Online Learning with a Hint

Ofer Dekel
Microsoft Research
oferd@microsoft.com

Arthur Flajolet
Operations Research Center
Massachusetts Institute of Technology
flajolet@mit.edu

Nika Haghtalab
Computer Science Department
Carnegie Mellon University
nika@cmu.edu

Patrick Jaillet
EECS Department
Massachusetts Institute of Technology
jaillet@mit.edu

Abstract

We study a variant of online linear optimization where the player receives a hint about the loss function at the beginning of each round. The hint is given in the form of a vector that is weakly correlated with the loss vector on that round. We show that the player can benefit from such a hint if the set of feasible actions is sufficiently round. Specifically, if the set is strongly convex, the hint can be used to guarantee a regret of $O(\log(T))$, and if the set is q-uniformly convex for $q \in (2, 3)$, the hint can be used to guarantee a regret of $o(\sqrt{T})$. In contrast, we establish $\Omega(\sqrt{T})$ lower bounds on regret when the set of feasible actions is a polyhedron.

1 Introduction

Online linear optimization is a canonical problem in online learning. In this setting, a player attempts to minimize an online adversarial sequence of loss functions while incurring a small *regret*, compared to the best offline solution. Many online algorithms exist that are designed to have a regret of $O(\sqrt{T})$ in the worst case and it has been known that one cannot avoid a regret of $\Omega(\sqrt{T})$ in the worst case. While this worst-case perspective on online linear optimization has lead to elegant algorithms and deep connections to other fields, such as boosting [9, 10] and game theory [4, 2], it can be overly pessimistic. In particular, it does not account for the fact that the player may have side-information that allows him to anticipate the upcoming loss functions and evade the $\Omega(\sqrt{T})$ regret. In this work, we go beyond this worst case analysis and consider online linear optimization when additional information in the form of a function that is correlated with the loss is presented to the player.

More formally, online convex optimization [22, 11] is a T-round repeated game between a player and an adversary. On each round, the player chooses an action x_t from a convex set of feasible actions $\mathcal{K} \subseteq \mathbb{R}^d$ and the adversary chooses a convex bounded loss function f_t. Both choices are revealed and the player incurs a loss of $f_t(x_t)$. The player then uses its knowledge of f_t to adjust its strategy for the subsequent rounds. The player's goal is to accumulate a small loss compared to the best fixed action in hindsight. This value is called *regret* and is a measure of success of the player's algorithm.

When the adversary is restricted to Lipschitz loss functions, several algorithms are known to guarantee $O(\sqrt{T})$ regret [22, 16, 11]. If we further restrict the adversary to strongly convex loss functions, the regret bound improves to $O(\log(T))$ [14]. However, when the loss functions are linear, no online algorithm can have a regret of $o(\sqrt{T})$ [5]. In this sense, linear loss functions are the most difficult convex loss functions to handle [22].

In this paper, we focus on the case where the adversary is restricted to linear Lipschitz loss functions. More specifically, we assume that the loss function $f_t(x)$ takes the form $c_t^\intercal x$, where c_t is a bounded loss vector in $\mathcal{C} \subseteq \mathbb{R}^d$. We further assume that the player receives a *hint* before choosing the action on each round. The hint in our setting is a vector that is guaranteed to be weakly correlated with the loss vector. Namely, at the beginning of round t, the player observes a unit-length vector $v_t \in \mathbb{R}^d$ such that $v_t^\intercal c_t \geq \alpha \|c_t\|_2$, and where α is a small positive constant. So long as this requirement is met, the hint could be chosen maliciously, possibly by an adversary who knows how the player's algorithm uses the hint. Our goal is to develop a player strategy that takes these hints into account, and to understand when hints of this type make the problem provably easier and lead to smaller regret.

We show that the player's ability to benefit from the hints depends on the geometry of the player's action set $\mathcal{K}$. Specifically, we characterize the roundness of the set $\mathcal{K}$ using the notion of (C, q)-uniform convexity for convex sets. In Section 3, we show that if $\mathcal{K}$ is a $(C, 2)$-uniformly convex set (or in other words, if $\mathcal{K}$ is a C-strongly convex set), then we can use the hint to design a player strategy that improves the regret guarantee to $O\big((C\alpha)^{-1}\log(T)\big)$, where our $O(\cdot)$ notation hides a polynomial dependence on the dimension d and other constants. Furthermore, as we show in Section 4, if $\mathcal{K}$ is a (C, q)-uniformly convex set for $q \in (2, 3)$, we can use the hint to improve the regret to $O\left((C\alpha)^{\frac{1}{1-q}} T^{\frac{2-q}{1-q}}\right)$, when the hint belongs to a small set of possible hints at every step.

In Section 5, we prove lower bounds on the regret of any online algorithm in this model. We first show that when $\mathcal{K}$ is a polyhedron, such as a L_1 or a L_∞ ball, even a stronger form of hint cannot guarantee a regret of $o(\sqrt{T})$. Furthermore, we prove a lower bound of $\Omega(\log(T))$ regret when $\mathcal{K}$ is strongly convex.

1.1 Comparison with Other Notions of Hints

The notion of hint that we introduce in this work generalizes some of the notions of predictability on online learning. Hazan and Megiddo [13] considered as an example a setting where the player knows the first coordinate of the loss vector at all rounds, and showed that when $|c_{t1}| \geq \alpha$ and when the set of feasible actions is the Euclidean ball, one can achieve a regret of $O(1/\alpha \cdot \log(T))$. Our work directly improves over this result, as in our setting a hint $v_t = \pm e_1$ also achieves $O(1/\alpha \cdot \log(T))$ regret, but we can deal with hints in different directions at different rounds and we allow for general uniformly convex action sets. Rakhlin and Sridharan [20] considered online learning with predictable sequences, with a notion of predictability that is concerned with the gradient of the convex loss functions. They show that if the player receives a hint M_t at round t, then the regret of the algorithm is at most $O(\sqrt{\sum_{t=1}^T \|\nabla f_t(x_t) - M_t\|_*^2})$. In the case of linear loss functions, this implies that having an estimate vector c'_t of the loss vector within distance σ of the true loss vector c_t results in an improved regret bound of $O(\sigma\sqrt{T})$. In contrast, we consider a notion of hint that pertains to the *direction of the loss vector* rather than its location. Our work shows that merely knowing whether the loss vector positively or negatively correlates with another vector is sufficient to achieve improved regret bound, when the set is uniformly convex. That is, rather than having access to an approximate value of c_t, we only need to have access to a halfspace that classifies c_t correctly with a margin. This notion of hint is weaker that the notion of hint in the work of Rakhlin and Sridharan [20] when the approximation error satisfies $\|c_t - c'_t\|_2 \leq \sigma \cdot \|c_t\|_2$ for $\sigma \in [0, 1)$. In this case one can use $c'_t / \|c'_t\|_2$ as the direction of the hint in our setting and achieve a regret of $O(\frac{1}{1-\sigma}\log T)$ when the set is strongly convex. This shows that when the set of feasible actions is strongly convex, a directional hint can improve the regret bound beyond what has been known to be achievable by an approximation hint. However, we note that our results require the hints to be always valid, whereas the algorithm of Rakhlin and Sridharan [19] can adapt to the quality of the hints.

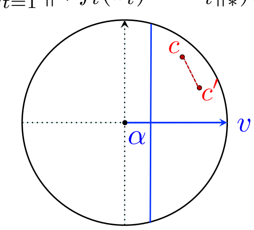

Figure 1: Comparison between notions of hint.

We discuss these works and other related works, such as [15], in more details in Appendix A.

2 Preliminaries

We begin with a more formal definition of online linear optimization (without hints). Let $\mathcal{A}$ denote the player's algorithm for choosing its actions. On round t the player uses $\mathcal{A}$ and all of the information it has observed so far to choose an action x_t in a convex compact set $\mathcal{K} \subseteq \mathbb{R}^d$. Subsequently, the adversary chooses a loss vector c_t in a compact set $\mathcal{C} \subseteq \mathbb{R}^d$. The player and the adversary reveal their actions and the player incurs the loss $c_t^\intercal x_t$. The player's regret is defined as

$$R(\mathcal{A}, c_{1:T}) \;=\; \sum_{t=1}^{T} c_t^\intercal x_t - \min_{x \in \mathcal{K}} \sum_{t=1}^{T} c_t^\intercal x.$$

In online linear optimization *with hints*, the player observes $v_t \in \mathbb{R}^d$ with $\|v_t\|_2 = 1$, before choosing x_t, with the guarantee that $v_t^\intercal c_t \geq \alpha \|c_t\|_2$, for some $\alpha > 0$.

We use *uniform convexity* to characterize the degree of convexity of the player's action set $\mathcal{K}$. Informally, uniform convexity requires that the convex combination of any two points x and y on the boundary of $\mathcal{K}$ be sufficiently far from the boundary. A formal definition is given below.

Definition 2.1 (Pisier [18])**.** *Let $\mathcal{K}$ be a convex set that is symmetric around the origin. $\mathcal{K}$ and the Banach space defined by $\mathcal{K}$ are said to be* uniformly convex *if for any $0 < \epsilon < 2$ there exists a $\delta > 0$ such that for any pair of points $x, y \in \mathcal{K}$ with $\|x\|_{\mathcal{K}} \leq 1$, $\|y\|_{\mathcal{K}} \leq 1$, $\|x - y\|_{\mathcal{K}} \geq \epsilon$, we have $\left\|\frac{x+y}{2}\right\|_{\mathcal{K}} \leq 1 - \delta$. The* modulus of uniform-convexity *$\delta_{\mathcal{K}}(\epsilon)$ is the best possible δ for that ϵ, i.e.,*

$$\delta_{\mathcal{K}}(\epsilon) = \inf\left\{1 - \left\|\frac{x+y}{2}\right\|_{\mathcal{K}} : \|x\|_{\mathcal{K}} \leq 1, \|y\|_{\mathcal{K}} \leq 1, \|x - y\|_{\mathcal{K}} \geq \epsilon\right\}.$$

For brevity, we say that $\mathcal{K}$ is (C, q)-uniformly convex when $\delta_{\mathcal{K}}(\epsilon) = C\epsilon^q$ and we omit C when it is clear from the context.

Examples of uniformly convex sets include L_p balls for any $1 < p < \infty$ with modulus of convexity $\delta_{L_p}(\epsilon) = C_p \epsilon^p$ for $p \geq 2$ and a constant C_p and $\delta_{L_p}(\epsilon) = (p-1)\epsilon^2$ for $1 < p \leq 2$. On the other hand, L_1 and L_∞ units balls are not uniformly convex. When $\delta_{\mathcal{K}}(\epsilon) \in \Theta(\epsilon^2)$, we say that $\mathcal{K}$ is *strongly convex*.

Another notion of convexity we use in this work is called *exp-concavity*. A function $f : \mathcal{K} \to \mathbb{R}$ is exp-concave with parameter $\beta > 0$, if $\exp(-\beta f(x))$ is a concave function of $x \in \mathcal{K}$. This is a weaker requirement than strong convexity when the gradient of f is uniformly bounded [14]. The next proposition shows that we can obtain regret bounds of order $\Theta(\log(T))$ in online convex optimization when the loss functions are exp-concave with parameter β.

Proposition 2.2 ([14])**.** *Consider online convex optimization on a sequence of loss functions $f_1, \ldots, f_T$ over a feasible set $\mathcal{K} \subseteq \mathbb{R}^d$, such that all t, $f_t : \mathcal{K} \to \mathbb{R}$ is exp-concave with parameter $\beta > 0$. There is an algorithm, with runtime polynomial in d, which we call $\mathcal{A}_{\mathrm{EXP}}$, with a regret that is at most $\frac{d}{\beta}(1 + \log(T+1))$.*

Throughout this work, we draw intuition from basic orthogonal geometry. Given any vector x and a hint v, we define $x^{\parallel v} = (x^\intercal v)v$ and $x^{\perp v} = x - (x^\intercal v)v$, as the parallel and the orthogonal components of x with respect to v. When the hint v is clear from the context we simply use $x^{\parallel}$ and $x^{\perp}$ to denote these vectors.

Naturally, our regret bounds involve a number of geometric parameters. Since the set of actions of the adversary $\mathcal{C}$ is compact, we can find $G \geq 0$ such that $\|c\|_2 \leq G$ for all $c \in \mathcal{C}$. When $\mathcal{K}$ is uniformly convex, we denote $\mathcal{K} = \{w \in \mathbb{R}^d \mid \|w\|_{\mathcal{K}} \leq 1\}$. In this case, since all norms are equivalent in finite dimension, there exist $R > 0$ and $r > 0$ such that $B_r \subseteq \mathcal{K} \subseteq B_R$, where B_r (resp. B_R) denote the L_2 unit ball centered at 0 with radius r (resp. R). This implies that $\frac{1}{R}\|\cdot\|_2 \leq \|\cdot\|_{\mathcal{K}} \leq \frac{1}{r}\|\cdot\|_2$.

3 Improved Regret Bounds for Strongly Convex $\mathcal{K}$

At first sight, it is not immediately clear how one should use the hint. Since v_t is guaranteed to satisfy $c_t^\intercal v_t \geq \alpha\|c_t\|_2$, moving the action x in the direction $-v_t$ always decreases the loss. One could hope to get the most benefit out of the hint by choosing x_t to be the extremal point in $\mathcal{K}$ in the direction $-v_t$. However, this naïve strategy could lead to a linear regret in the worst case. For example, say that $c_t = (1, \frac{1}{2})$ and $v_t = (0, 1)$ for all t and let $\mathcal{K}$ be the Euclidean unit ball. Choosing $x_t = -v_t$ would incur a loss of $-\frac{T}{2}$, while the best fixed action in hindsight, the point $(\frac{-2}{\sqrt{5}}, \frac{-1}{\sqrt{5}})$, would incur a loss of $\frac{-\sqrt{5}}{2}T$. The player's regret would therefore be $\frac{\sqrt{5}-1}{2}T$.

Intuitively, the flaw of this naïve strategy is that the hint does not give the player any information about the $(d-1)$-dimensional subspace orthogonal to v_t. Our solution is to use standard online learning machinery to learn how to act in this orthogonal subspace. Specifically, on round t, we use v_t to define the following *virtual loss function*:

$$\hat{c}_t(x) \;=\; \min_{w \in \mathcal{K}} \; c_t^\intercal w \quad \text{s.t.} \quad w^{\perp v_t} = x^{\perp v_t} \;\; .$$

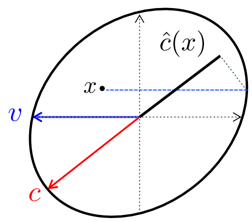

Figure 2: Virtual function $\hat{c}(\cdot)$.

In words, we consider the 1-dimensional subspace spanned by v_t and its $(d-1)$-dimensional orthogonal subspace separately. For any action $x \in \mathcal{K}$, we find another point, $w \in \mathcal{K}$, that equals x in the $(d-1)$-dimensional orthogonal subspace, but otherwise incurs the optimal loss. The value of the virtual loss $\hat{c}_t(x)$ is defined to be the value of the original loss function c_t at w. The virtual loss simulates the process of moving x as far as possible in the direction $-v_t$ without changing its value in any other direction (see Figure 2). This can be more formally seen by the following equation.

$$\operatorname*{arg\,min}_{w \in \mathcal{K}: w^\perp = \hat{x}^\perp} c_t^\intercal w = \operatorname*{arg\,min}_{w \in \mathcal{K}: w^\perp = \hat{x}^\perp} \left((c_t^\perp)^\intercal \hat{x}^\perp + (c_t^\parallel)^\intercal w^\parallel \right) = \operatorname*{arg\,min}_{w \in \mathcal{K}: w^\perp = \hat{x}^\perp} v_t^\intercal w, \tag{1}$$

where the last transition holds by the fact that $c_t^\parallel = \left\| c_t^\parallel \right\|_2 v_t$ since the hint is valid.

This provides an intuitive understanding of a measure of convexity of our virtual loss functions. When $\mathcal{K}$ is uniformly convex then the function $\hat{c}_t(\cdot)$ demonstrates convexity in the subspace orthogonal to v_t. To see that, note that for any x and y that lie in the space orthogonal to v_t, their mid point $\frac{x+y}{2}$ transforms to a point that is farther away in the direction of $-v_t$ than the midpoint of the transformations of x and y. As shown in Figure 3, the modulus of uniform convexity of $\mathcal{K}$ affects the degree of convexity of $\hat{c}_t(\cdot)$. We note, however, that $\hat{c}_t(\cdot)$ is not strongly convex in all directions. In fact, $\hat{c}_t(\cdot)$ is constant in the direction of v_t. Nevertheless, the properties shown here allude to the fact that $\hat{c}_t(\cdot)$ demonstrates some notion of convexity. As we show in the next lemma, this notion is indeed *exp-concavity*:

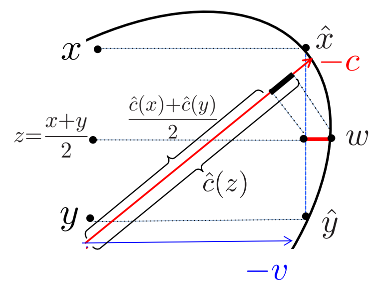

Figure 3: Uniform-convexity of the feasible set affects the convexity the virtual loss function.

Lemma 3.1. *If $\mathcal{K}$ is $(C, 2)$-uniformly convex, then $\hat{c}_t(\cdot)$ is $8\frac{\alpha \cdot C \cdot r}{G \cdot R^2}$-exp-concave.*

Proof. Let $\gamma = 8\frac{\alpha \cdot C \cdot r}{G \cdot R^2}$. Without loss of generality, we assume that $c_t \neq 0$, otherwise $\hat{c}_t(\cdot) = 0$ is a constant function and the proof follows immediately. Based on the above discussion, it is not hard to see that $\hat{c}_t(\cdot)$ is continuous (we prove this in more detail in the Appendix D.1). So, to prove that $\hat{c}_t(\cdot)$ is exp-concave, it is sufficient to show that

$$\exp\left(-\gamma \cdot \hat{c}_t\left(\frac{x+y}{2}\right)\right) \geq \frac{1}{2}\exp\left(-\gamma \cdot \hat{c}_t(x)\right) + \frac{1}{2}\exp\left(-\gamma \cdot \hat{c}_t(y)\right) \quad \forall (x,y) \in \mathcal{K}.$$

Consider $(x, y) \in \mathcal{K}$ and choose corresponding $(\hat{x}, \hat{y}) \in \mathcal{K}$ such that $\hat{c}_t(x) = c_t^\intercal \hat{x}$ and $\hat{c}_t(y) = c_t^\intercal \hat{y}$. Without loss of generality, we have $\|\hat{x}\|_\mathcal{K} = \|\hat{y}\|_\mathcal{K} = 1$, as we can always choose corresponding $\hat{x}, \hat{y}$ that are extreme points of $\mathcal{K}$. Since $\exp(-\gamma \hat{c}_t(\cdot))$ is decreasing in $\hat{c}_t(\cdot)$, we have

$$\exp\left(-\gamma \cdot \hat{c}_t\left(\frac{x+y}{2}\right)\right) = \max_{\substack{\|w\|_\mathcal{K} \leq 1 \\ w^{\perp v_t} = (\frac{x+y}{2})^{\perp v_t}}} \exp(-\gamma \cdot c_t^\intercal w). \tag{2}$$

Note that $w = \frac{\hat{x}+\hat{y}}{2} - \delta_\mathcal{K}(\|\hat{x}-\hat{y}\|_\mathcal{K}) \frac{v_t}{\|v_t\|_\mathcal{K}}$ satisfies $\|w\|_\mathcal{K} \leq 1$, since $\|w\|_\mathcal{K} \leq \left\|\frac{\hat{x}+\hat{y}}{2}\right\|_\mathcal{K} + \delta_\mathcal{K}(\|\hat{x}-\hat{y}\|_\mathcal{K}) \leq 1$ (see also Figure 3). Moreover, $w^{\perp v_t} = (\frac{x+y}{2})^{\perp v_t}$. So, by using this w in Equation (2), we have

$$\exp\left(-\gamma \cdot \hat{c}_t\left(\frac{x+y}{2}\right)\right) \geq \exp\left(-\frac{\gamma}{2} \cdot (c_t^\intercal \hat{x} + c_t^\intercal \hat{y}) + \gamma \cdot \frac{c_t^\intercal v_t}{\|v_t\|_\mathcal{K}} \cdot \delta_\mathcal{K}(\|\hat{x}-\hat{y}\|_\mathcal{K})\right). \tag{3}$$

On the other hand, since $\|v_t\|_{\mathcal{K}} \le \frac{1}{r}\|v_t\|_2 = \frac{1}{r}$ and $\|\hat{x}-\hat{y}\|_{\mathcal{K}} \ge \frac{1}{R}\|\hat{x}-\hat{y}\|_2$, we have

$$\begin{aligned}
\exp\left(\gamma \cdot \frac{c_t^\intercal v_t}{\|v_t\|_{\mathcal{K}}} \cdot \delta_{\mathcal{K}}(\|\hat{x}-\hat{y}\|_{\mathcal{K}})\right) &\ge \exp\left(\gamma \cdot r \cdot \alpha \cdot \|c_t\|_2 \cdot C \cdot \frac{1}{R^2} \cdot \|\hat{x}-\hat{y}\|_2^2\right) \\
&\ge \exp\left(\gamma \cdot \frac{\alpha \cdot C \cdot r}{R^2} \cdot \|c_t\|_2 \cdot \left(\frac{c_t^\intercal \hat{x}}{\|c_t\|_2} - \frac{c_t^\intercal \hat{y}}{\|c_t\|_2}\right)^2\right) \\
&\ge \exp\left(\frac{(\gamma/2)^2 \cdot (c_t^\intercal \hat{x} - c_t^\intercal \hat{y})^2}{2}\right) \\
&\ge \frac{1}{2} \cdot \exp\left(\frac{\gamma}{2} \cdot (c_t^\intercal \hat{x} - c_t^\intercal \hat{y})\right) + \frac{1}{2} \cdot \exp\left(\frac{\gamma}{2} \cdot (c_t^\intercal \hat{y} - c_t^\intercal \hat{x})\right),
\end{aligned}$$

where the penultimate inequality follows by the definition of γ and the last inequality is a consequence of the inequality $\exp(z^2/2) \ge \frac{1}{2}\exp(z) + \frac{1}{2}\exp(-z), \forall z \in \mathbb{R}$. Plugging the last inequality into (3) yields

$$\begin{aligned}
\exp\left(-\gamma \hat{c}_t(\tfrac{x+y}{2})\right) &\ge \frac{1}{2}\exp\left(-\frac{\gamma}{2}(c_t^\intercal \hat{x} + c_t^\intercal \hat{y})\right) \cdot \left\{\exp\left(\frac{\gamma}{2}(c_t^\intercal \hat{x} - c_t^\intercal \hat{y})\right) + \exp\left(\frac{\gamma}{2}(c_t^\intercal \hat{y} - c_t^\intercal \hat{x})\right)\right\} \\
&= \frac{1}{2}\exp\left(-\gamma \cdot c_t^\intercal \hat{y}\right) + \frac{1}{2}\exp\left(-\gamma \cdot c_t^\intercal \hat{x}\right) \\
&= \frac{1}{2}\exp\left(-\gamma \cdot \hat{c}_t(y)\right) + \frac{1}{2}\exp\left(-\gamma \cdot \hat{c}_t(x)\right),
\end{aligned}$$

which concludes the proof. □

Now, we use the sequence of virtual loss functions to reduce our problem to a standard online convex optimization problem (without hints). Namely, the player applies $\mathcal{A}_{\text{EXP}}$ (from Proposition 2.2), which is an online convex optimization algorithm known to have $O(\log(T))$ regret with respect to exp-concave functions, to the sequence of virtual loss functions. Then our algorithm takes the action $\hat{x}_t \in \mathcal{K}$ that is prescribed by $\mathcal{A}_{\text{EXP}}$ and moves it as far as possible in the direction of $-v_t$. This process is formalized in Algorithm 1.

Algorithm 1 $\mathcal{A}_{\text{hint}}$ FOR STRONGLY CONVEX $\mathcal{K}$

For $t = 1, \ldots, T$,

1. Use Algorithm $\mathcal{A}_{\text{EXP}}$ with the history $\hat{c}_\tau(\cdot)$ for $\tau < t$, and let $\hat{x}_t$ be the chosen action.
2. Let $x_t = \arg\min_{w \in \mathcal{K}}(v_t^\intercal w)$ s.t. $w^{\perp v_t} = \hat{x}_t^{\perp v_t}$. Play x_t and receive c_t as feedback.

Next, we show that the regret of algorithm $\mathcal{A}_{\text{EXP}}$ on the sequence of virtual loss functions is an upper bound on the regret of Algorithm 1.

Lemma 3.2. *For any sequence of loss functions $c_1, \ldots, c_T$, let $R(\mathcal{A}_{\text{hint}}, c_{1:T})$ be the regret of algorithm $\mathcal{A}_{\text{hint}}$ on the sequence $c_1, \ldots, c_T$, and $R(\mathcal{A}_{\text{EXP}}, \hat{c}_{1:T})$ be the regret of algorithm $\mathcal{A}_{\text{EXP}}$ on the sequence of virtual loss functions $\hat{c}_1, \ldots, \hat{c}_T$. Then, $R(\mathcal{A}_{\text{hint}}, c_{1:T}) \le R(\mathcal{A}_{\text{EXP}}, \hat{c}_{1:T})$.*

Proof. Equation (1) provides an equivalent definition $x_t = \arg\min_{w \in \mathcal{K}}(c_t^\intercal w)$ s.t. $w^{\perp v_t} = \hat{x}_t^{\perp v_t}$. Using this, we show that the loss of algorithm $\mathcal{A}_{\text{hint}}$ on the sequence $c_{1:T}$ is the same as the loss of algorithm $\mathcal{A}_{\text{EXP}}$ on the sequence $\hat{c}_{1:T}$.

$$\sum_{t=1}^T \hat{c}_t(\hat{x}_t) = \sum_{t=1}^T \min_{w \in \mathcal{K}: w^\perp = \hat{x}_t^\perp} c_t^\intercal w = \sum_{t=1}^T c_t^\intercal \Big(\arg\min_{w \in \mathcal{K}: w^\perp = \hat{x}_t^\perp} c_t^\intercal w\Big) = \sum_{t=1}^T c_t^\intercal x_t.$$

Next, we show that the offline optimal on the sequence $\hat{c}_{1:T}$ is more competitive that the offline optimal on the sequence $c_{1:T}$. First note that for any x and t, $\hat{c}_t(x) = \min_{w \in \mathcal{K}: w^\perp = x^\perp} c_t^\intercal w \le c_t^\intercal x$. Therefore, $\min_{x \in \mathcal{K}} \sum_{t=1}^T \hat{c}_t(x) \le \min_{x \in \mathcal{K}} \sum_{t=1}^T c_t^\intercal x$. The proof concludes by

$$R(\mathcal{A}_{\text{hint}}, c_{1:T}) = \sum_{t=1}^T c_t^\intercal x_t - \min_{x \in \mathcal{K}} \sum_{t=1}^T c_t^\intercal x \le \sum_{t=1}^T \hat{c}_t(\hat{x}_t) - \min_{x \in \mathcal{K}} \sum_{t=1}^T \hat{c}_t(x) = R(\mathcal{A}_{\text{EXP}}, \hat{c}_{1:T}).$$

□

Our main result follows from the application of Lemmas 3.1 and 3.2.

Theorem 3.3. *Suppose that $\mathcal{K} \subseteq \mathbb{R}^d$ is a $(C, 2)$-uniformly convex set that is symmetric around the origin, and $B_r \subseteq \mathcal{K} \subseteq B_R$ for some r and R. Consider online linear optimization with hints where the cost function at round t is $\|c_t\|_2 \leq G$ and the hint v_t is such that $c_t^\intercal v_t \geq \alpha \|c_t\|_2$, while $\|v_t\|_2 = 1$. Algorithm 1 in combination with $\mathcal{A}_{\mathrm{EXP}}$ has a worst-case regret of*

$$R(\mathcal{A}_{\mathrm{hint}}, c_{1:T}) \leq \frac{d \cdot G \cdot R^2}{8\alpha \cdot C \cdot r} \cdot (1 + \log(T + 1)).$$

Since $\mathcal{A}_{\mathrm{EXP}}$ requires the coefficient of exp-concavity to be given as an input, α needs to be known a priori to be able to use Algorithm 1. However, we can use a standard doubling trick to relax this requirement and derive the same asymptotic regret bound. We defer the presentation of this argument to Appendix B.

4 Improved Regret Bounds for (C, q)-Uniformly Convex $\mathcal{K}$

In this section, we consider any feasible set $\mathcal{K}$ that is (C, q)-uniformly convex for $q \geq 2$. Our results differ from the previous section in two aspects. First, our algorithm can be used with (C, q)-uniformly convex feasible sets for any $q \geq 2$ compared to the results of the previous section that only hold for strongly convex sets ($q = 2$). On the other hand, the approach in this section requires the hints to be restricted to a finite set of vectors $\mathcal{V}$. We show that when $\mathcal{K}$ is (C, q)-uniformly convex for $q > 2$, our regret is $O(T^{\frac{2-q}{1-q}})$. If $q \in (2, 3)$, this is an improvement over the worst case regret of $O(\sqrt{T})$ guaranteed in the absence of hints.

We first consider the scenario where the hint is always pointing in the same direction, i.e. $v_t = v$ for some v and all $t \in [T]$. In this case, we show how one can use a simple algorithm that picks the best performing action so far (a.k.a the Follow-The-Leader algorithm) to obtain improved regret bounds. We then consider the case where the hint belongs to a finite set $\mathcal{V}$. In this case, we instantiate one copy of the Follow-The-Leader algorithm for each $v \in \mathcal{V}$ and combine their outcomes in order to obtain improved regret bounds that depend on the cardinality of $\mathcal{V}$, which we denote by $|\mathcal{V}|$.

Lemma 4.1. *Suppose that $v_t = v$ for all $t = 1, \cdots, T$ and that $\mathcal{K}$ is (C, q)-uniformly convex that is symmetric around the origin, and $B_r \subseteq \mathcal{K} \subseteq B_R$ for some r and R. Consider the algorithm, called Follow-The-Leader (FTL), that at every round t, plays $x_t \in \arg\min_{x \in \mathcal{K}} \sum_{\tau < t} c_\tau^\intercal x$. If $\sum_{\tau=1}^t c_\tau^\intercal v \geq 0$ for all $t = 1, \cdots, T$, then the regret is bounded as follows,*

$$R(\mathcal{A}_{\mathrm{FTL}}, c_{1:T}) \leq \left(\frac{\|v\|_{\mathcal{K}} \cdot R^q}{2C} \right)^{1/(q-1)} \cdot \sum_{t=1}^{T} \left(\frac{\|c_t\|_2^q}{\sum_{\tau=1}^t c_\tau^\intercal v} \right)^{1/(q-1)}.$$

Furthermore, when v is a valid hint with margin α, i.e., $c_t^\intercal v \geq \alpha \cdot \|c_t\|_2$ for all $t = 1, \cdots, T$, the right-hand side can be further simplified to obtain the regret bound:

$$R(\mathcal{A}_{\mathrm{FTL}}, c_{1:T}) \leq \frac{1}{2\gamma} \cdot G \cdot (\ln(T) + 1) \qquad \text{if } q = 2$$

and

$$R(\mathcal{A}_{\mathrm{FTL}}, c_{1:T}) \leq \frac{1}{(2\gamma)^{1/(q-1)}} \cdot G \cdot \frac{q-1}{q-2} \cdot T^{\frac{q-2}{q-1}} \qquad \text{if } q > 2,$$

where $\gamma = \frac{C \cdot \alpha}{\|v\|_{\mathcal{K}} \cdot R^q}$.

Proof. We use a well-known inequality, known as FT(R)L Lemma (see e.g., [12, 17]), on the regret incurred by the FTL algorithm:

$$R(\mathcal{A}_{\mathrm{FTL}}, c_{1:T}) \leq \sum_{t=1}^{T} c_t^\intercal (x_t - x_{t+1}).$$

Without loss of generality, we can assume that $\|x_t\|_{\mathcal{K}} = \|x_{t+1}\|_{\mathcal{K}} = 1$ since the maximum of a linear function is attained at a boundary point. Since $\mathcal{K}$ is (C, q)-uniformly convex, we have

$$\left\| \frac{x_t + x_{t+1}}{2} \right\|_{\mathcal{K}} \leq 1 - \delta_{\mathcal{K}}(\|x_t - x_{t+1}\|_{\mathcal{K}}).$$

This implies that

$$\left\| \frac{x_t + x_{t+1}}{2} - \delta_{\mathcal{K}}(\|x_t - x_{t+1}\|_{\mathcal{K}}) \frac{v}{\|v\|_{\mathcal{K}}} \right\|_{\mathcal{K}} \leq 1.$$

Moreover, $x_{t+1} \in \arg\min_{x \in \mathcal{K}} x^\intercal \sum_{\tau=1}^{t} c_\tau$. So, we have

$$\left(\sum_{\tau=1}^{t} c_\tau\right)^\intercal \left(\frac{x_t + x_{t+1}}{2} - \delta_{\mathcal{K}}(\|x_t - x_{t+1}\|_{\mathcal{K}}) \frac{v}{\|v\|_{\mathcal{K}}}\right) \geq \inf_{x \in \mathcal{K}} x^\intercal \sum_{\tau=1}^{t} c_\tau = x_{t+1}^\intercal \sum_{\tau=1}^{t} c_\tau.$$

Rearranging this last inequality and using the fact that $\sum_{\tau=1}^{t} v^\intercal c_\tau \geq 0$, we obtain:

$$\left(\sum_{\tau=1}^{t} c_\tau\right)^\intercal \left(\frac{x_t - x_{t+1}}{2}\right) \geq \delta_{\mathcal{K}}(\|x_t - x_{t+1}\|_{\mathcal{K}}) \cdot \frac{\sum_{\tau=1}^{t} v^\intercal c_\tau}{\|v\|_{\mathcal{K}}} \geq \frac{C \cdot \|x_t - x_{t+1}\|_2^q}{\|v\|_{\mathcal{K}} \cdot R^q} \cdot \left(\sum_{\tau=1}^{t} v^\intercal c_\tau\right).$$

By definition of FTL, we have $x_t \in \arg\min_{x \in \mathcal{K}} x^\intercal \sum_{\tau=1}^{t-1} c_\tau$, which implies:

$$\left(\sum_{\tau=1}^{t-1} c_\tau\right)^\intercal \frac{x_{t+1} - x_t}{2} \geq 0.$$

Summing up the last two inequalities and setting $\gamma = \frac{C \cdot \alpha}{\|v\|_{\mathcal{K}} \cdot R^q}$, we derive:

$$c_t^\intercal \left(\frac{x_t - x_{t+1}}{2}\right) \geq \frac{\gamma}{\alpha} \cdot \left(\sum_{\tau=1}^{t} v^\intercal c_\tau\right) \cdot \|x_t - x_{t+1}\|_2^q \geq \frac{\gamma}{\alpha} \cdot \left(\sum_{\tau=1}^{t} v^\intercal c_\tau\right) \cdot \frac{(c_t^\intercal (x_t - x_{t+1}))^q}{\|c_t\|_2^q}.$$

Rearranging this last inequality and using the fact that $\sum_{\tau=1}^{t} v^\intercal c_\tau \geq 0$, we obtain:

$$|c_t^\intercal (x_t - x_{t+1})| \leq \frac{1}{(2\gamma/\alpha)^{1/(q-1)}} \cdot \left(\frac{\|c_t\|_2^q}{\sum_{\tau=1}^{t} v^\intercal c_\tau}\right)^{1/(q-1)}. \tag{4}$$

Summing (4) over all t completes the proof of the first claim. The regret bounds for when $v^\intercal c_t \geq \alpha \cdot \|c_t\|_2$ for all $t = 1, \cdots, T$ follow from the first regret bound. We defer this part of the proof to Appendix D.2. ☐

Note that the regret bounds become $O(T)$ when $q \to \infty$. This is expected because L_q balls are q-uniformly convex for $q \geq 2$ and converge to L_∞ balls as $q \to \infty$ and it is well-known that Follow-The-Leader yields $\Theta(T)$ regret in online linear optimization when $\mathcal{K}$ is a L_∞ ball.

Using the above lemma, we introduce an algorithm for online linear optimization with hints that belong to a set $\mathcal{V}$. In this algorithm, we instantiate one copy of the FTL algorithm for each possible direction of the hint. On round t, we invoke the copy of the algorithm that corresponds to the direction of the hint v_t, using the history of the game for rounds with hints in that direction. We show that the overall regret of this algorithm is no larger than the sum of the regrets of the individual copies.

Algorithm 2 $\mathcal{A}_{\text{set}}$: SET-OF-HINTS

For all $v \in \mathcal{V}$, let $T_v = \emptyset$.
For $t = 1, \ldots, T$,

1. Play $x_t \in \arg\min_{x \in \mathcal{K}} \sum_{\tau \in T_{v_t}} c_\tau^\intercal x$ and receive c_t as feedback.
2. Update $T_{v_t} \leftarrow T_{v_t} \cup \{t\}$.

Theorem 4.2. *Suppose that $\mathcal{K} \subseteq \mathbb{R}^d$ is a (C, q)-uniformly convex set that is symmetric around the origin, and $B_r \subseteq \mathcal{K} \subseteq B_R$ for some r and R. Consider online linear optimization with hints where the cost function at round t is $\|c_t\|_2 \leq G$ and the hint v_t comes from a finite set $\mathcal{V}$ and is such that $c_t^\intercal v_t \geq \alpha \|c_t\|_2$, while $\|v_t\|_2 = 1$. Algorithm 2 has a worst-case regret of*

$$R(\mathcal{A}_{\text{set}}, c_{1:T}) \leq |\mathcal{V}| \cdot \frac{R^2}{2C \cdot \alpha \cdot r} \cdot G \cdot (\ln(T) + 1), \qquad \textit{if } q = 2,$$

and

$$R(\mathcal{A}_{\text{set}}, c_{1:T}) \leq |\mathcal{V}| \cdot \left(\frac{R^q}{2C \cdot \alpha \cdot r}\right)^{1/(q-1)} \cdot G \cdot \frac{q-1}{q-2} \cdot T^{\frac{q-2}{q-1}} \qquad \textit{if } q > 2.$$

Proof. We decompose the regret as follows:

$$R(\mathcal{A}_{\mathrm{set}}, c_{1:T}) = \sum_{t=1}^{T} c_t^\intercal x_t - \inf_{x\in\mathcal{K}} \sum_{t=1}^{T} c_t^\intercal x \le \sum_{v\in\mathcal{V}} \left\{ \sum_{t\in T_v} c_t^\intercal x_t - \inf_{x\in\mathcal{K}} \sum_{t\in T_v} c_t^\intercal x \right\}$$
$$\le |\mathcal{V}| \cdot \max_{v\in\mathcal{V}} R(\mathcal{A}_{\mathrm{FTL}}, c_{T_v}).$$

The proof follows by applying Lemma 4.1 and by using $\|v_t\|_{\mathcal{K}} \le (1/r) \cdot \|v_t\|_2 = 1/r$.

□

Note that $\mathcal{A}_{\mathrm{set}}$ does not require α or $\mathcal{V}$ to be known a priori, as it can compile the set of hint directions as it sees new ones. Moreover, if the hints are not limited to finite set $\mathcal{V}$ a priori, then the algorithm can first discretize the L_2 unit ball with an $\alpha/2$-net and approximate any given hint with one of the hints in the discretized set. Using this discretization technique, Theorem 4.2 can be extended to the setting where the hints are not constrained to a finite set while having a regret that is linear in the size of the $\alpha/2$-net (exponential in the dimension d.) Extensions of Theorem 4.2 are discussed in more details in the Appendix C.

5 Lower Bounds

The regret bounds derived in Sections 3 and 4 suggest that the curvature of $\mathcal{K}$ can make up for the lack of curvature of the loss function to get rates faster than $O(\sqrt{T})$ in online convex optimization, provided we receive additional information about the next move of the adversary in the form of a hint. In this section, we show that the curvature of the player's decision set $\mathcal{K}$ is necessary to get rates better than $O(\sqrt{T})$, even in the presence of a hint.

As an example, consider the unit cube, i.e. $\mathcal{K} = \{x \mid \|x\|_\infty \le 1\}$. Note that this set is not uniformly convex. Since, the i^{th} coordinate of points in such a set, namely x_i, has no effect on the range of acceptable values for the other coordinates, revealing one coordinate does not give us any information about the other coordinates x_j for $j \ne i$. For example, suppose that c_t has each of its first two coordinates set to $+1$ or -1 with equal probability and all other coordinates set to 1. In this case, even after observing the last $d-2$ coordinates of the loss vector, the problem is reduced to a standard online linear optimization problem in the 2-dimensional unit cube. This choice of c_t is known to incur a regret of $\Omega(\sqrt{T})$ [1]. Therefore, online linear optimization with the set $\mathcal{K} = \{x \mid \|x\|_\infty \le 1\}$, even in the presence of hints, has a worst-case regret of $\Omega(\sqrt{T})$. As it turns out, this result holds for any polyhedral set of actions. We prove this by means of a reduction to the lower bounds established in [8] that apply to the online convex optimization framework (without hint). We defer the proof to the Appendix D.4.

Theorem 5.1. *If the set of feasible actions is a polyhedron then, depending on the set $\mathcal{C}$, either there exists a trivial algorithm that achieves zero regret or every online algorithm has worst-case regret $\Omega(\sqrt{T})$. This is true even if the adversary is restricted to pick a fixed hint $v_t = v$ for all $t = 1, \cdots, T$.*

At first sight, this result may come as a surprise. After all, since any L_p ball with $1 < p \le 2$ is strongly convex, one can hope to use a $L_{1+\nu}$ unit ball $\mathcal{K}'$ to approximate $\mathcal{K}$ when $\mathcal{K}$ is a L_1 ball (which is a polyhedron) and apply the results of Section 3 to achieve better regret bounds. The problem with this approach is that the constant in the modulus of convexity of $\mathcal{K}'$ deteriorates when $p \to 1$ since $\delta_{L_p}(\epsilon) = (p-1)\cdot\epsilon^2$, see [3]. As a result, the regret bound established in Theorem 3.3 becomes $O(\frac{1}{p-1} \cdot \log T)$. Since the best approximation of a L_1 unit ball using a L_p ball is of the form $\{x \in \mathbb{R}^d \mid d^{1-\frac{1}{p}}\|x\|_p \le 1\}$, the distance between the offline benchmark in the definition of regret when using $\mathcal{K}'$ instead of $\mathcal{K}$ can be as large as $(1 - d^{\frac{1}{p}-1}) \cdot T$, which translates into an additive term of order $(1 - d^{\frac{1}{p}-1}) \cdot T$ in the regret bound when using $\mathcal{K}'$ as a proxy for $\mathcal{K}$. Due to the inverse dependence of the regret bound obtained in Theorem 3.3 on $p-1$, the optimal choice of $p = 1 + \tilde{O}(\frac{1}{\sqrt{T}})$ leads to a regret of order $\tilde{O}(\sqrt{T})$.

Finally, we conclude with a result that suggests that $O(\log(T))$ is, in fact, the optimal achievable regret when $\mathcal{K}$ is strongly convex in online linear optimization with a hint. We defer the proof to the Appendix D.4.

Theorem 5.2. *If $\mathcal{K}$ is a L_2 ball then, depending on the set $\mathcal{C}$, either there exists a trivial algorithm that achieves zero regret or every online algorithm has worst-case regret $\Omega(\log(T))$. This is true even if the adversary is restricted to pick a fixed hint $v_t = v$ for all $t = 1, \cdots, T$.*

References

[1] J. Abernethy, P. L. Bartlett, A. Rakhlin, and A. Tewari. Optimal strategies and minimax lower bounds for online convex games. In *Proceedings of the 21st Conference on Learning Theory (COLT)*, pages 415–424, 2008.

[2] Sanjeev Arora, Elad Hazan, and Satyen Kale. The multiplicative weights update method: a meta-algorithm and applications. *Theory of Computing*, 8(1):121–164, 2012.

[3] Keith Ball, Eric A Carlen, and Elliott H Lieb. Sharp uniform convexity and smoothness inequalities for trace norms. *Inventiones mathematicae*, 115(1):463–482, 1994.

[4] Avrim Blum and Yishay Monsour. Learning, regret minimization, and equilibria. 2007.

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

[6] Chao-Kai Chiang and Chi-Jen Lu. Online learning with queries. In *Proceedings of the 21st Annual ACM-SIAM Symposium on Discrete Algorithms (SODA)*, pages 616–629, 2010.

[7] Chao-Kai Chiang, Tianbao Yang, Chia-Jung Lee, Mehrdad Mahdavi, Chi-Jen Lu, Rong Jin, and Shenghuo Zhu. Online optimization with gradual variations. In *Proceedings of the 25th Conference on Learning Theory (COLT)*, pages 6–1, 2012.

[8] A. Flajolet and P. Jaillet. No-regret learnability for piecewise linear losses. *arXiv preprint arXiv:1411.5649*, 2014.

[9] Yoav Freund. Boosting a weak learning algorithm by majority. *Information and computation*, 121(2): 256–285, 1995.

[10] Yoav Freund and Robert E Schapire. A decision-theoretic generalization of on-line learning and an application to boosting. *Journal of computer and system sciences*, 55(1):119–139, 1997.

[11] Elad Hazan. The convex optimization approach to regret minimization. *Optimization for machine learning*, pages 287–303, 2012.

[12] Elad Hazan and Satyen Kale. Extracting certainty from uncertainty: Regret bounded by variation in costs. In *Proceedings of the 23th Conference on Learning Theory (COLT)*, 2008.

[13] Elad Hazan and Nimrod Megiddo. Online learning with prior knowledge. In *Proceedings of the 20th Conference on Learning Theory (COLT)*, pages 499–513, 2007.

[14] Elad Hazan, Amit Agarwal, and Satyen Kale. Logarithmic regret algorithms for online convex optimization. *Machine Learning*, 69(2-3):169–192, 2007.

[15] R. Huang, T. Lattimore, A. György, and C. Szepesvári. Following the leader and fast rates in linear prediction: Curved constraint sets and other regularities. In *Proceedings of the 30th Annual Conference on Neural Information Processing Systems (NIPS)*, pages 4970–4978, 2016.

[16] Adam Kalai and Santosh Vempala. Efficient algorithms for online decision problems. *Journal of Computer and System Sciences*, 71(3):291–307, 2005.

[17] H Brendan McMahan. A survey of algorithms and analysis for adaptive online learning. *arXiv preprint arXiv:1403.3465*, 2014.

[18] Gilles Pisier. Martingales in banach spaces (in connection with type and cotype). *Manuscript., Course IHP, Feb*, pages 2–8, 2011.

[19] Alexander Rakhlin and Karthik Sridharan. Online learning with predictable sequences. In *Proceedings of the 25th Conference on Learning Theory (COLT)*, pages 993–1019, 2013.

[20] Alexander Rakhlin and Karthik Sridharan. Optimization, learning, and games with predictable sequences. In *Proceedings of the 27th Annual Conference on Neural Information Processing Systems (NIPS)*, pages 3066–3074, 2013.

[21] R. Vershynin. Introduction to the non-asymptotic analysis of random matrices. *arXiv preprint arXiv:1011.3027*, 2010.

[22] Martin Zinkevich. Online convex programming and generalized infinitesimal gradient ascent. In *Proceedings of the20thInternational Conference on Machine Learning (ICML)*, pages 928–936, 2003.

A Additional Related Works

The notion of hint introduced in this work is quite general and arises naturally in a variety of settings. Indeed, this notion generalizes some of the previous notions of predictability in online convex optimization.

Aside from the example, mentioned in Section 1.1, where a hint on the first coordinate of the loss vector is provided to the player, Hazan and Megiddo [13] also considered modeling the prior information in each round as a state space and measuring regret against a stronger benchmark that uses a mapping from the state space to the feasible set. Chiang and Lu [6] considered actively querying bits of the loss vector, but their result mainly improves the dependence of the regret on the dimension d.

Another notion of predictability is concerned with predictability of the entire loss vector rather than individual bits. Chiang et al. [7] considered online convex optimization with a sequence of loss functions that demonstrates a gradual change and they derived a regret bound in terms of the deviation of the loss functions. Rakhlin and Sridharan [19, 20] extended this line of work beyond sequences with gradual changes and showed that one can achieve an improved regret bound if the gradient of the loss function is predictable. They also applied this method to offline optimization problems such as Max Flow and to the problem of computing Nash equilibria in zero-sum games. In the latter case, they showed that when both players employ a variant of the Mirror Prox algorithm, they converge to the minimax equilibrium at a rate $O(\log(T))$. We compare our results to results derived in the literature on online convex optimization in more details below.

Comparison with [13] Hazan and Megiddo [13] considered as an example a setting where the player knows the first coordinate of the loss vector at all rounds, and showed that when $|c_{t1}| > 0$ and when the set of feasible actions is the Euclidean ball, one can achieve a regret of $O(d \cdot \log(T))$. Our work directly improves over this result, as in our setting a hint $v_t = \pm e_1$ also achieves $O(\log(T))$ regret, where the hidden factors are independent of d (see Theorem 4.2 with $\mathcal{V} = \{e_1, -e_1\}$). Moreover, we can deal with hints in different directions at different rounds and we allow for general uniformly convex action sets.

Connection with [19, 20] Suppose that we are provided with a vector $\tilde{c}_t$ at the beginning of time period t such that $\tilde{c}_t$ approximates c_t in the following sense: $\|c_t - \tilde{c}_t\|_2 \leq \sigma \|c_t\|_2$ for $\sigma \in [0, 1)$ (if $\sigma \geq 1$ this gives us essentially no information because $\tilde{c}_t = 0$ is a valid choice). Then $\frac{\tilde{c}_t}{\|\tilde{c}_t\|_2}$ (or 0 if $\tilde{c}_t = 0$) is a valid hint with margin $\alpha = (1 - \sigma)/(1 + \sigma)$. Indeed note that $\|\tilde{c}_t\|_2 \leq (1 + \sigma) \cdot \|c_t\|_2$ and $\|\tilde{c}_t\|_2 \geq (1 - \sigma) \cdot \|c_t\|_2$. Moreover:

$$\begin{aligned}
\frac{\tilde{c}_t^{\intercal} c_t}{\|\tilde{c}_t\|_2} &= \frac{1}{2} \cdot \frac{\|c_t\|_2^2 + \|\tilde{c}_t\|_2^2 - \|\tilde{c}_t - c_t\|_2^2}{\|\tilde{c}_t\|_2} \\
&\geq \frac{1}{2} \cdot (1 + (1 - \sigma)^2 - \sigma^2) \cdot \frac{\|c_t\|_2^2}{\|\tilde{c}_t\|_2} \\
&= (1 - \sigma) \cdot \frac{\|c_t\|_2}{\|\tilde{c}_t\|_2} \\
&\geq \frac{1 - \sigma}{1 + \sigma} \cdot \|c_t\|_2.
\end{aligned}$$

When $\mathcal{K}$ is strongly convex (e.g. a L_2 ball) and $\|c_t - \tilde{c}_t\|_2 \leq \sigma \|c_t\|_2$ for $\sigma \in [0, 1)$ for all time periods $t = 1, \cdots, T$, we get a regret bound of order $O((1 + \sigma)/(1 - \sigma) \cdot \log(T))$ which improves upon the bound $\sigma \cdot \sqrt{\sum_{t=1}^{T} \|c_t\|_2^2}$ obtained in [19]. However, the regret bounds that we get are not adaptive, i.e. we need to assume that $\|c_t - \tilde{c}_t\|_2 \leq \sigma \|c_t\|_2$ holds at all time periods to establish the regret bounds. This is in contrast with the analysis carried out in [19] where the regret bounds adapt to the sequence $(c_1, \cdots, c_T)$ at hand.

Comparison with [15] The regret bounds obtained for the Follow-The-Leader algorithm in Lemma 4.1 are stronger than the ones obtained in [15] for online linear optimization (without hints) in two ways. First, we consider general uniformly convex sets, as opposed to strongly convex sets in [15], which enables us to get intermediate rates that interpolate between $\log(T)$ and $\sqrt{T}$. Second, in the fully adversarial setting and when the set is strongly convex, their regret bound is guaranteed to be of order $O(\log(T))$ only if $0 \notin \text{conv}(\mathcal{C})$. Our condition is much weaker, in particular we get $O(\log(T))$ bound even if 0 is an extreme point of $\text{conv}(\mathcal{C})$. This generalization enables us to tackle the general setting of an arbitrary sequence of hints, which would not be possible with the analysis of [15].

B Doubling Trick when $\mathcal{K}$ is Strongly Convex and α is not Known a Priori.

We break down the horizon into phases where, during phase $i \in \mathbb{N}$, we run $\mathcal{A}_{\text{hint}}$ from scratch (discarding all previously observed values of the loss vectors c_t and the hints v_t) with exp-concavity parameter taken as $2^{i-1} \cdot \frac{8 \cdot C \cdot r}{G \cdot R^2}$. Phase $i+1$ begins at time t_{i+1} when the hint is no longer valid with margin $1/2^{i-1}$, i.e.:

$$t_{i+1} = \min\{\tau \geq t_i \mid c_\tau^\intercal v_\tau < (1/2^{i-1}) \cdot \|c_\tau\|_2\} + 1$$

(with the convention that $0/0 = 1$) and phase 1 begins at time $t_1 = 1$.

Lemma B.1. *When α is not known a priori, using the doubling trick yields a regret bound that is identical (up to a constant additive term) to the one we would obtain if we knew α a priori. Specifically, we have:*

$$R(\mathcal{A}_{\text{hint}}, c_{1:T}) \leq 2\log(\frac{1}{\alpha}) \cdot G \cdot R + \frac{d \cdot G \cdot R^2}{8\alpha \cdot C \cdot r} \cdot (1 + \log(1+T)).$$

Proof. Let N denote the number of phases and define $t_{N+1} = T+1$. Note that there are at most $N \leq \log(1/\alpha)$ phases since $\frac{c_t^\intercal v_t}{\|c_t\|_2} \in [\alpha, 1]$ for all time periods t by assumption. Observe that, for any phase i and for any time period $t = t_i, \cdots, t_{i+1}-2$, we have $\frac{c_t^\intercal v_t}{\|c_t\|_2} \geq 1/2^{i-1}$ so that $\hat{c}_t(\cdot)$ is $2^{i-1} \cdot \frac{8C \cdot r}{G \cdot R^2}$-exp-concave. Using Lemmas 3.1 and 3.2, we get:

$$\begin{aligned}
\sum_{t=1}^T c_t^\intercal x_t - \inf_{x \in \mathcal{K}} \sum_{t=1}^T c_t^\intercal x &\leq \sum_{i=1}^N \{ \sum_{t=t_i}^{t_{i+1}-1} c_t^\intercal x_t - \inf_{x \in \mathcal{K}} \sum_{t=t_i}^{t_{i+1}-1} c_t^\intercal x \} \\
&\leq 2N \cdot G \cdot R + \sum_{i=1}^N \{ \sum_{t=t_i}^{t_{i+1}-2} c_t^\intercal x_t - \inf_{x \in \mathcal{K}} \sum_{t=t_i}^{t_{i+1}-2} c_t^\intercal x \} \\
&\leq 2\log(\frac{1}{\alpha}) \cdot G \cdot R + \sum_{i=1}^N 2^{i-1} \cdot \frac{d \cdot G \cdot R^2}{8C \cdot r} \cdot (1 + \log(1 + t_{i+1} - t_i - 1)) \\
&\leq 2\log(\frac{1}{\alpha}) \cdot G \cdot R + 2^N \cdot \frac{d \cdot G \cdot R^2}{8C \cdot r} \cdot (1 + \log(1+T)) \\
&\leq 2\log(\frac{1}{\alpha}) \cdot G \cdot R + \frac{d \cdot G \cdot R^2}{8\alpha \cdot C \cdot r} \cdot (1 + \log(1+T)),
\end{aligned}$$

where we use $\|c_t\|_2 \leq G$ and $\|x_t\|_2 \leq R \cdot \|x_t\|_{\mathcal{K}} = R$ for the second inequality. This concludes the proof. □

C Extensions of Theorem 4.2.

Hints pointing in arbitrary directions When the directions of the hints are arbitrary, we can discretize the L_2 unit sphere using an $\alpha/2$-net (which contains at most $(1+\frac{4}{\alpha})^n$ points, see [21]), which we denote by $\tilde{\mathcal{V}}$. At any time $t = 1, \cdots, T$, the hint v_t is first mapped to its closest neighbor in $\mathcal{V}$, denoted by $\tilde{v}_t$, and we use $\mathcal{A}_{\text{set}}$ with $\mathcal{V} = \tilde{\mathcal{V}}$ and $\tilde{v}_t$ as "the" hint. We refer to this new algorithm as $\tilde{\mathcal{A}}_{\text{set}}$.

Theorem C.1. *Suppose that $\mathcal{K}$ is (C, q)-uniformly convex. If the hints come from a finite set $\mathcal{V}$, then $\tilde{\mathcal{A}}_{\text{set}}$ yields the same regret bounds as $\mathcal{A}_{\text{set}}$ up to a multiplicative factor 2. If the hints point in arbitrary directions then we have:*

$$R(\tilde{\mathcal{A}}_{\text{set}}, c_{1:T}) \leq (1+\frac{4}{\alpha})^d \cdot \frac{R^2}{C \cdot \alpha \cdot r} \cdot G \cdot (\ln(T)+1), \qquad \textit{if } q = 2,$$

and

$$R(\tilde{\mathcal{A}}_{\text{set}}, c_{1:T}) \leq (1+\frac{4}{\alpha})^d \cdot \left(\frac{R^q}{C\cdot\alpha\cdot r}\right)^{1/(q-1)} \cdot G \cdot \frac{q-1}{q-2} \cdot T^{\frac{q-2}{q-1}} \qquad \textit{if } q > 2.$$

Proof. Observe that, at any time period $t = 1, \cdots, T$, $\tilde{v}_t$ is a valid hint with margin $\alpha/2$. Indeed:

$$\begin{aligned} c_t^\intercal \tilde{v}_t &= c_t^\intercal v_t - c_t^\intercal (v_t - \tilde{v}_t) \\ &\geq \alpha \cdot \|c_t\|_2 - \|c_t\|_2 \cdot \|v_t - \tilde{v}_t\|_2 \\ &\geq \alpha/2 \cdot \|c_t\|_2 , \end{aligned}$$

by definition of $\tilde{\mathcal{V}}$. If the hints come from a finite set $\mathcal{V}$, the set of hints $\{\tilde{v}_t \mid t = 1, \cdots, T\}$ is also finite with cardinality at most $|\mathcal{V}|$ since the mapping $v_t \to \tilde{v}_t$ is independent of t. This shows the first part of the claim. If the hints point in arbitrary directions, then let, for each $v \in \tilde{\mathcal{V}}$, T_v be the subset of time periods such that v_t is mapped to v. We have:

$$\begin{aligned} R(\tilde{\mathcal{A}}_{\text{set}}, c_{1:T}) &= \sum_{t=1}^T c_t^\intercal x_t - \inf_{x\in\mathcal{K}} \sum_{t=1}^T c_t^\intercal x \\ &\leq \sum_{v\in\tilde{\mathcal{V}}} \{ \sum_{t\in T_v} c_t^\intercal x_t - \inf_{x\in\mathcal{K}} \sum_{t\in T_v} c_t^\intercal x \} \\ &\leq |\tilde{\mathcal{V}}| \cdot \max_{v\in\tilde{\mathcal{V}}} R(\mathcal{A}_{\text{FTL}}, c_{T_v}), \end{aligned}$$

which concludes the proof with Lemma 4.1. ☐

Random hints We consider an extension to a stochastic setting where the hint is not necessarily always valid at each round but rather in expectations. We next show that $\mathcal{A}_{\text{set}}$ yields regret bounds similar to the ones derived when the hint is always valid.

Theorem C.2. *Suppose that: (a) the hints come from a finite set $\mathcal{V}$, (b) $\mathcal{K}$ is (C, q)-uniformly convex, (c) $((c_t, v_t))_{t\in\mathbb{N}}$ is an independent stochastic process (but not necessarily i.i.d.), and (d) $\mathbb{E}[c_t^\intercal v_t \mid v_t = v] \geq \alpha \cdot \mathbb{E}[\|c_t\|_2 \mid v_t = v]$ for all $v \in \mathcal{V}$. Then, we have:*

$$\mathbb{E}[R(\mathcal{A}_{\text{set}}, c_{1:T})] \leq |\mathcal{V}| \cdot \frac{7R^2}{C\cdot\alpha\cdot r} \cdot G \cdot (\ln(T)+1) + O(1),$$

if $q = 2$ and

$$\mathbb{E}[R(\mathcal{A}_{\text{set}}, c_{1:T})] \leq |\mathcal{V}| \cdot \left(\frac{7R^q}{C\cdot\alpha\cdot r}\right)^{1/(q-1)} \cdot G \cdot \frac{q-1}{q-2} \cdot T^{\frac{q-2}{q-1}} + O(1),$$

if $q > 2$.

The proof is deferred to the Appendix D.3. Note that Theorem C.2 is a strict generalization of Theorem 4.2 since we allow the sequence $((c_t, v_t))_{t\in\mathbb{N}}$ not to be identically distributed.

To illustrate the applicability of Theorem C.2, we consider the setting studied in [15] where $((c_t))_{t\in\mathbb{N}}$ is an i.i.d. stochastic process with mean $\mu \neq 0$ and no hint is available. In this setting, we can take $v_t = \mu/\|\mu\|_2$ as the hint at any time period for the purpose of the analysis (but μ need not be known since we are using a single instance of FTL when the hint is the same at all time periods). Indeed, in this case, we have:

$$\begin{aligned} \mathbb{E}[c_t^\intercal v_t] &= \mathbb{E}[c_t]^\intercal \frac{\mu}{\|\mu\|_2} \\ &= \|\mu\|_2 \\ &\geq \frac{\|\mu\|_2}{G} \cdot \mathbb{E}[\|c_t\|_2] \end{aligned}$$

as long as $\|c_t\|_2 \leq G$. Hence, when the set $\mathcal{K}$ is a L_2 ball, we recover the regret bound $O(\frac{G^2}{\|\mu\|_2} \cdot \log(T))$ established in [15].

D Omitted Proofs

D.1 Omitted Proof for Lemma 3.1

Lemma D.1. *The function $\hat{c}_t(\cdot)$ is continuous for any $t \in \mathbb{N}$.*

Proof. Take a point $x \in \mathcal{K}$ and a sequence $x_n \in \mathcal{K} \to x$. For every n, there exists $w_n \in \mathcal{K}$ such that $w_n^\perp = x_n^\perp$ and $\hat{c}_t(x_n) = c_t^\intercal w_n$. Observe that the sequence $(c_t^\intercal w_n)_{n\in\mathbb{N}}$ is bounded since $\|w_n\|_{\mathcal{K}} \leq 1$. Hence, it is sufficient to show that $(c_t^\intercal w_n)_{n\in\mathbb{N}}$ has a unique limit point $\hat{c}_t(x)$. Consider a subsequence of $(c_t^\intercal w_n)_{n\in\mathbb{N}}$ that converges and, since $\mathcal{K}$ is compact, extract further a subsequence of $(w_n)_{n\in\mathbb{N}}$ that converges to some $w_\infty \in \mathcal{K}$. Without loss of generality, we continue to assume that these sequences are indexed by all $n \in \mathbb{N}$. Taking limits in $w_n^\perp = x_n^\perp$, we get $w_\infty^\perp = x^\perp$ (an orthogonal projection is a linear operator and thus is continuous in finite dimension). Consider $w \in \mathcal{K}$ such that $\|w\|_{\mathcal{K}} < 1$ and $w^\perp = x^\perp$. For n big enough we have $\left\|x^\perp - x_n^\perp\right\|_{\mathcal{K}} \leq 1 - \|w\|_{\mathcal{K}}$ and so $\left\|w + x_n^\perp - x^\perp\right\|_{\mathcal{K}} \leq 1$ and $(w + x_n^\perp - x^\perp)^\perp = x_n^\perp$. By definition of w_n, we get $c_t^\intercal w_n^\perp \leq c_t^\intercal w + c_t^\intercal(x_n^\perp - x^\perp)$. Taking limits, we get $c_t^\intercal w_\infty \leq c_t^\intercal w$ for all w such that $\|w\|_{\mathcal{K}} < 1$ and $w^\perp = x^\perp$. By continuity of linear functions this also holds for all $w \in \mathcal{K}$ such that $w^\perp = x^\perp$ and we get $\hat{c}_t(x) = c_t^\intercal w_\infty = \lim_{n\to\infty} c_t^\intercal w_n = \lim_{n\to\infty} \hat{c}_t(x_n)$. $\square$

D.2 Omitted Proofs for Lemma 4.1

As we showed in the main body of the paper, we have:

$$|c_t^\intercal(x_t - x_{t+1})| \leq \frac{1}{(2\gamma/\alpha)^{1/(q-1)}} \cdot \left(\frac{\|c_t\|_2^q}{\sum_{\tau=1}^t v^\intercal c_\tau}\right)^{1/(q-1)}.$$

In what follows, we further assume that v is a valid hint with margin α, i.e., $v^\intercal c_t \geq \alpha \cdot \|c_t\|_2$ for all $t = 1, \cdots, T$. We get:

$$|c_t^\intercal(x_t - x_{t+1})| \leq \frac{1}{(2\gamma)^{1/(q-1)}} \cdot \left(\frac{\|c_t\|_2^q}{\sum_{\tau=1}^t \|c_\tau\|_2}\right)^{1/(q-1)}.$$

Note that the right-hand side is finite even if $c_t = 0$. Plugging this last inequality back into the first regret bound, we derive:

$$\begin{aligned}
R(\mathcal{A}_{\mathrm{FTL}}, c_{1:T}) &\leq \frac{1}{(2\gamma)^{1/(q-1)}} \cdot \sum_{t=1}^T \left(\frac{\|c_t\|_2^q}{\sum_{\tau=1}^t \|c_\tau\|_2}\right)^{1/(q-1)} \\
&\leq \frac{1}{(2\gamma)^{1/(q-1)}} \cdot \sup_{(y_1,\cdots,y_T)\in[0,G]^T} \sum_{t=1}^T \left(\frac{y_t^q}{\sum_{\tau=1}^t y_\tau}\right)^{1/(q-1)}.
\end{aligned}$$

We prove below that for any $t = 1, \cdots, T$ and any fixed values $(y_1, \cdots, y_{t-1}, y_{t+1}, \cdots, y_T) \in [0, G]^{T-1}$, the function $y_t \to \sum_{n=1}^T \left(\frac{y_n^q}{\sum_{\tau=1}^n y_\tau}\right)^{1/(q-1)}$ is convex on $[0, G]$ and thus the maximum of this function is attained at an extreme point: either 0 or G. Repeating this process for $t = 1, \cdots, T$, we get:

$$\begin{aligned}
R(\mathcal{A}_{\mathrm{FTL}}, c_{1:T}) &\leq \frac{1}{(2\gamma)^{1/(q-1)}} \cdot \sup_{(y_1,\cdots,y_T)\in\{0,G\}^T} \sum_{t=1}^T \left(\frac{y_t^q}{\sum_{\tau=1}^t y_\tau}\right)^{1/(q-1)} \\
&\leq \frac{1}{(2\gamma)^{1/(q-1)}} \cdot \sup_{n=0,\cdots,T} \sum_{k=1}^n \left(\frac{G^q}{k \cdot G}\right)^{1/(q-1)} \\
&\leq \frac{1}{(2\gamma)^{1/(q-1)}} \cdot G \cdot \sum_{k=1}^T \frac{1}{k^{1/(q-1)}}.
\end{aligned}$$

This concludes the proof as $\sum_{k=1}^{T} \frac{1}{k^{1/(q-1)}} \leq \ln(T) + 1$ if $q = 2$ and $\sum_{k=1}^{T} \frac{1}{k^{1/(q-1)}} \leq \frac{q-1}{q-2} \cdot T^{\frac{q-2}{q-1}}$ if $q > 2$.

Let us now prove that, for any $t = 1, \cdots, T$ and any fixed values $(y_1, \cdots, y_{t-1}, y_{t+1}, \cdots, y_T) \in [0, G]^{T-1}$, the function $y_t \to \sum_{n=1}^{T} (\frac{y_n^q}{\sum_{\tau=1}^{n} y_\tau})^{1/(q-1)}$ is convex on $[0, G]$. Clearly, $y_t \to \sum_{n \neq t} (\frac{y_n^q}{\sum_{\tau=1}^{n} y_\tau})^{1/(q-1)}$ is convex on $[0, G]$ since y_t only appears in the denominator and $1/(q-1) \geq 0$. We use the shorthand $A = \sum_{\tau=1}^{t-1} y_\tau$. It remains to show that $\phi : y \to (\frac{y^q}{y+A})^{1/(q-1)}$ is convex. We have:

$$\phi^{''}(y) = \frac{q}{(q-1)^2} \cdot y^{q/(q-1)-2} \cdot A^2 \cdot (y + A)^{-1/(q-1)-2},$$

which is non-negative for $A \geq 0$ and $y > 0$ (for $y = 0$, we can directly show by hand that $\phi(\lambda \cdot 0 + (1-\lambda) \cdot z) \leq \lambda \cdot \phi(0) + (1-\lambda) \cdot \phi(z)$ for $\lambda \in [0, 1]$ and $z \geq 0$).

D.3 Proof of Theorem C.2

We will need the following concentration inequality.

Lemma D.2 (Chernoff-Hoeffding concentration inequality). *Let $(X_t)_{t=1,\cdots,T}$ be a sequence of jointly independent random variables in $[0, 1]$. We have, for any $\epsilon \in (0, 1)$:*

$$\mathbb{P}\left[\sum_{t=1}^{T} X_t > (1+\epsilon) \cdot \sum_{t=1}^{T} \mathbb{E}[X_t]\right] \leq \exp(-\frac{\epsilon^2}{3} \cdot \sum_{t=1}^{T} \mathbb{E}[X_t])$$

and

$$\mathbb{P}\left[\sum_{t=1}^{T} X_t < (1-\epsilon) \cdot \sum_{t=1}^{T} \mathbb{E}[X_t]\right] \leq \exp(-\frac{\epsilon^2}{2} \cdot \sum_{t=1}^{T} \mathbb{E}[X_t]).$$

To establish the claims of Theorem C.2, first observe that:

$$\begin{aligned}
\mathbb{E}[R(\mathcal{A}_{\text{set}}, c_{1:T})] &= \mathbb{E}[\sum_{t=1}^{T} c_t^\intercal x_t - \inf_{x \in \mathcal{K}} \sum_{t=1}^{T} c_t^\intercal x] \\
&\leq \sum_{v \in \mathcal{V}} \{\mathbb{E}[\sum_{t \in T_v} c_t^\intercal x_t - \inf_{x \in \mathcal{K}} \sum_{t \in T_v} c_t^\intercal x]\} \\
&\leq |\mathcal{V}| \cdot \max_{v \in \mathcal{V}} \mathbb{E}[R(\mathcal{A}_{\text{FTL}}, c_{T_v})].
\end{aligned}$$

Take $v \in \mathcal{V}$. In the same spirit as in Lemma 4.1, what remains to be done is to bound the regret incurred by Follow-The-Leader for any independent stochastic sequence $(c_t)_{t \in \mathbb{N}}$ such that $\mathbb{E}[c_t^\intercal v] \geq \alpha \cdot \mathbb{E}[\|c_t\|_2]$ for all $t = 1, \cdots, T$. We start with the same standard inequality on the regret incurred by Follow-The-Leader:

$$\mathbb{E}[R(\mathcal{A}_{\text{FTL}}, c_{1:T})] \leq \mathbb{E}[\sum_{t=1}^{T} c_t^\intercal (x_t - x_{t+1})]. \tag{5}$$

Without loss of generality, we can assume that $\|x_t\|_{\mathcal{K}} = \|x_{t+1}\|_{\mathcal{K}} = 1$. Hence, we have:

$$\left\|\frac{x_t + x_{t+1}}{2}\right\|_{\mathcal{K}} \leq 1 - \delta_{\mathcal{K}}(\|x_t - x_{t+1}\|_{\mathcal{K}}),$$

which implies that

$$\left\|\frac{x_t + x_{t+1}}{2} - \frac{6}{7} \cdot \delta_{\mathcal{K}}(\|x_t - x_{t+1}\|_{\mathcal{K}}) \cdot rv\right\|_{\mathcal{K}} \leq 1,$$

since $\|v\|_{\mathcal{K}} \leq 1/r \cdot \|v\|_2 = 1/r$. Since moreover we have $x_{t+1} \in \arg\min_{x \in \mathcal{K}} x^\intercal \sum_{\tau=1}^{t} c_\tau$, we derive, along the same lines as in Lemma 4.1, that:

$$\left(\sum_{\tau=1}^{t} c_\tau\right)^\intercal \frac{x_t - x_{t+1}}{2} \geq \frac{6}{7} \cdot \delta_{\mathcal{K}}(\|x_t - x_{t+1}\|_{\mathcal{K}}) \cdot r \cdot \sum_{\tau=1}^{t} v^\intercal c_\tau \tag{6}$$

Moreover, since $x_t \in \arg\min_{x\in\mathcal{K}} x^\intercal \sum_{\tau=1}^{t-1} c_\tau$, we have:

$$\left(\sum_{\tau=1}^{t-1} c_\tau\right)^\intercal \frac{x_{t+1}-x_t}{2} \geq 0.$$

Summing these two inequalities yields:

$$c_t^\intercal \frac{x_t - x_{t+1}}{2} \geq \frac{6}{7} \cdot \delta_{\mathcal{K}}(\|x_t - x_{t+1}\|_{\mathcal{K}}) \cdot r \cdot \sum_{\tau=1}^{t} v^\intercal c_\tau.$$

Similarly, since:

$$\left\| \frac{x_t + x_{t+1}}{2} \pm \delta_{\mathcal{K}}(\|x_t - x_{t+1}\|_{\mathcal{K}}) \frac{c_t}{\|c_t\|_{\mathcal{K}}} \right\|_{\mathcal{K}} \leq 1,$$

$x_{t+1} \in \arg\min_{x\in\mathcal{K}} x^\intercal \sum_{\tau=1}^{t} c_\tau$, and $x_t \in \arg\min_{x\in\mathcal{K}} x^\intercal \sum_{\tau=1}^{t-1} c_\tau$, we get:

$$(\sum_{\tau=1}^{t} c_\tau)^\intercal \frac{x_t - x_{t+1}}{2} \geq \delta_{\mathcal{K}}(\|x_t - x_{t+1}\|_{\mathcal{K}}) \cdot \frac{\sum_{\tau=1}^{t} c_t^\intercal c_\tau}{\|c_t\|_{\mathcal{K}}}$$

and

$$\left(\sum_{\tau=1}^{t-1} c_\tau\right)^\intercal \frac{x_{t+1} - x_t}{2} \geq \delta_{\mathcal{K}}(\|x_t - x_{t+1}\|_{\mathcal{K}}) \cdot \frac{\sum_{\tau=1}^{t-1} -c_t^\intercal c_\tau}{\|c_t\|_{\mathcal{K}}}.$$

Summing these last two inequalities yields:

$$\begin{aligned} c_t^\intercal \frac{x_t - x_{t+1}}{2} &\geq \delta_{\mathcal{K}}(\|x_t - x_{t+1}\|_{\mathcal{K}}) \cdot \frac{\|c_t\|_2^2}{\|c_t\|_{\mathcal{K}}} \\ &\geq \delta_{\mathcal{K}}(\|x_t - x_{t+1}\|_{\mathcal{K}}) \cdot r \cdot \|c_t\|_2 , \end{aligned} \tag{7}$$

since $\|c_t\|_{\mathcal{K}} \leq 1/r \cdot \|c_t\|_2$. Define the events:

$$A_t = \{\sum_{\tau=1}^{t-1} v^\intercal c_\tau < \frac{1}{2} \cdot \sum_{\tau=1}^{t-1} \mathbb{E}[v^\intercal c_\tau]\}$$

and

$$B_t = \{\sum_{\tau=1}^{t-1} \|c_\tau\|_2 > \frac{3}{2} \cdot \sum_{\tau=1}^{t-1} \mathbb{E}[\|c_\tau\|_2]\}.$$

Using Lemma D.2, we have:

$$\begin{aligned} \mathbb{P}[A_t] &\leq \exp(-\frac{1}{8G} \cdot \sum_{\tau=1}^{t-1} \mathbb{E}[v^\intercal c_\tau]) \\ &\leq \exp(-\frac{\alpha}{16G} \cdot \sum_{\tau=1}^{t-1} \mathbb{E}[\|c_\tau\|_2]), \end{aligned}$$

and

$$\mathbb{P}[B_t] \leq \exp(-\frac{1}{12G} \cdot \sum_{\tau=1}^{t-1} \mathbb{E}[\|c_\tau\|_2]),$$

since $|v^\intercal c_t| \leq \|c_t\|_2 \leq G$. Note that, conditioned on the events $A_t^\complement$ and $B_t^\complement$, summing inequalities (6) and (7) yields:

$$\begin{aligned} c_t^\intercal (x_{t+1} - x_t) &\geq \delta_{\mathcal{K}}(\|x_t - x_{t+1}\|_{\mathcal{K}}) \cdot r \cdot \left(\|c_t\|_2 + 6/7 v^\intercal c_t + 6/7 \sum_{\tau=1}^{t-1} v^\intercal c_\tau\right) \\ &\geq \delta_{\mathcal{K}}(\|x_t - x_{t+1}\|_{\mathcal{K}}) \cdot r \cdot \left(1/7 \|c_t\|_2 + \alpha/7 \sum_{\tau=1}^{t-1} \|c_\tau\|_2\right) \end{aligned}$$

$$\geq \frac{C \cdot r \cdot \alpha}{7 \cdot R^q} \cdot \left(\sum_{\tau=1}^{t} \|c_\tau\|_2 \right) \cdot \|x_t - x_{t+1}\|_2^q.$$

Following the same steps as in Lemma 4.1, we conclude that:

$$|c_t^\intercal (x_t - x_{t+1})| \leq \frac{1}{(\gamma)^{1/(q-1)}} \cdot \left(\frac{\|c_t\|_2^q}{\sum_{\tau=1}^t \|c_\tau\|_2} \right)^{1/(q-1)},$$

with $\gamma = \frac{C \cdot r \cdot \alpha}{7R^q}$. Plugging this last inequality back into (5), we get:

$$\begin{aligned}
\mathbb{E}[R(\mathcal{A}^{\mathrm{FTL}}, c_{1:T})] &\leq \sum_{t=1}^T \mathbb{E}\left[\|c_t\|_2 \cdot (\|x_t\|_2 + \|x_{t+1}\|_2) \cdot 1_{A_t \cup B_t}\right] \\
&\quad + \frac{1}{(\gamma)^{1/(q-1)}} \cdot \sum_{t=1}^T \mathbb{E}\left[\left(\frac{\|c_t\|_2^q}{\sum_{\tau=1}^t \|c_\tau\|_2} \right)^{1/(q-1)} \cdot 1_{A_t^\complement \cap B_t^\complement} \right] \\
&\leq 2R \cdot \sum_{t=1}^T \mathbb{E}[\|c_t\|_2 \cdot 1_{A_t \cup B_t}] \\
&\quad + \frac{1}{(\gamma)^{1/(q-1)}} \cdot \sup_{p_1, \cdots, p_T \in \mathcal{P}(\mathcal{C})} \mathbb{E}\left[\sum_{t=1}^T \left(\frac{\|c_t\|_2^q}{\sum_{\tau=1}^t \|c_\tau\|_2} \right)^{1/(q-1)} \right] \\
&= 2R \cdot \sum_{t=1}^T \mathbb{E}[\|c_t\|_2] \cdot (\mathbb{P}[A_t] + \mathbb{P}[B_t]) \\
&\quad + \frac{1}{(\gamma)^{1/(q-1)}} \cdot \sup_{y_1, \cdots, y_T \in [0,G]} \sum_{t=1}^T \left(\frac{y_t^q}{\sum_{\tau=1}^t y_\tau} \right)^{1/(q-1)},
\end{aligned}$$

where $\mathcal{P}(\mathcal{C})$ denotes the set of probability distributions on $\mathcal{C}$ and where we use $\|x_t\|_2 \leq R \cdot \|x_t\|_{\mathcal{K}} \leq R$ and $\|x_{t+1}\|_2 \leq R \cdot \|x_{t+1}\|_{\mathcal{K}} \leq R$ since $x_t, x_{t+1} \in \mathcal{K}$. The last equality is obtained by independence of c_t and $(c_1, \cdots, c_{t-1})$. The second term is bounded in the proof of Lemma 4.1. As for the first term, we have:

$$\begin{aligned}
&\sum_{t=1}^T \mathbb{E}[\|c_t\|_2] \cdot (\mathbb{P}[A_t] + \mathbb{P}[B_t]) \\
&\qquad \leq 2 \sum_{t=1}^T \mathbb{E}[\|c_t\|_2] \cdot \exp(-\frac{\alpha}{16G} \cdot \sum_{\tau=1}^{t-1} \mathbb{E}[\|c_\tau\|_2]) \\
&\qquad \leq 2 \sup_{y_1, \cdots, y_T \in [0,G]} \sum_{t=1}^T y_t \cdot \exp(-\frac{\alpha}{16G} \cdot \sum_{\tau=1}^{t-1} y_\tau).
\end{aligned}$$

Observe that, for any $t = 1, \cdots, T$ and for any fixed values $(y_1, \cdots, y_{t-1}, y_{t+1}, \cdots, y_T) \in [0, G]^{T-1}$, the function $y_t \to \sum_{n=1}^T y_n \cdot \exp(-\frac{\alpha}{16G} \cdot \sum_{\tau=1}^{n-1} y_\tau)$ is convex on $[0, G]$. Hence, its maximum is attained at an extreme point: either 0 or G. Repeating this process for $t = 1, \cdots, T$, we get:

$$\begin{aligned}
\sum_{t=1}^T \mathbb{E}[\|c_t\|_2] \cdot (\mathbb{P}[A_t] + \mathbb{P}[B_t]) &\leq 2 \sup_{y_1, \cdots, y_T \in \{0,G\}} \sum_{t=1}^T y_t \cdot \exp(-\frac{\alpha}{16G} \cdot \sum_{\tau=1}^{t-1} y_\tau) \\
&\leq 2G \cdot \sup_{n=0, \cdots, T} \sum_{k=0}^n \exp(-\frac{\alpha}{16} \cdot k) \\
&\leq \frac{2G}{1 - \exp(-\frac{\alpha}{16})} = O(1),
\end{aligned}$$

which concludes the proof.

D.4 Derivation of the Lower Bounds

For any given $\alpha \in (0, 1]$, we establish, depending on the curvature of $\mathcal{K}$, lower bounds on regret when the opponent adversarially chooses the hints $(v_1, \cdots, v_T)$ as well as the cost vectors $(c_1, \cdots, c_T) \in \mathcal{C}$. In fact, we establish the regret bounds in the case of a weaker adversary who has to pick a fixed hint v initially and to stick to it throughout the game (i.e. $v_t = v$ for all $t = 1, \cdots, T$). Since this is a weaker notion of adversary, the lower bounds carry over to the more adversarial setting where the adversary is free to pick a different hint at every time period. The minimax regret that can be achieved by an online algorithm in this setting is expressed as:

$$\mathcal{R}_T(\mathcal{C}, \mathcal{K}) = \sup_{v \in B_2} \inf_{x_1 \in \mathcal{K}} \sup_{c_1 \in \mathcal{C}:\ c_1^\intercal v \geq \alpha \cdot \|c_1\|_2} \inf_{x_T \in \mathcal{K}} \sup_{c_T \in \mathcal{C}:\ c_T^\intercal v \geq \alpha \cdot \|c_T\|_2} \left[\sum_{t=1}^{T} c_t^\intercal x_t - \inf_{x \in \mathcal{K}} \sum_{t=1}^{T} c_t^\intercal x \right],$$

where B_2 denotes the unit ball for the L_2 norm. Observe that:

$$\mathcal{R}_T(\mathcal{C}, \mathcal{K}) = \sup_{v \in B_2} \Phi(v),$$

where $\Phi(v)$ is the minimax regret that can be achieved by an online algorithm in online linear optimization without hints when the cost vectors c_t all lie in $\mathcal{C} \cap \{c \in \mathbb{R}^d \mid c^\intercal v \geq \alpha \cdot \|c\|_2\}$ and $x_t \in \mathcal{K}$, i.e:

$$\Phi(v) = \inf_{x_1 \in \mathcal{K}} \sup_{c_1 \in \mathcal{C}:\ c_1^\intercal v \geq \alpha \cdot \|c_1\|_2} \inf_{x_T \in \mathcal{K}} \sup_{c_T \in \mathcal{C}:\ c_T^\intercal v \geq \alpha \cdot \|c_T\|_2} \left[\sum_{t=1}^{T} c_t^\intercal x_t - \inf_{x \in \mathcal{K}} \sum_{t=1}^{T} c_t^\intercal x \right].$$

Given this characterization, all the lower bounds established in this section are derived by means of a reduction to the lower bounds established in [8] in the standard online linear optimization framework (without hints). Following Flajolet and Jaillet [8], we are first led to identify trivial settings where there is an obvious algorithm that achieves zero regret.

Definition D.3. *The "game with hints" is said to be trivial if and only if, for all hints $v \in B_2$, there exists $x(v) \in \mathcal{K}$ such that $c^\intercal x(v) = \min_{x \in \mathcal{K}} c^\intercal x$ for all $c \in \mathcal{C}$ that satisfy $c^\intercal v \geq \alpha \cdot \|c\|_2$.*

If the game with hints is trivial then the optimal strategy is to play $x_t = x(v_t)$ at any time period $t \in \mathbb{N}$ in order to get zero regret. As it turns out, this uniquely identifies trivial games with hints, as we next show.

Lemma D.4. *For any $T \in \mathbb{N}$, $\mathcal{R}_T(\mathcal{C}, \mathcal{K}) \geq 0$. Moreover, $\mathcal{R}_T(\mathcal{C}, \mathcal{K}) = 0$ if and only if the game with hints is trivial.*

Proof. For the first part, observe that either $\mathcal{C} = \{0\}$ in which case $\mathcal{R}_T(\mathcal{C}, \mathcal{K}) = 0$ or we can find $c \neq 0$ in $\mathcal{C}$ in which case the opponent can pick $v = \frac{c}{\|c\|_2}$ and $c_t = c$ at any time period t which yields:

$$\mathcal{R}_T(\mathcal{C}, \mathcal{K}) \geq \inf_{x_1 \in \mathcal{K}} \cdots \inf_{x_T \in \mathcal{K}} \left[\sum_{t=1}^{T} c^\intercal x_t - \inf_{x \in \mathcal{K}} \sum_{t=1}^{T} c^\intercal x \right] = T \cdot \inf_{x \in \mathcal{K}} c^\intercal x - T \cdot \inf_{x \in \mathcal{K}} c^\intercal x = 0.$$

For the second part, observe that if the game with hints is trivial, we can play $x_t = x(v_t)$ at any time period $t \in \mathbb{N}$ and the regret incurred is $\sum_{t=1}^{T} \inf_{x \in \mathcal{K}} c_t^\intercal x - \inf_{x \in \mathcal{K}} \sum_{t=1}^{T} c_t^\intercal x \leq 0$, which in combination with $\mathcal{R}_T(\mathcal{C}, \mathcal{K}) \geq 0$ shows that $\mathcal{R}_T(\mathcal{C}, \mathcal{K}) = 0$. Conversely, suppose that $\mathcal{R}_T(\mathcal{C}, \mathcal{K}) = 0$. Consider any $v \in B_2$. Using Lemma 3 of [8] for online linear optimization when the adversary's decision set is $\mathcal{Z} = \mathcal{C} \cap \{c \in \mathbb{R}^d \mid c^\intercal v \geq \alpha \cdot \|c\|_2\}$ and the player's decision set is $\mathcal{F} = \mathcal{K}$ (in their notations), we get that $\Phi(v) \geq 0$ and that $\Phi(v) = 0$ if and only if there exists a trivial algorithm for this online linear optimization problem with zero regret, i.e. if and only if there exists $x(v) \in \mathcal{K}$ such that $c^\intercal x(v) = \min_{x \in \mathcal{K}} c^\intercal x$ for all $c \in \mathcal{C}$ satisfying $c^\intercal v \geq \alpha \cdot \|c\|_2$. Since $0 = \mathcal{R}_T(\mathcal{C}, \mathcal{K}) = \sup_{v \in B_2} \Phi(v)$, this implies that the game with hints is trivial. □

Flajolet and Jaillet [8] show that it is essential for $\mathcal{K}$ to be sufficiently curved to get regret bounds better than $\sqrt{T}$ in online linear optimization. We extend this characterization and show that this is true even in the presence of hints. To establish the regret bound, the goal is to find a hint v^* such that the online linear optimization problem where $c_t \in \mathcal{C} \cap \{c \in \mathbb{R}^d \mid c^\intercal v^* \geq \alpha \cdot \|c\|_2\}$ and $x_t \in \mathcal{K}$ is non-trivial, in the sense that there is no single point x^* in $\mathcal{K}$ that belongs to $\arg\min_{x \in \mathcal{K}} c^\intercal x$ uniformly for all $c \in \mathcal{C} \cap \{c \in \mathbb{R}^d \mid c^\intercal v^* \geq \alpha \cdot \|c\|_2\}$. If we can find such a hint v^*, then Theorem 2 of [8] shows that $\Phi(v^*) = \Omega(\sqrt{T})$ and thus $\mathcal{R}_T(\mathcal{C}, \mathcal{K}) = \sup_{v \in B_2} \Phi(v) = \Omega(\sqrt{T})$.

Theorem D.5. *Suppose that $\mathcal{K}$ is a polyhedron, then either the game with hints is trivial or $\mathcal{R}_T(\mathcal{C}, \mathcal{K}) = \Omega(\sqrt{T})$.*

Proof. Suppose that the game with hints is not trivial. Then, there exists a hint $v^* \in B_2$ such that for all $x \in \mathcal{K}$ there exists $c \in \mathcal{C}$ such that $c^\intercal v^* \geq \alpha \cdot \|c\|_2$ and $c^\intercal x > \min_{x \in \mathcal{K}} c^\intercal x$. As a result, and borrowing the notations of Lemma D.4, $\Phi(v^*) = \Omega(\sqrt{T})$ follows from Theorem 2 of [8] applied to the online linear optimization problem where the adversary's decision set is $\mathcal{Z} = \mathcal{C} \cap \{c \in \mathbb{R}^d \mid c^\intercal v^* \geq \alpha \cdot \|c\|_2\}$ and the player's decision set is $\mathcal{F} = \mathcal{K}$ (in their notations). Since $\mathcal{R}_T(\mathcal{C}, \mathcal{K}) = \sup_{v \in B_2} \Phi(v)$, this concludes the proof. □

This shows that $o(\sqrt{T})$ regret bounds are not possible when $\mathcal{K}$ is not curved. It is a priori unclear however whether $\log(T)$ is the optimal growth rate when $\mathcal{K}$ is strongly convex. We show that this is indeed the case by means of a reduction to a standard online linear optimization problem where $\mathcal{K}$ is a L_2 ball and 0 does not lie in the convex hull of the adversary's decision set.

Theorem D.6. *If $\mathcal{K}$ is a L_2 ball, then either the game with hints is trivial or $\mathcal{R}_T(\mathcal{C}, \mathcal{K}) = \Omega(\log(T))$.*

Proof. Suppose that the game with hints is not trivial. There exists a hint $v^* \in B_2$ such that for all $x \in \mathcal{K}$, there exists $c \in \mathcal{C}$ such that $c^\intercal v^* \geq \alpha \cdot \|c\|_2$ and $c^\intercal x > \min_{x \in \mathcal{K}} c^\intercal x$. In other words, for all $x \in \mathcal{K}$, we can find a cost vector c that is valid with respect to the hint v^* and such that x is not the optimal solution for c. Let x_1, x_2 be two non-co-linear points in $\mathcal{K}$ and let c_1 as a c_2 be the corresponding valid cost vectors. Observe that, necessarily, c_1 and c_2 are non-co-linear and we have $0 \notin [c_1, c_2]$ since $(\lambda c_1 + (1-\lambda) c_2)^\intercal v^* \geq \alpha \cdot (\lambda \|c_1\|_2 + (1-\lambda)\|c_2\|_2) \geq \alpha \cdot \min(\|c_1\|_2, \|c_2\|_2) > 0$ for $\lambda \in [0, 1]$. Then $\Phi(v^*) = \Omega(\log(T))$ follows from Theorem 5 of [8] applied to the online linear optimization problem where the adversary's decision set is $\mathcal{Z} = \{c_1, c_2\}$ and the player's decision set is $\mathcal{F} = \mathcal{K}$ (in their notations). Note that the assumptions of Theorem 5 of [8] are satisfied since, as shown above, $0 \notin \mathrm{conv}(\mathcal{Z})$. Since $\mathcal{R}_T(\mathcal{C}, \mathcal{K}) = \sup_{v \in B_2} \Phi(v)$ this concludes the proof. □