy failed to hold in stochastic games[1], even in variants of Big-Match.

Guided by the history of stochastic games, we tackle the general model of stochastic games with vector payoffs by focusing, in a first step, on the class of absorbing games, and in particular Big-Match games. Indeed, to obtain a simple geometric characterization of approachable sets, it is helpful to consider an underlying class of games that admit an explicit characterization for the asymptotic value.

We call "generalized quitting games" the subclass of absorbing games we focus on. This terminology refers to quitting games, in which each player has exactly one quitting action and one non-quitting action. In contrast, in our case, one or both players may have none or many quitting actions. The game is repeated until a quitting action is chosen at some stage t^*, in which case it enters an absorbing state that may depend on both actions at stage t^*. When only player 1 (resp. player 2) has quitting actions, the game is called Big-Match of type I (resp. type II).

Main contributions: We introduce three strongly related simple geometric conditions on a convex target set $\mathcal{C}$. They are nested (the first implies the second, which implies the third) and they all have flavors of both Blackwell's condition [5] and Laraki's formula in [18] for the asymptotic value in absorbing games. We prove that the first condition is sufficient for player 1 to weakly approach $\mathcal{C}$ and that it can be used to build an approachability strategy, even though the explicit construction is delicate and relies on a calibration technique developed notably in Perchet [24]. The second condition is a useful intermediate condition, but it is neither necessary nor sufficient. The third condition is proven to be necessary: indeed, if $\mathcal{C}$ does not satisfy it, then $\mathcal{C}$ is weakly excludable by player 2. Finally, we show that there are convex sets that are neither weakly approachable nor weakly excludable.

We examine Big-Match games in detail. In Big-Match games of type I, our three conditions are shown to be equivalent and to coincide with Blackwell's condition. This provides a full characterization for *weak* approachability and proves that this class is *weakly* determined. This contrasts with the *uniform* indeterminacy, see Example 15, where we provide a 1-dimensional counter-example.

In Big-Match games of type II, the first two conditions are equivalent and they are proven to be necessary and sufficient for *uniform* approachability. Despite this full characterization, uniform determinacy fails. For weak approachability we show that none

[1]This remark was already made by Sorin in the eighties in a small but unpublished note; its flavor is provided in Example 15.

of the three conditions is both necessary and sufficient. We also develop in some cases, an approach based on differential game, similarly to Vieille [36].

To summarize, our analysis of Big-Match games reveals that: (1) in Big-Match games of type I, a simple full characterization is available for *weak* approachability; (2) in Big-Match games of type II, a simple full characterization is available for *uniform* approachability; (3) uniform determinacy fails in both types of Big-Match games; (4) weak determinacy holds for Big-Match games of type I. Weak determinacy for Big-Match games of type II remains an open problem.

Almost sure approachability: In the classical Blackwell model on convex sets, and in ergodic stochastic games with vector payoffs, weak, uniform, in expectation or almost sure approachability problems are equivalent. In our case, they all differ. In this paper we focus on weak and uniform approachability in expectation, as they appear to be very interesting and challenging in generalized quitting games. We refer to Section 2 and Appendix C for more details.

Related literature: Blackwell approachability is frequently used in the literature of repeated games. It was used first by Aumann and Maschler [2] to construct optimal strategies in zero-sum repeated games with incomplete information and perfect monitoring. Their construction has been extended by Kohlberg [17] to the imperfect monitoring case. Blackwell approachability was further used by Renault and Tomala [30] to characterize the set of communication equilibria in N-player repeated games with imperfect monitoring; by Hörner and Lovo [14] and Hörner, Lovo and Tomala [15] to characterize belief-free equilibria in N-player repeated games with incomplete information, and by Tomala [35] to characterize belief-free communication equilibria in N-player repeated games. Blackwell approachability has also been used to construct adaptive strategies leading to correlated equilibria (see Hart and Mas-Colell [13]), machine learning strategies minimizing regret (see Blackwell [6], Abernethy, Bartlett and Hazan [1]), and calibrating algorithms in prediction problems (see Dawid [8], Foster and Vohra [11], Perchet [26, 28]). In fact, one can show that Blackwell approachability, regret-minimization and calibration are formally equivalent (see for instance Abernethy, Bartlett and Hazan [1] or Perchet [26]).

Applications: Classical machine learning assumes that a one stage mistake has small consequences. Our paper allows to tackle realistic situations where the total payoff can be affected by one stage decisions. One could think of clinical trials between two treatments: at some point in time one of the two must be selected and prescribed to the rest of the patients. At a more theoretical level, as in Aumann and Maschler [2] for zero-sum repeated games with incomplete information, our paper may be a useful step towards a characterization of the asymptotic value of absorbing games with incomplete information and determining the optimal strategy of the non–informed player. A problem for which we know existence of the asymptotic value (see Rosenberg [31]), and have some explicit characterizations of the asymptotic value in 2×2 Big-Match games (see Sorin [33, 34]).

2 Model and Main Results

In this section, we describe the model of generalized quitting games, the problem of Blackwell approachability, and present our main results.

Generalized quitting games. We denote by $\mathbf{I} = \mathcal{I} \cup \mathcal{I}^\star$ the finite set of (pure) actions of player 1 and by $\mathbf{J} = \mathcal{J} \cup \mathcal{J}^\star$ the finite set of actions of player 2. The actions in $\mathcal{I}$ and $\mathcal{J}$ are called non-quitting, and the actions in $\mathcal{I}^\star$ and $\mathcal{J}^\star$ are called quitting. A payoff vector $g(i,j) \in \mathbb{R}^d$ is associated to each pair of actions $(i,j) \in \mathbf{I} \times \mathbf{J}$, and to ease notations we assume that $\|g(i,j)\|_2 \leq 1$.

The game is played at stages in $\mathbb{N}^\star$ as follows: at stage $t \in \mathbb{N}^\star$, the players choose actions simultaneously, say $i_t \in \mathbf{I}$ and $j_t \in \mathbf{J}$. If only non-quitting actions have been played before stage t, i.e. $i_{t'} \in \mathcal{I}$ and $j_{t'} \in \mathcal{J}$ for every $t' < t$, then player 1 is free to choose any action in $\mathbf{I}$ and player 2 is free to choose any action in $\mathbf{J}$. However, if a quitting action was played by either player at a stage prior to stage t, i.e. $i_{t'} \in \mathcal{I}^\star$ or $j_{t'} \in \mathcal{J}^\star$ for some $t' < t$, then the players are obliged to take $i_t = i_{t-1}$ and $j_t = j_{t-1}$. Another equivalent way to model this setup is to assume that, as soon as a quitting action is played, the game absorbs in a state where the payoff is constant.

When a player plays a quitting action and neither player has played a quitting action before, we say that this player quits and that play absorbs.

Mixed actions. A mixed action for a player is a probability distribution over his (pure) actions. We will denote mixed actions of player 1 by $\mathbf{x} \in \Delta(\mathcal{I} \cup \mathcal{I}^\star)$, $x \in \Delta(\mathcal{I})$, $x^\star \in \Delta(\mathcal{I}^\star)$. Thus, a bold letter stands for a mixed action over the full set of actions $\mathbf{I}$, a regular letter for a mixed action restricted to non-quitting actions in $\mathcal{I}$ and a letter with an asterisk for a mixed action over the set of quitting actions in $\mathcal{I}^\star$. Similarly, we denote mixed actions of player 2 by $\mathbf{y} \in \Delta(\mathcal{J} \cup \mathcal{J}^\star)$, $y \in \Delta(\mathcal{I})$ and $y^\star \in \Delta(\mathcal{I}^\star)$.

To introduce our conditions for a convex set to be approachable, it will be helpful to consider finite nonnegative measures on $\mathbf{I}$ and $\mathbf{J}$ instead of probability distributions. We shall denote them by $\alpha \in \mathcal{M}(\mathbf{I})$ for player 1 and by $\beta \in \mathcal{M}(\mathbf{J})$ for player 2.

The payoff mapping g is extended as usual multi-linearly to the set of mixed actions $\Delta(\mathbf{I})$ and $\Delta(\mathbf{J})$ and, more generally, to the set of measures $\mathcal{M}(\mathbf{I})$ and $\mathcal{M}(\mathbf{J})$.

We also introduce the "measure" or "probability of absorption" and the "expected absorption payoff" (which is not the expected payoff conditional to absorption), defined respectively by

$$p^\star(\alpha,\beta) = \sum_{(i,j)\,\in\,(\mathcal{I}^\star\times\mathbf{J})\cup(\mathbf{I}\times\mathcal{J}^\star)} \alpha_i\beta_j \quad \text{and} \quad g^\star(\alpha,\beta) = \sum_{(i,j)\,\in\,(\mathcal{I}^\star\times\mathbf{J})\cup(\mathbf{I}\times\mathcal{J}^\star)} \alpha_i\beta_j g(i,j).$$

Strategies. In our model of generalized quitting games, histories are defined as long as no quitting action is played. Thus, the set of histories H is the set of finite sequences in $\mathcal{I} \times \mathcal{J}$ (that is, $H = \cup_{t=1}^{\infty}(\mathcal{I} \times \mathcal{J})^{t-1}$). A strategy for player 1 is a mapping $\sigma : H \to \Delta(\mathbf{I})$, and a strategy for player 2 is a mapping $\tau : H \to \Delta(\mathbf{J})$.

Specific subclasses of games: Big-Match games. We shall consider two subclasses of generalized quitting games, in which only one of the players can quit.

Following the nomenclature of Gillette [12], a generalized quitting game is called a Big-Match game of type I if player 1 has at least one non-quitting action – to avoid degenerate cases – and at least one quitting action, but player 2 has no quitting action, i.e. $\mathcal{I} \neq \emptyset$, $\mathcal{I}^* \neq \emptyset$ and $\mathcal{J}^\star = \emptyset$.

A generalized quitting game is called a Big-Match game of type II if player 2 has at least one non-quitting action and at least one quitting action, but player 1 has no quitting action, i.e. $\mathcal{J} \neq \emptyset$, $\mathcal{J}^* \neq \emptyset$ and $\mathcal{I}^\star = \emptyset$.

Objectives. In short, the objective of player 1 is to construct a strategy σ such that, for any strategy τ of player 2, the expected average payoff $\mathbb{E}_{\sigma,\tau}\frac{1}{T}\sum_{t=1}^{T} g(i_t, j_t)$ is close to some exogenously given convex set $\mathcal{C} \subset \mathbb{R}^d$, called the "target set". Instead of the Cesaro average, we can also consider the expected discounted payoff $\mathbb{E}_{\sigma,\tau}\sum_{t=1}^{\infty}\lambda(1-\lambda)^{t-1}g(i_t, j_t)$ or even a general payoff evaluation $\mathbb{E}_{\sigma,\tau}\sum_{t=1}^{\infty}\theta_t g(i_t, j_t)$, where $\theta_t \in [0, 1]$ and $\sum_{t=1}^{\infty}\theta_t = 1$, with the interpretation that θ_t is the weight of stage t.

We emphasize here that we focus on the distance of the expected average payoff to $\mathcal{C}$ (and not on the expected distance of the average payoff to $\mathcal{C}$, corresponding to almost sure convergence, see e.g. Milman [23]), as it is might be more traditional and even challenging in stochastic games. Indeed, consider the toy game where player 1 has only two actions, both absorbing, and they give payoffs -1 and 1 respectively. In this game, $\{0\}$ is obviously not approachable in the almost sure sense, but is easily approachable in the expected sense by playing each action with probability $\frac{1}{2}$ at the first stage. We still quickly investigate almost sure approachability in Appendix C.

We can distinguish at least two different concepts of approachability, that we respectively call uniform approachability and weak approachability.

Specifically, we say that a convex set $\mathcal{C} \subset \mathbb{R}^d$ is uniformly approachable by player 1, if for every $\varepsilon > 0$ player 1 has a strategy such that after some stage $T_\varepsilon \in \mathbb{N}$, the expected average payoff is ε-close to $\mathcal{C}$, against any strategy of player 2. Stated with quantifiers

$$\mathcal{C} \text{ is uniformly app.} \iff \forall\varepsilon > 0, \exists\sigma, \exists T_\varepsilon \in \mathbb{N}, \forall T \geq T_\varepsilon, \forall\tau, d_{\mathcal{C}}\Big(\mathbb{E}_{\sigma,\tau}\frac{1}{T}\sum_{t=1}^{T} g(i_t, j_t)\Big) \leq \varepsilon.$$

Reciprocally, a convex set $\mathcal{C} \subset \mathbb{R}^d$ is uniformly excludable by player 2 if she can uniformly approach the complement of some δ neighborhood of $\mathcal{C}$, for some fixed $\delta > 0$.

A similar definition holds for general evaluations induced by a sequence of weights $\theta = (\theta_t)_{t\in\mathbb{N}}$ such that $\sum_{t=1}^{\infty}\theta_t = 1$ and $\theta_t \geq 0$ for all $t \in \mathbb{N}$. For every $\varepsilon > 0$, there must exist a threshold θ_ε so that if the sequence $\theta = (\theta_t)_{t\in\mathbb{N}}$ satisfies $\|\theta\|_2 = \sqrt{\sum_t \theta_t^2} \leq \theta_\varepsilon$ then the θ-evaluation of the payoffs is within distance ε of $\mathcal{C}$. We emphasize here that the Cesaro average corresponds to $\theta_t = 1/N$ for $t \in [N] = \{1, \ldots, N\}$ while the discounted evaluation, with discount factor $\lambda \in (0, 1]$ corresponds to $\theta_t = \lambda(1-\lambda)^{t-1}$ for all $t \geq 1$.

We then denote the accumulated θ-weighted average payoff up to stage $t \in \mathbb{N} \cup \{\infty\}$ as

$$\overline{g}_t^\theta = \sum_{s=1}^{t} \theta_s g(i_s, j_s), \quad \overline{g}_t^N = \frac{1}{N} \sum_{s=1}^{t \wedge N} g(i_s, j_s), \quad \text{and} \quad \overline{g}_t^\lambda = \sum_{s=1}^{t} \lambda(1-\lambda)^{s-1}(i_s, j_s).$$

We now focus on our main objective, weak approachability. We say that a convex set $\mathcal{C} \subset \mathbb{R}^d$ is weakly approachable by player 1, if for every $\varepsilon > 0$, if the horizon of the game is sufficiently large and known, player 1 has a strategy such that the expected payoff is ε-close to $\mathcal{C}$, against any strategy of player 2. Stated with quantifiers

$$\mathcal{C} \text{ is weakly app.} \iff \forall \varepsilon > 0, \exists T_\varepsilon \in \mathbb{N}, \forall T \geq T_\varepsilon, \exists \sigma_T, \forall \tau, d_{\mathcal{C}}\Big(\mathbb{E}_{\sigma_T,\tau} \frac{1}{T} \sum_{t=1}^{T} g(i_t, j_t)\Big) \leq \varepsilon$$

Reciprocally, a convex set $\mathcal{C} \subset \mathbb{R}^d$ is weakly excludable by player 2 if she can weakly approach the complement of some δ neighborhood of $\mathcal{C}$. This definition of weak approachability may be extended, just as above, for general evaluation, where the strategy of player 1 depends on $\theta = (\theta_t)_{t \in \mathbb{N}}$.

Observe that we can assume without loss of generality that the target set $\mathcal{C}$ is closed, because approaching a set or its closure are two equivalent problems.

We emphasize that, without an irreversible Markov chain structure, uniform approachability will be equivalent to weak approachability, because of the doubling trick, when the target set is convex. However, as we shall see, it is no longer the case in generalized quitting games.

Reminder on approachability in classical repeated games. Blackwell [5] proved that in classical repeated games (i.e., when $\mathcal{I}^\star = \mathcal{J}^\star = \emptyset$) there is a simple geometric necessary and sufficient condition under which a convex set is (uniformly and weakly) approachable. It reads as follows

$$\begin{aligned} \mathcal{C} \text{ is uniformly/weakly app.} &\iff \forall y \in \Delta(\mathcal{J}), \exists x \in \Delta(\mathcal{I}), g(x,y) \in \mathcal{C} \\ &\iff \max_{y \in \Delta(\mathcal{J})} \min_{x \in \Delta(\mathcal{I})} d_{\mathcal{C}}(g(x,y)) = 0. \end{aligned}$$

This immediately entails that a convex set is either weakly approachable or weakly excludable.

Approachability conditions. We aim at providing a similar geometric condition ensuring that a convex set $\mathcal{C}$ is weakly approachable (or weakly excludable). Inspired by a recent formula, obtained in Laraki [18], which characterizes the asymptotic value by making use of perturbations of mixed actions with measures, we introduce the following three conditions.

The strongest of the three conditions is:

$$\max_{\mathbf{y} \in \Delta(\mathbf{J})} \min_{\mathbf{x} \in \Delta(\mathbf{I})} \inf_{\alpha \in \mathcal{M}(\mathbf{I})} \sup_{\beta \in \mathcal{M}(\mathbf{J})} d_{\mathcal{C}}\Big(\frac{g(\mathbf{x},\mathbf{y}) + g^\star(\alpha,\mathbf{y}) + g^\star(\mathbf{x},\beta)}{1 + p^\star(\alpha,\mathbf{y}) + p^\star(\mathbf{x},\beta)}\Big) = 0. \tag{1}$$

The next condition will be shown to be a useful intermediate condition:

$$\max_{\mathbf{y}\in\Delta(\mathbf{J})}\min_{\mathbf{x}\in\Delta(\mathbf{I})}\sup_{\beta\in\mathcal{M}(\mathbf{J})}\inf_{\alpha\in\mathcal{M}(\mathbf{I})} d_{\mathcal{C}}\Big(\frac{g(\mathbf{x},\mathbf{y})+g^\star(\alpha,\mathbf{y})+g^\star(\mathbf{x},\beta)}{1+p^\star(\alpha,\mathbf{y})+p^\star(\mathbf{x},\beta)}\Big)=0. \tag{2}$$

Finally, the weakest of the three conditions is:

$$\max_{\mathbf{y}\in\Delta(\mathbf{J})}\sup_{\beta\in\mathcal{M}(\mathbf{J})}\min_{\mathbf{x}\in\Delta(\mathbf{I})}\inf_{\alpha\in\mathcal{M}(\mathbf{I})} d_{\mathcal{C}}\Big(\frac{g(\mathbf{x},\mathbf{y})+g^\star(\alpha,\mathbf{y})+g^\star(\mathbf{x},\beta)}{1+p^\star(\alpha,\mathbf{y})+p^\star(\mathbf{x},\beta)}\Big)=0. \tag{3}$$

We emphasize here that, in the above conditions, the maxima and minima are indeed attained since the mapping $(\mathbf{x},\alpha,\mathbf{y},\beta)\mapsto\frac{g(\mathbf{x},\mathbf{y})+g^\star(\alpha,\mathbf{y})+g^\star(\mathbf{x},\beta)}{1+p^\star(\alpha,\mathbf{y})+p^\star(\mathbf{x},\beta)}$ is Lipchitz.

Main results. We can already state our main results, which we will prove throughout the paper. In these results, approachability always refers to player 1 whereas excludability always refers to player 2.

Theorem 1 (Weak Approachability) *Let $\mathcal{C}\subset\mathbb{R}^d$ be a convex set.*

Sufficiency: *If Condition* (1) *is satisfied, then $\mathcal{C}$ is weakly approachable.*

Necessity: *If $\mathcal{C}$ is weakly approachable, then Condition* (3) *is satisfied. Indeed, if Condition* (3) *is not satisfied, then $\mathcal{C}$ is weakly excludable.*

The theorem above is proven in Section 3.

In the special class of Big-Match games, our findings for weak approachability are summarized by the next proposition.

Proposition 2 (Weak Approachability in Big-Match Games)

In Big-Match games of type I: *Conditions* (1), (2) *and* (3) *coincide with the classical Blackwell condition, i.e.,*

$$\mathcal{C} \text{ is weakly approachable} \iff \forall\mathbf{y}\in\Delta(\mathbf{J}),\exists\mathbf{x}\in\Delta(\mathbf{I}), g(\mathbf{x},\mathbf{y})\in\mathcal{C}.$$

Consequently, in this class, weak determinacy holds: a convex set is either weakly approachable or weakly excludable.

In Big-Match games of type II: *Conditions* (1) *and* (2) *coincide and generally differ with Condition* (3). *Moreover, none of the conditions is both sufficient and necessary for weak approachability.*

In the proposition above, the first part of the claim on Big-Match games of type I follows from Lemma 8. The second part is then a direct consequence of Theorem 1. Indeed, suppose that a convex set $\mathcal{C}$ is not weakly approachable. Then by Theorem 1, $\mathcal{C}$ does not satisfy Condition (1). Since Conditions (1) and (3) coincide, $\mathcal{C}$ does not satisfy Condition (3) either. Hence, by Theorem 1 once again, $\mathcal{C}$ is weakly excludable.

The claim on Big-Match games of type II follows from Lemma 6 and by Examples 16 and 17.

For uniform approachability we obtain the following results.

Proposition 3 (Uniform Approachability in Big-Match Games)

In Big-Match games of either type: *There are convex sets which are weakly approachable but not uniformly approachable; and so they are neither uniformly approachable nor uniformly excludable.*

In Big-Match games of type II: *Condition* (1) *(and hence Condition* (2)*) is necessary and sufficient for uniform approachability.*

The first claim is shown by Examples 15 and 16. The second claim on Big-Match games of type II follows from Proposition 12.

Notice that, in the results above, approachability conditions for Big-Match games of types I and II are drastically different. The necessary and sufficient weak approachability condition takes a simple form for type I, but not for type II, the situation being completely reversed for uniform approachability. The second consequence is that determinacy of convex sets is very specific to the original model of Blackwell [5]. We remark that determinacy also fails in the standard model of Blackwell, if player 1 has an imperfect observation on past actions of player 2, as proved by Perchet [25] by providing an example of a convex set that is neither approachable nor excludable.

Outline of the Paper. The remaining of the paper is organized as follows. In Section 3, we prove Theorem 1. In Section 4 we compare the notions of weak and uniform approachability, with a focus on Big-Match games. In Section 5, we present several examples. Additional results and examples can be found in the Appendices.

3 Necessary and Sufficient Conditions for Weak Approachabilty

In this section, we prove Theorem 1. First we shall prove that, assuming the sufficiency of Condition (1) for weak approachability, Condition (3) is indeed necessary. Then, we show that Condition (1) ensures weak approachability.

3.1 If Condition (1) is Sufficient, then Condition (3) is Necessary

As claimed, we will prove later in Proposition 11 that Condition (1) is sufficient for the weak approachability of convex sets in generalized quitting games. The purpose of this section is to demonstrate that this entails the necessity of Condition (3) by switching the role of players 1 and 2.

Proposition 4 *Assume that Condition* (1) *is sufficient for the weak approachability of convex sets. Then Condition* (3) *is necessary: If a convex set* $\mathcal{C}$ *does not satisfy Condition* (3)*, then* $\mathcal{C}$ *is weakly excludable by player 2.*

Proof. As Condition (3) is not satisfied for $\mathcal{C}$, there exists $\delta > 0$ such that

$$\max_{\mathbf{y}\in\Delta(\mathbf{J})} \sup_{\beta\in\mathcal{M}(\mathbf{J})} \min_{\mathbf{x}\in\Delta(\mathbf{I})} \inf_{\alpha\in\mathcal{M}(\mathbf{I})} d_{\mathcal{C}}\Big(\frac{g(\mathbf{x},\mathbf{y}) + g^\star(\alpha,\mathbf{y}) + g^\star(\mathbf{x},\beta)}{1 + p^\star(\alpha,\mathbf{y}) + p^\star(\mathbf{x},\beta)}\Big) \geq \delta.$$

Choose some $\mathbf{y}_0$ and β_0 that realize the supremum up to $\delta/2$. It is not difficult to see that

$$\Big\{\frac{g(\mathbf{x},\mathbf{y}_0) + g^\star(\alpha,\mathbf{y}_0) + g^\star(\mathbf{x},\beta_0)}{1 + p^\star(\alpha,\mathbf{y}_0) + p^\star(\mathbf{x},\beta_0)}\ ;\ \mathbf{x}\in\Delta(\mathbf{I}), \alpha\in\mathcal{M}(\mathbf{I})\Big\}$$

is a bounded convex set that is $\delta/2$ away from $\mathcal{C}$ and whose closure is denoted by $\mathcal{E}$.

To prove the convexity of the above set, suppose that

$$z = \sum_i \lambda_i \frac{g(\mathbf{x}_i,\mathbf{y}_0) + g^\star(\alpha_i,\mathbf{y}_0) + g^\star(\mathbf{x}_i,\beta_0)}{1 + p^\star(\alpha_i,\mathbf{y}_0) + p^\star(\mathbf{x}_i,\beta_0)}, \quad \text{with } \lambda_i \geq 0 \ \text{ and } \sum_i \lambda_i = 1\ .$$

Taking $\theta_i = \big(1 + p^\star(\alpha_i,\mathbf{y}_0) + p^\star(\mathbf{x}_i,\beta_0)\big)^{-1}$, we obtain that

$$z = \frac{g(\mathbf{x},\mathbf{y}_0) + g^\star(\alpha,\mathbf{y}_0) + g^\star(\mathbf{x},\beta_0)}{1 + p^\star(\alpha,\mathbf{y}_0) + p^\star(\mathbf{x},\beta_0)} \ \text{ with } \mathbf{x} = \frac{\sum_i \lambda_i\theta_i\mathbf{x}_i}{\sum_i \lambda_i\theta_i} \ \text{ and } \alpha = \frac{\sum_i \lambda_i\theta_i\alpha_i}{\sum_i \lambda_i\theta_i}\ .$$

Thus we have proved that

$$\max_{\mathbf{x}\in\Delta(\mathbf{I})} \min_{\mathbf{y}\in\Delta(\mathbf{J})} \inf_{\beta\in\mathcal{M}(\mathbf{J})} \max_{\alpha\in\mathcal{M}(\mathbf{I})} d_{\mathcal{E}}\Big(\frac{g(\mathbf{x},\mathbf{y}) + g^\star(\alpha,\mathbf{y}) + g^\star(\mathbf{x},\beta)}{1 + p^\star(\alpha,\mathbf{y}) + p^\star(\mathbf{x},\beta)}\Big) = 0,$$

and that $\mathcal{E}$ satisfies Condition (1), but stated from the point of view of player 2. Therefore, by assumption, player 2 can weakly approach $\mathcal{E}$, and hence she can weakly approach the complement of the $\delta/2$-neighborhood of $\mathcal{C}$. This means that $\mathcal{C}$ is weakly excludable by player 2 (and in particular not weakly approachable by player 1), as desired. ■

3.2 Condition (1) is Sufficient

We prove in this section that Condition (1) is sufficient for the weak approachability of a convex set $\mathcal{C} \subset \mathbb{R}^d$. Assuming that the target set satisfies Condition (1), the construction of the approachability strategy will be based on a calibrated algorithm, as introduced by Dawid [8]. Similar ideas can be found in the online learning literature (see, e.g., Foster and Vohra [11], Perchet [24], and Bernstein, Mannor and Shimkin [3]), where Blackwell approachability and calibration now play an increasingly important role (as evidenced by Abernethy, Bartlett and Hazan [1], Mannor and Perchet [20], Perchet [27], Rakhlin, Sridharan and Tewari [29], and Foster, Rakhlin, Sridharan and Tewari [10]).

For the sake of clarity, we divide this section in several parts. First, we introduce the auxiliary calibration tool, then we prove the sufficiency condition in Big-Match games of types II and I, and finally we show how the main idea generalizes.

3.2.1 An auxiliary tool: Calibration

In this subsection, we adapt a result of Mannor, Perchet and Stoltz [21] on calibration to the setup with the general payoff evaluation (thus not necessarily with Cesaro averages).

We recall that calibration is the following sequential decision problem. Consider a non-empty and finite set Ω, a finite ε-grid of the set of probability distributions on Ω denoted by $\{p_k \in \Delta(\Omega), k \in [K]\}$ where $[K] = \{1, \dots, K\}$ for $K \in \mathbb{N}$, and a sequence of weights $\{\theta_t \in \mathbb{R}_+\}_{t\in\mathbb{N}}$. At each stage t, Nature chooses a state $\omega_t \in \Omega$ and the decision maker predicts it by choosing simultaneously a point of the grid $p_t \in \Delta(\Omega)$. Once p_t is chosen, the state ω_t and the weight θ_t are revealed (we emphasize that the sequence θ is not necessarily known in advance by the decision maker).

We denote by $\mathbb{N}_t[k] = \{s \leq t \text{ s.t. } p_s = p_k\}$ the set of stages $s \leq t$ where p_k was predicted and by

$$\bar{\omega}_t[k] = \frac{\sum_{s\in\mathbb{N}_t[k]} \theta_s \delta_{\omega_s}}{\sum_{s\in\mathbb{N}_t[k]} \theta_s} \in \Delta(\Omega)$$

the empirical weighted distribution of the state on $\mathbb{N}_t[k]$.

In that setting, we say that an algorithm of the decision maker is calibrated if

$$\limsup_{t\to\infty} \max_{k\in[K]} \frac{\sum_{s\in\mathbb{N}_t[k]} \theta_s}{\sum_{s\leq t} \theta_s} \Big(\|p_k - \bar{\omega}_t[k]\| - \varepsilon \Big)_+ = 0$$

almost surely.

Lemma 5 *The decision maker has a calibrated algorithm such that, for all $t \in \mathbb{N}$,*

$$\mathbb{E}\left[\max_{k\in[K]} \Big(\sum_{s\in\mathbb{N}_t[k]} \theta_s \Big) \Big(\|p_k - \bar{\omega}_t[k]\| - \varepsilon \Big)_+ \right] \leq \sqrt{8|\Omega| \sum_{s\leq t} \theta_s^2}.$$

Proof. The proof is almost identical to the one in Mannor, Perchet and Stoltz [21], Appendix A, thus is omitted. ■

We mention here that the construction of a calibrated algorithm is actually often based on the construction of an approachability strategy, as in Foster [9], and Perchet [26, 28].

3.2.2 Condition (1) is sufficient in Big-Match games of type II

We first focus on Big-match games of type II, where only player 2 can quit. The following lemma exhibits a useful equivalence between some of the conditions.

Lemma 6 *In Big-Match games of type II, Condition* (1) *and Condition* (2) *are equivalent, and they are further equivalent to*

$$\forall y \in \Delta(\mathcal{J}), \exists x \in \Delta(\mathcal{I}), g(x, y) \in \mathcal{C} \text{ and } g(x, j^\star) \in \mathcal{C}, \forall j^\star \in \mathcal{J}^\star. \tag{4}$$

A consequence of (4) is that if player 2, at every stage, either plays a non-quitting action i.i.d. accordingly to $y \in \Delta(\mathcal{J})$ or decides to quit, then player 1 can approach $\mathcal{C}$ by playing i.i.d. accordingly to $x \in \Delta(\mathcal{I})$. The sufficiency of this condition means that it is not more complicated to approach $\mathcal{C} \subset \mathbb{R}^d$ against an opponent than against an i.i.d. process that could eventually quit at some (unknown) time.

Proof: We already know that Condition (1) implies Condition (2).

Now we prove that Condition (2) implies (4). So, assume that $\mathcal{C}$ satisfies Condition (2). Since $p^\star(\alpha, y) = 0$ for all $\alpha \in \mathcal{M}(\mathbf{I})$ and $y \in \Delta(\mathcal{J})$, Condition (2) implies

$$\forall y \in \Delta(\mathcal{J}), \exists x \in \Delta(\mathcal{I}), \forall \beta \in \mathcal{M}(\mathbf{J}), \frac{g(x,y) + g^\star(x,\beta)}{1 + p^\star(x,\beta)} \in \mathcal{C}.$$

Now (4) follows by taking $\beta = 0$ and respectively taking $\beta_c = c \cdot \delta_{j^\star}$ with c tending to infinity.

Finally, we prove that (4) implies Condition (1). So, assume that $\mathcal{C}$ satisfies (4). Let $\mathbf{y} \in \Delta(\mathbf{J})$. Decompose it as $\mathbf{y} = \gamma y + (1-\gamma) y^\star$, where $y \in \Delta(\mathcal{J})$, $y^\star \in \Delta(\mathcal{J}^\star)$ and $\gamma \in [0,1]$. For this y, let $x \in \Delta(\mathcal{I})$ be given by (4). Then

$$\frac{g(x,\mathbf{y}) + g^\star(x,\beta)}{1 + p^\star(x,\beta)} = \frac{\gamma g(x,y) + (1-\gamma)\sum_{j^\star \in \mathcal{J}^\star} y^\star_{j^\star} g(x,j^\star) + \sum_{j^\star \in \mathcal{J}^\star} \beta_{j^\star} g(x,j^\star)}{\gamma + (1-\gamma)\sum_{j^\star \in \mathcal{J}^\star} y^\star_{j^\star} + \sum_{j^\star \in \mathcal{J}^\star} \beta_{j^\star}} \in \mathcal{C},$$

because all involved payoffs, $g(x,y)$ and $g(x,j^\star)$, belong to the convex set $\mathcal{C}$. Since we can choose $\alpha = 0$, we have shown that Condition (1) holds, as desired. ■

Proposition 7 *In Big-Match games of type II, a convex set $\mathcal{C}$ is (weakly or uniformly) approachable by player 1 if Condition* (1) *is satisfied.*

Proof. As advertised, the approachability strategy we will consider is based on calibration (as it can be generalized to more complex settings). The main insight is that player 1 predicts, stage by stage, $y \in \Delta(\mathcal{J})$ using a calibrated procedure and plays the response given by Lemma 6. Let θ_t be the sequence of weights used for the general payoff evaluation (recall that Cesaro average corresponds to $\theta_t = 1/N$ while discounted evaluation is $\theta_t = \lambda(1-\lambda)^{t-1}$).

Let $\Big\{y_k, k \in \{1,\dots,K\}\Big\}$ be a finite ε-discretization of $\Delta(\mathcal{J})$ and x_k be given by Lemma 6 for every $k \in \{1,\dots,K\}$. Consider the calibration algorithm introduced in Lemma 5 with respect to the sequence of weights θ_t. The strategy in the Big-Match game of type II is defined as follows: whenever $y_k \in \Delta(\mathcal{J})$ is predicted by the calibration algorithm, player 1 actually plays accordingly to x_k.

Assume that player 2 has never chosen an action in $\mathcal{J}^\star$ before stage t. Then Lemma 5 ensures that

$$\mathbb{E}\left[\max_{k\in[K]} \Big(\sum_{s \in \mathbb{N}_t[k]} \theta_s\Big)\Big(\|y_k - \bar{\jmath}_t[k]\| - \varepsilon\Big)_+\right] \le \sqrt{8|\mathcal{J}| \sum_{s \le t} \theta_s^2}$$

where $\bar{\jmath}_t[k] \in \Delta(\mathcal{J})$ is the weighted empirical distribution of actions of player 2 on the set of stages where y_k was predicted. We recall that on each of these stages, player 1

played accordingly to $x_k \in \Delta(\mathcal{J})$ so that the average weighted expected payoffs on those stages is $g(x_k, \bar{\jmath}_t[k])$.

Summing over $k \in [K]$, we obtain that

$$d_{\mathcal{C}}\left(\mathbb{E}\left[\frac{\sum_{k\in[K]}\left(\sum_{s\in\mathbb{N}_t[k]}\theta_s\right)g(x_k,\bar{\jmath}_t[k])}{\sum_{s\leq t}\theta_s}\right]\right) \leq \varepsilon + K\frac{\sqrt{8|\mathcal{J}|\sum_{s\leq t}\theta_s^2}}{\sum_{s\leq t}\theta_s}.$$

We stress that the payoff on the left-hand side of the above equation is exactly the expected weighted average vectorial payoff obtained by player 1 up to stage t. As a consequence, if player 2 never uses quitting action in $\mathcal{J}^\star$, letting t to infinity in the above equation yields

$$d_{\mathcal{C}}\left(\mathbb{E}\left[\bar{g}^\theta_\infty\right]\right) \leq \varepsilon + K\sqrt{8|\mathcal{J}|}\|\theta_2\|.$$

It remains to consider the case where player 2 used some quitting action $j^\star \in \mathcal{J}^\star$ at stage $\tau^\star + 1 \in \mathbb{N}$. At that stage, player 1 played accordingly to x_k for some $k \in [K]$, which ensures that $g(x_k, j^\star) \in \mathcal{C}$. As a consequence, absorption took place and the expected absorption payoff belongs to $\mathcal{C}$. We therefore obtain that

$$\begin{aligned} d_{\mathcal{C}}\left(\mathbb{E}\left[\bar{g}^\theta_\infty\right]\right) &\leq d_{\mathcal{C}}\left(\frac{\sum_{s\leq\tau\star}\theta_s}{\sum_{s\in\mathbb{N}}\theta_s}\bar{g}^\theta_{\tau^\star} + (1-\frac{\sum_{s\leq\tau\star}\theta_s}{\sum_{s\in\mathbb{N}}\theta_s})g(x_k,j^\star)\right) \leq \frac{\sum_{s\leq\tau\star}\theta_s}{\sum_{s\in\mathbb{N}}\theta_s}d_{\mathcal{C}}\left(\mathbb{E}\left[\bar{g}^\theta_{\tau^\star}\right]\right) \\ &\leq \frac{\sum_{s\leq\tau^\star}\theta_s}{\sum_{s\in\mathbb{N}}\theta_s}\varepsilon + K\frac{\sqrt{8|\mathcal{J}|\sum_{s\leq\tau^\star}\theta_s^2}}{\sum_{s\in\mathbb{N}}\theta_s} \leq \varepsilon + K\sqrt{8|\mathcal{J}|}\|\theta\|_2, \end{aligned}$$

hence the result. ■

3.2.3 Condition (1) is sufficient in Big-Match games of type I

We now turn to the case of Big-match games of type I, where only player 1 can quit. In those games, we have the following useful equivalence result.

Lemma 8 *In Big-Match games of type I, Conditions* (1), (2) *and* (3) *are all equivalent, and they are further equivalent to the usual Blackwell condition stated as*

$$\forall \mathbf{y} \in \Delta(\mathbf{J}), \exists \mathbf{x} \in \Delta(\mathbf{I}), g(\mathbf{x},\mathbf{y}) \in \mathcal{C}, \tag{5}$$

which also reads, equivalently, as

$$\forall y \in \Delta(\mathcal{J}), \exists (x, x^\star, \gamma) \in \Delta(\mathcal{I}) \times \Delta(\mathcal{I}^\star) \times [0,1], (1-\gamma)g(x,y) + \gamma g(x^\star, y) \in \mathcal{C}. \tag{6}$$

A consequence of (6) is that if player 2 plays i.i.d. according to $y \in \Delta(\mathcal{J})$, then player 1 can approach $\mathcal{C} \subset \mathbb{R}^d$ by playing $x \in \Delta(\mathcal{I})$ "perturbed" by $x^\star \in \Delta(\mathcal{I}^\star)$ with an overall total probability of absorption of γ.

Proof: We decompose the proof in three main parts.

Part a. First we argue that Condition (1) implies the Blackwell condition (5). Decompose every $\mathbf{x} \in \Delta(\mathbf{I})$ as $\mathbf{x} = (1-\gamma_{\mathbf{x}})x + \gamma_{\mathbf{x}}x^\star$, where $(x, x^\star, \gamma_{\mathbf{x}}) \in \Delta(\mathcal{I}) \times \Delta(\mathcal{I}^\star) \times$

$[0,1]$. Similarly, decompose every $\alpha \in \mathcal{M}(\mathbf{I})$ into $\alpha_0 \in \mathcal{M}(\mathcal{I})$ and $\alpha^\star \in \mathcal{M}(\mathcal{I}^\star)$. Then the fraction in Condition (1) can be rewritten into

$$\frac{g(\mathbf{x},\mathbf{y}) + g(\alpha^\star,\mathbf{y}) + \gamma_{\mathbf{x}} g(x^\star,\beta)}{1 + \|\alpha^\star\|_1 + \gamma_{\mathbf{x}}\|\beta\|_1} \in \mathcal{C}. \tag{7}$$

Now suppose that $\mathcal{C}$ satisfies Condition (1). Let $\mathbf{y} \in \Delta(\mathbf{J})$, and take an $\mathbf{x} \in \Delta(\mathbf{I})$ that gives the minimum in Condition (1). We distinguish two cases.

Suppose first that $\gamma_{\mathbf{x}} > 0$. Then, by taking $\beta_c = c \cdot \delta_j$ in (7) and letting c tend to infinity, we find $g(x^\star, j) \in \mathcal{C}$ for all $j \in \mathcal{J} = \mathbf{J}$. Hence, $g(x^\star,\mathbf{y}) \in \mathcal{C}$.

Now assume that $\gamma_{\mathbf{x}} = 0$. Define $\hat{\mathbf{x}} = \frac{\mathbf{x}+\alpha^\star}{1+\|\alpha^\star\|_1} \in \Delta(\mathbf{I})$. Then in view of (7), we have

$$g(\hat{\mathbf{x}},\mathbf{y}) = \frac{g(\mathbf{x},\mathbf{y}) + g(\alpha^\star,\mathbf{y})}{1 + \|\alpha^\star\|_1} \in \mathcal{C}.$$

So, in both cases, $\mathcal{C}$ satisfies the Blackwell condition (5).

Part b. Now we prove that the Blackwell condition (5) implies Condition (1). So, assume that $\mathcal{C}$ satisfies (5). Then, for every $\mathbf{y} \in \Delta(\mathbf{J})$, we decompose again the associated $\mathbf{x} \in \Delta(\mathbf{I})$ as $\mathbf{x} = (1-\gamma_{\mathbf{x}})x + \gamma_{\mathbf{x}} x^\star$. The choice of (x,α) where $\alpha = (\frac{1}{\gamma_{\mathbf{x}}} - 1)x^\star$ ensures that $p^\star(x,\beta) = 0$ for all β, and hence $\mathcal{C}$ satisfies Condition (1).

Part c. We already know that Condition (1) implies Condition (2), which further implies Condition (3). Since Condition (1) is equivalent to the Blackwell condition (5), it only remains to verify that Condition (3) implies (5). This can be easily checked by taking $\beta = 0$ and using the same decomposition trick as above. ■

Proposition 9 *In Big-Match games of type I, a convex set $\mathcal{C}$ is weakly approachable by player 1 if Condition* (1) *is satisfied.*

Proof: The approachability strategy is rather similar to the one introduced for Big-Match games of type II. Given the finite ε-discretization of $\Delta(\mathcal{J})$ denoted by $\{y[k], k \in [K]\}$, Lemma 8 guarantees, for any $\eta > 0$, the existence of $x[k] \in \Delta(\mathcal{I})$, $x^\star[k] \in \Delta(\mathcal{I}^\star)$ and $\gamma[k] \in [0, 1-\eta]$ such that $(1-\gamma[k])g(x[k], y[k]) + \gamma[k]g(x^\star[k], y[k])$ is 2η-close to $\mathcal{C}$.

Based on an auxiliary calibration algorithm (to be adapted and described later) whose prediction at stage τ is some $y[k_\tau] \in \Delta(\mathcal{J})$, we consider the strategy of player 1 that dictates to play at this stage

$$\begin{aligned} &x^\star[k_\tau] \ \text{ with probability } \ \gamma_\tau[k_\tau] := \frac{\gamma[k_\tau]\theta_\tau}{(1-\gamma[k_\tau])\sum_{s=\tau}^{\infty}\theta_s + \gamma[k_\tau]\theta_\tau} \\ &\text{and } x[k_\tau] \ \text{ with probability } \ 1 - \gamma_\tau[k_\tau] \ . \end{aligned}$$

Thus at stage τ, with probability $\gamma_\tau[k_\tau]$ player 1 quits according to $x^\star[k_\tau]$ (then the expected absorption payoff is $g(x^\star[k_\tau], j_\tau)$), and with the remaining probability, which is positive, he plays $x[k_\tau]$ and play does not absorb at this stage. Since the cumulative weight of all the remaining stages is $\sum_{s=\tau}^{\infty}\theta_s$, the associated expected payoff of the

decision taken at stage τ is

$$\begin{aligned}
&\gamma_\tau[k_\tau]\sum_{s=\tau}^{\infty}\theta_s g(x^\star[k_\tau],j_\tau)+(1-\gamma_\tau[k_\tau])\theta_\tau g(x[k_\tau],j_\tau)\\
&=\frac{\theta_\tau\sum_{s=\tau}^{\infty}\theta_s}{(1-\gamma[k_\tau])\sum_{s=\tau}^{\infty}\theta_s+\gamma[k_\tau]\theta_\tau}\Big(\gamma[k_\tau]g(x^\star[k_\tau],j_\tau)+(1-\gamma[k_\tau])g(x[k_\tau],j_\tau)\Big)\\
&=\theta^\star_\tau\Big(\gamma[k_\tau]g(x^\star[k_\tau],j_\tau)+(1-\gamma[k_\tau])g(x[k_\tau],j_\tau)\Big),
\end{aligned}$$

where $\theta^\star_\tau := \theta_\tau(1-\gamma_\tau[k_\tau])/(1-\gamma[k_\tau])$. As a consequence, summing over $\tau\in\mathbb{N}$ and using the fact that the game is absorbed at stage τ with probability $\gamma_\tau[k_\tau]$, we obtain that

$$\begin{aligned}
\mathbb{E}\left[\overline{g}^\theta_\infty\right]&=\sum_{\tau=1}^{\infty}\prod_{s<\tau}\big(1-\gamma_s[k_s]\big)\theta^\star_\tau\Big(\gamma[k_\tau]g(x^\star[k_\tau],j_\tau)+(1-\gamma[k_\tau])g(x[k_\tau],j_\tau)\Big)\\
&:=\sum_{\tau=1}^{\infty}\hat\theta_\tau\Big(\gamma[k_\tau]g(x^\star[k_\tau],j_\tau)+(1-\gamma[k_\tau])g(x[k_\tau],j_\tau)\Big).
\end{aligned}$$

We stress the fact that the sequence $\hat\theta_\tau$ is predicable with respect to the filtration induced by the strategies of players 1 and 2 (i.e., it does not depend on the choices made at stages $s\geq\tau+1$).

To define the strategy of player 1, we consider an auxiliary algorithm calibrated with respect to the sequence of weights $\hat\theta_\tau$ (which is possible even though $\hat\theta_\tau$ depends on the past predictions). Using the fact that $\hat\theta_s\leq\theta^\star_s\leq\theta_s/\eta$, Lemma 5 and the same argument as in Big-match games of type II, we obtain that our strategy guarantees the following:

$$d_{\mathcal{C}}\left(\mathbb{E}\left[\overline{g}^\theta_\infty\right]\right)\leq 2\eta+\frac{\varepsilon}{\eta}+K\frac{\sqrt{8|\Omega|}}{\eta}\|\theta\|_2.$$

The result follows by, for instance, taking $\eta=\sqrt{\varepsilon}$. ■

3.2.4 Condition (1) is sufficient in all generalized quitting games

Using the tools introduced in the previous subsections for Big-Match games of types I and II, we are now able to give a simple proof of the main result, that Condition (1) is sufficient to ensure weak approachability in all generalized quitting games.

We start with a useful consequence of Condition (1).

Lemma 10 *In generalized quitting games, Condition* (1) *implies that at least one of the following conditions holds:*

(a) $\exists(x,x^\star,\gamma)\in\Delta(\mathcal{I})\times\Delta(\mathcal{I}^\star)\times(0,1]$ *with*

$$g(x^\star,j)\in\mathcal{C},\ \forall j\in\mathcal{J}\quad\text{and}\quad g((1-\gamma)x+\gamma x^\star,j^\star)\in\mathcal{C},\forall j^\star\in\mathcal{J}^\star. \tag{8}$$

(b) $\forall y\in\Delta(\mathcal{J}),\exists(x,x^\star,\gamma)\in\Delta(\mathcal{I})\times\Delta(\mathcal{I}^\star)\times[0,1]$ *with*

$$g(x,j^\star)\in\mathcal{C},\forall j^\star\in\mathcal{J}^\star\quad\text{and}\quad(1-\gamma)g(x,y)+\gamma g(x^\star,y)\in\mathcal{C}. \tag{9}$$

Proof: Let $y \in \Delta(\mathbf{J})$. Assume first that Condition (1) is satisfied with some $\mathbf{x} \in \Delta(\mathbf{I})$ that puts positive weight on $\mathcal{I}^\star$, i.e. $\mathbf{x} = (1-\gamma)x + \gamma x^\star$ with $\gamma \in (0,1]$. Then, taking $\beta = c \cdot \delta_j$ in the expression of Condition (1) with c going to infinity, for any $j \in \mathbf{J}$, yields condition (a).

Otherwise, Condition (1) is satisfied for some $x \in \Delta(\mathcal{I})$. Then, the same argument as in the proof of Lemma 8 yields condition (b). ■

Proposition 11 *In generalized quitting games, a convex set $\mathcal{C}$ is weakly approachable by player 1 if Condition* (1) *is satisfied.*

Proof. Assume that Condition (1) is satisfied. Then either condition (a) or condition (b) of Lemma 10 is satisfied.

First assume that condition (a) of Lemma 10 is satisfied. Then, player 1 just has to play i.i.d. according to $(1-\gamma)x+\gamma x^\star \in \Delta(\mathbf{I})$. Indeed, then the probability of absorption at each stage is at least γ, so absorption will eventually take place with probability 1, and by condition (a) the expected absorption payoff is in $\mathcal{C}$. As a consequence,

$$d_{\mathcal{C}}\left(\mathbb{E}\left[\overline{g}^{\theta}_{\infty}\right]\right) \leq \sum_{s=1}^{\infty}(1-\gamma)^s\theta_s \leq \frac{1-\gamma}{\sqrt{2\gamma-\gamma^2}}\|\theta\|_2,$$

hence the result.

Now assume that condition (b) of Lemma 10 is satisfied. We claim that the strategy defined in the proof of Proposition 9 is an approachability strategy. Indeed, as long as player 2 does not play an absorbing action $j^\star \in \mathcal{J}^\star$, the analysis is identical.

If, on the other hand, player 2 plays $j^\star \in \mathcal{J}^\star$ at stage $\tau^\star$, then the absorbing payoff is equal to

$$\begin{aligned} &g((1-\gamma_{\tau^\star}[k_{\tau^\star}])x[k_{\tau^\star}] + \gamma_{\tau^\star}[k_{\tau^\star}]x^\star[k_{\tau^\star}], j^\star) \\ &= g(x[k_{\tau^\star}], j^\star) + \gamma_{\tau^\star}[k_{\tau^\star}]\Big(g(x^\star[k_{\tau^\star}], j^\star) - g(x[k_{\tau^\star}], j^\star)\Big), \end{aligned}$$

which is therefore within a distance of $2\gamma_{\tau^\star}[k_{\tau^\star}]$ to $\mathcal{C}$. As a consequence,

$$\begin{aligned} d_{\mathcal{C}}\left(\mathbb{E}\left[\overline{g}^{\theta}_{\infty}\right]\right) &\leq \mathbb{E}\left[\eta + \frac{\varepsilon}{\eta} + K\sqrt{8|\Omega|\sum_{s<\tau^\star}\hat{\theta}_s^2} + 2\gamma_\tau[k_{\tau^\star}]\sum_{s=\tau^\star}^{\infty}\theta_s\right] \\ &\leq \eta + \frac{\varepsilon}{\eta} + K\frac{\sqrt{8|\Omega|\sum_{s\in\mathbb{N}}\theta_s^2}}{\eta} + 2\mathbb{E}\left[\frac{\theta_{\tau^\star}}{\eta}\right] \\ &\leq \eta + \frac{\varepsilon}{\eta} + K\frac{\sqrt{8|\Omega|}}{\eta}\|\theta\|_2 + 2\frac{\|\theta\|_2}{\eta}. \end{aligned}$$

And the result follows, by taking $\eta = \sqrt{\varepsilon}$. ■

4 Weak vs Uniform Approachability

In this section, we compare the notions of weak and uniform approachability in Big-Match games.

First we consider Big-Match games of type I. In this class of games, there are convex sets that are weakly approachable but not uniformly approachable, which is illustrated by Example 15 in Section 5. Hence, in this class of games, uniform determinacy fails. In contrast, weak determinacy holds true by Proposition 2.

Now we turn our attention to Big-Match games of type II. We have the following characterization result for uniform approachability.

Proposition 12 *In Big-Match games of type II, a convex set $\mathcal{C}$ is uniformly approachable by player 1 if and only if Condition* (1) *holds.*

In view of Lemma 6, for Big-Match games of type II, Condition (1) can be rewritten as

$$\forall \mathbf{y} \in \Delta(\mathbf{J}), \exists x \in \mathcal{X}_{\mathcal{C}}, g(x, \mathbf{y}) \in \mathcal{C}, \tag{10}$$

where

$$\mathcal{X}_{\mathcal{C}} = \{x \in \Delta(\mathcal{I}) \ \text{ s.t. } \ g(x, j^\star) \in \mathcal{C}, \forall j^\star \in \mathcal{J}^\star\}.$$

Note that (10) is exactly Blackwell's approachability condition for the set $\mathcal{C}$ in the following related repeated game (with no absorption): player 1 is restricted to play mixed actions in the set $\mathcal{X}_{\mathcal{C}}$, player 2 can choose actions in $\mathbf{J}$, and the payoff is given by g.

Proof: The sufficiency part is a direct consequence of Proposition 7.

Now assume that $\mathcal{C}$ does not satisfy (10), i.e., there exists $\mathbf{y}_0 \in \Delta(\mathbf{J})$ such that $g(x, \mathbf{y}_0) \notin \mathcal{C}$ for all $x \in \mathcal{X}_{\mathcal{C}}$. By the definition of $\mathcal{X}_{\mathcal{C}}$, this implies that there exists $y_0 \in \Delta(\mathcal{J})$ such that $g(x, y_0) \notin \mathcal{C}$ for all $x \in \mathcal{X}_{\mathcal{C}}$. Since $\mathcal{X}_{\mathcal{C}}$ is compact, there exists $\delta > 0$ such that $d_{\mathcal{C}}(g(x, y_0)) \geq \delta$ for all $x \in \mathcal{X}_{\mathcal{C}}$. By continuity, this also implies that $d_{\mathcal{C}}(g(x, y_0)) \geq \delta/2$ for all mixed actions $x \in \mathcal{X}_{\mathcal{C}}^{\eta}$, the η-neighborhood of $\mathcal{X}_{\mathcal{C}}$, for $\eta > 0$ small enough. By continuity, there also exists $\varepsilon > 0$ such that for all $x \notin \mathcal{X}_{\mathcal{C}}^{\eta}$, there exists $j^\star \in \mathcal{J}^\star$ with $d_{\mathcal{C}}(g(x, j^\star)) \geq \varepsilon$.

In conclusion, if player 1 only uses mixed actions in $\mathcal{X}_{\mathcal{C}}^{\eta}$ then y_0 ensures that the average payoff is asymptotically $\delta/2$-away from $\mathcal{C}$, whereas if player 1 uses, at some stage t, a mixed action $x_t \notin \mathcal{X}_{\mathcal{C}}^{\eta}$ then there is $j^\star \in \mathcal{J}^\star$ such that $g(x_t, j^\star)$ is ε-away from $\mathcal{C}$. Thus, $\mathcal{C}$ is not uniformly approachable. ■

In Big-Match games of type II, uniform determinacy fails: there are convex sets that are neither uniformly approachable nor uniformly excludable. This is demonstrated by Example 16 in Section 5. The question remains open whether weak determinacy holds true for Big-Match games of type II.

5 Examples

In this section we consider a number of examples. In all these examples, player 1 is the row player and player 2 is the column player. Quitting actions of the players are marked with a superscript $\star$. The payoffs are given by the corresponding matrices.

Example 13 The following generalized quitting game shows that Condition (1) is generally not necessary for weak (and hence for uniform) approachability.

	$L^\star$	R
$T^\star$	1	1
B	-1	-1

In this game, actions T and L are quitting, whereas actions B and R are non-quitting.

The set $\mathcal{C} = \{0\}$ does not satisfy Condition (1). Indeed, the distance in (1) can be made arbitrary close to 1 by player 2 as follows. Choose any $\mathbf{y} \in \Delta(\mathbf{J})$. Then given $\mathbf{x} \in \Delta(\mathbf{I})$ and $\alpha \in \mathcal{M}(\mathbf{I})$, choose $\beta = (0, r)$ with a large r if $\mathbf{x}$ puts a positive weight on action T, and choose $\beta = (r, 0)$ with a large r otherwise. This choice of β implies that the fraction in (1) is close to 1 in the former case (due to the absorption payoff of 1 in entry (T, R)), and close to -1 in the latter case (due to the absorption payoff of -1 in entry (B, L)). In either case, the distance in (1) is close to 1.

Yet, playing $1/2T + 1/2B$ at the first stage and B at all remaining stages approaches $\{0\}$, both weakly and uniformly.

We remark that Condition (1) is generally not necessary for weak (and hence uniform) approachability even in Big-Match games of type II. This is shown later by Example 16, which involves a more difficult proof. $\diamond$

Example 14 The next game demonstrates that weak (and hence uniform) determinacy fails in generalized quitting games, and that Conditions (2) and (3) are generally not sufficient for weak (and hence uniform) approachability.

	$L^\star$	$R^\star$
$T^\star$	1	0
$B^\star$	0	-1

In this game all actions are quitting. Note that this game is not a Big-Match game, as neither player has a non-quitting action. (Later, we will see versions of this game when some of the actions are non-quitting instead – see Examples 15 and 17.)

The set $\mathcal{C} = \{0\}$ satisfies Conditions (2) and (3), due to the fact that α is the last in the order of terms in (2) and (3). Yet, $\{0\}$ is trivially not approachable and not excludable, neither weakly nor uniformly.

Note that Blackwell's approachability condition is actually satisfied in this game: against a strategy that plays $yL + (1-y)R$ at the first stage, player 1 can ensure by playing $(1-y)T + yB$ that the expected payoff is at all stages equal to 0.

We remark that Condition (3) is generally not sufficient for weak (and hence uniform) approachability even in Big-Match games of type II. This is shown later by Example 17, which involves a more difficult proof. $\diamond$

Example 15 The following game shows that, in Big-Match games of type I, there are convex sets that are weakly approachable but not uniformly approachable, and thus uniform determinacy fails (a property already noted by S. Sorin).

	L	R
T^*	1	0
B	0	-1

Note that the payoffs are identical to those in Example 14. However, action B in this game is non-quitting, which increases player 1's possibilities tremendously. Indeed, playing action B gives player 1 the option to wait and see how player 2 behaves.

Consider the set $\mathcal{C} = \{0\}$. This set satisfies Condition (1), which follows easily from Lemma 8. Hence, by Proposition 9, $\{0\}$ is weakly approachable by player 1. This also implies that $\{0\}$ is not uniformly excludable by player 2.

The set $\{0\}$ is however not uniformly approachable. The main reason is that player 1 cannot use the quitting action T effectively, since it backfires if player 2 places positive probability on action L. A precise argument is given later in Proposition 21. Hence, the game is not uniformly determined. $\diamond$

Example 16 The following game shows that, in Big-Match games of type II, Conditions (1) and (2) are not necessary for weak approachability, and that there convex sets that are weakly approachable but not uniformly approachable, and thus uniform determinacy fails.

	$L^\star$	R
T	1	1
B	0	-1

Consider the set $\mathcal{C} = \{0\}$. Conditions (1) and (2) are not satisfied for $\{0\}$, which can be easily verified with the help of Lemma 6. Indeed, for $y = (1/2, 1/2)$, there is no $x \in \Delta(\mathcal{I})$ satisfying (4). Yet, as we show in Appendix A.2, player 1 can weakly approach $\mathcal{C}$. This shows that Conditions (1) and (2) are not necessary for weak approachability in Big-Match games of type II.

In view of Proposition 12, the set $\{0\}$ is not uniformly approachable by player 1. Since $\{0\}$ is weakly approachable by player 1, we also obtain that $\{0\}$ is not uniformly excludable by player 2. Hence, the game is not uniformly determined. $\diamond$

Example 17 The following game shows that, in Big-Match games of type II, Condition (3) is not sufficient for weak approachability.

	$L^\star$	R
T	1	0
B	0	-1

Consider the set $\mathcal{C} = \{0\}$. We argue that Condition (3) is satisfied. Indeed, take any $\mathbf{y} \in \Delta(\mathbf{J})$ and $\beta \in \mathcal{M}(\mathbf{J})$. Then by choosing

$$\mathbf{x} = \left(\frac{\mathbf{y}_R}{1+\beta_L}, \frac{\mathbf{y}_L + \beta_L}{1+\beta_L} \right) \quad \text{and} \quad \alpha = (0,0),$$

we obtain

$$g(\mathbf{x}, \mathbf{y}) + g^*(\alpha, \mathbf{y}) + g^*(\mathbf{x}, \beta) = \frac{\mathbf{y}_R}{1+\beta_L} \cdot (\mathbf{y}_L + \beta_L) - \frac{\mathbf{y}_L + \beta_L}{1+\beta_L} \cdot \mathbf{y}_R = 0,$$

which implies that Condition (3) holds indeed.

Clearly, Conditions (1) and (2) are not satisfied, which can be easily verified with the help of Lemma 6. We will show in Appendix A.1 that this prevents the weak approachability of $\{0\}$. $\diamond$

6 Conclusion

We have introduced the model of stochastic games with vector payoffs and have exhibited a sufficient condition and a strongly related necessary condition for the weak approachability of a convex set in the class of generalized quitting games. In Big-Match games of type I the conditions coincide, but generally they differ in Big-Match games of type II.

Some of our conditions are also useful for uniform approachability, though in a non-obvious way. In fact, as we have seen, weak and uniform approachability conditions drastically differ, even in Big-Match games.

When Condition (1) is satisfied, we have also provided explicit strategies for weak approachability based on an auxiliary calibration strategy (itself induced by a traditional approachability result). The question of optimal rates of convergence and efficient algorithms are left open: our techniques provide qualitative results, and, unfortunately the rates decrease with the dimension, as we need to consider ε-discretization of $\Delta(\mathcal{J})$.

References

[1] J. Abernethy, P.L. Bartlett, and E. Hazan. Blackwell approachability and low-regret learning are equivalent. *J. Mach. Learn. Res.: Workshop Conf. Proc.*, 19:27–46, 2011.

[2] R. J. Aumann and M. B. Maschler. *Repeated Games with Incomplete Information.* MIT Press, Cambridge, MA, 1995. With the collaboration of Richard E. Stearns.

[3] A. Bernstein, S. Mannor, and N. Shimkin. Opportunistic strategies for generalized no-regret problems. *J. Mach. Learn. Res.: Workshop Conf. Proc.*, 30:158–171, 2013.

[4] T. Bewley and E. Kohlberg. The asymptotic theory of stochastic games. *Mathematics of Operations Research*, 1(3):197–208, 1976.

[5] D. Blackwell. An analog of the minimax theorem for vector payoffs. *Pacific J. Math.*, 6:1–8, 1956.

[6] D. Blackwell. Controlled random walks. In *Proceedings of the International Congress of Mathematicians, 1954, Amsterdam, vol. III*, pages 336–338, 1956.

[7] D. Blackwell and T. Ferguson. The big match. *The Annals of Mathematical Statistics*, 39(1):159–163, 1968.

[8] A. P. Dawid. Self-calibrating priors do not exist: Comment. *J. Amer. Statist. Assoc.*, 80:340–341, 1985.

[9] D. P. Foster. A proof of calibration via blackwell's approachability theorem. *Games and Economic Behavior*, 29:73 – 78, 1999.

[10] D. P. Foster, A. Rakhlin, K. Sridharan, and A. Tewari. Complexity-based approach to calibration with checking rules. *J. Mach. Learn. Res.: Workshop Conf. Proc.*, 19:293–314, 2011.

[11] D. P. Foster and R. V. Vohra. Calibrated learning and correlated equilibrium. *Games Econom. Behav.*, 21:40–55, 1997.

[12] D. Gillette. Stochastic games with zero stop probabilities. In *Contributions to the Theory of Games 3*. 1957.

[13] S. Hart and A. Mas-Collel. A simple adaptive procedure leading to correlated equilibrium. *Econometrica*, 68(5):1127–1150, 2000.

[14] J. Hörner and Lovo S. Belief-free equilibria in games with incomplete information. *Econometrica*, 77(2):453–487, 2009.

[15] Lovo S. Hörner, J. and Tomala T. Belief-free equilibria in games with incomplete information: Characterization and existence. *Journal of Economic Theory*, 146(5):1770–1795, 2011.

[16] E. Kohlberg. Repeated games with absorbing states. 2(4):724–738, 1974.

[17] E. Kohlberg. Optimal strategies in repeated games with incomplete information. *International Journal of Game Theory*, 4(1):7–24, 1975.

[18] R. Laraki. Explicit formulas for repeated games with absorbing states. *International Journal of Game Theory*, 39(1):53–69, 2010.

[19] R. Laraki and S. Sorin. *Advances in Zero-Sum Dynamic Games*. 2014. Chapter 2 in Handbook of Game Theory IV, edited by H. P. Young and S. Zamir.

[20] S. Mannor and V. Perchet. Approachability, fast and slow. *Journal of Machine Learning Research: Workshop and Conference Proceedings (COLT)*, 30:474–488, 2013.

[21] S. Mannor, V. Perchet, and G. Stoltz. Set-valued approachability and online learning under partial monitoring. *Journal of Machine Learning Research*, page to appear, 2015.

[22] J. Mertens and A. Neyman. Stochastic games. *International Journal of Game Theory*, 10(2):53–66, 1981.

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

[24] V. Perchet. Calibration and internal no-regret with random signals. *Lecture Notes in Computer Science (ALT)*, 5809:68–82, 2009.

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

[26] V. Perchet. Approachability, regret and calibration. implications and equivalences. *Journal of Dynamics and Games*, 1:181–254, 2014.

[27] V. Perchet. Internal regret with partial monitoring : Calibration-based optimal algorithms. *Journal of Machine Learning Research*, 12:1893–1921, 2014.

[28] V. Perchet. Exponential weight approachability, applications to calibration and regret minimization. *Dynamic Games And Applications*, page to appear, 2015.

[29] A. Rakhlin, K. Sridharan, and A. Tewari. Online learning: Beyond regret. *J. Mach. Learn. Res.: Workshop Conf. Proc.*, 19:559–594, 2011.

[30] J. Renault and T. Tomala. Communication equilibrium payoffs in repeated games with imperfect monitoring. *Games and Economic Behavior*, 49:313–344, 2004.

[31] D. Roseberg. Zero sum absorbing games with incomplete information on one side : Asymptotic analysis. *SIAM Journal on Control and Optimization*, 39:208–225, 2000.

[32] L. S. Shapley. Stochastic games. *Proceedings of the National Academy of Sciences*, 39(10):1095–1100, 1953.

[33] S. Sorin. "big match" with lack of information on one side (part i). *International Journal of Game Theory*, 13(3):201–255, 1984.

[34] S. Sorin. "big match" with lack of information on one side (part ii). *International Journal of Game Theory*, 14(3):173–204, 1985.

[35] T. Tomala. Belief-free communication equilibria in repeated games. *Mathematics of Operations Research*, 38(4):617–637, 2013.

[36] N. Vieille. Weak approachability. *Math. Oper. Res.*, 17:781–791, 1992.

[37] G. Vigeral. A zero-sum stochastic game with compact action sets and no asymptotic value. *Dynamic Games and Applications*, 3(2):172–186, 2013.

[38] B. Ziliotto. Zero-sum repeated games: Counterexamples to the existence of the asymptotic value and the conjecture maxmin $= \lim v_n$. *The Annals of Probability*, 44(2):1107–1133, 2016.

A On weak approachability in Big-Match games of type II

We have proved that Condition (1) and Condition (3) are respectively sufficient and necessary for the weak approachability of a convex target set $\mathcal{C} \subset \mathbb{R}^d$ in generalized quitting games (cf. Theorem 1). Even though these two conditions are equivalent in Big-Match games of type I (cf. Lemma 8), generally there is still a gap between them.

The goal of this section is to provide further insight by considering specific examples in the calls of Big-Match games of type II. Recall that, in this class, Conditions (1) and (2) are equivalent (cf. Lemma 6).

In the next subsections, we prove that Condition (3) is not sufficient and that Conditions (1) and (2) are not necessary for weak approachability in Big-Match games of type II. We then provide some necessary and sufficient conditions, in the very specific case where player 2 has only one absorbing and one non-absorbing action, based on techniques in continuous time, similar to the one used by Vieille [36] when he proves weak determinacy for non convex sets in the classical Blackwell setting.

A.1 Condition (3) is not sufficient for weak approachability

We revisit the Big-Match game of type II from Example 17.

	$L^\star$	R
T	1	0
B	0	-1

Consider again the set $\mathcal{C} = \{0\}$. We will now show that $\{0\}$ is not weakly approachable.

For the sake of simplicity, we will consider λ-discounted payoffs. Given the strategies σ and τ, we denote

$$\bar{g}^\lambda_\infty(\sigma,\tau) = \mathbb{E}_{(\sigma,\tau)}\left(\sum_{t=1}^{\infty} \lambda(1-\lambda)^{j-1} g(i_t,j_t)\right) = \sum_{t=1}^{\infty} \lambda(1-\lambda)^{t-1}\mathbb{E}_{\sigma,\tau} g(i_t,j_t).$$

Similarly, the expected λ-discounted payoff up to period k is

$$\bar{g}^\lambda_k(\sigma,\tau) = \mathbb{E}_{(\sigma,\tau)}\left(\sum_{t=1}^{k} \lambda(1-\lambda)^{t-1} g(i_t,j_t)\right) = \sum_{t=1}^{k} \lambda(1-\lambda)^{t-1}\mathbb{E}_{\sigma,\tau} g(i_t,j_t).$$

Proposition 18 *There exists a $\lambda^\star \in (0,1)$ such that, for every $\lambda \in (0,\lambda^\star)$ and every strategy σ of player 1, there is a strategy τ of player 2 with the property that $\bar{g}^\lambda_\infty(\sigma,\tau) \notin [-\frac{1}{2e}, \frac{1}{2e}]$.*

Consequently, Condition (3) is not sufficient for weak approachability.

Idea of the proof: Denote $\varepsilon = \frac{1}{2e}$, and consider a small discount factor $\lambda \in (0,1)$. The main problem for player 1 is the following. Player 1 needs to guarantee an expected λ-discounted payoff close to 0 against the strategy R^∞ of player 2, which means that he

has to put large probabilities on action T at some point. However, a large probability on T could easily backfire if player 2 plays L instead. As our analysis will show, player 1's best chance is to gradually increase the probability on T. The maximal probability that player 1 can put on T at period 1 is ε, since higher probabilities would lead to an expected λ-discounted payoff higher than ε if player 2 plays L. Thus, player 1 should play T with probability ε at period 1, and this is safe against action L. If player 2 plays R at period 1, then the expected payoff at period 1 is negative, $-1+\varepsilon$ to be precise, and this allows player 1 to increase the probability on action T up to $\varepsilon + \frac{\lambda}{1-\lambda}$ at period 2. Indeed, if player 2 plays R at period 1 and plays L at period 2, then the expected λ-discounted payoff is exactly ε. By continuing so, we obtain a sequence $(\overline{z}_{\lambda k})_{k\in\mathbb{N}}$ for the probabilities on T. This sequence is strictly increasing until it reaches 1 and then it stays 1 forever. Let $\sigma^\star_\lambda$ denote the corresponding Markov strategy for player 1 which uses these probabilities on T during the game. We will show that this strategy $\sigma^\star_\lambda$ is indeed player 1's best chance. However, for small $\lambda \in (0,1)$, the probabilities on T do not converge fast enough to 1 and consequently, when player 2 always plays action R, the expected λ-discounted payoff stays below $-\varepsilon$.

Step 1: A reduction. We argue that it is sufficient to prove the statement of the proposition when we only consider Markov strategies (i.e. strategies where the probabilities on the actions depend only on the stage). That is, if there exists a $\lambda_M \in (0,1)$ such that, for every $\lambda \in (0,\lambda_M)$ and every Markov strategy σ for player 1, there is a Markov strategy τ for player 2 with the property that $\bar{g}^\lambda_\infty(\sigma,\tau) \notin [-\varepsilon,\varepsilon]$, then the proposition follows with $\lambda^\star = \lambda_M$.

Proof for step 1: Assume that such a $\lambda_M \in (0,1)$ exists. Consider an arbitrary strategy σ' for player 1. For every $k \in \mathbb{N}$, let p_k denote the probability, with respect σ' and R^∞, that player 1 plays action T at period k. Now define the Markov strategy σ for player 1 which prescribes to play action T with probability p_k at every period $k \in \mathbb{N}$. Let $\lambda \in (0,\lambda_M)$. Since σ is a Markov strategy, by our assumption there exists a Markov strategy τ for player 2 with the property that $\bar{g}^\lambda_\infty(\sigma,\tau) \notin [-\varepsilon,\varepsilon]$. Because (σ',τ) and (σ,τ) generate the same expected payoff for each period $k \in \mathbb{N}$, we have $\bar{g}^\lambda_\infty(\sigma',\tau) = \bar{g}^\lambda_\infty(\sigma,\tau)$. Hence, $\bar{g}^\lambda_\infty(\sigma',\tau) \notin [-\varepsilon,\varepsilon]$, and the proposition will then follow by choosing $\lambda^\star = \lambda_M$.

Step 2: The main strategy $\sigma^\star_\lambda$. Define

$$z_{\lambda k} = \varepsilon + (k-1)\frac{\lambda}{1-\lambda} \quad \text{and} \quad \overline{z}_{\lambda k} = \min\{z_{\lambda k}, 1\}$$

for every $\lambda \in (0,1)$ and every $k \in \mathbb{N}$. For every $\lambda \in (0,1)$, the sequence $(z_{\lambda k})_{k=1}^\infty$ is positive and strictly increasing, and it diverges to infinity. So, for every $\lambda \in (0,1)$, there is a unique $k_\lambda \in \mathbb{N}$ such that $z_{\lambda k_\lambda} \le 1 < z_{\lambda,k_\lambda+1}$.

For every $\lambda \in (0,1)$, let $\sigma^\star_\lambda$ be the Markov strategy for player 1 which prescribes to play action T with probability $\overline{z}_{\lambda k}$ at every period $k \in \mathbb{N}$. We argue that for every $\lambda \in (0,1)$

$$\bar{g}^\lambda_\infty(\sigma^\star_\lambda, R^\infty) = \varepsilon - \lambda(1-\lambda)^{k_\lambda - 1} - (1-\lambda)^{k_\lambda} z_{\lambda,k_\lambda} < \varepsilon - (1-\lambda)^{k_\lambda}, \tag{11}$$

that for every $\lambda \in (0,1)$ and $k \in \{1,\dots,k_\lambda\}$

$$\bar{g}^\lambda_\infty(\sigma^\star_\lambda, L[k]) = \varepsilon, \tag{12}$$

and that for every $\lambda \in (0,1)$ and $k \in \mathbb{N}$ with $k > k_\lambda$

$$\bar{g}^\lambda_\infty(\sigma^\star_\lambda, L[k]) < \varepsilon. \tag{13}$$

Proof for step 2: Fix an arbitrary $\lambda \in (0,1)$. We first prove

$$\bar{g}^\lambda_k(\sigma^\star_\lambda, R^\infty) = \varepsilon - \lambda(1-\lambda)^{k-1} - (1-\lambda)^k z_{\lambda k} \tag{14}$$

for every $k \in \{1,\dots,k_\lambda\}$ by induction on k. For $k = 1$ we have

$$\bar{g}^\lambda_1(\sigma^\star_\lambda, R^\infty) = \lambda(0 \cdot z_{\lambda 1} - 1 \cdot (1 - z_{\lambda 1})) = \lambda(-1+\varepsilon)$$

and

$$\varepsilon - \lambda - (1-\lambda)z_{\lambda 1} = \varepsilon - \lambda - (1-\lambda)\varepsilon = \lambda(-1+\varepsilon),$$

hence (14) is true for $k = 1$. Now, suppose that (14) is true for some k and that $k+1 \le k_\lambda$. Then

$$\begin{aligned}
\bar{g}^\lambda_{k+1}(\sigma^\star_\lambda, R^\infty) &= \bar{g}^\lambda_k(\sigma^\star_\lambda, R^\infty) + \lambda(1-\lambda)^k(-1 + z_{\lambda,k+1}) \\
&= \varepsilon - \lambda(1-\lambda)^{k-1} - (1-\lambda)^k z_{\lambda k} + \lambda(1-\lambda)^k(-1 + z_{\lambda,k+1}) \\
&= \varepsilon(1-(1-\lambda)^{k+1}) - \lambda(1-\lambda)^{k-1} - (k-1)\lambda(1-\lambda)^{k-1} - \lambda(1-\lambda)^k + k\lambda^2(1-\lambda)^{k-1} \\
&= \varepsilon(1-(1-\lambda)^{k+1}) - \lambda(1-\lambda)^k - k\lambda(1-\lambda)^{k-1} + k\lambda^2(1-\lambda)^{k-1} \\
&= \varepsilon(1-(1-\lambda)^{k+1}) - \lambda(1-\lambda)^k - k\lambda(1-\lambda)^k \\
&= \varepsilon - \lambda(1-\lambda)^k - (1-\lambda)^{k+1} z_{\lambda,k+1},
\end{aligned}$$

which proves (14) for $k+1$. So we have shown (14) for every $k \in \{1,\dots,k_\lambda\}$.

Since $\sigma^\star_\lambda$ prescribes action T with probability 1 from period $k_\lambda + 1$ onwards, (14) for k_λ implies

$$\begin{aligned}
\bar{g}^\lambda_\infty(\sigma^\star_\lambda, R^\infty) &= \bar{g}^\lambda_{k_\lambda}(\sigma^\star_\lambda, R^\infty) \\
&= \varepsilon - \lambda(1-\lambda)^{k_\lambda - 1} - (1-\lambda)^{k_\lambda} z_{\lambda,k_\lambda} \\
&< \varepsilon - \lambda(1-\lambda)^{k_\lambda - 1} - (1-\lambda)^{k_\lambda}(1 - \tfrac{\lambda}{1-\lambda}) \\
&= \varepsilon - (1-\lambda)^{k_\lambda},
\end{aligned}$$

which proves (11).

In view of (14), for any $k \in \{1,\dots,k_\lambda\}$ we have

$$\begin{aligned}
\bar{g}^\lambda_k(\sigma^\star_\lambda, R^\infty) + (1-\lambda)^k z_{\lambda,k+1} &= \varepsilon - \lambda(1-\lambda)^{k-1} - (1-\lambda)^k z_{\lambda k} + (1-\lambda)^k z_{\lambda,k+1} \\
&= \varepsilon - \lambda(1-\lambda)^{k-1} - (k-1)\lambda(1-\lambda)^{k-1} + k\lambda(1-\lambda)^{k-1} \\
&= \varepsilon. \qquad (15)
\end{aligned}$$

Note that $\bar{g}^\lambda_\infty(\sigma^\star_\lambda, L[1]) = z_{\lambda 1} = \varepsilon$, and for every $k \in \{2, \ldots, k_\lambda\}$ we have by (15) that

$$\bar{g}^\lambda_\infty(\sigma^\star_\lambda, L[k]) = \bar{g}^\lambda_{k-1}(\sigma^\star_\lambda, R^\infty) + (1-\lambda)^{k-1} z_{\lambda k} = \varepsilon.$$

Hence, we have proven (12).

Finally, assume that $k \in \mathbb{N}$ with $k > k_\lambda$. Since $\sigma^\star_\lambda$ puts probability 1 on action T from period $k_\lambda + 1$ onwards, by using (15) for k_λ we obtain

$$\begin{aligned}
\bar{g}^\lambda_\infty(\sigma^\star_\lambda, L[k]) &\leq \bar{g}^\lambda_\infty(\sigma^\star_\lambda, L[k_\lambda + 1]) \\
&= \bar{g}^\lambda_{k_\lambda}(\sigma^\star_\lambda, R^\infty) + (1-\lambda)^{k_\lambda} \\
&< \bar{g}^\lambda_{k_\lambda}(\sigma^\star_\lambda, R^\infty) + (1-\lambda)^{k_\lambda} z_{\lambda, k_\lambda + 1} \\
&= \varepsilon,
\end{aligned}$$

which proves (13).

Step 3: The main strategy $\sigma^\star_\lambda$ is the best against R^∞. We argue that, for every $\lambda \in (0,1)$ and every Markov strategy σ for player 1 for which $\bar{g}^\lambda_\infty(\sigma, L[k]) \leq \varepsilon$ holds for every $k \in \mathbb{N}$, we have

$$\bar{g}^\lambda_\infty(\sigma, R^\infty) \leq \bar{g}^\lambda_\infty(\sigma^\star_\lambda, R^\infty). \tag{16}$$

Proof for step 3: Let $\lambda \in (0,1)$ and let $\sigma = (p_k, 1-p_k)_{k=1}^\infty$ be such a Markov strategy. Since $\bar{g}^\lambda_\infty(\sigma, R^\infty) \leq \bar{g}^\lambda_{k_\lambda}(\sigma, R^\infty)$, it suffices to prove

$$\bar{g}^\lambda_{k_\lambda}(\sigma, R^\infty) \leq \bar{g}^\lambda_\infty(\sigma^\star_\lambda, R^\infty).$$

To prove this inequality, in view of (11), it suffices in turn to show that for every every $k \in \{1, \ldots, k_\lambda\}$

$$\bar{g}^\lambda_k(\sigma, R^\infty) \leq \varepsilon - \lambda(1-\lambda)^{k-1} - (1-\lambda)^k z_{\lambda k}. \tag{17}$$

We do so by induction on k. So first take $k = 1$. Since $u_\lambda(\sigma, L[1]) \leq \varepsilon$ by assumption and $u_\lambda(\sigma, L[1]) = p_1$, we have $p_1 \leq \varepsilon$. Hence,

$$\bar{g}^\lambda_1(\sigma, R^\infty) = \lambda(-1 + p_1) \leq -\lambda + \varepsilon\lambda = \varepsilon - \lambda - (1-\lambda) z_{\lambda 1},$$

where we used that $z_{\lambda 1} = \varepsilon$. Thus, (17) holds for $k = 1$. Now assume that (17) holds for some k and that $k + 1 \leq k_\lambda$. Since $\bar{g}^\lambda_\infty(\sigma, L[k+1]) \leq \varepsilon$ by assumption and

$$\bar{g}^\lambda_\infty(\sigma, L[k+1]) = \bar{g}^\lambda_k(\sigma, R^\infty) + (1-\lambda)^k p_{k+1},$$

we have

$$p_{k+1} \leq \frac{\varepsilon - \bar{g}^\lambda_k(\sigma, R^\infty)}{(1-\lambda)^k}.$$

Therefore,

$$\begin{aligned}
\bar{g}^\lambda_{k+1}(\sigma, R^\infty) &= \bar{g}^\lambda_k(\sigma, R^\infty) + \lambda(1-\lambda)^k(-1 + p_{k+1}) \\
&\leq \bar{g}^\lambda_k(\sigma, R^\infty) - \lambda(1-\lambda)^k + \varepsilon\lambda - \lambda u_{\lambda k}(\sigma, R^\infty) \\
&= (1-\lambda)\bar{g}^\lambda_k(\sigma, R^\infty) - \lambda(1-\lambda)^k + \varepsilon\lambda \\
&\leq (1-\lambda)[\varepsilon - \lambda(1-\lambda)^{k-1} - (1-\lambda)^k z_{\lambda k}] - \lambda(1-\lambda)^k + \varepsilon\lambda \\
&= \varepsilon - \lambda(1-\lambda)^k - (1-\lambda)^{k+1} z_{\lambda, k+1},
\end{aligned} \tag{18}$$

we the last equality follows from the definitions of $z_{\lambda k}$ and $z_{\lambda,k+1}$. Thus, (17) holds for $k+1$ too. The proof for step 3 is now complete.

Step 4: Conclusion of the proof of the proposition. We have

$$\lim_{\lambda\downarrow 0} (1-\lambda)^{(1-\varepsilon)\frac{1-\lambda}{\lambda}+1} = e^{-1+\varepsilon},$$

because

$$\begin{aligned}\lim_{\lambda\downarrow 0} \ln\left[(1-\lambda)^{(1-\varepsilon)\frac{1-\lambda}{\lambda}+1}\right] &= \lim_{\lambda\downarrow 0} \left[\left((1-\varepsilon)\tfrac{1-\lambda}{\lambda}+1\right) \ln(1-\lambda)\right] \\ &= (1-\varepsilon) \lim_{\lambda\downarrow 0} \frac{\ln(1-\lambda)}{\lambda} \\ &= -1+\varepsilon.\end{aligned}$$

Consequently, there is a $\lambda\star \in (0,1)$ so that for every $\lambda \in (0,\lambda\star)$

$$(1-\lambda)^{(1-\varepsilon)\frac{1-\lambda}{\lambda}+1} \geq e^{-1}.$$

Let $\lambda \in (0,\lambda\star)$. By step 1, it is sufficient to consider a Markov strategy σ for player 1. If $\bar{g}^{\lambda}_{\infty}(\sigma, L[k]) > \varepsilon$ for some $k \in \mathbb{N}$, then statement of the proposition is valid. So, suppose that $\bar{g}^{\lambda}_{\infty}(\sigma, L[k]) \leq \varepsilon$ for all $k \in \mathbb{N}$. Then, by (16) and (11), we obtain

$$\bar{g}^{\lambda}_{\infty}(\sigma, R^{\infty}) \leq \bar{g}^{\lambda}_{\infty}(\sigma^{\star}_{\lambda}, R^{\infty}) < \varepsilon - (1-\lambda)^{k_{\lambda}}.$$

Because $z_{\lambda,k_{\lambda}} \leq 1$, we obtain

$$k_{\lambda} \leq (1-\varepsilon)\tfrac{1-\lambda}{\lambda} + 1,$$

which implies that

$$\bar{g}^{\lambda}_{\infty}(\sigma, R^{\infty}) < \varepsilon - (1-\lambda)^{(1-\varepsilon)\frac{1-\lambda}{\lambda}+1} \leq \varepsilon - \frac{1}{e} = \frac{1}{2e} - \frac{1}{e} = -\frac{1}{2e} = -\varepsilon.$$

Since $\bar{g}^{\lambda}_{\infty}(\sigma, R^{\infty}) < -\varepsilon$, the proof of the proposition is complete. ■

A.2 Conditions (1) and (2) are not necessary for weak approachability

We revisit the Big-Match game of type II from Example 16.

	$L^{\star}$	R
T	1	1
B	0	-1

Consider the set $\mathcal{C} = \{0\}$. As discussed earlier, Conditions (1) and (2) are not satisfied for $\mathcal{C}$. Now we argue that player 1 can weakly approach $\mathcal{C}$.

The idea (formalized later on) behind the approachability strategy, for the T–times repeated game, is the following. At stage 1, player 1 plays action T with probability

$p_1 = 0$. If the game absorbs at stage 1, the payoff is 0 and is in $\mathcal{C}$. Otherwise, at stage 2, player 1 plays action T with probability $p_2 = 1/(T-1)$. If the game absorbs at stage 2, the total payoff is 0, while otherwise the average payoff up to stage 2 is

$$\frac{1}{2}(-1 - \frac{T-3}{T-1}) = -\frac{T-2}{T-1}.$$

By following this idea, the probability of playing T at stage 3 is then $p_3 = 2/(T-1)$ and the average payoff up to stage 3 is $-\frac{T-3}{T-1}$, etc.

This ensures that at stage T, the cumulative payoff is exactly equal to 0.

This technique can be generalized and formalized as follows. For a mixed action $\mathbf{x} \in \Delta(\mathbf{I})$ for player 1, we use the shorter notations $g_R(\mathbf{x}) = g(\mathbf{x}, R)$ and $g_L^*(\mathbf{x}) = g^*(\mathbf{x}, L)$. In the remaining part of this subsection, we do not need to assume that the target set $\mathcal{C}$ is convex.

Proposition 19 *If a set $\mathcal{C}$ is weakly approachable, then there is a measurable mapping $\xi : [0,1] \to \Delta(\mathbf{I})$ such that for almost every $t \in [0,1]$, $\int_0^t g_R(\xi(s))ds + (1-t)g_L^\star(\xi(t)) \in \mathcal{C}$.*

Proof. Suppose that player 1 can weakly approach $\mathcal{C}$. Then, for each $\varepsilon > 0$, there is T_ε, such that for every $T \geq T_\varepsilon$, there is $\{x^{T,\varepsilon}(k) \in \Delta(\mathbf{I}), \quad k = 1, ..., T\}$, such that for every $s \in [0,1]$:

$$\sum_{k=1}^{\lfloor sT \rfloor} \frac{g_R(x^{T,\varepsilon}(k))}{T} + (1 - \frac{\lfloor sT \rfloor}{T}) g_L^\star(x^{T,\varepsilon}(\lfloor sT \rfloor + 1)) \in \mathcal{C} + \varepsilon,$$

where $\lfloor r \rfloor$ is the integer part of r. Defining $\xi^{T,\varepsilon}(s) = x^{T,\varepsilon}(\lfloor sT \rfloor + 1)$, we obtain that for every $x \in [0,1]$:

$$\int_0^x g_R(\xi^{T,\varepsilon}(s))ds + (1 - \frac{\lfloor xT \rfloor}{T}) g_L^\star(\xi^{T,\varepsilon}(x)) \in \mathcal{C} + \varepsilon$$

We conclude by simply tending T to infinity and ε to zero. ■

Up to a continuity issue, we obtain a converse.

Proposition 20 *If there is a continuous mapping $\xi : [0,1] \to \Delta(\mathbf{I})$ such that for every $t \in [0,1]$, $\int_0^t g_R(\xi(s))ds + (1-t)g_L^\star(\xi(t)) \in \mathcal{C}$, then $\mathcal{C}$ is weakly approachable.*

Proof. For each $\varepsilon > 0$, let T_ε be sufficiently large so that for every $T \geq T_\varepsilon$ and every s and t in $[0,1]$, if $|s-t| \leq \frac{1}{T}$ then $\|\xi(s) - \xi(t)\|_1 \leq \varepsilon$. Now, for each T, defining $x^T(k) = \xi(\frac{k}{T})$ we obtain a strategy that satisfies, for every $K \in \mathbb{N}$:

$$\sum_{k=1}^{K} \frac{g_R(x^T(k))}{T} + (1 - \frac{K}{T}) g_L^\star(x^T(K+1)) \in \mathcal{C} + \varepsilon.$$

Thus, no matter the time K where player 2 plays L, the total average payoff will always be ε-close to $\mathcal{C}$. ■

The above condition in Proposition 20 seems not easy to check, as it is merely a rewriting of the approachability objectives in continuous time. However, it can be helpful in practice. For instance, it allows us to prove (and to find a strategy) that for any $p \geq 1$, player 1 can weakly approach $\{0\}$ in the following game (recall that, as shown in the previous subsection, player 1 cannot weakly approach $\{0\}$ when $p = 0$):

	L^*	R
T	1	p
B	0	-1

To prove our claim, it is sufficient to find a C^1 function $\xi : [0,1] \to [0,1]$ (where $\xi(s)$ is the probability of playing T at time s) such that for every t:

$$\int_0^t (\xi(s)p - (1-\xi(s))ds + (1-t)\xi(t) = 0,$$

which is equivalent, by differentiating, to $\xi(0) = 0$ and for every t:

$$\xi(t)(p+1) - 1 - \xi(t) + (1-t)\frac{d\xi(t)}{dt} = 0,$$

or equivalently

$$\xi(t)p - 1 + (1-t)\frac{d\xi(t)}{dt} = 0, \quad \xi(0) = 0.$$

This differential equation has a C^1 solution $\xi(t) = \frac{1}{p}(1-(1-t)^p))$ that belongs to $[0,1]$, satisfies $\xi(0) = 0$. This has an interpretation in terms of continuous strategy:

$$(1-t)^p\, B + (1-(1-t)^p)(\frac{1}{p}\, T + (1-\frac{1}{p})\, B).$$

In words, this strategy stipulates that player 1 starts playing $x_0 = B$ (the strategy such that $g_L^\star(x_0) = 0$) and then, with time, he increases slightly the probability of T until reaching $x_1 = \frac{1}{p}\, T + (1-\frac{1}{p})\, B$. Discretization of this continuous time strategy gives, as in the example when $p = 1$, weak approachability strategies.

Finally, since this game does not satisfy Condition (1), we deduce by Proposition 12 that $\{0\}$ is not uniformly approachable by player 1. Further, $\{0\}$ is not uniformly excludable by player 2 either, which follows directly from the fact that $\{0\}$ is weakly approachable by player 1. Consequently, this Big-Match game of type II is not uniformly determined.

Observe that the condition in Proposition 20 can easily extend to the case where player 2 has many quitting actions (and only one non-quitting action). We just need $\int_0^t g_R(\xi(s))ds + (1-t)g_{j^*}^\star(\xi(t)) \in \mathcal{C}$ to hold for all t and all quitting actions $j^* \in J^*$. If player 2 has more than one non-quitting action, the continuous time condition becomes more complex and is very related to Vieille's approach: one must prove that player 1 can weakly approach the set $\mathcal{C}$ if he can guarantee zero in a auxiliarry (and non classical) zero-sum differential game Γ played between initial time 0 and terminal

time 1 in which player 1 chooses a trajectory $x(t)$, and player 2 chooses a trajectory $y(t)$ (supported on non-quitting actions), chooses a quitting time $t \in [0,1]$ and a quitting distribution y^* (supported on quitting-actions). The payoff of player 1 is the distance of $\int_0^t g(x(s),y(s))ds + (1-t)g^\star(x(t),y^*)$ to the set $\mathcal{C}$. To prove that the game is determined, one should prove that Γ has a value. Those are still open problems.

B On uniform approachability in Big-Match games of type I

First we revisit Example 15.

	L	R
T^*	1	0
B	0	-1

In the next proposition we now show that $\{0\}$ is not uniformly approachable.

Proposition 21 *In the above game, for every strategy σ for player 1, there is a strategy τ for player 2 such that*

$$\liminf_{T\to\infty} \left|\mathbb{E}_{\sigma,\tau}\frac{1}{T}\sum_{t=1}^{T} g(i_t,j_t)\right| \geq \frac{1}{10},$$

thus $\{0\}$ is not uniformly approachable.

Proof. Take an arbitrary strategy σ for player 1. Let τ be the stationary strategy for player 2 which uses the mixed action $(\frac{1}{2},\frac{1}{2})$ at every period. Denote by $q^\star$ the probability, with respect to σ and τ, that play absorbs. If

$$\limsup_{T\to\infty} \mathbb{E}_{\sigma,\tau}\frac{1}{T}\sum_{t=1}^{T} g(i_t,j_t) < -\tfrac{1}{10}$$

then we are done. So assume that $\limsup_{T\to\infty} \mathbb{E}_{\sigma,\tau}\frac{1}{T}\sum_{t=1}^{T} g(i_t,j_t) \geq -\frac{1}{10}$. Since

$$\lim_{T\to\infty} \mathbb{E}_{\sigma,\tau}\frac{1}{T}\sum_{t=1}^{T} g(i_t,j_t) = \tfrac{1}{2}q^\star - \tfrac{1}{2}(1-q^\star) = q^\star - \tfrac{1}{2},$$

we have

$$q^\star \geq -\tfrac{1}{10} + \tfrac{1}{2} = \tfrac{4}{10}.$$

Now let $n \in \mathbb{N}$ be so large that the probability q_n that play absorbs with respect to σ and τ before period n is at least $\frac{3}{10}$. Denote by τ' the Markov strategy for player 2 which uses the mixed action $(\frac{1}{2},\frac{1}{2})$ at all periods before period n and chooses action L from period n onwards. Then

$$\limsup_{T\to\infty} \mathbb{E}_{\sigma,\tau'}\frac{1}{T}\sum_{t=1}^{T} g(i_t,j_t) \geq \tfrac{1}{2}q_n \geq \tfrac{3}{20} > \tfrac{1}{10},$$

so the proof is complete. ■

The Big-Match game of type I above illustrates the complexity of the sufficient condition needed to guarantee uniform approachability. Indeed, let us consider the point of view of player 2. To prove that $\{0\}$ is excludable, we must construct a strategy such that the average payoff is asymptotically ε away from 0, against all strategy of player 1. Let us denote by y_1 the probability of playing L at the first stage. Obviously, y_1 must be bigger than ε, otherwise playing T with probability 1 ensures that the asymptotic average payoff is $y_1 \leq \varepsilon$. If we denote by x_1 the probability of playing T at the first stage, then in order to approach $[-1, -\varepsilon] \cup [\varepsilon, 1]$, player 2 must be able to approach all the possible sets $[-1, -\frac{\varepsilon + x_1 y_1}{1-x_1}] \cup [\frac{\varepsilon - x_1 y_1}{1-x_1}, 1]$.

As a consequence, uniformly approachable set should be defined recursively: player 2 can approach a set if she can approach a certain family of sets, etc, etc. This idea has been observed and explored by Sorin [34] on 2×2 Big-Match games of type I.

C On almost-sure approachability in Big-Match games

We consider in this section almost sure approachability. A closed and convex set $\mathcal{C} \subset \mathbb{R}^d$ is uniformly almost surely approachable if the following condition holds

$$\forall \varepsilon > 0, \exists \sigma, \exists T_\varepsilon \in \mathbb{N}, \forall \tau, \mathbb{P}_{\sigma,\tau}\Big\{\exists\, T \geq T_\varepsilon,\ d_{\mathcal{C}}\Big(\frac{1}{T}\sum_{t=1}^{T} g(i_t, j_t)\Big) \geq \varepsilon\Big\} \leq \varepsilon.$$

Similarly, $\mathcal{C} \subset \mathbb{R}^d$ is weakly almost surely approachable if

$$\forall \varepsilon > 0, \exists T_\varepsilon \in \mathbb{N}, \forall T \geq T_\varepsilon, \exists \sigma_T, \forall \tau, \mathbb{P}_{\sigma_T,\tau}\Big\{d_{\mathcal{C}}\Big(\frac{1}{T}\sum_{t=1}^{T} g(i_t, j_t)\Big) \geq \varepsilon\Big\} \leq \varepsilon.$$

The condition for uniform almost surely approachability can be quite easily obtained.

Proposition 22 *Let $\mathcal{C} \subset \mathbb{R}^d$ be a closed and convex set, then in*

BM games of type I: *$\mathcal{C}$ is uniformly almost surely approachable if and only if there exists $i^\star \in \mathcal{I}^\star$ such that $g(i^*, j) \in \mathcal{C}$ for all $j \in \mathcal{J}$ or*

$$\forall y \in \Delta(\mathcal{J}), \exists x \in \Delta(\mathcal{I}), g(x, y) \in \mathcal{C}$$

BM games of type II: *Let $\mathcal{I}_{\mathcal{C}} := \Big\{i \in \mathcal{I}; g(i, j^\star) \in \mathcal{C},\ \forall j^\star \in \mathcal{J}^\star\Big\}$, then $\mathcal{C}$ is uniformly almost surely approachable if and only if*

$$\forall y \in \Delta(\mathcal{J}), \exists x \in \Delta(\mathcal{I}_{\mathcal{C}}), g(x, y) \in \mathcal{C}$$

Generalized quitting games: *If $\mathcal{I}_{\mathcal{C}} = \emptyset$ then $\mathcal{C}$ is uniformly almost surely approachable if and only if there exists $i^\star \in \mathcal{I}^\star$ such that $g(i^\star, j) \in \mathcal{C}$ for all $j \in \mathbf{J}$.*

If $\mathcal{I}_{\mathcal{C}} \neq \emptyset$, then $\mathcal{C}$ is uniformly almost surely approachable if and only if either there exists $i^\star \in \mathcal{I}^\star$ such that $g(i^\star, j) \in \mathcal{C}$ for all $j \in \mathcal{J}$ or

$$\forall y \in \Delta(\mathcal{J}), \exists x \in \Delta(\mathcal{I}_{\mathcal{C}}), g(x, y) \in \mathcal{C}$$

Proof. We consider each case independently.

BM games of type I: The fact that the condition is sufficient is immediate: in the first case, player 1 just has to play $i^\star$ at the first stage; in the second case, he just needs to follow the classical Blackwell strategy.

To prove that the condition is necessary, assume that it does not hold. Using Blackwell [5] results (see also Perchet [26]), this implies that there exists $y \in \Delta(\mathcal{J})$ with full support such that $g(y) := \{g(x,y) \ ; \ x \in \Delta(\mathcal{I})\}$ is δ-away from $\mathcal{C}$. We denote by $\eta > 0$ the smallest probability put on a pure action by y. Then Player 2 just has to play i.i.d. this mixed action. On the event where the game is not absorbed, then the average payoff converges to the set $g(y)$ thus is δ-away from $\mathcal{C}$. As a consequence, to approach $\mathcal{C}$, player 1 must enforce absorption with probability $1-\varepsilon$. But since for every $i^\star \in \mathcal{I}^\star$ there exists $j \in \mathcal{J}$ such that $g(i^\star, j) \notin \mathcal{C}$, this implies that with probability at least $(1-\varepsilon)\eta > \eta/2 > 0$ the payoff does not belong to $\mathcal{C}$.

BM games of type II: In this case again, the fact that the condition is sufficient is immediate. On the contrary, assume that it does not hold, then there exists $y \in \Delta(\mathcal{J})$ with full support (and minimal weight bigger than $\eta > 0$) such that $g^{\mathcal{C}}(y) := \{g(x,y) \ ; \ x \in \Delta(\mathcal{I}_{\mathcal{C}})\}$ is δ-away from $\mathcal{C}$. As a consequence, in order to approach $\mathcal{C}$, player 1 must use actions that do not belong to $\mathcal{I}_{\mathcal{C}}$ with total probability at least $\delta - \varepsilon$. In particular, it must exist one stage where player 1 uses an action $i^\star \notin \mathcal{I}_{\mathcal{C}}$ with probability at least $\delta - \varepsilon$; at that stage, player 2 just needs to play the associated action $j^\star \in \mathcal{J}^\star$ (so that $g(i^\star, j^\star) \notin \mathcal{C}$) to prevent almost sure uniform approachability.

Generalized quitting games: In that case, it is actually the sufficient condition which is tricky to prove if $\mathcal{I}_{\mathcal{C}} \neq \emptyset$ and there exists $i^\star \in \mathcal{I}^\star$ such that $g(i^\star, j) \in \mathcal{C}$ for all $j \in \mathcal{J}$ (the other cases are immediate). In that case, player 1 just needs to play i.i.d. any action in $\mathcal{I}_{\mathcal{C}}$ with probability $(1-\varepsilon)$ and the strategy $i^\star \in \mathcal{I}^\star$ with probability ε. If player 2 uses at any stage $j^\star \in \mathcal{J}^\star$, then the game is absorbed in $\mathcal{C}$ with probability at least $1-\varepsilon$; on the other hand, if he only uses actions in $\mathcal{J}$, then the game is absorbed in $\mathcal{C}$ with probability 1.

To prove that the condition is necessary, one just needs to combine the two above arguments.

■

We conclude this section by mentioning that those techniques of proof do not directly provide a necessary and sufficient condition for the weak almost sure approachability case, except in Big-Match of type I.

Indeed, consider a Big-Match game of type II and assume that the game is absorbed at some stage $t^\star$, and that the absorbing payoff lies outside $\mathcal{C}$. Obviously, this implies that the asymptotic average payoff (as t goes to infinity) lies outside $\mathcal{C}$, thus uniform approachability fails. On the other hand, this unfortunately does not imply that the average payoff at stage $2t^\star$ is outside $\mathcal{C}$.