Appendix A More Preliminaries

A.1 Convexity and Lipschitz Continuity

Let $X \subseteq \mathbb{R}^d$ be a convex set, that is, for any $x, y \in X$ and any $\lambda \in [0,1]$, $\lambda x + (1-\lambda)y \in X$. We say $f : X \to \mathbb{R}$ is a convex function if for any $\lambda \in [0,1]$ and for any $x, y \in X$

$$\lambda f(x) + (1-\lambda)f(y) \geq f(\lambda x + (1-\lambda)y).$$

An equivalent definition of convexity is the following [21]. f is convex if and only if

$$f(x) \geq f(y) + \nabla f(y)^\top (x-y) \quad \forall x, y \in X.$$

Here $\nabla f(y)$ denotes any element in the subdifferential of f at y.
We say $f : X \to \mathbb{R}$ is strongly convex with parameter $\beta > 0$ if and only if

$$f(x) \geq f(y) + \nabla f(y)^\top (x-y) + \frac{\beta}{2}||x-y||^2 \quad \forall x, y \in X.$$

We say f is G-Lipschitz continuous with respect to a norm $||\cdot||$ if for every $x, y \in X$, $|f(x) - f(y)| \leq G||x-y||$.

Lemma 5. *[27] [Ch. 2]Let $f : X \to \mathbb{R}$ be a convex function. Then, f is G-Lipschitz over X with respect to a norm $||\cdot||$ if and only if for all $x \in X$ and for all $\nabla f(x) \in \partial f(x)$ we have that $||\nabla f(x)||_* \leq G$, where $||\cdot||_*$ denotes the dual norm.*

Throughout this paper, whenever we say f is G-Lipschitz we mean f is G-Lipschitz with respect to $||\cdot||_2$ unless otherwise stated.

A.2 From OCO to to Bandit Feedback

We present a result from that allows us to transform regret bounds from OCO into expected regret bounds for Online Bandit Optimization.

Lemma 6. *[16][Ch. 6] Let u be a fixed point in X. Let $f_1, ..., f_T : X \to \mathbb{R}$ be a sequence of differentiable functions. Let $\mathcal{A}$ be a first order algorithm that ensures $Regret_T(\mathcal{A}) \leq B_{\mathcal{A}}(\nabla f_1(x_1), ..., \nabla f_T(x_T))$ in the full information setting. Define $\{x_t\}$ as: $x_1 \leftarrow \mathcal{A}(\emptyset)$, $x_t \leftarrow \mathcal{A}(g_1, ..., g_{t-1})$ where each g_t satisfies:*

$$\mathbb{E}[g_t | x_1, f_1, ..., x_t, f_t] = \nabla f_t(x_t)$$

Then, for every $u \in X$:

$$\mathbb{E}[\sum_{t=1}^{T} f_t(x_t)] - \sum_{t=1}^{T} f_t(u) \leq \mathbb{E}[B_{\mathcal{A}}(g_1, ... g_T)]$$

Moreover, Online Gradient Descent is a first order Algorithm [16][Ch. 6].

A.3 Some Useful Concentration Results

In this section we present results on how quickly random functions uniformly concentrate around their mean.

Lemma 7. *[26][Theorem 5] Let $\hat{F}(x) = \frac{1}{N}\sum_{n=1}^{N} f(x, \xi_n)$ where $f(\cdot, \xi)$ is L-Lipschitz with function values bounded by R and the set where it is defined has diameter B. Let $F(x) := \mathbb{E}_\xi[f(x, \xi)]$. Then*

$$P(\sup_{x \in X} |F(x) - \hat{F}(x)| \geq \epsilon) \leq O(d^2(\frac{LB}{\epsilon})^d \exp(-\frac{N\epsilon^2}{128LR})). \tag{7}$$

This result implies the following two lemmas.

Lemma 8. *With probability at least $1-\delta$, for any $x \in X$, over a sample size N*

$$|F(x) - \hat{F}(x)| \leq \tilde{O}(\sqrt{\frac{LRd\ln(\frac{1}{\delta})}{N}}).$$

Proof. Setting the right hand side of (7) equal to δ and solving for ϵ gives

$$\epsilon = \sqrt{\frac{128LR[2\ln(\frac{d}{\sqrt{\delta}}) + d\ln(LB) + d\ln(\frac{1}{\epsilon})]}{N}}$$

Since we must bound ϵ by above, we now bound $\ln(\frac{1}{\epsilon})$. Using the previous equality we have

$$\ln(\frac{1}{\epsilon}) = \frac{1}{2}\ln(\frac{N}{128LR[2\ln(\frac{d}{\sqrt{\delta}}) + d\ln(\frac{LB}{\epsilon})]})$$

since $\ln(\frac{LB}{\epsilon})$ is large and in the denominator, we have

$$\ln(\frac{1}{\epsilon}) \leq \frac{1}{2}\ln(\frac{N}{256LR\ln(\frac{d}{\sqrt{\delta}})})$$

this implies[1]

$$\begin{aligned}
\epsilon &\leq \sqrt{\frac{128LR[2\ln(\frac{d}{\sqrt{\delta}}) + d\ln(LB) + d\frac{1}{2}\ln(\frac{N}{256LR\ln(\frac{d}{\sqrt{\delta}})})]}{N}} \\
&= \sqrt{\frac{\kappa LRd\ln(\frac{dLBN}{\sqrt{\delta}256LR\ln(\frac{d}{\delta})})}{N}} \\
&= \tilde{O}(\sqrt{\frac{LRd\ln(\frac{1}{\delta})}{N}})
\end{aligned}$$

□

Lemma 9.

$$\mathbb{E}[\sup_{x\in X}|F(x) - \hat{F}(x)|] \leq \tilde{O}(\frac{\sqrt{LRd}}{\sqrt{N}})$$

Proof. Recall that for a nonnegative random variable X it holds that $\mathbb{E}[X] = \int_0^\infty P(X > t)dt$. We have from (7)

$$\begin{aligned}
P(\sup_{x\in Z}|F(x) - \hat{F}(x)| > \epsilon) &\leq O(d^2(\frac{LB}{\epsilon})^d\exp(-\frac{N\epsilon^2}{128LR})) \\
&= \exp[-(\frac{N\epsilon^2}{128LR} + d\ln(\epsilon) - 2\ln(d) - d\ln(LB))]
\end{aligned}$$

Let $\lambda(\epsilon) = a\epsilon^2 + d\ln(\epsilon)$ with $a := \frac{N}{128LR}$ and notice that when $\epsilon \geq \sqrt{\frac{d}{2a}}$ the second derivative of $\lambda(\cdot)$ is nonnegative and therefore the function is convex in that domain thus we can lower bound it with its first order Taylor approximation at $\sqrt{\frac{d}{2a}}$.

$$\lambda(\epsilon) \geq 2\sqrt{2ad}\epsilon - 2d + \frac{d}{2} + \frac{d}{2}\ln(\frac{d}{2a})$$

Therefore, for $\epsilon \geq \sqrt{\frac{d}{2a}}$

$$\begin{aligned}
P(\sup_{x\in Z}|F(x) - \hat{F}(x)| > \epsilon) &\leq \exp[-(2\sqrt{2ad}\epsilon - 2d + \frac{d}{2} + \frac{d}{2}\ln(\frac{d}{2a}) - 2\ln(d) - d\ln(LB))] \\
&\leq \exp[-(2\sqrt{2ad}\epsilon - 2d + \frac{d}{2}\ln(\frac{d}{2a}) - 2\ln(d) - d\ln(LB))] \\
&\leq \exp[-(2\sqrt{2ad}\epsilon - 2d - \frac{d}{2}\ln(2a) - 2\ln(d) - d\ln(LB))] \\
&= \exp[-(2\sqrt{2ad}\epsilon) + \theta]
\end{aligned}$$

[1]Throughout the paper we let κ be some universal constant that may change from line to line.

where $\theta := 2d + \frac{d}{2}\ln(2a) + 2\ln(d) + d\ln(LB)$. We have

$$\begin{aligned}
\mathbb{E}[\sup_{x\in X}|F(x)-\hat{F}(x)|] &\leq \int_0^\infty \min[1, \exp[-(2\sqrt{2ad}\epsilon)+\theta]]d\epsilon \\
&= \int_0^{\epsilon'} d\epsilon + \int_{\epsilon'}^\infty \exp[-2\sqrt{2ad}\epsilon+\theta]d\epsilon \quad \epsilon' = \frac{\theta}{2\sqrt{2ad}} \\
&= \epsilon' + \frac{\exp[\theta - 2\sqrt{2ad}\epsilon']}{2\sqrt{2ad}} \\
&= \frac{1}{2\sqrt{2ad}}[\theta+1] \\
&= \frac{\sqrt{128LR}}{2\sqrt{2dN}}[2d + \frac{d}{2}\ln(2a) + 2\ln(d) + d\ln(LB)+1] \\
&= \tilde{O}(\frac{\sqrt{LRd}}{\sqrt{N}})
\end{aligned}$$

□

A.4 Conditional Value at Risk

Proof of Theorem 1. For any fixed $x \in X$, we define $\phi(z) := z + \frac{1}{\alpha}E_{\xi\sim P}[f(x,\xi)-z]_+$ and $\widehat{\phi}(z) = \frac{1}{N}\sum_{n=1}^N z + \frac{1}{\alpha}[f(x,\xi_n)-z]_+$. By Lemma 8 we know that with probability at least $1-\delta$ for all $z \in [0,1]$

$$|\phi(z) - \widehat{\phi}(z)| \leq O(\sqrt{\frac{LR\ln(N/\delta)}{N}})$$

and it is easy to see that L, R are both $O(\frac{1}{\alpha})$.

It remains to show that $A := \{X_A = \sup_z |\phi(z) - \widehat{\phi}(z)| \leq \epsilon\}$ implies $B := \{X_B = |CVaR_\alpha[F](x) - \widehat{CVaR_\alpha}[F](x)| \leq \epsilon\}$. Indeed, we have that for any $z \in Z$

$$\phi(z) - \epsilon \leq \widehat{\phi}(z)$$

Therefore, if $\bar{z} = \arg\min_{z\in Z}\widehat{\phi}(z)$ we have:

$$CVaR_\alpha[F](x) - \epsilon \leq \phi(\bar{z}) - \epsilon \leq \widehat{\phi}(\bar{z}) = \widehat{CVaR}_\alpha[F](x)$$

The other side of the inequality follows by applying the same type of argument to $\widehat{\phi}(z) \leq \phi(z) + \epsilon$. □

Remark 1. *We make one last remark about the proof above. We showed that $A \implies B$ therefore $P(B') \leq P(A')$. Since for a nonnegative random variable X we can write $\mathbb{E}[X] = \int P(X > \epsilon)d\epsilon$ we can conclude that $\mathbb{E}[X_B] \leq \mathbb{E}[X_A]$, or which is the same, $\mathbb{E}[|CVaR_\alpha[F](x) - \widehat{CVaR_\alpha}[F](x)|] \leq \mathbb{E}[\sup_z|\phi(z) - \widehat{\phi}(z)|]$.*

Lemma 10. *Let ξ be a random variable supported in Ξ with probability distribution P. Let $f : X \times \Xi \to \mathbb{R}$ and assume $0 \leq f(x,\xi) \leq 1$ for all $x \in X$ and $\xi \in \Xi$. If $f(\cdot,\xi)$ is G-Lipschitz then so is $CVaR_\alpha[F](x)$.*

Proof. By Theorem 6.4 in [28] for any $x \in X$. We have

$$CVaR_\alpha[F](x) = \sup_{\xi\in\Theta}\mathbb{E}_\xi[f(x,\xi)]$$

where Θ is some family of probability distributions.

Since convex combinations of G-Lipschitz functions is G-Lipschitz we have that for any $x_1 \in X$

$$\mathbb{E}_{\xi\in\Theta_1^*}[f(x_1,\xi)] - \mathbb{E}_{\xi\in\Theta_1^*}[f(x_2,\xi)] \leq G||x_1 - x_2||$$

where Θ_1^* is the probability distribution that maximizes $\mathbb{E}_{\xi \in \Theta}[f(x_1, \xi)]$ (assuming it exists). Since

$$\mathbb{E}_{\xi \in \Theta_1^*}[f(x_1, \xi)] - \mathbb{E}_{\xi \in \Theta_2^*}[f(x_2, \xi)] \leq \mathbb{E}_{\xi \in \Theta_1^*}[f(x_1, \xi)] - \mathbb{E}_{\xi \in \Theta_1^*}[f(x_2, \xi)]$$

by combining the two inequalities we have

$$CVaR_\alpha[F](x_1) - CVaR_\alpha[F](x_2) \leq G||x_1 - x_2||$$

a symmetry argument yields the other side of the inequality, this concludes the proof. □

Lemma 11. *Let X be a convex set with diameter $D_{||\cdot||}$ that contains the origin, that is for all $x_1, x_2 \in X$, $||x_1 - x_2|| \leq D_{||\cdot||}$. Let $X_\delta := \{x : x \in (1-\delta)X\}$. For any $x \in X$ let $x_\delta := \Pi_{X_\delta}(x)$ where the projection is taken with respect to any norm $||\cdot||$. Then*

$$||x - x_\delta|| \leq \delta D_{||\cdot||} \tag{8}$$

Proof. Notice $(1-\delta)x \in X_\delta$

$$\begin{aligned} ||x - x_\delta|| &\leq ||x - (1-\delta)x|| \quad \text{By definition of } \Pi \\ &\leq \delta ||x|| \\ &\leq \delta D_{||\cdot||} \quad \text{since } X \text{ contains the origin} \end{aligned}$$

□

Lemma 12. *Let $x = [x_1, x_2]^\top$. Define $||x|| = ||x_1||_2 + ||x_2||_\infty$. Then*

$$||x||_* = \max\{||x_1||_2, ||x_2||_1\}$$

Proof. By definition of dual norm we have

$$\begin{aligned} ||x||_* &= \max_{||y|| \leq 1} x_1^\top y_1 + x_2^\top y_2 \\ &= \max_{||y_1||_2 + ||y_2||_\infty \leq 1} x^\top y \\ &= \max_{c_1 + c_2 \leq 1} c_1 ||x_1||_2 + c_2 ||x_2||_1 \\ &= \max\{||x_1||_2, ||x_2||_1\} \end{aligned}$$

□

A.5 Analysis of Algorithm 1

Lemma 13. *The function $\mathcal{L}_t(x, z) := z + \frac{1}{\alpha}[f_t(x) - z]_+$ is jointly convex, $G_\mathcal{L}$-Lipschitz continuous with $G_\mathcal{L} = \alpha^{-1}(G+1) + 1$, and the diameter of the set where it is defined $D_\mathcal{L} \leq D_X + 1$.*

Proof. We first prove convexity. The function $f_t(x) - z$ is jointly convex since both $f_t(x)$ and $-z$ are, and addition preserves convexity. Point-wise supremum over convex functions preserves convexity and since any constant function is convex we have that $[f_t(x) - z]_+$ is convex. Again, using the fact that addition preserves convexity we get the desired claim.

To prove the second part of the claim we notice:

$$\nabla_x \mathcal{L}_t(x, z) = \begin{cases} \frac{1}{\alpha} \nabla f_t(x) & \text{if } f_t(x) - z > 0 \\ 0 & \text{otherwise} \end{cases}$$

$$\nabla_z \mathcal{L}_t(x, z) = \begin{cases} 1 - \frac{1}{\alpha} & \text{if } f_t(x) - z > 0 \\ 1 & \text{otherwise} \end{cases}$$

Let $\nabla\mathcal{L}_t := [\nabla_x\mathcal{L}_t; \nabla_y\mathcal{L}_t]$ and recall that a function f is G-Lipschitz continuous if and only if $||\nabla f|| \leq G$. We have that We have that

$$\begin{aligned}||\mathcal{L}_t|| &\leq \max\{||[\bar{0};1]||, ||[\alpha^{-1}\nabla f; 1+\alpha^{-1}]||\} \\ &\leq \alpha^{-1}(G+1)+1 =: G_{\mathcal{L}}\end{aligned}$$

Where the last inequality follows by simple algebra.

The fact that $D_{\mathcal{L}} \leq D_X + 1$ follows from the definition of the diameter of a set. □

The key to prove Theorem 2 is to realize that Algorithm 1 is performing Online Gradient Descent using an estimate of the gradient of the smoothened function $\hat{\mathcal{L}}_t$ as in [9].

Next we prove a lemma assuming that for every $t = 1, ..., T$ $\nabla\mathcal{L}_t := \nabla\mathcal{L}_t(x_t, z_t)$ is revealed and we update according to

$$[x_{t+1}, z_{t+1}]^\top \leftarrow \Pi_{X\times Z}([x_t, z_t]^\top - \eta\nabla\mathcal{L}_t) \tag{9}$$

That is, we perform Zinkevich's Online gradient Descent (OGD) on functions $\mathcal{L}_t$ [32]. Due to Lemma 6 we will be able to use this guarantee when we have bandit feedback.

Lemma 14. *Applying OGD on sequence of functions $\{\mathcal{L}_t\}_{t=1}^T$ guarantees: for every $w = (x, z) \in \mathcal{W} := X \times Z$.*

$$\sum_{t=1}^T \mathcal{L}_t(w_t) - \sum_{t=1}^T \mathcal{L}_t(w) \leq \frac{D_{\mathcal{L}}}{2\eta} + \frac{\eta}{2}\sum_{t=1}^T ||\nabla\mathcal{L}_t||^2.$$

Proof. We follow Zinkevich's proof. By properties of projections we have:

$$\begin{aligned}||w_{t+1} - w||^2 &\leq ||w_t - \eta\nabla\mathcal{L}_t - w||^2 \\ &= ||w_t - w||^2 + \eta^2||\nabla\mathcal{L}_t||^2 - 2\eta\nabla\mathcal{L}_t^\top(w_t - w)\end{aligned}$$

Therefore:

$$2\eta\nabla\mathcal{L}_t^\top(w_t - w) \leq \frac{||w_t - w||^2 - ||w_{t+1} - w||^2}{\eta} + \eta||\nabla\mathcal{L}_t||^2$$

Using convexity and summing up the inequalities above for every t we have:

$$2(\sum_{t=1}^T \mathcal{L}_t(w_t) - \sum_{t=1}^T \mathcal{L}_t(w)) \leq \sum_{t=1}^T 2\eta\nabla\mathcal{L}_t^\top(w_t - w) \tag{10}$$

$$\leq \sum_{t=1}^T \frac{||w_t - w||^2 - ||w_{t+1} - w||^2}{\eta} + \eta\sum_{t=1}^T ||\nabla\mathcal{L}_t||^2 \tag{11}$$

$$\leq \frac{D_{\mathcal{L}}}{\eta} + \eta\sum_{t=1}^T ||\nabla\mathcal{L}_t||^2$$

Which yields the desired result. □

Lemma 15. *Let $\tilde{y}_t = (\tilde{x}_t, \tilde{z}_t)$ and $y^* = (x^*, z^*) :=$ argmin$_{x,z\in X\times Z}\sum_{t=1}^T \mathbb{E}_\xi[\mathcal{L}_t(x, z)]$, Algorithm 1 guarantees:*

$$\sum_{t=1}^T \mathbb{E}_{int}[\mathcal{L}_t(\tilde{y}_t)] - \sum_{t=1}^T \mathcal{L}_t(y^*) = O(\frac{dD_X G T^{3/4}}{\alpha})$$

Proof. Define $y^*_\delta = \Pi_{X_\delta}[y^*]$. By Lemma 11 in the Appendix, it holds that $||y^*_\delta - y^*|| \leq \delta D_{\mathcal{L}}$. Using a similar argument as in [9] we have:

$$\begin{aligned}
&\mathbb{E}_{int}[\sum_{t=1}^T \mathcal{L}_t(\tilde{y}_t) - \sum_{t=1}^T \mathcal{L}_t(y^*)] \\
&\leq \mathbb{E}_{int}[\sum_{t=1}^T \mathcal{L}_t(y_t) - \sum_{t=1}^T \mathcal{L}_t(y^*)] + \delta G_{\mathcal{L}} T \quad \text{by Lemma 1 and } ||y_t - \tilde{y}_t|| \leq \delta \\
&\leq \mathbb{E}_{int}[\sum_{t=1}^T \mathcal{L}_t(y_t) - \sum_{t=1}^T \mathcal{L}_t(y^*_\delta)] + \delta G_{\mathcal{L}} T + \delta G_{\mathcal{L}} D_{\mathcal{L}} T \\
&\leq \mathbb{E}_{int}[\sum_{t=1}^T \hat{\mathcal{L}}_t(y_t) - \sum_{t=1}^T \hat{\mathcal{L}}_t(y^*_\delta)] + 3\delta G_{\mathcal{L}} T + \delta G_{\mathcal{L}} D_{\mathcal{L}} T \quad \text{by Lemma 1} \\
&\leq \frac{\eta}{2} \sum_{t=1}^T \mathbb{E}_{int}[||g_t||^2] + \frac{D^2_{\mathcal{L}}}{2\eta} + 3\delta G_{\mathcal{L}} T + \delta D_{\mathcal{L}} G_{\mathcal{L}} T \quad \text{by Lemma 6} \\
&\leq \frac{\eta}{2} \frac{(d+1)^2}{\delta^2} \sum_{t=1}^T |\tilde{z}_t + \frac{1}{\alpha}[f_t(\tilde{x}_t) - \tilde{z}_t]|^2 + \frac{D^2_{\mathcal{L}}}{2\eta} + 3\delta G_{\mathcal{L}} T + \delta D_{\mathcal{L}} G_{\mathcal{L}} T \\
&\leq \frac{\eta}{2} \frac{(d+1)^2}{\delta^2 \alpha^2} T + \frac{D^2_{\mathcal{L}}}{2\eta} + 3\delta G_{\mathcal{L}} T + \delta D_{\mathcal{L}} G_{\mathcal{L}} T \\
&= O(\frac{d D_X G T^{3/4}}{\alpha})
\end{aligned}$$

Where we chose $\eta = O(\frac{D_X \alpha}{d T^{3/4}})$ and $\delta = O(\frac{1}{T^{1/4}})$. □

We are now ready to give a proof of Theorem 2.

Proof of Theorem 2. Notice that for all t, every $x \in X$ and every $z \in Z$, we have:

$$\mathbb{E}_{\xi \sim P}[\mathcal{L}_t(x, z)] = z + \frac{1}{\alpha} \mathbb{E}_{\xi \sim P}[f(x, \xi) - z]_+ \geq CVaR_\alpha[F](x).$$

The result then follows by taking $\mathbb{E}_{\xi \sim P}[\cdot]$ in both sides of the result in Lemma 15 and interchanging the expectations. The interchange can be done using Fubini's Theorem since for every $x \in X$ and for every $z \in Z$ we have that $\mathcal{L}_t(x, z) < O(\frac{1}{\alpha})$ almost surely. □

We are now ready to prove Theorem 3. We assume f_t is 1-Lipschitz continuous.

Proof of Theorem 3. Define concentration error $CE = C_\alpha[\{f_t(x^*)\}_{t=1}^T] - C_\alpha[\{f_t(\bar{x})\}_{t=1}^T]$, where $\bar{x} = \arg\min_{x \in X} C_\alpha[\{f_t(x)\}_{t=1}^T]$, let $x^* = \arg\min_{x \in X} C_\alpha[F](x)$, we have

$$
\begin{aligned}
&\mathbb{E}[C_\alpha[\{f_t(x_t)\}_{t=1}^T] \pm C_\alpha[\{f_t(x^*)\}_{t=1}^T]] - min_{x\in X} C_\alpha[\{f_t(x)\}_{t=1}^T] \\
&= \mathbb{E}[\min_y y + \frac{1}{\alpha T}\sum_{t=1}^T \max\{f_t(x_t) + f_t(x^*) - f_t(x^*) - y, 0\} - C_\alpha[\{f_t(x^*)\}_{t=1}^T]] + \mathbb{E}[CE] \\
&\le \mathbb{E}[\min_y y + \frac{1}{\alpha T}\sum_{t=1}^T \max\{f_t(x^*) + |f_t(x_t) - f_t(x^*)| - y, 0\} - C_\alpha[\{f_t(x^*)\}_{t=1}^T]] + \mathbb{E}[CE] \\
&\le \mathbb{E}[\min_y y + \frac{1}{\alpha T}\sum_{t=1}^T \max\{f_t(x^*) + |f_t(x_t) - f_t(x^*)| - y, |f_t(x_t) - f_t(x^*)|\} - C_\alpha[\{f_t(x^*)\}_{t=1}^T]] + \mathbb{E}[CE] \\
&= \mathbb{E}[\min_y y + \frac{1}{\alpha T}\sum_{t=1}^T \max\{f_t(x^*) - y, 0\} + \frac{1}{\alpha T}\sum_{t=1}^T + |f_t(x_t) - f_t(x^*)| - C_\alpha[\{f_t(x^*)\}_{t=1}^T]] + \mathbb{E}[CE] \\
&= \mathbb{E}[\frac{1}{\alpha T}\sum_{t=1}^T |f_t(x_t) - f(x^*)|] + \mathbb{E}[CE] \\
&\le \frac{1}{\alpha T}\sum_{t=1}^T \mathbb{E}_t[||x_t - x^*||] + \mathbb{E}[CE] \quad \text{since } f_t \text{ is 1-Lipschitz} \\
&\le \frac{1}{\alpha T}\sqrt{T}\sqrt{\sum_{t=1}^T \mathbb{E}_t[||x_t - x^*||]^2} + \mathbb{E}[CE] \quad \text{by Cauchy Schwartz} \\
&\le \frac{1}{\alpha T}\sqrt{T}\sqrt{\sum_{t=1}^T \mathbb{E}_t[\frac{2}{\beta}[C_\alpha[F](x_t) - C_\alpha[F](x^*)]]} + \mathbb{E}[CE] \quad \text{by strong convexity of } C_\alpha[F](\cdot) \text{ and KKT condition} \\
&= \frac{1}{\alpha T}\sqrt{T}\sqrt{\frac{2}{\beta}\mathbb{E}[\sum_{t=1}^T C_\alpha[F](x_t) - C_\alpha[F](x^*)]} + \mathbb{E}[CE] \\
&= O(\frac{d^{1/2}}{\alpha^{3/2}\beta^{1/2}T^{1/8}}) + \mathbb{E}[CE] \quad \text{by Theorem 2}
\end{aligned}
$$

We still need to bound the concentration error CE in expectation. Notice we can write

$$
CE = [C_\alpha[\{f_t(x^*)\}_{t=1}^T] - C_\alpha[F](x^*)] + [C_\alpha[F](x^*) - C_\alpha[F](\bar{x})] + [C_\alpha[F](\bar{x}) - C_\alpha[\{f_t(\bar{x})\}_{t=1}^T]
$$

and the second term is nonpositive. To bound CE in expectation we apply Lemma 9 on functions $\phi(x, y) = y + \frac{1}{\alpha}[f(x) - y]_+$ (notice $L \le O(\frac{1}{\alpha})$ and $R = O(\frac{1}{\alpha})$), by Remark 1 and the same reasoning as in the proof of Lemma 9 we have $\mathbb{E}[|C_\alpha[F](\bar{x}) - C_\alpha[\{f_t(\bar{x})\}_{t=1}^T|] \le \tilde{O}(\frac{\sqrt{d}}{\alpha\sqrt{T}})$. Thus $\mathbb{E}[CE] \le \tilde{O}(\frac{\sqrt{d}}{\alpha\sqrt{T}})$. This finishes the proof. □

A.6 Analysis of Algorithm 2 (1-D)

We proceed to formally analyze the algorithm following [2]. In this section, for ease of reading we refer to quantity $T\bar{\mathcal{R}}_T$ as the regret. We work conditioned on $\mathcal{E}$ which is defined as the event that for every epoch and for every round i, $h(x) \in [LB_{\gamma_i}(x), UB_{\gamma_i}(x)]$ for $x \in \{x_l, x_c, x_r\}$. We will first bound the regret in an epoch and then bound the total number of epochs. We do the previous in the next sequence of lemmas. Notice that by Theorem 1 we can obtain a γ-CI for $h(x)$ that holds with probability at least $1 - \frac{1}{T^2}$ with only $\frac{\kappa \ln(T/(\alpha\gamma))}{\alpha^2\gamma^2}$ samples. We first show that we never discard points that are near optimal.

Lemma 16. *If epoch τ ends in round i, then the interval $[l_{\tau+1}, r_{\tau+1}]$ contains every $x \in [l_\tau, r_\tau]$ such that $h(x) \le h(x^*) + \gamma_i$. In particular, $x^* \in [l_\tau, r_\tau]$ for all epochs τ.*

Proof. Assume epoch τ terminates in round i through Case 1. Then, either $LB_{\gamma_i}(x_l) \ge UB_{\gamma_i}(x_r) + \gamma_i$ or $LB_{\gamma_i}(x_r) \ge UB_{\gamma_i}(x_l) + \gamma_i$. We assume the former occurs. It then holds that

$$
h(x_l) \ge h(x_r) + \gamma_i.
$$

We must show that the points in the working feasible region to the left of x_l are not near optimal. That is, for every $x \in [l_\tau, l_{\tau+1}] = [l_\tau, x_l]$ we have $h(x) \geq h(x^*) + \gamma_i$. Pick $x \in [l_\tau, x_l]$ then, for some $t \in [0,1]$ we have $x_l = tx + (1-t)x_r$. Since h is convex we have

$$h(x_l) \leq th(x) + (1-t)h(x_r)$$

which implies

$$\begin{aligned} h(x) &\geq h(x_r) + \frac{h(x_l) - h(x_r)}{t} \\ &\geq h(x_r) + \frac{\gamma_i}{t} \\ &\geq h(x^*) + \gamma_i \end{aligned}$$

as required. If $LB_{\gamma_i}(x_r) \geq UB_{\gamma_i}(x_l) + \gamma_i$ had occurred the argument is analogous.

If epoch τ had terminated through case 2 then

$$\max\{LB_{\gamma_i}(x_l), LB_{\gamma_i}(x_r)\} \geq UB_{\gamma_i}(x_c) + \gamma_i.$$

We assume $LB_{\gamma_i}(x_l) \geq UB_{\gamma_i}(x_c) + \gamma_i$, then

$$h(x_l) \geq h(x_c) + \gamma_i.$$

The same argument as above with x_c instead of x_r guarantees $h(x_l) \geq h(x^*) + \gamma_i$. If $LB_{\gamma_i}(x_r) \geq UB_{\gamma_i}(x_c) + \gamma_i$ had occurred the argument is analogous. The fact that $x^* \in [l_\tau, r_\tau]$ for every epoch τ follows by induction. $\square$

We now show that if an epoch does not terminate in a given round i then the regret ($T\bar{\mathcal{R}}_T$) incurred in that epoch was not too high.

Lemma 17. *If epoch τ continues from round i to $i+1$ then the regret in round i is at most*

$$\frac{\kappa \ln(T/(\alpha\gamma_i))}{\alpha^2 \gamma_i}$$

Proof. The regret incurred in round i of epoch τ is

$$\frac{\kappa \ln(T/(\alpha\gamma_i))}{\alpha^2 \gamma_i^2}[(h(x_l) - h(x^*)) + (h(x_c) - h(x^*)) + (h(x_r) - h(x^*))]$$

It suffices to show that for every $x \in \{x_l, x_c, x_r\}$ it holds that

$$h(x) \leq h(x^*) + 12\gamma_i.$$

The algorithm continues from round i to round $i+1$ if and only if

$$\max\{LB_{\gamma_i}(x_l), LB_{\gamma_i}(x_r)\} < \min\{UB_{\gamma_i}(x_l), UB_{\gamma_i}(x_r)\} + \gamma_i$$

and

$$\max\{LB_{\gamma_i}(x_l), LB_{\gamma_i}(x_r)\} < UB_{\gamma_i}(x_c) + \gamma_i.$$

This implies that $h(x_l), h(x_c)$, and $h(x_r)$ are all contained in an interval of at most $3\gamma_i$. There are two cases for which the argument is essentially the same, either $x^* \leq x_c$ or $x^* > x_c$, we consider the former. Since by the previous lemma we know that $x^* \in [l_\tau, r_\tau]$, then there exists $t \in [0,1]$ such that $x^* = x_c + t(x_c - x_r)$. Therefore

$$x_c = \frac{1}{1+t}x^* + \frac{t}{1+t}x_r.$$

Since $|x_c - l_\tau| = w_\tau/2$ and $|x_r - x_c| = w_\tau/4$ we have

$$t = \frac{|x^* - x_c|}{|x_r - x_c|} \leq \frac{|l_\tau - x_c|}{|x_r - x_c|} = \frac{w_\tau/2}{w_\tau/4} = 2$$

Since h is convex

$$h(x_c) \le \frac{1}{1+t}h(x^*) + \frac{t}{1+t}h(x_r)$$

therefore

$$\begin{aligned} h(x^*) &\ge (1+t)\big(h(x_c) - \frac{t}{1+t}h(x_r)\big) \\ &= h(x_c) + (1+t)(h(x_c) - h(x_r)) \\ &\ge h(x_c) - (1+t)|h(x_c) - h(x_r)| \\ &\ge h(x_r) - (1+t)3\gamma_i \\ &\ge h(x_r) - 9\gamma_i \end{aligned}$$

So, for all $x \in \{x_l, x_c, x_r\}$ it holds that

$$h(x) \le h(x_r) + 3\gamma_i \le h(x^*) + 12\gamma_i.$$

□

We proceed to bound the regret in each epoch.

Lemma 18. *If epoch τ ends in round i the regret incurred in the epoch is no more than*

$$\frac{\kappa \ln(T/(\alpha\gamma_i))}{\alpha^2\gamma_i}.$$

Proof. If $i = 1$, since $h(x)$ is 1-Lipschitz and $X = [0,1]$ we have that for every $x \in \{x_l, x_c, x_r\}$ $h(x) - h(x^*) \le 1$. Therefore the regret in epoch τ is

$$\begin{aligned} &\frac{\kappa \ln(T/(\alpha^2\gamma_i^2))}{\alpha^2\gamma_i^2}\big(((h(x_l) - h(x^*)) + (h(x_c) - h(x^*)) + (h(x_r) - h(x^*))\big) \\ &\le \frac{6\kappa \ln(T/(\alpha^2\gamma_i^2))}{\alpha^2\gamma_1} \end{aligned}$$

If $i \ge 2$, by the previous lemma we have that the regret incurred in round j with $1 \le j \le i-1$ is no more than

$$\frac{\kappa \ln(T/(\alpha^2\gamma_i^2))}{\alpha^2\gamma_j}.$$

For round i the regret incurred is at most

$$3 \cdot 12\gamma_{i-1}\frac{\kappa \ln(T/(\alpha^2\gamma_i^2))}{\alpha^2\gamma_i^2} = \frac{\kappa 72 \ln(T/(\alpha^2\gamma_i^2))}{\alpha^2\gamma_i}.$$

It follows that the regret in epoch τ is

$$\begin{aligned} &\sum_{j=1}^{i-1}\frac{\kappa \ln(T/(\alpha^2\gamma_j^2))}{\alpha^2\gamma_j} + \frac{\kappa \ln(T/(\alpha^2\gamma_i^2))}{\alpha^2\gamma_i} \\ &= \sum_{j=1}^{i-1}\frac{\kappa \ln(T/(\alpha^2\gamma_j^2))}{\alpha^2} \cdot 2^j + \frac{\kappa \ln(T/(\alpha^2\gamma_i^2))}{\alpha^2\gamma_i} \\ &< \frac{\kappa \ln(T/(\alpha^2\gamma_i^2))}{\alpha^2} \cdot 2^i + \frac{\kappa \ln(T/(\alpha^2\gamma_i^2))}{\alpha^2\gamma_i} \\ &= \frac{\kappa \ln(T/(\alpha\gamma_i))}{\alpha^2\gamma_i}. \end{aligned}$$

□

We have bounded the regret that we incur in each epoch. We proceed to bound the number of epochs.

Lemma 19. *The total number of epochs τ satisfies*

$$\tau \leq \kappa \log_{4/3}(\frac{\alpha^2 T}{\ln(T)}).$$

Proof. The key is to observe that since the number of times we sample a point is bounded above by T then $\gamma_i \geq (\alpha^2 T/(\kappa \ln(T)))^{-1/2}$ for every round and every epoch. Let $\gamma_{min} := (\alpha^2 T/(\kappa \ln(T)))^{-1/2}$ and let $I := [x^* - \gamma_{min}, x^* + \gamma_{min}]$. Since h is 1-Lipschitz, for any $x \in I$

$$h(x) - h(x^*) \leq \gamma_{min}.$$

By Lemma 16 we have that for any round τ' which ends in round i'

$$I \subseteq \{x \in [0,1] : f(x) < f(x^*) + \gamma_{i'}\} \subseteq [l_{\tau'+1}, r_{\tau'+1}]$$

since $\gamma_{min} \leq \gamma_{i'}$. The previous implies

$$2\gamma_{min} \leq r_{\tau+1} - l_{\tau+1} = w_{\tau+1}.$$

By the definitions of $l_{\tau'+1}$, $r_{\tau'+1}$ and $w_{\tau'+1}$ we have that for any $\tau' \in \{1, ..., \tau\}$

$$w_{\tau'+1} \leq \frac{3}{4} w_{\tau'}.$$

Therefore,

$$2\gamma_{min} \leq w_{\tau+1} \leq (\frac{3}{4})^{\tau} w_1 \leq (\frac{3}{4})^{\tau}$$

which yields the result. □

We are now ready to prove Theorems 4 and 5 .

Proof of Theorem 4. The per epoch regret when epoch τ ends in round i is

$$\frac{\kappa \ln(T/(\alpha\gamma_i))}{\alpha^2 \gamma_i} \leq \frac{\kappa \ln(T/(\alpha\gamma_i))}{\alpha^2 \gamma_{\min}} \leq \frac{\kappa\sqrt{T} \ln(T/(\alpha\gamma_{\min}))}{\alpha} = \frac{\kappa\sqrt{T}\ln(T)}{\alpha}.$$

Using the previous lemma we know that the regret will not be more than

$$\frac{\kappa\sqrt{T}\ln(T)}{\alpha} \log_{4/3}(\frac{\alpha^2 T}{\ln(T)})$$

Recall we have been working conditioned on $\mathcal{E}$. We need an upper bound on $P(\mathcal{E}')$. We know that after $\frac{\kappa \ln(T/(\alpha\gamma))}{\alpha^2 \gamma_i}$ queries we have

$$P(|\hat{h}(x) - h(x)| \geq \gamma_i) \leq \frac{1}{T^2}.$$

Since there are at most T epochs a union bound gives

$$P(\mathcal{E}') \leq \frac{1}{T}$$

which yields the desired result. □

Proof of Theorem 5. The proof is very similar to that of Theorem 3 with the difference that we have to bound the concentration error $CE := C_\alpha[\{f_t(x^*)\}_{t=1}^T] - \min_{x\in X} C_\alpha[\{f_t(x)\}_{t=1}^T]$ with high probability. As explained in the proof of Theorem 3 we know

$$CE \leq |C_\alpha[\{f_t(x^*)\}_{t=1}^T] - C_\alpha[F](x^*)| + |C_\alpha[F](\bar{x}) - C_\alpha[\{f_t(\bar{x})\}_{t=1}^T]|$$

where $\bar{x} = \arg\min_{x\in X} C_\alpha[\{f_t(x)\}_{t=1}^T]$. To bound CE with high probability we apply Lemma 8 with $\delta = 1/T$ on functions $\phi(x,y) = y + \frac{1}{\alpha}[f(x) - y]_+$ (notice $L \leq O(\frac{1}{\alpha})$ and $R = O(\frac{1}{\alpha})$), by the same reasoning as in the proof of Theorem 1 we have that with probability at least $1 - \frac{1}{T}$, $|C_\alpha[F](\bar{x}) - C_\alpha[\{f_t(\bar{x})\}_{t=1}^T]| \leq \tilde{O}(\frac{1}{\alpha\sqrt{T}})$ and thus by a union bound we have that with probability at least $1 - \frac{2}{T}$, $CE \leq \tilde{O}(\frac{1}{\alpha\sqrt{T}})$. As in the proof of Theorem 3 we have

$$\mathcal{R}_T \leq \frac{\sqrt{T}}{\alpha T \beta^{1/2}}\sqrt{T\bar{\mathcal{R}}_T} + CE.$$

Using Theorem 4 to bound $\bar{\mathcal{R}}_T$, the argument in the previous paragraph to bound CE, and a union bound yields the result. □

A.7 Analysis of Algorithm 2 (d-D)

We first describe the algorithm informally. As in the special case from the previous section, Algorithm 2 proceeds in epochs. Let the initial working feasible region be $\mathcal{X}_0 = X$. The goal is that at the end of every epoch τ we will discard some portion of the working region $\mathcal{X}_\tau$ and end up with a smaller region $\mathcal{X}_{\tau+1}$ which contains at least one approximate optimum.

We now give a brief description of the algorithm. At the beginning of every epoch τ we apply an affine transformation to the current working region $\mathcal{X}_\tau$ such that the smallest ellipsoid that contains it is an Euclidean ball of radius R_τ which we denote $\mathcal{B}(R\tau)$. We assume that $R_1 \leq 1$. Let $r_\tau := R_\tau/(c_1 d)$ for some $c_1 \geq 1$ so that $\mathcal{B}(r_\tau) \subseteq \mathcal{X}_\tau$ (such a construction is always possible see Lecture 1 p. 2 of [3]). We refer to the enclosing ball $\mathcal{B}(R_\tau)$ as $\mathcal{B}_\tau$. Every epoch will consist of several rounds where γ_i is halved in every round.
Let x_0 be the center of $\mathcal{B}_\tau$. At the start of epoch τ, we build a simplex with center x_0 contained in $\mathcal{B}(r_\tau)$. We will play the vertices of the simplex $x_1, ..., x_{d+1}$ enough times so that the CI's at each vertex are of width γ_i and hold with high probability. The algorithm will then choose point y_1 for which $\hat{h}(x)_i$ is the largest, here $\hat{h}$ denotes the empirical estimate of h. By construction we are guaranteed that $h(y_1) \geq h(x_j) - \gamma_i$ for $j = 1, ..., d+1$.
The algorithm will now try to identify a region where the function value is high so that at the end of the epoch we can discard it. It will do this by constructing pyramids with parameter $\hat{\gamma}$ (always greater that γ) until a bad region is found, if this does not happen for the current value of γ it means that the algorithm did not incur to much regret (relative to how large γ was). The pyramid construction follows from Section 9.2.2 of [20]. The pyramids have angle 2ϕ at the apex where $\cos(\phi) = c_2/d$. The base of the pyramid has d vertices, $z_1, ..., z_d$ such that $z_i - x_0$ and $y_1 - z_i$ are orthogonal. The previous construction is always possible. Indeed, take a sphere with diameter $y_1 - x_0$ and arrange $z_1, ...z_d$ on its boundary such that the angle between $y_1 - x_0$ and $y_1 - z_i$ is ϕ. We now set $\hat{\gamma} = 1$ and play all the points $y_1, z_1, ...z_d$, and the center of the pyramid enough times until all the CI's are of width $\hat{\gamma}$. Let TOP and BOTTOM be the vertices of the pyramid (including y_1) with the largest and smallest values for $\hat{h}(x)$. Let $\Delta(\cdot), \bar{\Delta}(\cdot)$, be functions which are specified later. We then check for one of the following cases:

1. If $LB_{\hat{\gamma}}(\text{TOP}) \geq UB_{\hat{\gamma}}(\text{BOTTOM}) + \Delta_\tau(\hat{\gamma})$ then we proceed depending on what the separation between the CI's of TOP and APEX is.
   (a) If $LB_{\hat{\gamma}}(\text{TOP}) \geq UB_{\hat{\gamma}}(\text{APEX}) + \hat{\gamma}$, then with high probability
   $$h(\text{TOP}) \geq h(\text{APEX}) + \hat{\gamma} \geq h(\text{APEX}) + \gamma_i.$$
   We then build a new pyramid with apex equal to TOP, reset $\hat{\gamma} = 1$ and continue sampling on the new pyramid.
   (b) If $LB_{\hat{\gamma}}(\text{TOP}) < UB_{\hat{\gamma}}(\text{APEX}) + \hat{\gamma}$, then $LB_{\hat{\gamma}}(\text{APEX}) \geq UB_{\hat{\gamma}}(\text{BOTTOM}) + \Delta(\hat{\gamma}) - 2\hat{\gamma}$. We then conclude the epoch and pass the current apex to the cone-cutting subroutine.

2. If $LB_{\hat{\gamma}}(\text{TOP}) < UB_{\hat{\gamma}}(\text{BOTTOM}) + \Delta_\tau(\hat{\gamma})$, then one of the following things happen:
   (a) If $UB_{\hat{\gamma}}(\text{CENTER}) \geq LB_{\hat{\gamma}}(\text{BOTTOM}) - \bar{\Delta}_\tau(\hat{\gamma})$, then all the vertices of the pyramid and the center of the pyramid have function values in an interval of size $2\Delta_\tau(\hat{\gamma}) + 3\hat{\gamma}$. We can then set $\hat{\gamma} = \hat{\gamma}/2$. If $\hat{\gamma} < \gamma_i$, we start the next round with $\gamma_{i+1} = \gamma_i/2$. Otherwise we continue sampling with the new $\hat{\gamma}$.
   (b) If $UB_{\hat{\gamma}}(\text{CENTER}) < LB_{\hat{\gamma}}(\text{BOTTOM}) - \bar{\Delta}_\tau(\hat{\gamma})$. We conclude the epoch and pass the center and current apex to the hat-raising subroutine.

Hat-Raising: This occurs whenever the pyramid satisfies $LB_{\hat{\gamma}}(\text{TOP}) \leq UB_{\hat{\gamma}}(\text{BOTTOM}) + \Delta_\tau(\hat{\gamma})$ and $UB_{\hat{\gamma}}(\text{CENT}) \leq LB_{\hat{\gamma}}(\text{BOTTOM}) - \bar{\Delta}_\tau(\hat{\gamma})$. We will later show that if we move the apex a little from y_i to y_i', then the CI of y_i' is above the CI of TOP and the new angle ϕ' in not too much smaller than 2ϕ. In particular, we will let $y_i' = y_i + (y_i - \text{CENTER}_i)$.

Cone-cutting: This is the last step in a given epoch (notice this is the last step in the hat-raising subroutine). This subroutine receives a pyramid with apex y and base $z_1, ..., z_d$ with angle $2\bar{\phi}$ at the apex such that $\cos(\bar{\phi}) \leq 1/2d$. Define the cone

$$K_\tau = \{x : \exists\lambda > 0, \alpha_1, ..., \alpha_d > 0, \sum_{i=1}^{d} \alpha_i = 1 : x = y - \lambda \sum_{i=1}^{d} \alpha_i (z_i - y)\} \tag{12}$$

which is centered at y and is the reflection of the pyramid around the apex. By construction $\mathcal{K}_\tau$ has angle $2\bar{\phi}$ at the apex. Let $\mathcal{B}'_{\tau+1}$ be the minimum volume ellipsoid that contains $\mathcal{B}_\tau \setminus \mathcal{K}_\tau$ and let $\mathcal{X}_{\tau+1} = \mathcal{X}_\tau \cap \mathcal{B}'_{\tau+1}$. Finally, by applying an affine transformation to $\mathcal{B}'_{\tau+1}$ we obtain $\mathcal{B}_{\tau+1}$.

Before proving that the algorithm achieves low regret we discuss the computational aspects of the algorithm. The most computationally intensive steps are cone-cutting, and the isotropic transformation that transforms $B'_{\tau+1}$ into a sphere $B_{\tau+1}$. These steps are analogous to the implementation of the ellipsoid algorithm. In particular, there is an equation for $B'_{\tau+1}$ see [12]. The affine transformations can be computed via rank one matrix updates and therefore the computation of inverses can be done efficiently.

We follow [2] for the analysis of the algorithm. The main difference in the analysis is that we must build estimates of the $CVaR$ of the random loss at every point instead of building them for the expected loss. Because of this, we have to use different concentration results which directly affect how many times we must choose an action.

In this section we will first prove the correctness of the algorithm and then bound the regret. As in the 1-dimensional case we work conditioned on $\mathcal{E}$ which is defined as the event that for every epoch and every round i, $h(x) \in [LB_{\gamma_i}(x), UB_{\gamma_i}(x)]$ for all x played in that round. We will assume that

$$\Delta_\tau(\gamma) = \big(\frac{6c_1 d^4}{c_2^2} + 3\big)\gamma \text{ and } \bar{\Delta}_\tau(\gamma) = \big(\frac{6c_1 d^4}{c_2^2} + 5\big)\gamma \tag{13}$$

and $c_1 \geq 64$, $c_2 \leq 1/32$.

A.7.1 Correctness of the Algorithm

In the next sequence of lemmas we show that whenever the cone-cutting procedure is carried out we do not discard all the approximate optima of h. We also show that the hat-raising step does what we claim.

For the next two lemmas we assume that the distance from apex y of any Π built in epoch τ to the center of $\mathbb{B}(r_\tau)$ is at least r_τ/d. That the previous is true will be shown later.

Lemma 20. *Let $\mathcal{K}_\tau$ be the cone that will be discarded in epoch τ through case 1b) in round i. Let* BOTTOM *be the lowest CI of pyramid Π. Assume the distance from the apex y to the center of $\mathbb{B}(r_\tau)$ is at least r_τ/d. Then $h(x) \geq h(\text{BOTTOM}) + \gamma_i \ \forall x \in \mathcal{K}_\tau$.*

Proof. Let x be a point in $\mathcal{K}_\tau$. By construction, there exists a point z in the base of the pyramid such that $x = \alpha z + (1-\alpha)y$ for some $\alpha \in (0, 1]$. Using the convexity of h, the fact that z is in the base, and the fact that we are in case 1b), we have the two following inequalities

$$h(z) \leq h(\text{TOP}) \leq h(y) + 3\hat{\gamma}$$
$$h(y) \geq h(\text{BOTTOM}) + \Delta_\tau(\hat{\gamma}) - 2\hat{\gamma}$$

Algorithm 2 ($X \subset \mathbb{R}^d$)

Input: X, constants c_1 and c_2, functions $\Delta_\tau(\gamma)$ and $\hat{\Delta}_\tau(\gamma)$, and total number of time-steps T
Let $\mathcal{X}_1 = X$
for epoch $\tau = 1, 2, ...$ **do**
  Round $\mathcal{X}_t$ so $\mathcal{B}(r_\tau) \subseteq \mathcal{X}_\tau \subseteq \mathcal{R}(R_\tau)$, R_τ is minimized and $r_\tau := R_\tau/(c_1 d)$. Let $\mathcal{B}_\tau = \mathcal{B}(R_\tau)$.
  Build a simplex with vertices $x_1, ..., x_{d+1}$ on the surface of $\mathcal{B}(r_\tau)$.
  for round $i = 1, 2, ...$ **do**
    Let $\gamma_i := 2^{-i}$
    Play x_j for each $j = 1, ..., d+1$, $\kappa \frac{\ln(T/(\alpha\gamma))}{\alpha^2\gamma_i^2}$ times and build CI's: $[\hat{C}_\alpha[F](x_j) - \gamma_i, \hat{C}_\alpha[F](x_j) + \gamma_i]$
    Let $y_1 := \arg\max_{x_j} LB_{\gamma_i}(x_j)$
    for pyramid $k = 1, 2, ...$ **do**
      Construct pyramid Π_k with apex y_k; let $z_1, ..., z_d$ be the vertices of the base of Π_k and z_0 be the center of Π_k
      loop
        Play each of $\{y_k, z_0, z_1, ..., z_d\}$, $\kappa \frac{\ln(T/(\alpha\gamma))}{\alpha^2\gamma_i^2}$ times and build CI's
        Let CENTER $:= z_0$, APEX $:= y_k$, TOP be the vertex v of Π_k maximizing $LB_{\hat{\gamma}}(v)$, BOTTOM be the vertex v of Π_k minimizing $LB_{\hat{\gamma}}(v)$
        if $LB_{\hat{\gamma}}(\text{TOP}) \geq UB_{\hat{\gamma}}(\text{BOT}) + \Delta_\tau(\hat{\gamma})$ and $LB_{\hat{\gamma}}(\text{TOP}) \geq UB_{\hat{\gamma}}(\text{APEX}) + \hat{\gamma}$: (Case 1a)) **then**
          Let $y_{k+1} :=$ TOP, immediately continue to pyramid $k+1$
        else if $LB_{\hat{\gamma}}(\text{TOP}) \geq UB_{\hat{\gamma}}(\text{BOT}) + \Delta_\tau(\hat{\gamma})$ and $LB_{\hat{\gamma}}(\text{TOP}) < UB_{\hat{\gamma}}(\text{APEX}) + \hat{\gamma}$: (Case 1b)) **then**
          Set $(\mathcal{X}_{\tau+1}, \mathcal{B}_{\tau+1}) =$ CONE-CUTTING$(\Pi_k, \mathcal{X}_\tau, \mathcal{B}_\tau)$, proceed to epoch $\tau + 1$
        else if $LB_{\hat{\gamma}}(\text{TOP}) < UB_{\hat{\gamma}}(\text{BOT}) + \Delta_\tau(\hat{\gamma})$ and $UB_{\hat{\gamma}}(\text{CENT}) \geq LB_{\hat{\gamma}}(\text{BOT}) - \bar{\Delta}_\tau(\hat{\gamma})$: (Case 2a)) **then**
          Let $\hat{\gamma} := \hat{\gamma}/2$
          if $\hat{\gamma} < \gamma_i$ **then**
            Start next round $i + 1$
          end if
        else if $LB_{\hat{\gamma}}(\text{TOP}) < UB_{\hat{\gamma}}(\text{BOT}) + \Delta_\tau(\hat{\gamma})$ and $UB_{\hat{\gamma}}(\text{CENT}) < LB_{\hat{\gamma}}(\text{BOT}) - \bar{\Delta}_\tau(\hat{\gamma})$: (Case 2b)) **then**
          Set $(\mathcal{X}_{\tau+1}, \mathcal{B}_{\tau+1}) =$ HAT-RAISING$(\Pi_k, \mathcal{X}_\tau, \mathcal{B}_\tau)$ and proceed to epoch $\tau + 1$
        end if
      end loop
    end for
  end for
end for

Algorithm CONE-CUTTING

Input: pyramid Π with apex y, (rounded) feasible region $\mathcal{X}_\tau$ for each epoch τ, enclosing ball $\mathcal{B}_\tau$
1. Let $z_1, ..., z_d$ be the vertices of the base of Π, and ϕ the angle at its apex.
2. Define the cone $\mathcal{K}_\tau = \{x | \exists\lambda > 0, \alpha_1, ..., \alpha_d > 0, \sum_{i=1}^d \alpha_i = 1, x = y - \lambda \sum_{i=1}^d \alpha_i(z_i - y)\}$
3. Set $\mathcal{B}'_{\tau+1}$ to be the minimum volume ellipsoid containing $\mathcal{B}_\tau \setminus \mathcal{K}_\tau$
4. Set $\mathcal{X}_{\tau+1} = \mathcal{X}_\tau \cap \mathcal{B}'_{\tau+1}$

Output: Output: new feasible region $\mathcal{X}'_{\tau+1}$ and enclosing ellipsoid $\mathcal{B}'_{\tau+1}$

Algorithm HAT-RAISING

Input: pyramid Π with apex y, (rounded) feasible region $\mathcal{X}_\tau$ for each epoch τ, enclosing ball $\mathcal{B}_\tau$
1. Let CENT be the center of Π
2. Set $y' = y + (y - \text{CENT})$
3. Set Π' to be the pyramid with apex y' and same base as Π
4. Set $(\mathcal{X}_{\tau+1}, \mathcal{B}'_{\tau+1}) =$ CONE-CUTTING$(\Pi', \mathcal{X}_\tau, \mathcal{B}_\tau)$

Output: new feasible region $\mathcal{X}'_{\tau+1}$ and enclosing ellipsoid $\mathcal{B}'_{\tau+1}$

where $\hat{\gamma}$ is the CI level used for the pyramid. Since h is convex we have

$$h(y) \leq \alpha h(z) + (1-\alpha)h(x) \leq \alpha(h(y) + 3\hat{\gamma}) + (1-\alpha)h(x).$$

Which implies

$$h(x) \geq h(y) - 3\frac{\alpha}{1-\alpha}\hat{\gamma} > h(\text{BOTTOM}) + \Delta_\tau(\hat{\gamma}) - 3\frac{\alpha}{1-\alpha}\hat{\gamma} - 2\hat{\gamma}.$$

We know $\alpha/(1-\alpha) = ||y-x||/||y-z||$. Since $x \in \mathbb{B}(R_\tau)$, $||y-x|| \leq 2R_\tau = 2c_1 d r_\tau$. Moreover, $||y-z||$ is at least the height of Π, which by Lemma **??** in the Appendix, is at least $r_\tau c_2^2/d^3$. Thus

$$\frac{\alpha}{1-\alpha} = \frac{||y-x||}{||y-z||} \leq \frac{2c_1 d r_\tau}{r_\tau c_2^2/d^3}.$$

This implies

$$h(x) > h(\text{BOTTOM}) + \Delta_\tau(\hat{\gamma}) - 2\hat{\gamma} - \frac{6c_1 d^4}{c_2^2}\hat{\gamma} \geq h(\text{BOTTOM}) + \gamma_i$$

as required. □

Lemma 21. *Let Π' be the pyramid built using the hat-raising procedure with apex y' and the same base as Π in round i of epoch τ. let $\mathcal{K}'_\tau$ be the cone to be removed. Assume the distance from y, the apex of Π to the center of $\mathbb{B}(r_\tau)$ is at least r_τ/d. Then Π' has angle $\bar{\phi}$ at the apex with $\cos\bar{\phi} \leq 2c_2/d$, height at most $2r_\tau c_1^2/d^2$, and every point x in $\mathcal{K}'_\tau$ satisfies $h(x) \geq h(x^*) + \gamma_i$.*

Proof. Let $y' = y + (y - \text{CENTER})$ be the apex of Π'. Let g be the height of Π (the shortest distance from the apex to the base), let g' be the height of Π' and let b be the distance from any vertex in the base to the center of the base. By Lemma **??** in the Appendix we have $g' < 2g \leq 2r_\tau c_1^2/d^2$. Since $\cos\phi = g/\sqrt{h^2+b^2} = c_2/d$ we have $\cos\bar{\phi} = g'/\sqrt{g'^2+b^2} \leq 2g/\sqrt{g^2+b^2} = 2\cos\phi = 2c_2/d$.

We now show that for all $x \in \mathcal{K}'_\tau$ we have $h(x) \geq h(x^*) + \hat{\gamma}$. Since h is convex we have $h(y) \leq (h(y) + h(\text{CENTER}))/2$ therefore $h(y') \geq 2h(y) - h(\text{CENTER})$. Since we are in case 2b) we know $h(\text{CENTER}) \leq h(y) - \bar{\Delta}_\tau(\hat{\gamma})$, so

$$h(y') \geq h(y) + \bar{\Delta}_\tau(\hat{\gamma}). \tag{14}$$

Since we are under case 2b) we have $h(y) > h(\text{TOP}) - \Delta_\tau(\hat{\gamma}) - 2\hat{\gamma} > h(x) - \Delta_\tau(\hat{\gamma}) - 2\hat{\gamma}$ for all $x \in \Pi$. We therefore have that for any z in the base of Π,

$$h(y') > h(z) + \bar{\Delta}_\tau(\hat{\gamma}) - \Delta_\tau(\hat{\gamma}) - 2\hat{\gamma} \geq h(z), \tag{15}$$

where we used the settings of $\Delta_\tau(\hat{\gamma})$ and $\bar{\Delta}_\tau(\hat{\gamma})$. Finally, for any $x \in \mathcal{K}'_\tau$ there exists $\alpha \in [0,1)$ and z in the base of Π' such that $y' = \alpha z + (1-\alpha)x$, by convexity we have $h(y') \leq \alpha h(z) + (1-\alpha)h(x) \leq \alpha h(y') + (1-\alpha)h(x)$. The previous implies $h(x) \geq h(y') \geq h(y) + \bar{\Delta}_\tau(\hat{\gamma}) \geq h(x^*) + \gamma_i$. □

A.7.2 Regret Analysis

As in the 1-dimensional case, to bound the total pseudo-regret ($T\bar{\mathcal{R}}_T$) we must bound the regret incurred in a round and then bound the total number of epochs. In this section, for ease of reading we refer to quantity $T\bar{\mathcal{R}}_T$ as the regret.

A.7.3 Bounding the regret incurred in a round.

We first bound the regret in round i if case 2a) takes place. As before, we let Π be a pyramid built by the algorithm with angle ϕ, apex y, base $z_1, ..., z_d$ and center CENTER. recall that the pyramids built by the algorithm are such that the distance from the center to the base is at least $r_\tau c_2^2/d^3$.

Lemma 22. *Suppose the algorithm reaches case 2a) in round i of epoch τ, assume $x^* \in \mathcal{B}(R_\tau)$, where x^* minimizes h. Let Π be the current pyramid and $\hat{\gamma}$ be the current width of the CI. Assume the distance from the apex of Π to the center of $\mathcal{B}(r_\tau)$ is at least r_τ/d. Then the regret incurred while playing on Π in round i is no more than*

$$\frac{\kappa d \ln(T/(\alpha\hat{\gamma}))}{\alpha^2\hat{\gamma}}\big(\frac{4d^7 c_1}{c_2^3} + \frac{d(d+2)}{c_2}\big)\big(\frac{12c_1 d^4}{c_2^2} + 11\big).$$

Proof. The proof follows by convexity. We will first bound the variation of h in the pyramid and then bound the regret on the round depending on wether x^* is in Π or not.
Since Π is a convex set we know that the function value on any point in Π is bounded above by the maximum function value at the vertices. Case 2a) implies that for any vertex its function value is bounded above by $h(\text{CENTER}) + \Delta_\tau(\hat{\gamma}) + \bar{\Delta}_\tau(\hat{\gamma}) + 3\hat{\gamma}$. The previous implies that for all $x \in \Pi$ we have

$$h(x) \leq h(\text{CENTER}) + \Delta_\tau(\hat{\gamma}) + \hat{\Delta}_\tau(\hat{\gamma}) + 3\hat{\gamma}.$$

We let $\delta := \Delta_\tau(\hat{\gamma}) + \hat{\Delta}_\tau(\hat{\gamma}) + 3\hat{\gamma}$. Let $x \in \Pi$, let b be the a point in the base of Π such that $\text{CENTER} = \alpha x + (1-\alpha)b$ for some $\alpha \in [0,1]$. We know that $(1-\alpha)/\alpha = ||\text{CENTER} - x||/||\text{CENTER} - b||$. Since the furthest x can be from CENTER is when x is a vertex, and the distance from CENTER to b is at least the radius of the largest ball inscribed in Π with center CENTER, by Lemma **??** in the Appendix we have

$$\frac{1-\alpha}{\alpha} = \frac{||\text{CENTER} - x||}{||\text{CENTER} - b||} \leq \frac{d(d+1)}{c_2}$$

Since h is convex and we have a bound on all the function values over Π we have

$$h(\text{CENTER}) \leq \alpha h(x) + (1-\alpha)h(b) \leq \alpha h(x) + (1-\alpha)(h(\text{CENTER}) + \delta).$$

This implies

$$h(x) \geq h(\text{CENTER}) - \frac{d(d+1)\delta}{c_2}. \tag{16}$$

Combining the previous two equations we have that for any $x, x' \in \Pi$

$$|h(x) - h(x')| \leq \frac{d(d+2)\delta}{c_2}.$$

Consider the case when $x^* \in \Pi$. Since in a given round we sample $d+2$ points in the pyramid, each of them only $\kappa \ln(T/(\alpha\hat{\gamma}))/(\alpha^2\hat{\gamma}^2))$ we have that the total regret incurred when sampling the pyramid is no more than

$$(d+2)(\frac{d(d+2)\delta}{c_2})(\frac{\kappa \ln(T/(\alpha\hat{\gamma}))}{\alpha^2\hat{\gamma}^2}).$$

We now consider the case where $x^* \notin \Pi$. Recall that we always have $x^* \in \mathcal{B}_\tau$ by Lemma 20. Thus we can write $b = \alpha x^* + (1-\alpha)\text{CENTER}$, for some $\alpha \in [0,1]$ where b is a point in some face of the current pyramid. We know $\alpha = ||\text{CENTER} - b||/||\text{CENTER} - x^*||$. Using the triangle inequality we have $||\text{CENTER} - x^*|| \leq 2R_\tau = 2c_1 d r_\tau$. We also know that $||\text{CENTER} - b||$ is at least the radius of the largest ball inscribed in Π which by **??** in the Appendix is at least $r_\tau c_2^2/(2d^4)$. Using the convexity of h and Equation (16) we have

$$h(\text{CENTER}) - \frac{d(d+2)\delta}{c_2} \leq h(b) \leq \alpha h(x^*) + (1-\alpha)h(\text{CENTER}).$$

Thus, $\forall x \in \Pi$ we have

$$h(x^*) \geq h(\text{CENTER}) - \frac{d(d+1)\delta}{c_2\alpha} \geq h(\text{CENTER}) - \frac{4d^7c_1\delta}{c_2^3} \geq h(x) - \frac{4d^7c_1\delta}{c_2^3} - \frac{d(d+2)\delta}{c_2}.$$

Using the same argument as before we know that the regret incurred in the round while evaluating points in Π is no more than

$$(d+2)(\frac{4d^7c_1\delta}{c_2^3} + \frac{d(d+2)\delta}{c_2})(\frac{\kappa \ln(T/(\alpha\hat{\gamma}))}{\alpha^2\hat{\gamma}^2}).$$

Plugging in $\Delta_\tau(\hat{\gamma})$ and $\bar{\Delta}_\tau(\hat{\gamma})$ yields the result. $\square$

Lemma 22 is important because it implies that whenever we sample from a pyramid using $\hat{\gamma}$ we were in Case 2a) with $2\hat{\gamma}$ and the regret incurred is only $poly(d)/\hat{\gamma}$. The exception is when we are in the first round, however since h is 1-Lipschitz the previous claim holds trivially.

We now show that we only visit Case 1a) only a bounded number of times in every round. The intuition is that every time Case 1a) occurs and we build a new pyramid its center will be closer to the center of $\mathcal{B}(R_\tau)$ and at some point the pyramid will be inside the simplex we built at the beginning of the epoch for which we know h at its vertices.

Lemma 23. *At any round, the number of visits to Case 1a) is at most $2d^2\ln(d)/c_2^2$, and every pyramid build by the algorithm with apex y satisfies $||y - x_0|| \geq r_\tau/d$.*

Proof. By definition of Case 1a) $\text{TOP} \neq y$, without loss of generality we assume $\text{TOP} = z_1$. By construction we have

$$||z_1 - x_0|| = \sin(\phi)||y - x_0||.$$

Since this holds every time we enter Case 1a), we know that the total number of visits k satisfies

$$||z_1 - x_0|| = (\sin(\phi))^k r_\tau$$

where r_τ is the radius of the ball where the simplex is inscribed at the beginning of round τ. We also notice that for a simplex of radius r_τ the largest ball inscribed in it has radius r_τ/d. Additionally, by construction we have $\cos(\phi) = c_2/d$ and therefore $\sin(\phi) = \sqrt{1 - c_2^2/d} \leq 1 - c_2^2/(2d^2)$. Therefore, $k = 2d^2\ln(d)/c_2^2$ ensures $||z_1 - x_0|| \leq r_\tau/d$ which implies that z_1 lies inside the simplex we build at the beginning of round τ.

Let $y_1, ..., y_k$ be the apexes of the pyramids built in round τ. By construction we have

$$h(z_1) \geq h(\text{TOP}) \geq h(y_k)\gamma \geq h(y_{k-2})2\gamma \geq ... \geq h(y_1) + k\gamma.$$

On the other hand, by definition of y_1 we have $h(y_1) \geq h(x_i) - \gamma$ for all vertices of the simplex x_i. Since z_1 is in the simplex and h is convex we have

$$h(y_1) \geq h(z_1) - \gamma \geq h(y_1) + (k-1)\gamma$$

which is a contradiction unless $k \leq 1$. Therefore, if z_1 is not in the simplex it must be the case that $k \leq 2d^2\ln(d)/c_2^2$. ∎

Using the Lemma 23 we will bound the regret incurred in a round whenever it terminates in Case 2a).

Lemma 24. *For any round with CI width of γ that terminates in Case 2a) the total regret incurred in the round is no more than*

$$\frac{\kappa d\ln(T/(\alpha\gamma))}{\alpha^2\gamma}\Big(\frac{2d^2\ln(d)}{c_2^2} + 1\Big)\Big(\frac{4d^7c_1}{c_2^3} + \frac{d(d+2)}{c_2}\Big)\Big(\frac{12c_1d^4}{c_2^2} + 11\Big).$$

Proof. By Lemma 23 we have that for the given round, the total number of pyramids we have built is $k \leq 2d^2\ln(d)/c_2$. Then, by Lemma 22 we know that for any point in the k-th pyramid the instantaneous regret is no more than

$$\delta := \kappa\gamma d\Big(\frac{4d^7c_1}{c_2^3} + \frac{d(d+2)}{c_2}\Big)\Big(\frac{12c_1d^4}{c_2^2} + 11\Big).$$

We now show that the regret for any point we played during the round is at most δ. Indeed, by construction y_k is TOP of the $(k-1)$-th pyramid. By definition of Case 1a) we know that for any $x \in \Pi_{k-1}$ we have $f(x) \leq f(y_k) + \gamma$. Using this reasoning, we get that the function value at any vertex of any pyramid we have built during the round is also bounded by the function value at y_k. Additionally, as in the proof of the previous lemma, the function value at all the vertices of the simplex we built at the beginning of the epoch is also bounded by the function value at y_k. Since in every pyramid (and the initial simplex) we sample $d+2$ points we know that the total number of points we will play at is no more than $(d+2)(2d^2/(c_2^2\ln(d)) + 1)$. To bound the total number of times we play a point we notice that for a CI with width $\hat{\gamma}$ we play it $\kappa \ln(T/(\alpha\gamma))/(\alpha^2\hat{\gamma}^2)$. Suppose $\gamma = 2^{-i}$, since $\hat{\gamma}$ is geometrically decreased to γ we know that the total number of plays at any point is bounded by

$$\sum_{j=1}^{i} \frac{\kappa \ln(T/(\alpha\gamma))}{\alpha^2 2^{-2j}} \leq \frac{4\kappa \ln(T/(\alpha\gamma)) 2^{2i}}{\alpha^2} = \frac{4\kappa \ln(T/(\alpha\gamma))}{\alpha^2\gamma^2}$$

Putting everything together we get that the total regret incurred during the round is no more than

$$\frac{\kappa d \ln(T/(\alpha\gamma))}{\alpha^2\gamma}\Big(\frac{2d^2\ln(d)}{c_2^2}+1\Big)\Big(\frac{4d^7c_1}{c_2^3}+\frac{d(d+2)}{c_2}\Big)\Big(\frac{12c_1d^4}{c_2^2}+11\Big).$$

□

Using Lemma 24 we will now bound the total regret incurred at any round.

Lemma 25. *For any round that terminates in a CI with width γ, the total regret over the round is no more than*

$$\frac{\kappa d \ln(T/(\alpha\gamma))}{\alpha^2\gamma}\Big(\frac{2d^2\ln(d)}{c_2^2}+1\Big)\Big(\frac{4d^7c_1}{c_2^3}+\frac{d(d+2)}{c_2}\Big)\Big(\frac{12c_1d^4}{c_2^2}+11\Big).$$

Proof. We just need to bound the regret when the round ends in Case 1b) or 2b). By the definition of the algorithm, whenever a round has level γ it must be the case that in the previous round the level was 2γ and thus using the previous lemma we can bound the regret. The exception is in the first round when $\gamma = 1$, in this case using the Lipschitz assumption we know that the instantaneous regret is no more than 1.

Because of the previous we have that the instantaneous regret at any point of the simplex we build is no more than

$$2\gamma\Big(\frac{4d^7c_1}{c_2^3}+\frac{d(d+2)}{c_2}\Big)\Big(\frac{12c_1d^4}{c_2^2}+11\Big).$$

Now, if the algorithm was in Cases 1a), 1b) , or 2b) with level $\hat{\gamma}$, then it must have been in Case 2a) with level $2\hat{\gamma}$. And thus, using the bound on the regret whenever a round ends through Case 2a), we have that the instantaneous regret on the vertices any pyramid is no more than

$$2\hat{\gamma}\Big(\frac{4d^7c_1}{c_2^3}+\frac{d(d+2)}{c_2}\Big)\Big(\frac{12c_1d^4}{c_2^2}+11\Big),$$

and by using the same argument as in the proof of Lemma 24, the number of plays at a given point is bounded above by $\kappa \ln(T/(\alpha\gamma))/(\alpha^2\hat{\gamma}^2)$. Therefore, the total regret incurred at any pyramid built by the algorithm is no more than

$$\frac{\kappa d \ln(T/(\alpha\hat{\gamma}))}{\alpha^2\gamma}\Big(\frac{4d^7c_1}{c_2^3}+\frac{d(d+2)}{c_2}\Big)\Big(\frac{12c_1d^4}{c_2^2}+11\Big).$$

Recalling the bound on the total number of pyramids built in any round yields the result. □

Lemma 26. *The regret in any epoch which ends in level γ is at most*

$$\frac{\kappa d \ln(T/(\alpha\gamma))}{\alpha^2\gamma}\Big(\frac{2d^2\ln(d)}{c_2^2}+1\Big)\Big(\frac{4d^7c_1}{c_2^3}+\frac{d(d+2)}{c_2}\Big)\Big(\frac{12c_1d^4}{c_2^2}+11\Big).$$

Proof. From Lemma 25 we know that on any round with level γ, the regret is bounded by C/γ where C is some constant. Since γ is reduced geometrically, the net regret on an epoch where the largest level we encounter is γ is bounded by

$$\sum_{j=1}^{i} \frac{C}{2^{-j}} \leq 2C2^i = \frac{2C}{\gamma},$$

which yields the result. □

A.7.4 Bounding the Number of Epochs

To bound the number of epochs we must show that every time CONE-CUTTING is performed we discard a sufficiently large portion of the current ball. More specifically, we need to analyze the ratios of volumes of $\mathcal{B}_{\tau+1}$ and $\mathcal{B}_\tau$.

Lemma 27. *Let $\mathcal{B}_\tau$ be the smallest ball containing $\mathcal{X}_\tau$, let $\mathcal{B}'_{\tau+1}$ be the minimum volume ellipsoid containing $\mathcal{B}\tau \setminus \mathcal{K}_\tau$. Then, for small enough constants c_1, c_2, $vol(\mathcal{B}'_{\tau+1}) \leq \rho \cdot vol(\mathcal{B}_\tau)$ where $\rho = \exp(-\frac{1}{4(d+1)})$.*

Proof. This result is analogous to the volume reduction results for the ellipsoid method with a gradient oracle. It is easy to see that it suffices to consider the intersection of $\mathcal{B}_\tau$ with a half-space in order to understand the set $\mathcal{B}_\tau \setminus \mathcal{K}_\tau$. This is because if we were to discard only the spherical cap instead of the whole cone then the minimum enclosing ellipsoid would increase its volume.

The previous choices of c_1, c_2 guarantee that the distance from the center of $\mathcal{B}_\tau$ to the origin is at most $R_\tau/(4(d+1))$. The previous is true because by construction the apex of cone $\mathcal{K}_\tau$ is always contained in $\mathbb{B}(r_\tau)$, and the height of the cone is at most $R_\tau \cos(\bar{\phi}) \leq R_\tau/(8(d+1))$ again by construction. Thus, if $r_\tau \leq R_\tau/(32(d+1))$, then the distance of the hyperplane to the origin is at most $R_\tau/(4(d+1))$.
Therefore, $\mathcal{B}'_{\tau+1}$ is the minimum volume ellipsoid that contains the intersection of $\mathcal{B}_\tau$ with a hyperplane that is at most $R_\tau/(4(d+1))$ from its center. Using Theorem 2.1 from [12] (with $\alpha = -1/(4(d+1))$) we get the result. ☐

Lemma 28. *At any epoch with CI level γ, the instantaneous regret of any point in $\mathcal{K}_\tau$ is at least γ.*

Proof. Since every epoch terminates only through Cases 1b) or 2b) we only check the claim is true for these two cases. If the epoch ends through Case 1b) the proof of Lemma 20 gives the result. If the epoch ends through Case 2b), after HAT-RAISING we now that the apex y' of pyramid Π' satisfies $h(y') \geq h(z_i) + \gamma$ for all vertices $z_1, ..., z_d$ of the pyramid. Writing $y' = \alpha x + (1-\alpha)z$ with x in $\mathcal{K}_\tau$, z in the base of Π' and $\alpha \in [0,1]$, we can conclude that $h(x) \geq h(x^*) + \gamma$ just as we did in the proof of Lemma 21. ☐

We are now ready to bound the total number of epochs.

Lemma 29. *The total number of epochs in the algorithm is no more than $\frac{d\ln(T)}{\ln(1/\theta)}$ where $\theta = \exp(-\frac{1}{4(d+1)})$.*

Proof. Recall x^* is the minimizer of h. Since h is 1-Lipschitz, any point inside a ball or radius $1/\sqrt{T}$ centered around x^* has instantaneous regret of at most $1/\sqrt{T}$. The volume of this ball is $T^{-d/2}V_d$, with V_d equal to the volume of the unit ball in d-dimensions. Suppose the algorithm goes through k epochs. By Lemma 27 we know that the volume of $\mathcal{X}_\tau$ is bounded above by $\rho^k V_d$. By the previous lemma we know that the instantaneous regret of any point that was discarded had instantaneous regret at least $1/\sqrt{T}$. This is because at any given epoch and round we sample at $\frac{\kappa \ln(T/(\alpha\gamma))}{\alpha^2\gamma^2}$ and this quantity can not be more than T. Because of the previous, any point in the ball centered at x^* with radius $1/\sqrt{T}$ is never discarded. Therefore the algorithms stops whenever

$$\theta^k V_d \leq T^{-d/2} V_d$$

implying $k \leq \frac{d\ln(T)}{\ln(1/\theta)}$. ☐

We are now ready to prove Theorems 6 and 7.

Proof of Theorem 6. Using the bound on the regret incurred in an epoch and the fact that $\gamma \geq 1/\sqrt{T}$ we know the total regret on an epoch is no more than

$$\frac{\kappa d\sqrt{T}\ln(T/\alpha)}{\alpha^2}\big(\frac{2d^2\ln(d)}{c_2^2}+1\big)\big(\frac{4d^7c_1}{c_2^3}+\frac{d(d+2)}{c_2}\big)\big(\frac{12c_1d^4}{c_2^2}+11\big).$$

By the previous lemma we know the total number of epochs is no more than $d\ln(T)/\ln(1/\theta)$. Thus the total regret $T\bar{\mathcal{R}}_T$ is bounded above by

$$\frac{\kappa d^2\sqrt{T}\ln(T/\alpha)\ln(T)}{\alpha^2\ln(1/\theta)}\big(\frac{2d^2\ln(d)}{c_2^2}+1\big)\big(\frac{4d^7c_1}{c_2^3}+\frac{d(d+2)}{c_2}\big)\big(\frac{12c_1d^4}{c_2^2}+11\big).$$

Recall that we were working conditioned on $\mathcal{E}$. As in the proof of the 1-dimensional algorithm, we have $P(\mathcal{E}') \leq 1/T$. Plugging in the value of θ above yields the result. $\square$

Proof of Theorem 7. The proof is almost the same as the one of Theorem 5 with two slight differences. First, we use Theorem 6, instead of 4 to bound $\bar{\mathcal{R}}_T$. Second, using the same argument as in the proof of Theorem 5 we get that with probability at least $1 - \frac{2}{T}$, $CE = \tilde{O}(\frac{\sqrt{d}}{\alpha\sqrt{T}})$. $\square$

A.8 Analysis of Algorithm 3

The following algorithm, a generalization of Algorithm 1, will guarantee vanishing $\bar{\mathcal{R}}_T^\rho$ and $\mathcal{R}_T^\rho$ by exploiting the Kusuoka representation of risk measure ρ.

Algorithm 3

Input: $X \subset \mathbb{R}^d$, $x_1 \in X$, $z_1 \in Z$ step size η, δ
for $t = 1, ..., T$ **do**
  Sample $u \sim \mathbb{S}^{d+N}$
  Let $u_t^1 = [u_1; ...; u_d]$ and $u_t^2 = [u_{d+1}; ...; u_{d+N}]$
  Play $\tilde{x}_t := x_t + \delta u^1$, observe $f_t(\tilde{x}_t)$
  Let $\tilde{z}_t = z_t + \delta u^2$
  Let $g_t^1 := \frac{(d+N)}{\delta}(\mathcal{G}_t(\tilde{x}_t, \tilde{z}_t))u_t^1$
  Let $g_t^2 := \frac{(d+N)}{\delta}(\mathcal{G}_t(\tilde{x}_t, \tilde{z}_t))u_t^2$
  Update $x_{t+1} \leftarrow \Pi_{X_\delta}(x_t - \eta g_t^1)$
  Update $z_{t+1} \leftarrow \Pi_{Z_\delta}(z_t - \eta g_t^2)$
end for

Notice that due to Lemma 1, $g_t := [g_t^1; g_t^2]$ is a one point gradient estimator of the smoothened version of $\mathcal{G}$, $\hat{\mathcal{G}}$.

The proofs of Theorems 8 and 9 will be similar to that of Theorems 2 and 3, however we must be careful to make sure we do not introduce unnecessary factors of N, d and $\frac{1}{\alpha}$.

Lemma 30. $||\nabla\mathcal{G}|| \leq N(G+1)+1$

Proof.

$$\begin{aligned}
||\nabla\mathcal{G}|| &= \sqrt{\sum_{i=1}^d (\sum_{n=1}^N \mu_n \nabla_{x_i}\mathcal{L}_n)^2 + \sum_{n=1}^N (\mu_n \nabla_{z_n}\mathcal{L}_n)^2} \\
&\leq \sqrt{\sum_{i=1}^d (||\mu||_1 ||\nabla_{x_i}\mathcal{L}_n||_\infty)^2 + \sum_{n=1}^N (\mu_n \nabla_{z_n}\mathcal{L}_n)^2} \quad ||.||_\infty \text{ is over n=1,...,N} \\
&\leq \sqrt{\sum_{i=1}^d ||\nabla_{x_i}\mathcal{L}_n||_\infty^2 + \sum_{n=1}^N \mu_n \nabla_{z_n}\mathcal{L}_n^2} \quad \text{since } \sum_{n=1}^N \mu_n = 1, \text{ and } \mu_i \leq 1 \\
&\leq \sqrt{\sum_{i=1}^d ||\nabla_{x_i}\mathcal{L}_n||_\infty^2 + \sum_{n=1}^N \mu_n (1+N)^2} \\
&\leq \sqrt{\sum_{i=1}^d ||\nabla_{x_i}\mathcal{L}_n||_\infty^2} + \sqrt{\sum_{n=1}^N \mu_n (1+N)^2} \\
&\leq \sqrt{\sum_{i=1}^d ||N\nabla_{x_i} f||_\infty^2} + \sqrt{\sum_{n=1}^N \mu_n (1+N)^2} \\
&\leq NG + (1+N) \quad \text{since } \sum_{n=1}^N \mu_n = 1
\end{aligned}$$

□

Lemma 31. *Running online gradient descent on $\{\mathcal{G}_t\}_{t=1}^T$ ensures that for all $x \in X$ and all $z \in Z$*

$$2[\sum_{t=1}^T \mathcal{G}_t(x_t, z_t) - \sum_{t=1}^T \mathcal{G}_t(x, z)] \leq \frac{||x_T - x^*||^2 + \sum_{n=1}^d \mu_n ||z_{t,n} - z_n^*||^2}{\eta} + \eta[\sum_{t=1}^T (||\nabla_x \mathcal{G}_t(x_t, y_t) + \sum_{n=1}^N \mu_n |\nabla_{z_n} \mathcal{L}(x_t, z_t)|^2)].$$

Proof.

$$\begin{aligned}
&2[\sum_{t=1}^T \mathcal{G}_t(x_t, z_t) - \sum_{t=1}^T \mathcal{G}_t(x, z)] \\
&\leq 2\sum_{t=1}^T \nabla \mathcal{G}_t(x_t, z_t)^\top ([x_t; z_t] - [x; z]) \\
&= 2\sum_{t=1}^T \nabla_x \mathcal{G}_t(x_t, z_t)^\top (x_t - x) + 2\sum_{t=1}^T \sum_{n=1}^d \mu_n \nabla_z \mathcal{L}(x_t, z_t)(z_{t,n}.z_n) \\
&\leq \frac{||x_T - x||^2}{\eta} + \sum_{n=1}^N \mu_n \frac{||z_{T,n} - z_n||^2}{\eta} + \eta[\sum_{t=1}^T (||\nabla_x \mathcal{G}_t|| + \sum_{n=1}^d \mu_n (\nabla_z \mathcal{L}_{t,n})^2)] \quad \text{by Equations 10 and 11}
\end{aligned}$$

□

Lemma 32. *Let $y^* = (x^*, z^*) \in \arg\min \mathbb{E}_\xi[\sum_{t=1}^T \mathcal{G}_t(x, z)]$. With appropriate choice of parameters η, δ we have*

$$\mathbb{E}_{int}[\sum_{t=1}^T \mathcal{G}_t(\tilde{y}_t)] - \sum_{t=1}^T \mathcal{G}_t(y^*) \leq O(dN^{3/2}T^{3/4})$$

Proof. First we need a bound on $\sum_{t=1}^T \mathcal{G}_t(y_\delta^*) - \sum_{t=1}^T \mathcal{G}_t(y^*)$, where $y_\delta^* = \Pi_{X_\delta \times Z_\delta}(y^*)$. If $\mathcal{G}$ is Lipschitz L with respect to some norm $||\cdot||$, by Lemma 5 we have $||\nabla \mathcal{G}||_* \leq L$. For any $y = [x; z]$ with $x \in X$ and $z \in Z$, let us use $||y|| = ||x||_2 + ||z||_\infty$ with dual norm $||y||_* = \max\{||x||_2, ||z||_1\}$ (see Lemma 12 in the Appendix).

$$\begin{aligned}
\sum_{t=1}^T \mathcal{G}_t(y_\delta^*) - \sum_{t=1}^T \mathcal{G}_t(y^*) &\leq TL||y^* - y_\delta^*|| \\
&\leq \delta TLD_{\mathcal{G}}^{||\cdot||} \quad \text{by Lemma 11 in the Appendix} \\
&\leq O(\delta TGN).
\end{aligned}$$

The last inequality holds because of the following two facts, 1) $||\nabla \mathcal{G}||_* = \max\{||\nabla_x \mathcal{G}||_2, ||\nabla_z \mathcal{G}||_1\} \leq \max\{G, \sum_{n=1}^N \mu[1 + N]\} \leq G + 1 + N$ and 2) $||y_1 - y_2|| = ||x_1 - x_2||_2 + ||z_1 - z_2||_\infty \leq D_X + 2 := D_{\mathcal{G}}^{||\cdot||}$. Let $\mathbb{E}_{int}$ be the expectation taken with respect to the internal randomization of the algorithm. Following the proof of Lemma 15 we have

$$
\begin{aligned}
&\mathbb{E}_{int}[\sum_{t=1}^{T} \mathcal{G}_t(\tilde{y}_t)] - \sum_{t=1}^{T} \mathcal{G}_t(y^*) \\
&\leq \mathbb{E}_{int}[\sum_{t=1}^{T} \mathcal{G}_t(y_t)] - \sum_{t=1}^{T} \mathcal{G}_t(y^*) + \delta G_{\mathcal{G}} T \quad \mathcal{G} \text{ is } G_{\mathcal{G}}\text{-Lipschitz and } ||y - \tilde{y}|| \leq \delta \\
&\leq \mathbb{E}_{int}[\sum_{t=1}^{T} \mathcal{G}_t(y_t)] - \sum_{t=1}^{T} \mathcal{G}_t(y^*_\delta) + \delta G_{\mathcal{G}} T + O(\delta T G N) \\
&\leq \mathbb{E}_{int}[\sum_{t=1}^{T} \hat{\mathcal{G}}_t(y_t)] - \sum_{t=1}^{T} \hat{\mathcal{G}}_t(y^*_\delta) + 3\delta G_{\mathcal{G}} T + \delta D_{\mathcal{G}} G_{\mathcal{G}} T \quad |\mathcal{G}(y) - \hat{\mathcal{G}}(y)| < \delta G_{\mathcal{G}} \\
&\leq \frac{||x_T - x^*||_2^2}{2\eta} + \sum_{n=1}^{N} \mu_n \frac{||z_{t,n} - z_n^*||_2^2}{2\eta} + \mathbb{E}_{int}[2\eta[\sum_{t=1}^{T}(||g_t^1||_2 + \sum_{n=1}^{d} \mu_n (g_{t,n}^2)^2)]] + 3\delta G_{\mathcal{G}} T + O(\delta T G N) \\
&\quad \text{reduction to bandit feedback and Lemma 31} \\
&\leq \frac{D_X^2 + 2}{2\eta} + 2\eta \mathbb{E}_{int}[\sum_{t=1}^{T}(||g_t^1||_2^2 + \sum_{n=1}^{d} \mu_n (g_{t,n}^2)^2)] + 3\delta G_{\mathcal{G}} T + O(\delta T G N) \\
&\leq \frac{D_X^2 + 2}{2\eta} + 2\eta \frac{(d+N)^2 N^2}{\delta^2} T + 3\delta G_{\mathcal{G}} T + O(\delta T G N) \\
&\leq O(d N^{3/2} T^{3/4})
\end{aligned}
$$

where we chose $\eta = O(\frac{1}{dN^{3/2}T^{3/4}})$ and $\delta = O(\frac{N^{1/2}}{T^{1/4}})$ and plugged in the bound on $G_{\mathcal{G}}$ from Lemma 30. □

Proof of Theorem 8. Take $\mathbb{E}_\xi[\cdot]$ on both sides of the result in Lemma 32 and interchange the expectations (this can be done using Fubini's Theorem and the uniform bound on $\mathcal{G}_t$). Noting that for all $x \in X$ and all $z \in [0, 1]$ (in particular for every $(\tilde{x}_t, \tilde{z}_t)$) we have

$$
\mathbb{E}_{\xi \sim P}[\mathcal{L}_n^t(x, z)] = z + \frac{1}{n/N} \mathbb{E}_{\xi \sim P}[f(x, \xi) - z]_+ \geq CVaR_{n/N}[F](x),
$$

it follows that since $\mathcal{G}_t(x, z) := \sum_{n=1}^{N} \mu_n \mathcal{L}_n^t(x, z)$ we have $\mathbb{E}_{\xi \sim P}[\mathcal{G}_t(x, z)] \geq \rho[F](x)$. Noting that $\mathbb{E}_\xi[\sum_{t=1}^{T} \mathcal{G}_t(y^*)] = \min_{x \in X} \rho[F](x)$ we get the desired result. □

Proof of Theorem 9. We notice that strong convexity of $f(\cdot, \xi)$ implies strong convexity of $\rho[F](\xi)$ since each of the $C_{\alpha_i}[F](\cdot)$ in the Kusuoka representation of $\rho[F]$ is strongly convex. Let $x^* = argmin_{x \in X} \rho[F](x)$. We follow the proof of Theorem 3. Let the concentration error $CE = \rho[\{f_t(x^*)\}_{t=1}^T] - \min_{x\in X}\rho[\{f_t(x)\}_{t=1}^T]$.

$$\begin{aligned}
&\mathbb{E}[\rho[\{f_t(x_t)\}] - \min_{x\in X}\rho[\{f_t(x)\}]] \\
&= \mathbb{E}[\rho[\{f_t(x_t)\}] \pm \rho[\{f_t(x^*)\}] - \min_{x\in X}\rho[\{f_t(x)\}]] \\
&= \mathbb{E}[\sum_{n=1}^N \mu_n C_{n/N}[\{f_t(x_t)\}] - \rho[\{f_t(x^*)\}]] + \mathbb{E}[CE] \\
&\leq \mathbb{E}[\frac{N}{T}\sum_{t=1}^T |f_t(x_t) - f_t(x^*)|] + \mathbb{E}[CE] \quad \text{as in the last line of the proof of Theorem 3} \\
&\leq \frac{N}{T}\sum_{t=1}^T \mathbb{E}_t[||x_t - x^*||] + \mathbb{E}[CE] \\
&\leq \frac{N}{T}\sqrt{T}\sqrt{\sum_{t=1}^T \mathbb{E}_t[||x_t - x^*||^2]} + \mathbb{E}[CE] \\
&\leq \frac{N}{T}\sqrt{T}\sqrt{\sum_{t=1}^T \frac{2}{\beta}\mathbb{E}[\rho[F](x_t) - \rho[F](x^*)]} + \mathbb{E}[CE] \\
&\leq O(\frac{d^{1/2}N^{7/4}}{\beta^{1/2}T^{1/8}}) + \mathbb{E}[CE] \quad \square
\end{aligned}$$

The expectation of the concentration error can be bounded as in the proof of Theorem 3 by $\tilde{O}(\frac{N^{3/2}\sqrt{d}}{\sqrt{T}})$. This yields the result. □

A.9 Analysis of Algorithm 4

Recall Algorithm 4 is the modification of Algorithm 2 where we sample $\tilde{O}(\frac{N^2\ln(NT)}{\gamma^2})$ times a point (instead of $O(\frac{\ln(T/(\alpha\gamma))}{\alpha^2\gamma^2})$) to build a γ-CI. In this section we present the proofs of Theorems 10 and 11. We only need to show that $\tilde{O}(\frac{N^2\ln(NT)}{\gamma^2})$ samples are sufficient to build a γ-CI that holds with high probability. Afterwards it is easy to verify that the proofs of Theorems 6 and 7 go through.

Lemma 33. *To build a γ-CI for $\rho[F](x)$ that holds with probability at least $1 - \frac{1}{T^2}$ we need no more than $O(\frac{N\ln(N)\ln(\sqrt{N}T)}{\gamma^2})$ samples.*

Proof. Notice that

$$|\rho[X] - \hat{\rho}[X]| = |\sum_{n=1}^N \mu_n(C_{n/N}[X] - \hat{C}_{n/N}[X])| \leq \sum_{n=1}^N \mu_n|C_{n/N}[X] - \hat{C}_{n/N}|$$

Therefore, if we obtain γ-CI's for each term $|C_{n/N}[X] - \hat{C}_{n/N}|$ that hold with probability at least $1 - \frac{1}{NT^2}$ a union bound yields the result. From Theorem 1 we know that $O(\frac{N^2\ln(\sqrt{N}T)}{n^2\gamma^2})$ samples suffice to build a γ-CI for $C_{n/N}[X]$ that holds probability at least $1 - \frac{1}{NT^2}$. Summing up the number of samples, approximating the sum with an integral and using a union bound yields the result. □

We are now ready to prove the theorems.

Proof of Theorem 10. It is easy to see that the proof of Theorem 6 goes through if we set $h(\cdot) = \rho[F](\cdot)$ and we replace everywhere the number of times we sample a point $O(\frac{\ln(T/(\alpha\gamma))}{\alpha^2\gamma^2})$ with $\tilde{O}(\frac{N^2\ln(T)}{\gamma^2})$. □

Proof of Theorem 11. The proof follows from almost the same reasoning as in the proof of Theorem 7. We have

$$\begin{aligned}
&\rho[\{f_t(x_t)\}_{t=1}^T] - \min_{x \in X} \rho[\{f_t(x)\}_{t=1}^T] \\
&\leq \frac{N}{T}\sqrt{T}\sqrt{\frac{2}{\beta}\sum_{t=1}^{T} C_\alpha[F](x_t) - C_\alpha[F](x^*)} + CE \\
&\leq O(\frac{d^8 N^3}{\beta^{1/2} T^{1/4}}) + CE \quad \text{(with probability at least } 1 - \frac{1}{T})
\end{aligned}$$

where $CE = \rho[F](x^*) - \min_{x \in X} \rho[\{f_t(x)\}]$ and $x^* = argmin_{x \in X} \rho[F](x)$. Just as in the proof of Theorem 3 we can bound CE with probability at least $1 - \frac{2}{T}$ by $\tilde{O}(\frac{N^3/2\sqrt{d}}{\sqrt{T}})$. A union bound yields the result. □