Uniformly Supported Approximate Equilibria in Families of Games

Yehuda John Levy[*][†]

September 8, 2021

Abstract

This paper considers uniformly bounded classes of non-zero-sum strategic-form games with large finite or compact action spaces. The central class of games considered is assumed to be defined via a semi-algebraic condition. We show that for each $\varepsilon > 0$, the support size required for ε-equilibrium can be taken to be uniform over the entire class. As a corollary, the value of zero-sum games, as a function of a single-variable, is well-behaved in the limit. More generally, the result only requires that the collection of payoff functions considered, as functions of other players actions, have finite pseudo-dimension.

Keywords: Small-Support Equilibrium; Semi-Algebraic Classes; Vapnik-Chervonenkis Dimension
JEL Classifications: C72, C65, C13

1 Introduction

An important strand of literature in game theory is dedicated to computing how many strategies players must randomize over if they wish to be playing, approximately, an equilibrium. The primary technique was laid out in the the seminal paper [1] which studied, among another class of linear problems, small-support ε-optimal strategies in two-player zero-sum games. That paper demonstrates that in an $m \times m$ matrix game (payoffs in $[0, 1]$), the players possess ε-optimal strategies with support size $O(\frac{1}{\varepsilon^2} \ln(m))$, and in fact it is demonstrated that this required growth rate is tight. For each player, an approximately optimal strategy can be found, with high probability, by random sampling w.r.t. the distribution induced by an optimal strategy. Indeed, standard law of large number arguments show that such a random sampling of an optimal strategy would converge; the specific bounds are then obtained by bounds on the error, specifically Chernoff's inequality.

[*]Adam Smith Business School, University of Glasgow, Glasgow, G12 8QQ, UK.
[†]The author is very grateful to Yacov Babichenko for many useful discussions, and to Ilan Nehama and Miquel Oliu-Barton for helpful feedback.

This result was extended in [18] to ε-equilibria of any game with arbitrary (but fixed) number of players, and then again extended by [4] to games with variable number of players; specifically, to n-player games, m actions each, by showing that ε-equilibrium exists in which each player uses strategies with support $O(\frac{1}{\varepsilon^2}(\ln(m) + \ln(n) - \ln(\varepsilon)))$; see also [6], which improves the bound for bimatrix games which are sparse (i.e., relatively few profiles give non-zero payoffs). Both papers discuss the implications for computation of approximate Nash equilibria in quasi-polynomial time. Again, in both cases, the existence is proven by showing that random sampling w.r.t. an equilibrium distribution will again give, with high probability, an ε-equilibrium strategy profile. Although not the topic of this paper, it is interesting to note that is unknown if the dependence on the number of players in the later result is tight, or for that matter, if dependence on the number of players is needed at all. This open question is discussed further in [2].

One can naturally ask on the support size required if we allow the players to possess a continuum of strategies, say, the unit interval. In general one could not impose any bounds; indeed, any game with finitely many actions could be embedded into one in which players have a continuum of actions, even in such a way that gives continuous utility functions, while preserving (in an approriate sense) the (exact and approximate) equilibria; this point is elaborated on in Section 6. The question then is, for a class of games and given $\varepsilon > 0$, can the size of support required for the existence of ε-equilibrium be bounded by some function of $\varepsilon > 0$ and of the parameters or properties of the class?

A natural class of payoff functions to consider, a class which displays significant regularity properties, are the *semi-algebraic functions*. Semi-algebraic geometry has an intimate relationship with game theory, see, e.g., [22] or [20, Ch. 6]. Semi-algebraic sets are those subsets of Euclidean spaces defined using logical formulas whose atoms are polynomial equalities and inequalities; semi-algebraic functions are those with semi-algebraic graphs. The set of Nash equilibria of a game with finitely many actions is naturally a semi-algebraic set, as is the manifold of Nash equilibria (see [15]) and many other closely related sets, mappings, and correspondences. The regularity and decomposition properties of semi-algebraic sets can be used, as in the above papers, to derive various useful results concerning equilibria and how equilibria change as the game does, and in turn, Nash equilibria and the Nash equilibrium correspondence have been shown to be rich, universal is certain senses, in the appropriate classes of semi-algebraic objects; see, e.g., [17], [27], [5].

In this paper, one of our central goals is to study classes of games with semi-algebraic payoffs on semi-algebraic action spaces $X_1, \ldots, X_N$. The complexity (*class*, in this paper, following [7]) of a semi-algebraic payoff function on $X_1 \times \cdots \times X_N$ can be defined in terms of the number of polynomials s of degree at most r needed to define its graph. A particular case of such uniformly bounded class over a collection of games arises when the payoffs are defined by a function which depends, in addition to the actions, on a parameter from a semi-algebraic set Λ. Our paper's main theorems, which will be stated in greater precision later, can be summarized as follows:

Theorem 1.1. *Let $X^1, \ldots, X^N$ be compact and semi-algebraic sets. For each $\varepsilon > 0$, and $s, r \in \mathbb{N}$, there exists $m \in \mathbb{N}$, s.t. for each semi-algebraic game with payoffs in $[0,1]$ of class (s,r), there is an ε-equilibrium of support at most m. Furtheremore, if Λ is a semi-algebraic set, $\phi : \Lambda \times X^1 \times \cdots \times X^N \to [0,1]^N$ is semi-algebraic, there exists $m \in \mathbb{N}$ and a semi-algebraic mapping from Λ assigning to $\phi(\lambda, \cdot)$ an ε-equilibrium with support of size at most m for each $\lambda \in \Lambda$.*

We note, for the latter part of the theorem, that if Λ were compact as well, the result would following easily; however, this is not assumed. In particular, we deduce as a corollary that if $N = 2$ and the games are zero-sum, and Λ is an interval, the limits of the value at the end-points of the interval exist. This result is shown despite the fact that the value, it turns out, need not be a semi-algebraic function of the parameter; see [11]. We remark also that our proof relies crucially on the boundedness of ϕ, i.e., on the payoffs being uniformly bounded. If the corollary on zero-sum games were to be established without the boundedness assumption, where the limit would be required to exist in the extended real line $\mathbb{R} \cup \{-\infty, \infty\}$, it would imply the existence of the asymptotic value (that is, convergence of the discounted values) of stochastic games with finitely many states, compact action spaces, and semi-algebraic transitions, due to a result of Attia & Oliu-Barton, [3]. Indeed, that paper defines, from a stochastic game, a family of auxiliary strategic-form games $W^k_\lambda(z)$ (one for each state k) which depend on a parameter $\lambda \in (0,1]$ and on a real parameter $z \in \mathbb{R}$, and show that not only can the λ-discounted value of state k be characterized the unique solution z of $val(W^k_\lambda(z)) = 0$, but that if for each z, $val(W^k_\lambda(z))$ converges quickly enough as $\lambda \to 0$, then the limit of the discounted values exist. We elaborate in Section 7. Hopefully the techniques presented here can be extended to address this long-standing open question.

The primary tool we use is that of Vapnik-Chervonenkis (VC) dimension and its applications to uniform convergence of random sampling, as introduced in [24]. These results show that if a class of sets has a certain bounded complexity, formalized in terms of having a bounded VC dimension, then computing the measure of sets via random sampling converges in probabiltiy at a uniform rate over all sets in the class and over all probability measures. The results on uniform convergence for the VC classes (those classes of finite VC dimension) can replace, with a considerable amount of care, the use of laws on large deviations (in particular, Hoeffding's inequality) in the proofs of [1] and [4]. Indeed, we will show the assumption of the class of games being semi-algebraic of a specific class can be replaced with the weaker assumption of the collection of payoff functions considered, as a function of other players actions, having finite VC dimension.

The paper is outlined as followed. Section 2 presents background notions and results on games, semi-algebraic geometry, and VC dimension. Section 3 develops these concepts to auxilliary results that we will use, many of which may be of independant interest. Section 4 presents the results. Section 5 extends the discussion to classes of games defined in an o-minimal structure, a generalization of semi-algebraic structures. Section 6 has some discussion on the bounds derived, and presents a technique of going from games with finite actions to continuous actions. Section 7 discusses potential connections to stochastic games. Appendix A presents briefly a slightly different approach.

2 Background

2.1 Games and Equilibria

A *(strategic-form) game* consists of:

- Some finite number N of players.
- For each Player $i \in [N] = \{1, \ldots, N\}$, a set X^i of actions.
- For each Player $i \in [N] = \{1, \ldots, N\}$, a payoff function $u_i : \prod_{i \in [N]} X^i \to \mathbb{R}$.

For $\varepsilon \geq 0$, an *ε-equilibrium (in pure strategies)* of a game is a profile $x = (x_1, \ldots, x_N)$, with $x_i \in X^i$ for $i = 1, \ldots, N$, such that for each $i = 1, \ldots, N$ and each $y_i \in X^i$,

$$u_i(y_i, x_{-i}) \leq u_i(x) + \varepsilon$$

where $x_{-i} = (x_j)_{j \neq i}$. A 0-equilibrium is termed simply an *equilibrium*.

Assuming each X^i has a σ-algebra on it (if X^i is, say, Polish, this will be assumed to be the Borel σ-algebra), let $\Delta(X^i)$ denote the space of probability distributions on X^i, which are referred to as *mixed actions*. The *mixed extension* of a game is the game in which the set of actions X^i is replaced by $\Delta(X^i)$ for each agent, and the payoff function u_i is replaced by the payoff function (denoted by the same letter, by mildly abusive notation),

$$u_i(\sigma_1, \ldots, \sigma_N) = \int_{\prod_j X^j} u_i(\cdot) d(\sigma_1 \times \cdots \sigma_N)$$

If, e.g., each u_i is measurable and bounded, the payoffs in the mixed extension are well-defined.

An *equilibrium (resp. ε-equilibrium) in mixed strategies* - or, for the purposes of this paper, simply an *equilibrium (resp. ε-equilibrium)* - is an equilibrium (resp. ε-equilibrium) in pure strategies of the mixed extension.

A mixed strategy $s_i \in \Delta(X^i)$ of some player i is said to be *m-uniform*, where $m \in \mathbb{N}$, if there exist[1] $x_i^1, \ldots, x_i^m$ s.t. $s_i = \frac{1}{m}\sum_{j=1}^m \delta_{x_i^j}$, where δ_z denote the Dirac measure at z. Equivalently, s_i is finitely supported and $s_i(x) \in \frac{1}{m}\mathbb{Z}$ for each $x \in X^i$.

2.2 Semi-Algebraic Sets and Functions

Let $\mathbb{R}[x_1, \ldots, x_N]$ denote the ring[2] of polynomials in N variables, $x_1, \ldots, x_N$. Recall that a *Boolean algebra* of subsets of a space is a collection that is closed under finite unions, finite intersections, and complements. The semi-algebraic subsets of $\mathbb{R}^N$ are those sets in the smallest Boolean algebra containing all sets defined by polynomial equalities or inequalities, i.e., all sets of the form $\{x \in \mathbb{R}^N \mid P(x) * 0\}$, where $*$ can be equality or weak/strong inequality, and $P \in \mathbb{R}[x_1, \ldots, x_N]$.

Equivalently (e.g., [10, Ch. 2]), by the Tarsksi-Seidenberg theorem, semi-algebraic sets are those that can be expressed as a formula in first-order logic whose atoms are polynomial equalities or inequalities. In particular, we mention the Tarski-Seidenberg theorem:

Theorem 2.1. *Let $A \subseteq \mathbb{R}^N$ be semi-algebraic, let $\pi_K : \mathbb{R}^N \to \mathbb{R}^K$ denote the projection to a subset $K \subseteq \{1, \ldots, N\}$ of coordinates. Then $\pi_K(A)$ is semi-algebraic.*

A *semi-algebraic function* $f : A \to \mathbb{R}^K$, where $A \subseteq \mathbb{R}^N$, is one whose graph $Gr(f) := \{(x, y) \in A \times \mathbb{R}^K \mid y = f(x)\}$ is semi-algebraic: It follows from Theorem 2.1 that the domain A is semi-algebraic, and that the image / inverse image of a semi-algebraic set under a semi-algebraic function is also semi-algebraic; it also follows that the composition of semi-algebraic functions is semi-algebraic.

The following is the Monotonicity Theorem for semi-algebraic functions, e.g., [12, Ch. 3].

Theorem 2.2. *Let $I \subseteq \mathbb{R}$ be connected, and let $f : I \to \mathbb{R}$ be semi-algebraic. Then there is $n \in \mathbb{N}$ and $a_1 = \inf I < a_2 < \cdots < a_{n-1} < a_n = \sup I$ s.t. f is (weakly) monotonic and continuous in each sub-interval (a_i, a_{i+1}).*

The following is the Semi-Algebraic Selection Theorem, proved in e.g., [20, Thm. 4, Ch. 6]:

Theorem 2.3. *Let $X \subseteq \mathbb{R}^n \times \mathbb{R}^m$ be a semialgebraic set, and let π be the natural projection of $\mathbb{R}^n \times \mathbb{R}^m$ to $\mathbb{R}^n$. Then there is a semialgebraic function $f : \pi(X) \to \mathbb{R}^m$ s.t. $Gr(f) \subseteq X$.*

[1] Note that the $x_i^1, \ldots, x_i^m$ need not be distinct.

[2] A ring is an algebraic structure with operations of addition and multiplication satisfying certain axioms; we will not need to make use of the specific axioms, which can be found in any introductory text on abstract algebra.

2.3 (s, r) Class in $\mathbb{R}^N$

The following terminology follows (although differs slightly from[3]) [7]:

Definition 2.1. *A set $A \subseteq \mathbb{R}^N$ is said to be of class (s, r) if it can be defined via Boolean formula containing s distinct atomic predicates, each of which is a polynomial equality or inequality in N variables of degree at most r.*[4] *We will also say that A is of class (s, r) in $B \subseteq \mathbb{R}^N$ if there is such a Boolean formula ϕ s.t. $A = \{x \in B \mid \phi(x)\}$.*

2.4 The Vapnik-Chervonenkis Dimension

Definition 2.2. *Let X be a set, and let $\mathcal{C}$ be a collection of subsets of X. A set $S = \{x_1, \ldots, x_n\} \subseteq X$ is said to be shattered by $\mathcal{C}$ if for each $A \subseteq S$, there is $C \in \mathcal{C}$ s.t. $C \cap S = A$.*

Definition 2.3. *The Vapnik-Chervonenkis (VC) dimension of $\mathcal{C}$ as in the previous definition, denoted $VCdim(\mathcal{C})$, is the largest integer $n \in \mathbb{N}$ s.t. there exists a set $S \subseteq X$ of cardinality n which is shattered by $\mathcal{C}$. (If there is no such largest $n \in \mathbb{N}$, $VCdim(\mathcal{C}) = \infty$.)*

The following result is [14, Thm 2.2]:

Theorem 2.4. *Let $\overline{C} \subseteq \mathbb{R}^{k+n}$ be of class (s, r). Let $\mathcal{C}$ be the class of sets in $\mathbb{R}^n$ defined by*

$$\mathcal{C} = \left\{ \overline{C}_y \mid y \in \mathbb{R}^k \right\}, \text{ where } \overline{C}_y = \{x \in \mathbb{R}^n \mid (y, x) \in \overline{C}\} \tag{2.1}$$

Then $VCdim(\mathcal{C}) \leq 2k \cdot \log_2(8ers)$.

In Theorem 2.4, the bound is demonstrated to be tight up to a logarithmic factor.

2.5 The Pseudo-Dimension

Recall that 1_B is the indicator function of the set B.

Definition 2.4. *Let X be a measurable space, and let $\mathcal{F} \subseteq [0, 1]^X$ be a collection of functions. A set $S = \{x_1, \ldots, x_n\} \subseteq X$ is said to be P-shattered by $\mathcal{F}$ if there is $c \in [0, 1]^n$ s.t. for every $A \subseteq S$, there exists $f_A \in \mathcal{F}$ s.t. for $i = 1, \ldots, n$, $f_A(x_i) \geq c_i$ iff $x_i \in A$, i.e., $1_{[0,\infty)}(f_A(x_i) - c_i) = 1_A(x_i)$.*

The following concept origionated in [21]:

[3] That paper requires the atomic formulae to be strict polynomial inequalities, $P(x) > 0$, while we allow weak or strong inequalities.

[4] The *degree* of a polynomial here refers to *total degree*, that is, the maximal total degree of any of its monomials, where the total degree of a monomial is the sum of its exponents. E.g., the degree of $x^2y^2z^3$ is 7.

Definition 2.5. *For a collection $\mathcal{F} \subseteq [0,1]^X$ of functions, the pseudo-dimension (a.k.a. P-dimension, or Pollard dimension) of $\mathcal{F}$, denoted $P - dim(\mathcal{F})$, is the largest integer $n \in \mathbb{N}$ s.t. there exists a set $S \subseteq X$ of cardinality n which is P-shattered by $\mathcal{F}$. (If there is no such largest $n \in \mathbb{N}$, $P - dim(\mathcal{C}) = \infty$.)*

Lemma 4.1 of [25, Ch. 4] states:

Lemma 2.6. *For a collection $\mathcal{F} \subseteq [0,1]^X$ of functions, define a family $\mathcal{C}$ of subsets of $X \times [0,1]$ given by*

$$\mathcal{C} = \Big\{\{x \in X, c \in [0,1] \mid f(x) - c \geq 0\}\Big| f \in \mathcal{F}\Big\}$$

Then $P - dim(\mathcal{F}) = VCdim(\mathcal{C})$.

2.6 Uniform Convergence of Means

The following result can be found in, e.g., [25, Ch. 7, Thm 7.1]. Let X be a measurable space, $\Delta(X)$ the space of probability measures on X, and for each $P \in \Delta(X)$ and $m \in \mathbb{N}$, let P^m be the product measure $P \times \cdots \times P$ on X^m. For each family $\mathcal{F}$ of bounded measurable functions $X \to \mathbb{R}$, denote[5] for each $P \in \Delta(X)$, $m \in \mathbb{N}$ and $\varepsilon > 0$,

$$q(m, \varepsilon, \mathcal{F}, P) = P^m\Big((x^j)_{j=1}^m \in X^m \Big| \sup_{f \in \mathcal{F}} \Big|\frac{1}{m}\sum_{j=1}^m f(x^j) - \int_X f dP\Big| > \varepsilon\Big) \quad (2.2)$$

and by $q(m, \varepsilon, \mathcal{F}) = \sup_{P \in \Delta(X)} q(m, \varepsilon, \mathcal{F}, P)$.

Theorem 2.5. *If $\mathcal{F}$ is a family of measurable functions $X \to [0,1]$ of pseudo-dimension d, then for all $m \in \mathbb{N}$, and any $0 < \varepsilon < \frac{e \ln(2)}{4} (\approx 0.47)$,*

$$q(m, \varepsilon, \mathcal{F}) \leq 8\Big(\frac{16e}{\varepsilon} \ln\big(\frac{16e}{\varepsilon}\big)\Big)^d \cdot \exp\big(-\frac{m\varepsilon^2}{32}\big)$$

In particular, note that as d is smaller, $q(m, \varepsilon, \mathcal{F})$ is smaller, i.e. the convergence is quicker. This reflects the fact that a large value of d means the set of functions $\mathcal{F}$ is 'complex' in some sense, and a larger sampling is required to uniformly approximate (in mean) the family well.

3 Auxilliary Results

3.1 Miscellaneous Results

Lemma 3.1. *Let $\phi : (0, \delta) \to \mathbb{R}$ be a bounded function which can be uniformly approximated by semi-algebraic functions. Then $\lim_{\lambda \to 0} \phi(\lambda)$ exists.*

[5] To be precise, in general we should define (2.2) using the P^m-outer measure, as the event we are taking the probably of need not be P^m-measurable. Every textbook the author has seen on the subject has glossed over this point, which is easily overcome.

Proof. Suppose (ϕ_n) is a sequence of semi-algebraic functions $(0,\delta) \to \mathbb{R}$ converging to ϕ uniformly. By Theorem 2.2, $v_n := \lim_{\lambda\to 0} \phi_n(\lambda)$ exists for each n. $(v_n)_{n=1}$ is bounded; w.l.o.g., $v := \lim_{n\to\infty} v_n$ exists. We contend $v = \lim_{\lambda\to 0} \phi(\lambda)$. Let $\varepsilon > 0$, let n be such that $|\phi_n - \phi| < \frac{\varepsilon}{3}$ uniformly and $|v_n - v| < \frac{\varepsilon}{3}$, and let $\eta > 0$ be such that $|\phi_n(\lambda) - v_n| < \frac{\varepsilon}{3}$ for $\lambda \in (0,\eta)$. Hence for $\lambda \in (0,\eta)$, $|\phi(\lambda) - v| < \varepsilon$. ◻

Proposition 3.2. *Let $\Lambda \subseteq \mathbb{R}^k$ and $X \subseteq \mathbb{R}^n$ be semi-algebraic, and let $\phi : \Lambda \times X \to [0,1]$ be such that the graph of ϕ, $Gr(\phi) = \{(y,x,z) \in \Lambda \times X \times [0,1] \mid z = \phi(y,x)\}$, is of class (s,r), $r \geq 1$, in $\Lambda \times X \times \mathbb{R}$. Then the collection of functions $X \to [0,1]$ given by*

$$\mathcal{F} = \left\{\phi_y \mid y \in \mathbb{R}^k\right\},\ \textit{where } \phi_y(x) = \phi(y,x)$$

satisfies $P - dim(\mathcal{F}) \leq 2(k+1) \cdot \log_2(8er(s+1))$.

Proof. Let ψ be the formula in $y_1,\dots,y_k,x_1,\dots,x_n,z$ which witnesses that the graph $Gr(\phi)$ is of class (r,s) in $\Lambda \times X \times \mathbb{R}$. Define $\overline{D} \subseteq \mathbb{R}^{k+n+2}$

$$\overline{D} = \{(y,x,z,c) \in \Lambda \times X \times \mathbb{R} \times \mathbb{R} \mid \psi(y,x,z) \wedge (z - c \geq 0)\}$$

and

$$\begin{aligned} \mathcal{D} &= \{\overline{D}_{y,z} \mid y \in \Lambda, z \in \mathbb{R}\} \\ &= \big\{\{(x,c) \in X \times \mathbb{R} \mid \phi(y,x) - c \geq 0\} \mid y \in \Lambda\big\} \end{aligned} \tag{3.1}$$

where we recall the notation of (2.1), and define similarly $\overline{C} \subseteq \mathbb{R}^{k+n+2}$

$$\overline{C} = \{(y,x,z,c) \in \Lambda \times X \times \mathbb{R} \times [0,1] \mid \psi(y,x,z) \wedge (z - c \geq 0)\} \tag{3.2}$$

and

$$\begin{aligned} \mathcal{C} &= \{\overline{C}_{y,z} \mid y \in \Lambda, z \in \mathbb{R}\} \\ &= \big\{\{(x,c) \in X \times [0,1] \mid \phi(y,x) - c \geq 0\} \mid y \in \Lambda\big\} \end{aligned} \tag{3.3}$$

By Theorem 2.4, since $\Lambda \times \mathbb{R} \subseteq \mathbb{R}^{k+1}$ and $\overline{D}$ is of class $(s+1,r)$,

$$VCdim(\mathcal{D}) \leq 2(k+1)\log_2(8er(s+1)). \tag{3.4}$$

We claim[6] $VCdim(\mathcal{C}) \leq VCdim(\mathcal{D})$. Indeed, denoting $V = X \times \mathbb{R} \times [0,1]$, we have since $0 \leq \phi \leq 1$,

$$\mathcal{C} = \mathcal{D} \cap V := \{D \cap V \mid D \in \mathcal{D}\}$$

so any set shattered by $\mathcal{C}$ is also shattered by $\mathcal{D}$.

(3.3) is equivalent to:

$$\mathcal{C} = \big\{\{(x,c) \in X \times [0,1] \mid f(x) - c \geq 0\} \mid f \in \mathcal{F}\big\}$$

so $VCdim(\mathcal{C}) = P - dim(\mathcal{F})$ by Lemma 2.6. Combining this with (3.4) and the fact that $VCdim(\mathcal{C}) \leq VCdim(\mathcal{D})$ gives the desired result.

◻

[6]In fact, they are equal, but we do not need this stronger claim.

3.2 Uniform Convergence on Product Spaces

We need a generalization of Theorem 2.5, as the sampling is done via a product distribution, and we want our sampling approximation to be a product distribution as well. Fix $N \in \mathbb{N}$. Suppose X can be decomposed $X = X^1 \times \cdots \times X^N$, and we consider N classes of functions, $\mathcal{F}_1, \ldots, \mathcal{F}_N$, each consisting of functions $X^1 \times \cdots \times X^N \to [0,1]$. For each $i = 1, \ldots, N$, we modify the class $\mathcal{F}_i$ to the class $\mathcal{F}'_i$ of functions $\prod_{j \neq i} X^j \to [0,1]$, henceforth called the *induced classes*, given by

$$\mathcal{F}'_i = \{\phi(x_i, \cdot) \mid \phi \in \mathcal{F}_i,\ x_i \in X^i\} \tag{3.5}$$

i.e., all those functions derived from functions in $\mathcal{F}_i$ by specifying an arbitrary input from X^i. We use the notation that for $m \in \mathbb{N}$, $[m] := \{1, \ldots, m\}$. For each N-tuple of measures $P_i \in \Delta(X^i)$ for $i = 1, \ldots, N$, denote $P = \prod_{j=1}^N P_j$, $P^m = P \times \cdots \times P$, $P_{-i} = \prod_{j \neq i} P_j$, and, in the spirit of (2.2),

$$\tilde{q}(m,\varepsilon,(\mathcal{F}_i),(P_i)) = P^m\Big((x^j)_{j=1}^m \in X^m \Big| \max_{i=1,\ldots,N} \sup_{f \in \mathcal{F}_i} \sup_{x_i \in X^i} \Big|\frac{1}{m^{N-1}} \sum_{j_* \in [m]^{N-1}} f(x_i, x_{-i}^{j_*}) - \int_X f(x_i, \cdot) dP_{-i}\Big| > \varepsilon\Big)$$

where, for each $i = 1, \ldots, N$, in the relevant sum, we view j_* as an element of $\prod_{k \neq i}[m] \sim [m]^{N-1}$ and $x_{-i}^{j_*}$ is the $N-1$-tuple which specifies action $x_k^{j_*(k)}$ for player $k \neq i$. Further denote,

$$\tilde{q}(m,\varepsilon,(\mathcal{F}_i)) = \sup_{(P_i)_{i=1}^N \in \prod_{i=1}^N \Delta(X^i)} \tilde{q}(m,\varepsilon,(\mathcal{F}_i),(P_i))$$

The following proposition is likely the most significant piece of technical machinery developed in this paper:

Proposition 3.3. *If the induced classes $\mathcal{F}'_1,\ldots,\mathcal{F}'_N$ defined by (3.5) are each of pseudo-dimension at most d, then for all $m \in \mathbb{N}$, and any $0 < \varepsilon < \frac{e\ln(2)}{8}$,*

$$\tilde{q}(m,\varepsilon,(\mathcal{F}_i)) \leq \frac{16N}{\varepsilon}\Big(\frac{32e}{\varepsilon}\ln(\frac{32e}{\varepsilon})\Big)^d \cdot exp(-\frac{m\varepsilon^2}{128}) \tag{3.6}$$

The proof technique is similar to that of [4, Lem. 3.4], except that instead of appealing to Hoeffdings inequality, we appeal to Theorem 2.5, which results in additional care needed throughout; unlike [4], we also account for each of the classes $\mathcal{F}_1, \ldots, \mathcal{F}_N$ separately.

Proof. Fix probability measures $P_1 \in \Delta(X^1), \ldots, P_N \in \Delta(X^N)$, and $m \in \mathbb{N}$. Fix $0 < \varepsilon < \frac{e\ln(2)}{8}$. Let $z = (z(k))_{k=1}^m$, where $z(k) = (z_1(k), \ldots, z_N(k))$, be the random variable in $(X^1 \times \cdots \times X^N)^m$ which distributes by P^m. Fix a player, the first w.l.o.g.. For every $j_* = (j_2, \ldots, j_N) \in [m]^{N-1}$ and every $\ell \in [k]$, denote

$z_{-1}(j_*+\ell) = (z_2(j_2+\ell), \dots, z_N(j_N+\ell))$, where arithmetic of indices is modular. Hence, we can write, for each $\ell \in [m]$, each $x_1 \in X^1$ and $u_1 \in \mathcal{F}_1$,

$$\frac{1}{m^{N-1}} \sum_{j_* \in [m]^{N-1}} u_1(x_1, z_{-1}(j_*)) = \frac{1}{m^{N-1}} \sum_{j_* \in [m]^{N-1}} u_1(x_1, z_{-1}(j_* + \ell))$$

and therefore,

$$\frac{1}{m^{N-1}} \sum_{j_* \in [m]^{N-1}} u_1(x_1, z_{-1}(j_*)) = \frac{1}{m^N} \sum_{j_* \in [m]^{N-1}} \sum_{\ell=1}^{m} u_1(x_1, z_{-1}(j_* + \ell)) \quad (3.7)$$

Define for each $j_* \in [m]^{N-1}$ the random variable $d(j_*)$ by

$$d(j_*) = \begin{cases} 1 & \text{if} \sup_{u_1 \in \mathcal{F}_1, x_1 \in X^1} \left| \frac{1}{m} \sum_{\ell=1}^{m} u_1(x_1, z_{-1}(j_* + \ell)) - \int u_1(x_1, \cdot) dP_{-1} \right| > \frac{\varepsilon}{2} \\ 0 & \text{otherwise} \end{cases}$$

Hence, we have

$$d(j_*) + \frac{\varepsilon}{2} \geq \sup_{u_1 \in \mathcal{F}_1, x_1 \in X^1} \left| \frac{1}{m} \sum_{\ell=1}^{m} u_1(x_1, z_{-1}(j_* + \ell)) - \int u_1(x_1, \cdot) dP_{-1} \right|$$

Note also that for each fixed j_*, the random variables $z_{-1}(j_*+1), \dots, z_{-1}(j_*+m)$ are independent and distribute according to P_{-1}, which is just the marginal of P. Hence, by Theorem 2.5 (applied to the class $\{u_1(x_1, \cdot) \mid u_1 \in \mathcal{F}_1, x_1 \in X^1\}$, with $\frac{\varepsilon}{2}$ instead of ε),

$$E[d(j_*)] \leq 8\Big(\frac{32e}{\varepsilon} \ln(\frac{32e}{\varepsilon})\Big)^d \cdot exp(-\frac{m\varepsilon^2}{128})$$

Hence,

$$\begin{aligned}
T_1 &:= P\Big(\sup_{u_1 \in \mathcal{F}_1, x_1 \in X^1} \Big|\frac{1}{m^{N-1}} \sum_{j_* \in [m]^{N-1}} u_1(x_1, z_{-1}(j_*)) - \int u_1(x_1, \cdot) dP_{-1}\Big| \geq \varepsilon\Big) \\
&= P\Big(\sup_{u_1 \in \mathcal{F}_1, x_1 \in X^1} \Big|\frac{1}{m^N} \sum_{j_* \in [m]^{N-1}} \sum_{\ell=1}^{m} u_1(x_1, z_{-1}(j_* + \ell)) - \int u_1(x_1, \cdot) dP_{-1}\Big| \geq \varepsilon\Big) \\
&\leq P\Big(\frac{1}{m^{N-1}} \sum_{j_* \in [m]^{N-1}} \sup_{u_1 \in \mathcal{F}_1, x_1 \in X^1} \Big|\frac{1}{m} \sum_{\ell=1}^{m} u_1(x_1, z_{-1}(j_* + \ell)) - \int u_1(x_1, \cdot) dP_{-1}\Big| \geq \varepsilon\Big) \\
&\leq P\Big(\frac{1}{m^{N-1}} \sum_{j_* \in [m]^{N-1}} d(j_*) \geq \frac{\varepsilon}{2}\Big) \leq \frac{16}{\varepsilon}\Big(\frac{32e}{\varepsilon} \ln(\frac{32e}{\varepsilon})\Big)^d \cdot exp(-\frac{m\varepsilon^2}{128}) \quad (3.8)
\end{aligned}$$

where the equality follows from (3.7), the first inequality results from the triangle inequality, the second is from the definition of $d(j_*)$, and the last follows from Markov's inequality. The bound in (3.6), which differs from the bound in (3.8) by a factor N, now follows by taking the bound over all players, as, denoting $T_2, \dots, T_N$ similarly, we have $\tilde{q}(m, \varepsilon, (\mathcal{F}_i)) \leq \sum_{i=1}^{N} T_i$. $\square$

Hence, it follows immediately:

Proposition 3.4. *If $\mathcal{F}_1, \mathcal{F}_2, \ldots, \mathcal{F}_N$ are families of measurable functions $X^1 \times \cdots \times X^N \to [0,1]$ whose induced classes, defined by (3.5), are of pseudo-dimension each at most d, $0 < \varepsilon < \frac{e\ln(2)}{8}$, and*

$$m > \frac{128}{\varepsilon^2}\Big(\ln(16N) - \ln(\varepsilon) + d \cdot \ln\big(\frac{32e}{\varepsilon}\ln(\frac{32e}{\varepsilon})\big)\Big) \tag{3.9}$$

then for each $P = \prod_{i=1}^N P_i \in \prod_{i=1}^N \Delta(X^i)$, there exist $x_i^1, \ldots, x_i^m \in X^i$ for $i = 1, \ldots, N$, s.t., denoting $P'_i \in \Delta(X)$ defined by $P'_i = \frac{1}{m}\sum_{j=1}^m \delta_{x_i^j}$ and $P'_{-i} = \prod_{j\neq i} P'_j$,

$$|\int_{X_{-i}} f(x_i,\cdot)dP'_{-i} - \int_{X_{-i}} f(x_i,\cdot)dP_{-i}| \leq \varepsilon,\ \forall i = 1,\ldots,N, \forall f \in \mathcal{F}_i, \forall x_i \in X^i \tag{3.10}$$

and furtheremore, given $Z^i \subseteq X^i$ s.t. $P_i(Z^i) = 1$ for each $i = 1, \ldots, N$, $x_i^1, \ldots, x_i^m$ can be chosen to be in Z^i.

Proof. Choosing m large enough s.t. $\tilde{q}(m, \varepsilon, (\mathcal{F}_i)) < 1$ in Proposition 3.3 gives the existence of such $x_i^1, \ldots, x_i^m \in Z_i$ for $i = 1, \ldots, N$. □

4 Main Results

Theorem 4.1. *Let $X^1, \ldots, X^N$ be separable metric spaces. Let $0 < \varepsilon < \frac{e\ln(2)}{16}$, and let $\phi_1, \ldots, \phi_N$ be continuous functions $X^1 \times \cdots \times X^N \to [0,1]$, such that the game $(\phi_1, \ldots, \phi_N)$ possesses an equilibrium, and such that the induced classes*

$$\mathcal{F}'_i = \{\phi_i(x_i, \cdot) \mid x_i \in X^i\},\ i = 1, \ldots, N$$

each have pseudo-dimension at most d. Suppose

$$m > \frac{512}{\varepsilon^2}\Big(\ln(32N) - \ln(\varepsilon) + d \cdot \ln\big(\frac{64e}{\varepsilon}\ln(\frac{64e}{\varepsilon})\big)\Big) \tag{4.1}$$

Then the game $(\phi_1, \ldots, \phi_N)$ possesses an m-uniform ε-equilibrium.

When $X^1, \ldots, X^N$ are finite or compact metric, the existence of equilibrium follows from [13, Sec. 2]. The short proof of Theorem 4.1 resembles the proof of the main theorem in [4], but relying on our Proposition 3.4. Recall that the *support* of a probability measure on a separable metric space is the smallest closed set of full measure.

Proof. Let $\sigma = (\sigma_i)_{i=1}^N$ be a Nash equilibrium of the given game, $0 < \varepsilon < \frac{e\ln(2)}{16}$, and m as in (4.1). By Proposition 3.4 (with $\frac{\varepsilon}{2}$ replacing ε, and $Z_i = supp(\sigma_i)$) there are $(x_i^j)_{i=1,\ldots,N,j=1,\ldots,m}$ s.t. $x_i^1, \ldots, x_i^m$ are in $supp(\sigma_i)$ for $i = 1, \ldots, N$, and s.t. denoting $s_k = \frac{1}{m}\sum_{j=1}^m \delta_{x_k^j}$ for each player k, and writing (3.10) using the notation of mixed strategies,

$$|\phi_i(x_i, s_{-i}) - \phi_i(x_i, \sigma_{-i})| \leq \frac{\varepsilon}{2},\ \forall i = 1, \dots, N, \forall x_i \in X^i$$

Since σ is an equilibrium, and ϕ_i is continuous, for every $b_i \in X^i$,

$$\phi_i(b_i, \sigma_{-i}) \leq \phi_i(s_i, \sigma_{-i}) = \max \phi_i(\cdot, \sigma_{-i})$$

By choice of m and $s_1, \dots, s_N$, for any $b_i \in X^i$,

$$\phi_i(b_i, s_{-i}) \leq \phi_i(b_i, \sigma_{-i}) + \frac{\varepsilon}{2} \leq \phi_i(s_i, \sigma_{-i}) + \frac{\varepsilon}{2} \leq \phi_i(s_i, s_{-i}) + \varepsilon$$

Hence, $(s_1, \dots, s_N)$ is an m-uniform ε-equilibrium. □

Remark 4.1. In Theorem 4.1, if each player has k actions, then $d \leq \log_2(k)$, and equality can hold. Hence, up to a constant and the logarithmic factor $\ln\left(\frac{64e}{\varepsilon}\ln(\frac{64e}{\varepsilon})\right)$, we recover the result of [4] discussed in the introduction, as well as demonstrate the tightness of (4.1), up to said logarithmic factor, in actions. As discussed in the introduction, the tightness of dependence on the number of players is an open problem.

Remark 4.2. In the particular case of two-player zero-sum games, the full strength of Proposition 3.4 is not needed for Theorem 4.1; one can show by using simpler versions of our arguments that if $\varepsilon > 0$, d_1, d_2 denote the VC dimensions of the induced classes $\mathcal{F}'_1, \mathcal{F}'_2$, respectively, and m_1, m_2 satisfy (4.1) with d_2, d_1, respectively,[7] replacing d, then there are mixed actions s_1, s_2 which are m_1, m_2-uniform, respectively, which constitute a ε-equilibrium. Indeed, a result of this sort was pointed out in [19, p. 21:6] for two-player zero-sum games, with finitely many actions, in which the payoffs are $0-1$ valued.

Theorem 4.2. *Let $X^j \subseteq \mathbb{R}^{n_j}$ for $j = 1, \dots, N$, be semi-algebraic and compact, and let $\Lambda \subseteq \mathbb{R}^K$ be semi-algebraic.*

Let $\phi : \Lambda \times \prod_{j=1}^N X^j \to [0,1]^N$ be semi-algebraic, such that for each fixed $\lambda \in \Lambda$, $\phi(\lambda, \cdot)$ is continuous. Let $0 < \varepsilon < \frac{e\ln(2)}{16}$. Suppose $Gr(\phi^i)$ is of class (s, r), $r \geq 1$, in $\prod_i X^i \times \mathbb{R}$ for each $i = 1, \dots, N$, and suppose

$$m > \frac{512}{\varepsilon^2}\Big(\ln(32N) - \ln(\varepsilon) + 2\cdot(\max_i n_i + 1)\cdot\log_2(8er(s+1))\cdot\ln\big(\frac{64e}{\varepsilon}\ln(\frac{64e}{\varepsilon})\big)\Big) \tag{4.2}$$

Then there exists semi-algebraic mappings $\sigma_i^j : \Lambda \to X^i$, $i = 1, \dots, N$, $j = 1, \dots, m$, s.t. for each $\lambda \in \Lambda$, the profile $(s_1, \dots, s_n)$ defined by $s_i = \frac{1}{m}\sum_{j=1}^m \delta_{\sigma_i^j(\lambda)}$ for $i = 1, \dots, N$, is an m-uniform ε-equilibrium of $\phi(\lambda, \cdot)$.

[7] No, that's not a typo - the condition on m_1 uses d_2, and the condition on m_2 uses d_1.

Remark 4.3. Note that Theorem 4.2 allows for each ϕ^i to be the restriction to $\prod_{i=1}^N X^i$ of a function with graph of class (s,r) defined on a larger domain. This observation is useful, as the sets $X^1,\ldots,X^N$ may be complex (e.g., consisting of many isolated points) while ϕ's graph can be defined on a larger simpler domain (e.g., a hypercube); we may also wish to study approximate equilibria as we vary the domains $X^1,\ldots,X^N$.

Proof. As remarked above, the existence of Nash equilibrium in $\phi(\lambda,\cdot)$ for each $\lambda\in\Lambda$ follows from, e.g., [13, Sec. 2]. From Proposition 3.2, it follows that for each fixed $\lambda\in\Lambda$, and each $i=1,\ldots,N$, the family of functions $X_{-i}\to\mathbb{R}$ given by

$$\{\phi_i(\lambda,x_i,\cdot)\mid x_i\in X^i\}$$

is of pseudo-dimension at most $2(n_i+1)\cdot\log_2(8er(s+1))$. Hence, for m as in (4.2), the existence of an m-uniform ε-equilibrium in $\phi(\lambda,\cdot)$ for each $\lambda\in\Lambda$ follows from Theorem 4.1. The graph of the correspondence Φ which assigns to each $\lambda\in\Lambda$ its m-uniform ε-equilibria[8] is given by

$$\begin{aligned}\{(\lambda,(a_i^j)_{i=1,\ldots,N,j=1,\ldots,m})\mid \forall i=1,\ldots,N,\forall b_i\in X^i,\\ \frac{1}{m^{N-1}}\sum_{j_*\in[m]^{N-1}}\phi(\lambda,b_i,a_{-i}^{j_*})\le\frac{1}{m^N}\sum_{j\in[m]}\sum_{j_*\in[m]^{N-1}}\phi(\lambda,a_i^j,a_{-i}^{j_*})+\varepsilon\}\end{aligned}$$

where for $j_*=(j_*[k])_{k\neq i}\in[m]^{N-1}=\prod_{j\neq i}[m]$, $a_{-i}^{j_*}$ specifies $a_k^{j_*[k]}$ for each player $k\neq i$. Hence $Gr(\Phi)$ is semi-algebraic. Hence, by the semi-algebraic selection theorem, Theorem 2.3, there exists the desired σ. ☐

Proposition 4.4. *Let $X^j\subseteq\mathbb{R}^{n_j}$ for $j=1,\ldots,N$, be semi-algebraic and compact.*

Let $\phi:(0,\delta)\times X^1\times X^2\to\mathbb{R}$ be bounded and semi-algebraic, with X^1,X^2 compact and semi-algebraic, and for each $\lambda\in(0,\delta)$, let v_λ denote the value of $\phi(\lambda,\cdot)$. Then $\lim_{\lambda\to 0}v_\lambda$ exists.

The proposition follows immediately from Theorem 4.2 and from Lemma 3.1.

Remark 4.5. As pointed out in the introduction, if $\lim_{\lambda\to 0}v_\lambda$ were shown to exist in $\mathbb{R}\cup\{-\infty,\infty\}$ under the other assumptions Proposition 4.4 but *without* the boundedness condition, it would imply the existence of the asymptotic value (that is, convergence of the discounted values) of stochastic games with finitely many states, compact action spaces, and semi-algebraic transitions, due to a result of [3]; see Section 7.

[8] Formally, Φ assigns, to each λ, N m-tuples, possibly with repeats, which present the strategies each player should mix over; an m tuple $(a_i^1,\ldots,a_i^m)$ for Player i corresponds to the mixed strategy $\frac{1}{m}\sum_{j=1}^m\delta_{a_i^j}$.

5 Extensions

A generalization of semi-algebraic structures are o-minimal structures (see, e.g., [12]):

Definition 5.1. *An o-minimal structure on $\mathbb{R}$ is a sequence $(\mathcal{S}_n)_{n\in\mathbb{N}}$ such that for each n:*

1. *$\mathcal{S}_n$ is a Boolean algebra of subsets of $\mathbb{R}^n$.*
2. *$A \in \mathcal{S}_n \to A \times \mathbb{R} \in \mathcal{S}_{n+1}$ and $\mathbb{R} \times A \in \mathcal{S}_{n+1}$.*
3. *$\{(x_1, \ldots, x_n) \in \mathbb{R}^n \mid x_i = x_j\} \in \mathcal{S}_n$ for each $1 \leq i < j \leq n$.*
4. *$A \in \mathcal{S}_{n+1} \to \pi(A) \in \mathcal{S}_n$, where $\pi : \mathbb{R}^{n+1} \to \mathbb{R}^n$ is the usual projection map.*
5. *$\{(x, y) \in \mathbb{R}^2 \mid x < y\} \in \mathcal{S}_2$.*
6. *$\mathcal{S}_1$ consists precisely of finite unions of open intervals*[9] *and points.*

The semi-algebraic sets are an example of an o-minimal structure. But there are others: E.g., the *globally subanalytic sets*: There exists an o-minimal structure that contains all sets of the form $\{(x,t) \in [-1,1]^p \times \mathbb{R} \mid f(x) = t\}$ where $f : [-1,1]^p \to \mathbb{R}$, $p \in \mathbb{N}$, is a function which can be extended analytically to a neighborhood of $[-1,1]^p$. (See, e.g., [9]). For further examples, see [12] and the references therein.

Given fixed o-minimal structure, its elements are called *definable* (in that structure). A function is *definable* if its graph is definable. (Indeed, [11] discusses stochastic games, which have a long relation to semi-algebraic geometry - see, e.g., [20, Ch. 6] - in which the primitives of the game are definable.)

The monotonicity theorem, Theorem 2.2, and the selection theorem, Theorem 2.3, hold for o-minimal structures as well (see [12], Ch. 3 and Section 6.1, respectively). [16] shows that if $A \subseteq \mathbb{R}^n \times \mathbb{R}^m$ is defineable in an o-minimal structure, then the collection

$$\mathcal{A} = \Big\{\{(y,x) \in A\} \mid y \in \mathbb{R}^n\Big\}$$

is of finite VC dimension; see also [12, Ch. 5]. Using this, one can show as we have done for the semi-algebraic case:

Theorem 5.1. *Fix an o-minimal structure. Let $\Lambda \subseteq \mathbb{R}^K$ be definable, let $X^1, \ldots, X^N$ be definable and compact, and let $\phi : \Lambda \times \prod_{j=1}^N X^j \to [0,1]^N$ be definable, such that for each fixed $\lambda \in \Lambda$, $\phi(\lambda, \cdot)$ is continuous. Let $0 < \varepsilon < \frac{e\ln(2)}{16}$. Then, for $m \in \mathbb{N}$ large enough, there exists definable mappings $\sigma_j^i : \Lambda \to X^i$, $i = 1, \ldots, N$, $j = 1, \ldots, m$, s.t. for each $\lambda \in \Lambda$, the profile $(s_1, \ldots, s_n)$ defined by $s_i = \frac{1}{m}\sum_{j=1}^m \delta_{\sigma_i^j(\lambda)}$ for $i = 1, \ldots, N$, is an m-uniform ε-equilibrium of $\phi(\lambda, \cdot)$.*

[9] Infinite endpoints allowed.

6 Discussion on Bounds, From Finite Actions to Continuum

In this section we show how to embed a game with finite action spaces into a continuous game with action spaces being the unit interval, such that if the original game did not have small support ε-equilibria, then neither does the resulting game.

First we describe the transformation. Essentially, one embeds the finite actions spaces into the unit interval, and extends the payoffs piece-wise multi-linearly. Formally: Let $N \in \mathbb{N}$, and let $A^1, \ldots, A^N$ be finite actions for N players, $G : \prod_{i=1}^N A^i \to \mathbb{R}^N$ a game, and embeddings $A^i \to [0,1]$, which includes $0, 1$ in each of their ranges, identifying A^i as a subset of $[0,1]$. Define a game $\tilde{G} : [0,1]^N \to [0,1]^N$ in the following manner: For each $i = 1, \ldots, N$ and each $x^i \in [0,1]$, let $\phi^{\pm}(x^i)$ be defined by

$$\phi^-(x^i) = \max\{a^i \in A^i \mid a^i \le x^i\},\ \phi^+(x^i) = \min\{a^i \in A^i \mid a^i \ge x^i\}$$

Hence, $x^i \in [\phi^-(x^i), \phi^+(x^i)]$; $\phi^-(x^i) = \phi^+(x^i) = x^i$ if $x^i \in A^i$; and if $x^i \notin A^i$, $(\phi^-(x^i), \phi^+(x^i)) \cap A^i = \emptyset$. We let $\tilde{G}$ be the piece-wise multi-linear extension of G, i.e., for each $i = 1, \ldots, N$ and each $x^i \in [0,1]$, let $\lambda^i(x^i) \in [0,1]$ be s.t.

$$x^i = \lambda^i(x^i) \cdot \phi^+(x^i) + (1 - \lambda^i(x^i)) \cdot \phi^-(x^i)$$

if $x^i \notin A^i$, and arbitrarily $\in [0,1]$ if $x^i \in A^i$; and then set

$$\tilde{G}(x^1, \ldots, x^N) = \sum_{s \in \{\pm\}^N} \Big[\prod_{j \mid s[j]=+} \lambda^j(x^j) \prod_{j \mid s[j]=-} (1-\lambda^j(x^j)) \Big] G(\phi^{s[1]}(x^1), \ldots, \phi^{s[N]}(x^N)) \tag{6.1}$$

Such $\tilde{G}$ is easily seen to be semi-algebraic and Lipshitz.

Now, fix $\varepsilon \ge 0$, and suppose the origional game G has no ε-equilibrium with strategies which have support $\le K$. We contend that $\tilde{G}$ has no ε-equilibrium with strategies which have support $\le \lfloor \frac{K}{2} \rfloor$. Suppose it had such a ε-equilibrium $(s^1, \ldots, s^N)$. Intuitively, weight α on some $x^i \in [0,1]$ corresponds to putting weights $\alpha\lambda(x^i), \alpha(1 - \lambda(x^i))$, on $\phi^-(x^i), \phi^+(x^i) \in A^i$, respectively. More formally: For each $i = 1, \ldots, N$, define t^i from s^i by replacing each action used with the closest actions in A^i with the appropriate proportions, i.e.,

$$s^i = \sum_{j=1}^{M_i} \delta_{x^i_j} \to t_i := \sum_{j=1}^{M_i} \Big[\lambda^i(x^i_j)\delta_{\phi^+(x^i_j)} + (1 - \lambda^i(x^i_j))\delta_{\phi^-(x^i_j)} \Big]$$

Then it follows

$$\tilde{G}^i(\cdot, s_{-i}) \equiv \tilde{G}^i(\cdot, t_{-i})$$

and

$$\max_{A^i} G^i(\cdot, t_{-i}) = \max_{A^i} \tilde{G}^i(\cdot, t_{-i}) = \max_{[0,1]} \tilde{G}^i(\cdot, t_{-i})$$

So $(t^1, \dots, t^N)$ is a ε-equilibrium with support $\leq 2 \cdot \lfloor \frac{K}{2} \rfloor \leq K$, a contradiction. Clearly also

$$\max G^i = \max \tilde{G}^i, \;\; \min G^i = \min \tilde{G}^i, \;\; \forall i = 1, \dots, N$$

and if G is a two-player zero-sum game, then so is $\tilde{G}$.

In particular, using the results of [1, Sec. 5] which provides zero-sum games which require arbitrarily large support for ε-equilibria:

Theorem 6.1. *For each $K \in \mathbb{N}$, there is $\varepsilon > 0$ and a semi-algebraic Lipshitz two-player zero-sum game $\tilde{G} : [0,1]^2 \to [0,1]$ s.t. there do not exist ε-optimal strategies for each player with support of size at most K.*

Let us unbox this theorem a little; we will use the 'Big O' notation, where$f = O(g)$ means $f \geq C \cdot g$ for asymtotically for some $C > 0$; the 'Big Omega' notation, where $f = \Omega(g)$ means $g = O(f)$; and the 'Big Theta' notation, where $f = \Theta(g)$ means $f = O(g)$ and $g = O(f)$.

Recall that [1, Sec. 5] shows that there is $C > 0$ s.t. for n large enough and ε, there exists a zero-sum game $G_{n,\varepsilon}$, with n strategies for each players and payoffs $0-1$-valued, s.t. there do not exist for both players ε-optimal strategies in $G_{n,\varepsilon}$ with support size $\leq C \frac{\log(n)}{\varepsilon^2}$. Let $\tilde{G}_{n,\varepsilon}$ denote the transformation of $G_{n,\varepsilon}$, as described above, to a zero-sum game with strategy space $[0,1]$ for each player. The above argument shows there do not exist for both players ε-optimal strategies in $G_{n,\varepsilon}$ with support size $\leq \frac{C}{2} \cdot \frac{\log(n)}{\varepsilon^2}$. Hence, the dependence on ε in Theorem 4.2, which is $\Theta(\frac{1}{\varepsilon^2} \log(\frac{1}{\varepsilon}))$, is tight up to the factor $\log(\frac{1}{\varepsilon})$, which is much less significant than the factor $\frac{1}{\varepsilon^2}$.[10]

Let us unbox further: The resulting game $\tilde{G}_{n,\varepsilon}$ have semi-algebraic payoffs easily seen to be of class (at most) $(O(n^2), O(1))$. Hence, for fixed $\varepsilon > 0$, Theorem 4.2 shows that ε-optimal strategies exist with support $\Theta(\log(n))$. Hence, although it is possible that the class of the payoffs of $\tilde{G}_{n,\varepsilon}$ may be of class $(s_n, O(1))$ with s_n much less than $\Theta(n^2)$, we must have $\log(s_n) = \Omega(\log(n))$.

7 Potential Connections to Stochastic Games

The seminal paper [23] studies the following model of zero-sum stochastic games:

- K - A finite set of states.
- I, J - Finite action sets.
- $r : K \times I \times J \to \mathbb{R}$ - A payoff function.

[10] This additional logarithmic factor is already presented in the bounds in [4], and it is not clear if it is required.

- $Q : I \times J \to \mathbb{R}^{K\times K}$ - A mapping to stochastic matrices.
- A discount rate $\lambda \in (0,1)$.

The game is played in discrete time. If at some stage of the game and the players select an actions $i \in I, j \in J$, the probability of transition from k to ℓ is given by $Q_{k,\ell}(i,j)$. The payoff starting at stage $z_1 = k \in K$, if z_t is stage at t, under stationary strategies $\sigma : K \to \Delta(I), \tau : K \to \Delta(J)$, is

$$\gamma_\lambda^k(\sigma,\tau) = E_{\sigma,\tau}[\lambda \sum_{t=1}^{\infty}(1-\lambda)^{t-1} r(z_t, \sigma(z_t), \tau(z_t)) \mid z_1 = z]$$

and $\gamma_\lambda = (\gamma_\lambda^k)_{k\in K}$. Let $v_\lambda \in \mathbb{R}^K$ denote the λ-discounted value, where v_λ^k is the value in the game with initial state k.

[3] derives a characterization for the asymptotic value $\lim_{\lambda\to 0} v_\lambda$; the limit had been known to exist by the results of [8]. The essence of the convergence of $\lim_{\lambda\to 0} v_\lambda$ is that v_λ is semi-algebraic when K, I, J are finite, and hence the limit exists by piece-wise monotonicity.

The analysis in [3] begins as follows. By standard dynamic programming results, e.g. [20, Ch. 2], the payoffs (as a function of initial state) is given by for pure strategies $i \in I^K, j \in J^K$:

$$\gamma_\lambda(i,j) = \lambda r(i,j) + (1-\lambda)Q(i,j)\gamma_\lambda(i,j)$$

where $r(i,j) = (r^k(i^k, j^k))_{k\in K}$ and $Q(i,j) = (Q_{k,k'}(i^k, j^k))_{k,k'\in K}$. Denoting

$$d_\lambda^0(i,j) = det(I - (1-\lambda)Q(i,j))$$

and d_λ^k is similarly defined by replacing k-th column of $I - (1-\lambda)Q(i,j)$ with $\lambda r(i,j)$, Cramer's rule gives

$$d_\lambda^k(i,j) - \gamma_\lambda^k(i,j) \cdot d_\lambda^0(i,j) = 0$$

Hence, defining for each $z \in \mathbb{R}$, $\lambda \in (0,1)$, $k \in K$, the game with action spaces I^K, J^K,

$$W_\lambda^k(z)[i,j] = d_\lambda^k(i,j) - z \cdot d_\lambda^0(i,j)$$

we see that $W_\lambda^k(\gamma_\lambda^k(i,j))[i,j] = 0$ for every pair of pure strategies $i \in I^K, j \in J^K$. Theorem 1 of [3] shows:

Theorem 7.1. *For each $k \in K$, v_λ^k is unique solution to $val(W_\lambda^k(z)) = 0$.*

Not only this, but v_λ^k is a 0 of $val(W_\lambda^k(\cdot))$ in a strong sense. Theorem 2 of [3] shows:

Theorem 7.2. *Denote for $z \in \mathbb{R}$, $k \in K$,*

$$F^k(z) := \lim_{\lambda \to 0} \frac{valW_\lambda^k(z)}{\lambda^K} \in \mathbb{R} \cup \{\pm\infty\} \tag{7.1}$$

Then $\lim_{\lambda \to 0} v_\lambda = w \in \mathbb{R}^K$, where for each $k \in K$, w^k is unique solution to

$$\begin{aligned} z > w^k \to F^k(z) < 0 \\ z < w^k \to F^k(z) > 0 \end{aligned} \tag{7.2}$$

For I, J, K finite, the limit in (7.1) exists in $\mathbb{R} \cup \{\pm\infty\}$ by piece-wise monotonicity, Theorem 2.2. Section 6B of [3] discusses extensions to the case when I, J are compact metric spaces (and r, q are continuous); K remains finite. It is pointed out there that Theorem 7.1 extends readily to this case. However, Theorem 7.2 may fail as the limit (7.1) need not exist. Indeed, [26] shows that for such games, $\lim_{\lambda \to 0} v_\lambda$ need not exist. However, a variation of Theorem 7.2 readily follows:

Theorem 7.3. *Assume that the limit (7.1) exists for each $k \in K$, $z \in \mathbb{R}$. Then $\lim_{\lambda \to 0} v_\lambda = w \in \mathbb{R}^K$, where for each $k \in K$, w^k is unique solution to (7.2).*

Hence, the question arises - for which stochastic games does the the limit(7.1) exist? More generally:

Given $\phi : (0,1) \times X \times Y \to \mathbb{R}$ continuous and semi-algebraic, does $\lim_{\lambda \to 0} val\phi(\lambda, \cdot, \cdot)$ exists in $\mathbb{R} \cup \{\pm\infty\}$?

In our paper, we have shown in Proposition 4.4 that the answer is affirmative when ϕ is bounded. However, when $\phi_z : (0,1) \times I \times J \to \mathbb{R}$ is give by $\phi_z(\lambda, i, j) = \frac{1}{\lambda^K} W_\lambda^k(z)[i,j]$ for some $k \in K$ and $z \in \mathbb{R}$, the resulting ϕ_z is clearly in general *not* bounded, and this case is still open.[11]

8 Appendix: Alternative Approach

The following result is derived using similar techniques in [7]:[12]

Theorem 8.1. *Let $\mathcal{A}(n, s, r)$ be the collection of all semi-algebraic sets in $\mathbb{R}^n$ of class (s, r). Then $VCdim(\mathcal{A}(n, s, r)) \leq 4s\binom{r+n}{r} \cdot \log_2 \left(2s(2s+1)\binom{r+n}{r}\right)$.*

Like for Theorem 2.4, the bound in Theorem 8.1 is demonstrated to be tight up to a logarithmic factor.

[11] It will be bounded for all $z \in \mathbb{R}$ and $k \in K$ if, e.g., the game remains in the same state at all times regardless of actions; it is not clear if this is the only case.

[12] The statement in that paper, Claim 8 of p. 217, gives a slightly better bound but considers only Boolean combinations of sets defined using *strict* polynomial inequalities. To be consistent with the rest of the paper, we've modified the result to allow equalities and weak inequalities, which requires at worst doubling the number of polynomials needed.

Proposition 8.1. *For $X \subseteq \mathbb{R}^n$, let $\mathcal{F}(n,s,r)$ denote the class of all functions $X \to [0,1]$ whose graphs in $X \times \mathbb{R}$ are of class (s,r). For $r \geq 1$,*

$$P - dim(\mathcal{F}(n,s,r)) \leq 4(s+1)\binom{r+n+1}{r} \cdot \log_2\big(2(s+1)(2s+3)\binom{r+n+1}{r}\big).$$

The proof mimics that of Proposition 3.2, except uses Theorem 8.1 instead of Theorem 2.4.

Theorem 8.2. *Let $X^j \subseteq \mathbb{R}^{n_j}$ for $j = 1, \ldots, N$, be semi-algebraic and compact. Let $0 < \varepsilon < \frac{e\ln(2)}{16}$, and let $\phi_1, \cdots, \phi_N$ be payoff functions $X^1 \times \cdots X^N \to [0,1]$ with graphs of class (s,r) in $X^1 \times \cdots X^N \times [0,1]$, $r \geq 1$; denote $n_0 = \max_i \sum_{j \neq i} n_j$, and suppose*

$$\begin{aligned} m > & \frac{512}{\varepsilon^2}\Big(\ln(32N) - \ln(\varepsilon) \\ & + 4(s+1)\binom{r+n_0+1}{r} \cdot \log_2\big(2(s+1)(2s+3)\binom{r+n_0+1}{r}\big) \cdot \ln\big(\frac{64e}{\varepsilon}\ln(\frac{64e}{\varepsilon})\big)\Big). \end{aligned} \tag{8.1}$$

Then the game $(\phi_1, \ldots, \phi_N)$ posseses an m-uniform ε-equilibrium.

Proof. As remarked above, the existence of Nash equilibrium in each such game follows from, e.g., [13, Sec. 2]. From Proposition 8.1, it follows that for each $i = 1, \ldots, N$, the family of functions $X_{-i} \to \mathbb{R}$ given by

$$\{\phi_i(\lambda, x_i, \cdot) \mid \lambda \in \Lambda, x_i \in X^i\}$$

is of pseudo-dimension at most $4(s+1)\binom{r+n_0+1}{r} \cdot \log_2\big(2(s+1)(2s+3)\binom{r+n_0+1}{r}\big)$. Hence, for m as in (8.1), the existence of an m-uniform ε-equilibrium follows from Theorem 4.1. □

References

[1] Ingo Althofer. On sparse approximations to randomized strategies and convex combinations. *Linear Algebra and its Applications*, 199:339 – 355, 1994. Special Issue Honoring Ingram Olkin.

[2] Itai Arieli and Yakov Babichenko. Simple approximate equilibria in games with many players. In *Proceedings of the 2017 ACM Conference on Economics and Computation*, EC 17, page 681691, New York, NY, USA, 2017. Association for Computing Machinery.

[3] Luc Attia and Miquel Oliu-Barton. A formula for the value of a stochastic game. *Proceedings of the National Academy of Sciences*, 116(52):26435–26443, 2019.

[4] Yakov Babichenko, Siddharth Barman, and Ron Peretz. Simple approximate equilibria in large games. In *Proceedings of the Fifteenth ACM Conference on Economics and Computation*, EC 14, page 753770, New York, NY, USA, 2014. Association for Computing Machinery.

[5] Dieter Balkenborg and Dries Vermeulen. On the topology of the set of nash equilibria. *Games and Economic Behavior*, 118:1 – 6, 2019.

[6] Siddharth Barman. Approximating nash equilibria and dense bipartite subgraphs via an approximate version of caratheodorys theorem. In *Proceedings of the Forty-Seventh Annual ACM Symposium on Theory of Computing*, STOC 15, page 361369, New York, NY, USA, 2015. Association for Computing Machinery.

[7] Shai Ben-David and Michael Lindenbaum. Localization vs. identification of semi-algebraic sets. *Machine Learning*, 32(3):207–224, 1998.

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

[9] Edward Bierstone and Pierre D. Milman. Semianalytic and subanalytic sets. *Publications Mathématiques de l'Institut des Hautes Études Scientifiques*, 67(1):5–42, 1988.

[10] J. Bochnak, M. Coste, and M.-F. Roy. *Géométrie Algébrique Réelle*. Springer-Verlag, New York, 1987.

[11] Jrme Bolte, Stphane Gaubert, and Guillaume Vigeral. Definable zero-sum stochastic games. *Mathematics of Operations Research*, 40(1):171–191, 2015.

[12] L. P. D. van den Dries. *Tame Topology and O-minimal Structures*. London Mathematical Society Lecture Note Series. Cambridge University Press, 1998.

[13] I.L. Glicksberg. A further generalization of the Kakutani fixed point theorem with applications to Nash equilibrium. *Proceedings of the American Mathematical Society*, 3:170–174, 1952.

[14] Paul W. Goldberg and Mark R. Jerrum. Bounding the vapnik-chervonenkis dimension of concept classes parameterized by real numbers. *Machine Learning*, 18(2):131–148, 1995.

[15] E. Kohlberg and J.-F. Mertens. On the strategic stability of equilibria. *Econometrica*, 54:1003–1038, 1986.

[16] Michael C. Laskowski. Vapnik-Chervonenkis Classes of Definable Sets. *Journal of the London Mathematical Society*, s2-45(2):377–384, 04 1992.

[17] Yehuda John Levy. Projections and functions of nash equilibria. *International Journal of Game Theory*, 45(1):435–459, 2016.

[18] Richard J. Lipton, Evangelos Markakis, and Aranyak Mehta. Playing large games using simple strategies. In *Proceedings of the 4th ACM Conference on Electronic Commerce*, EC 03, page 3641, New York, NY, USA, 2003. Association for Computing Machinery.

[19] Shay Moran and Amir Yehudayoff. Sample compression schemes for vc classes. *J. ACM*, 63(3), June 2016.

[20] Abraham Neyman and Sylvain Sorin. *Stochastic Games and Applications*. Kluwer Academic Publishers, 2003.

[21] D. Pollard. *Convergence of Stochastic Processes*. Clinical Perspectives in Obstetrics and Gynecology. Springer New York, 1984.

[22] S. Schanuel, L.K. Simon, and W.R. Zame. The algebraic geometry of games and the tracing procedure. In R. Selten, editor, *Game Equilibrium Models, II: Methods, Morals and Markets*, pages 9–44. Springer-Verlag, Berlin, 1991.

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

[24] V. N. Vapnik and A. Ya. Chervonenkis. On the uniform convergence of relative frequencies of events to their probabilities. *Theory of Probability & Its Applications*, 16(2):264–280, 1971.

[25] Mathukumalli Vidyasagar. *Learning and Generalization: With Applications to Neural Networks*. Springer Publishing Company, Incorporated, 2nd edition, 2010.

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

[27] Guillaume Vigeral and Yannick Viossat. Semi-algebraic sets and equilibria of binary games. *Oper. Res. Lett.*, 44:19–24, 2016.