

# Swap Regret Minimization Through Response-Based Approachability

**Ioannis Anagnostides**

*Carnegie Mellon University*

IANAGNOS@CS.CMU.EDU

**Gabriele Farina**

*Massachusetts Institute of Technology*

GFARINA@MIT.EDU

**Maxwell Fishelson**

*Massachusetts Institute of Technology*

MAXFISH@MIT.EDU

**Haipeng Luo**

*University of Southern California and Google Research*

HAIPENGL@USC.EDU

**Jon Schneider**

*Google Research*

JSCHNEI@GOOGLE.COM

**Editors:** Steve Hanneke and Tor Lattimore

## Abstract

We consider the problem of minimizing different notions of swap regret in online optimization. These forms of regret are tightly connected to correlated equilibrium concepts in games, and have been more recently shown to guarantee non-manipulability against strategic adversaries. The only computationally efficient algorithm for minimizing *linear swap regret* over a general convex set in  $\mathbb{R}^d$  was developed recently by Daskalakis, Farina, Fishelson, Pipis, and Schneider (STOC '25). However, it incurs a highly suboptimal regret bound of  $\Omega(d^4\sqrt{T})$  and also relies on computationally intensive calls to the ellipsoid algorithm at each iteration.

In this paper, we develop a significantly simpler, computationally efficient algorithm that guarantees  $O(d\sqrt{T})$  linear swap regret for a general convex set that has been preconditioned via the John ellipsoid. Our algorithm leverages the powerful response-based approachability framework of Bernstein and Shimkin (JMLR '15)—previously overlooked in the line of work on swap regret minimization—and simultaneously minimizes *profile swap regret*, which was recently shown to guarantee non-manipulability. Moreover, we establish a matching information-theoretic lower bound: any learner must incur in expectation  $\Omega(d\sqrt{T})$  linear swap regret for large enough  $T$ , even when the set is centrally symmetric. This also shows that the classic algorithm of Gordon, Greenwald, and Marks (ICML '08) is existentially optimal for minimizing linear swap regret, although it is computationally inefficient. Finally, we extend our approach to minimize regret with respect to the set of swap deviations with polynomial dimension, unifying and strengthening recent results in equilibrium computation and online learning.

## 1. Introduction

Online optimization provides a framework for sequential decision making under uncertainty. In this setting, a learner repeatedly selects a strategy from a convex set  $\mathcal{P} \subset \mathbb{R}^d$ , while an adversary specifies a loss function. Although standard algorithms, such as online gradient descent (Zinkevich, 2003) and multiplicative weights update (Littlestone and Warmuth, 1994), minimize *external regret*—comparing the learner’s performance relative to the best *fixed* strategy in hindsight—this metric is often inadequate in complex dynamic environments. To address this, various strengthenings of external regret have been put forward.  $\Phi$ -regret stands out as a central notion for capturing advanced forms of hindsight rationality (Stoltz and Lugosi, 2007; Greenwald and Jafari, 2003).

Unlike external regret, which juxtaposes the performance of the learner against a static comparator,  $\Phi$ -regret compares against a dynamic sequence of strategies produced by applying a transformation  $\phi : \mathcal{P} \rightarrow \mathcal{P}$  from a set  $\Phi$  to the learner’s history of play. A learner experiences low  $\Phi$ -regret if no consistent modification of their past behavior—induced by some transformation  $\phi \in \Phi$ —would significantly reduce their cumulative loss. When  $\Phi$  comprises all affine transformations mapping  $\mathcal{P}$  to itself, this is known as *linear swap regret*.  $\Phi$ -regret has an intrinsic appeal as a general and robust performance measure in online learning, and has found many important applications in other areas.

Indeed,  $\Phi$ -regret has been a mainstay in algorithmic game theory, where it is deeply connected to fundamental solution concepts. A celebrated result establishes that if all players employ algorithms that minimize swap regret, the average correlated distribution of play converges to the set of *correlated equilibria* (Foster and Vohra, 1997; Hart and Mas-Colell, 2000; Stoltz and Lugosi, 2005; Blum and Mansour, 2007), a seminal notion introduced by Aumann (1974). More broadly, no- $\Phi$ -regret learning guarantees convergence to the set of  $\Phi$ -*equilibria* (Stoltz and Lugosi, 2007; Greenwald and Jafari, 2003). While the early foundational work focused on the probability simplex, corresponding to *normal-form* games, there has been a surge of recent research exploring  $\Phi$ -regret minimization in games with more complex strategy sets, most notably extensive-form games (see also Section A).

Swap regret has also been linked to *non-manipulability* against a strategic adversary (Deng et al., 2019). In fact, Arunachaleswaran et al. (2024) showed that, in stark contrast to mean-based algorithms such as multiplicative weights update, no-swap-regret algorithms are, in some sense, *Pareto optimal*. The scope of this line of work has expanded to general polytope games (Mansour et al., 2022), culminating in the recent introduction of *profile swap regret* by Arunachaleswaran et al. (2025). Crucially, while minimizing full swap regret is intractable in the online learning setting (Daskalakis et al., 2024), Arunachaleswaran et al. (2025) established that there are efficient algorithms for minimizing profile swap regret. Moreover, minimizing this metric suffices to guarantee that the learner cannot be manipulated by a strategic adversary.

From a computational point of view, developing efficient online algorithms to minimize  $\Phi$ -regret has been a formidable challenge. The first polynomial-time algorithm for minimizing linear swap regret over general convex sets was established only recently by Daskalakis et al. (2025). However, their algorithm yields a weak regret bound of  $\Omega(d^4\sqrt{T})$ ; this stands in stark contrast to the  $O(\sqrt{Td}\log d)$  bound achievable on the probability simplex (Blum and Mansour, 2007). Furthermore, their approach relies on invoking the ellipsoid algorithm at each iteration, rendering it computationally intensive. Subsequent work extending the scope of Daskalakis et al. (2025) beyond linear swap regret (Zhang et al., 2025b; Arunachaleswaran et al., 2025) inherits these computational and information-theoretic limitations, leaving a significant gap in the literature.

## 1.1. Our results

We introduce a new approach for minimizing linear swap regret and beyond over general convex sets. Our first main result is summarized below.

**Theorem 1** *If  $\mathcal{P} \subset \mathbb{R}^d$  is a convex body, there is a computationally efficient algorithm that guarantees  $O(d\sqrt{T})$  linear swap regret.*

Prior to our work, the only computationally efficient algorithm for minimizing linear swap regret for general convex sets was the one by Daskalakis et al. (2025), which required  $\Omega(d^8/\epsilon^2)$  rounds to

guarantee that the learner incurs at most  $\epsilon$  average linear swap regret; Theorem 1 provides a considerable improvement by a factor of  $d^5$ . Furthermore, our algorithm is significantly more efficient in terms of the per-iteration complexity.<sup>1</sup> By virtue of the connection between linear swap regret and *linear correlated equilibria* (Farina and Pipis, 2023), Theorem 1 yields the fastest known algorithm for that problem in general convex games. Last but not least, our algorithm *simultaneously* minimizes profile swap regret, thereby being non-manipulable (Arunachaleswaran et al., 2025).

Moreover, we establish a matching lower bound, showing that Theorem 1 is existentially optimal for centrally symmetric convex bodies.

**Theorem 2** *For any  $T = \Omega(d^4)$ , there is a centrally symmetric convex body  $\mathcal{P} \subset \mathbb{R}^d$  and an adversary that guarantees that the expected linear swap regret of the learner is at least  $\Omega(d\sqrt{T})$ .*

This establishes a strict separation between minimizing linear swap regret over the simplex and a general convex set. It also shows that the classic algorithm of Gordon et al. (2008) is information-theoretically optimal for minimizing linear swap regret (Proposition 3). One limitation of the algorithm behind Theorem 1 is that it relies on a potentially inefficient preprocessing step, which we separate from the running time of the algorithm itself; an interesting question is whether the  $O(d\sqrt{T})$  rate attained by Gordon et al. (2008) can be matched without that limitation.

Finally, we extend our approach beyond linear (or profile) swap regret. We generalize Theorem 1 to the class of swap deviations with *polynomial dimension* (Zhang et al., 2025b; Gordon et al., 2008), which encompasses *low-degree* swap deviations (Zhang et al., 2024a). As before, we significantly improve over those prior results in terms of the regret bounds. In fact, our algorithm simultaneously controls both the set of swap deviations with polynomial dimension and profile swap regret, unifying the two. The resulting equilibrium concept is the strongest currently known notion that, under suitable oracle access, can be efficiently computed in general convex games.

**Technical approach** From a technical point of view, we crucially leverage the response-based algorithm of Bernstein and Shimkin (2015). The relevance of their framework has been overlooked in recent developments on minimizing  $\Phi$ -regret. Our work brings it back to the forefront.

To obtain the  $O(d\sqrt{T})$  upper bound on linear swap regret (Theorem 1), we also perform a preprocessing step whereby the strategy set is brought into *John’s position* (Algorithm 2). This allows us to tightly bound the norm of affine endomorphisms (Lemma 12) and the norm of points in the approachability space (Lemma 14), paving the way to Theorem 1. Similarly, our extension to the set of swap deviations with polynomial dimension extends the basic framework by crucially leveraging *mixed strategies* (Algorithm 3).

Regarding our lower bound (Theorem 2), we base our construction on the product strategy set  $\mathcal{P} = B_1 \times B_\infty$ , where  $B_1$  and  $B_\infty$  denote the  $\ell_1$  and  $\ell_\infty$  balls, respectively. The core argument exploits a basic tradeoff: an adversary can compel the learner’s strategy in the  $B_1$  component to either incur large linear swap regret directly, or, in an effort to evade this, consistently cover a large fraction of the vertices (Lemma 21). In the latter case, we show that this enforced strategy diversity in  $B_1$  strengthens the power of hindsight deviations available in the  $B_\infty$  component (Lemma 22): linear swap regret becomes a stronger benchmark as the learner’s strategy sequence fluctuates.

---

1. The exact running time of our algorithm depends on the type of oracles we assume. Using linear optimization oracles (Section C), it is amenable to standard online algorithms such as follow the perturbed leader (e.g., Abernethy et al., 2016) since it requires approximately solving a (bilinear) minimax game in each iteration. This is unlike the algorithm of Daskalakis et al. (2025) which requires repeatedly invoking the ellipsoid algorithm, resulting in a running time growing as a very high degree polynomial.

## 2. Preliminaries

We denote by  $\mathcal{P}, \mathcal{L} \subset \mathbb{R}^d$  the convex and compact strategy and loss sets, respectively.  $\text{conv}(\cdot)$  represents the convex hull. We make the usual normalization assumption  $|\langle \ell, p \rangle| \leq 1$  for all  $\ell \in \mathcal{L}$  and  $p \in \mathcal{P}$ . In particular,  $\mathcal{L}$  can be assumed to be the polar cone of  $\text{conv}(\mathcal{P}, -\mathcal{P})$ , namely  $\{\ell \in \mathcal{L} : \max_{p \in \text{conv}(\mathcal{P}, -\mathcal{P})} \langle p, \ell \rangle \leq 1\}$ . For the set  $\mathcal{P}$ , we can assume access to a linear optimization oracle (Grötschel et al., 1993), which also yields a membership oracle for its polar cone  $\mathcal{P}^\circ$ . We can also assume throughout that  $0 \in \mathcal{P}$ . We denote the tensor product of  $\ell$  and  $p$  by  $\ell \otimes p = \ell p^\top$ . For a matrix  $A$ , we use the shorthand notation  $A^{-\top} = (A^{-1})^\top$ .

**Linear swap regret** We say that a function  $\phi : \mathcal{P} \rightarrow \mathbb{R}^d$  is an *endomorphism* if  $\phi(\mathcal{P}) \subseteq \mathcal{P}$ . We denote by  $\text{End}(\mathcal{P})$  the set of *affine* endomorphisms with respect to  $\mathcal{P}$ . A function  $\phi \in \text{End}(\mathcal{P})$  can be represented as a matrix-vector pair  $(M, a)$ , so that  $\phi(p) = Mp + a$ . For a sequence of strategies  $(p_t)_{t=1}^T$  and losses  $(\ell_t)_{t=1}^T$ , the *linear swap regret* (Gordon et al., 2008) of the learner is defined as<sup>2</sup>

$$\text{LinearSwapReg}_T := \sum_{t=1}^T \langle \ell_t, p_t \rangle - \min_{\phi \in \text{End}(\mathcal{P})} \sum_{t=1}^T \langle \ell_t, \phi(p_t) \rangle.$$

A general framework for minimizing  $\Phi$ -regret was developed by Gordon et al. (2008). It distills the basic ingredients of earlier constructions—in the special case where  $\mathcal{P}$  is a probability simplex—by Blum and Mansour (2007) and Stoltz and Lugosi (2005). Specifically, the reduction of Gordon et al. (2008) posits access to i) a fixed-point oracle, which takes as input an endomorphism  $\phi$  and returns a fixed point thereof; and ii) an external regret minimizer *over the set*  $\Phi$ . For linear swap regret, each function  $\phi$  is affine, so a fixed point can be computed efficiently. Beyond linear swap regret, one can do away with the fixed-point oracle required by Gordon et al. (2008) by using what Zhang et al. (2024a) refer to as *expected* fixed points, which can be computed efficiently for any endomorphism. This is relevant to our extension to nonlinear swap deviations (Section 4). On the other hand, minimizing external regret with respect to the set of affine endomorphisms turns out to be computationally intractable, as shown recently by Daskalakis et al. (2025).

From an information-theoretic standpoint, we point out that the algorithm of Gordon et al. (2008) can yield a linear swap regret bound of  $\tilde{O}(d\sqrt{T})$ . Specifically, one can obtain  $\tilde{O}(\sqrt{DT})$  external regret over a general  $D$ -dimensional convex set using FTRL endowed with the *entropic barrier* function (Abernethy et al., 2008; Bubeck and Eldan, 2014), also known as continuous exponential weights (Bubeck, 2011, Section 3).<sup>3</sup> Since the set of affine endomorphisms has dimension  $d(d+1)$ , we arrive at the following.

**Proposition 3** *Assuming that  $|\langle p, \ell \rangle| \leq 1$  for all  $p \in \mathcal{P}$  and  $\ell \in \mathcal{L}$ , there is an algorithm that guarantees  $\text{LinearSwapReg}_T \leq \tilde{O}(d\sqrt{T})$  for any sequence of losses  $\ell_1, \dots, \ell_T \in \mathcal{L}$ .*

Our lower bound (Theorem 20) implies that this guarantee is information-theoretically optimal in the regime where  $T$  is large enough.

- 
2. Gordon et al. (2008) did not include the affine component in their definition. When  $0 \in \mathcal{P}$ , which is always the case when  $\mathcal{P}$  is centrally symmetric, selecting  $p_t = 0$  guarantees zero linear swap regret. For this reason, we include the affine component.
  3. The loss vector produced by the reduction of Gordon et al. (2008) under affine endomorphisms can be cast as  $L_t = (\ell_t \otimes p_t, \ell_t)$ , so the condition  $|\langle \phi_t, L_t \rangle| \leq 1$  is met.

**Profile swap regret** Profile swap regret was recently introduced by [Arunachaleswaran et al. \(2025\)](#), who showed that the property of having no profile swap regret is equivalent to the property of being non-manipulable by a dynamic optimizer. Profile swap regret is defined with respect to the *correlated strategy profile (CSP)*  $\bar{\kappa}_T = \frac{1}{T} \sum_{t=1}^T \ell_t \otimes p_t$ . If  $R : \mathcal{P} \times \mathcal{L} \ni (p, \ell) \mapsto \max_{p^* \in \mathcal{P}} \langle p - p^*, \ell \rangle$  measures the *instantaneous regret*, we have

$$\text{ProfileSwapReg}_T(\bar{\kappa}_T) = T \left( \min_{(p(t), \ell(t), \lambda_t)} \sum_{t=1}^T \lambda_t R(p(t), \ell(t)) \right),$$

where the minimum above is over all valid convex decompositions of  $\bar{\kappa}_T$  into  $\sum_{t=1}^T \lambda_t (\ell(t) \otimes p(t))$  such that  $\lambda_1, \dots, \lambda_T \geq 0$  with  $\sum_{t=1}^T \lambda_t = 1$ ,  $\ell(1), \dots, \ell(T) \in \mathcal{L}$ , and  $p(1), \dots, p(T) \in \mathcal{P}$ .

### 3. Efficient linear swap regret minimization via approachability

In this section, we provide a computationally efficient algorithm whose linear swap regret is bounded as  $O(d\sqrt{T})$  (Theorem 15). Our approach circumvents the obstacle of minimizing external regret over  $\text{End}(\mathcal{P})$  by making a connection to earlier work in the approachability literature, specifically the framework of [Bernstein and Shimkin \(2015\)](#), together with preconditioning techniques from convex geometry. All omitted proofs are deferred to Section D.

#### 3.1. Reducing linear swap regret to approachability

We begin by reducing linear swap regret to a suitable approachability problem (Lemma 4), so that we can leverage the response-based algorithm of [Bernstein and Shimkin \(2015\)](#). We define the best-response function as  $b(\ell) = \arg \min_{p \in \mathcal{P}} \langle \ell, p \rangle$ , with ties broken arbitrarily. Furthermore, we let  $\mathcal{K} := \text{conv}\{(\ell \otimes p, \ell) : \ell \in \mathcal{L}, p \in \mathcal{P}\}$ , and define the *target set*  $\mathcal{S} \subseteq \mathcal{K}$  as  $\mathcal{S} = \text{conv}\{(\ell \otimes b(\ell), \ell) : \ell \in \mathcal{L}\}$ . For a sequence of strategies  $(p_t)_{t=1}^T$  and losses  $(\ell_t)_{t=1}^T$ , we denote by  $\kappa_t = (\ell_t \otimes p_t, \ell_t)$  and  $\bar{\kappa}_T = \frac{1}{T} \sum_{t=1}^T \kappa_t$  the average correlated strategy profile. With this notation at hand, linear swap regret can be equivalently expressed as

$$\sum_{t=1}^T \langle \ell_t, p_t \rangle - \min_{\phi \in \text{End}(\mathcal{P})} \sum_{t=1}^T \langle \ell_t, \phi(p_t) \rangle = T \max_{\phi = (M, a) \in \text{End}(\mathcal{P})} \langle (I_d, 0) - (M, a), \bar{\kappa}_T \rangle,$$

where the inner product is component-wise. By convention, we write  $\|(M, a)\|_F^2 = \|M\|_F^2 + \|a\|_2^2$ . We define the *approachability loss*, measured in the Euclidean norm, as

$$\text{AppLoss}_T := \min_{s \in \mathcal{S}} \|\bar{\kappa}_T - s\|_F.$$

The following lemma shows that minimizing linear swap regret reduces to minimizing the approachability loss.

**Lemma 4**  $\text{LinearSwapReg}_T \leq 2T \text{AppLoss}_T \cdot \max_{\phi \in \text{End}(\mathcal{P})} \|\phi\|_F$ .

Profile swap regret also reduces to minimizing the same approachability loss. Specifically, following [Arunachaleswaran et al. \(2025\)](#), the goal is to bound  $\text{ProfileSwapDist}_T(\bar{\kappa}_T)$ , the *profile swap distance*, which measures the distance of  $\bar{\kappa}_T$  from a CSP  $s$  such that  $\text{ProfileSwapReg}_T(s) \leq 0$ . Since this is equivalent to  $s \in \mathcal{S}$ , profile swap distance is equal to the approachability loss.

---

**Algorithm 1** Response-based approachability (Bernstein and Shimkin, 2015)
 

---

**Input:** Horizon  $T$ , convex sets  $\mathcal{P}, \mathcal{L} \subset \mathbb{R}^d$ , best-response map  $b : \mathcal{L} \rightarrow \mathcal{P}$ 
**Output:** Sequence of strategies  $p_1, \dots, p_T$ 

- 1 Initialize  $U_0 \leftarrow 0 \in \mathbb{R}^{d \times (d+1)}$
  - 2 **for**  $t = 1$  **to**  $T$  **do**
  - 3     Compute a pair of minimax strategies  $(p_t, \ell_t^*)$  of the (bilinear) zero-sum game
 
$$\min_{p \in \mathcal{P}} \max_{\ell \in \mathcal{L}} \langle U_{t-1}, (\ell \otimes p, \ell) \rangle$$
  - 4     Set  $s_t \leftarrow (\ell_t^* \otimes b(\ell_t^*), \ell_t^*)$
  - 5     **Play** the strategy  $p_t$  and **observe** the loss  $\ell_t$
  - 6     Set  $\kappa_t \leftarrow (\ell_t \otimes p_t, \ell_t)$
  - 7     Update  $U_t \leftarrow U_{t-1} + (\kappa_t - s_t)$
  - 8 **end**
- 

### 3.2. Response-based approachability

Having reduced minimizing swap regret to an approachability problem (Lemma 4), we now leverage the algorithm of Bernstein and Shimkin (2015), which in our setting can be cast as Algorithm 1.

**Algorithm description and analysis** As with other approachability algorithms, the response-based algorithm of Bernstein and Shimkin (2015) fundamentally relies on the *Pythagorean lemma*, recalled next. Its proof follows simply by induction.

**Lemma 5** Let  $v_1, \dots, v_T \in \mathbb{R}^d$  be a sequence of points each with Euclidean norm bounded by  $B$ . Suppose further that  $\langle \sum_{\tau=1}^{t-1} v_\tau, v_t \rangle \leq 0$  for all  $t \in [T]$ . Then  $\left\| \sum_{\tau=1}^T v_\tau \right\|_2 \leq B\sqrt{T}$ .

In light of this lemma, the key of the algorithm is to ensure that  $\langle \bar{\kappa}_{t-1} - \bar{s}_{t-1}, \kappa_t - s_t \rangle \leq 0$ . Indeed, by Lemma 5, this guarantees a bound on the approachability loss:  $\min_{s \in \mathcal{S}} \|\bar{\kappa}_T - s\|_2 \leq \|\bar{\kappa}_T - \bar{s}_T\|_2 \leq 2B/\sqrt{T}$ , for  $B = \max_{p \in \mathcal{P}, \ell \in \mathcal{L}} \|(\ell \otimes p, \ell)\|_F = \max_{\ell \in \mathcal{L}} \|\ell\|_2 \max_{p \in \mathcal{P}} \sqrt{\|p\|_2^2 + 1}$ .

To guarantee the invariance  $\langle \bar{\kappa}_{t-1} - \bar{s}_{t-1}, \kappa_t - s_t \rangle \leq 0$ , Algorithm 1 proceeds as follows. At each time step  $t \in [T]$ , it maintains internally a point  $s_t \in \mathcal{S}$ . We let  $\bar{s}_t = \frac{1}{t} \sum_{\tau=1}^t s_\tau$ ,  $\bar{\kappa}_t = \frac{1}{t} \sum_{\tau=1}^t \kappa_\tau$ , and  $U_t := \sum_{\tau=1}^t (\kappa_\tau - s_\tau)$ . To select  $p_t$  and  $s_t$  in such a way that, for any adversarial selection of  $\ell_t$ ,  $\langle \bar{\kappa}_{t-1} - \bar{s}_{t-1}, \kappa_t - s_t \rangle \leq 0$ , Algorithm 1 first solves a (bilinear) minimax game:  $p_t \in \arg \min_{p \in \mathcal{P}} \max_{\ell \in \mathcal{L}} \langle U_{t-1}, (\ell \otimes p, \ell) \rangle$ ,  $\ell_t^* \in \arg \max_{\ell \in \mathcal{L}} \min_{p \in \mathcal{P}} \langle U_{t-1}, (\ell \otimes p, \ell) \rangle$ . Then it selects  $s_t = (\ell_t^* \otimes b(\ell_t^*), \ell_t^*)$ . The correctness of the approach follows from the minimax theorem.

**Lemma 6** The points  $p_t$  and  $s_t$  selected by Algorithm 1 satisfy  $\langle U_{t-1}, \kappa_t - s_t \rangle \leq 0$  for any adversarial loss  $\ell_t \in \mathcal{L}$ , where  $\kappa_t = (\ell_t \otimes p_t, \ell_t)$ .

We summarize the main guarantee of Algorithm 1 concerning linear swap regret below. The proof follows by combining Lemmas 4 to 6.

**Theorem 7** For any sequence of losses  $\ell_1, \dots, \ell_T \in \mathcal{L}$ , Algorithm 1 guarantees

$$\text{LinearSwapReg}_T \leq 4\sqrt{T} \left( \max_{p \in \mathcal{P}, \ell \in \mathcal{L}} \|\ell\|_2 \sqrt{\|p\|_2^2 + 1} \right) \left( \max_{\phi \in \text{End}(\mathcal{P})} \|\phi\|_F \right). \quad (1)$$

Given that  $\text{ProfileSwapDist}_T = \text{AppLoss}_T$ , we also obtain the following.

**Corollary 8** *For any sequence of losses  $\ell_1, \dots, \ell_T \in \mathcal{L}$ , Algorithm 1 guarantees*

$$\text{ProfileSwapDist}_T \leq 2\sqrt{T} \left( \max_{p \in \mathcal{P}, \ell \in \mathcal{L}} \|\ell\|_2 \sqrt{\|p\|_2^2 + 1} \right).$$

For a comparison between Algorithm 1 and Blackwell's approachability we refer to Section B. Our next goal is to show that (1) can be upper bounded by  $O(d\sqrt{T})$  using a suitable preconditioner.

### 3.3. Geometric preconditioning

We now employ a suitable preconditioning step in conjunction with Algorithm 1 to obtain improved upper bounds on linear swap regret. We first point out a simple lemma, showing that if  $\mathcal{P}$  is mapped to  $A\mathcal{P}$  and  $\mathcal{L}$  to  $A^{-\top}\mathcal{L}$  for an invertible linear transformation  $A$ , the linear swap regret is preserved.

**Lemma 9** *Let  $A \in \mathbb{R}^{d \times d}$  be an invertible matrix. Consider the transformed strategy set  $\mathcal{P}' = A\mathcal{P} := \{Ap : p \in \mathcal{P}\}$  and loss set  $\mathcal{L}' = A^{-\top}\mathcal{L} := \{A^{-\top}\ell : \ell \in \mathcal{L}\}$ . For any sequence of strategies  $(p_t)_{t=1}^T$  in  $\mathcal{P}$  and sequence of losses  $(\ell_t)_{t=1}^T$  in  $\mathcal{L}$ , define the transformed sequences  $p'_t = Ap_t$  and  $\ell'_t = A^{-\top}\ell_t$  for all  $t \in [T]$ . Then the linear swap regret  $\text{LinearSwapReg}_T$  of  $(p_t)_{t=1}^T$  with respect to  $(\ell_t)_{t=1}^T$  is equal to the linear swap regret  $\text{LinearSwapReg}'_T$  of  $(p'_t)_{t=1}^T$  with respect to  $(\ell'_t)_{t=1}^T$ .*

In particular, if  $|\langle p, \ell \rangle| \leq 1$  for all  $p \in \mathcal{P}$  and  $\ell \in \mathcal{L}$ , then  $|\langle p', \ell' \rangle| = |\langle Ap, A^{-\top}\ell \rangle| \leq 1$  for all  $p' \in \mathcal{P}'$  and  $\ell' \in \mathcal{L}'$ . This lemma allows us to adapt the geometry of  $\mathcal{P}$  and  $\mathcal{L}$  so as to minimize the right-hand side of (1) in Theorem 7.

**Frobenius norm of endomorphisms** We begin by bounding the Frobenius norm of  $\phi \in \text{End}(\mathcal{P})$ . Our basic approach is to bring  $\mathcal{P}$  into *John's position*. A convex body  $\mathcal{P}$  is said to be in John's position if the maximum volume ellipsoid contained in  $\mathcal{P}$  is the  $d$ -dimensional Euclidean unit ball centered at 0, denoted by  $B_2$ . Without loss of generality, we will assume that  $\mathcal{P}$  is fully dimensional.

**Theorem 10 (John's Theorem)** *If  $\mathcal{P}$  is in John's position, there exist contact points  $\xi_1, \dots, \xi_m \in \partial\mathcal{P} \cap \partial B_2$  and weights  $c_1, \dots, c_m > 0$  such that*

$$\sum_{i=1}^m c_i \xi_i \otimes \xi_i = I_d \quad \text{and} \quad \sum_{i=1}^m c_i \xi_i = 0. \quad (2)$$

Furthermore, if  $\mathcal{P}$  is centrally symmetric,  $B_2 \subseteq \mathcal{P} \subseteq \sqrt{d}B_2$ .

In this context, we will first show that if  $\mathcal{P}$  is centrally symmetric and is in John's position, the Frobenius norm of each  $\phi \in \text{End}(\mathcal{P})$  is upper bounded by  $O(\sqrt{d})$  (Lemma 12). First, we point out that using the second implication of John's theorem, namely  $B_2 \subseteq \mathcal{P} \subseteq \sqrt{d}B_2$ , does *not* suffice to obtain this bound; a lopsided ellipsoid provides a counterexample.

**Lemma 11** *There exists a convex set  $\mathcal{P} \subset \mathbb{R}^d$  satisfying  $B_2 \subseteq \mathcal{P} \subseteq \sqrt{d}B_2$  and a linear map  $M$  such that  $M(\mathcal{P}) \subseteq \mathcal{P}$  but  $\|M\|_F = \Omega(d)$ .*

Moreover, if we only knew that  $B_2 \subseteq \mathcal{P} \subseteq O(d)B_2$ , which is the case when  $\mathcal{P}$  is put into isotropic position (Lovász and Vempala, 2006), we could only guarantee a spectral norm bound of the form  $\|M\|_2 \leq O(d)$ ; in turn, this would yield  $\|M\|_F \leq O(d^{3/2})$ . To obtain an improved  $O(\sqrt{d})$  upper bound, we need to make crucial use of John’s decomposition—(2) in Theorem 10—as formalized below.

**Lemma 12** *Let  $\mathcal{P} \subset \mathbb{R}^d$  be a convex body in John’s position and  $\phi(p) = Mp + a$  be an affine endomorphism of  $\mathcal{P}$ . Then we have  $\|M\|_F \leq \sqrt{d + \|a\|_2^2}$ . In particular, if  $\mathcal{P}$  is centrally symmetric,  $\|a\|_2 \leq \sqrt{d}$  and  $\|M\|_F \leq \sqrt{2d}$ .*

While this  $O(\sqrt{d})$  bound hinges on  $\mathcal{P}$  being centrally symmetric, we observe that one can more generally put  $\text{conv}(\mathcal{P}, -\mathcal{P})$  into John’s position, which is centrally symmetric by definition. The following lemma justifies that transformation.

**Lemma 13** *Let  $\phi : p \mapsto Mp + a$  be an affine endomorphism on  $\mathcal{P}$  with  $0 \in \mathcal{P}$ . Then  $\psi : p \mapsto \frac{1}{3}Mp + \frac{1}{3}a$  is an endomorphism on  $\text{conv}(\mathcal{P}, -\mathcal{P})$ .*

As a result, if  $\text{conv}(\mathcal{P}, -\mathcal{P})$  is in John’s position, Lemma 12 implies an  $O(\sqrt{d})$  bound on the Frobenius norm of endomorphisms on  $\mathcal{P}$ , as desired.

**Frobenius norm of  $\mathcal{K}$**  Our next goal is to upper bound the Frobenius norm of points in  $\mathcal{K}$ . If we know that  $|\langle p, \ell \rangle| \leq 1$ , the naive bound on  $\|\ell \otimes p\|_F$  is linear in the dimension  $d$ . We significantly improve this by again making use of John’s position and Theorem 10.

**Lemma 14** *If  $\mathcal{P}$  is a convex body such that  $\text{conv}(\mathcal{P}, -\mathcal{P})$  is in John’s position and  $\mathcal{L}$  is such that  $|\langle p, \ell \rangle| \leq 1$  for all  $p \in \mathcal{P}$  and  $\ell \in \mathcal{L}$ , we have  $\max_{p \in \mathcal{P}, \ell \in \mathcal{L}} \|(\ell \otimes p, \ell)\|_F \leq \sqrt{d + 1}$ .*

Combining Lemmas 9 and 12 to 14 with Theorem 7, we obtain an online algorithm whose linear swap regret grows as  $O(d\sqrt{T})$ . This is given as Algorithm 2, which is a preconditioned version of Algorithm 1.

**Theorem 15** *Let  $\mathcal{P} \subset \mathbb{R}^d$  be a convex body and a set of losses  $\mathcal{L} \subset \mathbb{R}^d$  such that  $|\langle p, \ell \rangle| \leq 1$  for all  $p \in \mathcal{P}$  and  $\ell \in \mathcal{L}$ . For any sequence of losses  $\ell_1, \dots, \ell_T \in \mathcal{L}$ , Algorithm 2 guarantees*

$$\text{LinearSwapReg}_T \leq O(d\sqrt{T}).$$

In terms of computational efficiency, the per-iteration complexity is dominated by the computation of minimax strategies, which can be solved efficiently using linear optimization oracles. Moreover, we show that Theorem 15 extends when  $(p_t, \ell_t^*)$  is an  $\epsilon$ -approximate equilibria with degradation proportional to  $\epsilon T$  (Theorem 25 in Section C); when  $\epsilon \propto 1/\sqrt{T}$ , this error term becomes negligible. The problem of approximate minimax equilibrium computation can be solved in a scalable fashion via (external) no-regret learning (Freund and Schapire, 1997).

One limitation of Theorem 15 is that the John ellipsoid cannot be computed efficiently with just oracle access to the set. (On the other hand, it is well-known that the John ellipsoid can be computed in polynomial time for explicitly represented polytopes using convex optimization (Kumar and Yildirim, 2005; Todd and Yildirim, 2007; Cao et al., 2022).) This preprocessing step is only applied once, and we separate that from the running time of the algorithm itself. It is an interesting

---

**Algorithm 2** Preconditioned response-based approachability
 

---

**Input:** Horizon  $T$ , convex body  $\mathcal{P} \subset \mathbb{R}^d$  and loss set  $\mathcal{L} \subset \mathbb{R}^d$  such that  $\langle p, \ell \rangle \leq 1$  for all  $p \in \mathcal{P}$  and  $\ell \in \mathcal{L}$ , best-response map  $b : \mathcal{L} \rightarrow \mathcal{P}$

**Output:** Sequence of strategies  $p_1, \dots, p_T \in \mathcal{P}$

// 1. Preconditioning Step

1 Compute an invertible matrix  $A$  such that  $\text{conv}(A\mathcal{P}, -A\mathcal{P})$  is in John's position

2 Let  $\mathcal{P}' \leftarrow A\mathcal{P}$  and  $\mathcal{L}' \leftarrow A^{-\top}\mathcal{L}$

3 Define the transformed best-response map  $b'(\ell') \leftarrow Ab(A^\top \ell')$

4 Initialize  $U_0 \leftarrow 0 \in \mathbb{R}^{d \times (d+1)}$

5 **for**  $t = 1$  **to**  $T$  **do**

// Solve the minimax game in the transformed space

6 Compute a pair of minimax strategies  $(p'_t, \ell'^*_t)$  of the (bilinear) zero-sum game

$$\min_{p' \in \mathcal{P}'} \max_{\ell' \in \mathcal{L}'} \langle U_{t-1}, (\ell' \otimes p', \ell') \rangle$$

7 Set  $s'_t \leftarrow (\ell'^*_t \otimes b'(\ell'^*_t), \ell'^*_t)$

// Play in the original space

8 **Play** the strategy  $p_t \leftarrow A^{-1}p'_t \in \mathcal{P}$  and **observe** the loss  $\ell_t \in \mathcal{L}$

// Update using transformed loss

9 Set transformed loss  $\ell'_t \leftarrow A^{-\top}\ell_t$

10 Set  $\kappa'_t \leftarrow (\ell'_t \otimes p'_t, \ell'_t)$

11 Update  $U_t \leftarrow U_{t-1} + (\kappa'_t - s'_t)$

12 **end**

---

question whether the preprocessing step can also be made efficient while maintaining the  $O(d\sqrt{T})$  bound on linear swap regret.

Prior to our work, the only computationally efficient algorithm for minimizing linear swap regret over general convex sets was the one developed by [Daskalakis et al. \(2025\)](#), which gave a bound of  $O(d^4\sqrt{T})$ . In terms of the number of iterations needed to guarantee at most  $\epsilon$  linear swap regret, their algorithm requires  $T = \Omega(d^8/\epsilon^2)$ , whereas ours requires at most  $16d^3/\epsilon^2$ . This is an improvement by a factor of  $d^5$ . Furthermore, Algorithm 2 is significantly more efficient per iteration.

#### 4. Swap deviations with polynomial dimension

In this section, we extend our approach beyond linear and profile swap regret. Specifically, we consider the class of swap deviations with *polynomial dimension*, introduced in the form below by [Zhang et al. \(2025b\)](#), although it goes back to [Gordon et al. \(2008\)](#).

**Definition 16 (Swap deviations with polynomial dimension; Zhang et al., 2025b)** Let  $m : \mathcal{P} \rightarrow \mathbb{R}^D$  be a feature map.  $\Phi^m$  is the set of all endomorphisms  $\phi$  that can be expressed by the matrix-vector product  $Mm(p)$  for some  $M = M(\phi) \in \mathbb{R}^{d \times D}$ . The set  $\Phi^m$  has dimension at most  $k := D \cdot d$ .

We do not include the affine component since one can simply append an additional coordinate of 1 to  $m(p)$  to capture that. A canonical example of Definition 16 is  $\ell$ -degree polynomials, which can

be obtained when  $m(p)$  outputs all  $\ell$ -wise (and lower) products of  $p$  and  $M$  contains the coefficients of the polynomials. In what follows, we assume that the first  $d$  coordinates of  $m(p)$  contain  $p$  itself.

Our goal is to minimize  $\Phi^m$ -regret with respect to a set  $\Phi^m$  with dimension  $k \leq \text{poly}(d)$ . Even when minimizing regret with respect to low-degree swap deviations, Zhang et al. (2024a) leveraged an earlier reduction by Hazan and Kale (2007) to show that the problem is intractable—specifically, PPAD-hard—unless the learner can commit to *mixed* strategies, that is, distributions in  $\mathcal{P}$ . As a result, it will be essential to consider mixed strategies. For a sequence of losses  $\ell_1, \dots, \ell_T \in \mathcal{L}$ , the  $\Phi^m$ -regret of a learner who plays a sequence of mixed strategies  $\mu_1, \dots, \mu_T \in \Delta(\mathcal{P})$  is defined as

$$\text{PolyDimSwapReg}_T = \sum_{t=1}^T \mathbb{E}_{p_t \sim \mu_t} \langle \ell_t, p_t \rangle - \min_{\phi \in \Phi^m} \sum_{t=1}^T \mathbb{E}_{p_t \sim \mu_t} \langle \ell_t, \phi(p_t) \rangle. \quad (3)$$

#### 4.1. Reducing to approachability

To reduce this problem into an approachability instance, we define the best-response map  $B(\ell) = m(b(\ell))$ , where, as before,  $b(\ell) = \arg \min_{p \in \mathcal{P}} \langle \ell, p \rangle$ , and the target set is defined as  $\mathcal{S} = \text{conv}\{\ell \otimes B(\ell) : \ell \in \mathcal{L}\}$ . Furthermore, for a sequence of distributions  $\mu_1, \dots, \mu_T \in \Delta(\mathcal{P})$  and losses  $\ell_1, \dots, \ell_T \in \mathcal{L}$ , we let  $\kappa_t = \mathbb{E}_{p_t \sim \mu_t} \ell_t \otimes m(p_t)$  and  $\bar{\kappa}_T = \frac{1}{T} \sum_{t=1}^T \kappa_t$ . Let  $J_d \in \mathbb{R}^{D \times d}$  be a matrix that contains 1 in the diagonal and 0 everywhere else. We can equivalently write (3) as

$$\begin{aligned} \text{PolyDimSwapReg}_T &= \sum_{t=1}^T \mathbb{E}_{p_t \sim \mu_t} \langle \ell_t, J_d m(p_t) \rangle - \min_{\phi \in \Phi^m} \sum_{t=1}^T \mathbb{E}_{p_t \sim \mu_t} \langle \ell_t, \phi m(p_t) \rangle \\ &= T \max_{M \in \Phi^m} \langle J_d - M, \bar{\kappa}_T \rangle, \end{aligned}$$

where the above derivation used the assumption that the first  $d$  coordinates of  $m(p)$  contain  $p$  itself. As before, we define the approachability loss measured in the Euclidean norm as  $\text{AppLoss}_T = \min_{s \in \mathcal{S}} \|\bar{\kappa}_T - s\|_F$ . The following lemma follows analogously to Lemma 4.

**Lemma 17**  $\text{PolyDimSwapReg}_T \leq 2T \text{AppLoss}_T \cdot \max_{M \in \Phi^m} \|M\|_F$ .

In particular, combining with Lemma 4, the approachability loss in this instance upper bounds both  $\text{PolyDimSwapReg}_T$  and  $\text{ProfileSwapReg}_T$ , unifying the notions introduced by Zhang et al. (2025b) and Arunachaleswaran et al. (2025). This gives rise to a stronger notion of correlated equilibrium in games.

To solve the induced approachability problem, we again rely on Algorithm 1. As before, the key is to maintain the invariance  $\langle \bar{\kappa}_{t-1} - \bar{s}_{t-1}, \kappa_t - s_t \rangle \leq 0$ . To do so, it suffices to show that for any matrix  $U \in \mathbb{R}^{d \times D}$ , there exists  $s \in \mathcal{S}$  and  $\mu \in \Delta(\mathcal{P})$  such that, no matter the choice of  $\ell \in \mathcal{L}$ ,  $\mathbb{E}_{p \sim \mu} \langle U, \ell \otimes m(p) - s \rangle \leq 0$ . We can obtain this guarantee by computing a pair of minimax equilibrium strategies  $(\mu_t, \ell_t^*)$  of the zero-sum game

$$\min_{\mu \in \Delta(\mathcal{P})} \max_{\ell \in \mathcal{L}} \mathbb{E}_{p \sim \mu} \ell^\top U_{t-1} m(p),$$

where  $U_{t-1} = \sum_{\tau=1}^{t-1} (\kappa_\tau - s_\tau)$ . The key point is that this is bilinear in terms of  $\mu$  and  $\ell$ , and the minimax theorem can be applied. This is where working with mixed strategies is crucial; otherwise, we would be dealing with the problem  $\min_{p \in \mathcal{P}} \max_{\ell \in \mathcal{L}} \ell^\top U m(p)$ , for which we cannot use the

---

**Algorithm 3** Response-based approachability for nonlinear swap deviations
 

---

**Input:** Horizon  $T$ , convex and compact sets  $\mathcal{P}, \mathcal{L} \subset \mathbb{R}^d$ , feature map  $m : \mathcal{P} \rightarrow \mathbb{R}^D$ , best-response map  $b : \mathcal{L} \rightarrow \mathcal{P}$

**Output:** Sequence of mixed strategies  $\mu_1, \dots, \mu_T \in \Delta(\mathcal{P})$

- 1 Initialize  $U_0 \leftarrow 0 \in \mathbb{R}^{d \times D}$
  - 2 **for**  $t = 1$  **to**  $T$  **do**
    - // Compute a minimax equilibrium in mixed strategies
    - 3 Compute a pair of minimax strategies  $(\mu_t, \ell_t^*)$  of the (bilinear) zero-sum game
 
$$\min_{\mu \in \Delta(\mathcal{P})} \max_{\ell \in \mathcal{L}} \mathbb{E}_{p \sim \mu} \ell^\top U_{t-1} m(p)$$
    - 4 Set  $s_t \leftarrow \ell_t^* \otimes m(b(\ell_t^*))$
    - 5 **Play** the mixed strategy  $\mu_t$  and **observe** the loss  $\ell_t$
    - 6 Set  $\kappa_t \leftarrow \mathbb{E}_{p_t \sim \mu_t} [\ell_t \otimes m(p_t)]$
    - 7 Update  $U_t \leftarrow U_{t-1} + (\kappa_t - s_t)$
  - 8 **end**
- 

minimax theorem. This is not a limitation of the approachability framework: Zhang et al. (2024a) showed that producing a sequence  $p_1, \dots, p_T$  that minimizes regret even with respect to degree-2 swap deviations is intractable. Our construction mirrors Algorithm 1, and is given in Algorithm 3.

**Lemma 18** *The distribution  $\mu_t \in \Delta(\mathcal{P})$  and the point  $s_t$  selected by Algorithm 3 satisfy  $\langle U_{t-1}, \kappa_t - s_t \rangle \leq 0$  for any adversarial loss  $\ell_t \in \mathcal{L}$ .*

Combining Lemmas 5, 17 and 18, we arrive at the following result.

**Theorem 19** *For any sequence of losses  $\ell_1, \dots, \ell_t \in \mathcal{L}$ , Algorithm 3 guarantees*

$$\text{PolyDimSwapReg}_T \leq 4\sqrt{T} \left( \max_{p \in \mathcal{P}} \|m(p)\|_2 \max_{\ell \in \mathcal{L}} \|\ell\|_2 \right) \left( \max_{M \in \Phi^m} \|M\|_F \right).$$

This significantly improves over the algorithm of Zhang et al. (2025b) in terms of the regret bound. On the other hand, one question we leave open for future work is under what conditions we can efficiently solve the bilinear saddle-point problem in Algorithm 3 with only oracle access to  $\mathcal{P}$ ; the challenge is that optimizing over the set of distributions  $\Delta(\mathcal{P})$  could be more complex.

## 5. A matching lower bound for linear swap regret

Finally, we construct an adaptive adversary that forces any learner to incur  $\Omega(d\sqrt{T})$  linear swap regret in expectation. We show this even when  $\mathcal{P}$  is centrally symmetric, matching Theorem 15. Our lower bound also implies that the algorithm of Gordon et al. (2008) is existential optimal algorithm for minimizing linear swap regret (with or without central symmetry) in view of Proposition 3.

To put this into context, we recall that for the simplex  $\mathcal{P} = \Delta^d$ , the most well-studied strategy set in the context of minimizing swap regret, there are algorithms that guarantee  $O(\sqrt{Td \log d})$  (linear) swap regret (Blum and Mansour, 2007), so our lower bound establishes a strict separation.

**Theorem 20** *Let  $T = \Omega(d^4)$  and  $\mathcal{P} \subset \mathbb{R}^d$  a centrally symmetric convex body. There exists an adaptive adversary that guarantees  $\mathbb{E}[\text{LinearSwapReg}_T] \geq \Omega(d\sqrt{T})$ .*

We now sketch the main steps in the proof of Theorem 20; details are available in Section D.

**The setup** The strategy set we consider is  $\mathcal{P} = B_1 \times B_\infty$ ; here,  $B_1$  is the  $\ell_1$  ball with radius 1 centered at 0 and  $B_\infty$  is the  $\ell_\infty$  ball with radius 1 centered at 0. We denote the strategy of the learner in each round  $t \in [T]$  by  $(x_t, p_t)$ , where  $x_t \in B_1$  and  $p_t \in B_\infty$ . Furthermore, the loss is written as  $(y_t, \ell_t)$ , where  $y_t \in B_\infty$  and  $\ell_t \in B_1$ . In other words,  $x_t$  is pitted against  $y_t$  and  $p_t$  against  $\ell_t$ . This guarantees that  $\langle x_t, y_t \rangle \leq 1$  and  $\langle p_t, \ell_t \rangle \leq 1$ .<sup>4</sup> Furthermore,  $\mathcal{P}$  is centrally symmetric.

Our construction relies on a subset of  $\text{End}(\mathcal{P})$ , namely transformations of the form

$$\phi(x, p) = \begin{bmatrix} A & 0 \\ M & 0 \end{bmatrix} \begin{bmatrix} x \\ p \end{bmatrix} + \begin{bmatrix} x^* \\ 0 \end{bmatrix}, \quad (4)$$

where  $M \in [-1, 1]^{d \times d}$  and  $A$  and  $x^*$  are such that  $Ax + x^* \in B_1$  for all  $x \in B_1$ . For any  $M \in [-1, 1]^{d \times d}$  and  $x \in B_1$ , it follows that  $Mx \in B_\infty$ , so any function per (4) is indeed a valid endomorphism. Let  $\Phi \subset \text{End}(\mathcal{P})$  be the set of all such functions. The  $\Phi$ -regret of the learner can be expressed as

$$\Phi\text{-Regret}_T = \sum_{t=1}^T \langle y_t, x_t - (Ax_t + x^*) \rangle + \sum_{t=1}^T \langle \ell_t, p_t - Mx_t \rangle.$$

The idea in our construction is this: the adversary we will use the sequence of losses  $y_1, \dots, y_T \in B_\infty$  to force the sequence  $x_1, \dots, x_T$  to have two key properties: i) a large fraction of the strategies are close to a vertex of  $\Delta^d$  (local concentration), and ii) a large fraction of the vertices are played often with considerable probability (global anti-concentration). Any learner whose linear swap regret is sublinear will be forced to exhibit these properties (Lemma 21). Furthermore, when these conditions are satisfied, we will show that the second term in  $\Phi$ -regret,  $\sum_{t=1}^T \langle \ell_t, p_t - Mx_t \rangle$ , will be  $\Omega(d\sqrt{T})$  in expectation (Lemma 22). In other words, the more the first part of the strategy  $x_t$  mixes between vertices of  $\Delta^d$ , the harder it becomes to minimize linear swap regret in terms of  $p_t$ : (linear) swap regret becomes a stronger metric the more the learner changes their strategy.

**Enforcing vertex coverage** We first construct an adversary with the following property.

**Lemma 21** *Let  $T \geq 16d^2(d+1)^2$ . There exists an adaptive adversary who produces a sequence of losses  $y_1, \dots, y_T \in B_\infty$  such that one of the following holds.*

1.  $\sum_{t=1}^T \langle y_t, x_t - (Ax_t + x^*) \rangle \geq d\sqrt{T}$  for some endomorphism  $(A, x^*)$  on  $B_1$ .
2. If  $I := \{i \in [d] : \sum_{t=1}^T \mathbb{1}\{x_{t,i} \geq \frac{1}{8}\} \geq \frac{T}{16d}\}$ , then  $|I| \geq d/15$ .

The other sequence of losses we construct  $\ell_1, \dots, \ell_T$  guarantees  $\mathbb{E}\langle p_t, \ell_t \rangle = 0$ , so we have

$$\mathbb{E}[\Phi\text{-Regret}_T] \geq \sum_{t=1}^T \langle y_t, x_t - (Ax_t + x^*) \rangle.$$

4. Technically, this only guarantees  $\langle x_t, y_t \rangle + \langle p_t, \ell_t \rangle \leq 2$ , but one can simply rescale the losses by a factor of 2. Similarly, while the set we construct is embedded in  $\mathbb{R}^{2d}$ , the lower bound can be obtained in terms of  $d' := 2d$ .

This means that Item 1 implies  $\mathbb{E}[\Phi\text{-Regret}_T] \geq d\sqrt{T}$ . The adversary behind Lemma 21 makes each of the  $d$  vertices of  $\Delta^d$  highly profitable in each of  $d$  periods of equal length, while making all previously profitable vertices highly costly.

**Punishing vertex coverage** To complete the construction, we show that if the sequence  $x_1, \dots, x_T$  satisfies Item 2, the adversary can guarantee  $\mathbb{E}[\sum_{t=1}^T \langle \ell_t, p_t - Mx_t \rangle] \geq \Omega(d\sqrt{T})$ .

**Lemma 22** *Let  $T \geq 16d^2(d+1)^2$ . If the set  $I := \{i \in [d] : \sum_{t=1}^T \mathbb{1}\{x_{t,i} \geq \frac{1}{8}\} \geq \frac{T}{16d}\}$  satisfies  $|I| = \Omega(d)$ , there is an oblivious adversary who can produce a sequence of losses  $\ell_1, \dots, \ell_T \in B_1$  such that*

$$\mathbb{E} \left[ \max_{M \in [-1,1]^{d \times d}} \sum_{t=1}^T \langle \ell_t, p_t - Mx_t \rangle \right] \geq \Omega(d\sqrt{T}).$$

Combining Lemma 21 and Lemma 22, Theorem 20 follows.

## 6. Conclusions and future research

We introduced a new unifying approach for minimizing different notions of swap regret by connecting to the response-based approachability framework of Bernstein and Shimkin (2015). Compared to previous work of Daskalakis et al. (2025), the resulting algorithm attains an improved regret bound and is also computationally more efficient. Moreover, we established a matching  $\Omega(d\sqrt{T})$  lower bound for linear swap regret, establishing for the first time that the classic algorithm of Gordon et al. (2008) is existentially optimal for that metric, although it is computationally intractable (Daskalakis et al., 2025). An interesting question is to obtain stronger lower bounds for a set of swap deviations with polynomial dimension, such as low-degree polynomials. Moreover, our focus has been on the regime where  $T$  is significantly larger than the dimension. It is possible that one can improve over the  $\Omega(d\sqrt{T})$  barrier in the high dimensional regime by leveraging recent advances (Dagan et al., 2024; Peng and Rubinstein, 2024).

## Acknowledgments

We are grateful to Brian Hu Zhang for insightful discussions. HL is supported by NSF award IIS-1943607. GF was supported in part by the National Science Foundation award CCF-2443068, the Office of Naval Research grant N000142512296, and the AI2050 Early Career Fellowship.

## References

- Jacob Abernethy, Elad Hazan, and Alexander Rakhlin. Competing in the dark: An efficient algorithm for bandit linear optimization. In *Conference on Learning Theory (COLT)*, 2008.
- Jacob Abernethy, Peter L Bartlett, and Elad Hazan. Blackwell approachability and no-regret learning are equivalent. In *Conference on Learning Theory (COLT)*, 2011.
- Jacob Abernethy, Chansoo Lee, and Ambuj Tewari. Perturbation techniques in online learning and optimization. *Perturbations, Optimization, and Statistics*, 233:17, 2016.
- Eshwar Ram Arunachaleswaran, Natalie Collina, and Jon Schneider. Pareto-optimal algorithms for learning in games. In *Conference on Economics and Computation (EC)*, 2024.

- Eshwar Ram Arunachaleswaran, Natalie Collina, Yishay Mansour, Mehryar Mohri, Jon Schneider, and Balasubramanian Sivan. Swap regret and correlated equilibria beyond normal-form games. In *Conference on Economics and Computation (EC)*, 2025.
- Robert Aumann. Subjectivity and correlation in randomized strategies. *Journal of Mathematical Economics*, 1:67–96, 1974.
- Keith Ball. Volume ratios and a reverse isoperimetric inequality. *Journal of the London Mathematical Society*, 2(2):351–359, 1991.
- Keith Ball. Convex geometry and functional analysis. In *Handbook of the geometry of Banach spaces*, volume 1, pages 161–194. 2001.
- Andrey Bernstein and Nahum Shimkin. Response-based approachability with applications to generalized no-regret problems. *Journal of Machine Learning Research*, 16(1):747–773, 2015.
- David Blackwell. An analog of the minmax theorem for vector payoffs. *Pacific Journal of Mathematics*, 6:1–8, 1956.
- Avrim Blum and Yishay Mansour. From external to internal regret. *Journal of Machine Learning Research*, 8:1307–1324, 2007.
- Sébastien Bubeck. Introduction to online optimization. *Lecture notes*, 2(1):1–86, 2011.
- Sébastien Bubeck and Ronen Eldan. The entropic barrier: a simple and optimal universal self-concordant barrier. *arXiv:1412.1587*, 2014.
- Yang Cao, Xiaoyu Li, Zhao Song, Xin Yang, and Tianyi Zhou. Faster algorithm for structured john ellipsoid computation. *arXiv:2211.14407*, 2022.
- Yair Carmon, Yujia Jin, Aaron Sidford, and Kevin Tian. Variance reduction for matrix games. In *Conference on Neural Information Processing Systems (NeurIPS)*, 2019.
- Yair Carmon, Yujia Jin, Aaron Sidford, and Kevin Tian. Coordinate methods for matrix games. In *Foundations of Computer Science (FOCS)*, 2020.
- Yuval Dagan, Constantinos Daskalakis, Maxwell Fishelson, and Noah Golowich. From external to swap regret 2.0: An efficient reduction for large action spaces. In *Symposium on Theory of Computing (STOC)*, 2024.
- Christoph Dann, Yishay Mansour, Mehryar Mohri, Jon Schneider, and Balasubramanian Sivan. Rate-preserving reductions for Blackwell approachability. In *Conference on Learning Theory (COLT)*, 2025.
- Constantinos Daskalakis, Gabriele Farina, Noah Golowich, Tuomas Sandholm, and Brian Hu Zhang. A lower bound on swap regret in extensive-form games. *arXiv:2406.13116*, 2024.
- Constantinos Daskalakis, Gabriele Farina, Maxwell Fishelson, Charilaos Pipis, and Jon Schneider. Efficient learning and computation of linear correlated equilibrium in general convex games. In *Symposium on Theory of Computing (STOC)*, 2025.

- Victor De la Peña and Evarist Giné. *Decoupling: from dependence to independence*. Springer Science & Business Media, 2012.
- Yuan Deng, Jon Schneider, and Balasubramanian Sivan. Strategizing against no-regret learners. In *Conference on Neural Information Processing Systems (NeurIPS)*, 2019.
- Gabriele Farina and Charilaos Pipis. Polynomial-time linear-swap regret minimization in imperfect-information sequential games. In *Conference on Neural Information Processing Systems (NeurIPS)*, 2023.
- Gabriele Farina, Tommaso Bianchi, and Tuomas Sandholm. Coarse correlation in extensive-form games. In *Conference on Artificial Intelligence (AAAI)*, volume 34, pages 1934–1941, 2020.
- 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.
- Maxwell Fishelson, Noah Golowich, Mehryar Mohri, and Jon Schneider. High-dimensional calibration from swap regret. In *Conference on Neural Information Processing Systems (NeurIPS)*, 2025.
- Dean Foster and Rakesh Vohra. Calibrated learning and correlated equilibrium. *Games and Economic Behavior*, 21:40–55, 1997.
- Yoav Freund and Robert Schapire. A decision-theoretic generalization of on-line learning and an application to boosting. *Journal of Computer and System Sciences*, 55(1):119–139, 1997.
- Kaito Fujii. Bayes correlated equilibria, no-regret dynamics in bayesian games, and the price of anarchy. In *Conference on Learning Theory (COLT)*, 2025.
- Parikshit Gopalan, Adam Tauman Kalai, Omer Reingold, Vatsal Sharan, and Udi Wieder. Omnipredictors. In *Innovations in Theoretical Computer Science Conference (ITCS)*, 2021.
- Parikshit Gopalan, Michael Kim, and Omer Reingold. Swap agnostic learning, or characterizing omniprediction via multicalibration. In *Conference on Neural Information Processing Systems (NeurIPS)*, 2023.
- Geoffrey J Gordon, Amy Greenwald, and Casey Marks. No-regret learning in convex games. In *International Conference on Machine Learning (ICML)*, 2008.
- Amy Greenwald and Amir Jafari. A general class of no-regret learning algorithms and game-theoretic equilibria. In *Conference on Learning Theory (COLT)*, 2003.
- Michael D. Grigoriadis and Leonid G. Khachiyan. A sublinear-time randomized approximation algorithm for matrix games. *Operations Research Letters*, 18(2):53–58, 1995.
- Martin Grötschel, László Lovász, and Alexander Schrijver. *Geometric algorithms and combinatorial optimization*. Springer, 1993.
- Sergiu Hart and Andreu Mas-Colell. A simple adaptive procedure leading to correlated equilibrium. *Econometrica*, 68:1127–1150, 2000.

- Elad Hazan and Satyen Kale. Computational equivalence of fixed points and no regret algorithms, and convergence to equilibria. In *Conference on Neural Information Processing Systems (NIPS)*, 2007.
- Elad Hazan and Zohar Karnin. Volumetric spanners: an efficient exploration basis for learning. *The Journal of Machine Learning Research*, 17(1):4062–4095, 2016.
- Ursula Hébert-Johnson, Michael Kim, Omer Reingold, and Guy Rothblum. Multicalibration: Calibration for the (computationally-identifiable) masses. In *International Conference on Machine Learning (ICML)*, 2018.
- Fritz John. Extremum problems with inequalities as subsidiary conditions, studies and essays presented to R. Courant on his 60th birthday, january 8, 1948, 1948.
- Piyush Kumar and E Alper Yildirim. Minimum-volume enclosing ellipsoids and core sets. *Journal of Optimization Theory and applications*, 126(1):1–21, 2005.
- Yin Tat Lee and Aaron Sidford. Path finding methods for linear programming: Solving linear programs in  $o$  (vrank) iterations and faster algorithms for maximum flow. In *Symposium on Foundations of Computer Science (FOCS)*, 2014.
- Nick Littlestone and M. K. Warmuth. The weighted majority algorithm. *Information and Computation*, 108(2):212–261, 1994.
- 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.
- Erwin Lutwak, Deane Yang, and Gaoyong Zhang. John ellipsoids. *Proceedings of the London Mathematical Society*, 90(2):497–520, 2005.
- Yishay Mansour, Mehryar Mohri, Jon Schneider, and Balasubramanian Sivan. Strategizing against learners in bayesian games. In *Conference on Learning Theory (COLT)*, 2022.
- Aleksandar Nikolov, Kunal Talwar, and Li Zhang. The geometry of differential privacy: the sparse and approximate cases. In *Symposium on Theory of computing (STOC)*, 2013.
- Binghui Peng. High dimensional online calibration in polynomial time. In *Foundations of Computer Science (FOCS)*, 2025.
- Binghui Peng and Aviad Rubinfeld. Fast swap regret minimization and applications to approximate correlated equilibria. In *Symposium on Theory of Computing (STOC)*, 2024.
- Alexander Rakhlin and Karthik Sridharan. Optimization, learning, and games with predictable sequences. In *Conference on Neural Information Processing Systems (NeurIPS)*, 2013.
- Gilles Stoltz and Gábor Lugosi. Internal regret in on-line portfolio selection. *Machine Learning*, 59(1-2):125–159, 2005.
- Gilles Stoltz and Gábor Lugosi. Learning correlated equilibria in games with compact sets of strategies. *Games and Economic Behavior*, 59(1):187–208, 2007.

- Yukai Tang, Jean-Bernard Lasserre, and Heng Yang. Uncertainty quantification of set-membership estimation in control and perception: Revisiting the minimum enclosing ellipsoid. In *Learning for Dynamics & Control Conference (LADC)*, 2024.
- Michael J Todd. *Minimum-volume ellipsoids: Theory and algorithms*. SIAM, 2016.
- Michael J Todd and E Alper Yildirim. On Khachiyan’s algorithm for the computation of minimum-volume enclosing ellipsoids. *Discrete Applied Mathematics*, 155(13):1731–1744, 2007.
- John von Neumann. Zur Theorie der Gesellschaftsspiele. *Mathematische Annalen*, 100:295–320, 1928.
- Bernhard von Stengel and Françoise Forges. Extensive-form correlated equilibrium: Definition and computational complexity. *Mathematics of Operations Research*, 33(4):1002–1022, 2008.
- Brian Hu Zhang, Ioannis Anagnostides, Gabriele Farina, and Tuomas Sandholm. Efficient  $\Phi$ -regret minimization with low-degree swap deviations in extensive-form games. In *Conference on Neural Information Processing Systems (NeurIPS)*, 2024a.
- Brian Hu Zhang, Gabriele Farina, and Tuomas Sandholm. Mediator interpretation and faster learning algorithms for linear correlated equilibria in general extensive-form games. In *International Conference on Learning Representations (ICLR)*, 2024b.
- Brian Hu Zhang, Ioannis Anagnostides, Emanuel Tewelde, Ratip Emin Berker, Gabriele Farina, Vincent Conitzer, and Tuomas Sandholm. Expected variational inequalities. In *International Conference on Machine Learning (ICML)*, 2025a.
- Brian Hu Zhang, Ioannis Anagnostides, Emanuel Tewelde, Ratip Emin Berker, Gabriele Farina, Vincent Conitzer, and Tuomas Sandholm. Learning and computation of  $\Phi$ -equilibria at the frontier of tractability. In *Conference on Economics and Computation (EC)*, 2025b.
- Martin Zinkevich. Online convex programming and generalized infinitesimal gradient ascent. In *International Conference on Machine Learning (ICML)*, 2003.

## Appendix A. Further related work

The notion of linear swap regret—as a relaxation of full swap regret—was introduced by [Gordon et al. \(2008\)](#). The first efficient algorithm for minimizing linear swap regret in extensive-form games was provided more recently by [Farina and Pipis \(2023\)](#). In that setting, linear swap regret subsumes weaker notions of regret ([Farina et al., 2022](#)) linked to extensive-form correlated and coarse correlated equilibria ([von Stengel and Forges, 2008](#); [Farina et al., 2020](#)). A natural interpretation of linear correlated equilibria based on a mediator was later provided by [Zhang et al. \(2024b\)](#) (*cf.* [Fujii, 2025](#)). [Daskalakis et al. \(2025\)](#) observed that when the underlying strategy set  $\mathcal{P}$  is given explicitly as a polytope, there is an efficient membership oracle that ascertains whether a purported affine endomorphism is indeed one. By standard reductions ([Grötschel et al., 1993](#)), this enables efficient optimization over  $\text{End}(\mathcal{P})$ , and thereby efficient regret minimization. When  $\mathcal{P}$  is given explicitly as a polytope, [Zhang et al. \(2025a\)](#) provided an explicit representation of the set  $\text{End}(\mathcal{P})$ , which enables efficient algorithms without relying on ellipsoid.

Algorithms for minimizing  $\Phi$ -regret have traditionally either followed the template of [Gordon et al. \(2008\)](#), or rely on the approachability framework of [Blackwell \(1956\)](#). While these approaches are in a certain sense equivalent ([Abernethy et al., 2011](#)), there is nuance to this equivalence ([Dann et al., 2025](#)). A new algorithmic approach for minimizing swap regret was recently introduced by [Peng and Rubinstein \(2024\)](#) and [Dagan et al. \(2024\)](#), but their reduction is only meaningful in the very high dimensional regime. Interestingly, our  $\Omega(d\sqrt{T})$  lower bound for linear swap regret only applies when  $T \geq \Omega(d^4)$ ; it is possible that in the high dimensional regime one can improve over our  $O(d\sqrt{T})$  upper bound using some variation of the approach of [Peng and Rubinstein \(2024\)](#); [Dagan et al. \(2024\)](#) tailored to linear swap regret.

Tracing back to the work of [Foster and Vohra \(1997\)](#), it has been known that swap regret is intimately connected to calibration. This connection was further cultivated in recent work that addressed the complexity of calibration in high dimensions ([Fishelson et al., 2025](#); [Peng, 2025](#)). Relatedly, the notion of swap regret has found fertile ground in the line of work on multicalibration ([Hébert-Johnson et al., 2018](#)) and omniprediction ([Gopalan et al., 2021, 2023](#)).

The famous theorem of [John \(1948\)](#) states that every compact and convex set with nonempty interior has a unique maximum-volume inscribed ellipsoid—known as the John ellipsoid. This concept has been a mainstay in convex geometry ([Ball, 1991, 2001](#); [Lutwak et al., 2005](#); [Todd, 2016](#)), and has many important algorithmic applications (*e.g.*, [Lee and Sidford, 2014](#); [Hazan and Karnin, 2016](#); [Nikolov et al., 2013](#); [Tang et al., 2024](#)).

## Appendix B. Comparison of different approachability algorithms

In this section, we contrast Algorithm 1 with Blackwell’s algorithm. While both rely at their core on the Pythagorean lemma (Lemma 5), the key distinction is that Blackwell’s algorithm separates  $s$  from  $(\ell \otimes p, \ell)$  via the supporting hyperplane whereas Algorithm 1 separates via a hyperplane produced by solving the auxiliary minimax game (Figure 1).

Figure 1 depicts both  $\mathcal{S}$  and  $\mathcal{K}$ .  $U$  is pointing to the right, so the goal is guarantee that  $(\ell \otimes p, \ell)$  is left of  $s$ . We show two hyperplanes orthogonal to  $U$ : one is tangent to  $\mathcal{S}$  (support) and the other (minimax) corresponds to  $\langle U, \kappa \rangle = v$  where  $v$  is the value of the minimax game  $\min_{p \in \mathcal{P}} \max_{\ell \in \mathcal{L}} \langle U, (\ell \otimes p, \ell) \rangle$ . The minimax hyperplane is necessarily left of the supporting hyperplane.

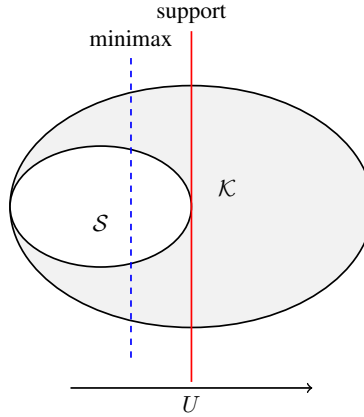


Figure 1: Illustration of how Algorithm 1 produces a separating hyperplane (minimax) versus Blackwell's choice (support).

In Blackwell's algorithm,  $U_t = \bar{\kappa}_t - \text{Proj}_S(\bar{\kappa}_t)$ , and  $s_t \in \arg \max_{s' \in S} \langle U_{t-1}, s' \rangle$  is selected as the point of tangency of the supporting hyperplane. Then  $p_t \in \mathcal{P}$  is chosen such that for any adversarial loss  $\ell_t \in \mathcal{L}$ ,  $(\ell_t \otimes p_t, \ell_t)$  is left of the supporting hyperplane; the existence of such a point  $p_t$  is guaranteed by the preconditions of approachability.

The main advantage of Algorithm 1 is that it only relies on a best-response oracle  $b : \mathcal{L} \rightarrow \mathcal{P}$ . Compared to Blackwell's algorithm, the choice of  $s_t$  is in a certain sense weaker, but the selection of  $p_t$  still ensures that  $\kappa_t$  will be on the left of the minimax hyperplane. Crucially, this only involves optimizing over  $\mathcal{P}$ .

### Appendix C. Extension under approximate minimax equilibria

We are concerned in this section with the approximate version of Algorithm 1 whereby  $(p_t, \ell_t^*)$  is an  $\epsilon_t$ -approximate minimax equilibrium (Algorithm 4). We say that  $(p_t, \ell_t^*)$  is an  $\epsilon_t$ -approximate minimax equilibrium if the duality gap—the sum of the players' best-response gaps—is at most  $\epsilon_t$ . Compared to Algorithm 1, Algorithm 4 normalizes the payoff matrix.

To account for this error in the analysis, we begin by extending the Pythagorean lemma (Lemma 5).

**Lemma 23** *Let  $v_1, \dots, v_T \in \mathbb{R}^d$  be a sequence of points each with Euclidean norm bounded by  $B$ . Suppose further that for all  $t \geq 2$ ,*

$$\left\langle \frac{1}{t-1} \sum_{\tau=1}^{t-1} v_\tau, v_t \right\rangle \leq \epsilon_t$$

for some  $\epsilon_t \geq 0$ . Then  $\left\| \sum_{\tau=1}^T v_\tau \right\|_2 \leq \sqrt{TB^2 + 2 \sum_{t=2}^T (t-1)\epsilon_t}$ .

**Proof** Let  $S_t := \sum_{\tau=1}^t v_\tau$ . We have

$$\begin{aligned} \|S_t\|_2^2 &= \|S_{t-1} + v_t\|_2^2 \\ &= \|S_{t-1}\|_2^2 + \|v_t\|_2^2 + 2 \langle S_{t-1}, v_t \rangle. \end{aligned}$$

For  $t = 1$ ,  $\|S_1\|_2^2 = \|v_1\|_2^2 \leq B^2$ . For  $t \geq 2$ , we use the assumption that  $\langle S_{t-1}, v_t \rangle \leq (t-1)\epsilon_t$  to bound

$$\|S_t\|_2^2 \leq \|S_{t-1}\|_2^2 + B^2 + 2(t-1)\epsilon_t.$$

Summing this inequality over  $t = 1, \dots, T$ , the sum telescopes, yielding

$$\|S_T\|_2^2 \leq \sum_{t=1}^T B^2 + 2 \sum_{t=2}^T (t-1)\epsilon_t = TB^2 + 2 \sum_{t=2}^T (t-1)\epsilon_t.$$

Taking the square root completes the proof.  $\blacksquare$

We next adapt Lemma 6.

**Lemma 24** *The points  $p_t$  and  $s_t$  selected by Algorithm 4 satisfy  $\langle \bar{U}_{t-1}, \kappa_t - s_t \rangle \leq \epsilon_t$  for any adversarial loss  $\ell_t \in \mathcal{L}$ , where  $\kappa_t = (\ell_t \otimes p_t, \ell_t)$ .*

**Proof** Let  $g_t(p, \ell) := \langle \bar{U}_{t-1}, (\ell \otimes p, \ell) \rangle$  be the payoff function of the auxiliary minimax game at step  $t$ , as defined in Algorithm 4. Since  $\mathcal{P}$  and  $\mathcal{L}$  are both convex and compact and  $g$  is bilinear, the minimax theorem applies (von Neumann, 1928). We let  $V_t$  be the value of the game,

$$V_t := \min_{p \in \mathcal{P}} \max_{\ell \in \mathcal{L}} g_t(p, \ell) = \max_{\ell \in \mathcal{L}} \min_{p \in \mathcal{P}} g_t(p, \ell).$$

We let  $\epsilon_{p_t}$  and  $\epsilon_{\ell_t}$  be the best-response gap of the two players, respectively, which satisfy  $\epsilon_{p_t} + \epsilon_{\ell_t} = \epsilon_t$ . It follows that for any (realized) loss  $\ell_t \in \mathcal{L}$ ,

$$\langle \bar{U}_{t-1}, \kappa_t \rangle = g_t(p_t, \ell_t) \leq V_t + \epsilon_{p_t}.$$

Furthermore,

$$V_t \leq \min_{p \in \mathcal{P}} \langle \bar{U}_{t-1}, (\ell_t^* \otimes p, \ell_t^*) \rangle + \epsilon_{\ell_t} \leq \langle \bar{U}_{t-1}, (\ell_t^* \otimes b(\ell_t^*), \ell_t^*) \rangle + \epsilon_{\ell_t} = \langle \bar{U}_{t-1}, s_t \rangle + \epsilon_{\ell_t}.$$

Combining these inequalities completes the proof.  $\blacksquare$

Combining with Lemma 23, it follows that

$$\min_{s \in \mathcal{S}} \|\bar{\kappa}_T - s\|_2 \leq \|\bar{\kappa}_T - \bar{s}_T\|_2 \leq \sqrt{\frac{4B^2}{T} + \epsilon} \leq \frac{2B}{\sqrt{T}} + \epsilon,$$

where  $B = \max_{\ell \in \mathcal{L}} \|\ell\|_2 \max_{p \in \mathcal{P}} \sqrt{\|p\|_2^2 + 1}$  and we assumed  $\epsilon_t = \epsilon$  for all  $t \in [T]$ . We thus arrive at the following refinement of Theorem 7.

**Theorem 25** *For any sequence of losses  $\ell_1, \dots, \ell_t \in \mathcal{L}$ , Algorithm 4 with  $\epsilon_t = \epsilon$  for all  $t \in [T]$  guarantees*

$$\text{LinearSwapReg}_T \leq 2 \left( 2\sqrt{T} \left( \max_{p \in \mathcal{P}, \ell \in \mathcal{L}} \|\ell\|_2 \sqrt{\|p\|_2^2 + 1} \right) + \epsilon T \right) \left( \max_{\phi \in \text{End}(\mathcal{P})} \|\phi\|_F \right).$$

*In particular, if  $\mathcal{P}$  is a centrally symmetric convex body in John's position,*

$$\text{LinearSwapReg}_T \leq 2\sqrt{2d} \left( 2\sqrt{T} \left( \sqrt{d+1} \right) + \epsilon T \right).$$

---

**Algorithm 4** Approximate response-based approachability
 

---

**Input:** Horizon  $T$ , convex sets  $\mathcal{P}, \mathcal{L} \subset \mathbb{R}^d$ , best response map  $b : \mathcal{L} \rightarrow \mathcal{P}$ 
**Output:** Sequence of strategies  $p_1, \dots, p_T$ 

- 1 Initialize  $\bar{U}_0 \leftarrow 0 \in \mathbb{R}^{d \times (d+1)}$
  - 2 **for**  $t = 1$  **to**  $T$  **do**
  - 3     Compute a pair of  $\epsilon_t$ -approximate minimax strategies  $(p_t, \ell_t^*)$  of the (bilinear) zero-sum game
 
$$\min_{p \in \mathcal{P}} \max_{\ell \in \mathcal{L}} \langle \bar{U}_{t-1}, (\ell \otimes p, \ell) \rangle,$$
  - 4     Set  $s_t \leftarrow (\ell_t^* \otimes b(\ell_t^*), \ell_t^*)$
  - 5     **Play** the strategy  $p_t$  and **observe** the loss  $\ell_t$
  - 6     Set  $\kappa_t \leftarrow (\ell_t \otimes p_t, \ell_t)$
  - 7     Update  $U_t \leftarrow U_{t-1} + (\kappa_t - s_t)$  and  $\bar{U}_t = \frac{1}{t} U_t$
  - 8 **end**
- 

As a result, if we want to guarantee that the average linear swap regret incurred by the learner is at most  $\delta > 0$ , it suffices to run  $T = O(\frac{d^2}{\delta^2})$  iterations each with precision  $\epsilon = O(\frac{\delta}{\sqrt{d}})$ .

As discussed in the main body, the problem of approximate minimax equilibrium computation is amenable to (external) no-regret learning (Freund and Schapire, 1997). The best algorithm for this problem depends on the desired precision. For example, the sublinear algorithm of Grigoriadis and Khachiyan (1995) computes an  $\epsilon$ -approximate minimax equilibrium in time  $\tilde{O}(d/\epsilon^2)$  in the  $\ell_1$ - $\ell_1$  geometry. In a moderately higher precision regime, improved algorithms are known (Rakhlin and Sridharan, 2013; Carmon et al., 2019, 2020). More concretely, under linear optimization oracles for  $\mathcal{P}$  and  $\mathcal{L}$ , an  $\epsilon$ -approximate minimax equilibrium can be computed through follow the perturbed leader (e.g., Abernethy et al., 2016). Interestingly, the underlying sequence of minimax games to be solved changes slowly; it is possible that this can be leveraged to obtain further algorithmic improvements.

From a more practical standpoint, reducing linear swap regret to a sequence of minimax problems offers a significant advantage. Recent progress in practical zero-sum game solving can now be leveraged for the problem of minimizing linear swap regret.

## Appendix D. Omitted proofs

This section contains the proofs omitted from the main body. We begin with Lemma 4.

**Lemma 4**  $\text{LinearSwapReg}_T \leq 2T \text{AppLoss}_T \cdot \max_{\phi \in \text{End}(\mathcal{P})} \|\phi\|_F$ .

**Proof** Let  $s^* = \arg \min_{s \in \mathcal{S}} \|\bar{\kappa}_T - s\|_F$ . By definition,  $s^*$  can be decomposed as  $s^* = \sum_{\ell \in \mathcal{L}} w_\ell (\ell \otimes b(\ell), \ell)$ . For any affine endomorphism  $\phi = (M, a) \in \text{End}(\mathcal{P})$ ,

$$\begin{aligned} \langle (I_d, 0) - (A, c), s^* \rangle &= \sum_{\ell \in \mathcal{L}} w_\ell (\langle (I_d, 0), (\ell \otimes b(\ell), \ell) \rangle - \langle (M, a), (\ell \otimes b(\ell), \ell) \rangle) \\ &= \sum_{\ell \in \mathcal{L}} w_\ell (\langle \ell, b(\ell) \rangle - (\langle \ell, Mb(\ell) \rangle + \langle \ell, a \rangle)) \\ &= \sum_{\ell \in \mathcal{L}} w_\ell (\langle \ell, b(\ell) \rangle - \langle \ell, \phi(b(\ell)) \rangle) \leq 0 \end{aligned} \quad (5)$$

since  $b(\ell)$  minimizes  $\langle \ell, p \rangle$  (by definition) and  $\phi(b(\ell)) \in \mathcal{P}$  ( $\phi \in \text{End}(\mathcal{P})$ ). As a result, using (5),

$$\begin{aligned} \text{LinearSwapReg}_T &= T \max_{(M, a) \in \text{End}(\mathcal{P})} \langle (I_d, 0) - (M, a), \bar{\kappa}_T \rangle \\ &\leq T \max_{(M, a) \in \text{End}(\mathcal{P})} \langle (I_d, 0) - (M, a), \bar{\kappa}_T - s^* \rangle \\ &\leq T \max_{(M, a) \in \text{End}(\mathcal{P})} \|(I_d, 0) - (M, a)\|_F \|\bar{\kappa}_T - s^*\|_F, \end{aligned}$$

where the last inequality is Cauchy-Schwarz. This completes the proof since  $\|(I_d, 0) - (M, a)\|_F \leq \|(I_d, 0)\|_F + \|(M, a)\|_F$  and  $(I_d, 0) \in \text{End}(\mathcal{P})$ .  $\blacksquare$

The correctness of Algorithm 1 hinges on the Pythagorean lemma (Lemma 5). The main precondition of that lemma is satisfied by virtue of the minimax theorem, as formalized below.

**Lemma 6** *The points  $p_t$  and  $s_t$  selected by Algorithm 1 satisfy  $\langle U_{t-1}, \kappa_t - s_t \rangle \leq 0$  for any adversarial loss  $\ell_t \in \mathcal{L}$ , where  $\kappa_t = (\ell_t \otimes p_t, \ell_t)$ .*

**Proof** Let  $g_t(p, \ell) := \langle U_{t-1}, (\ell \otimes p, \ell) \rangle$  be the payoff function of the auxiliary minimax game at step  $t$ , as defined in Algorithm 1. Since  $\mathcal{P}$  and  $\mathcal{L}$  are both convex and compact and  $g$  is bilinear, the minimax theorem applies (von Neumann, 1928). We let  $V_t$  be the value of the game,

$$V_t := \min_{p \in \mathcal{P}} \max_{\ell \in \mathcal{L}} g_t(p, \ell) = \max_{\ell \in \mathcal{L}} \min_{p \in \mathcal{P}} g_t(p, \ell).$$

By the definition of the learner's strategy  $p_t \in \arg \min_{p \in \mathcal{P}} \max_{\ell \in \mathcal{L}} g_t(p, \ell)$ , we have that for any (realized) loss  $\ell_t \in \mathcal{L}$ ,

$$\langle U_{t-1}, \kappa_t \rangle = g_t(p_t, \ell_t) \leq V_t. \quad (6)$$

Furthermore, since  $\ell_t^* \in \arg \max_{\ell \in \mathcal{L}} \min_{p \in \mathcal{P}} g_t(p, \ell)$ ,

$$V_t = \min_{p \in \mathcal{P}} g_t(p, \ell_t^*).$$

Now, the target point is chosen as  $s_t = (\ell_t^* \otimes b(\ell_t^*), \ell_t^*) \in \mathcal{S}$ . We can thus obtain the upper bound

$$V_t = \min_{p \in \mathcal{P}} \langle U_{t-1}, (\ell_t^* \otimes p, \ell_t^*) \rangle \leq \langle U_{t-1}, (\ell_t^* \otimes b(\ell_t^*), \ell_t^*) \rangle = \langle U_{t-1}, s_t \rangle. \quad (7)$$

Combining (6) and (7),

$$\langle U_{t-1}, \kappa_t \rangle \leq V_t \leq \langle U_{t-1}, s_t \rangle \implies \langle U_{t-1}, \kappa_t - s_t \rangle \leq 0.$$

■

We continue by establishing a basic invariance that allows suitably transforming the strategy and loss sets; this is essential for our preconditioning approach.

**Lemma 9** *Let  $A \in \mathbb{R}^{d \times d}$  be an invertible matrix. Consider the transformed strategy set  $\mathcal{P}' = A\mathcal{P} := \{Ap : p \in \mathcal{P}\}$  and loss set  $\mathcal{L}' = A^{-\top}\mathcal{L} := \{A^{-\top}\ell : \ell \in \mathcal{L}\}$ . For any sequence of strategies  $(p_t)_{t=1}^T$  in  $\mathcal{P}$  and sequence of losses  $(\ell_t)_{t=1}^T$  in  $\mathcal{L}$ , define the transformed sequences  $p'_t = Ap_t$  and  $\ell'_t = A^{-\top}\ell_t$  for all  $t \in [T]$ . Then the linear swap regret  $\text{LinearSwapReg}_T$  of  $(p_t)_{t=1}^T$  with respect to  $(\ell_t)_{t=1}^T$  is equal to the linear swap regret  $\text{LinearSwapReg}'_T$  of  $(p'_t)_{t=1}^T$  with respect to  $(\ell'_t)_{t=1}^T$ .*

**Proof** First, we observe that the instantaneous loss is preserved. That is, for any  $t$ ,  $\langle \ell'_t, p'_t \rangle = \langle A^{-\top}\ell_t, Ap_t \rangle = \langle \ell_t, p_t \rangle$ .

Next, we establish a correspondence between  $\text{End}(\mathcal{P})$  and  $\text{End}(\mathcal{P}')$ . We consider the mapping  $\Psi : \text{End}(\mathcal{P}) \rightarrow \text{End}(\mathcal{P}')$  defined by  $\Psi(\phi) = A \circ \phi \circ A^{-1}$ . For any  $\phi \in \text{End}(\mathcal{P})$  and any  $p' \in \mathcal{P}'$ , we have  $A^{-1}p' \in \mathcal{P}$ , so  $\phi(A^{-1}p') \in \mathcal{P}$ , which implies  $\Psi(\phi)(p') = A(\phi(A^{-1}p')) \in \mathcal{P}'$ . Thus,  $\Psi$  maps into  $\text{End}(\mathcal{P}')$ . Since  $A$  is invertible, this mapping is a bijection; for any  $\phi' \in \text{End}(\mathcal{P}')$ , the inverse mapping is given by  $\phi = A^{-1} \circ \phi' \circ A$ .

Now, for any  $\phi \in \text{End}(\mathcal{P})$  and its corresponding  $\phi' = \Psi(\phi) \in \text{End}(\mathcal{P}')$ , we have

$$\sum_{t=1}^T \langle \ell'_t, \phi'(p'_t) \rangle = \sum_{t=1}^T \langle A^{-\top}\ell_t, (A \circ \phi \circ A^{-1})(Ap_t) \rangle = \sum_{t=1}^T \langle A^{-\top}\ell_t, A\phi(p_t) \rangle = \sum_{t=1}^T \langle \ell_t, \phi(p_t) \rangle.$$

This implies that  $\text{LinearSwapReg}_T = \text{LinearSwapReg}'_T$ , completing the proof. ■

We next bound the Frobenius norm of endomorphisms. We first point out that the containment  $B_2 \subseteq \mathcal{P} \subseteq \sqrt{d}B_2$  guaranteed by John's theorem (Theorem 10) does not suffice on its own.

**Lemma 11** *There exists a convex set  $\mathcal{P} \subset \mathbb{R}^d$  satisfying  $B_2 \subseteq \mathcal{P} \subseteq \sqrt{d}B_2$  and a linear map  $M$  such that  $M(\mathcal{P}) \subseteq \mathcal{P}$  but  $\|M\|_F = \Omega(d)$ .*

**Proof** We define  $\mathcal{P}$  to be an ellipsoid. Specifically, we choose axis lengths

$$s_1 = \dots = s_k = \sqrt{d}, \quad s_{k+1} = \dots = s_d = 1,$$

where  $k = \lfloor d/2 \rfloor$ , so that

$$\mathcal{P} := \left\{ p \in \mathbb{R}^d : \sum_{i=1}^d \frac{p_i^2}{s_i^2} \leq 1 \right\}.$$

Since  $s_i \geq 1$  for each  $i$ , any point  $p \in \mathbb{R}^d$  such that  $\|p\|_2 \leq 1$  satisfies  $\sum_{i=1}^d p_i^2/s_i^2 \leq 1$ , so  $B_2 \subseteq \mathcal{P}$ . Furthermore, since  $s_i \leq \sqrt{d}$  for each  $i$ , it follows that  $\mathcal{P} \subseteq \sqrt{d}B_2$ .

Now, let  $D = \text{diag}(s_1, \dots, s_d)$ , so that  $\mathcal{P} = \{p : \|D^{-1}p\| \leq 1\}$ . For an orthogonal matrix  $U$ , we define

$$M = DUD^{-1}.$$

Since  $D^{-1}MD = U$  is orthogonal, we have that for any  $p \in \mathcal{P}$ ,

$$\|D^{-1}Mp\|_2 = \|U(D^{-1}p)\|_2 = \|D^{-1}p\|_2 \leq 1.$$

Thus,  $M(\mathcal{P}) \subseteq \mathcal{P}$ , so  $M$  is an endomorphism on  $\mathcal{P}$ .

To make  $\|M\|_F$  large, we choose  $U$  to be a permutation matrix that maps the  $k$  indices with  $s_j = \sqrt{d}$  into distinct indices with  $s_i = 1$ . We can then write

$$M_{i,j} = \frac{s_i}{s_j} U_{i,j},$$

so each column has exactly one nonzero entry. Moreover, there are  $k$  columns in which the magnitude of the nonzero entry is  $\sqrt{d}$ . As a result, it follows that  $\|M\|_F^2 \geq k(\sqrt{d})^2 = d \lfloor d/2 \rfloor$ .  $\blacksquare$

For this reason, our next proof makes crucial use of John's decomposition.

**Lemma 12** *Let  $\mathcal{P} \subset \mathbb{R}^d$  be a convex body in John's position and  $\phi(p) = Mp + a$  be an affine endomorphism of  $\mathcal{P}$ . Then we have  $\|M\|_F \leq \sqrt{d + \|a\|_2^2}$ . In particular, if  $\mathcal{P}$  is centrally symmetric,  $\|a\|_2 \leq \sqrt{d}$  and  $\|M\|_F \leq \sqrt{2d}$ .*

**Proof** John's theorem (Theorem 10) guarantees the existence of a finite set of points  $\xi_1, \dots, \xi_m \in \partial\mathcal{P} \cap \partial B_2$  and weights  $c_1, \dots, c_m > 0$  such that  $\sum_{i=1}^m c_i \xi_i \otimes \xi_i = I_d$  and  $\sum_{i=1}^m c_i \xi_i = 0$ . Since  $\|\xi_i\|_2 = 1$  for each  $i \in [m]$ ,

$$d = \text{tr}(I_d) = \text{tr} \left( \sum_{i=1}^m c_i \xi_i \otimes \xi_i \right) = \sum_{i=1}^m c_i \text{tr}(\xi_i \otimes \xi_i) = \sum_{i=1}^m c_i \|\xi_i\|_2^2 = \sum_{i=1}^m c_i.$$

For each contact point  $\xi_i$ , it holds that  $\langle p, \xi_i \rangle \leq 1$  for any  $p \in \mathcal{P}$ . In particular, since  $\phi(\mathcal{P}) \subseteq \mathcal{P}$ ,

$$\max_{p \in \mathcal{P}} \langle Mp + a, \xi_i \rangle \leq 1.$$

Thus, since  $p = M^\top \xi_i / \|M^\top \xi_i\|_2 \in \mathcal{P}$  (by the fact that  $B_2 \subseteq \mathcal{P}$ ),

$$\|M^\top \xi_i\|_2 \leq \max_{p \in \mathcal{P}} \langle p, M^\top \xi_i \rangle \leq 1 - \langle a, \xi_i \rangle.$$

Finally,

$$\begin{aligned} \|M\|_F^2 &= \text{tr}(M^\top I_d M) = \text{tr} \left( M^\top \left( \sum_{i=1}^m c_i \xi_i \otimes \xi_i \right) M \right) = \sum_{i=1}^m c_i \text{tr}(M^\top \xi_i \xi_i^\top M) \\ &= \sum_{i=1}^m c_i \|M^\top \xi_i\|_2^2. \end{aligned}$$

Using the bound  $\|M^\top \xi_i\|_2 \leq 1 - \langle a, \xi_i \rangle$ ,

$$\begin{aligned} \|M\|_F^2 &\leq \sum_{i=1}^m c_i (1 - \langle a, \xi_i \rangle)^2 = \sum_{i=1}^m c_i - 2 \sum_{i=1}^m c_i \langle a, \xi_i \rangle + \sum_{i=1}^m c_i \langle a, \xi_i \rangle^2 \\ &= d - 2 \left\langle a, \sum_{i=1}^m c_i \xi_i \right\rangle + a^\top \left( \sum_{i=1}^m c_i \xi_i \otimes \xi_i \right) a \\ &= d + \|a\|_2^2. \end{aligned}$$

When  $\mathcal{P}$  is centrally symmetric, it follows that  $M \cdot 0 + a \in \mathcal{P}$  since  $0 \in \mathcal{P}$ , so  $\|a\|_2 \leq \sqrt{d}$  by Theorem 10.  $\blacksquare$

**Lemma 13** *Let  $\phi : p \mapsto Mp + a$  be an affine endomorphism on  $\mathcal{P}$  with  $0 \in \mathcal{P}$ . Then  $\psi : p \mapsto \frac{1}{3}Mp + \frac{1}{3}a$  is an endomorphism on  $\text{conv}(\mathcal{P}, -\mathcal{P})$ .*

**Proof** It suffices to show that  $\psi(\mathcal{P}) \subseteq \text{conv}(\mathcal{P}, -\mathcal{P})$  and  $\psi(-\mathcal{P}) \subseteq \text{conv}(\mathcal{P}, -\mathcal{P})$ . For any point  $p \in \mathcal{P}$ ,  $\psi(p) = \frac{1}{3}\phi(p) \in \frac{1}{3}\mathcal{P} \in \text{conv}(\mathcal{P}, -\mathcal{P})$ . Moreover,  $\psi(-p) = -\frac{1}{3}Mp + \frac{1}{3}a = -\frac{1}{3}(Mp + a) + \frac{2}{3}a = \frac{1}{3}(-\phi(p)) + \frac{2}{3}\phi(0) \in \text{conv}(\mathcal{P}, -\mathcal{P})$ . ■

To complete the proof of Theorem 15, it suffices to establish the following bound.

**Lemma 14** *If  $\mathcal{P}$  is a convex body such that  $\text{conv}(\mathcal{P}, -\mathcal{P})$  is in John's position and  $\mathcal{L}$  is such that  $|\langle p, \ell \rangle| \leq 1$  for all  $p \in \mathcal{P}$  and  $\ell \in \mathcal{L}$ , we have  $\max_{p \in \mathcal{P}, \ell \in \mathcal{L}} \|(\ell \otimes p, \ell)\|_F \leq \sqrt{d+1}$ .*

**Proof** By assumption, it follows that  $\mathcal{L}$  is the polar set of  $\text{conv}(\mathcal{P}, -\mathcal{P})$ . By Theorem 10, since  $\text{conv}(\mathcal{P}, -\mathcal{P})$  is in John's position, we have  $\|p\|_2 \leq \sqrt{d}$  for any  $p \in \mathcal{P}$ . Furthermore, since  $\text{conv}(\mathcal{P}, -\mathcal{P}) \supseteq B_2$ , it follows that  $\mathcal{L} \subseteq B_2^\circ = B_2$ . We conclude that  $\|\ell\|_2 \leq 1$ , and the claim follows. ■

We continue with the proofs from Section 4, beginning with Lemma 17.

**Lemma 17**  $\text{PolyDimSwapReg}_T \leq 2T\text{AppLoss}_T \cdot \max_{M \in \Phi^m} \|M\|_F$ .

**Proof** Let  $s^* = \arg \min_{s \in \mathcal{S}} \|\bar{\kappa}_T - s\|_F$ . By the definition of the target set  $\mathcal{S}$ , there exist weights  $w_\ell \geq 0$ ,  $\sum_{\ell \in \mathcal{L}} w_\ell = 1$  such that  $s^* = \sum_{\ell \in \mathcal{L}} w_\ell (\ell \otimes B(\ell))$ . For any function  $M \in \Phi^m$ ,

$$\langle J_d - M, s^* \rangle = \sum_{\ell \in \mathcal{L}} w_\ell (\langle \ell, b(\ell) \rangle - \langle \ell, MB(\ell) \rangle) \leq 0$$

since  $b(\ell) = \arg \min_{p \in \mathcal{P}} \langle \ell, p \rangle$ ,  $J_d B(\ell) = J_d m(b(\ell)) = b(\ell)$ , and  $MB(\ell) \in \mathcal{P}$ . As a result,

$$\begin{aligned} \text{PolyDimSwapReg}_T &= T \max_{M \in \Phi^m} \langle J_d - M, \bar{\kappa}_T \rangle \\ &\leq T \max_{M \in \Phi^m} \langle J_d - M, \bar{\kappa}_T - s^* \rangle \\ &\leq T \max_{M \in \Phi^m} \|J_d - M\|_F \|\bar{\kappa}_T - s^*\|_F. \end{aligned}$$

Given that  $\|J_d - M\|_F \leq \|J_d\|_F + \|M\|_F$  and  $J_d \in \Phi^m$ , the claim follows. ■

The proof of the following lemma is similar to that of Lemma 9, making crucial use of the minimax theorem.

**Lemma 18** *The distribution  $\mu_t \in \Delta(\mathcal{P})$  and the point  $s_t$  selected by Algorithm 3 satisfy  $\langle U_{t-1}, \kappa_t - s_t \rangle \leq 0$  for any adversarial loss  $\ell_t \in \mathcal{L}$ .*

**Proof** Let  $g_t(\mu, \ell) := \mathbb{E}_{p \sim \mu} \ell^\top U_{t-1} m(p_t)$  be the payoff function of the auxiliary zero-sum game. By the minimax theorem,

$$V_t := \min_{\mu \in \Delta(\mathcal{P})} \max_{\ell \in \mathcal{L}} g_t(\mu, \ell) = \max_{\ell \in \mathcal{L}} \min_{\mu \in \Delta(\mathcal{P})} g_t(\mu, \ell).$$

By the definition of the learner's strategy  $\mu_t \in \arg \min_{\mu} \max_{\ell} g_t(\mu, \ell)$ , we have that for any (realized) loss  $\ell_t \in \mathcal{L}$ ,

$$\mathbb{E}_{p_t \sim \mu_t} \ell_t^\top U_{t-1} m(p_t) = g_t(\mu_t, \ell_t) \leq V_t. \quad (8)$$

Moreover,

$$V_t = \min_{\mu \in \Delta(\mathcal{P})} \mathbb{E}_{p \sim \mu} \langle U_{t-1}, \ell_t^* \otimes m(p) \rangle \leq \langle U_{t-1}, \ell_t^* \otimes m(b(\ell_t^*)) \rangle = \langle U_{t-1}, s_t \rangle. \quad (9)$$

Combining (8) and (9),

$$\langle U_{t-1}, \kappa_t \rangle \leq V_t \leq \langle U_{t-1}, s_t \rangle. \quad \blacksquare$$

We conclude with the proofs from Section 5, leading to the proof of Theorem 20.

**Lemma 21** *Let  $T \geq 16d^2(d+1)^2$ . There exists an adaptive adversary who produces a sequence of losses  $y_1, \dots, y_T \in B_\infty$  such that one of the following holds.*

1.  $\sum_{t=1}^T \langle y_t, x_t - (Ax_t + x^*) \rangle \geq d\sqrt{T}$  for some endomorphism  $(A, x^*)$  on  $B_1$ .
2. If  $I := \{i \in [d] : \sum_{t=1}^T \mathbb{1}\{x_{t,i} \geq \frac{1}{8}\} \geq \frac{T}{16d}\}$ , then  $|I| \geq d/15$ .

**Proof** We consider the following adaptive adversary. We divide  $T$  into  $d$  periods, each comprising  $T/d$  rounds; we can assume without any loss of generality that  $T$  is divisible by  $d$ . In the first period, the adversary consistently selects the loss  $y_t = (-1, -\frac{1}{2}, \dots, -\frac{1}{2})$ . At the end of each period, the adversary calculates  $\sum_{\tau=1}^t \langle y_\tau, x_\tau - (Ax_\tau + x^*) \rangle$ , which is the cumulative linear swap regret up to that time  $t$ . If it exceeds  $d\sqrt{T}$ , it proceeds by selecting the loss  $y_t = (0, 0, \dots, 0)$  thereafter; we call this scenario *termination*. This is the only place where our adversary is adaptive. By construction, termination always implies that  $\sum_{t=1}^T \langle y_t, x_t - (Ax_t + x^*) \rangle \geq d\sqrt{T}$ . Assuming there is no termination, in the second period, the adversary selects the loss  $y_t = (0, -1, -\frac{1}{2}, \dots, -\frac{1}{2})$ . Subject to not experiencing termination, the loss in the last period is set to  $(0, 0, \dots, 0, -1)$ . This sequence of losses is designed to entice the learner to play each vertex of  $\Delta^d$  with high probability, in the sense of Lemma 21.

Let us assume that there was no termination. The first observation is that there is a fixed hindsight deviation  $x^* \in \Delta^d$  that secures (at most)  $-\frac{1}{2}t$  cumulative loss after  $t$  rounds, namely  $x_d^* = 1$ . As a result, if  $t \in \{T/d, 2T/d, \dots, T\}$ , corresponding to the end of a period, we have

$$\sum_{\tau=1}^t \langle x_\tau, y_\tau \rangle \leq -\frac{1}{2}t + d\sqrt{T}.$$

Otherwise, the learner would experience higher than  $d\sqrt{T}$  linear swap regret, which would result in termination. We use  $\mathcal{T}_k$  to denote all rounds in the  $k$ th period. By the definition of the sequence of losses  $y_1, \dots, y_T$ , we have that if  $t = Tk/d$ ,

$$\sum_{\tau=1}^t \langle x_\tau, y_\tau \rangle = \sum_{\kappa=1}^k \sum_{\tau \in \mathcal{T}_\kappa} \left( -x_{\tau, \kappa} - \frac{1}{2} \sum_{i > \kappa} x_{\tau, i} \right) \leq -\frac{t}{2} + d\sqrt{T}. \quad (10)$$

Furthermore, again for  $t = Tk/d$ , we consider the linear swap deviation

$$A : (x_1, \dots, x_{k-1}, x_k, x_{k+1}, \dots, x_d) \mapsto \left( x_1, \dots, x_{k-1}, \sum_{i \geq k} x_i, 0, \dots, 0 \right),$$

which clearly satisfies  $Ax \in B_1$  for all  $x \in B_1$ . Then

$$\sum_{\tau=1}^t \langle y_\tau, x_\tau - Ax_\tau \rangle = \sum_{\kappa=1}^k \sum_{\tau \in \mathcal{T}_\kappa} \left( \sum_{i > k} y_{\tau,i} x_{\tau,i} - y_{\tau,k} \sum_{i > k} x_{\tau,i} \right).$$

For  $\kappa < k$  and  $\tau \in \mathcal{T}_\kappa$ , it holds that  $y_{\tau,i} = y_{\tau,i'}$  for any  $i, i' \geq k$ . So,

$$\begin{aligned} \sum_{\tau=1}^t \langle y_\tau, x_\tau - Ax_\tau \rangle &= \sum_{\tau \in \mathcal{T}_k} \left( \sum_{i > k} y_{\tau,i} x_{\tau,i} - y_{\tau,k} \sum_{i > k} x_{\tau,i} \right) = \sum_{\tau \in \mathcal{T}_k} \left( -\frac{1}{2} \sum_{i > k} x_{\tau,i} + \sum_{i > k} x_{\tau,i} \right) \\ &= \frac{1}{2} \sum_{\tau \in \mathcal{T}_k} \sum_{i > k} x_{\tau,i}. \end{aligned}$$

As a result, for any  $k = 1, \dots, d$  and  $t = Tk/d$ ,

$$\sum_{\tau \in \mathcal{T}_k} \sum_{i > k} x_{\tau,i} \leq 2d\sqrt{T}.$$

So, continuing from (10),

$$-\sum_{\kappa=1}^k \sum_{\tau \in \mathcal{T}_\kappa} x_{\tau,\kappa} \leq -\frac{t}{2} + \frac{1}{2} \sum_{\kappa=1}^k \sum_{\tau \in \mathcal{T}_\kappa} \sum_{i > k} x_{\tau,i} + d\sqrt{T} \leq -\frac{t}{2} + d(k+1)\sqrt{T}.$$

In particular,

$$\sum_{k=1}^d \sum_{t \in \mathcal{T}_k} x_{t,k} \geq \frac{1}{2}T - d(d+1)\sqrt{T}.$$

When  $T \geq 16d^2(d+1)^2$ ,

$$\sum_{k=1}^d \sum_{t \in \mathcal{T}_k} x_{t,k} \geq \frac{1}{4}T.$$

Using the obvious inequality  $x_{t,k} \leq \mathbb{1}\{x_{t,k} \geq \frac{1}{8}\} + \frac{1}{8}$ ,

$$\sum_{k=1}^d \sum_{t \in \mathcal{T}_k} \mathbb{1}\left\{x_{t,k} \geq \frac{1}{8}\right\} \geq \frac{1}{8}T. \quad (11)$$

We define

$$I := \left\{ i \in [d] : \sum_{t=1}^T \mathbb{1}\left\{x_{t,i} \geq \frac{1}{8}\right\} \geq \frac{T}{16d} \right\}.$$

Based on this definition, we can write

$$\sum_{k=1}^d \sum_{t \in \mathcal{T}_k} \mathbb{1} \left\{ x_{t,k} \geq \frac{1}{8} \right\} \leq |I| \frac{T}{d} + (d - |I|) \frac{T}{16d}.$$

Combining with (11), the claim follows.  $\blacksquare$

**Lemma 22** *Let  $T \geq 16d^2(d+1)^2$ . If the set  $I := \{i \in [d] : \sum_{t=1}^T \mathbb{1} \{x_{t,i} \geq \frac{1}{8}\} \geq \frac{T}{16d}\}$  satisfies  $|I| = \Omega(d)$ , there is an oblivious adversary who can produce a sequence of losses  $\ell_1, \dots, \ell_T \in B_1$  such that*

$$\mathbb{E} \left[ \max_{M \in [-1,1]^{d \times d}} \sum_{t=1}^T \langle \ell_t, p_t - Mx_t \rangle \right] \geq \Omega(d\sqrt{T}).$$

**Proof** For each time  $t \in [T]$ , the adversary selects uniformly at random a pair  $j \neq j' \sim [d]$  and sets  $\ell_{t,j} = \frac{1}{2}$ ,  $\ell_{t,j'} = -\frac{1}{2}$ , and  $\ell_{t,j''} = 0$  for  $j'' \neq j, j'$ . Then  $\mathbb{E}[\langle \ell_t, p_t \rangle] = \langle \mathbb{E}[\ell_t], p_t \rangle = 0$ .

Furthermore, we write

$$-\sum_{t=1}^T \langle \ell_t, Mx_t \rangle = -\sum_{i=1}^d \sum_{j=1}^d M_{j,i} \sum_{t=1}^T (x_{t,i} \ell_{t,j}).$$

We select  $M \in [-1, 1]^{d \times d}$  so that

$$\max_{M \in [-1,1]^{d \times d}} -\sum_{t=1}^T \langle \ell_t, Mx_t \rangle = \sum_{i=1}^d \sum_{j=1}^d \left| \sum_{t=1}^T x_{t,i} \ell_{t,j} \right|.$$

Specifically,

$$M_{j,i} = \begin{cases} 1 & \text{if } \sum_{t=1}^T (x_{t,i} \ell_{t,j}) < 0, \\ -1 & \text{if } \sum_{t=1}^T (x_{t,i} \ell_{t,j}) \geq 0. \end{cases}$$

For a fixed  $i \in [d]$  and  $j \in [d]$ , let  $Z_{i,j} = \sum_{t=1}^T x_{t,i} \ell_{t,j}$ .  $(x_{t,i} \ell_{t,j})_{t \geq 1}$  is a conditionally symmetric martingale difference sequence with respect to the filtration  $\mathcal{F}_{t-1}$  generated by the history of play. Indeed,  $\mathbb{E}[x_{t,i} \ell_{t,j} \mid \mathcal{F}_{t-1}] = x_{t,i} \mathbb{E}[\ell_{t,j}] = 0$ . Let  $(\epsilon_1, \dots, \epsilon_T)$  be a sequence of i.i.d. Radamacher random variables independent of the play. By symmetry,  $Z_{i,j}$  has the same distribution as the decoupled sum  $\sum_{t=1}^T \epsilon_t x_{t,i} \ell_{t,j}$ . For any fixed realization of  $(x_{t,i} \ell_{t,j})_{t=1}^T$ , the Khintchine inequality (e.g. [De la Peña and Giné, 2012](#)) implies that there exists an absolute constant  $C > 0$  such that

$$\mathbb{E}_{\epsilon_1, \dots, \epsilon_T} \left[ \left| \sum_{t=1}^T \epsilon_t x_{t,i} \ell_{t,j} \right| \right] \geq C \sqrt{\sum_{t=1}^T x_{t,i}^2 \ell_{t,j}^2}.$$

As a result, taking the expectation over all realizations of  $(x_{t,i} \ell_{t,j})_{t=1}^T$ ,

$$\mathbb{E}[|Z_{i,j}|] \geq C \mathbb{E} \left[ \sqrt{\sum_{t=1}^T x_{t,i}^2 \ell_{t,j}^2} \right].$$

Now, by definition of the loss, we have  $\mathbb{E}[\ell_{t,j}^2] = \frac{2}{d} \frac{1}{4} = \frac{1}{2d}$ . For a given set  $\mathcal{T}_i \subseteq [T]$  with  $|\mathcal{T}_i| \geq T/(16d)$ , a Chernoff bound yields

$$\mathbb{P} \left( \sum_{t \in \mathcal{T}_i} \ell_{t,j}^2 \leq \frac{1}{2} \frac{|\mathcal{T}_i|}{2d} \right) \leq \exp \left( -\frac{|\mathcal{T}_i|}{16d} \right) \leq \exp \left( -\frac{T}{256d^2} \right) \leq \exp \left( -\frac{(d+1)^2}{16} \right) \leq \frac{1}{2},$$

where we used the fact that  $|\mathcal{T}_i| \geq T/(16d)$ ,  $T \geq 16d^2(d+1)^2$ , and  $d \geq 3$ . This in turn implies

$$\mathbb{E} \left[ \sqrt{\sum_{t \in \mathcal{T}_i} \ell_{t,j}^2} \right] \geq \sqrt{\frac{|\mathcal{T}_i|}{4d}} \mathbb{P} \left( \sum_{t \in \mathcal{T}_i} \ell_{t,j}^2 \geq \frac{|\mathcal{T}_i|}{4d} \right) \geq \frac{1}{4} \sqrt{\frac{|\mathcal{T}_i|}{d}}.$$

As a result, for any  $i \in I$ ,

$$\mathbb{E}[|Z_{i,j}|] \geq \frac{C}{8} \left( \frac{1}{d} \sqrt{T} \right).$$

Combining, we have

$$\max_{M \in [-1,1]^{d \times d}} - \sum_{t=1}^T \langle \ell_t, Mx_t \rangle \geq \sum_{i \in I} \sum_{j=1}^d \mathbb{E}[|Z_{i,j}|] \geq \Omega(d\sqrt{T}).$$

■

Combining Lemma 21 and Lemma 22, Theorem 20 follows.