Sparsity-Based Interpolation of External, Internal and Swap Regret

Zhou Lu*
Princeton University
zhoul@princeton.edu

Y. Jennifer Sun*
Princeton University
ys7849@princeton.edu

Zhiyu Zhang*
Harvard University
zhiyuz@seas.harvard.edu

Abstract

Focusing on the expert problem in online learning, this paper studies the interpolation of several performance metrics via ϕ-regret minimization, which measures the performance of an algorithm by its regret with respect to an arbitrary action modification rule ϕ. With d experts and $T \gg d$ rounds in total, we present a single algorithm achieving the instance-adaptive ϕ-regret bound

$$\tilde{O}\left(\min\left\{\sqrt{d - d_\phi^{\mathrm{unif}} + 1}, \sqrt{d - d_\phi^{\mathrm{self}}}\right\} \cdot \sqrt{T}\right),$$

where d_ϕ^{unif} is the maximum amount of experts modified identically by ϕ, and d_ϕ^{self} is the amount of experts that ϕ trivially modifies to themselves. By recovering the optimal $O(\sqrt{T \log d})$ external regret bound when $d_\phi^{\mathrm{unif}} = d$, the standard $\tilde{O}(\sqrt{T})$ internal regret bound when $d_\phi^{\mathrm{self}} = d - 1$ and the optimal $\tilde{O}(\sqrt{dT})$ swap regret bound in the worst case, we improve existing results in the intermediate regimes. In addition, the same algorithm achieves the optimal quantile regret bound, which corresponds to even easier settings of ϕ than the external regret.

Building on the classical reduction from ϕ-regret minimization to external regret minimization on stochastic matrices, our main idea is to further convert the latter to online linear regression using Haar-wavelet-inspired matrix features. Then, we apply a particular L_1-version of comparator-adaptive online learning algorithms to exploit the sparsity in this regression subroutine.

1 Introduction

We consider the distributional version of *Learning from Expert Advice* (LEA), which is a two-player repeated game between a learning algorithm and an adversary. Let $d \in \mathbb{Z}$, $d > 1$ be the total number of experts, and L be the loss range which is known by the algorithm in advance. In each (t-th) round, the two players interact as follows.

1. The learning algorithm picks a distribution $p_t \in \Delta(d)$ over the d experts, where $\Delta(d) \subset \mathbb{R}^d$ denotes the probability simplex, i.e., $\Delta(d) = \{p \in \mathbb{R}^d : p \geq 0, \|p\|_1 = 1\}$.
2. The adversary observes p_t and picks a vector $l_t \in [-L, L]^d$ specifying the losses of all experts.
3. The algorithm observes l_t and suffers a loss defined by the inner product $\langle p_t, l_t \rangle$.
4. The adversary determines whether the game terminates. If so, let T be the total number of rounds.

Targeting the regime of $T \gg d$, our goal is to design a learning algorithm that suffers low total loss $\sum_{t=1}^T \langle p_t, l_t \rangle$, without any additional assumption on the adversary.

It is well-known that such an objective can be approached by minimizing certain comparative performance metrics. Specifically, we study a class of performance metrics called the *ϕ-regret*, due to [GJ03]. Let $\mathcal{S}(d)$ be the collection of all linear functions mapping $\Delta(d)$ to itself, which can be equivalently expressed as d-by-d right

*Equal contribution, alphabetical order.

stochastic matrices (entries are nonnegative and each row sums up to 1). With any $\phi \in \mathcal{S}(d)$ called an *action modification rule*, the ϕ-regret is defined as

$$\mathrm{Regret}_T(\phi) := \sum_{t=1}^{T} \langle p_t, l_t \rangle - \sum_{t=1}^{T} \langle \phi(p_t), l_t \rangle .$$

Intuitively, the ϕ-regret compares the total loss of the algorithm at the end of the game to the total loss it would have obtained, had it transformed all its past actions according to ϕ (while assuming the adversary's past actions remain the same). If an algorithm guarantees that $\mathrm{Regret}_T(\phi) \leq f(\phi)$ for some function f, then by definition, its total loss can be upper-bounded by the oracle inequality

$$\sum_{t=1}^{T} \langle p_t, l_t \rangle \leq \min_{\phi \in \mathcal{S}(d)} \left[\sum_{t=1}^{T} \langle \phi(p_t), l_t \rangle + f(\phi) \right]. \tag{1}$$

The ϕ-regret can be viewed as a smooth interpolation of several important performance metrics with different power. Suppressing the dependence on L for simplicity,

- The most common *external regret* (or simply known as "the" regret) equals the supremum of $\mathrm{Regret}_T(\phi)$ over all the ϕ's that are constant functions. The celebrated *Multiplicative Weight Update* (MWU) algorithm [LW94, CBFH+97] achieves the optimal external regret upper bound, $O(\sqrt{T \log d})$.
- The *internal regret* [FV99] equals the supremum of $\mathrm{Regret}_T(\phi)$ over all the ϕ's that map exactly $d-1$ vertices of $\Delta(d)$ to themselves.[1] This is often motivated by game-theoretic applications, while [SL05] proved its quantitative advantage over the external regret in certain prediction problems. Running MWU over the targeted ϕ-class guarantees the standard internal regret bound, $O(\sqrt{T \log d})$ [CBL06, Chapter 4.4].
- The *swap regret* [BM07] is defined as the supremum of $\mathrm{Regret}_T(\phi)$ over the entire $\mathcal{S}(d)$. By construction it is at most d times the internal regret bound, but an even better $O(\sqrt{dT \log d})$ upper bound can be achieved via a well-known, computationally efficient *swap-to-external reduction* [BM07]. For the distributional version of LEA, this is recently shown to be optimal in the regime of $T \gg d$ we consider [DDFG24, PR24].
- The *quantile regret* [CFH09] compares the total loss of the algorithm to the total loss of the $\lceil \varepsilon d \rceil$-th best expert. In the language of the ϕ-regret, this amounts to considering the ϕ's that are not only constant, but also close to outputting the uniform distribution. It is known that the optimal upper bound is $O(\sqrt{T \log \varepsilon^{-1}})$ [OP16, NBC+21], which recovers the $O(\sqrt{T \log d})$ external regret as its own worst case ($\varepsilon = 1/d$).

Main result In this paper, we present a single ϕ-regret minimization algorithm (Algorithm 1 in Appendix A) that ties the above regimes together. Compared to the naive approach of aggregating the above specialized algorithms by MWU, our proposed algorithm achieves a ϕ-regret bound that depends on suitable complexity measures of each ϕ instance, leading to a sharper oracle inequality, Eq.(1). Specifically, with $e^{(i)} \in \mathbb{R}^d$ representing the canonical basis vector along the i-th coordinate, the complexity of ϕ is measured by the following two definitions.

Definition 1 (Uniformity). *The* uniformity *of $\phi \in \mathcal{S}(d)$, denoted by d_ϕ^{unif}, is defined as the frequency of the most frequent element in the size-d multiset $\{\phi(e^{(1)}), \ldots, \phi(e^{(d)})\}$.*

Definition 2 (Degree of self-map). *The* degree of self-map *of $\phi \in \mathcal{S}(d)$, denoted by d_ϕ^{self}, is defined as the amount of indices i satisfying $\phi(e^{(i)}) = e^{(i)}$.*

In plain words, d_ϕ^{unif} measures the maximum number of experts modified identically by ϕ, while d_ϕ^{self} measures the number of experts modified trivially by ϕ (i.e., modified to themselves). One could also interpret d_ϕ^{unif} as the similarity between ϕ and constant functions, while d_ϕ^{self} represents the similarity between ϕ and the *self-map*. In their easiest settings, $d_\phi^{\mathrm{unif}} = d$ and $d_\phi^{\mathrm{self}} = d - 1$ respectively recover the ϕ-classes of the external regret and the

[1] In other words, such a ϕ transforms the mass of p_t on one of the experts to the others.

internal regret, therefore accordingly, one would expect the ϕ-regret to be decreasing with respect to d_ϕ^{unif} and d_ϕ^{self}. In the other extreme we have $d_\phi^{\text{unif}} = 1$ and $d_\phi^{\text{self}} = 0$, and a sensible algorithm needs to still guarantee the optimal swap regret bound, $\tilde{O}(\sqrt{dT})$.

As a concrete realization of this reasoning, our algorithm guarantees the ϕ-regret bound

$$\text{Regret}_T(\phi) = \tilde{O}\left(\min\left\{\sqrt{d - d_\phi^{\text{unif}} + 1}, \sqrt{d - d_\phi^{\text{self}}}\right\} \cdot \sqrt{T}\right),$$

as well as the optimal quantile regret bound $O(\sqrt{T \log \varepsilon^{-1}})$. This corresponds to a natural interpolation of the $O(\sqrt{T \log d})$ external regret, the $\tilde{O}(\sqrt{T})$ internal regret and the $\tilde{O}(\sqrt{dT})$ swap regret, only sacrificing (multiplicatively) a constant factor compared to MWU and a polylog factor compared to [CBL06, Chapter 4.4] and [BM07], in their respective specialized regimes. Computationally, both the runtime and the memory of our algorithm are of the same order as [BM07].

Techniques Our algorithm is based on the following simple idea. It is already known that ϕ-regret minimization in LEA can be reduced to an external regret minimization problem on the stochastic matrix space $\mathcal{S}(d)$ [GGM08], for which solutions can be built as variants of mirror descent. Deviating from the latter part of this procedure, we apply a linear transform on the unconstrained matrix space $\mathbb{R}^{d\times d}$, converting all the comparing ϕ's to their corresponding *transform domain coefficients*. If ϕ is sparse on the transform domain, then we see it as "simple" – just like the intuition from Fourier transform, where time series consisting of very few frequencies are considered simple. Then, since learning a hypothesis ϕ is equivalent to learning its transform domain coefficient, we can apply a *sparsity-adaptive* online learning algorithm to perform this task, for which there are out-of-the-box options available [Ora23, Chapter 9].

Putting these together, our approach amounts to using a sparsity-adaptive online learning algorithm to solve a particular matrix linear regression subroutine. The crucial step is designing the *features* here (equivalently, the linear transform on the matrix space $\mathbb{R}^{d\times d}$), as we need to ensure its consistency with the inductive bias of our targeted regret bounds. Our construction is based on the *Haar wavelet* [Mal08], whose ability to sparsely represent low-variation signals has enabled several recent advances in online learning [BW19, ZCP23, JO24]. Along the way we address a number of technical challenges to be outlined shortly.

1.1 Related Work

Φ-regret The ϕ-regret we study is an instance-dependent version of a better-known concept called the Φ-regret, due to [GJ03]. With respect to any class $\Phi \subset \mathcal{S}(d)$ of action modification rules, the Φ-regret is defined as the supremum of $\text{Regret}_T(\phi)$ over all $\phi \in \Phi$. Due to its generality unifying external, internal and swap regret, further developments have been presented in numerous works afterwards, particularly regarding its connection to various equilibrium concepts in game theory [SL07, RST11, PRO$^+$22, BCM$^+$23, CDL$^+$24, ZAFS24].

Technically, we build on the well-known reduction from swap regret to external regret on the extended domain $\mathcal{S}(d)$, pioneered by [SL05, BM07, GGM08] and further developed by [Ito20]. In a conceptually different manner, one may also analyze the swap regret through the *subsequence regret* [Leh03, Rot23], and unifying algorithms have been proposed based on certain multi-objective formulations of online learning [LNPR22, HPY23, NRRX23] related to the L_∞-norm *Blackwell approachability* of the negative orthant [Bla56, Per15, Shi16]. The key idea here is to convert the subsequence regret to the regret of a "meta" LEA algorithm that reweighs different subsequences. In particular, utilizing the meta algorithm of [CLW21], the approach of [Rot23] can achieve an instance-dependent refinement of the $\tilde{O}(\sqrt{T})$ internal regret bound.

A recent breakthrough of [DDFG24] and [PR24] showed that the d-dependence of the time-averaged $\tilde{O}(\sqrt{d/T})$ swap regret bound can be improved exponentially, at the price of an exponentially worse dependence on T. This is orthogonal to the regime of $T \gg d$ we consider, but very intriguingly, their algorithms are also based on some sort of multi-resolution analysis. Closer to our setting, they showed that the $\tilde{O}(\sqrt{dT})$ swap regret is optimal in the distributional version of LEA with $T \gg d$, answering an open problem from earlier works.

Features and sparsity Our techniques are inspired by recent advances in *dynamic regret* minimization, which is the hardest regret notion that compares to an arbitrary sequence of predictions selected in hindsight. [ZCP23] presented a reduction from dynamic regret to external regret on the extended domain $\mathbb{R}^T$, which is reminiscent of the swap-to-external reduction discussed above; also see [JO24]. Central to this technique is the use of features to associate the complexity of a hypothesis to the sparsity of its linear representations. This suggests viewing the features of [BM07] as the canonical matrix basis, and our main conceptual takeaway is that the Haar wavelet, introduced to online learning by [BW19], produces matrix features that are better-aligned with the inherent structure of the external and internal regret.

Technical challenges We now highlight the technical challenges of this work along the lines of the above discussion. Regardless of the choice of features, the dynamic regret bound of [ZCP23] is $\tilde{O}(\sqrt{T}\cdot\sqrt{T})$ in the worst case, where one of the $\sqrt{T}$ is the iconic rate of online linear optimization, and the other one comes from the dimensionality of the extended domain $\mathbb{R}^T$. By default, it means the analogous approach for our problem would result in the suboptimal $\tilde{O}(\sqrt{d^2}\cdot\sqrt{T})$ swap regret. The key subtlety here is that the surrogate gradients from the swap-to-external reduction are well-structured, such that one could use *"first-order" gradient-adaptivity* to further shave off a $\sqrt{d}$ factor [BM07]. But this does not hold true for arbitrary features anymore.

Regarding this issue, we show that the Haar wavelet is particularly "compatible" with the gradient structure from the swap-to-external reduction, such that the optimal $\tilde{O}(\sqrt{dT})$ swap regret bound can still be recovered after incorporating specific Haar-wavelet-type matrix features. Our approach also involves a complexity-preserving augmentation of the hypothesis class $\mathcal{S}(d)$ (Section 2.1), as well as a projection technique that enforces the constraint $\mathcal{S}(d)$ (Section 2.4). Both are nontrivial constructions absent from [ZCP23].

1.2 Notation

Throughout this paper, $\langle x, y\rangle$ denotes the inner product of generic x and y. Specifically when acting on two matrices, it equals the Euclidean inner product of their vectorizations. $x\otimes y$ denotes the outer product of vectors. An interval (of integers) $[a:b]$ is defined as $\{a, a+1, \dots, b-1, b\}$. log denotes the natural logarithm if the base is omitted. $\boldsymbol{I}_d$ represents the d-by-d identity matrix. We use the following indexing rule: on a matrix x, x_i represents its i-th row while $x_{i,j}$ represents its (i,j)-th entry; on a vector x, x_i represents its i-th entry.

2 Technical Ingredients

The starting point of our approach is the celebrated ϕ-to-external reduction of [GGM08]. This is remarkably simple, with similar ideas also presented in several works around the same time [SL05, BM07]. First, let us assume access to a computational oracle which, given any input $\phi\in\mathcal{S}(d)$, returns a fixed point $p\in\Delta(d)$ satisfying $p=\phi(p)$. The existence of p is due to Brouwer's fixed-point theorem, and in practice, it can be found via a generic linear system solver.

Now, regarding the LEA problem, suppose we have some procedure that generates a linear map $\phi_t\in\mathcal{S}(d)$ in each round. By querying the fixed point oracle, our LEA algorithm simply predicts its output p_t such that $p_t=\phi_t(p_t)$. The instantaneous regret of this algorithm with respect to any comparing action modification rule ϕ^* can be expressed as

$$\langle p_t-\phi^*(p_t), l_t\rangle = \langle \phi_t(p_t)-\phi^*(p_t), l_t\rangle = \langle p_t\otimes l_t, \phi_t-\phi^*\rangle\,,$$

where the last step interprets ϕ_t and ϕ^* as right stochastic matrices. Then, by using the notation $g_t := p_t\otimes l_t$ and taking the summation over $t\in[1:T]$, the ϕ-regret can be rewritten as

$$\mathrm{Regret}_T(\phi^*) = \sum_{t=1}^{T}\langle g_t, \phi_t-\phi^*\rangle\,. \tag{2}$$

Notice that the RHS is the external regret of our ϕ_t-predicting procedure in a hypothetical *Online Linear Optimization* (OLO) problem: in each round it predicts $\phi_t\in\mathcal{S}(d)$, observes the rank-1 loss gradient $g_t\in\mathbb{R}^{d\times d}$, and suffers the linear loss $\langle g_t,\phi_t\rangle$. The total loss is compared to that of ϕ^*.

Given this reduction, the convention of the field is to proceed with *Online Mirror Descent* (OMD). For example, by running a separate copy of MWU on each row of $\mathcal{S}(d)$ (which corresponds to OMD with a "row-separable" Bregman divergence), one could obtain an $O(d\sqrt{T\log d})$ swap regret bound via a straightforward summation. A better approach is using a *gradient-adaptive* OMD algorithm to exploit the structure of g_t, leading to the optimal $O(\sqrt{dT\log d})$ swap regret bound [BM07]. This could be seen as a computationally efficient realization of running MWU on all vertices of $\mathcal{S}(d)$, which also intuitively justifies the necessity of the $\sqrt{d}$ factor.

In this paper we deviate from this convention. Instead of using OMD, our algorithm (Algorithm 1) will be a modular composition of several new components introduced next.

2.1 Preprocessing

As the first step, we need to assign an ordering to the set of experts and augment their indices into a sequence of dyadic length (i.e., power of 2). This is due to our later use of the Haar wavelet: the size of the Haar matrix is naturally dyadic, and the ordering of its input affects our intermediate, variation-based regret bound (Theorem 3).

Concretely, we define $S := \lceil \log_2 d \rceil$ and $\bar{d} := 2^S$, which means that $\bar{d}$ is the smallest power of 2 that is greater or equal to d, the total number of experts. To associate the *augmented index set* $[1:\bar{d}]$ with the original index set $[1:d]$, our algorithm requires a *relabeling function* $\mathcal{I}$ as the user's input, mapping any $i \in [1:d]$ to a set $\mathcal{I}(i) \subset [1:\bar{d}]$. Three conditions should be satisfied:

- For all $i \in [1:d]$, $\mathcal{I}(i)$ only contains consecutive integers in $[1:\bar{d}]$.
- For all $i \in [1:d]$, the cardinality of $\mathcal{I}(i)$ satisfies $1 \leq |\mathcal{I}(i)| \leq 2$.
- The collection of sets $\{\mathcal{I}(i); i \in [1:d]\}$ is disjoint, and $\cup_{i=1}^{d}\mathcal{I}(i) = [1:\bar{d}]$.

It is clear that such an $\mathcal{I}$ exists.[2] Besides, the third condition enables the definition of the "inverse" function $\mathcal{I}^{-1} : [1:\bar{d}] \to [1:d]$, such that for all $\bar{i} \in [1:\bar{d}]$ we have $\bar{i} \in \mathcal{I}(\mathcal{I}^{-1}(\bar{i}))$.

Using this construction, we then consider the following *augmented* LEA problem with $\bar{d}$ experts. Let $\bar{l}_t \in \mathbb{R}^{\bar{d}}$ be the vector whose $\bar{i}$-th entry equals the $\mathcal{I}^{-1}(\bar{i})$-th entry of l_t, and any decision $\bar{p}_t \in \Delta(\bar{d})$ for the augmented LEA problem suffers the loss $\langle \bar{p}_t, \bar{l}_t \rangle$. The point is that such an augmented problem can be made equivalent to the original one:

- By letting the i-th coordinate of $p_t \in \Delta(d)$ be the total mass of $\bar{p}_t$ on $\mathcal{I}(i)$, we always have $\langle p_t, l_t \rangle = \langle \bar{p}_t, \bar{l}_t \rangle$.
- For any $\phi^* \in \mathcal{S}(d)$, there exists $\bar{\phi}^* \in \mathcal{S}(\bar{d})$ that preserves the complexity of ϕ^* and satisfies $\langle \phi^*(p_t), l_t \rangle = \langle \bar{\phi}^*(\bar{p}_t), \bar{l}_t \rangle$ (Lemma B.1 and B.2).

Therefore based on the reduction of [GGM08], it suffices to consider external regret minimization on the *augmented stochastic matrix space* $\mathcal{S}(\bar{d})$ instead.

To proceed, we use bar-equipped notations analogous to Eq.(2): let $\bar{\phi}_t$ be our matrix prediction on $\mathcal{S}(\bar{d})$, with $\bar{p}_t \in \Delta(\bar{d})$ being its fixed point. Then, with $\bar{g}_t := \bar{p}_t \otimes \bar{l}_t$, our goal next is to design a $\bar{\phi}_t$-predicting procedure which ensures that for all $\bar{\phi}^* \in \mathcal{S}(\bar{d})$, $\sum_{t=1}^{T} \langle \bar{g}_t, \bar{\phi}_t - \bar{\phi}^* \rangle$ is upper-bounded by an appropriate function of $\bar{\phi}^*$.

2.2 Wavelet-Inspired Matrix Features

Our second step requires pretending that *improper* matrix predictions are allowed, i.e., $\bar{\phi}_t$ can be anything on $\mathbb{R}^{\bar{d}\times\bar{d}}$ rather than just $\mathcal{S}(\bar{d})$. Although its fixed point does not exist in general (which is problematic from the perspective of [GGM08]), we will show afterwards that the constraint $\mathcal{S}(\bar{d})$ can be enforced through an additional projection step (Section 2.4).

This leads to the main idea of this work: given any collection of matrices $\mathcal{B}$ that spans $\mathbb{R}^{\bar{d}\times\bar{d}}$, all elements of $\mathbb{R}^{\bar{d}\times\bar{d}}$ can be expressed as a linear combination of $\mathcal{B}$, which then facilitates the use of sparsity to measure the complexity of each comparator instance $\bar{\phi}^*$. There are just two desiderata when choosing $\mathcal{B}$: first, it should properly associate different matrix entries such that the resulting external and internal regret is low; and second, it needs to be congruent with the structure of the gradient $\bar{g}_t$, such that the $\tilde{O}(\sqrt{dT})$ swap regret is preserved in the worst case. We will show that the classical idea of *Haar wavelet* [Mal08] addresses both.

Specialized to Euclidean spaces, wavelets refer to a systematic construction of orthogonal *multi-resolution* bases, with numerous applications across signal processing, nonparametric statistics and machine learning. Here

[2] A simple construction is to keep the ordering of the original d experts unchanged, and make a duplicate for each expert (placed next to it) until the total number of expert is $\bar{d}$.

multi-resolution means that if we examine the inner product of any "data" vector with all wavelet basis vectors (i.e., taking the *wavelet transform*), then some of these inner products (i.e., transform domain coefficients) capture the global characteristics of the data, while others capture the local characteristics. Such a representation is extremely useful for common modalities such as images: a typical image would consists of global scenes, local details and pixel-level noises, therefore separately extracting these information would enable applications like denoising and compression. Our algorithm exemplifies the same intuition in an online context: extracting multi-resolution characteristics of the gradient $\bar{g}_t$ enables better decision making.

Haar wavelet Concretely, we use the Haar wavelet basis $\mathcal{H}$ defined as follows. Recall that our augmented dimensionality satisfies $\bar{d} = 2^S$ for some positive integer S. Given any *scale* parameter $s \in [1 : S]$ and any *location* parameter $l \in [1 : 2^{S-s}]$, we define a Haar basis vector $h^{(s,l)} \in \mathbb{R}^{\bar{d}}$ whose i-th entry is

$$h_i^{(s,l)} := \begin{cases} 1, & i \in [2^s(l-1)+1 : 2^s(l-1)+2^{s-1}]; \\ -1, & i \in [2^s(l-1)+2^{s-1}+1 : 2^s l]; \\ 0, & \text{else}. \end{cases}$$

Notice that $h^{(s,l)}$ is only nonzero on its *support* $I^{(s,l)} := [2^s(l-1)+1 : 2^s l]$; on the first half of $I^{(s,l)}$ it equals 1, and on the second half of $I^{(s,l)}$ it equals -1. For each s there are 2^{S-s} valid choices of l, which means that in total there are $\bar{d}-1$ pairs of (s,l). Then, we additionally define $h^{(S,0)}$ as the all-one vector in $\mathbb{R}^{\bar{d}}$, which completes the size-$\bar{d}$ collection of vectors,

$$\mathcal{H} := \left\{h^{(S,0)}\right\} \cup \left\{h^{(s,l)}; s \in [1 : S], l \in [1 : 2^{S-s}]\right\}.$$

One could verify that $\mathcal{H}$ is indeed an orthogonal basis of $\mathbb{R}^{\bar{d}}$: the all-one vector $h^{(S,0)}$ is orthogonal to all $h^{(s,l)}$; vectors on the same scale have disjoint supports; and for two vectors on different scales, the larger-scale vector remains constant over the support of the smaller-scale vector. Consequently, any vector $v \in \mathbb{R}^{\bar{d}}$ can be uniquely represented by all the inner products $\langle v, h\rangle$, $\forall h \in \mathcal{H}$. The most important property for our use is that $\langle v, h\rangle$ captures the *variability* of v over the support of h: if v remains constant over such an interval, then $\langle v, h\rangle = 0$. In other words, the Haar wavelet basis can sparsely represent low-variation signals.

Matrix features Starting from $\mathcal{H}$, we construct the following collection $\mathcal{B}$ of *matrix features*, and this will be used in an online linear regression subroutine introduced shortly. Let $\mathcal{E} = \{e^{(i)}; i \in [1 : \bar{d}]\}$ be the collection of $\bar{d}$-dimensional unit coordinate vectors. We define

$$\mathcal{B} := \{\boldsymbol{I}_{\bar{d}}\} \cup \{h \otimes e; h \in \mathcal{H}, e \in \mathcal{E}\},$$

where $\boldsymbol{I}_{\bar{d}}$ denotes the $\bar{d}$-by-$\bar{d}$ identity matrix. Without $\boldsymbol{I}_{\bar{d}}$, the rest of $\mathcal{B}$ is an orthogonal basis of the matrix space $\mathbb{R}^{\bar{d}\times\bar{d}}$, which can sparsely represent any comparator $\bar{\phi}^* \in \mathcal{S}(\bar{d})$ with large $d^{\mathrm{unif}}_{\bar{\phi}^*}$. The role of $\boldsymbol{I}_{\bar{d}}$ is to help in cases with large $d^{\mathrm{self}}_{\bar{\phi}^*}$.

To summarize, we will work with this slightly *overcomplete* collection of matrix features. It means representations are not unique, but this is fine – as long as there *exists* a sparse representation for the comparator $\bar{\phi}^*$, i.e., $\bar{\phi}^* = \sum_{b\in\mathcal{B}} \Phi^{*,(b)} b$ for some sparse coefficients $\{\Phi^{*,(b)}; b \in \mathcal{B}\}$, we can use the known algorithms introduced next to adapt to it.

2.3 First-Order Sparsity-Adaptive Oracle

Continuing from our improper matrix prediction problem on $\mathbb{R}^{\bar{d}\times\bar{d}}$ (the beginning of Section 2.2; predictions and loss gradients are denoted by $\bar{\phi}^{\mathrm{improper}}_t$ and $\bar{g}^{\mathrm{improper}}_t$), our third step is to solve it by learning $\{\Phi^{*,(b)}; b \in \mathcal{B}\}$, the representation of the comparator $\bar{\phi}^*$ on $\mathcal{B}$. Since $\bar{\phi}^*$ can be arbitrary, the "learning" here more precisely means suffering low regret in the following equivalent problem "transformed" by $\mathcal{B}$. In each round we first predict a coefficient $\Phi^{(b)}_t \in \mathbb{R}$ for all matrix features $b \in \mathcal{B}$, which results in an improper matrix prediction

$$\bar{\phi}^{\mathrm{improper}}_t := \sum_{b\in\mathcal{B}} \Phi^{(b)}_t b \in \mathbb{R}^{\bar{d}\times\bar{d}}. \tag{3}$$

After observing the corresponding loss gradient $\bar{g}_t^{\text{improper}} \in \mathbb{R}^{\bar{d}\times\bar{d}}$, we suffer the linear loss $\left\langle \bar{g}_t^{\text{improper}}, \bar{\phi}_t^{\text{improper}} \right\rangle = \sum_{b\in\mathcal{B}} \left\langle \bar{g}_t^{\text{improper}}, b \right\rangle \Phi_t^{(b)}$. The previous matrix-based regret definition $\sum_{t=1}^T \left\langle \bar{g}_t^{\text{improper}}, \bar{\phi}_t^{\text{improper}} - \bar{\phi}^* \right\rangle$ is equivalent to the sum of a one-dimensional regret,

$$\sum_{b\in\mathcal{B}} \sum_{t=1}^T \left\langle \bar{g}_t^{\text{improper}}, b \right\rangle (\Phi_t^{(b)} - \Phi^{*,(b)}).$$

In other words, the problem of improper matrix prediction is converted to *online linear regression with linear losses*.

Suppose we simply run gradient descent on each coefficient $\Phi_t^{(b)}$ separately. With the optimal constant learning rate, it guarantees the well-known "prototypical" bound $\sum_{t=1}^T \left\langle \bar{g}_t^{\text{improper}}, b \right\rangle (\Phi_t^{(b)} - \Phi^{*,(b)}) = O\left(\left| \Phi^{*,(b)} \right| \sqrt{\sum_{t=1}^T \left\langle \bar{g}_t^{\text{improper}}, b \right\rangle^2} \right)$ (see, e.g., [Ora23, Theorem 2.13]), and therefore

$$\sum_{t=1}^T \left\langle \bar{g}_t^{\text{improper}}, \bar{\phi}_t^{\text{improper}} - \bar{\phi}^* \right\rangle = O\left(\sum_{b\in\mathcal{B}} \left| \Phi^{*,(b)} \right| \sqrt{\sum_{t=1}^T \left\langle \bar{g}_t^{\text{improper}}, b \right\rangle^2} \right). \tag{4}$$

This is essentially what we need as the RHS depends on the number of nonzero elements within $\{\Phi^{*,(b)}; b \in \mathcal{B}\}$. The only issue (but critical) is the lack of *adaptivity*: the required constant learning rate on each $\Phi_t^{(b)}$ would depend on $\Phi^{*,(b)}$ (more specifically, its absolute value), but we need a single algorithm that simultaneously guarantees the desirable one-dimensional regret bound for all possible values of $\Phi^{*,(b)}$. No "oracle tuning" is allowed.

We address this issue using *comparator-adaptive* OLO algorithms [SM12, MO14, OP16], for which an excellent tutorial is given by [Ora23, Chapter 9]. The idea is that by replacing the one-dimensional gradient descent by suitable instances of *Follow the Regularized Leader* (FTRL), we can bypass the choice of the learning rate, thus concretely achieving a suitable weaker form of the above regret bounds. Specifically, the one-dimensional algorithm we use comes from the following lemma combining [ZCP22, Theorem 4] and [Cut18, Theorem 5.8].

Lemma 2.1 ([Cut18, ZCP22])**.** *Consider the one-dimensional OLO problem where in each round an algorithm makes a decision $x_t \in \mathbb{R}$ and then observes the loss gradient $c_t \in [-G, G]$ for some known Lipschitz constant G. Given any hyperparameter $\varepsilon > 0$, there exists an algorithm that guarantees*

$$\sum_{t=1}^T c_t(x_t - u) \le \sqrt{G^2 + G\sum_{t=1}^T |c_t|} \left[2\varepsilon + 2\sqrt{2}\,|u| \sqrt{\log\left(1 + \frac{|u|}{\sqrt{2}\varepsilon}\right)} + 4\sqrt{2}\,|u| \right],$$

for all time horizon $T \in \mathbb{N}_+$, all comparator $u \in \mathbb{R}$ and all possible $c_{1:T}$ sequence.

We remark that [ZCP22, Theorem 4] is the state-of-the-art result for comparator-adaptive OLO *without* gradient adaptivity, while [Cut18, Theorem 5.8] enhances it to *first-order gradient adaptivity* in a black-box manner (i.e., G^2T in the regret bound is improved to $G^2 + G\sum_{t=1}^T |c_t|$, following the notations of Lemma 2.1). As shown by [BM07], first-order gradient adaptivity is crucial for achieving the optimal $\tilde{O}(\sqrt{dT})$ swap regret bound. Our analysis additionally shows that such gradient adaptivity works harmoniously with the Haar wavelet, such that the $\tilde{O}(\sqrt{dT})$ swap regret bound can still be achieved with suitable matrix features.

2.4 Enforcing the Constraint

Our final step is to close the only remaining gap in the above reasoning: $\bar{\phi}_t^{\text{improper}}$ does not belong to $\mathcal{S}(\bar{d})$. Addressing this issue requires a suitable wrapper: given the algorithm from the previous step equipped with an upper bound on $\sum_{t=1}^T \left\langle \bar{g}_t^{\text{improper}}, \bar{\phi}_t^{\text{improper}} - \bar{\phi}^* \right\rangle$, we design mappings $\bar{\phi}_t^{\text{improper}} \to \bar{\phi}_t \in \mathcal{S}(\bar{d})$ called a *projection oracle* and $\bar{g}_t \to \bar{g}_t^{\text{improper}}$ called a *gradient processing oracle*, such that (Lemma B.5)

$$\sum_{t=1}^T \left\langle \bar{g}_t, \bar{\phi}_t - \bar{\phi}^* \right\rangle \le \sum_{t=1}^T \left\langle \bar{g}_t^{\text{improper}}, \bar{\phi}_t^{\text{improper}} - \bar{\phi}^* \right\rangle.$$

The LHS is equivalent to $\mathrm{Regret}_T(\phi^*)$, while the RHS adapts to the complexity of ϕ^* (measured with respect to $\mathcal{B}$). Furthermore, the processed gradient $\bar{g}_t^{\mathrm{improper}}$ should be low enough in a suitable notion of magnitude (Lemma B.6).

To this end, we employ the following two-stage procedure that extends a classical OLO-to-LEA wrapper[3] [LS15, OP16] to the matrix setting.

$$\bar{\phi}_t^{\mathrm{improper}} \in \mathbb{R}^{\bar{d}\times\bar{d}} \xrightarrow{\text{Stage 1}} \bar{\phi}_t^{+} \in \mathbb{R}_+^{\bar{d}\times\bar{d}} \xrightarrow{\text{Stage 2}} \bar{\phi}_t \in \mathcal{S}(\bar{d}) \tag{Projection}$$

$$\Downarrow$$

$$\bar{g}_t^{\mathrm{improper}} \xleftarrow{\text{Stage 1}} \bar{g}_t^{+} \xleftarrow{\text{Stage 2}} \bar{g}_t \tag{Gradient processing}$$

The first stage enforces the positivity constraint by mapping $\bar{\phi}_t^{\mathrm{improper}}$ to an intermediate prediction $\bar{\phi}_t^{+} \in \mathbb{R}_+^{\bar{d}\times\bar{d}}$, while the second stage enforces the stochastic matrix constraint by mapping $\bar{\phi}_t^{+}$ to $\bar{\phi}_t$. Accordingly, the gradient processing is also performed in two stages, $\bar{g}_t \to \bar{g}_t^{+} \to \bar{g}_t^{\mathrm{improper}}$. The details are deferred to Appendix A.1 due to space constraints.

With that we have presented all the technical ingredients of our algorithm, as well as the high-level analytical strategy. The next step is to combine all the pieces into our main results.

3 Main Result

Our main results are Algorithm 1 and its ϕ-regret bound (Theorem 1). The pseudocode is deferred to Appendix A, while Figure 1 illustrates its main idea. We note that since $|\mathcal{B}| = \Theta(d^2)$, both the runtime and the memory of our algorithm are of the same order as [BM07]. Although it requires the relabeling function $\mathcal{I}$ as the user's input (Section 2.1), all results in this section hold uniformly for all $\mathcal{I}$ (i.e., the multiplying constants c in the theorems do not depend on $\mathcal{I}$). The role of good $\mathcal{I}$ is discussed in Remark 3.1.

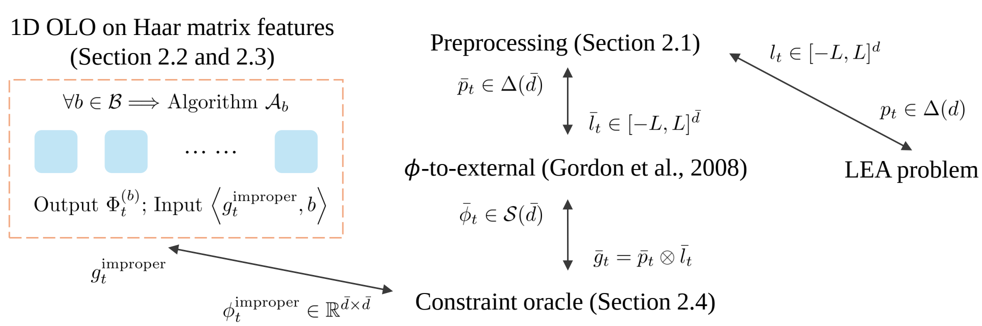

Figure 1: An illustration of Algorithm 1; see Appendix A for details.

Theorem 1 (Main result)**.** *There is an absolute constant $c > 0$ such that for all $T \in \mathbb{N}_+$ and all comparing action modification rule $\phi^* \in \mathcal{S}(d)$, Algorithm 1 guarantees*

$$\mathrm{Regret}_T(\phi^*) \leq c \cdot \left(L\sqrt{\min\left\{d - d_{\phi^*}^{\mathrm{unif}} + 1, d - d_{\phi^*}^{\mathrm{self}}\right\}(T+d)} \cdot (\log d)^{3/2} \right).$$

Assuming $T \gg d$, Theorem 1 recovers the $\tilde{O}(\sqrt{T})$ external regret in the case of $d_{\phi^*}^{\mathrm{unif}} = d$, the $\tilde{O}(\sqrt{T})$ internal regret in the case of $d_{\phi^*}^{\mathrm{self}} = d - 1$ and the $\tilde{O}(\sqrt{dT})$ swap regret in the worst case, matching the results achieved by specialized algorithms (modulo polylog factors). Although this basic requirement can also be

[3][LS15, OP16] converted the problem of LEA (i.e., OLO on the domain $\Delta(d)$) to "unconstrained" OLO on the domain $\mathbb{R}^d$. A more general treatment is given by [CO18].

achieved by simply aggregating specialized algorithms using MWU,[4] concrete improvements of our approach are manifested in the intermediate regimes. For example, suppose $d^{\mathrm{unif}}_{\phi^*} = d - k + 1$, then Algorithm 1 guarantees $\mathrm{Regret}_T(\phi^*) = \tilde{O}(\sqrt{kT})$ while being computationally efficient and agnostic to k. This improves both

- generic swap regret minimization algorithms achieving $\tilde{O}(\sqrt{dT})$ [BM07]; and
- the computationally inefficient, k-dependent baseline which runs MWU over an appropriate ϕ-class (specifically, all zero-one stochastic matrices ϕ satisfying $d^{\mathrm{unif}}_{\phi} = d - k + 1$). MWU can guarantee the $\tilde{O}(\sqrt{kT})$ ϕ-regret despite the cardinality of the hypothesis class being exponential in k, but computationally this is impractical when k is large.

Similar cases of improvement can be constructed for large $d^{\mathrm{self}}_{\phi^*}$ as well, although it requires a more subtle comparison to related works [Rot23, NRRX23]. This is deferred to Appendix D.

Proof sketch of Theorem 1 For simplicity, suppose Eq.(4) can indeed be achieved (the rigorous analysis using Lemma 2.1 is similar). Combining the components from Section 2, the LHS of Eq.(4) upper-bounds $\mathrm{Regret}_T(\phi^*)$. The RHS of Eq.(4) is a summation over $b \in \mathcal{B}$, and excluding the edge cases we may only consider those b expressed as $h^{(s,l)} \otimes e^{(j)}$, where $s \in [1:S]$ is the scale parameter, $l \in [1:2^{S-s}]$ is the location parameter, and $j \in [1:\bar{d}]$ is the column index. The remaining task becomes upper-bounding

$$\sum_{s=1}^{S} \sum_{l=1}^{2^{S-s}} \sum_{j=1}^{\bar{d}} \left|\Phi^{*,(s,l,j)}\right| \sqrt{\sum_{t=1}^{T} \left\langle \bar{g}_t^{\mathrm{improper}}, h^{(s,l)} \otimes e^{(j)} \right\rangle^2}. \tag{5}$$

More precisely, the "transform domain coefficient" $\Phi^{*,(s,l,j)} \in \mathbb{R}$ here is defined such that $\Phi^{*,(s,l,j)} \cdot h^{(s,l)} \otimes e^{(j)}$ is the projection of either $\bar{\phi}^*$ or $\bar{\phi}^* - \boldsymbol{I}_{\bar{d}}$ to $h^{(s,l)} \otimes e^{(j)}$. Recall that such non-uniqueness is due to $\mathcal{B}$ being overcomplete.

To proceed, there are two crucial steps.

- Lemma B.6 shows that $\left|\left\langle \bar{g}_t^{\mathrm{improper}}, h^{(s,l)} \otimes e^{(j)} \right\rangle\right| \leq 2L \sum_{i \in I^{(s,l)}} \bar{p}_{t,i}$, where $I^{(s,l)} \subset [1:\bar{d}]$ is the support of $h^{(s,l)}$. Since on each row the sum of $\bar{\phi}^*$ entries is 1, we also have $\sum_{j=1}^{\bar{d}} \left|\Phi^{*,(s,l,j)}\right| \leq \mathrm{const}$ (Lemma B.3). It means even very crudely,

$$\begin{aligned}\text{Eq.(5)} &\lesssim \sum_{s=1}^{S} \sum_{l=1}^{2^{S-s}} \sum_{j=1}^{\bar{d}} \left|\Phi^{*,(s,l,j)}\right| \sqrt{\sum_{t=1}^{T} \sum_{i \in I^{(s,l)}} \bar{p}_{t,i}} \lesssim \sum_{s=1}^{S} \sum_{l=1}^{2^{S-s}} \sqrt{\sum_{t=1}^{T} \sum_{i \in I^{(s,l)}} \bar{p}_{t,i}} \\ &\overset{\mathrm{Cauchy-Schwarz}}{\lesssim} \sum_{s=1}^{S} \sqrt{2^{S-s}} \sqrt{\sum_{t=1}^{T} \sum_{l=1}^{2^{S-s}} \sum_{i \in I^{(s,l)}} \bar{p}_{t,i}} \overset{\bar{p}_t \in \Delta(\bar{d})}{\lesssim} \sum_{s=1}^{S} \sqrt{2^{S-s} T} \overset{S \lesssim \log d}{\lesssim} \sqrt{dT} \log d,\end{aligned}$$

  which recovers the nonadaptive, $\tilde{O}(\sqrt{dT})$ swap regret bound.

- The adaptive bound requires a sparsity-based refinement. Let us reconsider $\Phi^{*,(s,l,j)}$ projected from $\bar{\phi}^*$: if for all $i \in I^{(s,l)}$ the i-th row of $\bar{\phi}^*$ equals the same vector, then due to the property of the Haar wavelet, $\sum_{j=1}^{\bar{d}} \left|\Phi^{*,(s,l,j)}\right| = 0$. Now the question is, for any fixed s, how many different l make this sum nonzero? By definition, the answer is at most $\mathrm{RowSwitch}(\bar{\phi}^*) := \sum_{i=1}^{\bar{d}-1} \mathbf{1}[\bar{\phi}^*_i \neq \bar{\phi}^*_{i+1}]$, which means the same Cauchy-Schwarz argument gives us the $\tilde{O}\left(\sqrt{\mathrm{RowSwitch}(\bar{\phi}^*) \cdot T}\right)$ bound rather than $\tilde{O}(\sqrt{dT})$. The last step is to construct the augmented comparator $\bar{\phi}^*$ in a complexity-preserving manner, i.e., $\mathrm{RowSwitch}(\bar{\phi}^*) \lesssim d - d^{\mathrm{unif}}_{\phi^*}$ (Lemma B.1).

  The proof of the $d^{\mathrm{self}}_{\phi^*}$-dependent bound is similar. The key observation is that $\mathrm{RowSwitch}(\bar{\phi}^* - \boldsymbol{I}_{\bar{d}})$ is small if $d^{\mathrm{self}}_{\bar{\phi}^*}$ is large. Since everything in this proof concerns the *same algorithm* (just with different constructions of $\bar{\phi}^*$ and $\Phi^{*,(s,l,j)}$), the $d^{\mathrm{unif}}_{\phi^*}$ and $d^{\mathrm{self}}_{\phi^*}$-dependent bounds are achieved simultaneously. □

[4] One could simultaneously maintain three algorithms targeting the external, internal and swap regret respectively, treat them as "meta"-experts, and aggregate them using a three-expert MWU algorithm on top. The weakness is that this baseline does not achieve a better-than-$\tilde{O}(\sqrt{dT})$ ϕ-regret bound with respect to a generic ϕ.

Remark 3.1 (Role of relabeling)**.** *Compared to typical LEA algorithms, Algorithm 1 has a notable difference: the algorithm generates different outputs under a permutation of the expert indices. Theorem 1 proves that regardless of the permutation its ϕ-regret bound improves [BM07], but as shown in the proof sketch, this is actually a* relaxation *of an order-dependent result (Theorem 3). In this regard, the relabeling function $\mathcal{I}$ could be seen as the user's prior: if it renders* $\mathrm{RowSwitch}(\bar{\phi}^*)$ *small, then the order-dependent bound would further improve Theorem 1.*

As an example, consider $d = \bar{d}$, with the first $d/2$ rows of ϕ^* being some vector v while the rest of the rows being $w \neq v$. Such a ϕ^* itself possesses certain simplicity, and we suppose the user happens to preserve this simplicity by "trivially" picking $\mathcal{I}(i) = i; \forall i \in [1:d]$, which ensures $\bar{\phi}^* = \phi^*$ by the construction of Lemma B.1. In this case, although Theorem 1 guarantees the same $\mathrm{Regret}_T(\phi^*) = \tilde{O}(\sqrt{dT})$ *as [BM07], Theorem 3 sharpens it to $\tilde{O}\left(\sqrt{\mathrm{RowSwitch}(\bar{\phi}^*) \cdot T}\right) = \tilde{O}(\sqrt{T})$. We expect that reaping such benefits in "ordered" downstream problems could be an intriguing direction for future works.*

Finally, we can characterize the quantile regret of Algorithm 1, which is formally defined as follows. Imagine that at the end of the game the cumulative loss of all experts $(\sum_{t=1}^{T} l_{t,i}; \forall i \in [1:d])$ are sorted from the lowest to the highest, with ties broken arbitrarily. Then, for any $\varepsilon \in [d^{-1}, 1]$, let i_ε be the index of the $\lceil \varepsilon d \rceil$-th element in this sorted list. The ε-quantile regret is defined as

$$\mathrm{Regret}_T(\varepsilon) := \sum_{t=1}^{T} \langle p_t, l_t \rangle - \sum_{t=1}^{T} l_{t,i_\varepsilon}.$$

Theorem 2. *There exists an absolute constant $c > 0$ such that for all $\varepsilon \in [d^{-1} : 1]$, Algorithm 1 guarantees the quantile regret bound*

$$\mathrm{Regret}_T(\varepsilon) \leq c \cdot \left(L\sqrt{T \log \varepsilon^{-1}}\right).$$

This is optimal [NBC+21], and importantly, it sharpens the $\tilde{O}(\sqrt{T})$ external regret bound from Theorem 1 to the optimal rate $O(\sqrt{T \log d})$, including the log factor. The proof hinges on the observation that the quantile regret corresponds to a special case of the ϕ-regret where the rows of the comparing ϕ matrix are all equal. It means we may only analyze a subset of $\mathcal{B}$ generated by the all-one feature $h^{(S,0)} \in \mathcal{H}$, and the rest of the proof is fairly standard [Ora23, Chapter 9.6].

4 Conclusion

Focusing on the LEA problem with $T \gg d$, this paper leverages the ideas of features and sparsity to refine the classical reduction approaches in ϕ-regret and swap regret minimization [BM07, GGM08]. We present an LEA algorithm with an adaptive ϕ-regret upper bound, matching the external, internal and swap regret of specialized algorithms while improving them in the intermediate regimes. Conceptually, our key observation is that incorporating certain Haar-wavelet-inspired matrix features can introduce inductive biases that are well-aligned with important special cases of the ϕ-regret. Technically, we show that the Haar wavelet is naturally congruent with the benign gradient structure from the ϕ-to-external reduction, such that the optimal $\tilde{O}(\sqrt{dT})$ swap regret bound can still be achieved after incorporating such matrix features. We leave a thorough study of the downstream benefits to future works.

References

[BCM+23] Martino Bernasconi, Matteo Castiglioni, Alberto Marchesi, Francesco Trovo, and Nicola Gatti. Constrained phi-equilibria. In *International Conference on Machine Learning*, pages 2184–2205. PMLR, 2023.

[Bla56] David Blackwell. An analog of the minimax theorem for vector payoffs. *Pacific Journal of Mathematics*, 6(1):1–8, 1956.

[BM07] Avrim Blum and Yishay Mansour. From external to internal regret. *Journal of Machine Learning Research*, 8(6), 2007.

[BW19] Dheeraj Baby and Yu-Xiang Wang. Online forecasting of total-variation-bounded sequences. *Advances in Neural Information Processing Systems*, 32, 2019.

[CBFH+97] Nicolo Cesa-Bianchi, Yoav Freund, David Haussler, David P Helmbold, Robert E Schapire, and Manfred K Warmuth. How to use expert advice. *Journal of the ACM*, 44(3):427–485, 1997.

[CBL06] Nicolo Cesa-Bianchi and Gábor Lugosi. *Prediction, learning, and games.* Cambridge university press, 2006.

[CDL+24] Yang Cai, Constantinos Daskalakis, Haipeng Luo, Chen-Yu Wei, and Weiqiang Zheng. On tractable Φ-equilibria in non-concave games. *Advances in Neural Information Processing Systems*, 37, 2024.

[CFH09] Kamalika Chaudhuri, Yoav Freund, and Daniel J Hsu. A parameter-free hedging algorithm. *Advances in Neural Information Processing Systems*, 22, 2009.

[CLW21] Liyu Chen, Haipeng Luo, and Chen-Yu Wei. Impossible tuning made possible: A new expert algorithm and its applications. In *Conference on Learning Theory*, pages 1216–1259. PMLR, 2021.

[CO18] Ashok Cutkosky and Francesco Orabona. Black-box reductions for parameter-free online learning in banach spaces. In *Conference On Learning Theory*, pages 1493–1529. PMLR, 2018.

[Cut18] Ashok Cutkosky. *Algorithms and Lower Bounds for Parameter-free Online Learning*. Stanford University, 2018.

[DDFG24] Yuval Dagan, Constantinos Daskalakis, Maxwell Fishelson, and Noah Golowich. From external to swap regret 2.0: An efficient reduction for large action spaces. In *Proceedings of the 56th Annual ACM Symposium on Theory of Computing*, pages 1216–1222, 2024.

[FV99] Dean P Foster and Rakesh Vohra. Regret in the on-line decision problem. *Games and Economic Behavior*, 29(1-2):7–35, 1999.

[GGM08] Geoffrey J Gordon, Amy Greenwald, and Casey Marks. No-regret learning in convex games. In *International Conference on Machine Learning*, pages 360–367, 2008.

[GJ03] Amy Greenwald and Amir Jafari. A general class of no-regret learning algorithms and game-theoretic equilibria. In *Learning Theory and Kernel Machines: 16th Annual Conference on Learning Theory and 7th Kernel Workshop*, pages 2–12. Springer, 2003.

[HPY23] Nika Haghtalab, Chara Podimata, and Kunhe Yang. Calibrated stackelberg games: Learning optimal commitments against calibrated agents. *Advances in Neural Information Processing Systems*, 36, 2023.

[Ito20] Shinji Ito. A tight lower bound and efficient reduction for swap regret. *Advances in Neural Information Processing Systems*, 33:18550–18559, 2020.

[JO24] Andrew Jacobsen and Francesco Orabona. An equivalence between static and dynamic regret minimization. *Advances in Neural Information Processing Systems*, 37, 2024.

[Leh03] Ehud Lehrer. A wide range no-regret theorem. *Games and Economic Behavior*, 42(1):101–115, 2003.

[LNPR22] Daniel Lee, Georgy Noarov, Mallesh Pai, and Aaron Roth. Online minimax multiobjective optimization: Multicalibeating and other applications. *Advances in Neural Information Processing Systems*, 35:29051–29063, 2022.

[LS15] Haipeng Luo and Robert E Schapire. Achieving all with no parameters: AdaNormalHedge. In *Conference on Learning Theory*, pages 1286–1304. PMLR, 2015.

[LW94] Nick Littlestone and Manfred K Warmuth. The weighted majority algorithm. *Information and computation*, 108(2):212–261, 1994.

[Mal08] Stephane Mallat. *A Wavelet Tour of Signal Processing: The Sparse Way.* Academic Press, 2008.

[MO14] H Brendan McMahan and Francesco Orabona. Unconstrained online linear learning in Hilbert spaces: Minimax algorithms and normal approximations. In *Conference on Learning Theory*, pages 1020–1039. PMLR, 2014.

[NBC+21] Jeffrey Negrea, Blair Bilodeau, Nicolò Campolongo, Francesco Orabona, and Dan Roy. Minimax optimal quantile and semi-adversarial regret via root-logarithmic regularizers. *Advances in Neural Information Processing Systems*, 34:26237–26249, 2021.

[NRRX23] Georgy Noarov, Ramya Ramalingam, Aaron Roth, and Stephan Xie. High-dimensional prediction for sequential decision making. *arXiv preprint arXiv:2310.17651*, 2023.

[OP16] Francesco Orabona and Dávid Pál. Coin betting and parameter-free online learning. *Advances in Neural Information Processing Systems*, 29, 2016.

[Ora23] Francesco Orabona. A modern introduction to online learning. *arXiv preprint arXiv:1912.13213*, 2023.

[Per15] Vianney Perchet. Exponential weight approachability, applications to calibration and regret minimization. *Dynamic Games and Applications*, 5:136–153, 2015.

[PR24] Binghui Peng and Aviad Rubinstein. Fast swap regret minimization and applications to approximate correlated equilibria. In *Proceedings of the 56th Annual ACM Symposium on Theory of Computing*, pages 1223–1234, 2024.

[PRO+22] Georgios Piliouras, Mark Rowland, Shayegan Omidshafiei, Romuald Elie, Daniel Hennes, Jerome Connor, and Karl Tuyls. Evolutionary dynamics and phi-regret minimization in games. *Journal of Artificial Intelligence Research*, 74:1125–1158, 2022.

[Rot23] Aaron Roth. Learning in games (and games in learning). *Unpublished Lecture Note, https://www.cis.upenn.edu/~aaroth/GamesInLearning.pdf*, 2023.

[RST11] Alexander Rakhlin, Karthik Sridharan, and Ambuj Tewari. Online learning: Beyond regret. In *Conference on Learning Theory*, pages 559–594, 2011.

[Shi16] Nahum Shimkin. An online convex optimization approach to blackwell's approachability. *Journal of Machine Learning Research*, 17(129):1–23, 2016.

[SL05] Gilles Stoltz and Gábor Lugosi. Internal regret in on-line portfolio selection. *Machine Learning*, 59:125–159, 2005.

[SL07] Gilles Stoltz and Gábor Lugosi. Learning correlated equilibria in games with compact sets of strategies. *Games and Economic Behavior*, 59(1):187–208, 2007.

[SM12] Matthew Streeter and Brendan Mcmahan. No-regret algorithms for unconstrained online convex optimization. *Advances in Neural Information Processing Systems*, 25, 2012.

[ZAFS24] Brian Hu Zhang, Ioannis Anagnostides, Gabriele Farina, and Tuomas Sandholm. Efficient Φ-regret minimization with low-degree swap deviations in extensive-form games. *Advances in Neural Information Processing Systems*, 37, 2024.

[ZCP22] Zhiyu Zhang, Ashok Cutkosky, and Ioannis Ch. Paschalidis. PDE-based optimal strategy for unconstrained online learning. In *International Conference on Machine Learning*, pages 26085–26115. PMLR, 2022.

[ZCP23] Zhiyu Zhang, Ashok Cutkosky, and Ioannis Ch. Paschalidis. Unconstrained dynamic regret via sparse coding. *Advances in Neural Information Processing Systems*, 36, 2023.

Appendix

Appendix A presents the pseudocode of Algorithm 1 alongside the algorithmic details omitted in the main paper. To analyze this algorithm, Appendix B presents several lemmas on its technical ingredients introduced in Section 2. These are then used in Appendix C to prove our main results. Appendix D contains detailed comparison to existing results.

A Details of the Algorithm

A.1 Constraint Oracle

Our approach to enforce the constraint $\mathcal{S}(d)$ (Section 2.4) is described below.

Stage 1: $\mathbb{R}^{\bar{d}\times\bar{d}} \to \mathbb{R}_+^{\bar{d}\times\bar{d}}$. Starting from $\bar{\phi}_t^{\text{improper}}$, we define the intermediate prediction $\bar{\phi}_t^+$ whose (i,j)-th entry $\bar{\phi}_{t,i,j}^+$ equals $\max\left\{\bar{\phi}_{t,i,j}^{\text{improper}}, 0\right\}$, the maximum of the (i,j)-th entry of $\bar{\phi}_t^{\text{improper}}$ and 0. This is performed for all $i, j \in [1:\bar{d}]$, resulting in a positive matrix $\bar{\phi}_t^+ \in \mathbb{R}_+^{\bar{d}\times\bar{d}}$.

Regarding gradient processing, given the intermediate gradient $\bar{g}_t^+ \in \mathbb{R}^{\bar{d}\times\bar{d}}$, we define the (i,j)-th entry of $\bar{g}_t^{\text{improper}}$ as

$$\bar{g}_{t,i,j}^{\text{improper}} := \begin{cases} \bar{g}_{t,i,j}^+, & \bar{\phi}_{t,i,j}^+ = \bar{\phi}_{t,i,j}^{\text{improper}}, \\ \min\left\{\bar{g}_{t,i,j}^+, 0\right\}, & \text{else.} \end{cases}$$

The intuition is the following. On each entry,

- If the "unprojected prediction" $\bar{\phi}_{t,i,j}^{\text{improper}}$ is already positive, then no projection is needed since $\bar{\phi}_{t,i,j}^+ = \max\left\{\bar{\phi}_{t,i,j}^{\text{improper}}, 0\right\} = \bar{\phi}_{t,i,j}^{\text{improper}}$. Naturally the gradient is kept unchanged, $\bar{g}_{t,i,j}^{\text{improper}} = \bar{g}_{t,i,j}^+$.
- If $\bar{\phi}_{t,i,j}^{\text{improper}}$ is negative, then it is projected to $\bar{\phi}_{t,i,j}^+ = 0$, which means "positive predictions are favored". To encourage that we only send negative $\bar{g}_{t,i,j}^{\text{improper}}$ to the wrapped algorithm that generates $\bar{\phi}_{t,i,j}^{\text{improper}}$.

Stage 2: $\mathbb{R}_+^{\bar{d}\times\bar{d}} \to \mathcal{S}(\bar{d})$. Given $\bar{\phi}_t^+ \in \mathbb{R}_+^{\bar{d}\times\bar{d}}$ from Stage 1, we scale its i-th row $\bar{\phi}_{t,i}^+$ by $1/\left\|\bar{\phi}_{t,i}^+\right\|_1$, for all i. If $\left\|\bar{\phi}_{t,i}^+\right\|_1 = 0$ then the scaled vector is defined as $[1/d, \ldots, 1/d]$. The matrix after the scaling is defined as $\bar{\phi}_t \in \mathcal{S}(\bar{d})$.

Regarding gradient processing, given $\bar{g}_t \in \mathbb{R}^{\bar{d}\times\bar{d}}$, we define the intermediate gradient $\bar{g}_t^+ \in \mathbb{R}^{\bar{d}\times\bar{d}}$ whose i-th row is

$$\bar{g}_{t,i}^+ := \bar{g}_{t,i} - \left\langle \bar{g}_{t,i}, \bar{\phi}_{t,i} \right\rangle.$$

Here $\bar{\phi}_{t,i} \in \Delta(\bar{d})$ denotes the i-th row of the proper matrix prediction $\bar{\phi}_t$. The intuition is that $\bar{g}_{t,i}^+$ represents the "centered" version of $\bar{g}_{t,i}$ with respect to $\bar{\phi}_t$. The obtained intermediate gradient $\bar{g}_t^+$ is then passed to Stage 1.

A.2 Pseudocode

The pseudocode of our algorithm is presented as Algorithm 1. It only requires the relabeling function $\mathcal{I}$ as the user's input.

B Details of the Technical Ingredients

The subsections below provide the detailed analysis of the algorithmic components from Section 2, following the order there.

Algorithm 1 Wavelet-based LEA with adaptive ϕ-regret guarantee.

Require: A user-specified relabeling function $\mathcal{I}$ (Section 2.1). Besides, recall the several components introduced previously:

- A fixed point oracle (the beginning of Section 2).
- The collection $\mathcal{B}$ of matrix features (Section 2.2).
- The one-dimensional OLO algorithm from Lemma 2.1, denoted by $\mathcal{A}_{1\mathrm{d}}$. It requires a Lipschitz constant G and a positive constant ε as hyperparameters.
- The projection oracle and the gradient processing oracle (Section 2.4 and Appendix A.1).

For all $b \in \mathcal{B}$, initiate a copy of $\mathcal{A}_{1\mathrm{d}}$, denoted by $\mathcal{A}_{1\mathrm{d}}^{(b)}$. We set $G = 2L$ for all $\mathcal{A}_{1\mathrm{d}}^{(b)}$. The choice of ε depends on b, which we denote as ε_b:

- If $b = \boldsymbol{I}_{\bar{d}}$, then $\varepsilon_b = 1$.
- For other $b \in \mathcal{B}$, we can write $b = h \otimes e^{(j)}$ for some $h \in \mathcal{H}$ and $e^{(j)} \in \mathcal{E}$. If $h = h^{(S,0)}$ then $\varepsilon_b = \bar{d}(d^2 \left|\mathcal{I}(\mathcal{I}^{-1}(j))\right|)^{-1}$; otherwise $\varepsilon_b = (d^2 \left|\mathcal{I}(\mathcal{I}^{-1}(j))\right|)^{-1}$.

for $t = 1, 2, \ldots,$ **do**

  For all $b \in \mathcal{B}$, query the t-th prediction $\Phi_t^{(b)} \in \mathbb{R}$ of $\mathcal{A}_{1d}^{(b)}$.

  Following Eq.(3), let

$$\bar{\phi}_t^{\text{improper}} = \sum_{b \in \mathcal{B}} \Phi_t^{(b)} b \in \mathbb{R}^{\bar{d} \times \bar{d}}.$$

  Using the projection oracle, compute $\bar{\phi}_t^{\text{improper}} \to \bar{\phi}_t \in \mathcal{S}(\bar{d})$.

  Using the fixed point oracle, compute $\bar{p}_t \in \Delta(\bar{d})$ such that $\bar{p}_t = \bar{\phi}_t(\bar{p}_t)$.

  Using the relabeling function, compute $p_t \in \Delta(d)$ whose i-th entry equals the total mass of $\bar{p}_t$ on $\mathcal{I}(i)$, i.e.,

$$p_{t,i} = \sum_{\bar{i} \in \mathcal{I}(i)} \bar{p}_{t,\bar{i}}, \quad \forall i \in [1 : d].$$

  Output p_t as the t-th decision in LEA, and observe the loss vector $l_t \in \mathbb{R}^d$.

  Using the relabeling function, define $\bar{l}_t \in \mathbb{R}^{\bar{d}}$ whose $\bar{i}$-th entry equals the $\mathcal{I}^{-1}(\bar{i})$-th entry of l_t, i.e.,

$$\bar{l}_{t,\bar{i}} = l_{t,\mathcal{I}^{-1}(\bar{i})}, \quad \forall \bar{i} \in [1 : \bar{d}].$$

  Let $\bar{g}_t = \bar{p}_t \otimes \bar{l}_t \in \mathbb{R}^{\bar{d} \times \bar{d}}$.

  Using the gradient processing oracle (which depends on $\bar{\phi}_t^{\text{improper}}$), compute $\bar{g}_t \to \bar{g}_t^{\text{improper}} \in \mathbb{R}^{\bar{d} \times \bar{d}}$.

  For all $b \in \mathcal{B}$, return $\left\langle \bar{g}_t^{\text{improper}}, b \right\rangle$ as the t-th loss gradient to $\mathcal{A}_{1\mathrm{d}}^{(b)}$.

end for

B.1 Preprocessing

Based on Section 2.1, since this subsection only handles the comparing action modification rules, we will remove the superscript star from the notations ϕ^* and $\bar{\phi}^*$ for conciseness.

By construction, Algorithm 1 satisfies the loss condition $\langle p_t, l_t \rangle = \langle \bar{p}_t, \bar{l}_t \rangle$. We now show that for any comparing action modification rule $\phi \in \mathcal{S}(d)$, there exists $\bar{\phi} \in \mathcal{S}(\bar{d})$ that preserves the complexity of ϕ and further satisfies $\langle \phi(p_t), l_t \rangle = \langle \bar{\phi}(\bar{p}_t), \bar{l}_t \rangle$. This means that the ϕ-regret of Algorithm 1 can be rewritten as $\sum_{t=1}^T \langle p_t - \phi(p_t), l_t \rangle = \sum_{t=1}^T \langle \bar{p}_t - \bar{\phi}(\bar{p}_t), \bar{l}_t \rangle$, and it suffices to upper-bound the RHS as a higher dimensional problem.

First, we use $d - d_\phi^{\text{unif}}$ to measure the complexity of ϕ. We note that the following lemma critically uses the requirement that for all i, the set $\mathcal{I}(i)$ only contains consecutive integers. The cardinality condition $1 \le |\mathcal{I}(i)| \le 2$ is not used.

Lemma B.1. *For any $\phi \in \mathcal{S}(d)$, there exists $\bar{\phi} \in \mathcal{S}(\bar{d})$ such that the following two conditions hold.*

- *For all $i \in [1:\bar{d}]$, let $\bar{\phi}_i$ represent the i-th row of $\bar{\phi}$, and let $\mathbf{1}[\cdot]$ be the indicator function. Then,*

$$\sum_{i=1}^{\bar{d}-1} \mathbf{1}[\bar{\phi}_i \neq \bar{\phi}_{i+1}] \leq 2(d - d_\phi^{\mathrm{unif}}).$$

- *In Algorithm 1 we have $\langle \phi(p_t), l_t \rangle = \langle \bar{\phi}(\bar{p}_t), \bar{l}_t \rangle$.*

Proof of Lemma B.1. We construct $\bar{\phi}$ entrywise as the following. For all $i, j \in [1:\bar{d}]$,

$$\bar{\phi}_{i,j} = \frac{\phi_{\mathcal{I}^{-1}(i),\mathcal{I}^{-1}(j)}}{|\mathcal{I}(\mathcal{I}^{-1}(j))|},$$

where the denominator denotes the cardinality of the set $\mathcal{I}(\mathcal{I}^{-1}(j))$. It can be verified that

$$\sum_{j=1}^{\bar{d}} \bar{\phi}_{i,j} = \sum_{j=1}^{\bar{d}} \frac{\phi_{\mathcal{I}^{-1}(i),\mathcal{I}^{-1}(j)}}{|\mathcal{I}(\mathcal{I}^{-1}(j))|} = \sum_{\mathcal{I}^{-1}(j)=1}^{d} \phi_{\mathcal{I}^{-1}(i),\mathcal{I}^{-1}(j)} = 1,$$

therefore $\bar{\phi} \in \mathcal{S}(\bar{d})$.

For later use, consider the size-d multiset consisting of all the rows of ϕ: we write $k^* \in [1:d]$ as the index of an arbitrary element that has the highest frequency in this multiset (ties are broken arbitrarily). Since for any generic $k \in [1:d]$ and any $i, j \in \mathcal{I}(k)$, the i-th row and the j-th row of $\bar{\phi}$ are exactly the same, we denote this shared row as the vector $\bar{\phi}_{\mathcal{I}(k)} \in \mathbb{R}^{\bar{d}}$. In particular, $\bar{\phi}_{\mathcal{I}(k^*)}$ has the intuitive interpretation of a "frequent row".

Now, consider the "row-variational" quantity $\sum_{i=1}^{\bar{d}-1} \mathbf{1}[\bar{\phi}_i \neq \bar{\phi}_{i+1}]$. Due to the requirements on the relabeling function $\mathcal{I}$, we can decompose the augmented index set $[1:\bar{d}]$ into d segments, each corresponding to $\mathcal{I}(k)$ for some different $k \in [1:d]$. For any $i \in [1:\bar{d}-1]$ such that $i, i+1 \in \mathcal{I}(k)$, $\bar{\phi}_i = \bar{\phi}_{i+1} = \bar{\phi}_{\mathcal{I}(k)}$ thus $\mathbf{1}[\bar{\phi}_i \neq \bar{\phi}_{i+1}] = 0$. Therefore if we define Switch $\subset [1:\bar{d}-1]$ as the collection of all indices $i \in [1:d-1]$ such that i and $i+1$ belong to different segments, then $\sum_{i=1}^{\bar{d}-1} \mathbf{1}[\bar{\phi}_i \neq \bar{\phi}_{i+1}] = \sum_{i \in \mathrm{Switch}} \mathbf{1}[\bar{\phi}_i \neq \bar{\phi}_{i+1}]$. In plain words, the row-variation of $\bar{\phi}$ only depends on the variation at the edge of segments.

Notice that for all $i \in [1:\bar{d}-1]$, $\mathbf{1}[\bar{\phi}_i \neq \bar{\phi}_{i+1}] \leq \mathbf{1}[\bar{\phi}_i \neq \bar{\phi}_{\mathcal{I}(k^*)}] + \mathbf{1}[\bar{\phi}_{i+1} \neq \bar{\phi}_{\mathcal{I}(k^*)}]$. Therefore we can further upper-bound $\sum_{i \in \mathrm{Switch}} \mathbf{1}[\bar{\phi}_i \neq \bar{\phi}_{i+1}]$ "segment-wise", by $2\sum_{k=1}^{d} \mathbf{1}[\bar{\phi}_{\mathcal{I}(k)} \neq \bar{\phi}_{\mathcal{I}(k^*)}]$. Due to Definition 1, there are $d - d_\phi^{\mathrm{unif}}$ different values of k such that $\bar{\phi}_{\mathcal{I}(k)} \neq \bar{\phi}_{\mathcal{I}(k^*)}$. Combining above proves that $\sum_{i=1}^{\bar{d}-1} \mathbf{1}[\bar{\phi}_i \neq \bar{\phi}_{i+1}] \leq 2(d - d_\phi^{\mathrm{unif}})$.

Finally we verify the second condition in the lemma.

$$\begin{aligned}
\langle \bar{\phi}(\bar{p}_t), \bar{l}_t \rangle &= \sum_{i=1}^{\bar{d}} \bar{p}_{t,i} \sum_{j=1}^{\bar{d}} \bar{l}_{t,j} \frac{\phi_{\mathcal{I}^{-1}(i),\mathcal{I}^{-1}(j)}}{|\mathcal{I}(\mathcal{I}^{-1}(j))|} = \sum_{i=1}^{\bar{d}} \bar{p}_{t,i} \sum_{j=1}^{\bar{d}} l_{t,\mathcal{I}^{-1}(j)} \frac{\phi_{\mathcal{I}^{-1}(i),\mathcal{I}^{-1}(j)}}{|\mathcal{I}(\mathcal{I}^{-1}(j))|} \\
&= \sum_{i=1}^{\bar{d}} \bar{p}_{t,i} \sum_{\mathcal{I}^{-1}(j)=1}^{d} l_{t,\mathcal{I}^{-1}(j)} \phi_{\mathcal{I}^{-1}(i),\mathcal{I}^{-1}(j)} = \sum_{\mathcal{I}^{-1}(j)=1}^{d} l_{t,\mathcal{I}^{-1}(j)} \sum_{i=1}^{\bar{d}} \bar{p}_{t,i} \phi_{\mathcal{I}^{-1}(i),\mathcal{I}^{-1}(j)} \\
&= \sum_{\mathcal{I}^{-1}(j)=1}^{d} l_{t,\mathcal{I}^{-1}(j)} \sum_{\mathcal{I}^{-1}(i)=1}^{d} \left(\sum_{n \in \mathcal{I}(\mathcal{I}^{-1}(i))} \bar{p}_{t,n} \right) \phi_{\mathcal{I}^{-1}(i),\mathcal{I}^{-1}(j)} \\
&= \sum_{\mathcal{I}^{-1}(j)=1}^{d} l_{t,\mathcal{I}^{-1}(j)} \sum_{\mathcal{I}^{-1}(i)=1}^{d} p_{t,\mathcal{I}^{-1}(i)} \phi_{\mathcal{I}^{-1}(i),\mathcal{I}^{-1}(j)} = \sum_{i=1}^{d} \sum_{j=1}^{d} p_{t,i} \phi_{i,j} l_{t,j} = \langle \phi(p_t), l_t \rangle. \quad \square
\end{aligned}$$

The following lemma constructs $\bar{\phi}$ which preserves a different notion of complexity based on d_ϕ^{self}. Different from the previous lemma, we use the cardinality condition $1 \leq |\mathcal{I}(i)| \leq 2$ rather than the requirement that $\mathcal{I}(i)$ only contains consecutive integers.

Lemma B.2. *For any $\phi \in \mathcal{S}(d)$, there exists $\bar{\phi} \in \mathcal{S}(\bar{d})$ such that the following two conditions hold.*

- $\bar{d} - d_{\bar{\phi}}^{\text{self}} \le 2(d - d_{\phi}^{\text{self}})$.
- *In Algorithm 1 we have* $\langle \phi(p_t), l_t \rangle = \langle \bar{\phi}(\bar{p}_t), \bar{l}_t \rangle$.

Proof of Lemma B.2. Again we construct $\bar{\phi}$ entry-wise. Define the collection of "important indices" $\mathcal{I}^* \subset [1:\bar{d}]$ as $\{\min \mathcal{I}(k); k \in [1:d]\}$, and by construction $|\mathcal{I}^*| = d$. For all $i, j \in [1:\bar{d}]$, we define the (i,j)-th entry of $\bar{\phi}$ as

$$\bar{\phi}_{i,j} = \begin{cases} \phi_{\mathcal{I}^{-1}(i),\mathcal{I}^{-1}(j)}, & j = i; \\ \phi_{\mathcal{I}^{-1}(i),\mathcal{I}^{-1}(j)}, & j \notin \mathcal{I}(\mathcal{I}^{-1}(i)), j \in \mathcal{I}^*; \\ 0, & \text{else.} \end{cases}$$

It is simple to verify that $\bar{\phi} \in \mathcal{S}(\bar{d})$.

Recall that $e^{(i)}$ denotes the unit vector along the i-th coordinate, whose dimensionality depends on the context. For all $i \in [1:\bar{d}]$, we now consider $\bar{\phi}_i$, the i-th row of $\bar{\phi}$. If $\phi_{\mathcal{I}^{-1}(i)} = e^{(\mathcal{I}^{-1}(i))}$, then $\bar{\phi}_{i,i} = \phi_{\mathcal{I}^{-1}(i),\mathcal{I}^{-1}(i)} = 1$, which means $\bar{\phi}_i = e^{(i)}$. Since there are $d - d_{\phi}^{\text{self}}$ different values of $k \in [1:d]$ such that $\phi_k \ne e^{(k)}$, we can further use the cardinality condition on the relabeling function $\mathcal{I}$ to show that there are at most $2(d - d_{\phi}^{\text{self}})$ different values of $i \in [1:\bar{d}]$ such that $\bar{\phi}_i \ne e^{(i)}$. Equivalently, $\bar{d} - d_{\bar{\phi}}^{\text{self}} \le 2(d - d_{\phi}^{\text{self}})$.

Next we verify the second condition in the lemma.

$$\begin{aligned}
\langle \bar{\phi}(\bar{p}_t), \bar{l}_t \rangle &= \sum_{i=1}^{\bar{d}} \sum_{j=1}^{\bar{d}} \bar{p}_{t,i} \bar{\phi}_{i,j} \bar{l}_{t,j} \\
&= \sum_{i=1}^{\bar{d}} \bar{p}_{t,i} \left(\phi_{\mathcal{I}^{-1}(i),\mathcal{I}^{-1}(i)} l_{t,\mathcal{I}^{-1}(i)} + \sum_{j \notin \mathcal{I}(\mathcal{I}^{-1}(i)), j \in \mathcal{I}^*} \phi_{\mathcal{I}^{-1}(i),\mathcal{I}^{-1}(j)} l_{t,\mathcal{I}^{-1}(j)} \right) \\
&= \sum_{i=1}^{\bar{d}} \bar{p}_{t,i} \left(\phi_{\mathcal{I}^{-1}(i),\mathcal{I}^{-1}(i)} l_{t,\mathcal{I}^{-1}(i)} + \sum_{\mathcal{I}^{-1}(j) \ne \mathcal{I}^{-1}(i)} \phi_{\mathcal{I}^{-1}(i),\mathcal{I}^{-1}(j)} l_{t,\mathcal{I}^{-1}(j)} \right) \\
&= \sum_{i=1}^{\bar{d}} \bar{p}_{t,i} \sum_{\mathcal{I}^{-1}(j)=1}^{d} \phi_{\mathcal{I}^{-1}(i),\mathcal{I}^{-1}(j)} l_{t,\mathcal{I}^{-1}(j)} \\
&= \sum_{\mathcal{I}^{-1}(i)=1}^{d} \left(\sum_{n \in \mathcal{I}(\mathcal{I}^{-1}(i))} \bar{p}_{t,n} \right) \sum_{\mathcal{I}^{-1}(j)=1}^{d} \phi_{\mathcal{I}^{-1}(i),\mathcal{I}^{-1}(j)} l_{t,\mathcal{I}^{-1}(j)} \\
&= \sum_{\mathcal{I}^{-1}(j)=1}^{d} \sum_{\mathcal{I}^{-1}(i)=1}^{d} p_{t,\mathcal{I}^{-1}(i)} \phi_{\mathcal{I}^{-1}(i),\mathcal{I}^{-1}(j)} l_{t,\mathcal{I}^{-1}(j)} \\
&= \sum_{i=1}^{d} \sum_{j=1}^{d} p_{t,i} \phi_{i,j} l_{t,j} \\
&= \langle \phi(p_t), l_t \rangle .
\end{aligned}$$

□

B.2 Matrix Features

Based on Section 2.2, we now consider the representation of an arbitrary augmented action modification rule $\bar{\phi}^* \in \mathcal{S}(\bar{d})$ using $\mathcal{B}$, i.e., the choice of $\Phi^{*,(b)}$ such that $\bar{\phi}^* = \sum_{b \in \mathcal{B}} \Phi^{*,(b)} b$. Specifically, we upper-bound $\left|\Phi^{*,(b)}\right|$ for two different representations defined as follows, which appears in the proof of Theorem 1.

- **Representation 1.** If $b = \boldsymbol{I}_{\bar{d}}$, then $\Phi^{*,(b)} = 0$. For all $b \in \mathcal{B} \backslash \{\boldsymbol{I}_{\bar{d}}\}$, $\Phi^{*,(b)} = \frac{\langle \bar{\phi}^*, b \rangle}{\langle b, b \rangle}$.
- **Representation 2.** If $b = \boldsymbol{I}_{\bar{d}}$, then $\Phi^{*,(b)} = 1$. For all $b \in \mathcal{B} \backslash \{\boldsymbol{I}_{\bar{d}}\}$, $\Phi^{*,(b)} = \frac{\langle \bar{\phi}^* - \boldsymbol{I}_{\bar{d}}, b \rangle}{\langle b, b \rangle}$.

Due to $\mathcal{B}\backslash\{\boldsymbol{I}_{\bar{\boldsymbol{d}}}\}$ being an orthogonal basis of $\mathbb{R}^{\bar{d}\times\bar{d}}$, we have $\bar{\phi}^* = \sum_{b\in\mathcal{B}} \Phi^{*,(b)} b$ in both cases. Furthermore, since all $b \in \mathcal{B}\backslash\{\boldsymbol{I}_{\bar{\boldsymbol{d}}}\}$ can be equivalently expressed as $b = h \otimes e^{(j)}$ for some $h \in \mathcal{H}$ and $e^{(j)} \in \mathcal{E}$, we also write the associated $\Phi^{*,(b)}$ equivalently as $\Phi^{*,(h,j)}$. Let $I^{(h)} \subset [1 : \bar{d}]$ be the support of h, i.e., the collection of indices i such that the entry $h_i \neq 0$.

Lemma B.3. *For all $\bar{\phi}^* \in \mathcal{S}(\bar{d})$ and $h \in \mathcal{H}$,*

$$\sum_{j=1}^{\bar{d}} \frac{\left|\left\langle \bar{\phi}^*, h \otimes e^{(j)} \right\rangle\right|}{\left\langle h \otimes e^{(j)}, h \otimes e^{(j)} \right\rangle} \leq 1, \quad \text{and,} \quad \sum_{j=1}^{\bar{d}} \frac{\left|\left\langle \bar{\phi}^* - \boldsymbol{I}_{\bar{d}}, h \otimes e^{(j)} \right\rangle\right|}{\left\langle h \otimes e^{(j)}, h \otimes e^{(j)} \right\rangle} \leq 2.$$

That is, $\sum_{j=1}^{\bar{d}} \left|\Phi^{*,(h,j)}\right|$ is at most a constant for both the two representations introduced above.

Proof of Lemma B.3. Throughout the proof we write $b = h \otimes e^{(j)}$ for conciseness. By construction $\langle b, b\rangle = \langle h, h\rangle = \left|I^{(h)}\right|$, therefore

$$\begin{aligned}\sum_{j=1}^{\bar{d}} \frac{\left|\left\langle \bar{\phi}^*, b \right\rangle\right|}{\langle b, b\rangle} = \frac{1}{\left|I^{(h)}\right|} \sum_{j=1}^{\bar{d}} \left|\left\langle \bar{\phi}^*, b \right\rangle\right| = \frac{1}{\left|I^{(h)}\right|} \sum_{j=1}^{\bar{d}} \left| \sum_{i\in I^{(h)}} \bar{\phi}^*_{i,j} h_i \right| &\leq \frac{1}{\left|I^{(h)}\right|} \sum_{j=1}^{\bar{d}} \sum_{i\in I^{(h)}} \left|\bar{\phi}^*_{i,j} h_i\right| \\ &= \frac{1}{\left|I^{(h)}\right|} \sum_{j=1}^{\bar{d}} \sum_{i\in I^{(h)}} \bar{\phi}^*_{i,j} = \frac{1}{\left|I^{(h)}\right|} \sum_{i\in I^{(h)}} \left(\sum_{j=1}^{\bar{d}} \bar{\phi}^*_{i,j}\right) = 1,\end{aligned}$$

Since $\boldsymbol{I}_{\bar{d}} \in \mathcal{S}(\bar{d})$, we also have $\sum_{j=1}^{\bar{d}} \frac{|\langle \boldsymbol{I}_{\bar{d}}, b\rangle|}{\langle b,b\rangle} \leq 1$, thus

$$\sum_{j=1}^{\bar{d}} \frac{\left|\left\langle \bar{\phi}^* - \boldsymbol{I}_{\bar{d}}, b \right\rangle\right|}{\langle b, b\rangle} \leq \sum_{j=1}^{\bar{d}} \frac{\left|\left\langle \bar{\phi}^*, b \right\rangle\right|}{\langle b, b\rangle} + \sum_{j=1}^{\bar{d}} \frac{\left|\langle \boldsymbol{I}_{\bar{d}}, b\rangle\right|}{\langle b, b\rangle} \leq 2. \qquad \square$$

The following lemma characterizes the most important property of the Haar wavelet, for our use case. We remark that the $\bar{\phi}$ matrix below does not have to be a stochastic matrix, and we let $\bar{\phi}_i$ denote its i-th row.

Lemma B.4. *Consider an arbitrary $\bar{\phi} \in \mathbb{R}^{\bar{d}\times\bar{d}}$ and $h \in \mathcal{H}$. If for all $i_1, i_2 \in I^{(h)}$ we have $\bar{\phi}_{i_1} = \bar{\phi}_{i_2}$, then for all $j \in \bar{d}$, $\left|\left\langle \bar{\phi}, h \otimes e^{(j)} \right\rangle\right| = 0$.*

Proof of Lemma B.4. From the condition in the lemma we have $\bar{\phi}_{i,j}; i \in I^{(h)}$ being equal to some $c_j \in \mathbb{R}$ that does not depend on i. Therefore,

$$\left|\left\langle \bar{\phi}, h \otimes e^{(j)} \right\rangle\right| = \left| \sum_{i\in I^{(h)}} \bar{\phi}_{i,j} h_i \right| = \left| c_j \sum_{i\in I^{(h)}} h_i \right| = |c_j| \left| \sum_{i\in I^{(h)}} h_i \right| = 0. \qquad \square$$

B.3 Constraint Oracle

This subsection analyzes our procedure from Section 2.4 which enforces the constraint $\mathcal{S}(\bar{d})$. We specifically place our analysis in the context of Algorithm 1. The following lemma closely mirrors the analysis of [LS15, OP16], and we provide the proof for completeness.

Lemma B.5 (Appendix D of [OP16], adapted)**.** *Regarding the quantities $\bar{g}_t^{\text{improper}}$, $\bar{\phi}_t^{\text{improper}}$, $\bar{g}_t$ and $\bar{\phi}_t$ in Algorithm 1, we have for all $\bar{\phi}^* \in \mathcal{S}(\bar{d})$,*

$$\left\langle \bar{g}_t, \bar{\phi}_t - \bar{\phi}^* \right\rangle \leq \left\langle \bar{g}_t^{\text{improper}}, \bar{\phi}_t^{\text{improper}} - \bar{\phi}^* \right\rangle.$$

Proof of Lemma B.5. Recall our notation that adding a subscript i to a matrix denotes its i-th row, and adding a pair of subscripts i, j to a matrix denotes its (i, j)-th entry. Throughout this proof we will also use $\mathbf{1}$ to represent the all-one vector.

Starting from the Stage 2 of the wrapper, we have

$$\begin{aligned}
\left\langle \bar{g}_t, \bar{\phi}_t - \bar{\phi}^* \right\rangle - \left\langle \bar{g}_t^+, \bar{\phi}_t^+ - \bar{\phi}^* \right\rangle &= \sum_{i\in[1:\bar{d}]} \left(\left\langle \bar{g}_{t,i}, \bar{\phi}_{t,i} - \bar{\phi}_i^* \right\rangle - \left\langle \bar{g}_{t,i}^+, \bar{\phi}_{t,i}^+ - \bar{\phi}_i^* \right\rangle \right) \\
&= \sum_{i\in[1:\bar{d}]} \left(\left\langle \bar{g}_{t,i}, \bar{\phi}_{t,i} - \bar{\phi}_i^* \right\rangle - \left\langle \bar{g}_{t,i} - \left\langle \bar{g}_{t,i}, \bar{\phi}_{t,i} \right\rangle \mathbf{1}, \bar{\phi}_{t,i}^+ - \bar{\phi}_i^* \right\rangle \right) \quad \text{(Definition of } \bar{g}_{t,i}^+\text{)} \\
&= \sum_{i\in[1:\bar{d}]} \left(\left\langle \bar{g}_{t,i}, \bar{\phi}_{t,i} - \bar{\phi}_{t,i}^+ \right\rangle + \left\langle \bar{g}_{t,i}, \bar{\phi}_{t,i} \right\rangle \left\langle \mathbf{1}, \bar{\phi}_{t,i}^+ - \bar{\phi}_i^* \right\rangle \right).
\end{aligned}$$

Regarding the summand on the RHS, there are two cases.

- If $\left\|\bar{\phi}_{t,i}^+\right\|_1 = 0$, then since $\bar{\phi}_i^* \in \Delta(\bar{d})$,

$$\left\langle \bar{g}_{t,i}, \bar{\phi}_{t,i} - \bar{\phi}_{t,i}^+ \right\rangle + \left\langle \bar{g}_{t,i}, \bar{\phi}_{t,i} \right\rangle \left\langle \mathbf{1}, \bar{\phi}_{t,i}^+ - \bar{\phi}_i^* \right\rangle = \left\langle \bar{g}_{t,i}, \bar{\phi}_{t,i} \right\rangle + \left\langle \bar{g}_{t,i}, \bar{\phi}_{t,i} \right\rangle \left\langle \mathbf{1}, -\bar{\phi}_i^* \right\rangle = 0.$$

- If $\left\|\bar{\phi}_{t,i}^+\right\|_1 \neq 0$, then using $\bar{\phi}_{t,i} = \bar{\phi}_{t,i}^+ / \left\|\bar{\phi}_{t,i}^+\right\|_1$ and $\bar{\phi}_{t,i}^+ \in \mathbb{R}_+^{\bar{d}}$,

$$\begin{aligned}
\left\langle \bar{g}_{t,i}, \bar{\phi}_{t,i} - \bar{\phi}_{t,i}^+ \right\rangle + \left\langle \bar{g}_{t,i}, \bar{\phi}_{t,i} \right\rangle \left\langle \mathbf{1}, \bar{\phi}_{t,i}^+ - \bar{\phi}_i^* \right\rangle &= \left\langle \bar{g}_{t,i}, \frac{\bar{\phi}_{t,i}^+}{\left\|\bar{\phi}_{t,i}^+\right\|_1} - \bar{\phi}_{t,i}^+ \right\rangle + \left\langle \bar{g}_{t,i}, \frac{\bar{\phi}_{t,i}^+}{\left\|\bar{\phi}_{t,i}^+\right\|_1} \right\rangle \left\langle \mathbf{1}, \bar{\phi}_{t,i}^+ - \bar{\phi}_i^* \right\rangle \\
&= \left\langle \bar{g}_{t,i}, \frac{\bar{\phi}_{t,i}^+}{\left\|\bar{\phi}_{t,i}^+\right\|_1} - \bar{\phi}_{t,i}^+ \right\rangle + \left\langle \bar{g}_{t,i}, \frac{\bar{\phi}_{t,i}^+}{\left\|\bar{\phi}_{t,i}^+\right\|_1} \right\rangle \left(\left\|\bar{\phi}_{t,i}^+\right\|_1 - 1 \right) \\
&= 0.
\end{aligned}$$

Therefore $\left\langle \bar{g}_t, \bar{\phi}_t - \bar{\phi}^* \right\rangle = \left\langle \bar{g}_t^+, \bar{\phi}_t^+ - \bar{\phi}^* \right\rangle$.

Next, consider Stage 1 of the wrapper.

$$\begin{aligned}
&\left\langle \bar{g}_t^+, \bar{\phi}_t^+ - \bar{\phi}^* \right\rangle - \left\langle \bar{g}_t^{\text{improper}}, \bar{\phi}_t^{\text{improper}} - \bar{\phi}^* \right\rangle \\
&= \sum_{i,j\in[1:\bar{d}]} \left[\bar{g}_{t,i,j}^+ (\bar{\phi}_{t,i,j}^+ - \bar{\phi}_{i,j}^*) - \bar{g}_{t,i,j}^{\text{improper}} (\bar{\phi}_{t,i,j}^{\text{improper}} - \bar{\phi}_{i,j}^*) \right] \\
&= \sum_{i,j;\bar{\phi}_{t,i,j}^{\text{improper}}<0} \left[\bar{g}_{t,i,j}^+ (\bar{\phi}_{t,i,j}^+ - \bar{\phi}_{i,j}^*) - \bar{g}_{t,i,j}^{\text{improper}} (\bar{\phi}_{t,i,j}^{\text{improper}} - \bar{\phi}_{i,j}^*) \right] \\
&= \sum_{i,j;\bar{\phi}_{t,i,j}^{\text{improper}}<0} \left[\bar{g}_{t,i,j}^+ (-\bar{\phi}_{i,j}^*) - \min\left\{ \bar{g}_{t,i,j}^+, 0 \right\} (\bar{\phi}_{t,i,j}^{\text{improper}} - \bar{\phi}_{i,j}^*) \right] \\
&= \sum_{i,j;\bar{\phi}_{t,i,j}^{\text{improper}}<0} \left[-\min\left\{ \bar{g}_{t,i,j}^+, 0 \right\} \bar{\phi}_{t,i,j}^{\text{improper}} + \bar{\phi}_{i,j}^* (\min\left\{ \bar{g}_{t,i,j}^+, 0 \right\} - \bar{g}_{t,i,j}^+) \right] \\
&\leq 0.
\end{aligned}$$

Combining it with the result of Stage 2 above completes the proof. □

In the next lemma, for all $b \in \mathcal{B} \backslash \{\boldsymbol{I}_{\bar{d}}\}$ we write it equivalently as $b = h \otimes e^{(j)}$ for some $h \in \mathcal{H}$ and $e^{(j)} \in \mathcal{E}$. Let $I^{(h)} \subset [1:\bar{d}]$ be the support of the vector h. We emphasize that the lemma hinges on the structure of the Haar wavelet.

Lemma B.6. *With $b = h \otimes e^{(j)}$, the quantity $\left\langle \bar{g}_t^{\text{improper}}, b \right\rangle$ in Algorithm 1 satisfies*

$$\left| \left\langle \bar{g}_t^{\text{improper}}, b \right\rangle \right| \leq 2L \sum_{i\in I^{(h)}} \bar{p}_{t,i}.$$

Furthermore, $\left| \left\langle \bar{g}_t^{\text{improper}}, \boldsymbol{I}_{\bar{d}} \right\rangle \right| \leq 2L$.

Proof of Lemma B.6. First, consider the general case of $b = h \otimes e^{(j)}$.

$$\left|\left\langle \bar{g}_t^{\text{improper}}, b\right\rangle\right| = \left|\left\langle \bar{g}_t^{\text{improper}}, h \otimes e^{(j)}\right\rangle\right| = \left|\sum_{i \in I^{(h)}} \bar{g}_{t,i,j}^{\text{improper}} h_i\right| \leq \sum_{i \in I^{(h)}} \left|\bar{g}_{t,i,j}^{\text{improper}} h_i\right| = \sum_{i \in I^{(h)}} \left|\bar{g}_{t,i,j}^{\text{improper}}\right|,$$

where the last equality is due to the entries of h being ± 1 on its support. From Stage 1 of the gradient processing oracle, we have $\left|\bar{g}_{t,i,j}^{\text{improper}}\right|$ being either $\left|\bar{g}_{t,i,j}^{+}\right|$ or 0. Therefore,

$$\sum_{i \in I^{(h)}} \left|\bar{g}_{t,i,j}^{\text{improper}}\right| \leq \sum_{i \in I^{(h)}} \left|\bar{g}_{t,i,j}^{+}\right|.$$

From Stage 2 of the gradient processing oracle, and using $\bar{g}_t = \bar{p}_t \otimes \bar{l}_t$,

$$\sum_{i \in I^{(h)}} \left|\bar{g}_{t,i,j}^{+}\right| = \sum_{i \in I^{(h)}} \left|\bar{g}_{t,i,j} - \left\langle \bar{g}_{t,i}, \bar{\phi}_{t,i}\right\rangle\right| = \sum_{i \in I^{(h)}} \bar{p}_{t,i} \left|\bar{l}_{t,j} - \left\langle \bar{l}_t, \bar{\phi}_{t,i}\right\rangle\right| \leq 2L \sum_{i \in I^{(h)}} \bar{p}_{t,i}.$$

As for the case of $b = \boldsymbol{I}_{\bar{d}}$,

$$\begin{aligned}\left|\left\langle \bar{g}_t^{\text{improper}}, \boldsymbol{I}_{\bar{d}}\right\rangle\right| &= \left|\sum_{i \in [1:\bar{d}]} \bar{g}_{t,i,i}^{\text{improper}}\right| \leq \sum_{i \in [1:\bar{d}]} \left|\bar{g}_{t,i,i}^{\text{improper}}\right| \leq \sum_{i \in [1:\bar{d}]} \left|\bar{g}_{t,i,i}^{+}\right| \\ &= \sum_{i \in [1:\bar{d}]} \left|\bar{g}_{t,i,i} - \left\langle \bar{g}_{t,i}, \bar{\phi}_{t,i}\right\rangle\right| = \sum_{i \in [1:\bar{d}]} \bar{p}_{t,i} \left|\bar{l}_{t,i} - \left\langle \bar{l}_t, \bar{\phi}_{t,i}\right\rangle\right| \leq 2L \sum_{i \in [1:\bar{d}]} \bar{p}_{t,i} = 2L. \quad \square\end{aligned}$$

C Proof of Main Results

Our main result (Theorem 1) is a combination of Theorem 3 and 4 proved below.

Theorem 3. *For any comparing action modification rule $\phi^* \in \mathcal{S}(d)$, let $\bar{\phi}^* \in \mathcal{S}(\bar{d})$ be the construction from Lemma B.1. Then, there is an absolute constant $c > 0$ such that for all $T \in \mathbb{N}_+$ and $\phi^* \in \mathcal{S}(d)$, Algorithm 1 guarantees*

$$\text{Regret}_T(\phi^*) \leq c \cdot \left(L \sqrt{(T+d)\left(1 + \sum_{i=1}^{\bar{d}-1} \mathbf{1}[\bar{\phi}_i^* \neq \bar{\phi}_{i+1}^*]\right)} \cdot (\log d)^{3/2}\right).$$

Proof of Theorem 3. We structure the proof into the following three steps.

Step 1 Putting together the guarantees of individual components, arriving at the combined regret bound, Eq.(7).

Due to Lemma B.1, $\langle p_t - \phi^*(p_t), l_t\rangle = \left\langle \bar{p}_t - \bar{\phi}^*(\bar{p}_t), \bar{l}_t\right\rangle$, therefore regarding our objective we have

$$\text{Regret}_T(\phi^*) = \sum_{t=1}^{T} \left\langle \bar{p}_t - \bar{\phi}^*(\bar{p}_t), \bar{l}_t\right\rangle. \tag{6}$$

Next, for all $b \in \mathcal{B}$ we consider $\Phi^{*,(b)} \in \mathbb{R}$ defined according to Representation 1 in Appendix B.2: if $b = \boldsymbol{I}_{\bar{d}}$, then $\Phi^{*,(b)} = 0$; for all $b \in \mathcal{B} \backslash \{\boldsymbol{I}_{\bar{d}}\}$, $\Phi^{*,(b)} = \frac{\langle \bar{\phi}^*, b\rangle}{\langle b, b\rangle}$. This ensures $\bar{\phi}^* = \sum_{b \in \mathcal{B}} \Phi^{*,(b)} b$. In addition, since all $b \in \mathcal{B} \backslash \{\boldsymbol{I}_{\bar{d}}\}$ can be expressed as $b = h \otimes e^{(j)}$ for some $h \in \mathcal{H}$ and $e^{(j)} \in \mathcal{E}$, we will write their corresponding $\Phi^{*,(b)}$ equivalently as $\Phi^{*,(h,j)}$. With any fixed h, $\sum_{j=1}^{\bar{d}} \left|\Phi^{*,(h,j)}\right| \leq 1$ due to Lemma B.3.

Now consider an arbitrary $b \in \mathcal{B}$. Due to Lemma B.6, the one-dimensional gradients that Algorithm 1 sends to $\mathcal{A}_{1\text{d}}^{(b)}$ satisfy $\left|\left\langle \bar{g}_t^{\text{improper}}, b\right\rangle\right| \leq 2L$, which means Lemma 2.1 can be applied with $G = 2L$. This yields the one-dimensional regret bound with respect to $\Phi^{*,(b)}$,

$$\sum_{t=1}^{T}\left\langle\bar{g}_t^{\text{improper}},b\right\rangle(\Phi_t^{(b)}-\Phi^{*,(b)})$$

$$\leq\sqrt{8L^2+4L\sum_{t=1}^{T}\left|\left\langle\bar{g}_t^{\text{improper}},b\right\rangle\right|}\left[\sqrt{2}\varepsilon_b+2\left|\Phi^{*,(b)}\right|\sqrt{\log\left(1+\frac{\left|\Phi^{*,(b)}\right|}{\sqrt{2}\varepsilon_b}\right)}+4\left|\Phi^{*,(b)}\right|\right].$$

Taking a summation over $b\in\mathcal{B}$, the LHS becomes

$$\begin{aligned}\sum_{b\in\mathcal{B}}\sum_{t=1}^{T}\left\langle\bar{g}_t^{\text{improper}},b\right\rangle(\Phi_t^{(b)}-\Phi^{*,(b)})&=\sum_{t=1}^{T}\left\langle\bar{g}_t^{\text{improper}},\sum_{b\in\mathcal{B}}(\Phi_t^{(b)}-\Phi^{*,(b)})b\right\rangle\\&=\sum_{t=1}^{T}\left\langle\bar{g}_t^{\text{improper}},\bar{\phi}_t^{\text{improper}}-\bar{\phi}^*\right\rangle\\&\geq\sum_{t=1}^{T}\left\langle\bar{g}_t,\bar{\phi}_t-\bar{\phi}^*\right\rangle&&\text{(Lemma B.5)}\\&=\sum_{t=1}^{T}\left\langle\bar{p}_t-\bar{\phi}^*(\bar{p}_t),\bar{l}_t\right\rangle&&(\phi\text{-to-external reduction})\\&=\text{Regret}_T(\phi^*).&&\text{(Eq.(6))}\end{aligned}$$

Therefore,

$$\begin{aligned}\text{Regret}_T(\phi^*)\leq&\sum_{b\in\mathcal{B}}\sqrt{8L^2+4L\sum_{t=1}^{T}\left|\left\langle\bar{g}_t^{\text{improper}},b\right\rangle\right|}\left[\sqrt{2}\varepsilon_b+2\left|\Phi^{*,(b)}\right|\sqrt{\log\left(1+\frac{\left|\Phi^{*,(b)}\right|}{\sqrt{2}\varepsilon_b}\right)}+4\left|\Phi^{*,(b)}\right|\right]\\=&\sqrt{16L^2+8L\sum_{t=1}^{T}\left|\left\langle\bar{g}_t^{\text{improper}},\boldsymbol{I}_{\bar{d}}\right\rangle\right|}\\&+\sum_{j=1}^{\bar{d}}\sqrt{8L^2+4L\sum_{t=1}^{T}\left|\left\langle\bar{g}_t^{\text{improper}},h^{(S,0)}\otimes e^{(j)}\right\rangle\right|}\left[\frac{\sqrt{2}\bar{d}}{d^2\left|\mathcal{I}(\mathcal{I}^{-1}(j))\right|}\right.\\&\left.+2\left|\Phi^{*,(h^{(S,0)},j)}\right|\sqrt{\log\left(1+\frac{\left|\Phi^{*,(h^{(S,0)},j)}\right|}{\sqrt{2}\bar{d}(d^2\left|\mathcal{I}(\mathcal{I}^{-1}(j))\right|)^{-1}}\right)}+4\left|\Phi^{*,(h^{(S,0)},j)}\right|\right]\\&+\sum_{h\in\mathcal{H}\setminus\{h^{(S,0)}\}}\sum_{j=1}^{\bar{d}}\sqrt{8L^2+4L\sum_{t=1}^{T}\left|\left\langle\bar{g}_t^{\text{improper}},h\otimes e^{(j)}\right\rangle\right|}\left[\frac{\sqrt{2}}{d^2\left|\mathcal{I}(\mathcal{I}^{-1}(j))\right|}\right.\\&\left.+2\left|\Phi^{*,(h,j)}\right|\sqrt{\log\left(1+\frac{\left|\Phi^{*,(h,j)}\right|}{\sqrt{2}(d^2\left|\mathcal{I}(\mathcal{I}^{-1}(j))\right|)^{-1}}\right)}+4\left|\Phi^{*,(h,j)}\right|\right].\end{aligned}\tag{7}$$

Step 2 Separately bounding the three parts in Eq.(7), corresponding to three different types of b.

We now separately bound the three terms on the RHS of Eq.(7). Denote them as PartOne (the first line), PartTwo (the second and third line), and PartThree (the last two lines).

- Due to Lemma B.6, $\left|\left\langle\bar{g}_t^{\text{improper}},\boldsymbol{I}_{\bar{d}}\right\rangle\right|\leq 2L$, therefore

$$\text{PartOne}=\sqrt{16L^2+8L\sum_{t=1}^{T}\left|\left\langle\bar{g}_t^{\text{improper}},\boldsymbol{I}_{\bar{d}}\right\rangle\right|}\leq 4L\sqrt{T+1}.$$

- Due to Lemma B.6, $\left|\left\langle \bar{g}_t^{\text{improper}}, h^{(S,0)} \otimes e^{(j)} \right\rangle\right| \le 2L \sum_{i=1}^{\bar{d}} \bar{p}_{t,i} = 2L$, therefore,

$$
\begin{aligned}
\text{PartTwo} &\le 2\sqrt{2}L\sqrt{T+1} \sum_{j=1}^{\bar{d}} \Bigg[\frac{\sqrt{2}\bar{d}}{d^2 \left|\mathcal{I}(\mathcal{I}^{-1}(j))\right|} \\
&\quad + 2\left|\Phi^{*,(h^{(S,0)},j)}\right| \sqrt{\log\left(1 + \frac{\left|\Phi^{*,(h^{(S,0)},j)}\right|}{\sqrt{2}\bar{d}(d^2 \left|\mathcal{I}(\mathcal{I}^{-1}(j))\right|)^{-1}}\right)} + 4\left|\Phi^{*,(h^{(S,0)},j)}\right| \Bigg] \\
&= 2\sqrt{2}L\sqrt{T+1} \Bigg[\sqrt{2}\bar{d}d^{-1} \\
&\quad + 2\sum_{j=1}^{\bar{d}} \left|\Phi^{*,(h^{(S,0)},j)}\right| \sqrt{\log\left(1 + \frac{\left|\Phi^{*,(h^{(S,0)},j)}\right|}{\sqrt{2}\bar{d}(d^2 \left|\mathcal{I}(\mathcal{I}^{-1}(j))\right|)^{-1}}\right)} + 4\sum_{j=1}^{\bar{d}} \left|\Phi^{*,(h^{(S,0)},j)}\right| \Bigg] \\
&\le 2\sqrt{2}L\sqrt{T+1} \left[2\sqrt{2} + 4 + 2\sum_{j=1}^{\bar{d}} \left|\Phi^{*,(h^{(S,0)},j)}\right| \sqrt{\log\left(1 + \frac{\left|\Phi^{*,(h^{(S,0)},j)}\right|}{\sqrt{2}\bar{d}(d^2 \left|\mathcal{I}(\mathcal{I}^{-1}(j))\right|)^{-1}}\right)} \right]. \qquad (8)
\end{aligned}
$$

For now we use a crude bound: $d \le \bar{d}$, $\left|\mathcal{I}(\mathcal{I}^{-1}(j))\right| \le 2$ and $\left|\Phi^{*,(h^{(S,0)},j)}\right| \le 1$, therefore

$$
\begin{aligned}
\text{PartTwo} &\le 2\sqrt{2}L\sqrt{T+1} \left[2\sqrt{2} + 4 + 2\sum_{j=1}^{\bar{d}} \left|\Phi^{*,(h^{(S,0)},j)}\right| \sqrt{\log\left(1+\sqrt{2}d\right)} \right] \\
&\le 2\sqrt{2}L\sqrt{T+1} \left[2\sqrt{2} + 4 + 2\sqrt{\log\left(1+\sqrt{2}d\right)} \right] \\
&= O(L\sqrt{T \log d}).
\end{aligned}
$$

Later on we will return to Eq.(8) when analyzing the quantile regret.

- Due to Lemma B.6, $\left|\left\langle \bar{g}_t^{\text{improper}}, h \otimes e^{(j)} \right\rangle\right| \le 2L \sum_{i \in I^{(h)}} \bar{p}_{t,i}$ for all $h \in \mathcal{H}$. We also have $\left|\mathcal{I}(\mathcal{I}^{-1}(j))\right| \le 2$ and $\left|\Phi^{*,(h,j)}\right| \le 1$. Therefore similar to the above,

$$
\begin{aligned}
&\text{PartThree} \\
&\le 2\sqrt{2}L \sum_{h \in \mathcal{H}\backslash\{h^{(S,0)}\}} \sum_{j=1}^{\bar{d}} \sqrt{1 + \sum_{t=1}^{T} \sum_{i \in I^{(h)}} \bar{p}_{t,i}} \left[\frac{\sqrt{2}}{d^2 \left|\mathcal{I}(\mathcal{I}^{-1}(j))\right|} + 2\left|\Phi^{*,(h,j)}\right| \left(\sqrt{\log\left(1+\sqrt{2}d^2\right)} + 2 \right) \right] \\
&= 2\sqrt{2}L \sum_{h \in \mathcal{H}\backslash\{h^{(S,0)}\}} \sqrt{1 + \sum_{t=1}^{T} \sum_{i \in I^{(h)}} \bar{p}_{t,i}} \left[\frac{\sqrt{2}}{d} + 2\sum_{j=1}^{\bar{d}} \left|\Phi^{*,(h,j)}\right| \left(\sqrt{\log\left(1+\sqrt{2}d^2\right)} + 2 \right) \right] \\
&\le 4L\sqrt{T+1} + 4\sqrt{2}\left(\sqrt{\log\left(1+\sqrt{2}d^2\right)} + 2 \right) L \underbrace{\sum_{h \in \mathcal{H}\backslash\{h^{(S,0)}\}} \left(\sum_{j=1}^{\bar{d}} \left|\Phi^{*,(h,j)}\right| \right) \sqrt{1 + \sum_{t=1}^{T} \sum_{i \in I^{(h)}} \bar{p}_{t,i}}}_{=:\Diamond}.
\end{aligned}
$$

The key step of our analysis is bounding this $\Diamond$ term.

Step 3 Exploiting sparsity.

To this end, notice that for all $h \in \mathcal{H}\backslash\{h^{(S,0)}\}$, we can write it as $h^{(s,l)}$ for some scale parameter $s \in [1:S]$ and some location parameter $l \in [1:2^{S-s}]$. Accordingly, we will write $\Phi^{*,(h,j)}$ equivalently as $\Phi^{*,(s,l,j)}$. Then,

$$
\Diamond = \sum_{s=1}^{S} \sum_{l=1}^{2^{S-s}} \left(\sum_{j=1}^{\bar{d}} \left|\Phi^{*,(s,l,j)}\right| \right) \sqrt{1 + \sum_{t=1}^{T} \sum_{i \in I^{(s,l)}} \bar{p}_{t,i}}
$$

$$= \sum_{s=1}^{S} \sum_{l=1}^{2^{S-s}} \left(\sum_{j=1}^{\bar{d}} \frac{\left| \left\langle \bar{\phi}^*, h^{(s,l)} \otimes e^{(j)} \right\rangle \right|}{\left\langle h^{(s,l)} \otimes e^{(j)}, h^{(s,l)} \otimes e^{(j)} \right\rangle} \right) \sqrt{1 + \sum_{t=1}^{T} \sum_{i \in I^{(s,l)}} \bar{p}_{t,i}}$$
$$= \sum_{s=1}^{S} \sum_{l=1}^{2^{S-s}} \left(\frac{1}{\left| I^{(s,l)} \right|} \sum_{j=1}^{\bar{d}} \left| \left\langle \bar{\phi}^*, h^{(s,l)} \otimes e^{(j)} \right\rangle \right| \right) \sqrt{1 + \sum_{t=1}^{T} \sum_{i \in I^{(s,l)}} \bar{p}_{t,i}}.$$

Now consider fixing s and letting l vary. Due to Lemma B.4, if $\bar{\phi}^*$ remains the same for all its rows with indices in $I^{(s,l)}$, then $\sum_{j=1}^{\bar{d}} \left| \left\langle \bar{\phi}^*, h^{(s,l)} \otimes e^{(j)} \right\rangle \right| = 0$. Since the number of row-changes is $\sum_{i=1}^{\bar{d}-1} \mathbf{1}[\bar{\phi}_i^* \neq \bar{\phi}_{i+1}^*]$, there are at most this amount of l such that $\sum_{j=1}^{\bar{d}} \left| \left\langle \bar{\phi}^*, h^{(s,l)} \otimes e^{(j)} \right\rangle \right| \neq 0$. Therefore using Cauchy-Schwarz,

$$\begin{aligned}
&\sum_{l=1}^{2^{S-s}} \left(\frac{1}{\left| I^{(s,l)} \right|} \sum_{j=1}^{\bar{d}} \left| \left\langle \bar{\phi}^*, h^{(s,l)} \otimes e^{(j)} \right\rangle \right| \right) \sqrt{1 + \sum_{t=1}^{T} \sum_{i \in I^{(s,l)}} \bar{p}_{t,i}} \\
&= \sum_{l; l \in [1:2^{S-s}], \sum_{j=1}^{\bar{d}} \left| \left\langle \bar{\phi}^*, h^{(s,l)} \otimes e^{(j)} \right\rangle \right| \neq 0} \left(\frac{1}{\left| I^{(s,l)} \right|} \sum_{j=1}^{\bar{d}} \left| \left\langle \bar{\phi}^*, h^{(s,l)} \otimes e^{(j)} \right\rangle \right| \right) \sqrt{1 + \sum_{t=1}^{T} \sum_{i \in I^{(s,l)}} \bar{p}_{t,i}} \\
&\leq \sqrt{\left| \left\{ l; l \in [1:2^{S-s}], \sum_{j=1}^{\bar{d}} \left| \left\langle \bar{\phi}^*, h^{(s,l)} \otimes e^{(j)} \right\rangle \right| \neq 0 \right\} \right|} \\
&\quad \cdot \sqrt{\sum_{\substack{l: l \in [1:2^{S-s}], \\ \sum_{j=1}^{\bar{d}} \left| \left\langle \bar{\phi}^*, h^{(s,l)} \otimes e^{(j)} \right\rangle \right| \neq 0}} \left(\frac{1}{\left| I^{(s,l)} \right|} \sum_{j=1}^{\bar{d}} \left| \left\langle \bar{\phi}^*, h^{(s,l)} \otimes e^{(j)} \right\rangle \right| \right)^2 \left(1 + \sum_{t=1}^{T} \sum_{i \in I^{(s,l)}} \bar{p}_{t,i} \right)} \\
&\leq \sqrt{\sum_{i=1}^{\bar{d}-1} \mathbf{1}[\bar{\phi}_i^* \neq \bar{\phi}_{i+1}^*]} \sqrt{\sum_{l=1}^{2^{S-s}} \left(\frac{1}{\left| I^{(s,l)} \right|} \sum_{j=1}^{\bar{d}} \left| \left\langle \bar{\phi}^*, h^{(s,l)} \otimes e^{(j)} \right\rangle \right| \right)^2 \left(1 + \sum_{t=1}^{T} \sum_{i \in I^{(s,l)}} \bar{p}_{t,i} \right)} \\
&\leq \sqrt{\sum_{i=1}^{\bar{d}-1} \mathbf{1}[\bar{\phi}_i^* \neq \bar{\phi}_{i+1}^*]} \sqrt{\sum_{l=1}^{2^{S-s}} \left(1 + \sum_{t=1}^{T} \sum_{i \in I^{(s,l)}} \bar{p}_{t,i} \right)} \qquad \text{(Lemma B.3)} \\
&\leq \sqrt{\sum_{i=1}^{\bar{d}-1} \mathbf{1}[\bar{\phi}_i^* \neq \bar{\phi}_{i+1}^*]} \sqrt{\bar{d} + \sum_{t=1}^{T} \sum_{l=1}^{2^{S-s}} \sum_{i \in I^{(s,l)}} \bar{p}_{t,i}} \\
&\leq \sqrt{\sum_{i=1}^{\bar{d}-1} \mathbf{1}[\bar{\phi}_i^* \neq \bar{\phi}_{i+1}^*]} \sqrt{2d + T}.
\end{aligned}$$

Since this holds for all $s \in [1:S]$ and $S = O(\log d)$, overall we have

$$\Diamond = O\left(\sqrt{(T+d) \sum_{i=1}^{\bar{d}-1} \mathbf{1}[\bar{\phi}_i^* \neq \bar{\phi}_{i+1}^*] \cdot \log d} \right).$$

Combining everything above completes the proof. $\square$

Theorem 4. *There is an absolute constant $c > 0$ such that for all $T \in \mathbb{N}_+$ and $\phi^* \in \mathcal{S}(d)$, Algorithm 1 guarantees*

$$\mathrm{Regret}_T(\phi^*) \leq c \cdot \left(L \sqrt{(d - d_{\phi^*}^{\mathrm{self}})(T+d)} \cdot (\log d)^{3/2} \right).$$

Proof of Theorem 4. We start by following the same proof as Theorem 3, with only two differences:

- For any $\phi^* \in \mathcal{S}(d)$, we let $\bar{\phi}^* \in \mathcal{S}(\bar{d})$ be the construction from Lemma B.2.
- $\Phi^{*,(b)}$ is defined according to Representation 2 in Appendix 2.2.

Then, analogous to Eq.(7), we have

$$\text{Regret}_T(\phi^*) \leq \text{PartOne} + \text{PartTwo} + \text{PartThree},$$

where PartOne $= O(L\sqrt{T})$, PartTwo $= O(L\sqrt{T \log d})$, and

$$\text{PartThree} \leq O(L\sqrt{T}) + O(L\sqrt{\log d}) \underbrace{\sum_{h \in \mathcal{H} \setminus \{h^{(S,0)}\}} \left(\sum_{j=1}^{\bar{d}} \left| \Phi^{*,(h,j)} \right| \right) \sqrt{1 + \sum_{t=1}^{T} \sum_{i \in I^{(h)}} \bar{p}_{t,i}}}_{=:\lozenge}.$$

Everything here is only different from the corresponding term in Theorem 3 by constant multiplying factors.

The analysis of $\lozenge$ is also similar to Step 3 in the proof of Theorem 3, with $\bar{\phi}^*$ replaced by $\bar{\phi}^* - \boldsymbol{I}_{\bar{d}}$. By the definition of $d^{\text{self}}_{\bar{\phi}^*}$, only $\bar{d} - d^{\text{self}}_{\bar{\phi}^*}$ rows of $\bar{\phi}^* - \boldsymbol{I}_{\bar{d}}$ are nonzero. Accordingly, there are at most $2(\bar{d} - d^{\text{self}}_{\bar{\phi}^*})$ row-changes, therefore

$$\lozenge = O\left(\sqrt{\left(\bar{d} - d^{\text{self}}_{\bar{\phi}^*}\right)(T + d) \cdot \log d} \right).$$

Due to Lemma B.2, $\bar{d} - d^{\text{self}}_{\bar{\phi}^*}$ here can be further replaced by $d - d^{\text{self}}_{\phi^*}$.

Combining the above proves the theorem in the case of $d^{\text{self}}_{\bar{\phi}^*} \leq d - 1$. Finally observe that $\text{Regret}_T(\phi^*)$ trivially equals 0 when $d^{\text{self}}_{\bar{\phi}^*} = d$ (i.e., $\phi^* = \boldsymbol{I}_d$), therefore the regret bound in the theorem holds in that case as well. $\square$

Next, we restate and prove the quantile regret bound of Algorithm 1.

Theorem 2. *There exists an absolute constant $c > 0$ such that for all $\varepsilon \in [d^{-1} : 1]$, Algorithm 1 guarantees the quantile regret bound*

$$\text{Regret}_T(\varepsilon) \leq c \cdot \left(L\sqrt{T \log \varepsilon^{-1}} \right).$$

Proof of Theorem 2. We start by considering the ϕ-regret $\text{Regret}_T(\phi^*)$, where $\phi^* = \mathbf{1} \otimes u \in \mathcal{S}(d)$ for some $u \in \Delta(d)$ (here $\mathbf{1}$ denotes the length-d all-one vector). Following exactly the proof of Theorem 3 until Eq.(7), we have

$$\text{Regret}_T(\phi^*) \leq \text{PartOne} + \text{PartTwo} + \text{PartThree},$$

where PartOne $\leq 4L\sqrt{T+1}$, and PartThree $\leq 4L\sqrt{T+1}$ due to $\lozenge = 0$ (the definition of $\lozenge$ is the same as in the proof of Theorem 3). It only remains to analyze PartTwo.

Now consider $\Phi^{*,(h^{(S,0)},j)}$ that appears in PartTwo.

$$\Phi^{*,(h^{(S,0)},j)} = \frac{\left\langle \bar{\phi}^*, h^{(S,0)} \otimes e^{(j)} \right\rangle}{\left\langle h^{(S,0)} \otimes e^{(j)}, h^{(S,0)} \otimes e^{(j)} \right\rangle} = d^{-1} \sum_{i=1}^{\bar{d}} \bar{\phi}^*_{i,j},$$

where due to the construction of Lemma B.1, $\bar{\phi}^*_{i,j} = \frac{\phi^*_{\mathcal{I}^{-1}(i),\mathcal{I}^{-1}(j)}}{|\mathcal{I}(\mathcal{I}^{-1}(j))|}$. Therefore,

$$\Phi^{*,(h^{(S,0)},j)} = d^{-1} \sum_{i=1}^{\bar{d}} \frac{\phi^*_{\mathcal{I}^{-1}(i),\mathcal{I}^{-1}(j)}}{|\mathcal{I}(\mathcal{I}^{-1}(j))|} = \bar{d}d^{-1} \frac{u_{\mathcal{I}^{-1}(j)}}{|\mathcal{I}(\mathcal{I}^{-1}(j))|}.$$

Continuing from Eq.(8),

$$\text{PartTwo} \leq 2\sqrt{2}L\sqrt{T+1} \left[2\sqrt{2} + 4 + 2 \sum_{j=1}^{\bar{d}} \left| \Phi^{*,(h^{(S,0)},j)} \right| \sqrt{\log \left(1 + \frac{\left| \Phi^{*,(h^{(S,0)},j)} \right|}{\sqrt{2}\bar{d}(d^2 \left| \mathcal{I}(\mathcal{I}^{-1}(j)) \right|)^{-1}} \right)} \right],$$

where

$$\sum_{j=1}^{\bar{d}} \left|\Phi^{*,(h^{(S,0)},j)}\right| \sqrt{\log\left(1+\frac{\left|\Phi^{*,(h^{(S,0)},j)}\right|}{\sqrt{2}(d\left|\mathcal{I}(\mathcal{I}^{-1}(j))\right|)^{-1}}\right)} = \bar{d}d^{-1}\sum_{j=1}^{\bar{d}} \frac{u_{\mathcal{I}^{-1}(j)}}{\left|\mathcal{I}(\mathcal{I}^{-1}(j))\right|}\sqrt{\log\left(1+d\frac{u_{\mathcal{I}^{-1}(j)}}{\sqrt{2}}\right)}$$

$$= \bar{d}d^{-1}\sum_{\mathcal{I}^{-1}(j)=1}^{d} u_{\mathcal{I}^{-1}(j)}\sqrt{\log\left(1+d\frac{u_{\mathcal{I}^{-1}(j)}}{\sqrt{2}}\right)}$$

$$\le 2\sum_{i=1}^{d} u_i\sqrt{\log\left(1+\frac{u_i}{\sqrt{2}d^{-1}}\right)}.$$

Putting things together,

$$\mathrm{Regret}_T(\phi^*) = O\left(L\sqrt{T}\sum_{i=1}^{d} u_i\sqrt{\log\left(1+\frac{u_i}{\sqrt{2}d^{-1}}\right)}\right).$$

From this, bounding the quantile regret $\mathrm{Regret}_T(\varepsilon)$ follows from the standard trick [Ora23, Remark 9.16]. $\mathrm{Regret}_T(\varepsilon) = \mathrm{Regret}_T(\phi^*)$ when $u \in \Delta(d)$ is defined entrywise as follows: if the i-th expert is among the $\lceil \varepsilon d\rceil$ best experts (denoted as the index set I^*), then $u_i = 1/\lceil \varepsilon d\rceil$; otherwise $u_i = 0$. Then,

$$\mathrm{Regret}_T(\varepsilon) = O\left(L\sqrt{T}\sum_{i\in I^*}\frac{1}{\lceil \varepsilon d\rceil}\sqrt{\log\left(1+\frac{d}{\sqrt{2}\lceil \varepsilon d\rceil}\right)}\right) = O\left(L\sqrt{T\log \varepsilon^{-1}}\right). \qquad \square$$

D Discussion

In this section we discuss the strength of our result in the regime of large $d^{\mathrm{self}}_{\phi^*}$, complementing the regime of large $d^{\mathrm{unif}}_{\phi^*}$ discussed in Section 3.

Let us consider $\mathrm{Regret}_T(\phi^*)$ where the comparing action modification rule ϕ^* satisfies $d^{\mathrm{self}}_{\phi^*} = d - k$. In other words, ϕ^* only nontrivially modifies k of the d experts; the rest are kept unchanged. Due to Theorem 1, Algorithm 1 guarantees $\mathrm{Regret}_T(\phi^*) = \tilde{O}(\sqrt{kT})$ while being computationally efficient and agnostic to k. This improves

- the $\tilde{O}(\sqrt{dT})$ bound achieved by generic swap regret minimization [BM07];
- the $\tilde{O}(k\sqrt{T})$ bound achieved by standard internal regret minimization [CBL06, Chapter 4.4], since the considered $\mathrm{Regret}_T(\phi^*)$ is at most k times the internal regret; and
- the computationally inefficient and k-dependent approach to achieve the $\tilde{O}(\sqrt{kT})$ bound, which runs MWU over all zero-one stochastic matrices ϕ satisfying $d^{\mathrm{self}}_{\phi} = d - k$.

Additionally, we devote special attention to the following baseline. This is a fairly straightforward corollary of [Rot23, Section 3.3], despite not being formally stated there. The baseline concerns internal regret minimization: following [Rot23, Definition 32], it formulates the internal regret as a special case of the subsequence regret [Rot23, Definition 28 and 29], which is then handled by a generic reduction from [Rot23, Theorem 14]. To be more specific, this reduction requires a "meta" LEA algorithm as input,[5] and [Rot23, Theorem 14] showed that the associated subsequence regret can be bounded by the coordinate-wise external regret of the input "meta" LEA algorithm. [NRRX23] further suggested initiating this reduction with a "coordinate-wise, second-order adaptive" LEA algorithm from [CLW21]. Combining everything, this yields a computationally efficient algorithm for our setting, whose internal regret bound is

$$\sum_{t=1}^{T} p_{t,i}(l_{t,i}-l_{t,j}) = O\left(\sqrt{\sum_{t=1}^{T} p_{t,i}^2\cdot\log(dT)}\right), \quad \forall i,j\in[1:d].$$

[5]Intuitively, the "meta" LEA algorithm treats each subsequence as a "meta" expert, and reweighs their importance according to their performance (i.e., subsequence regret). The higher the subsequence regret is, the larger weight the corresponding "meta" expert would receive. This "scalarizes" the multi-objective problem of minimizing all subsequence regret to a single-objective problem of minimizing the weighted sum of these regrets.

Here, $p_{t,i}$ denotes the i-th coordinate of the prediction $p_t \in \Delta(d)$, and $l_{t,i}, l_{t,j}$ denote the i-th and the j-th coordinate of the loss vector l_t. This result sharpens the standard $O(\sqrt{T \log d})$ internal regret bound of [CBL06, Chapter 4.4], modulo log factors.

Now, consider $\mathrm{Regret}_T(\phi^*)$ again with $d^{\mathrm{self}}_{\phi^*} = d - k$. Without knowing k, the above internal regret minimization baseline guarantees

$$\begin{aligned}
\mathrm{Regret}_T(\phi^*) &= O\left(\sum_{i;\phi^*(e^{(i)})\neq e^{(i)}} \sqrt{\sum_{t=1}^{T} p_{t,i}^2 \cdot \log(dT)}\right) \\
&= O\left(\sqrt{k} \cdot \sqrt{\sum_{i;\phi^*(e^{(i)})\neq e^{(i)}} \sum_{t=1}^{T} p_{t,i}^2 \cdot \log(dT)}\right) && \text{(Cauchy-Schwarz)} \\
&= O\left(\sqrt{k} \cdot \sqrt{\sum_{t=1}^{T} \sum_{i=1}^{d} p_{t,i}^2 \cdot \log(dT)}\right) \\
&= O\left(\sqrt{kT} \cdot \log(dT)\right).
\end{aligned}$$

Although also being $\tilde{O}(\sqrt{kT})$, its difference with our result is the log factor: the above log factor depends on T, whereas the log factor in our Theorem 1 does not. Recall that in our setting, T can be arbitrarily larger than d.

Finally, we remark that the last line of the above bound shows the second-order adaptivity of [CLW21] is a slight overkill. Using the "coordinate-wise first-order adaptive" LEA algorithm of [LS15] would guarantee $\mathrm{Regret}_T(\phi^*) = O\left(\sqrt{kT} \cdot (\log d + \log\log T)\right)$, but we are not aware of any existing approach that completely removes the T-dependence of the log factor.