Supplementary: Pseudo-1d Bandit Convex Optimization

A. Proofs

A.1. Proof of Theorem 1

Proof. **Problem instance construction.** Divide the time interval $[T]$ into d equal length sub intervals (hence each of length $\frac{T}{d}$) $T_1, \ldots, T_d$. Assume $T_0 = \emptyset$.

For $i \in [d]$: Choose $\sigma_i \sim \text{Ber}(\pm 1)$, and set $\mathbf{x}_i = \mathbf{e}_i$. Denote $\boldsymbol{\sigma} = (\sigma_1, \ldots, \sigma_d)$.

At any time $t \in T_i = \left\{ \frac{T}{d}(i-1)+1, \ldots, \frac{T}{d}i \right\}$, $i \in [d]$,

1. Choose $g_t(\mathbf{w}; \mathbf{x}_t) = \mathbf{w}^\top \mathbf{x}_t$. Clearly $\nabla_{\mathbf{w}}(g_t(\cdot; \mathbf{x}_t)) = \mathbf{x}_t \in \{0,1\}^d$ which is revealed to the learner at the beginning of round t. We choose $\mathbf{x}_t = \mathbf{x}_i$.
2. Loss function $f_t(\mathbf{w}) = \ell_t(\mathbf{w}^\top \mathbf{x}_t) + \varepsilon_t = \mu\sigma_i(\mathbf{w}^\top \mathbf{x}_i) + \varepsilon_t$, where $\varepsilon_t \sim \mathcal{N}(0, \frac{1}{16})$, for some constant $\mu > 0$ (to be decided later), $\forall \mathbf{w} \in \mathcal{W}$.
3. Learner plays $\mathbf{w}_t = [\mathbf{w}_t(1), \ldots, \mathbf{w}_t(d)] \in \mathcal{W}$.

Denote $\bar{\mathbf{w}}_i := \frac{1}{T_d}\sum_{t \in T_i} \mathbf{w}_t$, where $T_d = \frac{T}{d}$.

Remark 10 (Optimum Point)**.** *Note for any fixed $\mathbf{w} \in \mathcal{W}$, the total expected loss is $\mathbb{E}\Big[\sum_{i=1}^d \sum_{t \in T_i} f_t(\mathbf{w})\Big] = \frac{\mu T}{d}\sum_{i=1}^d (\sigma_i \mathbf{x}_i^\top)\mathbf{w} = \frac{T}{d}(\tilde{\boldsymbol{\sigma}}^\top \mathbf{w})$, where $\tilde{\sigma}(i) = \mu\sigma_i,\ \forall i \in [d]$. Thus clearly the best point (i.e. the minimizer) $\mathbf{w}^* = -\frac{\boldsymbol{\sigma}}{\sqrt{d}}$. Note $\mathbf{w}^* \in \mathcal{W}$.*

The expected regret of any $\mathcal{A}$:

$$
\begin{aligned}
\mathbb{E}[R_T] &= \sum_{i=1}^d \sum_{t \in T_i} \mu[(\sigma_i \mathbf{x}_i^\top)\mathbf{w}_t - (\sigma_i \mathbf{x}_i^\top)\mathbf{w}^*] = \sum_{i=1}^d \mu T_d \left[\mathbb{E}[\sigma_i \mathbf{x}_i^\top \bar{\mathbf{w}}_i] - (\sigma_i \mathbf{x}_i^\top)\mathbf{w}^*\right] \\
&= \sum_{i=1}^d T_d \mathbb{E}\Big[\mu\sigma(i)[\bar{\mathbf{w}}_i(i) - \mathbf{w}^*(i)]\Big] \\
&= \sum_{i=1}^d T_d \mathbb{E}\Big[\mu\sqrt{d}\mathbf{w}^*(i)[\mathbf{w}^*(i) - \bar{\mathbf{w}}_i(i)]\Big] = \sum_{i=1}^d T_d \mathbb{E}\Big[\mu\sqrt{d}\Big((\mathbf{w}^*(i))^2 - \bar{\mathbf{w}}_i(i)\mathbf{w}^*(i)\Big)\Big] \\
&= \sum_{i=1}^d T_d \mathbb{E}\Big[\mu\sqrt{d}\Big(\frac{1}{d} + \frac{\sigma_i}{\sqrt{d}}\bar{\mathbf{w}}_i(i)\Big)\Big] \\
&= \sum_{i=1}^d T_d \left[\frac{2\mu}{\sqrt{d}} Pr(\sigma_i \bar{\mathbf{w}}_i(i) > 0)\right] \text{ since } \bar{\mathbf{w}}_i(i) \in \{-1/\sqrt{d}, 1/\sqrt{d}\}
\end{aligned} \tag{7}
$$

Now for any $i \in [d]$:

$$
\begin{aligned}
Pr\big(\sigma_i \bar{\mathbf{w}}_i(i) > 0\big) &= \frac{1}{2}Pr\big(\bar{\mathbf{w}}_i(i) > 0 \mid \sigma_i = +1\big) + \frac{1}{2}Pr\big(\bar{\mathbf{w}}_i(i) < 0 \mid \sigma_i = -1\big) \\
&= \frac{1}{2}\Big(Pr\big(\bar{\mathbf{w}}_i(i) > 0 \mid \sigma_i = +1\big) + 1 - Pr\big(\bar{\mathbf{w}}_i(i) > 0 \mid \sigma_i = -1\big)\Big) \\
&\geq \frac{1}{2}\Big(1 - |Pr\big(\bar{\mathbf{w}}_i(i) > 0 \mid \sigma_i = +1\big) - Pr\big(\bar{\mathbf{w}}_i(i) > 0 \mid \sigma_i = -1\big)|\Big),
\end{aligned}
$$

Assumption 1. *For proving the lower bound we assume that $\bar{\mathbf{w}}_i(i)$ is a deterministic function of the observed function values $\{f_t\}_{t\in T_i}$, respectively at $\{\mathbf{w}_t\}_{t\in T_i}$. Note that this assumption is without loss of generality, since any random querying strategy can be seen as a randomization over deterministic querying strategies. Thus, a lower bound which holds uniformly for any deterministic querying strategy would also hold over a randomization. Let us denote: $f([T_i]) = \{f_t\}_{t\in T_i}$.*

Then since the randomness of $\bar{\mathbf{w}}_i(i)$ only depends on $f([T_i])$, applying Pinsker's inequality, we get:

$$\begin{aligned} Pr\big(\sigma_i\bar{\mathbf{w}}_i(i) > 0\big) &\geq \frac{1}{2}\Big(1 - \big|Pr\big(\sigma_i\bar{\mathbf{w}}_i(i) > 0 \mid \sigma_i = +1\big) - Pr\big(\sigma_i\bar{\mathbf{w}}_i(i) < 0 \mid \sigma_i = -1\big)\big|\Big) \\ &\geq \frac{1}{2}\bigg(1 - \sqrt{2KL\Big(P(f([T_i]) \mid \sigma_i = +1)||P(f([T_i]) \mid \sigma_i = -1\Big)}\bigg) \end{aligned}$$

and further applying the chain rule of KL-divergence, we have:

$$\begin{aligned} Pr\big(\sigma_i\bar{\mathbf{w}}_i(i) > 0\big) &\geq \frac{1}{2}\bigg(1 - \sqrt{2\sum_{t\in T_i} KL\Big(P(f_t \mid \sigma_i = +1, \{f_\tau\}_{\tau\in[t-1]\setminus T_{i-1}})||P(f_t \mid \sigma_i = -1, \{f_\tau\}_{\tau\in[t-1]\setminus T_{i-1}}\Big)}\bigg) \\ &\geq \frac{1}{2}\bigg(1 - \sqrt{2\sum_{t\in T_i}\frac{4\mu^2\sigma_i^2\mathbf{w}_t(i)^2}{\frac{2}{16}}}\bigg) = \frac{1}{2}\bigg(1 - \sqrt{\frac{64\mu^2 T_d}{d}}\bigg) \text{ since } \mathbf{w}_t(i)^2 = \frac{1}{d} \text{ and } \sigma_i^2 = 1 \end{aligned}$$

where the last inequality follows by noting $P(f_t \mid \sigma_i, \{f_\tau\}_{\tau\in[t-1]\setminus T_{i-1}}) \sim \mathcal{N}(\mu\sigma_i\mathbf{w}_t(i), \frac{1}{16})$, and $KL(\mathcal{N}(\mu_1,\sigma^2)||\mathcal{N}(\mu_2,\sigma^2)) = \frac{(\mu_1-\mu_2)^2}{2\sigma^2}$ (for bounding the each individual KL-divergence terms).

Case 1 ($d \leq 16\sqrt{T}$)

Combining the above claims with Eq. (7):

$$\begin{aligned} \mathbb{E}[R_T] = \sum_{i=d} T_d\left[\frac{2\mu}{\sqrt{d}}Pr(\sigma_i\bar{\mathbf{w}}_i(i) > 0)\right] &\geq \sum_{i=d} T_d\left[\frac{\mu}{\sqrt{d}}\Big(1 - 8\mu\sqrt{\frac{T_d}{d}}\Big)\right], \\ &\geq \sum_{i=d} T_d\frac{1}{16\sqrt{T}_d}\Big(1 - \frac{1}{2}\Big)\left(\text{setting } \mu = \frac{\sqrt{d}}{16\sqrt{T}_d} \leq 1\right) = \frac{\sqrt{dT}}{32}. \end{aligned}$$

Note that for any $t \in [T]$, f_t s are 1-lipschitz for $d \leq 16\sqrt{T}$, as desired to understand the dependency of lower bound to the lipschitz constant.

Case 2 ($d > 16\sqrt{T}$)

In this case $T < \frac{d^2}{256}$. Let us denote $d' = 16\sqrt{T} < d$, and let us use the above problem construction for dimension d' (we can simply ignore decision coordinates $\mathbf{w}(d'+1), \ldots, \mathbf{w}(d)$, i.e. for any $\mathbf{w} \in \mathcal{W} \subseteq \mathbb{R}^d$, denoting $\mathbf{w}_{[d']} = (\mathbf{w}_1, \ldots, \mathbf{w}_{d'})$, we can construct $f_t(\mathbf{w}) = f_t(\mathbf{w}_{[d']})$).

Now for the above problem suppose there exists an algorithm $\mathcal{A}$ such that $\mathbb{E}[R_T(\mathcal{A})] \leq \frac{\sqrt{d'T}}{32} = \frac{T^{3/4}}{32}$, then this violates the lower bound derived in **Case** 1. Thus the lower bound for **Case** 2 is must be at least $\frac{T^{3/4}}{32}$.

Combining the lower bounds of **Case** 1 **and** 2 concludes the proof. □

A.2. Proof of Lemma 5 and additional claims

Useful definitions and notation. Before proceeding to the proof, we define relevant notation that will be used throughout this section. For the kernel $\mathbf{K}'_t$ (Definition 4), we define a linear operator $\mathbf{K}_t^{'*}$ on the space of functions $\mathcal{G}_t \mapsto \mathbb{R}$ as follows. For any function $\ell : \mathcal{G}_t \mapsto \mathbb{R}$:

$$\mathbf{K}_t^{'*}\ell(y) := \int_{y'\in\mathcal{G}_t} \ell(y')\mathbf{K}'_t(y', y)dy \quad \forall y \in \mathcal{G}_t, \tag{8}$$

We also denote by $\mathcal{P}$ and $\mathcal{Q}_t$ the set of all probability measures on $\mathcal{W}$ and $\mathcal{G}_t$ respectively; and by $\boldsymbol{\delta}_y \in \mathcal{Q}_t$, $\boldsymbol{\delta}_{\mathbf{w}} \in \mathcal{P}$ the dirac mass at $y \in \mathcal{G}_t$ and at $\mathbf{w} \in \mathcal{W}$ respectively. For $\mathbf{q} \in \mathcal{Q}_t$, define:

$$\big\langle \mathbf{q}, \ell \big\rangle = \int_{y \in \mathcal{G}_t} \ell(y)\mathbf{q}(y)dy$$

As noted in (Bubeck et al., 2017), a useful observation on the operator (8) is that for any $\mathbf{q} \in \mathcal{Q}_t$:

$$\big\langle \mathbf{K}_t'\mathbf{q}, \ell_t \big\rangle = \big\langle {\mathbf{K}_t'}^{*}\ell_t, \mathbf{q} \big\rangle. \tag{9}$$

Proof of Lemma 5.

Proof. For ease, we abbreviate $g_t(\mathbf{w}_t; \mathbf{x}_t)$ as $g_t(\mathbf{w}_t)$ throughout the proof. We start by analyzing the expected regret w.r.t. the optimal point $\mathbf{w}^* \in \mathcal{W}$ (denote $y_t^* = g_t(\mathbf{w}^*)$ for all $t \in [T]$). Define $\forall y \in \mathcal{G}_t,\ \tilde{\ell}_t(y) := \tilde{f}_t(\mathbf{w})$, for any $\mathbf{w} \in \mathcal{W}(y)$. Also let $\mathcal{H}_t = \boldsymbol{\sigma}\big(\{\mathbf{x}_\tau, \mathbf{p}_\tau, \mathbf{w}_\tau, f_\tau\}_{\tau=1}^{t-1} \cup \{\mathbf{x}_t, \mathbf{p}_t\}\big)$ denote the sigma algebra generated by the history till time t. Then the expected cumulative regret of Algorithm 2 over T time steps can be bounded as:

$$\begin{aligned}
&\mathbb{E}[R_T(\mathbf{w}^*)] := \mathbb{E}\bigg[\sum_{t=1}^{T}\Big(f_t(\mathbf{w}_t) - f_t(\mathbf{w}^*)\Big)\bigg] = \mathbb{E}\bigg[\sum_{t=1}^{T}\Big(\ell_t(g_t(\mathbf{w}_t)) - \ell_t(g_t(\mathbf{w}^*))\Big)\bigg]\\
&= \mathbb{E}\bigg[\sum_{t=1}^{T}\Big(\ell_t(y_t) - \ell_t(y_t^*)\Big)\bigg] = \mathbb{E}\bigg[\sum_{t=1}^{T}\big\langle \mathbf{K}_t'\mathbf{q}_t - \boldsymbol{\delta}_{y_t^*}, \ell_t\big\rangle\bigg] \ \text{[since } y_t \sim K_t'\mathbf{q}_t]\\
&\le \mathbb{E}\bigg[\sum_{t=1}^{T}\frac{3\epsilon L}{\lambda} + \frac{1}{\lambda}\big\langle \mathbf{K}_t'(\mathbf{q}_t - \boldsymbol{\delta}_{y_t^*}), \ell_t\big\rangle\bigg] \ \big[\text{from Property\#2 of Lemma 11}\big]\\
&\le 6\epsilon LT + 2\sum_{t=1}^{T}\mathbb{E}\Big[\big\langle \mathbf{K}_t'(\mathbf{q}_t - \boldsymbol{\delta}_{y_t^*}), \ell_t\big\rangle\Big] \ \big[\text{we can choose } \lambda = 1/2, \text{ see proof of Lemma 11}\big]\\
&\overset{\text{by (9)}}{=} 6\epsilon LT + 2\sum_{t=1}^{T}\mathbb{E}\bigg[\sum_{t=1}^{T}\big\langle {\mathbf{K}'}_t^{*}\ell_t, (\mathbf{q}_t - \boldsymbol{\delta}_{y_t^*})\big\rangle\bigg]\\
&\overset{(a)}{=} 6\epsilon LT + 2\sum_{t=1}^{T}\mathbb{E}\bigg[\sum_{t=1}^{T}\mathbb{E}_{y_t\sim \mathbf{K}_t'\mathbf{q}_t}\Big[\big\langle \mathbf{q}_t - \boldsymbol{\delta}_{y_t^*}, \tilde{\ell}_t\big\rangle \mid \mathcal{H}_t\Big]\bigg]\\
&= 6\epsilon LT + 2\sum_{t=1}^{T}\mathbb{E}\bigg[\sum_{t=1}^{T}\mathbb{E}_{y_t\sim \mathbf{K}_t'\mathbf{q}_t}\Big[\big\langle \mathbf{p}_t - \boldsymbol{\delta}_{\mathbf{w}^*}, \tilde{f}_t\big\rangle \mid \mathcal{H}_t\Big]\bigg]
\end{aligned} \tag{10}$$

where the last equality follows by Lemma 9, and by $\big\langle \boldsymbol{\delta}_{\mathbf{w}^*}, \tilde{f}_t\big\rangle = \tilde{f}_t(\mathbf{w}^*) = \tilde{\ell}_t(y_t^*) = \big\langle \boldsymbol{\delta}_{y_t^*}, \tilde{\ell}_t\big\rangle$; the penultimate equality (a) follows noting that for any $y' \in \mathcal{G}_t$:

$$\mathbb{E}_{y_t\sim \mathbf{K}_t'\mathbf{q}_t}[\tilde{\ell}_t(y')] = \int_{y_t\in\mathcal{G}_t}\mathbf{K}_t'\mathbf{q}_t(y_t)\frac{\ell_t(y_t)}{\mathbf{K}_t'\mathbf{q}_t(y_t)}\mathbf{K}_t'(y_t, y')dy_t = \int_{y_t\in\mathcal{G}_t}\ell_t(y_t)\mathbf{K}_t'(y_t, y')dy_t = {\mathbf{K}_t'}^{*}\ell_t(y').$$

Let us denote by $\mathbf{p}^*$ a uniform measure on the set $\mathcal{W}_\kappa := \{\mathbf{w} \mid \mathbf{w} = (1-\kappa)\mathbf{w}^* + \kappa\mathbf{w}', \text{ for any } \mathbf{w}' \in \mathcal{W}\}$ for some $\kappa \in (0,1)$. Note, this implies $\mathbf{p}^*(\mathbf{w}) = \begin{cases}\frac{1}{\kappa^d \text{vol}(\mathcal{W})}, & \text{if } \mathbf{w} \in \mathcal{W}_\kappa\\ 0 \ \text{ otherwise}\end{cases}$.

Then note that:

$$\begin{aligned}
\sum_{t=1}^{T}\mathbb{E}_{y_t\sim \mathbf{K}_t'\mathbf{q}_t}\big\langle \mathbf{p}_t - \boldsymbol{\delta}_{\mathbf{w}^*}, \tilde{f}_t\big\rangle &= \sum_{t=1}^{T}\mathbb{E}_{y_t\sim \mathbf{K}_t'\mathbf{q}_t}[\big\langle \mathbf{p}_t, \tilde{f}_t\big\rangle - \big\langle \boldsymbol{\delta}_{\mathbf{w}^*}, \tilde{f}_t\big\rangle]\\
&\overset{(a)}{=} \sum_{t=1}^{T}\mathbb{E}_{y_t\sim \mathbf{K}_t'\mathbf{q}_t}[\big\langle \mathbf{p}_t, \tilde{f}_t\big\rangle] - {\mathbf{K}_t'}^{*}\ell_t(g_t(\mathbf{w}^*))
\end{aligned}$$

$$\overset{(b)}{\leq} \sum_{t=1}^{T} \mathbb{E}_{y_t \sim \mathbf{K}_t' \mathbf{q}_t}[\langle \mathbf{p}_t, \tilde{f}_t \rangle] + \sum_{t=1}^{T} \Big[\kappa LDW - \langle \mathbf{p}^*, \mathbf{K}_t'^* \ell_t(g_t(\cdot)) \rangle \Big]$$
$$= \sum_{t=1}^{T} \mathbb{E}_{y_t \sim \mathbf{K}_t' \mathbf{q}_t}[\langle \mathbf{p}_t, \tilde{f}_t \rangle - \langle \mathbf{p}^*, \tilde{f}_t \rangle] + \kappa LDWT$$

where (a) follows since $\mathbb{E}_{y_t \sim \mathbf{K}_t'\mathbf{q}_t} \langle \boldsymbol{\delta}_{\mathbf{w}^*}, \tilde{f}_t \rangle = \mathbb{E}_{y_t \sim \mathbf{K}_t'\mathbf{q}_t}[\tilde{f}_t(\mathbf{w}^*)] = \mathbb{E}_{y_t \sim \mathbf{K}_t'\mathbf{q}_t}[\tilde{\ell}_t(g_t(\mathbf{w}^*))] = \mathbf{K}_t'^* \ell_t(g_t(\mathbf{w}^*))$ as shown above; (b) follows since by assumption g_t is D lipschitz and so by definition of $\mathcal{W}_\kappa$ for any $\mathbf{w} \in \mathcal{W}_\kappa$ we have $|g_t(\mathbf{w}) - g_t(\mathbf{w}^*)| \leq DW$ (since $W = \text{Diam}(\mathcal{W})$). But from the Property #1 of Lemma 11 we have that the function $\mathbf{K}_t'^* \ell_t(\cdot)$ is L-lipschitz, which in turn implies for any $\mathbf{w} \in \mathcal{W}_\kappa$, $|\mathbf{K}_t'^* \ell_t(g_t(\mathbf{w})) - \mathbf{K}_t'^* \ell_t(g_t(\mathbf{w}^*))| \leq L|g_t(\mathbf{w}) - g_t(\mathbf{w}^*)| \leq \kappa LDW$. The last equality follows by applying the reverse logic used for (a).

Combining above claims with (10) we further get:

$$\mathbb{E}[R_T(\mathbf{w}^*)] \leq 6\epsilon LT + 2\Bigg(\kappa LDWT + \mathbb{E}\Bigg[\sum_{t=1}^{T} \mathbb{E}_{y_t \sim \mathbf{K}_t'\mathbf{q}_t} \Big[\langle \mathbf{p}_t - \mathbf{p}^*, \tilde{f}_t \rangle \mid \mathcal{H}_t \Big] \Bigg] \Bigg). \tag{11}$$

From Lemma 10 we get:

$$\sum_{t=1}^{T} \langle \mathbf{p}_t - \mathbf{p}^*, \tilde{f}_t \rangle \leq \frac{KL(\mathbf{p}^* || \mathbf{p}_1)}{\eta} + \frac{\eta}{2} \langle \mathbf{p}_t, \tilde{f}_t^2 \rangle = \frac{KL(\mathbf{p}^* || \mathbf{p}_1)}{\eta} + \frac{\eta}{2} \langle \mathbf{q}_t, \tilde{\ell}_t^2 \rangle, \tag{12}$$

where the equality $\langle \mathbf{p}_t, \tilde{f}_t^2 \rangle = \langle \mathbf{q}_t, \tilde{\ell}_t^2 \rangle$ follows from a similar derivation as shown in Lemma 9. Now, note that:

$$\mathbb{E}_{y_t \sim \mathbf{K}_t'\mathbf{q}_t} \Big[\langle \mathbf{q}_t, \tilde{\ell}_t^2 \rangle \Big] = \int_{y_t \in \mathcal{G}_t} \mathbf{K}_t'\mathbf{q}_t(y_t) \langle \mathbf{q}_t, \tilde{\ell}_t^2 \rangle dy_t$$
$$= \int_{y_t \in \mathcal{G}_t} \mathbf{K}_t'\mathbf{q}_t(y_t) \Big[\int_{y \in \mathcal{G}_t} \mathbf{q}_t(y) \frac{(\ell_t(y_t))^2}{(\mathbf{K}_t'\mathbf{q}_t(y_t))^2} (\mathbf{K}_t'(y_t, y))^2 dy \Big] dy_t$$
$$\leq C^2 \int_{y_t \in \mathcal{G}_t} \frac{\mathbf{K}_t'^{(2)} \mathbf{q}_t(y_t)}{\mathbf{K}_t'\mathbf{q}_t(y_t)} dy_t \leq BC^2, \tag{13}$$

where the last inequality follows from Property #3 of Lemma 11 with $B = 2\Big(1 + \ln \frac{1}{\epsilon} + \ln \Big(\beta_{\mathcal{W}} - \alpha_{\mathcal{W}} \Big) \Big)$.

Finally, by definition of $\mathbf{p}^*$, we can bound the KL divergence term as:

$$KL(\mathbf{p}^* || \mathbf{p}_1) = d \log \frac{1}{\kappa} \tag{14}$$

Substituting (13) and (14) in (12), letting $L' = LDW$, and setting $\kappa = \frac{1}{L'T}$, $\epsilon = \frac{1}{3LT}$, (11) yields:

$$\mathbb{E}[R_T(\mathbf{w}^*)] \leq 2 + 2\Bigg(1 + \frac{KL(\mathbf{p}^* || \mathbf{p}_1)}{\eta} + \frac{\eta}{2} \mathbb{E}\Bigg[\sum_{t=1}^{T} \mathbb{E}_{y_t \sim \mathbf{K}_t'\mathbf{q}_t} \langle \mathbf{q}_t, \tilde{\ell}_t^2 \rangle \mid \mathcal{H}_t \Bigg] \Bigg)$$
$$= 4 + 2\Bigg(\frac{d \log L'T}{\eta} + \frac{\eta BC^2 T}{2} \Bigg)$$
$$= 4 + 2\sqrt{2}\Bigg(\sqrt{dBC^2 T \log(L'T)} \Bigg),$$

where the last equality follows by choosing $\eta = \Big(\frac{2d \log(L'T)}{BC^2 T} \Big)^{\frac{1}{2}}$. This concludes the proof. □

Statements and proofs of additional lemmas used above:

Lemma 8. *In Algorithm 2, at any round t, both $\mathbf{q}_t \in \mathcal{Q}_t$ and $\mathbf{K}'_t\mathbf{q}_t \in \mathcal{Q}_t$.*

Proof. Firstly note that, $\mathbf{p}_1 \in \mathcal{P}$ simply by its initialization, and for any subsequent iteration $t = 2, 3, \ldots, T$, $\mathbf{p}_t \in \mathcal{P}$ by its update rule.

Now for any $t \in [T]$ and $y \in \mathcal{G}_t$, by definition $\mathbf{q}_t(y) > 0$, as $\mathbf{p}_t \in \mathcal{P}$. The only remaining thing to prove is that $\int_{\mathcal{G}_t} d\mathbf{q}_t(y) = 1$, which simply follows as:

$$\int_{y\in\mathcal{G}_t} \mathbf{q}_t(y)dy = \int_{y\in\mathcal{G}_t}\int_{\mathcal{W}_t(y)} \mathbf{p}_t(\mathbf{w})d\mathbf{w} = \int_{\mathcal{W}} \mathbf{p}_t(\mathbf{w})d\mathbf{w} = 1 \ \ [\text{since } \mathbf{p}_t \in \mathcal{P}].$$

Now, consider $\mathbf{K}'_t\mathbf{q}_t$. By definition, $\forall y \in \mathcal{G}_t, \mathbf{K}'_t\mathbf{q}_t(y) = \int_{\mathcal{G}_t} \mathbf{K}'_t(y, y')d\mathbf{q}_t(y') > 0$ since by construction $\mathbf{K}'_t(y, \cdot) > 0$ and $\mathbf{q}_t \in \mathcal{Q}_t$. Further, since $\int_{\mathcal{G}_t} \mathbf{K}'_t(y, y')dy = 1$ for every $y' \in \mathcal{G}_t$ (by construction), it is easy to show $\int_{\mathcal{G}_t} \mathbf{K}_t\mathbf{q}_t(y)dy = 1$ as follows:

$$\int_{\mathcal{G}_t} \mathbf{K}'_t\mathbf{q}_t(y)dy = \int_{\mathcal{G}_t}\Big[\int_{\mathcal{G}_t} \mathbf{K}'_t(y, y')d\mathbf{q}_t(y')\Big]dy = \int_{\mathcal{G}_t}\Big[\int_{\mathcal{G}_t} \mathbf{K}'_t(y, y')dy\Big]d\mathbf{q}_t(y') = \int_{\mathcal{G}_t} d\mathbf{q}_t(y') = 1.$$

□

Lemma 9. *At any round $t \in [T]$ of Algorithm 2, $\langle \mathbf{p}_t, \tilde{f}_t\rangle = \langle \mathbf{q}_t, \tilde{\ell}_t\rangle$.*

Proof. The claim follows from the straightforward analysis:

$$\begin{aligned}\langle \mathbf{p}_t, \tilde{f}_t\rangle &= \int_{\mathbf{w}\in\mathcal{W}} \mathbf{p}_t(\mathbf{w})\tilde{f}_t(\mathbf{w})d\mathbf{w} = \int_{y\in\mathcal{G}_t}\int_{\mathbf{w}\in\mathcal{W}_t(y)} \mathbf{p}_t(\mathbf{w})\tilde{f}_t(\mathbf{w})d\mathbf{w} \\ &= \int_{y\in\mathcal{G}_t}\int_{\mathbf{w}\in\mathcal{W}_t(y)} \mathbf{p}_t(\mathbf{w})\tilde{\ell}_t(y)d\mathbf{w} = \int_{y\in\mathcal{G}_t} \tilde{\ell}_t(y)\int_{\mathbf{w}\in\mathcal{W}_t(y)} \mathbf{p}_t(\mathbf{w})d\mathbf{w} \\ &= \int_{y\in\mathcal{G}_t} \tilde{\ell}_t(y)\mathbf{q}_t(y)dy = \langle \mathbf{q}_t, \tilde{\ell}_t\rangle.\end{aligned}$$

□

Lemma 10. *Consider any sequence of functions $f_1, f_2, \ldots f_T$ such that $f_t : \mathcal{D} \mapsto \mathbb{R}$ for all $t \in [T]$, $\mathcal{D} \subset \mathbb{R}^d$ for some $d \in \mathbb{N}_+$. Suppose $\mathcal{P}$ denotes the set of probability measure over $\mathcal{D}$. Then for any $\mathbf{p} \in \mathcal{P}$, and given any $\mathbf{p}_1 \in \mathcal{P}$, the sequence $\{\mathbf{p}_t\}_{t=2}^T$ is defined as $\mathbf{p}_{t+1}(\mathbf{w}) := \dfrac{\mathbf{p}_t(\mathbf{w})\exp\big(-\eta f_t(\mathbf{w})\big)}{\int_{\tilde{\mathbf{w}}} \mathbf{p}_t(\tilde{\mathbf{w}})\exp\big(-\eta f_t(\tilde{\mathbf{w}})\big)d\tilde{\mathbf{w}}}$, for all $\mathbf{w} \in \mathcal{D}$. Then it can be shown that:*

$$\sum_{t=1}^T \langle \mathbf{p}_t - \mathbf{p}, f_t\rangle \le \frac{KL(\mathbf{p}||\mathbf{p}_1)}{\eta} + \frac{\eta}{2}\sum_{t=1}^T \langle \mathbf{p}_t, f_t^2\rangle,$$

where $KL(\mathbf{p}||\mathbf{p}_1)$ denotes the KL-divergence between the two probability distributions $\mathbf{p}$ and $\mathbf{p}_1$.

Proof. We start by noting that by definition of KL-divergence:

$$KL(\mathbf{p}||\mathbf{p}_t) - KL(\mathbf{p}||\mathbf{p}_{t+1}) = \int_{\mathcal{W}} \mathbf{p}(\mathbf{w})\ln\Big(\frac{\mathbf{p}_{t+1}(\mathbf{w})}{\mathbf{p}_t(\mathbf{w})}\Big)d\mathbf{w}.$$

Moreover, by definition of $\mathbf{p}_{t+1}$, $\frac{1}{\eta}\Big(KL(\mathbf{p}||\mathbf{p}_t) - KL(\mathbf{p}||\mathbf{p}_{t+1})\Big) = \frac{1}{\eta}\Big(\int_{\mathcal{W}} \mathbf{p}(\mathbf{w})\ln\Big(\frac{\mathbf{p}_{t+1}(\mathbf{w})}{\mathbf{p}_t(\mathbf{w})}\Big)\Big) = -\mathbb{E}_{\mathbf{p}}[f_t(\mathbf{w})] - \frac{1}{\eta}\ln\mathbb{E}_{\mathbf{p}_t}[e^{-\eta f_t(\mathbf{w})}]$ for any $t = 1, 2, \ldots, T$. Then summing over T rounds,

$$\sum_{t=1}^T\Bigg[-\mathbb{E}_{\mathbf{p}}[f_t(\mathbf{w})] - \frac{1}{\eta}\ln\mathbb{E}_{\mathbf{p}_t}[e^{-\eta f_t(\mathbf{w})}]\Bigg] = \frac{1}{\eta}\Big(KL(\mathbf{p}||\mathbf{p}_1) - KL(\mathbf{p}||\mathbf{p}_{T+1})\Big).$$

Now adding $\sum_{t=1}^T f_t(\mathbf{w}_t)$ to both sides, this further gives:

$$\sum_{t=1}^{T}\Bigg[f_t(\mathbf{w}_t) - \mathbb{E}_{\mathbf{p}}[f_t(\mathbf{w})]\Bigg] = \frac{1}{\eta}\Big(KL(\mathbf{p}||\mathbf{p}_1) - KL(\mathbf{p}||\mathbf{p}_{T+1})\Big) + \sum_{t=1}^{T}\Big(f_t(\mathbf{w}_t) + \frac{1}{\eta}\ln \mathbb{E}_{\mathbf{p}_t}[e^{-\eta f_t(\mathbf{w})}]\Big)$$
$$\implies \sum_{t=1}^{T}\Bigg[f_t(\mathbf{w}_t) - \mathbb{E}_{\mathbf{p}}[f_t(\mathbf{w})]\Bigg] \le \frac{KL(\mathbf{p}||\mathbf{p}_1)}{\eta} + \sum_{t=1}^{T}\Big(f_t(\mathbf{w}_t) + \frac{1}{\eta}\ln \mathbb{E}_{\mathbf{p}_t}[e^{-\eta f_t(\mathbf{w})}]\Big)$$
$$\implies \sum_{t=1}^{T}\mathbb{E}_{\mathbf{w}_t\sim\mathbf{p}_t}\Bigg[f_t(\mathbf{w}_t) - \mathbb{E}_{\mathbf{p}}[f_t(\mathbf{w})]\Bigg] \le \frac{KL(\mathbf{p}||\mathbf{p}_1)}{\eta} + \frac{1}{\eta}\sum_{t=1}^{T}\mathbb{E}_{\mathbf{w}_t\sim\mathbf{p}_t}\Bigg[\eta f_t(\mathbf{w}_t) + \ln \mathbb{E}_{\mathbf{p}_t}[e^{-\eta f_t(\mathbf{w})}]\Bigg]$$
$$\implies \sum_{t=1}^{T}\Bigg[\Big\langle(\mathbf{p}_t - \mathbf{p}), f_t\Big\rangle\Bigg] \le \frac{KL(\mathbf{p}||\mathbf{p}_1)}{\eta} + \frac{1}{\eta}\sum_{t=1}^{T}\mathbb{E}_{\mathbf{w}_t\sim\mathbf{p}_t}\Bigg[\eta f_t(\mathbf{w}_t) + \mathbb{E}_{\mathbf{p}_t}[e^{-\eta f_t(\mathbf{w})}] - 1\Bigg]$$
$$\le \frac{KL(\mathbf{p}||\mathbf{p}_1)}{\eta} + \frac{1}{\eta}\sum_{t=1}^{T}\mathbb{E}_{\mathbf{w}_t\sim\mathbf{p}_t}\Bigg[\eta f_t(\mathbf{w}_t) + 1 - \eta\mathbb{E}_{\mathbf{w}\sim\mathbf{p}_t}[f_t(\mathbf{w})] + \mathbb{E}_{\mathbf{w}\sim\mathbf{p}_t}[\frac{\eta^2 f_t^2(\mathbf{w})}{2}] - 1\Bigg]$$
$$= \frac{KL(\mathbf{p}||\mathbf{p}_1)}{\eta} + \frac{\eta}{2}\sum_{t=1}^{T}\big\langle \mathbf{p}_t, f_t^2\big\rangle,$$

which concludes the proof. The last two inequalities above follow from $\ln s \le s-1,\ \forall s>0$ and $e^{-s} \le 1 - s + s^2/2,\ \forall s > 0$. □

Lemma 11. *For any convex and L-Lipschitz function, $\ell : \mathcal{G}_t \mapsto \mathbb{R}_+$, such that $\mathcal{G}_t = [\alpha,\beta] \subseteq \mathbb{R}$, $\mathbf{q} \in \mathcal{Q}_t$, and any $y \in \mathcal{G}_t$, the kernel $\boldsymbol{K}'_t : \mathcal{G}_t \times \mathcal{G}_t \mapsto \mathbb{R}_+$ satisfies:*

1. *The function ${\boldsymbol{K}'_t}^*\ell(\cdot)$ is L-Lipschitz.*
2. *${\boldsymbol{K}'_t}^*\ell(y) \le (1-\lambda)\big\langle \boldsymbol{K}'_t\mathbf{q}, \ell\big\rangle + \lambda\ell(y) + 3\epsilon L$, where λ is a constant.*
3. *For any $\mathbf{q} \in \mathcal{Q}_t$, define operator ${\boldsymbol{K}'_t}^{(2)}\mathbf{q} : \mathcal{G}_t \mapsto \mathbb{R}$ as:*

$$ {\boldsymbol{K}'_t}^{(2)}\mathbf{q}(y) := \int_{y'\in\mathcal{G}_t} (\boldsymbol{K}'_t(y,y'))^2 d\mathbf{q}(y') \quad \forall y \in \mathcal{G}_t,$$

then $\displaystyle\int_{y\in\mathcal{G}_t} \frac{{\boldsymbol{K}'_t}^{(2)}\mathbf{q}(y)}{\boldsymbol{K}'_t\mathbf{q}y}dy \le B$, *where* $B = 2\Big(1 + \ln\frac{1}{\epsilon} + \ln\Big(\beta-\alpha\Big)\Big)$.

Proof. 1. For the first part, let us denote $\bar{y} = \mathbb{E}_{y\sim\mathbf{q}}[y]$. Then note that:

$$ {\mathbf{K}'_t}^*\ell(y) = \big\langle \mathbf{K}'_t\delta_y, \ell\big\rangle = \begin{cases}\mathbb{E}_{U\sim\text{unif}[0,1]}\big[\ell(U\bar{y} + (1-U)y)\big], & \text{if } |y-\bar{y}| \ge \epsilon \\ \mathbb{E}_{U\sim\text{unif}[0,1]}\big[\ell(\bar{y} - \epsilon U)\big], & \text{if } |y-\bar{y}| < \epsilon\end{cases}, \tag{15}$$

which immediately implies the function ${\mathbf{K}'_t}^*\ell(\cdot)$ has the same Lipschitz parameter that of $\ell(\cdot)$.

2. We prove this part considering two cases separately:

Case 1. $|y - \bar{y}| \ge \epsilon$: By construction of $\mathbf{K}'_t$ (see Definition 4), we note that expectation of y w.r.t. $\mathbf{q}$ and $\mathbf{K}'_t\mathbf{q}$, i.e. respectively $\bar{y} = \mathbb{E}_{y\sim\mathbf{q}}[y]$ and $\mathbb{E}_{y\sim\mathbf{K}'_t\mathbf{q}}[y]$ can differ at most by 2ϵ, i.e. $|\mathbb{E}_{y\sim\mathbf{q}}[y] - \mathbb{E}_{y\sim\mathbf{K}'_t\mathbf{q}}[y]| \le 2\epsilon$ (Bubeck et al., 2017). We write, $\mathbb{E}_{y\sim\mathbf{K}'_t\mathbf{q}}[y] = \mathbb{E}_{y\sim\mathbf{q}}[y] + \psi$, clearly $\psi \in [-2\epsilon, 2\epsilon]$. Hence:

$$\ell(\bar{y}) = \ell(\mathbb{E}_{y\sim\mathbf{K}'_t\mathbf{q}}[y] - \psi)$$
$$\le \ell\Big(\int_{y\in\mathcal{G}_t} y\mathbf{K}'_t\mathbf{q}(y)dy\Big) + \psi L \le \int_{y\in\mathcal{G}_t} \ell(y)\mathbf{K}'_t\mathbf{q}(y)dy + 2\epsilon L$$

$$= \big\langle \mathbf{K}_t'\mathbf{q}, \ell \big\rangle + 2\epsilon L \tag{16}$$

where the first inequality follows using the L-lipschitzness of ℓ and the second inequality follows using Jensen's inequality (since ℓ is convex). Now consider the case $|y-\bar{y}| \geq \epsilon$ in (15):

$$\mathbf{K}_t'^{*}\ell(y) = \mathbb{E}_{U\sim \text{unif}[0,1]}\big[\ell(U\bar{y} + (1-U)y)\big] \leq \frac{\ell(\bar{y}) + \ell(y)}{2}$$
$$\overset{\text{by (16)}}{\leq} \frac{\big\langle \mathbf{K}_t'\mathbf{q}, \ell\big\rangle + \ell(y)}{2} + \epsilon L$$

This shows that for this case the claim of Part (2) holds for $\lambda = \frac{1}{2}$.
Case 2. $|y-\bar{y}| < \epsilon$: Note $\bar{y} - \epsilon U \in [\bar{y}-\epsilon, \bar{y}]$ in (15). And in this case $\ell(\bar{y}) \leq \ell(y) + \epsilon L$. Using the fact that $\ell(\cdot)$ is convex and L-lipschitz, by similar arguments used to obtain (16) above, we have:

$$\mathbf{K}_t'^{*}\ell(y) \leq \ell(\bar{y}) + \epsilon L = \ell(\bar{y})/2 + \ell(\bar{y})/2 + \epsilon L \leq \langle \mathbf{K}_t'\mathbf{q}, \ell\rangle/2 + (\ell(y)+\epsilon L)/2 + 2\epsilon L$$

which implies for this case as well, the claim of Part (2) holds for $\lambda = 1/2$.

3. For this part, note that:

$$\int_{y\in\mathcal{G}_t} \frac{\mathbf{K}_t'^{(2)}\mathbf{q}(y)}{\mathbf{K}_t'\mathbf{q}(y)} dy \overset{(a)}{\leq} \int_\alpha^\beta \frac{1}{\max(|y-\bar{y}|,\epsilon)} dy$$
$$= \int_\alpha^{\bar{y}-\epsilon} \frac{1}{\max(|y-\bar{y}|,\epsilon)} dy + \int_{\bar{y}-\epsilon}^{\bar{y}+\epsilon} \frac{1}{\max(|y-\bar{y}|,\epsilon)} dy + \int_{\bar{y}+\epsilon}^{\beta} \frac{1}{\max(|y-\bar{y}|,\epsilon)} dy$$
$$= \int_\alpha^{\bar{y}-\epsilon} \frac{1}{\bar{y}-y} dy + \int_{\bar{y}-\epsilon}^{\bar{y}+\epsilon} \frac{1}{\epsilon} dy + \int_{\bar{y}+\epsilon}^{\beta} \frac{1}{y-\bar{y}} dy$$
$$= \frac{1}{\epsilon}\int_{\bar{y}-\epsilon}^{\bar{y}+\epsilon} dy + 2\ln\frac{1}{\epsilon} + \ln(\beta-\bar{y}) + \ln(\bar{y}-\alpha)$$
$$\leq 2\bigg(1 + \ln\frac{1}{\epsilon} + \ln\Big(\beta-\alpha\Big)\bigg) \;\; (\text{since } \alpha \leq \bar{y} \leq \beta)$$

where (a) follows noting $\mathbf{K}_t'(y,y') \leq \frac{1}{\max(|y-\bar{y}|,\epsilon)}, \forall y, y' \in \mathcal{G}_t$ which implies $\mathbf{K}_t'^{(2)}\mathbf{q}(y) \leq \frac{\mathbf{K}_t'\mathbf{q}(y)}{\max(|y-\bar{y}|,\epsilon)}$. □

A.3. Proof of Lemma 6

Proof. For any $\ell_t : \mathbb{R} \to [0,C]$, $t \in [T]$, define $\hat{\ell}_t : \mathbb{R} \mapsto [0,C]$ such that $\hat{\ell}_t(y) = \mathbb{E}_{u\sim\mathbf{U}\left(\mathcal{B}_1(1)\right)}\ell_t(y+\delta u)$, for any $y \in \mathbb{R}$. Let us also define $\hat{f}_t(\mathbf{w}) = \hat{\ell}_t(g_t(\mathbf{w};\mathbf{x}_t)), \forall \mathbf{w} \in \mathcal{W}$. Let $y_t = g_t(\mathbf{w}_t;\mathbf{x}_t), \forall t \in [T]$.

Then given any fixed $\mathbf{w} \in \mathcal{W}$ and $\mathbf{x} \in \mathbb{R}^d$, by chain rule $\nabla_\mathbf{w}\hat{f}_t(\mathbf{w}) = \frac{d\hat{f}_t(y)}{dy}\nabla_\mathbf{w}(g_t(\mathbf{w};\mathbf{x}_t)) = \dfrac{d\hat{\ell}_t(y)}{dy}\nabla_\mathbf{w}(g_t(\mathbf{w}_t;\mathbf{x}_t))$. Consider the RHS of the lemma equality:

$$\mathbb{E}_{u\sim\mathbf{U}(\mathcal{S}_1(1))}\Big[\frac{1}{\delta}\ell_t\big(g_t(\mathbf{w}_t;\mathbf{x}_t) + \delta u\big)u \mid \mathbf{w}_t\Big]\nabla_\mathbf{w}(g_t(\mathbf{w}_t;\mathbf{x}_t))$$
$$= \frac{d\hat{\ell}_t(y_t)}{dy_t}\nabla_\mathbf{w}(g_t(\mathbf{w}_t;\mathbf{x}_t)) = \nabla_\mathbf{w}\hat{f}_t(\mathbf{w}_t) = \nabla_\mathbf{w}\mathbb{E}_u\big[\ell_t(g_t(\mathbf{w}_t;\mathbf{x}_t) + \delta u)\big],$$

where the first equality is due to Lemma 1 of (Flaxman et al., 2005) applied to the 1-dimensional ball $\mathcal{B}_1(1)$. □

A.4. Proof of Lemma 7

Proof. We start by recalling Lemma 2 of (Flaxman et al., 2005) that uses the online gradient descent analysis by (Zinkevich, 2003) with unbiased random gradient estimates. We restate the result below for convenience:

Lemma 12 (Lemma 2, (Flaxman et al., 2005)). *Let $S \subset \mathcal{B}_d(R) \subset \mathbb{R}^d$ be a convex set, $f_1, f_2, \ldots, f_T : S \mapsto \mathbb{R}$ be a sequence of convex, differentiable functions. Let $\mathbf{w}_1, \mathbf{w}_2, \ldots, \mathbf{w}_T \in S$ be a sequence of predictions defined as $\mathbf{w}_1 = 0$ and $\mathbf{w}_{t+1} = \mathbf{P}_S(\mathbf{w}_t - \eta h_t)$, where $\eta > 0$, and $h_1, h_2, \ldots, h_T$ are random variables such that $\mathbb{E}[h_t \big| \mathbf{w}_t] = \nabla f_t(\mathbf{w}_t)$, and $\|h_t\|_2 \le G$, for some $G > 0$ then, for $\eta = \frac{R}{G\sqrt{T}}$ the expected regret incurred by above prediction sequence is:*

$$\mathbb{E}\bigg[\sum_{t=1}^{T} f_t(\mathbf{w}_t)\bigg] - \min_{\mathbf{w}\in S}\sum_{t=1}^{T} f_t(\mathbf{w}) \le RG\sqrt{T}.$$

Coming back to our problem setup, let us first denote $\hat{f}_t(\mathbf{w}) = \hat{\ell}_t(g_t(\mathbf{w}; \mathbf{x}_t))$, for all $\mathbf{w} \in \mathcal{W}$, $t \in [T]$ (recall from the proof of Lemma 6, we define $\hat{\ell}_t : \mathbb{R} \mapsto [0, C]$ such that $\hat{\ell}_t(y) = \mathbb{E}_{u \sim \mathbf{U}(\mathcal{B}_1(1))}\ell_t(y + \delta u)$, for any $y \in \mathbb{R}$). We can now apply Lemma 12 in the setting of Algorithm 3 on the sequence of convex (by **(A1) (ii)**), differentiable functions $\hat{f}_1, \hat{f}_2, \ldots \hat{f}_T : \mathcal{W}_\alpha \mapsto [0, C]$, with $h_t = \frac{1}{\delta}\big(\ell_t(a_t)u\big)\nabla g_t(\mathbf{w}_t; \mathbf{x}_t)$, with $u \sim \mathcal{B}_1(1)$ (note that Lemma 6 implies $\mathbb{E}[h_t \big| \mathbf{w}_t] = \nabla_{\mathbf{w}} \hat{f}_t(\mathbf{w}_t) = \nabla_{\mathbf{w}}\mathbb{E}_u\big[\ell_t(g_t(\mathbf{w}_t; \mathbf{x}_t) + \delta u)\big]$). We get:

$$\mathbb{E}\bigg[\sum_{t=1}^{T} \hat{f}_t(\mathbf{w}_t)\bigg] - \min_{\mathbf{w}\in \mathcal{W}_\alpha}\sum_{t=1}^{T} \hat{f}_t(\mathbf{w}) \le \frac{WDC\sqrt{T}}{\delta}, \tag{17}$$

as in this case $R \le (1-\alpha)W < W$, and, by **(A3) (ii)**, $\|h_t\| = \|\frac{1}{\delta}\big(\ell_t(a_t)u\big)\nabla(g_t(\mathbf{w}_t; \mathbf{x}_t))\| \le \frac{DC}{\delta}$, so $G = \frac{DC}{\delta}$, and $\eta = \frac{W\delta}{DC\sqrt{T}}$. Further, since $\ell_t(\cdot)$s are assumed to be L-Lipschitz, (17) yields:

$$\begin{aligned}
&\mathbb{E}\bigg[\sum_{t=1}^{T}(f_t(\mathbf{w}_t) - \delta L)\bigg] - \min_{\mathbf{w}\in \mathcal{W}_\alpha}\sum_{t=1}^{T}(f_t(\mathbf{w}) + \delta L) \le \frac{WDC\sqrt{T}}{\delta},\\
\implies &\mathbb{E}\bigg[\sum_{t=1}^{T} f_t(\mathbf{w}_t)\bigg] - \min_{\mathbf{w}\in \mathcal{W}_\alpha}\sum_{t=1}^{T} f_t(\mathbf{w}) \le \frac{WDC\sqrt{T}}{\delta} + 2\delta LT\\
\implies &\mathbb{E}\bigg[\sum_{t=1}^{T} f_t(\mathbf{w}_t)\bigg] - \min_{\mathbf{w}\in \mathcal{W}}\sum_{t=1}^{T} f_t(\mathbf{w}) \le \frac{WDC\sqrt{T}}{\delta} + 2\delta LT + \alpha LT,\\
\implies &\mathbb{E}\bigg[\sum_{t=1}^{T} f_t(\mathbf{w}_t)\bigg] - \min_{\mathbf{w}\in \mathcal{W}}\sum_{t=1}^{T} f_t(\mathbf{w}) \le \frac{WDC\sqrt{T}}{\delta} + 3\delta LT,
\end{aligned}$$

setting $\alpha = \delta$. The claim follows minimizing the RHS above w.r.t. δ. Setting $\delta = \Big(\frac{WDC}{3L\sqrt{T}}\Big)^{1/2}$ gives:

$$\mathbb{E}[\mathcal{R}_T(\mathcal{A})] = \mathbb{E}\bigg[\sum_{t=1}^{T} f_t(\mathbf{w}_t)\bigg] - \min_{\mathbf{w}\in \mathcal{W}}\sum_{t=1}^{T} f_t(\mathbf{w}) \le 2\sqrt{3WLDC}T^{3/4},$$

which concludes the proof. □

B. Appendix for Simulations (Section 5)

Implementation details of Algorithm 2. The main challenge in implementing Kernelized Exponential Weights for PBCO (Algorithm 2) is to handle the continuous 'action space' $\mathcal{W}$; in particular, to maintain and update the probability distribution $\mathbf{p}_t$ over $\mathcal{W}$, and to sample from $\mathbf{p}_t$ given y_t at round t. Towards this we use an epsilon-net trick to discretize $\mathcal{W}$ into finitely many points—specifically, since we choose $\mathcal{W} = \mathcal{B}_d(1)$, we discretize the $[0, 1]$ interval every d direction with a grid size of $O(1/d)$, and consider only the points inside $\mathcal{B}_d(1)$. This reduces the action space $\mathcal{W}$ into finitely many points (say N), and we now proceed by maintaining and updating probabilities on every such discrete point following the steps of Algorithm 2 (we initialize $\mathbf{p}_1 \leftarrow 1/N$ for all N points in the epsilon net).