

# Learning and Computation of $\Phi$ -Equilibria at the Frontier of Tractability

Brian Hu Zhang<sup>\*1</sup>, Ioannis Anagnostides<sup>\*1</sup>, Emanuel Tewolde<sup>1,2</sup>, Ratip Emin Berker<sup>1,2</sup>,  
Gabriele Farina<sup>3</sup>, Vincent Conitzer<sup>1,2,4</sup>, and Tuomas Sandholm<sup>1,5</sup>

<sup>1</sup>Carnegie Mellon University

<sup>2</sup>Foundations of Cooperative AI Lab (FOCAL)

<sup>3</sup>Massachusetts Institute of Technology

<sup>4</sup>University of Oxford

<sup>5</sup>Additional affiliations: Strategy Robot, Inc., Strategic Machine, Inc., Optimized Markets, Inc.

{bh Zhang, ianagnos, etewolde, rberker, conitzer, sandholm}@cs.cmu.edu, gfarina@mit.edu

March 3, 2025

## Abstract

$\Phi$ -*equilibria*—and the associated notion of  $\Phi$ -*regret*—are a powerful and flexible framework at the heart of online learning and game theory, whereby enriching the set of deviations  $\Phi$  begets stronger notions of rationality. Recently, Daskalakis, Farina, Fishelson, Pipis, and Schneider (STOC '24)—abbreviated as DFFPS—settled the existence of efficient algorithms when  $\Phi$  contains only linear maps under a general,  $d$ -dimensional convex constraint set  $\mathcal{X}$ . In this paper, we significantly extend their work by resolving the case where  $\Phi$  is  $k$ -dimensional; degree- $\ell$  polynomials constitute a canonical such example with  $k = d^{O(\ell)}$ . In particular, positing only oracle access to  $\mathcal{X}$ , we obtain two main positive results:

- a  $\text{poly}(n, d, k, \log(1/\epsilon))$ -time algorithm for computing  $\epsilon$ -approximate  $\Phi$ -equilibria in  $n$ -player multilinear games, and
- an efficient online algorithm that incurs average  $\Phi$ -regret at most  $\epsilon$  using  $\text{poly}(d, k)/\epsilon^2$  rounds.

We also show nearly matching—up to constant factors in the exponents—lower bounds parameterized by  $k$  in the online learning setting, thereby obtaining for the first time a family of deviations that captures the learnability of  $\Phi$ -regret.

From a technical standpoint, we extend the framework of DFFPS from linear maps to the more challenging case of maps with polynomial dimension. At the heart of our approach is a polynomial-time algorithm for computing an *expected fixed point* of any  $\phi : \mathcal{X} \rightarrow \mathcal{X}$ —that is, a distribution  $\mu \in \Delta(\mathcal{X})$  such that  $\mathbb{E}_{\mathbf{x} \sim \mu}[\phi(\mathbf{x}) - \mathbf{x}] \approx 0$ —based on the seminal *ellipsoid against hope (EAH)* algorithm of Papadimitriou and Roughgarden (JACM '08). In particular, our algorithm for computing  $\Phi$ -equilibria is based on executing EAH in a nested fashion—each step of EAH itself being implemented by invoking a separate call to EAH.

---

\*Equal contribution.

# Contents

<b>1</b>	<b>Introduction</b>	<b>1</b>
1.1	Our results: $\Phi$ -equilibria at the frontier of tractability . . . . .	2
1.1.1	Learning . . . . .	3
1.1.2	Computation . . . . .	3
1.2	Technical approach . . . . .	3
1.3	Further related work . . . . .	5
<b>2</b>	<b>Preliminaries</b>	<b>6</b>
2.1	Online learning and $\Phi$ -regret . . . . .	7
2.2	Ellipsoid against hope . . . . .	9
<b>3</b>	<b>Sets of deviations with polynomial dimension</b>	<b>11</b>
<b>4</b>	<b>Polynomial-time expected fixed points and semi-separation</b>	<b>12</b>
<b>5</b>	<b>A polynomial-time algorithm for <math>\Phi^m</math>-equilibria in games</b>	<b>14</b>
<b>6</b>	<b>An efficient online algorithm for minimizing <math>\Phi^m</math>-regret</b>	<b>16</b>
6.1	Shell ellipsoid . . . . .	16
6.2	Shell gradient descent . . . . .	17
6.3	Shell projection . . . . .	18
6.4	Putting everything together . . . . .	18
<b>7</b>	<b>Conclusions and open problems</b>	<b>20</b>
<b>A</b>	<b>Sufficiency of regret minimization in isotropic position</b>	<b>26</b>
<b>B</b>	<b>Geometric properties of <math>\Phi^m</math></b>	<b>26</b>
<b>C</b>	<b>Further omitted proofs</b>	<b>27</b>

# 1 Introduction

What constitutes a solution to a game? Descriptive and prescriptive theories in economics brought together in response to such questions chiefly revolve around *equilibria*—strategically stable outcomes. The underpinning of strategic stability can be naturally conceived in more ways than one, but none is more standard than the one pregnant in Nash’s theorem [49], which single-handedly propelled the rapid development of noncooperative game theory in the last century. This is reflected in, for example, the exulting words of the Nobel prize-winning economist Roger Myerson, who juxtaposed the “fundamental and pervasive impact of Nash equilibrium” in economics and the social sciences to “the discovery of the DNA double helix in the biological sciences [48].” This profound influence notwithstanding, the concept of Nash equilibrium has also been subjected to fierce criticism. Some of the most vocal critiques have been articulated in the computer science community, vehemently objecting to its apparent intractability [16, 13, 55, 54, 22].

Aumann’s pathbreaking *correlated equilibrium* concept [2] offers a compelling alternative that promises to alleviate such concerns: unlike Nash equilibria, which amount to fixed points of general functions, correlated equilibria can be expressed as linear programs. This, in turn, enables their efficient computation, or at least so the common narrative goes. The reality is more nuanced, and is dictated by the underlying *game representation*. In particular, strategic interactions encountered in practice often unfold sequentially. Such settings do not lend themselves to the usual one-shot (aka. normal-form) representation—not without blowing up the description of the game, that is. Instead, sequential games are usually represented in *extensive form* (or other pertinent paradigms [41]). In stark contrast to normal-form games, wherein polynomial-time algorithms have long been established [51, 38], the complexity of computing correlated equilibria in extensive-form games remains an enigmatic open problem.

Daskalakis et al. [17] recently provided evidence in the negative through the prism of the *no-regret* framework. In particular, there is an intimate nexus between correlated equilibria and online learning: learners minimizing *swap regret*—a powerful notion of hindsight rationality—converge to the set of correlated equilibria [6, 56]. In this context, Daskalakis et al. [17] established an exponential lower bound for the number of iterations needed so that a learner incurs at most  $\epsilon$  average swap regret when facing an adversary; more precisely, that result applies in the regime where  $\epsilon$  is inversely polynomial (*cf.* Dagan et al. [14], Peng and Rubinstein [52]). While this result does not rule out the existence of efficient algorithms beyond the adversarial regime, it does immediately bring to the fore a well-studied but pressing question: *what notions of hindsight rationality are efficiently learnable?*

Hindsight rationality in online learning can be understood through a set of functions,  $\Phi$ , so that no deviation according to a function in  $\Phi$  can retrospectively improve the cumulative utility; such a learner is said to be consistent with minimizing  $\Phi$ -*regret* [31, 57, 29]. The broader the set of deviations  $\Phi$ , the more appealing the ensuing concept of hindsight rationality. The usual notion of external regret is an instantiation of that framework for which  $\Phi$  contains solely constant functions—referred to as *coarse* deviations. On the other end of the spectrum, when  $\Phi$  contains all possible deviations, one finds the powerful notion of swap regret—associated with (normal-form) correlated equilibria. The fundamental question thus is to characterize the structure of  $\Phi$  that enables efficient learnability—and, indeed, computation.

Much of the recent research in the context of learning in extensive-form games has focused on this exact problem. This can be traced back to the influential work of Zinkevich et al. [61], who introduced *counterfactual regret minimization (CFR)*—an algorithm that was at the heart of recent landmark results in AI benchmarks such as poker [9, 10, 8, 44]. CFR is an online algorithm

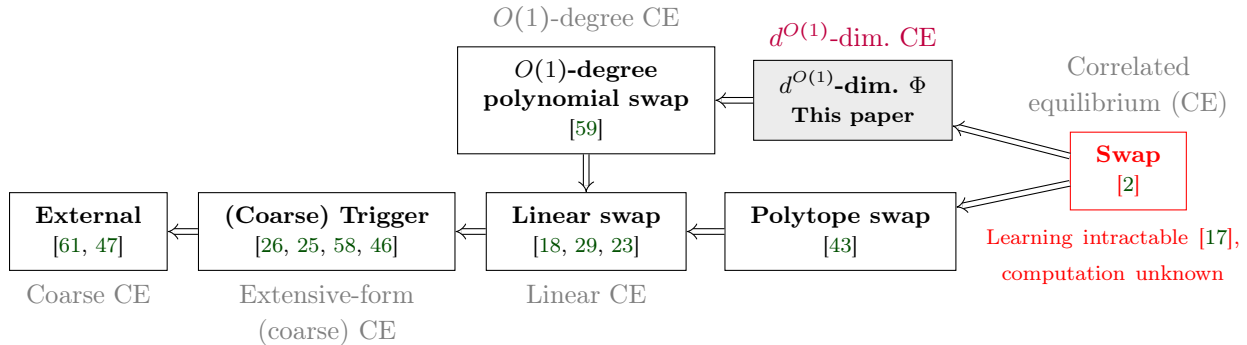


Figure 1: The arrows  $A \implies B$  denote that minimizing the notion of regret  $A$  implies minimizing the notion of regret  $B$ . In other words,  $A$  defines a superset of deviations that the learner considers compared to  $B$ , and hence leads to a stronger notion of equilibrium. The gray text below or above a notion of regret denotes the name of the corresponding notion of equilibrium, if applicable.

for minimizing external regret—associated with (normal-form) coarse correlated equilibria. Moving forward, efficient algorithms eventually emerged for *extensive-form correlated equilibria (EFCE)* [37, 26, 20, 3] (*cf.* Morrill et al. [45, 46]), and more broadly, when  $\Phi$  contains solely *linear* functions [24, 23]—corresponding to *linear correlated equilibria (LCE)*. Daskalakis et al. [18] recently took a step even further by strengthening those results whenever the underlying constraint set  $\mathcal{X} \subseteq \mathbb{R}^d$  admits a membership oracle. [Figure 1](#) summarizes the landscape that has emerged.

## 1.1 Our results: $\Phi$ -equilibria at the frontier of tractability

Table 1: Our main results for the  $k$ -dimensional set  $\Phi^m$  of [Definition 1.1](#);  $k \gg d$ .

	Upper bound	Lower bound
<b>Learning</b>	$\text{poly}(k)/\epsilon^2$ ( <a href="#">Theorem 1.2</a> )	$\min\{\sqrt{k}/4, \exp(\Omega(\epsilon^{-1/6}))\}$ ( <a href="#">Theorem 1.6</a> )
<b>Computation</b>	$\text{poly}(k, \log(1/\epsilon))$ ( <a href="#">Theorem 1.3</a> )	Open question

The primary focus of this paper is to expand the scope of that prior research beyond linear-swap regret—associated with linear correlated equilibria—toward the frontier of tractability. While handling all swap deviations is impossible in light of the recent lower bound of Daskalakis et al. [17], discussed above, we are able to cope with the broad class of functions introduced below.

**Definition 1.1.** Given a map  $m : \mathcal{X} \rightarrow \mathbb{R}^{k'}$ , the set of deviations  $\Phi^m$  is defined as the set of all maps  $\phi : \mathcal{X} \rightarrow \mathcal{X}$  that can be expressed by the matrix-vector product  $\mathbf{K}(\phi)m(\mathbf{x}) + \mathbf{c}(\phi)$  for some  $\mathbf{K}(\phi) \in \mathbb{R}^{d \times k'}$  and  $\mathbf{c}(\phi) \in \mathbb{R}^d$ . The set of functions  $\Phi^m$  has dimension at most  $k := k' \cdot d + d$ .

We think of  $k$  as a measure of the complexity of  $\Phi^m$ ; in what follows, one may imagine  $k \leq \text{poly}(d)$ . There is a clear sense in which going beyond [Definition 1.1](#) is daunting: even representing such

functions becomes prohibitive. Indeed, we also establish lower bounds that preclude going beyond the set of deviations in [Definition 1.1](#), showing that our results cannot be significantly improved.

As a canonical example, one can capture degree- $\ell$  polynomials by taking  $m(\mathbf{x})$  to be the function that outputs all  $\ell$ -wise (and lower) products of entries in  $\mathbf{x}$  (hence  $k = d^{O(\ell)}$ ), and  $\mathbf{K}$  the matrix of coefficients of the polynomial. (For technical reasons, we actually consider a certain orthonormal basis for polynomials introduced formally in [Definition 3.4](#).)

### 1.1.1 Learning

We begin by stating our results in the usual no-regret framework in online learning, and we then proceed with the centralized model of computation. A key reference point here is the recent paper of Zhang et al. [59], who provided online algorithms in extensive-form games when  $\Phi$  contains low-degree polynomials. However, the complexity of their algorithm—both the per-iteration running time and the total number of rounds—depends on the depth of the game tree; in particular, for general extensive-form games, their bounds become vacuous even for constant-degree polynomials (for example, when the game tree is lopsided). Our first main result addresses that limitation. What is more, it encompasses a more general class than extensive-form decision problems, while at the same time going beyond low-degree swap deviations.

**Theorem 1.2** (Online learning; precise version in [Theorem 6.4](#)). *Suppose that  $\mathcal{X} \subseteq \mathbb{R}^d$  admits a membership oracle and  $\Phi^m$  is  $k$ -dimensional per [Definition 1.1](#). There is an online algorithm that guarantees at most  $\epsilon$  average  $\Phi^m$ -regret after  $\text{poly}(k)/\epsilon^2$  rounds with  $\text{poly}(k, 1/\epsilon)$  running time.*

### 1.1.2 Computation

The result above holds even when the learner is facing an adversary, thereby being readily applicable when learning in  $n$ -player *multilinear games*. In such games, each player  $i \in [n]$  has a convex and compact strategy set  $\mathcal{X}_i \subseteq \mathbb{R}^{d_i}$  and utility function  $u_i : \mathcal{X}_1 \times \dots \times \mathcal{X}_n \rightarrow \mathbb{R}$  that is linear in  $\mathcal{X}_i$ , so that  $u_i(\mathbf{x}) = \langle \mathbf{g}_i, \mathbf{x}_i \rangle$  for some  $\mathbf{g}_i = \mathbf{g}_i(\mathbf{x}_{-i}) \in \mathbb{R}^{d_i}$ . (Extensive-form games constitute a canonical example of this framework.) In this context, [Theorem 1.2](#) implies a fully polynomial-time algorithm (FPTAS) for computing  $\epsilon$ -approximate  $\Phi^m$ -equilibria in convex games. Our next result establishes a polynomial-time algorithm—that is, with running time growing polynomially in  $\log(1/\epsilon)$  as opposed to  $1/\epsilon$ —for that problem.

**Theorem 1.3** (Computation; precise version in [Theorem 5.1](#)). *Consider an  $n$ -player multilinear game  $\Gamma$  such that, for each player  $i \in [n]$ , we are given  $\text{poly}(n, k)$ -time algorithms for the following:*

- *an oracle to compute the gradient, that is, the vector  $\mathbf{g}_i = \mathbf{g}_i(\mathbf{x}_{-i}) \in \mathbb{R}^{d_i}$  for which  $\langle \mathbf{g}_i(\mathbf{x}_{-i}), \mathbf{x}_i \rangle = u_i(\mathbf{x})$  for all  $\mathbf{x} \in \mathcal{X}_1 \times \dots \times \mathcal{X}_n$  (polynomial expectation property); and*
- *a membership oracle for the strategy set  $\mathcal{X}_i$ .*

*Suppose further that each  $\Phi^{m_i}$  is  $k_i$ -dimensional per [Definition 1.1](#) and  $\|\mathbf{g}_i\| \leq B$ . Then, an  $\epsilon$ -approximate  $\Phi^m$ -equilibrium of  $\Gamma$  can be computed in  $\text{poly}(n, k, \log(B/\epsilon))$  time.*

## 1.2 Technical approach

[Theorems 1.2](#) and [1.3](#) build on and extend certain recent developments due to Daskalakis et al. [18] and Zhang et al. [59]. Below, we outline our key technical contributions.

**Expected fixed points** Our approach crucially hinges on the notion of an *expected fixed point*: a distribution  $\mu \in \Delta(\mathcal{X})$  such that  $\mathbb{E}_{\mathbf{x} \sim \mu}[\phi(\mathbf{x}) - \mathbf{x}] \approx 0$ . Taking a step back, earlier approaches for minimizing  $\Phi$ -regret were based on computing an actual fixed point of a function  $\phi \in \Phi$  [29, 6, 56]. For normal-form games, computing a fixed point of such a function boils down to determining the stationary distribution of a certain Markov chain, which is directly amenable to linear programming—this holds more generally when  $\Phi$  contains linear functions. However, when considering nonlinear functions, this standard approach immediately becomes prohibitive since fixed points are hard to compute. Zhang et al. [59] bypass this obstacle by showing that a fixed point *in expectation*, as introduced earlier, is actually sufficient for minimizing  $\Phi$ -regret. In fact, building on an earlier result due to Hazan and Kale [35], Zhang et al. [59] observed that approximating expected fixed points also reduces to minimizing  $\Phi$ -regret, establishing a certain equivalence between the two. Crucially, there is a simple,  $O(1/\epsilon)$ -time algorithm for computing an  $\epsilon$ -expected fixed point of any  $\phi \in \Phi$  by taking the uniform distribution over the sequence  $\mathbf{x}, \phi(\mathbf{x}), \phi(\phi(\mathbf{x})), \dots$ ; the claimed guarantee follows by a telescopic summation. In this context, an important question left open by Zhang et al. [59]—which, as we shall see, is the crux in proving [Theorem 1.3](#)—concerns the complexity of computing expected fixed points in the regime where  $\epsilon$  is exponentially small. We address this question by showing that expected fixed points can be computed in time polynomial in the dimension and  $\log(1/\epsilon)$ .

**Theorem 1.4** (Expected fixed points). *Given oracle access to  $\mathcal{X}$  and  $\phi : \mathcal{X} \rightarrow \mathcal{X}$ , there is a  $\text{poly}(d, \log(1/\epsilon))$ -time algorithm that computes an  $\epsilon$ -expected fixed point of  $\phi$ .*

We have described so far the role of expected fixed points when learning in an online environment (*cf.* [Theorem 1.2](#)). Going back to [Theorem 1.3](#), expected fixed points also serve a crucial purpose in that context. Namely, [Theorem 1.3](#) relies on the *ellipsoid against hope (EAH)* algorithm of Papadimitriou and Roughgarden [51], which in turn is based on running the ellipsoid algorithm on an infeasible program—the rationale being that a correlated equilibrium can be extracted from the certificate of infeasibility of that program. Now, to execute ellipsoid, one needs a separation oracle. For normal-form games, this amounts to a fixed point oracle: for any  $\phi$ , compute  $\mathbf{x} \in \mathcal{X}$  such that  $\phi(\mathbf{x}) = \mathbf{x}$ . And, as we saw earlier,  $\phi$  is just a stochastic matrix, and so it suffices to identify a stationary distribution of the corresponding Markov chain.

However, there are two main obstacles, which manifest themselves in each iteration of the ellipsoid, when  $\mathcal{X}$  is a general convex set and  $\Phi$  is allowed to contain nonlinear endomorphisms:

1. computing a fixed point of a nonlinear  $\phi$  is intractable; and
2. separating over the set  $\Phi$  is also intractable even with respect to linear endomorphisms [18], let alone under [Definition 1.1](#).

With regard to [Item 1](#), we show that, during the execution of the ellipsoid, one might as well use *expected* fixed points, which are tractable by virtue of [Theorem 1.4](#) we described earlier. What is intriguing is that our proof of [Theorem 1.4](#) also relies on (a different instantiation of) EAH, and so the overall algorithm that we develop uses EAH in a nested fashion—each separation oracle as part of the execution of EAH is internally implemented via EAH!

For [Item 2](#), we build on the framework of Daskalakis et al. [18]. In light of the inability to efficiently separate over linear endomorphisms, they observed that one can still execute EAH with access to a weaker oracle, which they refer to as a *semi-separation oracle*. Moreover, they developed a polynomial-time semi-separation oracle with respect to the set of linear endomorphisms. Building

on [Theorem 1.4](#), we significantly extend their result, establishing a semi-separation oracle for general functions, not just linear ones.

**Theorem 1.5** (Semi-separation oracle for general functions). *Given oracle access to  $\mathcal{X}$  and  $\phi: \mathcal{X} \rightarrow \mathbb{R}^d$ , there is a  $\text{poly}(d, \log(1/\epsilon))$ -time algorithm that either computes an  $\epsilon$ -expected fixed point of  $\phi$ , or identifies a point  $\mathbf{x} \in \mathcal{X}$  such that  $\phi(\mathbf{x}) \notin \mathcal{X}$ .*

(In particular, in the usual case where  $\phi$  maps to  $\mathcal{X}$ , the algorithm above always returns an  $\epsilon$ -expected fixed point of  $\phi$ .)

We now turn to [Theorem 1.2](#), which revolves around the online learning setting. Equipped with our semi-separation oracle for general functions, we show that the framework of Daskalakis et al. [18] can be extended from linear endomorphisms to ones satisfying [Definition 1.1](#). The technical pieces underpinning [Theorem 1.2](#) are exposed in depth in [Section 6](#). Importantly, the dimension of  $\Phi$  turns out to be a fundamental barrier for no-regret learning in the following sense.

**Theorem 1.6.** *For any  $k$  and any  $d \geq \Theta(\log^{14} k)$ , there is an online decision problem with dimension  $d$  and an adversary such that the  $\Phi$ -regret of the learner with respect to a  $k$ -dimensional  $\Phi$  is at least  $\epsilon$  when  $T < \min\{\sqrt{k}/4, \exp(\Omega(\epsilon^{-1/6}))\}$ .*

This is a straightforward refinement of recent lower bounds [17, 14, 52]. What is important is that, up to constant factors in the exponent of  $k$ , [Theorem 1.6](#) matches the upper bound of [Theorem 1.2](#). In doing so, we establish for the first time a class of deviations that characterizes—in the previous sense—no- $\Phi$ -regret learning in the adversarial setting.

Finally, we remark that we did not attempt to optimize the polynomial dependencies on  $d$  and  $k$  throughout this paper; improving the overall complexity is an interesting direction for future work.

### 1.3 Further related work

The existence of no-regret algorithms goes back to the pioneering work of Blackwell [5]; the stronger notion of swap regret was crystallized and analyzed more recently [6, 56, 34]. Part of the impetus of that line of work revolves around the connection to correlated equilibria, highlighted earlier. Unfortunately, beyond online decision problems on the simplex, such no-swap-regret algorithms become inefficient when the number of pure strategies is exponential in the natural parameters of the problem—as is the case, for example, in Bayesian games, wherein  $\mathcal{X} := [0, 1]^d$ . Recent breakthrough results by Dagan et al. [14] and Peng and Rubinfeld [52] establish a new algorithmic paradigm for minimizing swap regret, applicable even when the number of pure strategies is exponential. However, that comes at the expense of introducing an exponential dependence on  $1/\epsilon$ , which is unavoidable in the adversarial regime [17]. Our main interest here is in online algorithms with complexity scaling polynomially in both the dimension and  $1/\epsilon$ .

Besides the game-theoretic implications concerning convergence to correlated equilibria, swap regret is a fundamental concept in its own right, being intimately tied to the notion of *calibration*; namely, it has been known since the foundational work of Foster and Vohra [27] that best-responding to calibrated forecasters guarantees no-swap-regret (*cf.* Foster and Hart [28]); in relation to that connection, it is worth noting an important, recent body of work that bypasses the intractability of calibration in high dimensions [50, 53, 36]. Swap regret is also more robust against exploitation, in a sense formalized in a series of recent papers [1, 43, 19, 33].

In particular, within that line of work, Mansour et al. [43] introduced the notion of *polytope swap regret*, which comprises deviations that allow the vertices of the underlying polytope to be

swapped with each other—points inside the polytope are mapped in accordance with the (worst-case) convex combination of vertices. It is currently unknown whether there is an efficient algorithm for minimizing polytope swap regret.

The more flexible framework of  $\Phi$ -regret, which has been gaining considerable traction in recent years, allows one to circumvent the recent lower bound of Daskalakis et al. [17] by restricting the set of deviations. In addition to the research highlighted above, chiefly in the context of extensive-form games, we now provide some further pointers for the interested reader. Bernasconi et al. [4] considered the more challenging setting of so-called “pseudo-games,” wherein players have joint constraint sets.  $\Phi$ -equilibria in such settings have certain counterintuitive properties; for example, they are not necessarily convex.  $\Phi$ -equilibria have also garnered attention in the context of Markov (aka. stochastic) games, going back to the work of Greenwald and Hall [30]; for more recent research, we refer to Jin et al. [39], Erez et al. [21], Cai et al. [12], and references therein. Even more broadly, we refer to Cai et al. [11], Şeref Ahunbay [62] for efficient solution concepts in nonconcave games [15].

Finally, we remark that the hardness result of Daskalakis et al. [18] for separating over linear endomorphisms does not apply to polytopes represented with a polynomial number of constraints. Indeed, it is relatively straightforward to implement a membership oracle for such polytopes [18]. In contrast, it is computationally hard to decide membership for low-degree polynomials [59].

## 2 Preliminaries

**Notation** We use boldface lowercase letters, such as  $\mathbf{x}$  and  $\mathbf{y}$ , to denote vectors in a Euclidean space. Matrices are represented with capital boldface letters, such as  $\mathbf{K}$ . For a vector  $\mathbf{x} \in \mathbb{R}^d$ , we denote by  $\|\mathbf{x}\| := \sqrt{\langle \mathbf{x}, \mathbf{x} \rangle}$  its Euclidean norm, where  $\langle \cdot, \cdot \rangle$  is the inner product. The  $j$ th coordinate of  $\mathbf{x}$  is accessed by  $\mathbf{x}[j]$ . For a matrix  $\mathbf{K}$ , we use  $\|\mathbf{K}\|_F$  to denote its Frobenius norm.  $\mathbf{I}_{d \times d} \in \mathbb{R}^{d \times d}$  represents the identity matrix.  $\mathcal{B}_r(\mathbf{x})$  is the (closed) Euclidean ball centered at  $\mathbf{x}$  with radius  $r > 0$ . An *endomorphism* on  $\mathcal{X}$  is a function mapping  $\mathcal{X}$  to  $\mathcal{X}$ .

**Oracle access** Throughout this paper, we assume that we have access to the convex and compact constraint set  $\mathcal{X}$  via an oracle. (For multi-player games, oracle access is posited for the constraint set  $\mathcal{X}_i$  of each player, which can be thereby extended to  $\mathcal{X} := \mathcal{X}_1 \times \dots \times \mathcal{X}_n$ .) In particular, the following three types of oracles are commonly considered in the literature:

- *membership*: given a point  $\mathbf{x} \in \mathbb{R}^d$ , decide whether  $\mathbf{x} \in \mathcal{X}$ ;
- *separation*: given a point  $\mathbf{x} \in \mathbb{R}^d$ , decide whether  $\mathbf{x} \in \mathcal{X}$ , and if not, output a hyperplane  $\mathbf{w} \in \mathbb{R}^d$  separating  $\mathbf{x}$  from  $\mathcal{X}$ :  $\langle \mathbf{x}, \mathbf{w} \rangle > \langle \mathbf{x}', \mathbf{w} \rangle$  for all  $\mathbf{x}' \in \mathcal{X}$ ;
- *linear optimization*: given a point  $\mathbf{u} \in \mathbb{R}^d$ , output any point in  $\operatorname{argmax}_{\mathbf{x} \in \mathcal{X}} \langle \mathbf{x}, \mathbf{u} \rangle$ .

Under the assumption that  $\mathcal{B}_r(\cdot) \subseteq \mathcal{X} \subseteq \mathcal{B}_R(\mathbf{0})$ , the three oracles described above are known to be (polynomially) equivalent [32]—up to logarithmic factors in  $R$  and  $1/r$ . The previous geometric condition can always be met by bringing  $\mathcal{X}$  into *isotropic position*, which means that, for a uniformly sampled  $\mathbf{x} \sim \mathcal{X}$ , we have  $\mathbb{E}[\mathbf{x}] = \mathbf{0}$  and  $\mathbb{E}[\mathbf{x}\mathbf{x}^\top] = \mathbf{I}_{d \times d}$ . This can be achieved in polynomial time through an affine transformation [42, 7, 40]; it is easy to see that minimizing  $\Phi$ -regret after applying that transformation suffices in order to minimize  $\Phi$ -regret in the original space (formally shown in [Appendix A](#)). As a result, we can assume throughout that, for example,  $\mathcal{B}_1(\mathbf{0}) \subseteq \mathcal{X} \subseteq \mathcal{B}_d(\mathbf{0})$  [42].



**Remark 2.1** (Weak oracles). When dealing with general convex sets, the oracles posited above can return points supported on irrational numbers. To address this issue in the usual Turing model of computation, it suffices to consider weaker versions of those oracles that allow for some small slackness  $\epsilon > 0$ . Our analysis in the sequel can be extended to account for such imprecision.

## 2.1 Online learning and $\Phi$ -regret

In the usual framework of online learning, a learner interacts with an environment over a sequence of  $T \in \mathbb{N}$  rounds. In each round  $t \in [T]$ , the learner selects a strategy  $\mathbf{x}^{(t)} \in \mathcal{X}$ , and then observes as feedback from the environment a utility function  $\mathbf{x} \mapsto \langle \mathbf{x}, \mathbf{u}^{(t)} \rangle$ , for some utility vector  $\mathbf{u}^{(t)} \in [-1, 1]^d$ ; the utility of the learner in the  $t$ th round is given by  $\langle \mathbf{x}^{(t)}, \mathbf{u}^{(t)} \rangle$ . For the purpose of our work, we will allow the learner to output a *mixed* strategy,  $\mu^{(t)} \in \Delta(\mathcal{X})$ , so that the (expected) utility at the  $t$ th round reads  $\mathbb{E}_{\mathbf{x}^{(t)} \sim \mu^{(t)}} \langle \mathbf{x}^{(t)}, \mathbf{u}^{(t)} \rangle$ . As shown by Zhang et al. [59], restricting the learner to output strategies in  $\mathcal{X}$ —as opposed to  $\Delta(\mathcal{X})$ —makes the problem of minimizing  $\Phi$ -regret PPAD-hard, even with respect to low-degree polynomials, and so employing mixed strategies will be essential for our purposes.

In this context, a canonical measure of performance in online learning is  $\Phi$ -regret, defined as

$$\Phi\text{-Reg}^{(T)} := \sup_{\phi \in \Phi} \sum_{t=1}^T \left\langle \mathbf{u}^{(t)}, \mathbb{E}_{\mathbf{x}^{(t)} \sim \mu^{(t)}} [\phi(\mathbf{x}^{(t)}) - \mathbf{x}^{(t)}] \right\rangle.$$

The *average*  $\Phi$ -regret is defined as  $\frac{1}{T} \Phi\text{-Reg}^{(T)}$ . Perhaps the most common instantiation of  $\Phi$ -regret is *external* regret, whereby  $\Phi$  contains solely constant transformations. We are interested in characterizing the broadest set of deviations  $\Phi$  that allows for efficient learnability. Our starting point is the template of Gordon et al. [29].

**The algorithm of Gordon et al. [29]** Gordon et al. [29] (*cf.* Blum and Mansour [6], Stoltz and Lugosi [56]) crystallized a basic template for minimizing  $\Phi$ -regret. It comprises two basic components:

1. a fixed-point oracle  $\text{FP}(\phi)$  that takes as input any transformation  $\Phi \ni \phi : \mathcal{X} \rightarrow \mathcal{X}$  and outputs a fixed point thereof; that is, a point  $\mathbf{x} \in \mathcal{X}$  such that  $\phi(\mathbf{x}) = \mathbf{x}$ .
2. an *external* regret minimizer  $\mathfrak{R}_\Phi$  operating over the set  $\Phi$ .

With access to the above components, a  $\Phi$ -regret minimizer  $\mathfrak{R}$  operating over  $\mathcal{X}$ —without the need to resort to mixed strategies—can be constructed as follows. At any time  $t \in [T]$ , upon selecting a strategy  $\mathbf{x}^{(t)} \in \mathcal{X}$  and observing  $\mathbf{u}^{(t)}$ , provide as input to  $\mathfrak{R}_\Phi$  the utility function  $\phi \mapsto \langle \mathbf{u}^{(t)}, \phi(\mathbf{x}^{(t)}) \rangle$ . Suppose that  $\phi^{(t+1)} \in \Phi$  is the next strategy of  $\mathfrak{R}_\Phi$ . The learner  $\mathfrak{R}$  can then output as its next strategy  $\mathbf{x}^{(t+1)}$  any fixed point of  $\phi^{(t+1)}$ ; that is,  $\mathbf{x}^{(t+1)} := \text{FP}(\phi^{(t+1)})$ . By definition, it follows that the  $\Phi$ -regret of  $\mathfrak{R}$  is equal to the external regret of  $\mathfrak{R}_\Phi$  (*cf.* Theorem 2.2).

However, our main result in the online learning setting hinges on relaxing both oracles posited in [Items 1 and 2](#) in the framework of Gordon et al. [29]. With regard to [Item 1](#), when operating over *mixed* strategies, Zhang et al. [59] observed that it suffices to output an  $\epsilon$ -*expected fixed point* (EFP) of  $\phi$ , that is, a distribution  $\mu$  such that  $\|\mathbb{E}_{\mathbf{x} \sim \mu} [\phi(\mathbf{x}) - \mathbf{x}]\|_1 \leq \epsilon$ . Unlike actual fixed points, which are marred by computational intractability, Zhang et al. [59] observed that there is a simple,

$O(1/\epsilon)$ -time algorithm for computing an  $\epsilon$ -expected fixed point: simply take the uniform distribution over the sequence  $\mathbf{x}, \phi(\mathbf{x}), \phi(\phi(\mathbf{x})), \dots$  for  $O(1/\epsilon)$  steps. (In fact, one of our main results—namely, [Theorem 1.4](#)—provides a polynomial-time algorithm for that problem.) The overall scheme resulting from replacing [Item 1](#) with an approximate EFP is given in [Algorithm 1](#).

---

**ALGORITHM 1:** Minimizing  $\Phi$ -regret with EFPs [29, 59]

---

**Input:**

- An external regret minimizer  $\mathfrak{R}_\Phi$  for the set  $\Phi$
- A convex and compact strategy set  $\mathcal{X}$
- Precision parameter  $\epsilon > 0$

**Output:** An  $\Phi$ -regret minimizer  $\mathfrak{R}$  for the set  $\mathcal{X}$ ;

Initialize  $\phi^{(1)} \in \Phi$ ;

**for**  $t = 1, \dots, T$  **do**

- Set  $\mu^{(t)} \in \Delta(\mathcal{X})$  to be an  $\epsilon$ -expected fixed point of  $\phi^{(t)}$ ;
- Output  $\mu^{(t)} \in \Delta(\mathcal{X})$  as the next mixed strategy of  $\mathfrak{R}$ ;
- Receive as feedback  $\mathbf{u}^{(t)}$ ;
- Give as input to  $\mathfrak{R}_\Phi$  the utility function  $\phi \mapsto \mathbb{E}_{\mathbf{x}^{(t)} \sim \mu^{(t)}} \langle \phi(\mathbf{x}^{(t)}), \mathbf{u}^{(t)} \rangle$ ;
- Let  $\phi^{(t+1)} \in \Phi$  be the next strategy of  $\mathfrak{R}_\Phi$ ;

---

**Theorem 2.2** ([29, 59]). *Let  $\Phi\text{-Reg}^{(T)}$  be the  $\Phi$ -regret of  $\mathfrak{R}$  and  $\text{Reg}_\Phi^{(T)}$  the external regret of  $\mathfrak{R}_\Phi$  in [Algorithm 1](#) with precision  $\epsilon > 0$ . Then, for any  $T \in \mathbb{N}$ ,*

$$\Phi\text{-Reg}^{(T)} \leq \text{Reg}_\Phi^{(T)} + \epsilon T.$$

In particular, taking, say,  $\epsilon \propto 1/\sqrt{T}$ , [Theorem 2.2](#) reduces  $\Phi$ -regret minimization on  $\mathcal{X}$  to external regret minimization on  $\Phi$ .

*Proof of [Theorem 2.2](#).* For any  $\phi \in \Phi$ , we have

$$\sum_{t=1}^T \left\langle \mathbf{u}^{(t)}, \mathbb{E}_{\mathbf{x}^{(t)} \sim \mu^{(t)}} [\phi(\mathbf{x}^{(t)}) - \mathbf{x}^{(t)}] \right\rangle \leq \sum_{t=1}^T \mathbb{E}_{\mathbf{x}^{(t)} \sim \mu^{(t)}} \langle \phi(\mathbf{x}^{(t)}) - \phi^{(t)}(\mathbf{x}^{(t)}), \mathbf{u}^{(t)} \rangle + \epsilon T, \quad (1)$$

where we used the fact that  $\mu^{(t)}$  is an  $\epsilon$ -expected fixed point of  $\phi^{(t)}$  for all  $t \in [T]$ . The right-hand side of (1) can be in turn bounded by the external regret of  $\mathfrak{R}_\Phi$  plus the slackness term  $\epsilon T$ .  $\square$

This paradigm for minimizing  $\Phi$ -regret has been ubiquitous in prior work in this area. And yet, [Daskalakis et al. \[18\]](#) recently demonstrated that it is insufficient even when  $\Phi$  contains all linear endomorphisms of a general convex set. [Section 6](#) covers in detail the framework of [Daskalakis et al. \[18\]](#)—relaxing [Item 2](#) of [Gordon et al. \[29\]](#)—that will be the basis for our approach as well.

**$\Phi$ -equilibria** As we highlighted earlier, there is a celebrated connection between  $\Phi$ -regret and the game-theoretic solution concept of (correlated)  $\Phi$ -*equilibrium*. More precisely, in the context of multilinear games as introduced in [Section 1.1](#), we recall the following central definition [[57](#), [31](#)].

**Definition 2.3.** An  $\epsilon$ -approximate  $\Phi$ -*equilibrium* of an  $n$ -player multilinear  $\Gamma$  is a (correlated) distribution  $\mu \in \Delta(\mathcal{X}_1 \times \dots \times \mathcal{X}_n)$  such that for every player  $i \in [n]$  and deviation  $\phi_i \in \Phi_i \subseteq \mathcal{X}_i^{\mathcal{X}_i}$ ,

$$\mathbb{E}_{\mathbf{x} \sim \mu} [u_i(\phi_i(\mathbf{x}_i), \mathbf{x}_{-i}) - u_i(\mathbf{x})] \leq \epsilon.$$

A direct consequence of this definition is that if players repeatedly interact in a game and all incur sublinear  $\Phi$ -regret, the average distribution of play converges to the set of  $\Phi$ -equilibria.

## 2.2 Ellipsoid against hope

The *ellipsoid against hope (EAH)* algorithm was famously introduced by Papadimitriou and Roughgarden [51] to compute correlated equilibria in succinct, multi-player games—under the polynomial expectation property. A further crucial assumption in the approach of Papadimitriou and Roughgarden [51] is that the game is of *polynomial type*, in that the number of actions (or pure strategies) is polynomial in the representation of the game. In contrast to normal-form games, extensive-form games—and many other natural classes of games—are *not* of polynomial type. Farina and Pipis [24] recently showed how to apply EAH in the context of extensive-form games—albeit only for LCE; as we have seen, the complexity of NFCE remains open. We begin by recalling their framework, which crystallizes the approach of Papadimitriou and Roughgarden [51]. We then proceed by introducing the more powerful approach of Daskalakis et al. [18], which is crucial to compute LCE under general convex constraint sets, and which will form the basis for our approach as well.

Consider an arbitrary optimization problem of the form

$$\text{find } \mu \in \Delta(\mathcal{X}) \quad \text{s.t.} \quad \mathbb{E}_{\mathbf{x} \sim \mu} \langle \mathbf{y}, G(\mathbf{x}) \rangle \geq 0 \quad \forall \mathbf{y} \in \mathcal{Y}, \quad (2)$$

where  $\mathcal{X} \subseteq \mathbb{R}^d$ ,  $\mathcal{Y} \subseteq \mathbb{R}^k$ , and  $G : \mathcal{X} \rightarrow \mathbb{R}^k$  is a function such that  $\|G(\mathbf{x})\| \leq B$  for all  $\mathbf{x} \in \mathcal{X}$ . The crux in (2) lies in the fact that  $\mu$  resides in a high-dimensional (indeed, an infinite-dimensional) space, making standard approaches of little use. EAH addresses that challenge, as we describe next.

Suppose that we are given a  $\text{poly}(d, k)$ -time evaluation oracle for  $G$  and a separation oracle (SEP) for  $\mathcal{Y}$ , assumed to be *well-bounded*:  $\mathcal{B}_r(\cdot) \subseteq \mathcal{Y} \subseteq \mathcal{B}_R(\mathbf{0})$ . In addition, we assume that we have access to a *good-enough-response* (GER) oracle, which, given any  $\mathbf{y} \in \mathcal{Y}$ , returns  $\mathbf{x} \in \mathcal{X}$  such that  $\langle \mathbf{y}, G(\mathbf{x}) \rangle \geq 0$ . The EAH algorithm allows us to solve problems of the form (2) with just the above tools. In particular, EAH proceeds by considering an  $\epsilon$ -approximate version of the dual of (2).

$$\text{find } \mathbf{y} \in \mathcal{Y} \quad \text{s.t.} \quad \langle \mathbf{y}, G(\mathbf{x}) \rangle \leq -\epsilon \quad \forall \mathbf{x} \in \mathcal{X}. \quad (3)$$

Since a GER oracle exists, (3) is infeasible. Moreover, a certificate of infeasibility of (3) provides an  $\epsilon$ -approximate solution to (2). Thus, it suffices to run the ellipsoid algorithm on (3) and extract a certificate of infeasibility. This is precisely what EAH does, as formalized in [Theorem 2.4](#); the overall scheme is [Algorithm 8](#) in [Appendix C](#) ([Algorithm 2](#) below is a more general version thereof).

**Theorem 2.4** (Generalized form of EAH [24, 51]). *Suppose that we have  $\text{poly}(d, k)$ -time algorithms for the following:*

- an evaluation oracle for  $G$ , where  $\|G(\mathbf{x})\| \leq B$  for all  $\mathbf{x} \in \mathcal{X}$ ;
- a GER oracle for (2); and
- a separation oracle (SEP) for the well-bounded set  $\mathcal{Y}$ .

*Then, there is an algorithm that runs in time  $\text{poly}(d, k, \log(B/\epsilon))$  and returns an  $\epsilon$ -approximate solution to (2).*

However, when it comes to computing LCE under general constraint sets, [Theorem 2.4](#) is not enough: Daskalakis et al. [18] showed that separating over  $\mathcal{Y}$ —the set of linear endomorphisms—is hard. In light of this fact, their key observation was that an  $\epsilon$ -approximate solution to (2) can still

---

**ALGORITHM 2:** Ellipsoid against hope (EAH) under GERorSEP oracle [18]

---

**Input:**

- Parameters  $R_y, r_y > 0$  such that  $\mathcal{B}_{r_y}(\cdot) \subseteq \mathcal{Y} \subseteq \mathcal{B}_{R_y}(\mathbf{0})$
- Precision parameter  $\epsilon > 0$
- Parameter  $B > 0$  such that  $\|G(\mathbf{x})\| \leq B$  for all  $\mathbf{x} \in \mathcal{X}$
- A GERorSEP oracle ([Definition 2.5](#))

**Output:** A sparse,  $\epsilon$ -approximate solution  $\mu \in \Delta(\mathcal{X})$  of (2);

Initialize the ellipsoid  $\mathcal{E} := \mathcal{B}_{R_y}(\mathbf{0})$ ;

Initialize  $\tilde{\mathcal{Y}} := \mathcal{B}_{R_y}(\mathbf{0})$ ;

**while**  $\text{vol}(\mathcal{E}) \geq \text{vol}(\mathcal{B}_{\epsilon/B}(\cdot))$  **do**

Query the GERorSEP oracle on the center of  $\mathcal{E}$ ;

**if** it returns a good-enough-response  $\mathbf{x} \in \mathcal{X}$  **then**

Update  $\mathcal{E}$  to the minimum volume ellipsoid containing  $\mathcal{E} \cap \{\mathbf{y} \in \mathbb{R}^k : \langle \mathbf{y}, G(\mathbf{x}) \rangle \leq 0\}$ ;

**else**

Let  $H$  be the halfspace that separates  $\mathbf{y}$  from  $\mathcal{Y}$ ;

Update  $\mathcal{E}$  to the minimum volume ellipsoid containing  $\mathcal{E} \cap H$ ;

Update  $\tilde{\mathcal{Y}} := \tilde{\mathcal{Y}} \cap H$ ;

Let  $\mathbf{x}^{(1)}, \dots, \mathbf{x}^{(T)}$  be the GER oracle responses produced in the process above;

Define  $\mathbf{G} := [G(\mathbf{x}^{(1)}) \mid \dots \mid G(\mathbf{x}^{(T)})] \in \mathbb{R}^{k \times T}$ ;

Compute a solution  $\boldsymbol{\lambda}$  to the convex program

$$\text{find } \boldsymbol{\lambda} \in \Delta^T \quad \text{s.t.} \quad \min_{\mathbf{y} \in \tilde{\mathcal{Y}}} \boldsymbol{\lambda}^\top \mathbf{G}^\top \mathbf{y} \geq -\epsilon$$

**return**  $\Delta(\mathcal{X}) \ni \mu := \sum_{t=1}^T \lambda^{(t)} \mu(\mathbf{x}^{(t)})$

---

be computed given access to a weaker oracle. Namely, instead of requiring both a GER and a SEP oracle, as in [Theorem 2.4](#), Daskalakis et al. [18] showed that it suffices to implement the following oracle: for any given  $\mathbf{y} \in \mathbb{R}^k$  (not necessarily in  $\mathcal{Y}$ ),

1. *either* compute a good-enough response  $\mathbf{x} \in \mathcal{X}$ ,
2. *or* a hyperplane separating  $\mathbf{y}$  from  $\mathcal{Y}$ .

Although separating over  $\mathcal{Y}$  is hard, this weaker oracle suffices to recover the guarantee of [Theorem 2.4](#), and this is enough to compute linear correlated equilibria in games. Yet, for our purposes, it will be necessary to relax the aforescribed oracle even further, as formalized below.

**Definition 2.5** (GERorSEP). Consider problem (2). The oracle  $\epsilon$ -GERorSEP works as follows. It takes as input  $\mathbf{y} \in \mathbb{R}^k$ , and it

1. *either* computes an  $\epsilon$ -approximate good-enough-response  $\mu \in \Delta(\mathcal{X})$ ,  $\mathbb{E}_{\mathbf{x} \sim \mu} \langle \mathbf{y}, G(\mathbf{x}) \rangle \geq -\epsilon$ , such that  $\text{supp}(\mu) \leq \text{poly}(d, k, \log(1/\epsilon))$ ,
2. *or* a hyperplane  $\epsilon$ -approximately separating  $\mathbf{y}$  from  $\mathcal{Y}$ .

Compared to the oracle described earlier ([Items 1 and 2](#)), [Definition 2.5](#) makes two further concessions: first, the good-enough-response can now be a distribution, so long as it has polynomial support; and second, both GER and SEP can have some small slack  $\epsilon > 0$ . Both of those relaxations will be essential for our applications. We now summarize the key guarantee.

**Theorem 2.6** ([18]; generalization of [Theorem 2.4](#)). *Suppose that we have  $\text{poly}(d, k, \log(1/\epsilon))$ -time algorithms for the following:*

- *an  $\epsilon$ -GERorSEP oracle with respect to the well-bounded set  $\mathcal{Y}$ , and*
- *an evaluation oracle for  $G$ , where  $\|G(\mathbf{x})\| \leq B$  for all  $\mathbf{x} \in \mathcal{X}$ .*

*Then, there is an algorithm that runs in time  $\text{poly}(d, k, \log(B/\epsilon))$  and returns an  $\epsilon$ -approximate solution to (2).*

[Algorithm 2](#) depicts the overall scheme under a GERorSEP oracle. (The last line of the algorithm uses the notation  $\mu(\mathbf{x})$  for the distribution supported solely on  $\mathbf{x} \in \mathcal{X}$ .)

### 3 Sets of deviations with polynomial dimension

In this section, we formally introduce the assumptions we make concerning the feature map  $m$  of [Definition 1.1](#), and we then provide a canonical example that satisfies our blanket assumptions.

**Assumption 3.1.** We make the following assumptions regarding  $\Phi^m$  and  $m$  of [Definition 1.1](#):

- $m : \mathcal{X} \rightarrow \mathbb{R}^{k'}$  is computable in  $\text{poly}(k)$  time.
- $\text{conv } m(\mathcal{B}_1(\mathbf{0})) \supseteq \mathcal{B}_\delta(\mathbf{0})$  for some  $\delta \geq \text{poly}(1/k)$ .
- $\|m(\mathbf{x})\| \leq \text{poly}(k)$  for all  $\mathbf{x} \in \mathcal{X}$ , and  $m(\mathbf{0}) = \mathbf{0}$ .
- $\Phi^m$  contains the identity map.

**Remark 3.2** (Functions on the vertices). Let  $\mathcal{V}$  be the set of extreme points of  $\mathcal{X}$ . Our positive results ([Theorems 5.1](#) and [6.4](#)) only evaluate  $\phi$  at extreme points, so they would operate identically if we instead defined our maps  $\phi$  to be  $\mathcal{V} \rightarrow \mathcal{X}$ .

The definition above places some minimal assumptions on the feature mapping  $m$  to ensure that  $\Phi^m$  is geometrically well behaved. Indeed, we first show that the set of transformations  $\Phi^m$  under [Assumption 3.1](#) is well-bounded; the proof is provided in [Appendix B](#).

**Lemma 3.3.** *Let  $\mathcal{X} \subseteq \mathbb{R}^d$  be a convex and compact set such that  $\mathcal{B}_r(\mathbf{0}) \subseteq \mathcal{X} \subseteq \mathcal{B}_R(\mathbf{0})$ , with  $R \geq 1$  and  $r < R$ . Suppose further that  $\|m(\mathbf{x})\| \leq M$  for all  $\mathbf{x} \in \mathcal{B}_R(\mathbf{0})$ , with  $M = M(R) \geq 1$ ;  $\text{conv } m(\mathcal{B}_r(\mathbf{0})) \supseteq \mathcal{B}_\delta(\mathbf{0})$  for some  $\delta = \delta(r) > 0$ ; and  $m(\mathbf{0}) = \mathbf{0}$ . Then,*

$$\mathcal{B}_{r'}(\mathbf{0}) \subseteq \Phi^m \subseteq \mathcal{B}_{R'}(\mathbf{0}),$$

where  $r' := r/2M(R)$  and  $R' := R\left(\frac{2\sqrt{d}}{\delta(r)} + 1\right)$ .

We are now ready to provide a canonical, concrete example of deviations that satisfy [Definition 1.1](#) under [Assumption 3.1](#). As we alluded to earlier in our introduction, it is the family of low-degree polynomials; in particular, it will be convenient to work with the Legendre basis.

**Definition 3.4.** Let  $P_0(x) = 1$  and  $P_1(x) = x$ . The  $(\ell + 1)$ th Legendre polynomial is given by the recurrence  $(\ell + 1)P_{\ell+1}(x) - (2\ell + 1)xP_\ell(x) + \ell P_{\ell-1}(x) = 0$ .

These polynomials have a convenient orthogonality property over  $[-1, 1]$ :

$$\int_{-1}^1 P_\ell(x) P_{\ell'}(x) dx = \begin{cases} \frac{1}{2\ell+1} & \text{if } \ell = \ell', \\ 0 & \text{otherwise.} \end{cases} \quad (4)$$

For convenience, we shall consider the rescaled polynomial  $\bar{P}_\ell := \sqrt{2\ell+1} P_\ell$ , so that  $\int_{-1}^1 \bar{P}_\ell(x)^2 dx = 1$ . We now define

$$m(\mathbf{x}) := \left( \prod_{j=1}^d \bar{P}_{\ell_j}(\mathbf{x}[j]) \right)_{1 \leq \ell_1 + \dots + \ell_d \leq \ell}. \quad (5)$$

We establish (in [Appendix C](#)) that [Assumption 3.1](#) encompasses the above mapping.

**Proposition 3.5.** *Let  $m : \mathcal{X} \rightarrow \mathbb{R}^{k'}$  per (5), where  $k' = \binom{d+\ell}{\ell} - 1$ .  $\bar{m} : \mathbf{x} \mapsto m(\sqrt{d}\mathbf{x})$  satisfies [Assumption 3.1](#) with  $M \leq d^{O(\ell)}$  and  $\delta = 1/M$ .*

## 4 Polynomial-time expected fixed points and semi-separation

We will now start connecting the framework we laid out in [Section 2.2](#) with the problem of computing  $\Phi$ -equilibria. That a  $\Phi$ -equilibrium can be cast as (2)—by linking  $\mathcal{Y}$  to the set of deviations  $\Phi$ —is not hard to see, and will be spelled out in the next section. This section concerns the question of implementing  $\epsilon$ -GERorSEP (per [Definition 2.5](#)), which is the main precondition of [Theorem 2.4](#).

The key to implementing the  $\epsilon$ -GERorSEP oracle, and the main subject of this section, is the notion of an  $\epsilon$ -expected fixed point (EFP) ([Definition 4.1](#)). This relaxes the usual notion of a fixed point that was employed by Zhang et al. [59], who observed that, for minimizing  $\Phi$ -regret, one can replace a fixed-point oracle—as in the canonical framework of Gordon et al. [29]—by an expected relaxation (per [Definition 4.1](#)) thereof. This is crucial because, unlike fixed points which are intractable beyond linear maps, there is a simple,  $O(1/\epsilon)$ -time algorithm for computing  $\epsilon$ -expected fixed points. When it comes to computing  $\Phi$ -equilibria in games, our contribution here is twofold.

1. we give a  $\text{poly}(d, \log(1/\epsilon))$ -time algorithm for computing an  $\epsilon$ -expected fixed point, and
2. we show that expected fixed points can be naturally coupled with the EAH framework, and in particular, with the recent generalization of Daskalakis et al. [18].

This section establishes [Item 1](#), while the next section formalizes [Item 2](#). Going back to [Section 2.2](#) and the GERorSEP oracle, the connection with (expected) fixed points lies in the observation that, when it comes to  $\Phi$ -equilibria in games, the GER part of the oracle can be implemented by computing an expected fixed point. This will become clear in the upcoming section.

**Definition 4.1** (Expected fixed points). Let  $\mathcal{X} \subseteq \mathbb{R}^d$  be convex and compact and a function  $\phi : \mathcal{X} \rightarrow \mathcal{X}$  to which we are given oracle access. The  $\epsilon$ -expected fixed point (EFP) problem asks for a distribution  $\mu \in \Delta(\mathcal{X})$  such that<sup>1</sup>

$$\left\| \mathbb{E}_{\mathbf{x} \sim \mu} [\phi(\mathbf{x}) - \mathbf{x}] \right\|_1 \leq \epsilon.$$

<sup>1</sup>The choice of the 1-norm (instead of, say, another  $p$ -norm) here is unimportant, because one can always take  $\epsilon$  to be exponentially small.

This definition departs from the usual notion of a fixed point by measuring the fixed-point error *in expectation* over samples  $\mathbf{x}$  from  $\mu$ . [Definition 4.1](#) is natural in its own right, but our key motivation is computational: as we shall see, an expected fixed point can be computed in polynomial time ([Theorem 4.3](#)).

We first observe that any function  $\phi : \mathcal{X} \rightarrow \mathcal{X}$  admits an *exact* expected fixed point; crucially, unlike Brouwer’s fixed-point theorem, we do not assume that  $\phi$  is continuous, and so expected fixed points exist even when fixed points do not.

**Proposition 4.2.** *Every function  $\phi : \mathcal{X} \rightarrow \mathcal{X}$  admits an exact solution to the EFP problem.*

*Proof.* Let  $G(\mathbf{x}) := \phi(\mathbf{x}) - \mathbf{x}$ . We want to show that  $\text{conv } G(\mathcal{X})$  contains the origin. Suppose not. Then, by the separating hyperplane theorem, there is a vector  $\mathbf{y} \in \mathbb{R}^d$  with  $\langle \mathbf{y}, G(\mathbf{x}) \rangle < 0$  for every  $\mathbf{x} \in \mathcal{X}$ . But this is impossible: simply consider  $\mathbf{x}^* \in \text{argmin}_{\mathbf{x} \in \mathcal{X}} \langle \mathbf{y}, \mathbf{x} \rangle$ .  $\square$

We now turn to our main computational result regarding EFPs; namely, a polynomial-time algorithm based on EAH—in particular, its incarnation in [Algorithm 8](#).

**Theorem 4.3.** *Given oracle access to  $\mathcal{X}$  and any  $\phi : \mathcal{X} \rightarrow \mathcal{X}$ , there exists a  $\text{poly}(d, \log(1/\epsilon))$ -time algorithm that computes an  $\epsilon$ -EFP of  $\phi$ .*

*Proof.* We observe that an EFP can be equivalently expressed through the optimization problem

$$\text{find } \mu \in \Delta(\mathcal{X}) \quad \text{s.t.} \quad \mathbb{E}_{\mathbf{x} \sim \mu} \langle \mathbf{y}, \phi(\mathbf{x}) - \mathbf{x} \rangle \geq 0 \quad \forall \mathbf{y} \in [-1, 1]^d.$$

We will now apply [Theorem 2.4](#). The set  $\mathcal{Y} := [-1, 1]^d$  clearly admits a separation oracle (SEP). Further, for any  $\mathbf{y} \in [-1, 1]^d$ , taking  $\mathbf{x}^* = \text{argmin}_{\mathbf{x} \in \mathcal{X}} \langle \mathbf{y}, \mathbf{x} \rangle$  (using an optimization oracle for  $\mathcal{X}$ ) guarantees  $\langle \mathbf{y}, \phi(\mathbf{x}^*) - \mathbf{x}^* \rangle \geq 0$  since  $\phi(\mathbf{x}^*) \in \mathcal{X}$ , thereby implementing the GER oracle. We thus find that the preconditions of [Theorem 2.4](#) are satisfied, and EAH ([Algorithm 8](#)) returns  $\mu \in \Delta(\mathcal{X})$ , with  $\text{supp}(\mu) \leq \text{poly}(d, \log(1/\epsilon))$ , such that

$$\mathbb{E}_{\mathbf{x} \sim \mu} \langle \mathbf{y}, \phi(\mathbf{x}) - \mathbf{x} \rangle \geq -\epsilon \quad \forall \mathbf{y} \in [-1, 1]^d.$$

Taking  $\mathbf{y} = \text{sign}(\mathbb{E}_{\mathbf{x} \sim \mu} (\mathbf{x} - \phi(\mathbf{x})))$  (coordinate-wise) completes the proof.  $\square$

As it will become clear, [Theorem 4.3](#) yields a polynomial-time implementation of the GER oracle in the context of [Section 2.2](#), which can be employed in EAH. With a slight modification in the proof of [Theorem 4.3](#), we shall see how one can also recover an  $\epsilon$ -GERorSEP oracle ([Definition 2.5](#)), which will then enable us to harness [Theorem 2.6](#) for computing  $\Phi$ -equilibria in games. Following the nomenclature of Daskalakis et al. [18], we refer to this oracle as a *semi-separation oracle*.

**Definition 4.4** (Semi-separation oracle). The *semi-separation* problem is the following. Given a convex and compact  $\mathcal{X}$  and a function  $\phi : \mathcal{X} \rightarrow \mathbb{R}^d$ , compute

1. *either* a distribution  $\mu \in \Delta(\mathcal{X})$  such that  $\|\mathbb{E}_{\mathbf{x} \sim \mu} [\phi(\mathbf{x}) - \mathbf{x}]\|_1 \leq \epsilon$ ,
2. *or* a point  $\mathbf{x} \in \mathcal{X}$  with  $\phi(\mathbf{x}) \notin \mathcal{X}$ .

Unlike [Definition 4.1](#), here we allow  $\phi$  to map outside of  $\mathcal{X}$ . This more general framing is essential to arrive at the GERorSEP oracle. In particular, we note that [Item 2](#) yields a hyperplane separating  $\phi$  from the set of endomorphisms on  $\mathcal{X}$ . Namely, since  $\phi(\mathbf{x}) \notin \mathcal{X}$ , we can use the separation oracle on  $\mathcal{X}$  to separate  $\mathcal{X}$  from  $\phi(\mathbf{x})$ ; that is, there is a  $\mathbf{w}$  such that  $\langle \phi(\mathbf{x}), \mathbf{w} \rangle > \langle \mathbf{x}, \mathbf{w} \rangle$  for all  $\mathbf{x} \in \mathcal{X}$ . But this also implies that  $\langle \phi(\mathbf{x}), \mathbf{w} \rangle > \langle \phi'(\mathbf{x}), \mathbf{w} \rangle$  for any endomorphism  $\phi'$ , as promised.

**Theorem 4.5.** *Given oracle access to  $\mathcal{X}$  and  $\phi$ , there is a  $\text{poly}(d, \log(1/\epsilon))$ -time algorithm for implementing the semi-separation oracle of Definition 4.4.*

*Proof.* As in the proof of Theorem 4.3, we proceed by running the ellipsoid algorithm (per Algorithm 8) on the problem

$$\text{find } \mathbf{y} \in [-1, 1]^d \quad \text{s.t.} \quad \langle \mathbf{y}, \phi(\mathbf{x}) - \mathbf{x} \rangle \leq -\epsilon \quad \forall \mathbf{x} \in \mathcal{X}. \quad (6)$$

For any  $\mathbf{y} \in [-1, 1]^d$  during the execution of the ellipsoid, take  $\mathbf{x}^*(\mathbf{y}) \in \text{argmin}_{\mathbf{x} \in \mathcal{X}} \langle \mathbf{y}, \mathbf{x} \rangle$ . If  $\phi(\mathbf{x}^*(\mathbf{y})) \notin \mathcal{X}$ , the algorithm can terminate and return  $\mathbf{x}^*(\mathbf{y})$ . Otherwise, it follows that  $\langle \mathbf{y}, \phi(\mathbf{x}^*(\mathbf{y})) - \mathbf{x}^*(\mathbf{y}) \rangle \geq 0$ , by definition of  $\mathbf{x}^*$ , and so we can use  $\mathbf{x}^*$  to get a separation oracle for (6).

Now, if every  $\mathbf{x}^*(\mathbf{y}) \in \mathcal{X}$  generated above satisfies the constraint  $\phi(\mathbf{x}^*(\mathbf{y})) \in \mathcal{X}$ , then Algorithm 8 returns a certificate of infeasibility for (6) in  $\text{poly}(d, \log(1/\epsilon))$  time, which is an  $\epsilon$ -expected fixed point of  $\phi$ . On the other hand, if at some point there is  $\mathbf{y} \in [-1, 1]^d$  such that  $\phi(\mathbf{x}^*(\mathbf{y})) \notin \mathcal{X}$ , then the algorithm returns a point  $\mathbf{x}^*(\mathbf{y}) \in \mathcal{X}$  such that  $\phi(\mathbf{x}^*(\mathbf{y})) \notin \mathcal{X}$ . This completes the proof.  $\square$

This semi-separation oracle amounts to the  $\epsilon$ -GERorSEP oracle needed in Theorem 2.6, as we shall see next in the context of games. Compared to the semi-separation oracle of Daskalakis et al. [18] that only works for linear functions, ours (Theorem 4.5) places no restrictions on  $\phi$ .

## 5 A polynomial-time algorithm for $\Phi^m$ -equilibria in games

Armed with the powerful semi-separation oracle of Theorem 4.5, we now establish a polynomial-time algorithm for computing  $\Phi^m$ -equilibria in general multilinear games (Theorem 5.1).

Let us recall the basic setting of an  $n$ -player multilinear game  $\Gamma$ . Each player  $i \in [n]$  has a convex and compact strategy set  $\mathcal{X}_i \subseteq \mathbb{R}^{d_i}$  in isotropic position (Section 2). Player  $i$  has a utility function  $u_i : \mathcal{X}_1 \times \dots \times \mathcal{X}_n \rightarrow \mathbb{R}$  that is linear in  $\mathcal{X}_i$ , so that  $u_i(\mathbf{x}) = \langle \mathbf{g}_i, \mathbf{x}_i \rangle$  for some  $\mathbf{g}_i = \mathbf{g}_i(\mathbf{x}_{-i}) \in \mathbb{R}^{d_i}$ . Furthermore, for each player  $i \in [n]$ , we let  $\Phi^{m_i} \subseteq \mathcal{X}_i^{\mathcal{X}_i}$  be the  $k_i$ -dimensional set of deviations in the sense of Definition 1.1; that is, there exists a function  $m_i \in \mathcal{X}_i \rightarrow \mathbb{R}^{k'_i}$ , with  $k_i = k'_i \cdot d_i + d_i$ , such that for each  $\phi_i \in \Phi^{m_i}$  and  $\mathbf{x}_i \in \mathcal{X}_i$ , the function output  $\phi_i(\mathbf{x}_i)$  can be expressed as the matrix-vector product  $\mathbf{K}_i(\phi_i)m_i(\mathbf{x}_i) + \mathbf{c}_i$  for some matrix  $\mathbf{K}_i \in \mathbb{R}^{d_i \times k'_i}$  and  $\mathbf{c}_i \in \mathbb{R}^{d_i}$ . It is assumed throughout that  $\Phi^{m_i}$  contains the identity map. For notational simplicity, we let  $k := \sum_{i=1}^n k_i$  and  $d := \sum_{i=1}^n d_i$ .

In this context, we next state the main result of this section, and proceed with its proof.

**Theorem 5.1** (Precise version of Theorem 1.3). *Consider an  $n$ -player multilinear game  $\Gamma$  such that, for each player  $i \in [n]$ , we are given  $\text{poly}(n, k)$ -time algorithms for the following:*

- *an oracle to compute the gradient, that is, the vector  $\mathbf{g}_i = \mathbf{g}_i(\mathbf{x}_{-i}) \in \mathbb{R}^{d_i}$  for which  $\langle \mathbf{g}_i(\mathbf{x}_{-i}), \mathbf{x}_i \rangle = u_i(\mathbf{x})$  for all  $\mathbf{x} \in \mathcal{X}_1 \times \dots \times \mathcal{X}_n$  (polynomial expectation property); and*
- *a membership oracle for the strategy set  $\mathcal{X}_i$ , assumed to be in isotropic position.*

*Suppose further that each  $k_i$ -dimensional set  $\Phi^{m_i}$  satisfies Assumption 3.1 and  $\|\mathbf{g}_i\| \leq B$ . Then, an  $\epsilon$ -approximate  $\Phi^m$ -equilibrium of  $\Gamma$  can be computed in  $\text{poly}(n, k, \log(B/\epsilon))$  time.*

*Proof.* An  $\epsilon$ -approximate  $\Phi^m$ -equilibrium of  $\Gamma$  is a distribution  $\mu \in \Delta(\mathcal{X}_1 \times \dots \times \mathcal{X}_n)$  such that

$$\mathbb{E}_{\mathbf{x} \sim \mu} [u_i(\phi_i(\mathbf{x}_i), \mathbf{x}_{-i}) - u_i(\mathbf{x})] \leq \epsilon$$



for every player  $i \in [n]$  and deviation  $\phi_i \in \Phi^{m_i}$ . Using multilinearity and [Definition 1.1](#), it suffices to find a distribution  $\mu \in \Delta(\mathcal{X}_1 \times \cdots \times \mathcal{X}_n)$  satisfying

$$\mathbb{E}_{\mathbf{x} \sim \mu} \left[ \sum_{i=1}^n \langle \mathbf{g}_i(\mathbf{x}_{-i}), \mathbf{K}_i \mathbf{m}_i(\mathbf{x}_i) + \mathbf{c}_i - \mathbf{x}_i \rangle \right] \leq \epsilon$$

for every  $(\mathbf{K}_1(\phi_1), \dots, \mathbf{K}_n(\phi_n))$  and  $(\mathbf{c}_1(\phi_1), \dots, \mathbf{c}_n(\phi_n))$ , where  $(\phi_1, \dots, \phi_n) \in \Phi^m$ . (This derivation uses the fact that  $\Phi^{m_i}$  contains the identity map.) We will now apply [Theorem 2.6](#) with respect to  $\mathbb{R}^d \supseteq \mathcal{X} := \mathcal{X}_1 \times \cdots \times \mathcal{X}_n$  and

$$\mathbb{R}^k \supseteq \mathcal{Y} := \{(\mathbf{K}_1, \mathbf{c}_1, \dots, \mathbf{K}_n, \mathbf{c}_n) : \mathbf{K}_i \mathbf{m}_i(\mathbf{x}_i) + \mathbf{c}_i \in \mathcal{X}_i \quad \forall \mathbf{x}_i \in \mathcal{X}_i\}.$$

By the polynomial expectation property, we can evaluate the term  $\sum_{i=1}^n \langle \mathbf{g}_i(\mathbf{x}_{-i}), \mathbf{K}_i \mathbf{m}_i(\mathbf{x}_i) + \mathbf{c}_i - \mathbf{x}_i \rangle$ , for each  $\mathbf{x} \in \mathcal{X}$ , in  $\text{poly}(n, k)$  time. It thus suffices to show how to implement the  $\epsilon$ -GERorSEP oracle, which yields a separation oracle for the program

$$\begin{aligned} & \text{find } \mathbf{K}_1, \mathbf{c}_1, \dots, \mathbf{K}_n, \mathbf{c}_n \quad \text{s.t.} \\ & \sum_{i=1}^n \langle \mathbf{g}_i(\mathbf{x}_{-i}), \mathbf{K}_i \mathbf{m}_i(\mathbf{x}_i) + \mathbf{c}_i - \mathbf{x}_i \rangle \geq -\epsilon \quad \forall \mathbf{x} \in \mathcal{X}_1 \times \cdots \times \mathcal{X}_n, \\ & \mathbf{K}_i \mathbf{m}_i(\mathbf{x}_i) + \mathbf{c}_i \in \mathcal{X}_i \quad \forall \mathbf{x}_i \in \mathcal{X}_i. \end{aligned}$$

Consider any  $\mathbb{R}^k \ni \phi = (\mathbf{K}_1, \mathbf{c}_1, \dots, \mathbf{K}_n, \mathbf{c}_n)$ . We apply the semi-separation oracle of [Theorem 4.5](#) for each function  $\mathbf{x}_i \mapsto \mathbf{K}_i \mathbf{m}_i(\mathbf{x}_i) + \mathbf{c}_i$ . This returns *either* an  $\epsilon'$ -expected fixed point, that is, a distribution  $\nu_i \in \Delta(\mathcal{X}_i)$  such that

$$\left\| \mathbb{E}_{\mathbf{x}_i \sim \nu_i} [\mathbf{K}_i \mathbf{m}_i(\mathbf{x}_i) + \mathbf{c}_i - \mathbf{x}_i] \right\|_1 \leq \epsilon',$$

or a point  $\mathbf{x}_i \in \mathcal{X}_i$  such that  $\mathbf{K}_i \mathbf{m}_i(\mathbf{x}_i) + \mathbf{c}_i \notin \mathcal{X}_i$ . If any of those semi-separation oracles returned  $\mathbf{x}_i \in \mathcal{X}_i$  with  $\mathbf{K}_i \mathbf{m}_i(\mathbf{x}_i) + \mathbf{c}_i \notin \mathcal{X}_i$ , we can use it to obtain a hyperplane separating  $(\mathbf{K}, \mathbf{c})$  from the set of deviations  $\mathcal{Y}$ . Otherwise, let  $\nu := \nu_1 \times \cdots \times \nu_n \in \Delta(\mathcal{X}_1) \times \cdots \times \Delta(\mathcal{X}_n)$  be the induced product distribution. Then, we have

$$\mathbb{E}_{\mathbf{x} \sim \nu} \sum_{i=1}^n \langle \mathbf{g}_i(\mathbf{x}_{-i}), \mathbf{K}_i \mathbf{m}_i(\mathbf{x}_i) + \mathbf{c}_i - \mathbf{x}_i \rangle = \sum_{i=1}^n \left\langle \mathbb{E}_{\mathbf{x}_{-i} \sim \nu} \mathbf{g}_i(\mathbf{x}_{-i}), \mathbb{E}_{\mathbf{x}_i \sim \nu_i} [\mathbf{K}_i \mathbf{m}_i(\mathbf{x}_i) + \mathbf{c}_i - \mathbf{x}_i] \right\rangle \leq nB\epsilon', \quad (7)$$

where we used the fact that  $\nu$  is a product distribution in the equality above. Thus, we have identified an  $(\epsilon'nB)$ -approximate good-enough-response, yielding an  $\epsilon$ -GERorSEP oracle by rescaling  $\epsilon'$ , and the proof follows from [Theorem 2.6](#).  $\square$

It is worth stressing that it is crucial for our proof that the *expected* VI problem (*cf.* Zhang et al. [60]) above corresponds to a game. It allows each player to be treated *independently*, which yields a *product distribution*  $\nu = \nu_1 \times \cdots \times \nu_n$  when we apply the semi-separation oracle of [Theorem 4.5](#) (for each player). That  $\nu$  is a product distribution is crucial to implement the separation oracle for the dual because it allows us to push the expectation into the inner product in (7), as we saw in the last step of the proof.

---

**ALGORITHM 3:** Polynomial-time algorithm for  $\Phi^m$ -equilibria

---

**Input:**

- An  $n$ -player multilinear game  $\Gamma$
- A precision parameter  $\epsilon > 0$
- A membership oracle for each  $\mathcal{X}_i$
- An oracle for computing the gradient  $\mathbf{g}_i = \mathbf{g}_i(\mathbf{x}_{-i}) \in \mathbb{R}^{d_i}$  for each  $i \in [n]$
- A  $k_i$ -dimensional set  $\Phi^{m_i}$  under [Assumption 3.1](#) for each  $i \in [n]$

**Output:** An  $\epsilon$ -approximate  $\Phi^m$ -equilibrium of  $\Gamma$  in  $\text{poly}(k, \log(1/\epsilon))$  time

Define  $G : \mathbb{R}^d \rightarrow \mathbb{R}^k$  such that  $\langle G(\mathbf{x}), (\mathbf{K}, \mathbf{c}) \rangle = \sum_{i=1}^n \langle \mathbf{g}_i(\mathbf{x}_{-i}), \mathbf{K}_i m_i(\mathbf{x}_i) + \mathbf{c}_i - \mathbf{x}_i \rangle$ ;

Use the semi-separation oracle of [Theorem 4.5](#) to construct an  $\epsilon$ -GERorSEP oracle  $\mathcal{O}$ ;

Apply [Algorithm 2](#) with  $\mathcal{O}$  as the  $\epsilon$ -GERorSEP oracle

---

## 6 An efficient online algorithm for minimizing $\Phi^m$ -regret

We now switch gears to the online learning setting, recalled in [Section 2.1](#). Our main result, [Theorem 6.4](#), is an efficient online algorithm for minimizing  $\Phi^m$ -regret with respect to any  $\text{poly}(d)$ -dimensional set  $\Phi^m$  (under [Assumption 3.1](#)), which applies even in the adversarial regime.

In what follows, we build on the framework of Daskalakis et al. [18], itself a refinement of the template of Gordon et al. [29]. As we have seen, Daskalakis et al. [18] showed that separating even over the set of linear endomorphisms is hard. In light of this, they proceed as follows. Instead of operating over the set of linear endomorphisms, their key idea is to consider a sequence of “shell sets,” each of which contains the original set. Each shell set must also satisfy two basic properties:

- it is sufficiently structured so that it is possible to optimize over that set, and
- it contains a transformation with a fixed point inside  $\mathcal{X}$ .

Here, we show that by replacing fixed points with *expected* fixed points in the above template, it is possible to extend their main result to handle any  $\text{poly}(d)$ -dimensional set under [Assumption 3.1](#).

**Overview** Our main construction is [Algorithm 7](#). It is an instantiation of ShellGD ([Section 6.2](#)), which is projected gradient descent but with the twist that the constraint set is changing over time—reflecting the fact that a new shell set is computed at every round. To execute ShellGD, ShellProject ([Section 6.3](#)) provides an efficient projection oracle together with an approximate expected fixed point thereof, which is ultimately the output of our  $\Phi^m$ -regret minimizer. ShellProject crucially relies on ShellEllipsoid, introduced next in [Section 6.1](#). It strengthens our semi-separation oracle of [Theorem 4.5](#) by again using expected fixed points. [Section 6.4](#) combines those ingredients to arrive at our main result ([Theorem 6.4](#)).

### 6.1 Shell ellipsoid

Continuing from our semi-separation oracle of [Theorem 4.5](#), ShellEllipsoid ([Algorithm 4](#)) takes a step further: it takes as input a convex set of transformations  $\mathcal{F} \subseteq \mathcal{B}_D(\mathbf{0})$ —for which we have efficient oracle access, unlike  $\Phi^m$ —and returns *either* a function  $\phi \in \mathcal{F}$  and an  $\epsilon$ -expected fixed point thereof in  $\Delta(\mathcal{X})$ , *or* it provides a certificate—in the form of a polytope expressed as the intersection of a polynomial number of halfspaces—establishing that  $\text{vol}(\mathcal{F} \cap \Phi^m) \approx 0$ . ShellEllipsoid will be used later as part of the ShellProject algorithm so as to shrink the shell set.

**Lemma 6.1.** For any  $k$ -dimensional convex set  $\mathcal{F} \subseteq \mathcal{B}_D(\mathbf{0})$  with efficient oracle access and  $\epsilon > 0$ ,  $\text{ShellEllipsoid}(\mathcal{F})$  (Algorithm 4) runs in time  $\text{poly}(k, \log(1/\epsilon), \log D)$ , and

- either it returns a transformation  $\phi \in \mathcal{F}$  with an  $\epsilon$ -expected fixed point in  $\mathcal{X}$ ,
- or it returns a polytope  $\mathcal{Q} \subseteq \mathbb{R}^k$ , specified as the intersection of at most  $\text{poly}(k, \log(1/\epsilon), \log D)$  halfspaces, such that  $\Phi^m \subseteq \mathcal{Q}$  and  $\text{vol}(\mathcal{Q} \cap \mathcal{F}) < \epsilon$ .

Coupled with Theorem 4.5 pertaining to the semi-separation oracle, the proof of correctness of Lemma 6.1 is immediate. That  $\mathcal{Q}$  can be expressed as the intersection of a polynomial number of halfspaces follows from the usual analysis of ellipsoid, as in Daskalakis et al. [18, Lemma 4.2].

---

**ALGORITHM 4:**  $\text{ShellEllipsoid}(\mathcal{F})$

---

**Input:**

- Oracle access to convex set  $\mathcal{X} \subseteq \mathbb{R}^d$
- Oracle access to a  $k$ -dimensional convex set  $\mathcal{F} \subseteq \mathcal{B}_D(\mathbf{0})$
- Precision parameter  $\epsilon > 0$

Initialize  $\mathcal{E} := \mathcal{B}_D(\mathbf{0})$  and  $\mathcal{Q} := \mathbb{R}^k$ ;

**while**  $\text{vol}(\mathcal{E}) \geq \epsilon$  **do**

    Set  $\phi \in \mathcal{Q} \cap \mathcal{F}$  as the center of  $\mathcal{E}$ ;

    Run the semi-separation oracle of Theorem 4.5 with respect to  $\phi$ ;

**if** it returned an  $\epsilon$ -expected fixed point  $\mu \in \Delta(\mathcal{X})$  of  $\phi$  **then**

**return**  $\phi$ ;

**else**

        Let  $H$  be the halfspace returned by Theorem 4.5 that separates  $\phi$  from  $\Phi^m$ ;

        Set  $\mathcal{Q} := \mathcal{Q} \cap H$ ;

    Set  $\mathcal{E}$  to be the minimum volume ellipsoid containing  $\mathcal{Q} \cap \mathcal{F}$

**return**  $\mathcal{Q}$

---

## 6.2 Shell gradient descent

Instead of minimizing external regret with respect to the set  $\Phi^m$ , which is hard even under linear endomorphisms [18], the overarching idea is to run (projected) gradient descent but with respect to a sequence of changing shell sets,  $(\tilde{\mathcal{Y}}^{(t)})_{t=1}^T$ , of  $\Phi^m$  (each of which contains  $\Phi^m$ ); this process, called ShellGD, is given in Algorithm 5. So long as  $\Phi^m \subseteq \tilde{\mathcal{Y}}^{(t)}$ , ShellGD indeed minimizes external regret with respect to deviations in  $\Phi^m$ —of course, ShellGD is not a genuine regret minimizer for  $\Phi^m$  in that it is allowed to output strategies not in  $\Phi^m$ , but Lemma 6.2 below is in fact enough for the purpose of minimizing  $\Phi^m$ -regret.

**Lemma 6.2** ([18]). Suppose that the sequence of shell sets  $(\tilde{\mathcal{Y}}^{(t)})_{t=1}^T$  is such that  $\Phi^m \subseteq \tilde{\mathcal{Y}}^{(t)} \subseteq \mathcal{B}_D(\mathbf{0})$  for all  $t \in [T]$ . For any sequence of utilities  $\mathbf{U}^{(1)}, \dots, \mathbf{U}^{(T)} \in [-1, 1]^k$ , ShellGD (Algorithm 5) satisfies

$$\max_{\mathbf{y}^* \in \Phi^m} \sum_{t=1}^T \langle \mathbf{y}^* - \mathbf{y}^{(t)}, \mathbf{U}^{(t)} \rangle \leq \frac{D^2}{2\eta} + \eta \sum_{t=1}^T \|\mathbf{U}^{(t)}\|^2.$$

---

**ALGORITHM 5:** ShellGD [18]

---

**Input:** Learning rate  $\eta$ , convex and compact sets  $\tilde{\mathcal{Y}}^{(1)}, \dots, \tilde{\mathcal{Y}}^{(T)} \subseteq \mathcal{B}_D(\mathbf{0})$ ;

Initialize  $\mathbf{y}^{(0)} \in \tilde{\mathcal{Y}}^{(1)}$  and  $\mathbf{U}^{(0)} := \mathbf{0}$ ;

**for**  $t = 1, \dots, T$  **do**

    Obtain efficient oracle access to  $\tilde{\mathcal{Y}}^{(t)}$ ;

    Update  $\mathbf{y}^{(t)} := \Pi_{\tilde{\mathcal{Y}}^{(t)}}(\mathbf{y}^{(t-1)} + \eta \mathbf{U}^{(t-1)})$ ;

    Output  $\mathbf{y}^{(t)}$  as the next strategy and receive feedback  $\mathbf{U}^{(t)} \in [-1, 1]^k$

---

### 6.3 Shell projection

To implement ShellGD, we will make use of ShellProject, the algorithm that is the subject of this subsection. There are two main desiderata for the sequence of shell sets taken as input in ShellGD. First, each shell set must be structured or simple enough to allow projecting onto it—this is the whole rationale of expanding  $\Phi^m$  through shell sets. But, of course, this is not enough, for one could just consider the entire space. The second crucial concern is that each transformation given by ShellGD needs to admit (approximate) expected fixed points, so as to use the framework of Gordon et al. [29] (Theorem 2.2) and minimize  $\Phi^m$ -regret. Lemma 6.3 below, concerning ShellProject, shows how to accomplish that goal; its proof is similar to that of Daskalakis et al. [18, Theorem 4.4].

**Lemma 6.3.** *Let  $\mathcal{X}$  be a convex and compact set such that  $\mathcal{B}_r(\mathbf{0}) \subseteq \mathcal{X} \subseteq \mathcal{B}_R(\mathbf{0})$  and  $\mathcal{M}$  be a convex set such that  $\Phi^m \subseteq \mathcal{M} \subseteq \mathcal{B}_D(\mathbf{0})$ . For any  $\phi \in \mathcal{B}_D(\mathbf{0}) \subseteq \mathbb{R}^k$  and  $\epsilon > 0$ , ShellProject (Algorithm 6) runs in time  $\text{poly}(k, 1/\epsilon, R/r, D)$  and returns*

1. a shell set  $\tilde{\Phi}$  satisfying  $\Phi^m \subseteq \tilde{\Phi}$ , expressed by intersecting  $\mathcal{M}$  with at most  $\text{poly}(d, k, 1/\epsilon, R/r, D)$  halfspaces, and
2. a transformation  $\tilde{\phi} \in \tilde{\Phi}$  such that  $\|\tilde{\phi} - \Pi_{\tilde{\Phi}}(\phi)\| \leq \epsilon$ , together with an  $\epsilon$ -expected fixed point of  $\tilde{\phi}$ ,  $\mu \in \Delta(\mathcal{X})$ .

### 6.4 Putting everything together

We now combine all the previous pieces to obtain an efficient algorithm for minimizing  $\Phi^m$ -regret—when  $\Phi^m$  is  $\text{poly}(d)$ -dimensional—under a general convex and compact set  $\mathcal{X}$ . The overall construction is depicted in Algorithm 7. In effect, it runs ShellGD with respect to the sequence of shell sets  $(\tilde{\Phi}^{(t)})_{t=1}^T$ . Indeed, by the correctness guarantee of ShellProject (Item 1 of Lemma 6.3), we have the invariance  $\Phi(\mathcal{X}) \subseteq \tilde{\Phi}^{(t)}$  for all  $t \in [T]$ . Furthermore, Item 2 of Lemma 6.3 implies that  $(\mathbf{K}^{(t+1)}, \mathbf{c}^{(t+1)}) \in \tilde{\Phi}^{(t+1)}$ , returned by ShellProject in Algorithm 7, is within distance  $\epsilon$  of the projection prescribed by ShellGD. As a result, we can apply Lemma 6.2 (up to some slackness proportional to  $\epsilon$ ) to bound the external regret  $\text{Reg}_{\Phi^m}^{(T)}$  of  $((\mathbf{K}^{(t)}, \mathbf{c}^{(t)}))_{t=1}^T$  with respect to comparators from  $\Phi^m$ ; combined with the fact that  $\mu^{(t)} \in \Delta(\mathcal{X})$  is an  $\epsilon$ -expected fixed point of the function  $\mathbf{x} \mapsto \mathbf{K}^{(t)}m(\mathbf{x}) + \mathbf{c}^{(t)}$  (as promised by Item 2), it follows that the  $\Phi^m$ -regret of the learner (Algorithm 7) can be bounded by  $\text{Reg}_{\Phi^m}^{(T)} + \epsilon T$  (as in Theorem 2.2). We thus arrive at our main result.

**Theorem 6.4** (Precise version of Theorem 1.2). *Let  $\mathcal{X} \subseteq \mathbb{R}^d$  be a convex and compact set in isotropic position for which we have a membership oracle. Algorithm 7 has per-round running time*

---

**ALGORITHM 6:** ShellProject $_{\tilde{\Phi}}(\phi)$  projects  $\phi$  to a shell of  $\tilde{\Phi}$

---

**Input:**

- Convex and compact set  $\mathcal{X} \subseteq \mathbb{R}^d$  such that  $\mathcal{B}_r(\mathbf{0}) \subseteq \mathcal{X} \subseteq \mathcal{B}_R(\mathbf{0})$
- Convex set  $\mathcal{M}$  such that  $\Phi^m \subseteq \mathcal{M} \subseteq \mathcal{B}_D(\mathbf{0})$
- Transformation  $\phi \in \mathcal{B}_D(\mathbf{0})$
- Precision parameter  $\epsilon > 0$

**Output:**

- Convex set  $\tilde{\Phi}$  such that  $\Phi^m \subseteq \tilde{\Phi} \subseteq \mathcal{M}$
- Transformation  $\tilde{\phi} \in \tilde{\Phi}$  such that  $\|\tilde{\phi} - \Pi_{\tilde{\Phi}}(\phi)\| \leq \epsilon$
- An  $\epsilon$ -expected fixed point  $\mu \in \Delta(\mathcal{X})$  of  $\tilde{\phi}$

Set  $\epsilon' = \frac{\epsilon r}{32M(R)D^2}$ ;

Initialize  $\tilde{\Phi} := \mathcal{M}$ ;

**for**  $q = 0, \dots$  incremented by  $\delta := \epsilon/4D$  **do**

- Run ShellEllipsoid( $\tilde{\Phi} \cap \mathcal{B}_q(\phi)$ ) with precision  $\text{vol}(\mathcal{B}_{\epsilon'}(\cdot))$ ;
  - if** it finds  $\tilde{\phi}$  with an  $\epsilon$ -expected fixed point  $\mu \in \Delta(\mathcal{X})$  **then**
    - └ **return**  $\tilde{\Phi}, \tilde{\phi}, \mu$
  - else**
    - └ Let  $\mathcal{Q}$  be the polytope returned by ShellEllipsoid;
    - └ Set  $\tilde{\Phi} := \tilde{\Phi} \cap \mathcal{Q}$
- 

of  $\text{poly}(k, T)$  and guarantees average  $\Phi^m$ -regret of at most  $\text{poly}(k)/\sqrt{T}$ , where  $k$  is the dimension of  $\Phi^m$  under [Assumption 3.1](#).

Unlike the algorithm of Daskalakis et al. [18], a salient aspect of [Algorithm 7](#) is that it outputs a sequence of *mixed* strategies in  $\Delta(\mathcal{X})$ . As we saw earlier in [Section 2.1](#), this turns out to be necessary: Zhang et al. [59] showed that a learner restricted to output strategies in  $\mathcal{X}$  cannot efficiently minimize  $\Phi$ -regret even with respect to low-degree swap deviations (assuming  $\text{PPAD} \neq \text{P}$ ).

---

**ALGORITHM 7:**  $\Phi^m$ -regret minimizer for convex strategy sets

---

**Input:**

- Convex and compact set  $\mathcal{X} \subseteq \mathbb{R}^d$  in isotropic position
- $k$ -dimensional set  $\Phi^m$  under [Assumption 3.1](#) with respect to  $m : \mathcal{X} \rightarrow \mathbb{R}^{k'}$ , where  $k = k' \cdot d + d$
- time horizon  $T \in \mathbb{N}$

**Output:** An efficient  $\Phi^m$ -regret minimizer for  $\mathcal{X}$

Set the learning rate  $\eta \propto \frac{1}{\sqrt{T}}$  and  $\epsilon = 1/\text{poly}(k, T)$  to be sufficiently small;

Initialize  $\mu^{(1)} \in \Delta(\mathcal{X})$  and  $\mathbf{K}^{(1)} := \mathbf{I}_{d \times k'}$  to be the identity map and  $\mathbf{c}^{(1)} := \mathbf{0}$ ;

Initialize  $\mathcal{M} := \mathcal{B}_R(\mathbf{0})$  for a large enough  $R \leq \text{poly}(k)$ ;

**for**  $t = 1, \dots, T$  **do**

- Output  $\mu^{(t)} \in \Delta(\mathcal{X})$  and receive feedback  $\mathbf{u}^{(t)} \in [-1, 1]^d$ ;
  - Define  $\mathbb{R}^{d \times k' + d} \ni \mathbf{U}^{(t)} := (\mathbb{E}_{\mathbf{x}^{(t)} \sim \mu^{(t)}} \mathbf{u}^{(t)} \otimes m(\mathbf{x}^{(t)}), \mathbf{u}^{(t)})$ ;
  - Set  $\tilde{\Phi}^{(t+1)}, (\mathbf{K}^{(t+1)}, \mathbf{c}^{(t+1)}), \mu^{(t+1)} := \text{ShellProject}_{\tilde{\Phi}}((\mathbf{K}^{(t)}, \mathbf{c}^{(t)}) + \eta \mathbf{U}^{(t)})$  with input  $\mathcal{M}$  and precision  $\epsilon$ , where  $\mu^{(t+1)} \in \Delta(\mathcal{X})$  is an  $\epsilon$ -expected fixed point of  $\mathbf{x} \mapsto \mathbf{K}^{(t+1)} m(\mathbf{x}) + \mathbf{c}^{(t+1)}$
- 

Finally, we conclude by providing a lower bound that matches our upper bound ([Theorem 6.4](#)) up to a constant factor in the exponent of  $k$ . It is based on the following normal-form lower bound due to Dagan et al. [14] and Peng and Rubinfeld [52].

**Theorem 6.5** ([14, 52]). *Consider a learner operating on the simplex  $\Delta(\mathcal{A})$ . For any  $T < |\mathcal{A}|/4$ , there is an adversary that forces the swap regret of the learner to be  $\Omega(\log^{-6} T)$ .*

We observe that there is a simple way to parameterize the above lower bound in terms of the dimension of the set of deviations:

**Corollary 6.6.** *Consider a learner operating on the simplex  $\Delta(\mathcal{A})$ . There is a  $k$ -dimensional set of deviations  $\Phi \subseteq \Delta(\mathcal{A})^{\Delta(\mathcal{A})}$  such that for any  $T < \sqrt{k}/4$ , there is an adversary that forces the  $\Phi$ -regret of the learner to be  $\Omega(\log^{-6} T)$ .*

Indeed, one can first identify an arbitrary subset  $\mathcal{A}'$  of  $\mathcal{A}$  with cardinality  $\sqrt{k}$ , and then employ the adversary of [Theorem 6.5](#) with respect to  $\mathcal{A}'$  while rendering all other actions dominated by assigning to them very small utility. That  $\Phi$  is  $k$ -dimensional in this case follows because the set of stochastic matrices mapping  $\Delta(\mathcal{A}')$  to  $\Delta(\mathcal{A}')$ —which contains all relevant swap deviations—is  $(\sqrt{k})^2$ -dimensional.

Combining [Corollary 6.6](#) with the recent reduction of Daskalakis et al. [17], which embeds the normal-form game lower bound of [Theorem 6.5](#) into an extensive-form game, we arrive at [Theorem 1.6](#), which we restate below.

**Theorem 1.6.** *For any  $k$  and any  $d \geq \Theta(\log^{14} k)$ , there is an online decision problem with dimension  $d$  and an adversary such that the  $\Phi$ -regret of the learner with respect to a  $k$ -dimensional  $\Phi$  is at least  $\epsilon$  when  $T < \min\{\sqrt{k}/4, \exp(\Omega(\epsilon^{-1/6}))\}$ .*

## 7 Conclusions and open problems

In summary, we established efficient algorithms for minimizing  $\Phi$ -regret and computing  $\Phi$ -equilibria with respect to any set of deviations with a polynomial dimension. For the online learning setting, our upper bounds are tight up to constant factors in the exponents, crystallizing for the first time a family of deviations that characterizes the learnability of  $\Phi$ -regret.

There are many important avenues for future research. First, we did not attempt to optimize the (polynomial) dependence of the running time (in [Theorems 5.1](#) and [6.4](#)) on  $k$  and  $d$ ; improving the overall complexity of our algorithms is an interesting direction. Moreover, developing more practical algorithms—that refrain from using ellipsoid—would also be a valuable contribution. In particular, are there polynomial-time algorithms for computing  $\Phi$ -equilibria without resorting to the EAH framework? But the most pressing open question is to understand the complexity of computing (normal-form) correlated equilibria in the centralized model.

## Acknowledgments

T.S. is supported by the Vannevar Bush Faculty Fellowship ONR N00014-23-1-2876, National Science Foundation grants RI-2312342 and RI-1901403, ARO award W911NF2210266, and NIH award A240108S001. B.H.Z. is supported by the CMU Computer Science Department Hans Berliner PhD Student Fellowship. E.T., R.E.B., and V.C. thank the Cooperative AI Foundation, Polaris Ventures (formerly the Center for Emerging Risk Research) and Jaan Tallinn’s donor-advised fund at Founders Pledge for financial support. E.T. and R.E.B. are also supported in part by the Cooperative AI PhD Fellowship. G.F. is supported by the National Science Foundation grant CCF-2443068. We are indebted to Constantinos Daskalakis and Noah Golowich for many insightful discussions concerning the complexity of computing expected fixed points.

## References

- [1] Angelos Assos, Yuval Dagan, and Constantinos Daskalakis. Maximizing utility in multi-agent environments by anticipating the behavior of other learners. In *Proceedings of the Annual Conference on Neural Information Processing Systems (NeurIPS)*, 2024.
- [2] Robert Aumann. Subjectivity and correlation in randomized strategies. *Journal of Mathematical Economics*, 1:67–96, 1974.
- [3] Yu Bai, Chi Jin, Song Mei, Ziang Song, and Tiancheng Yu. Efficient phi-regret minimization in extensive-form games via online mirror descent. In *Proceedings of the Annual Conference on Neural Information Processing Systems (NeurIPS)*, 2022.
- [4] Martino Bernasconi, Matteo Castiglioni, Alberto Marchesi, Francesco Trovò, and Nicola Gatti. Constrained phi-equilibria. In *International Conference on Machine Learning (ICML)*, 2023.
- [5] David Blackwell. An analog of the minmax theorem for vector payoffs. *Pacific Journal of Mathematics*, 6:1–8, 1956.
- [6] Avrim Blum and Yishay Mansour. From external to internal regret. *J. Mach. Learn. Res.*, 8: 1307–1324, 2007.
- [7] Jean Bourgain. Random points in isotropic convex sets. *Convex geometric analysis*, 34:53–58, 1996.
- [8] Michael Bowling, Neil Burch, Michael Johanson, and Oskari Tammelin. Heads-up limit hold'em poker is solved. *Science*, 347(6218):145–149, 2015.
- [9] Noam Brown and Tuomas Sandholm. Superhuman AI for heads-up no-limit poker: Libratus beats top professionals. *Science*, 359(6374):418–424, 2018.
- [10] Noam Brown and Tuomas Sandholm. Superhuman AI for multiplayer poker. *Science*, 365(6456):885–890, 2019.
- [11] Yang Cai, Constantinos Daskalakis, Haipeng Luo, Chen-Yu Wei, and Weiqiang Zheng. On tractable  $\phi$ -equilibria in non-concave games. In *Proceedings of the Annual Conference on Neural Information Processing Systems (NeurIPS)*, 2024.
- [12] Yang Cai, Haipeng Luo, Chen-Yu Wei, and Weiqiang Zheng. Near-optimal policy optimization for correlated equilibrium in general-sum markov games. In *International Conference on Artificial Intelligence and Statistics (AISTATS)*, 2024.
- [13] Xi Chen, Xiaotie Deng, and Shang-Hua Teng. Settling the complexity of computing two-player Nash equilibria. *Journal of the ACM*, 2009.
- [14] Yuval Dagan, Constantinos Daskalakis, Maxwell Fishelson, and Noah Golowich. From external to swap regret 2.0: An efficient reduction for large action spaces. In *Proceedings of the Annual Symposium on Theory of Computing (STOC)*, 2024.
- [15] Constantinos Daskalakis. Non-concave games: A challenge for game theory’s next 100 years. In *Nobel symposium" One Hundred Years of Game Theory: Future Applications and Challenges*, 2022.

- [16] Constantinos Daskalakis, Paul Goldberg, and Christos Papadimitriou. The complexity of computing a Nash equilibrium. *SIAM Journal on Computing*, 2008.
- [17] Constantinos Daskalakis, Gabriele Farina, Noah Golowich, Tuomas Sandholm, and Brian Hu Zhang. A lower bound on swap regret in extensive-form games, 2024.
- [18] Constantinos Daskalakis, Gabriele Farina, Maxwell Fishelson, Charilaos Pipis, and Jon Schneider. Efficient learning and computation of linear correlated equilibrium in general convex games. In *Proceedings of the Annual Symposium on Theory of Computing (STOC)*, 2025.
- [19] Yuan Deng, Jon Schneider, and Balasubramanian Sivan. Strategizing against no-regret learners. In *Proceedings of the Annual Conference on Neural Information Processing Systems (NeurIPS)*, 2019.
- [20] Miroslav Dudík and Geoffrey J. Gordon. A sampling-based approach to computing equilibria in succinct extensive-form games. In *Conference on Uncertainty in Artificial Intelligence (UAI)*, 2009.
- [21] Liad Erez, Tal Lancewicki, Uri Sherman, Tomer Koren, and Yishay Mansour. Regret minimization and convergence to equilibria in general-sum markov games. In *International Conference on Machine Learning (ICML)*, 2023.
- [22] Kousha Etessami and Mihalis Yannakakis. On the complexity of Nash equilibria and other fixed points (extended abstract). In *Proceedings of the Annual Symposium on Foundations of Computer Science (FOCS)*, 2007.
- [23] Gabriele Farina and Charilaos Pipis. Polynomial-time linear-swap regret minimization in imperfect-information sequential games. In *Proceedings of the Annual Conference on Neural Information Processing Systems (NeurIPS)*, 2023.
- [24] Gabriele Farina and Charilaos Pipis. Polynomial-time computation of exact phi-equilibria in polyhedral games. In *Proceedings of the Annual Conference on Neural Information Processing Systems (NeurIPS)*, 2024.
- [25] Gabriele Farina, Tommaso Bianchi, and Tuomas Sandholm. Coarse correlation in extensive-form games. In *Conference on Artificial Intelligence (AAAI)*, 2020.
- [26] Gabriele Farina, Andrea Celli, Alberto Marchesi, and Nicola Gatti. Simple uncoupled no-regret learning dynamics for extensive-form correlated equilibrium. *Journal of the ACM*, 69(6):41:1–41:41, 2022.
- [27] Dean Foster and Rakesh Vohra. Calibrated learning and correlated equilibrium. *Games and Economic Behavior*, 21:40–55, 1997.
- [28] Dean P. Foster and Sergiu Hart. Smooth calibration, leaky forecasts, finite recall, and nash dynamics. *Games and Economic Behavior*, 109:271–293, 2018.
- [29] Geoffrey J Gordon, Amy Greenwald, and Casey Marks. No-regret learning in convex games. In *Proceedings of the 25<sup>th</sup> international conference on Machine learning*, pages 360–367. ACM, 2008.



- [30] Amy Greenwald and Keith Hall. Correlated Q-learning. In *International Conference on Machine Learning (ICML)*, 2003.
- [31] Amy Greenwald and Amir Jafari. A general class of no-regret learning algorithms and game-theoretic equilibria. In *Conference on Learning Theory (COLT)*, 2003.
- [32] M. Grotschel, L. Lovász, and A. Schrijver. *Geometric Algorithms and Combinatorial Optimizations*. Springer-Verlag, 1993.
- [33] Guru Guruganesh, Yoav Kolumbus, Jon Schneider, Inbal Talgam-Cohen, Emmanouil-Vasileios Vlatakis-Gkaragkounis, Joshua R. Wang, and S. Matthew Weinberg. Contracting with a learning agent. In *Proceedings of the Annual Conference on Neural Information Processing Systems (NeurIPS)*, 2024.
- [34] Sergiu Hart and Andreu Mas-Colell. A simple adaptive procedure leading to correlated equilibrium. *Econometrica*, 68:1127–1150, 2000.
- [35] Elad Hazan and Satyen Kale. Computational equivalence of fixed points and no regret algorithms, and convergence to equilibria. In *Proceedings of the Annual Conference on Neural Information Processing Systems (NIPS)*, 2007.
- [36] Lunjia Hu and Yifan Wu. Predict to minimize swap regret for all payoff-bounded tasks. In *Proceedings of the Annual Symposium on Foundations of Computer Science (FOCS)*, 2024.
- [37] Wan Huang and Bernhard von Stengel. Computing an extensive-form correlated equilibrium in polynomial time. In *International Workshop On Internet And Network Economics (WINE)*, 2008.
- [38] Albert Jiang and Kevin Leyton-Brown. Polynomial-time computation of exact correlated equilibrium in compact games. In *Proceedings of the ACM Conference on Electronic Commerce (EC)*, 2011.
- [39] Chi Jin, Qinghua Liu, Yuanhao Wang, and Tiancheng Yu. V-learning - A simple, efficient, decentralized algorithm for multiagent reinforcement learning. *Mathematics of Operations Research*, 49(4):2295–2322, 2024.
- [40] Ravi Kannan, László Lovász, and Miklós Simonovits. Random walks and an  $O^*(n^5)$  volume algorithm for convex bodies. *Random Struct. Algorithms*, 11(1):1–50, 1997.
- [41] Michael Littman. Markov games as a framework for multi-agent reinforcement learning. In *International Conference on Machine Learning (ICML)*, pages 157–163, 1994.
- [42] László Lovász and Santosh S. Vempala. Simulated annealing in convex bodies and an  $O^*(n^4)$  volume algorithm. *J. Comput. Syst. Sci.*, 72(2):392–417, 2006.
- [43] Yishay Mansour, Mehryar Mohri, Jon Schneider, and Balasubramanian Sivan. Strategizing against learners in bayesian games. In *Conference on Learning Theory (COLT)*, 2022.
- [44] Matej Moravčík, Martin Schmid, Neil Burch, Viliam Lisý, Dustin Morrill, Nolan Bard, Trevor Davis, Kevin Waugh, Michael Johanson, and Michael Bowling. Deepstack: Expert-level artificial intelligence in heads-up no-limit poker. *Science*, 356(6337):508–513, 2017.

- [45] Dustin Morrill, Ryan D’Orazio, Marc Lanctot, James R. Wright, Michael Bowling, and Amy R. Greenwald. Efficient deviation types and learning for hindsight rationality in extensive-form games. In *International Conference on Machine Learning (ICML)*, 2021.
- [46] Dustin Morrill, Ryan D’Orazio, Reza Sarfati, Marc Lanctot, James R. Wright, Amy R. Greenwald, and Michael Bowling. Hindsight and sequential rationality of correlated play. In *Conference on Artificial Intelligence (AAAI)*, 2021.
- [47] H. Moulin and J.-P. Vial. Strategically zero-sum games: The class of games whose completely mixed equilibria cannot be improved upon. *International Journal of Game Theory*, 7(3-4): 201–221, 1978.
- [48] Roger B. Myerson. Nash equilibrium and the history of economic theory. *Journal of Economic Literature*, 37(3):1067–1082, 1999.
- [49] John Nash. Non-cooperative games. *Annals of Mathematics*, 54:289–295, 1951.
- [50] Georgy Noarov, Ramya Ramalingam, Aaron Roth, and Stephan Xie. High-dimensional prediction for sequential decision making, 2023.
- [51] Christos H. Papadimitriou and Tim Roughgarden. Computing correlated equilibria in multi-player games. *Journal of the ACM*, 55(3):14:1–14:29, 2008.
- [52] Binghui Peng and Aviad Rubinstein. Fast swap regret minimization and applications to approximate correlated equilibria. In *Proceedings of the Annual Symposium on Theory of Computing (STOC)*, 2024.
- [53] Aaron Roth and Mirah Shi. Forecasting for swap regret for all downstream agents. In *Proceedings of the ACM Conference on Economics and Computation (EC)*, 2024.
- [54] Aviad Rubinstein. Inapproximability of Nash equilibrium. In *Proceedings of the Annual Symposium on Theory of Computing (STOC)*, 2015.
- [55] Aviad Rubinstein. Settling the complexity of computing approximate two-player Nash equilibria. In *Proceedings of the Annual Symposium on Foundations of Computer Science (FOCS)*, 2016.
- [56] Gilles Stoltz and Gábor Lugosi. Internal regret in on-line portfolio selection. *Machine Learning*, 59(1-2):125–159, 2005.
- [57] Gilles Stoltz and Gábor Lugosi. Learning correlated equilibria in games with compact sets of strategies. *Games Econ. Behav.*, 59(1):187–208, 2007.
- [58] Bernhard von Stengel and Françoise Forges. Extensive-form correlated equilibrium: Definition and computational complexity. *Mathematics of Operations Research*, 33(4):1002–1022, 2008.
- [59] Brian Hu Zhang, Ioannis Anagnostides, Gabriele Farina, and Tuomas Sandholm. Efficient  $\phi$ -regret minimization with low-degree swap deviations in extensive-form games. In *Proceedings of the Annual Conference on Neural Information Processing Systems (NeurIPS)*, 2024.

- [60] Brian Hu Zhang, Ioannis Anagnostides, Emanuel Tewolde, Ratip Emin Berker, Gabriele Farina, Vincent Conitzer, and Tuomas Sandholm. Expected variational inequalities, 2025.
- [61] Martin Zinkevich, Michael Bowling, Michael Johanson, and Carmelo Piccione. Regret minimization in games with incomplete information. In *Proceedings of the Annual Conference on Neural Information Processing Systems (NIPS)*, 2007.
- [62] Mete Şeref Ahunbay. First-order (coarse) correlated equilibria in non-concave games, 2025.

## A Sufficiency of regret minimization in isotropic position

Throughout the paper, we have assumed that we are minimizing  $\Phi$ -regret with respect to a convex set  $\mathcal{X}$  that is in isotropic position. [Lemma A.1](#) below shows that this is without any loss. The argument here is similar to Daskalakis et al. [18, Lemma A.1], with the minor modification that we need to account for mixed strategies.

**Lemma A.1.** *Let  $\mathcal{X} \subseteq \mathbb{R}^d$  be a convex and compact set such that  $\mathcal{X} \subseteq \mathcal{B}_R(\mathbf{0})$ . Let  $\psi : \mathcal{X} \rightarrow \tilde{\mathcal{X}}$  be an invertible affine transformation such that  $\tilde{\mathcal{X}} = \psi(\mathcal{X})$  is in isotropic position. Suppose that we have a regret minimizer  $\mathfrak{R}_{\tilde{\mathcal{X}}}$  for  $\tilde{\mathcal{X}}$  that incurs  $\Phi\text{-Reg}_{\tilde{\mathcal{X}}}^{(T)}$ . Then, using  $\text{poly}(d)$  time in each round, we can construct a regret minimizer  $\mathfrak{R}_{\mathcal{X}}$  for  $\mathcal{X}$  that incurs  $\Phi\text{-Reg}_{\mathcal{X}}^{(T)} \leq 2R\sqrt{d} \cdot \Phi\text{-Reg}_{\tilde{\mathcal{X}}}^{(T)}$ .*

*Proof.* Let  $\psi(\mathbf{x}) := \mathbf{A}\mathbf{x} + \mathbf{b}$  for an invertible  $\mathbf{A} \in \mathbb{R}^{d \times d}$  and  $\mathbf{b} \in \mathbb{R}^d$ . Let  $\mathbf{u}^{(1)}, \dots, \mathbf{u}^{(T)}$  be the sequence of utilities given as input to  $\mathfrak{R}_{\mathcal{X}}$ . We then provide as input to  $\mathfrak{R}_{\tilde{\mathcal{X}}}$  the sequence of utilities

$$\tilde{\mathbf{u}}^{(t)} := \frac{1}{2R\sqrt{d}} (\mathbf{A}^{-1})^\top \mathbf{u}^{(t)} \quad t = 1, \dots, T. \quad (8)$$

Since  $\tilde{\mathcal{X}}$  contains the ball  $\mathcal{B}_1(\mathbf{0})$  and  $\mathcal{X} \subseteq \mathcal{B}_R(\mathbf{0})$ , it follows that  $\|\mathbf{A}^{-1}\mathbf{x}\| \leq 2R$  for any  $\mathbf{x} \in \mathcal{B}_1(\mathbf{0})$ , which implies that the spectral norm of  $\mathbf{A}^{-1}$  is at most  $2R$ . As a result, assuming that  $\mathbf{u}^{(t)} \in [-1, 1]^d$  for all  $t \in [T]$ , it follows that the utilities constructed in (8) are also in  $[-1, 1]^d$ .

Now, suppose that  $\mathfrak{R}_{\tilde{\mathcal{X}}}$  returns the sequence of strategies  $\tilde{\mu}^{(1)}, \dots, \tilde{\mu}^{(T)} \in \Delta(\tilde{\mathcal{X}})$ . We define, for each  $t \in [T]$ ,  $\mu^{(t)} := \psi^{-1}(\tilde{\mu}^{(t)})$  as the next strategy. Consider any  $\phi \in \Phi$ , and define  $\tilde{\phi} : \tilde{\mathcal{X}} \ni \mathbf{x} \mapsto \psi(\phi(\psi^{-1}(\mathbf{x}))) \in \Phi(\tilde{\mathcal{X}})$ . Then,

$$\begin{aligned} & \sum_{t=1}^T \left\langle \mathbf{u}^{(t)}, \mathbb{E}_{\tilde{\mathbf{x}}^{(t)} \sim \tilde{\mu}^{(t)}} \psi^{-1}(\tilde{\mathbf{x}}^{(t)}) - \phi(\psi^{-1}(\tilde{\mathbf{x}}^{(t)})) \right\rangle \\ &= 2R\sqrt{d} \sum_{t=1}^T \left\langle \frac{1}{2R\sqrt{d}} (\mathbf{A}^{-1})^\top \mathbf{u}^{(t)}, \mathbb{E}_{\tilde{\mathbf{x}}^{(t)} \sim \tilde{\mu}^{(t)}} \mathbf{A}\psi^{-1}(\tilde{\mathbf{x}}^{(t)}) - \mathbf{A}\phi(\psi^{-1}(\tilde{\mathbf{x}}^{(t)})) \right\rangle \\ &= 2R\sqrt{d} \sum_{t=1}^T \left\langle \frac{1}{2R\sqrt{d}} (\mathbf{A}^{-1})^\top \mathbf{u}^{(t)}, \mathbb{E}_{\tilde{\mathbf{x}}^{(t)} \sim \tilde{\mu}^{(t)}} (\mathbf{A}\psi^{-1}(\tilde{\mathbf{x}}^{(t)}) + \mathbf{b}) - (\mathbf{A}\phi(\psi^{-1}(\tilde{\mathbf{x}}^{(t)})) + \mathbf{b}) \right\rangle \\ &= 2R\sqrt{d} \sum_{t=1}^T \left\langle \tilde{\mathbf{u}}^{(t)}, \mathbb{E}_{\tilde{\mathbf{x}}^{(t)} \sim \tilde{\mu}^{(t)}} \tilde{\mathbf{x}}^{(t)} - \tilde{\phi}(\tilde{\mathbf{x}}^{(t)}) \right\rangle \leq 2R\sqrt{d} \cdot \Phi\text{-Reg}_{\tilde{\mathcal{X}}}^{(T)}. \end{aligned}$$

□

## B Geometric properties of $\Phi^m$

In this section, we establish that the set  $\Phi^m$  per [Assumption 3.1](#) is geometrically well behaved, which is necessary to execute the ellipsoid algorithm (as well as the online learning setting). In particular, our goal is to prove [Lemma 3.3](#).

Below, for convex and compact  $\mathcal{A}, \mathcal{B} \subseteq \mathbb{R}^d$ , we use the notation

$$\Phi^m(\mathcal{A}, \mathcal{B}) := \{(\mathbf{K}, \mathbf{c}) \in \mathbb{R}^{k+d} : \mathbf{K}m(\mathbf{x}) + \mathbf{c} \in \mathcal{B} \quad \forall \mathbf{x} \in \mathcal{A}\}.$$

**Lemma B.1.** Let  $\mathcal{A}, \mathcal{B}, \mathcal{C}, \mathcal{D}$  be convex and compact sets. If  $\mathcal{A} \supseteq \mathcal{C}$  and  $\mathcal{B} \subseteq \mathcal{D}$ , then  $\Phi^m(\mathcal{A}, \mathcal{B}) \subseteq \Phi^m(\mathcal{C}, \mathcal{D})$ .

*Proof.* Consider any  $(\mathbf{K}, \mathbf{c}) \in \Phi^m(\mathcal{A}, \mathcal{B})$ . By definition, it holds that  $\mathbf{K}\mathbf{x} + \mathbf{c} \in \mathcal{B}$  for all  $\mathbf{x} \in \mathcal{A}$ . Since  $\mathcal{C} \subseteq \mathcal{A}$ , it follows that  $\mathbf{K}\mathbf{x} + \mathbf{c} \in \mathcal{B}$  for all  $\mathbf{x} \in \mathcal{C}$ , and in particular,  $\mathbf{K}\mathbf{x} + \mathbf{c} \in \mathcal{D}$  since  $\mathcal{B} \subseteq \mathcal{D}$ .  $\square$

**Lemma B.2.** Let  $\mathcal{X} \subseteq \mathbb{R}^d$  be a convex and compact set such that  $\mathcal{B}_r(\mathbf{0}) \subseteq \mathcal{X} \subseteq \mathcal{B}_R(\mathbf{0})$ , with  $R \geq 1$ . Suppose further that  $\|m(\mathbf{x})\| \leq M$  for all  $\mathbf{x} \in \mathcal{B}_R(\mathbf{0})$ , where  $M = M(R) \geq 1$ . Then,  $\Phi^m \supseteq \mathcal{B}_{r'}(\phi_0)$ , where  $r' := r/2M(R)$  and  $\phi_0 := (\mathbf{0}, \mathbf{0})$  is the constant transformation  $\mathbf{x} \mapsto \mathbf{0}$ .

*Proof.* By Lemma B.1, it suffices to prove  $\mathcal{B}_{r'}(\phi_0) \subseteq \Phi^m(\mathcal{X}, \mathcal{B}_r(\mathbf{0}))$ . Consider any  $(\mathbf{K}, \mathbf{c}) \in \mathcal{B}_{r'}(\phi_0)$ , which means that  $\|\mathbf{K}\|_F^2 + \|\mathbf{c}\|^2 \leq (r/2M(R))^2$ . Then, for any  $\mathbf{x} \in \mathcal{X}$ ,

$$\|\mathbf{K}m(\mathbf{x}) + \mathbf{c}\| \leq \|\mathbf{K}\|_F \|m(\mathbf{x})\| + \|\mathbf{c}\| \leq r.$$

This means that  $\mathcal{B}_{r'}(\phi_0) \subseteq \Phi^m(\mathcal{X}, \mathcal{B}_r(\mathbf{0}))$ , and the proof follows.  $\square$

**Lemma B.3.** Suppose that  $\text{conv } m(\mathcal{B}_r(\mathbf{0})) \supseteq \mathcal{B}_\delta(\mathbf{0})$  for some  $\delta = \delta(r) > 0$  and  $m(\mathbf{0}) = \mathbf{0}$ . Then, assuming that  $r < R$ ,

$$\Phi^m(\mathcal{B}_r(\mathbf{0}), \mathcal{B}_R(\mathbf{0})) \subseteq \mathcal{B}_{R'}(\mathbf{0}),$$

where  $R' := R\left(\frac{2\sqrt{d}}{\delta(r)} + 1\right)$ .

*Proof.* Consider any  $(\mathbf{K}, \mathbf{c}) \in \Phi^m(\mathcal{B}_r(\mathbf{0}), \mathcal{B}_R(\mathbf{0}))$ . By definition, we have  $\|\mathbf{K}m(\mathbf{x}) + \mathbf{c}\| \leq R$  for all  $\mathbf{x} \in \mathcal{B}_r(\mathbf{0})$ . Since  $m(\mathbf{0}) = \mathbf{0}$ , it follows that  $\|\mathbf{c}\| \leq R$ . Thus,  $\|\mathbf{K}m(\mathbf{x})\| \leq \|\mathbf{K}m(\mathbf{x}) + \mathbf{c}\| + \|\mathbf{c}\| \leq 2R$  for all  $\mathbf{x} \in \mathcal{B}_r(\mathbf{0})$ . Now, let  $\mathbf{x}' \in \mathbb{R}^{k'}$  with  $\|\mathbf{x}'\| = 1$  be such that  $\|\mathbf{K}\mathbf{x}'\| = \|\mathbf{K}\|$ , where  $\|\mathbf{K}\|$  is the spectral norm of  $\mathbf{K}$ . Since we have assumed that  $\text{conv } m(\mathcal{B}_r(\mathbf{0})) \supseteq \mathcal{B}_\delta(\mathbf{0})$ , it follows that there exist  $\lambda_1, \dots, \lambda_{k'+1}$ , with  $\lambda_1, \dots, \lambda_{k'+1} \geq 0$  and  $\sum_{j=1}^{k'+1} \lambda_j = 1$ , and  $\mathbf{x}_1, \dots, \mathbf{x}_{k'+1} \in \mathcal{B}_r(\mathbf{0})$  (by Carathéodory's theorem) such that  $\sum_{j=1}^{k'+1} \lambda_j m(\mathbf{x}_j) = \delta \mathbf{x}'$ . As a result,

$$\delta \|\mathbf{K}\| = \|\mathbf{K}(\delta \mathbf{x}')\| = \left\| \mathbf{K} \left( \sum_{j=1}^{k'+1} \lambda_j m(\mathbf{x}_j) \right) \right\| \leq \sum_{j=1}^{k'+1} \lambda_j \|\mathbf{K}m(\mathbf{x}_j)\| \leq 2R.$$

Finally, we have  $\|\mathbf{K}\|_F \leq \sqrt{d} \|\mathbf{K}\|$ , and the claim follows.  $\square$

*Proof of Lemma 3.3.* The claim follows directly by combining Lemmas B.1 to B.3.  $\square$

## C Further omitted proofs

Finally, this section provides the proof of Proposition 3.5, which we restate below. For completeness, we have also included the usual version of EAH (Algorithm 8, subsumed by Algorithm 2), which we used earlier in Theorems 4.3 and 4.5.

**Proposition 3.5.** Let  $m : \mathcal{X} \rightarrow \mathbb{R}^{k'}$  per (5), where  $k' = \binom{d+\ell}{\ell} - 1$ .  $\bar{m} : \mathbf{x} \mapsto m(\sqrt{d}\mathbf{x})$  satisfies Assumption 3.1 with  $M \leq d^{O(\ell)}$  and  $\delta = 1/M$ .

For the proof, we will use a simple, auxiliary lemma.

**Lemma C.1.** *Let  $X$  be a random variable such that  $\mathbb{E}[X] = 0$ ,  $\mathbb{V}[X] = 1$ , and  $X \in [-R, R]$  almost surely. Then,  $\mathbb{P}[X \geq 1/R] > 0$ .*

*Proof of Proposition 3.5.* It is clear that  $\overline{m}(\mathbf{0}) = \mathbf{0}$ . The bound on  $M$  is also immediate. We thus focus on proving that  $\delta := 1/M$  suffices.

For the sake of contradiction, suppose that  $\text{conv} \overline{m}(\mathcal{B}_1(\mathbf{0}))$  does not contain  $\mathbf{x}' \in \mathbb{R}^{k'}$  with  $\|\mathbf{x}'\| \leq \delta$ . Then, we consider a hyperplane that separates  $\text{conv} \overline{m}(\mathcal{B}_1(\mathbf{0}))$  from  $\mathbf{x}'$ , and we let  $\mathbf{v}$  be the normal vector to that hyperplane, so that  $\langle \mathbf{v}, \mathbf{x}' \rangle > \langle \mathbf{v}, \overline{m}(\mathbf{x}) \rangle$  for all  $\mathbf{x} \in \mathcal{B}_1(\mathbf{0})$ . Now, let  $\mathcal{U}$  be the uniform product distribution over  $[-1, 1]^d$ . By (4), we have  $\mathbb{E}_{\mathbf{x} \sim \mathcal{U}}[m(\mathbf{x})] = \mathbf{0}$  and  $\mathbb{E}_{\mathbf{x} \sim \mathcal{U}}[m(\mathbf{x})m(\mathbf{x})^\top] = \mathbf{I}_{k' \times k'}$  (by orthogonality). As a result, we have  $\mathbb{E}_{\mathbf{x} \sim \mathcal{U}}[\langle \mathbf{v}, m(\mathbf{x}) \rangle^2] = \mathbb{V}_{\mathbf{x} \sim \mathcal{U}}[\langle \mathbf{v}, m(\mathbf{x}) \rangle] = \|\mathbf{v}\|^2 = 1$ . Lemma C.1, applied for the random variable  $\langle \mathbf{v}, m(\mathbf{x}) \rangle$  with range  $[-M, M]$ , implies that there exists  $\mathbf{x} \in [-1, 1]^d$  such that  $\langle \mathbf{v}, m(\mathbf{x}) \rangle \geq 1/M$ , which in turn implies that there exists  $\overline{\mathbf{x}} \in \mathcal{B}_1(\mathbf{0})$ —namely,  $\overline{\mathbf{x}} := \mathbf{x}/\sqrt{d}$ —such that  $\langle \mathbf{v}, \overline{m}(\overline{\mathbf{x}}) \rangle \geq 1/M$ . But this yields  $\delta \leq \langle \mathbf{v}, \overline{m}(\overline{\mathbf{x}}) \rangle < \langle \mathbf{v}, \mathbf{x}' \rangle \leq \|\mathbf{v}\|\|\mathbf{x}'\| = \delta$ , a contradiction.  $\square$

---

**ALGORITHM 8:** Ellipsoid against hope (EAH) [51]

---

**Input:**

- Parameters  $R_y, r_y > 0$  such that  $\mathcal{B}_{r_y}(\cdot) \subseteq \mathcal{Y} \subseteq \mathcal{B}_{R_y}(\mathbf{0})$
- Precision parameter  $\epsilon > 0$
- Parameter  $B > 0$  such that  $\|G(\mathbf{x})\| \leq B$  for all  $\mathbf{x} \in \mathcal{X}$
- GER oracle for (2)
- SEP oracle for  $\mathcal{Y}$

**Output:** A sparse,  $\epsilon$ -approximate solution  $\mu \in \Delta(\mathcal{X})$  of (2) Initialize the ellipsoid  $\mathcal{E} := \mathcal{B}_{R_y}(\mathbf{0})$ ;

**while**  $\text{vol}(\mathcal{E}) \geq \text{vol}(\mathcal{B}_{\epsilon/B}(\cdot))$  **do**

Let  $\mathbf{y} \in \mathbb{R}^k$  be the center of  $\mathcal{E}$  ;

**if**  $\mathbf{y} \in \mathcal{Y}$  **then**

Let  $\mathbf{x} \in \mathcal{X}$  be a good-enough-response with respect to  $\mathbf{y}$  (via the GER oracle);

Update  $\mathcal{E}$  to the minimum volume ellipsoid containing  $\mathcal{E} \cap \{\mathbf{y} \in \mathbb{R}^k : \langle \mathbf{y}, G(\mathbf{x}) \rangle \leq 0\}$ ;

**else**

Let  $H$  be the halfspace that separates  $\mathbf{y}$  from  $\mathcal{Y}$  (via the SEP oracle) ;

Update  $\mathcal{E}$  to the minimum volume ellipsoid containing  $\mathcal{E} \cap H$ ;

Let  $\mathbf{x}^{(1)}, \dots, \mathbf{x}^{(T)}$  be the GER oracle responses produced in the process above;

Define  $\mathbf{G} := [G(\mathbf{x}^{(1)}) \mid \dots \mid G(\mathbf{x}^{(T)})] \in \mathbb{R}^{k \times T}$ ;

Compute a solution  $\boldsymbol{\lambda}$  to the convex program

$$\text{find } \boldsymbol{\lambda} \in \Delta^T \quad \text{s.t.} \quad \min_{\mathbf{y} \in \mathcal{Y}} \boldsymbol{\lambda}^\top \mathbf{G}^\top \mathbf{y} \geq -\epsilon$$

**return**  $\Delta(\mathcal{X}) \ni \mu := \sum_{t=1}^T \lambda^{(t)} \mu(\mathbf{x}^{(t)})$

---