A. Proofs

A.1. Proof of Theorem

**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 $x_i = e_i$. Denote $\sigma = (\sigma_1, \ldots, \sigma_d)$.

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 $g_t(w; x_i) = w^\top x_i$. Clearly $\nabla_w g_t(\cdot; x_i) = x_i \in \{0,1\}^d$ which is revealed to the learner at the beginning of round $t$. We choose $x_i = x_i$.

2. Loss function $f_t(w) = \ell_t(w^\top x_i) + \varepsilon_t = \mu \sigma_i(w^\top x_i) + \varepsilon_t$, where $\varepsilon_t \sim \mathcal{N}(0, \frac{1}{16})$, for some constant $\mu > 0$ (to be decided later), $\forall w \in \mathcal{W}$.

3. Learner plays $w_t = [w_t(1), \ldots, w_t(d)] \in \mathcal{W}$.

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

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

The expected regret of any $A$:

\begin{align*}
\mathbb{E}[R_I] &= \sum_{i=1}^d \sum_{t \in T_i} \mu [(\sigma_i x_i^\top) w_t - (\sigma_i x_i^\top) w^*] = \sum_{i=1}^d \mu T_d \mathbb{E}[\sigma_i x_i^\top \bar{w}_i] - (\sigma_i x_i^\top) w^*] \\
&= \sum_{i=1}^d T_d \mathbb{E} \left[ \mu \sigma(i) (\bar{w}_i - w^*) \right] \\
&= \sum_{i=1}^d T_d \mathbb{E} \left[ \mu \sqrt{d} w^* (\bar{w}_i - w^*) \right] = \sum_{i=1}^d T_d \mathbb{E} \left[ \mu \sqrt{d} \left( (w^*)^2 - (\bar{w}_i) w^*(i) \right) \right] \\
&= \sum_{i=1}^d T_d \mathbb{E} \left[ \mu \sqrt{d} \left( \frac{1}{d} + \frac{\sigma_i}{\sqrt{d}} w_i(i) \right) \right] \\
&= \sum_{i=1}^d T_d \left[ \frac{2\mu}{\sqrt{d}} \mathbb{P}(\sigma_i \bar{w}_i(i) > 0) \right] \text{ since } \bar{w}_i(i) \in \{-1/\sqrt{d}, 1/\sqrt{d}\} \tag{7}
\end{align*}

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

\begin{align*}
\mathbb{P}(\sigma_i \bar{w}_i(i) > 0) &= \frac{1}{2} \left( \mathbb{P}(\bar{w}_i(i) > 0 \mid \sigma_i = +1) + \mathbb{P}(\bar{w}_i(i) < 0 \mid \sigma_i = -1) \right) \\
&= \frac{1}{2} \left( \mathbb{P}(\bar{w}_i(i) > 0 \mid \sigma_i = +1) + 1 - \mathbb{P}(\bar{w}_i(i) > 0 \mid \sigma_i = -1) \right) \\
&\geq \frac{1}{2} \left( 1 - \mathbb{P}(\bar{w}_i(i) > 0 \mid \sigma_i = +1) - \mathbb{P}(\bar{w}_i(i) > 0 \mid \sigma_i = -1) \right),
\end{align*}
Assumption 1. For proving the lower bound we assume that \( \tilde{w}_i(i) \) is a deterministic function of the observed function values \( \{f_t\}_{t \in T_i} \), respectively at \( \{w_t\}_{t \in T_i} \). 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 \( \tilde{w}_i(i) \) only depends on \( f([T_i]) \), applying Pinsker’s inequality, we get:

\[
Pr(\sigma_i \tilde{w}_i(i) > 0) \geq \frac{1}{2} \left( 1 - Pr(\sigma_i \tilde{w}_i(i) > 0 \mid \sigma_i = +1) - Pr(\sigma_i \tilde{w}_i(i) < 0 \mid \sigma_i = -1) \right)
\]

\[
\geq \frac{1}{2} \left( 1 - \sqrt{2KL\left(P(f([T_i]) \mid \sigma_i = +1)\|P(f([T_i]) \mid \sigma_i = -1)\right)} \right)
\]

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

\[
Pr(\sigma_i \tilde{w}_i(i) > 0) \geq \frac{1}{2} \left( 1 - \sum_{t \in T_i} KL\left(P(f_t \mid \sigma_i = +1, \{f_r\}_{r \in [t-1]\setminus{T_i}}} \| P(f_t \mid \sigma_i = -1, \{f_r\}_{r \in [t-1]\setminus{T_i}}) \right) \right)
\]

\[
\geq \frac{1}{2} \left( 1 - \frac{2\mu^2 \sigma_i^2 \tilde{w}_i(i)^2}{16} \right) = \frac{1}{2} \left( 1 - \frac{64\mu^2 T_d}{d} \right) \text{ since } \tilde{w}_i(i)^2 = \frac{1}{d} \text{ and } \sigma_i^2 = 1
\]

where the last inequality follows by noting \( P(f_t \mid \sigma_i, \{f_r\}_{r \in [t-1]\setminus{T_i}}) \sim \mathcal{N}(\mu \sigma_i, \tilde{w}_i(i)^2) \), and \( KL(\mathcal{N}(\mu_1, \sigma^2_1)\|\mathcal{N}(\mu_2, \sigma^2_2)) = \frac{(\mu_1 - \mu_2)^2}{2\sigma^2_2} \) (for bounding the each individual KL-divergence terms).

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

Combining the above claims with Eq. (7):

\[
\mathbb{E}[R_T] = \sum_{i=d}^{T} T_d \left[ \frac{2\mu}{\sqrt{d}} Pr(\sigma_i \tilde{w}_i(i) > 0) \right] \geq \sum_{i=d}^{T} T_d \left[ \frac{\mu}{\sqrt{d}} (1 - 8\mu \sqrt{\frac{T_d}{d}}) \right]
\]

\[
\geq \sum_{i=d}^{T} T_d \frac{1}{16\sqrt{T_d}} \left( 1 - \frac{1}{2} \right) \left( \text{setting } \mu = \frac{\sqrt{d}}{16\sqrt{T_d}} \leq 1 \right) = \frac{\sqrt{dT}}{32}.
\]

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 \( w(d' + 1), \ldots, w(d) \), i.e. for any \( w \in W \subseteq \mathbb{R}^d \), denoting \( w_{[d']} = (w_1, \ldots, w_{d'}) \), we can construct \( f_t(w) = f_t(w_{[d']}) \)).

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

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 \( K'_{\ell} \) (Definition 4), we define a linear operator \( K'_{\ell}^{**} \) on the space of functions \( G_t \mapsto \mathbb{R} \) as follows. For any function \( \ell : G_t \mapsto \mathbb{R} \):

\[
K'_{\ell}^{**}(y') := \int_{y' \in G_t} \ell(y') K'_{\ell}(y', y) dy \quad \forall y \in G_t,
\]

(8)
We also denote by \( P \) and \( Q_t \) the set of all probability measures on \( W \) and \( G_t \) respectively; and by \( \delta_y \in Q_t \), \( \delta_w \in P \) the dirac mass at \( y \in G_t \) and at \( w \in W \) respectively. For \( q \in Q_t \), define:

\[
\langle q, \ell \rangle = \int_{y \in G_t} \ell(y)q(y)dy
\]

As noted in (Bubeck et al., 2017), a useful observation on the operator (8) is that for any \( q \in Q_t \):

\[
\langle K' q, \ell \rangle = \langle K^* q, \ell \rangle.
\]

**Proof of Lemma 5**

**Proof.** For ease, we abbreviate \( g_t(w_t; x_t) \) as \( g_t(w) \) throughout the proof. We start by analyzing the expected regret w.r.t. the optimal point \( w^* \in W \) (denote \( y^*_t = g_t(w^*) \) for all \( t \in [T] \)). Define \( \forall y \in G_t, \hat{\ell}_t(y) := f_t(w) \), for any \( w \in W(y) \). Also let \( H_t = \sigma(\{x_r, P_r, w_r, f_r\}_{r=1}^{t-1} \cup \{x_t, p_t\}) \) 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:

\[
\mathbb{E}[R_T(w^*)] := \mathbb{E}\left[ \sum_{t=1}^{T} (f_t(w_t) - f_t(w^*)) \right] = \mathbb{E}\left[ \sum_{t=1}^{T} (\hat{\ell}_t(g_t(w_t)) - \hat{\ell}_t(g_t(w^*))) \right]
\]

\[
= \mathbb{E}\left[ \sum_{t=1}^{T} (\hat{\ell}_t(y) - \hat{\ell}_t(y^*_t)) \right] = \mathbb{E}\left[ \sum_{t=1}^{T} \langle K'_t q_t - \delta_{y^*_t}, \hat{\ell}_t \rangle \right] \quad \text{[since } y_t \sim K'_t q_t \text{]}
\]

\[
\leq \mathbb{E}\left[ \sum_{t=1}^{T} 3\epsilon L + \frac{1}{\lambda} \langle K'_t (q_t - \delta_{y^*_t}), \hat{\ell}_t \rangle \right] \quad \text{[from Property\#2 of Lemma 11]}
\]

\[
\leq 6\epsilon LT + 2 \sum_{t=1}^{T} \mathbb{E}\left[ \langle K'_t (q_t - \delta_{y^*_t}), \hat{\ell}_t \rangle \right] \quad \text{[we can choose } \lambda = 1/2, \text{ see proof of Lemma 11]}
\]

\[
\overset{(a)}{=} 6\epsilon LT + 2 \sum_{t=1}^{T} \mathbb{E}\left[ \sum_{t=1}^{T} \langle K''_{t+1, t} \hat{\ell}_t, (q_t - \delta_{y^*_t}) \rangle \right]
\]

\[
= 6\epsilon LT + 2 \sum_{t=1}^{T} \mathbb{E}\left[ \sum_{t=1}^{T} \mathbb{E}_{y_t \sim K'_t q_t} \left[ \langle \hat{\ell}_t - \delta_{w^* - \hat{\ell}_t}, \hat{f}_t \rangle \mid H_t \right] \right]
\]

where the last equality follows by Lemma 5 and by \( \langle \delta_{w^* - \hat{\ell}_t}, \hat{f}_t \rangle = \hat{f}_t(w) = \hat{f}_t(y^*_t) = \langle \delta_{y^*_t}, \hat{\ell}_t \rangle \); the penultimate equality (a) follows noting that for any \( y' \in G_t \):

\[
\mathbb{E}_{y_t \sim K'_t q_t} [\hat{f}_t(y)] = \int_{y_t \in G_t} K'_t q_t(y) \frac{\hat{\ell}_t(y_t)}{K'_t q_t(y_t)} \mathcal{K}'(y_t, y') dy_t = \int_{y_t \in G_t} \hat{\ell}_t(y_t) \mathcal{K}'(y_t, y') dy_t = \mathcal{K}'^* \ell_t(y^*_t).
\]

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

Then note that:

\[
\sum_{t=1}^{T} \mathbb{E}_{y_t \sim K'_t q_t} [\langle p_t - \delta_{w^* - \hat{\ell}_t}, \hat{f}_t \rangle] = \sum_{t=1}^{T} \mathbb{E}_{y_t \sim K'_t q_t} [\langle p_t, \hat{f}_t \rangle - \langle \delta_{w^* - \hat{\ell}_t}, \hat{f}_t \rangle]
\]

\[
\overset{(a)}{=} \sum_{t=1}^{T} \mathbb{E}_{y_t \sim K'_t q_t} [\langle p_t, \hat{f}_t \rangle] - K''_{t+1, t} \hat{\ell}_t(g_t(w^*))
\]

**Pseudo-1d Bandit Convex Optimization**
\[
(b) \leq \sum_{t=1}^{T} E_{y_t \sim \mathcal{K}_t q_t} [\langle p_t, \tilde{f}_t \rangle] + \sum_{t=1}^{T} \left[ \kappa LDW - \langle p^*, K_t^* \ell_t (g_t(\cdot)) \rangle \right] \\
= \sum_{t=1}^{T} E_{y_t \sim \mathcal{K}_t q_t} [\langle p_t, \tilde{f}_t \rangle - \langle p^*, \tilde{f}_t \rangle] + \kappa LDWT
\]

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

Combining above claims with \[10\] we further get:

\[
\mathbb{E} [R_T (w^*)] \leq 6 \epsilon LT + 2 \left( \kappa LDWT + \mathbb{E} \left[ \sum_{t=1}^{T} E_{y_t \sim \mathcal{K}_t q_t} [\langle p_t - p^*, \tilde{f}_t \rangle | \mathcal{H}_t] \right] \right). \tag{11}
\]

From Lemma \[10\] we get:

\[
\sum_{t=1}^{T} \langle p_t - p^*, \tilde{f}_t \rangle \leq \frac{KL (p^* || p_t)}{\eta} + \frac{\eta}{2} \langle p_t, \tilde{f}_t \rangle = \frac{KL (p^* || p_t)}{\eta} + \frac{\eta}{2} \langle q_t, \tilde{\ell}_t \rangle, \tag{12}
\]

where the equality \(\langle p_t, \tilde{f}_t \rangle = \langle q_t, \tilde{\ell}_t \rangle\) follows from a similar derivation as shown in Lemma \[9\] Now, note that:

\[
\mathbb{E}_{y_t \sim \mathcal{K}_t q_t} [\langle q_t, \tilde{\ell}_t \rangle] = \int_{y_t \in G_t} K_t q_t (y_t) \langle q_t, \tilde{\ell}_t \rangle dy_t \\
= \int_{y_t \in G_t} K_t q_t (y_t) \left[ \int_{y_t \in G_t} q_t (y) \left( \frac{\ell_t (y_t)}{K_t (q_t (y_t))} \right)^2 dy_t \right] dy_t \\
\leq C^2 \int_{y_t \in G_t} K_t (q_t (y_t)) \langle q_t, \tilde{\ell}_t \rangle dy_t \leq BC^2, \tag{13}
\]

where the last inequality follows from Property \#3 of Lemma \[11\] with \(B = 2 \left( 1 + \frac{1}{\epsilon} + \text{ln} \left( \beta_{\mathcal{W}} - \alpha_{\mathcal{W}} \right) \right)\). Finally, by definition of \(p^*\), we can bound the KL divergence term as:

\[
KL (p^* || p_t) = d \log \frac{1}{\kappa}. \tag{14}
\]

Substituting \[13\] and \[14\] in \[12\], letting \(L' = LDW\), and setting \(\kappa = \frac{1}{LT}, \epsilon = \frac{1}{3LT}, \) \[11\] yields:

\[
\mathbb{E} [R_T (w^*)] \leq 2 + 2 \left( 1 + \frac{KL (p^* || p_t)}{\eta} + \frac{\eta}{2} \mathbb{E} \left[ \sum_{t=1}^{T} E_{y_t \sim \mathcal{K}_t q_t} [\langle q_t, \tilde{\ell}_t \rangle | \mathcal{H}_t] \right] \right) \\
= 4 + 2 \left( \frac{d \log L'T}{\eta} + \frac{\eta BC^2 T}{2} \right) \\
= 4 + 2 \sqrt{2} \left( \frac{d \log (L'T)}{\eta} + \frac{\eta BC^2 T}{2} \right),
\]

where the last equality follows by choosing \(\eta = \left( \frac{2d \log (L'T)}{BC^2 T} \right)^{\frac{1}{2}}\). This concludes the proof. \qed
Statements and proofs of additional lemmas used above:

**Lemma 8.** In Algorithm 2 at any round \( t \), both \( q_t \in Q_t \) and \( K_t'q_t \in Q_t \).

**Proof.** Firstly note that, \( p_t \in \mathcal{P} \) simply by its initialization, and for any subsequent iteration \( t = 2, 3, \ldots, T \), \( p_t \in \mathcal{P} \) by its update rule.

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

\[
\int_{y \in \mathcal{G}_t} q_t(y)dy = \int_{y \in \mathcal{G}_t} \int_{\mathcal{W}_t(y)} p_t(w)dw = \int_{\mathcal{W}_t} p_t(w)dw = 1 \quad \text{[since } p_t \in \mathcal{P}.\]

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

\[
\int_{\mathcal{G}_t} K_t'q_t(y)dy = \int_{\mathcal{G}_t} \left[ \int_{\mathcal{G}_t} K_t'(y, y')dq_t(y') \right] dy = \int_{\mathcal{G}_t} K_t'(y)dy = \int_{\mathcal{G}_t} dq_t(y) = 1.
\]

**Lemma 9.** At any round \( t \in [T] \) of Algorithm 2 \( \langle p_t, \tilde{f}_t \rangle = \langle q_t, \tilde{\ell}_t \rangle \).

**Proof.** The claim follows from the straightforward analysis:

\[
\langle p_t, \tilde{f}_t \rangle = \int_{w \in \mathcal{W}} p_t(w)\tilde{f}_t(w)dw = \int_{y \in \mathcal{G}_t} \int_{w \in \mathcal{W}_t(y)} p_t(w)\tilde{f}_t(w)dw = \int_{y \in \mathcal{G}_t} \tilde{\ell}_t(y)\int_{w \in \mathcal{W}_t(y)} p_t(w)dw = \int_{y \in \mathcal{G}_t} \tilde{\ell}_t(y)q_t(y)dy = \langle q_t, \tilde{\ell}_t \rangle.
\]

**Lemma 10.** Consider any sequence of functions \( f_1, f_2, \ldots, f_T \) such that \( f_t : \mathcal{D} \rightarrow \mathbb{R} \) 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 \( p \in \mathcal{P} \), and given any \( p_1 \in \mathcal{P} \), the sequence \( \{p_t\}_{t=2}^{T} \) is defined as \( p_{t+1}(w) := \frac{p_t(w)\exp\left(-\eta f_t(w)\right)}{\int_{\mathcal{D}} p_t(w)\exp\left(-\eta f_t(w)\right)dw} \), for all \( w \in \mathcal{D} \). Then it can be shown that:

\[
\sum_{t=1}^{T} \langle p_t - p, f_t \rangle \leq \frac{KL(p||p_1)}{\eta} + \frac{\eta}{2} \sum_{t=1}^{T} \langle p_t, f_t^2 \rangle,
\]

where \( KL(p||p_1) \) denotes the KL-divergence between the two probability distributions \( p \) and \( p_1 \).

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

\[
KL(p||p_t) = KL(p||p_{t+1}) = \int_{\mathcal{D}} p(w)\ln\left(\frac{p_{t+1}(w)}{p_t(w)}\right)dw.
\]

Moreover, by definition of \( p_{t+1}, \frac{1}{\eta} \left( KL(p||p_t) - KL(p||p_{t+1}) \right) = \frac{1}{\eta} \left( \int_{\mathcal{D}} p(w)\ln\left(\frac{p_{t+1}(w)}{p_t(w)}\right) \right) = -E_p[f_t(w)] - \frac{1}{\eta} \ln E_p[e^{-\eta f_t(w)}] \) for any \( t = 1, 2, \ldots, T \). Then summing over \( T \) rounds,

\[
\sum_{t=1}^{T} \left[ -E_p[f_t(w)] - \frac{1}{\eta} \ln E_p[e^{-\eta f_t(w)}] \right] = \frac{1}{\eta} \left( KL(p||p_1) - KL(p||p_{T+1}) \right).
\]
Now adding $\sum_{t=1}^T f_t(w_t)$ to both sides, this further gives:

$$
\sum_{t=1}^T \left[ f_t(w_t) - \mathbb{E}_p[f_t(w)] \right] = \frac{1}{\eta} \left( K L(p||p_1) - K L(p||p_{T+1}) \right) + \sum_{t=1}^T \left( f_t(w_t) + \frac{1}{\eta} \ln \mathbb{E}_p[e^{-\eta f_t(w)}] \right)
$$

$$
\implies \sum_{t=1}^T \left[ f_t(w_t) - \mathbb{E}_p[f_t(w)] \right] \leq \frac{K L(p||p_1)}{\eta} + \sum_{t=1}^T \left( f_t(w_t) + \frac{1}{\eta} \ln \mathbb{E}_p[e^{-\eta f_t(w)}] \right)
$$

$$
\implies \sum_{t=1}^T \sum_{w_i} f_t(w_t) - \mathbb{E}_p[f_t(w)] \leq \frac{K L(p||p_1)}{\eta} + \frac{1}{\eta} \sum_{t=1}^T \mathbb{E}_{w_i} \left[ \eta f_t(w_t) + \ln \mathbb{E}_p[e^{-\eta f_t(w)}] - 1 \right]
$$

$$
\leq \frac{K L(p||p_1)}{\eta} + \frac{1}{\eta} \sum_{t=1}^T \sum_{w_i} \left[ \eta f_t(w_t) + 1 - \eta \mathbb{E}_{w_i} \left[ f_t(w) \right] + \mathbb{E}_{w_i} \left[ \eta^2 f_t^2(w) \right] \right]
$$

$$
= \frac{K L(p||p_1)}{\eta} + \frac{1}{\eta} \sum_{t=1}^T \left< p_t, f_t^2 \right>,
$$

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

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

1. The function $K_t^* \ell(\cdot)$ is $L$-Lipschitz.
2. $K_t^* \ell(y) \leq (1 - \lambda) \left( K_t^* q, \ell \right) + \lambda \ell(y) + 3\epsilon L$, where $\lambda$ is a constant.
3. For any $q \in \mathcal{Q}_t$, define operator $K_t^{(2)} q : \mathcal{G}_t \mapsto \mathbb{R}$ as:

$$
K_t^{(2)} q(y) := \int_{y' \in \mathcal{G}_t} (K_t(y, y'))^2 d\mathbb{Q}(y') \quad \forall y \in \mathcal{G}_t,
$$

then $\int_{y \in \mathcal{G}_t} \frac{K_t^{(2)} q(y)}{K_t q(y)} dy \leq B$, where $B = 2 \left( 1 + \ln \frac{1}{2} + \ln \left( \beta - \alpha \right) \right)$.

**Proof.**

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

$$
K_t^* \ell(y) = \left< K_t^* \delta_y, \ell \right> = \begin{cases} \mathbb{E}_{U \sim \text{unif}[0,1]} [\ell(U \bar{y} + (1 - U)y)], & \text{if } |y - \bar{y}| \geq \epsilon \\ \mathbb{E}_{U \sim \text{unif}[0,1]} [\ell(\bar{y} - \epsilon U)], & \text{if } |y - \bar{y}| < \epsilon \end{cases},
$$

which immediately implies the function $K_t^* \ell(\cdot)$ has the same Lipschitz parameter that of $\ell(\cdot)$.

2. We prove this part considering two cases separately:

**Case 1.** $|y - \bar{y}| \geq \epsilon$: By construction of $K_t^*$ (see Definition 4), we note that expectation of $y$ w.r.t. $q$ and $K_t^* q$, i.e. respectively $\bar{y} = \mathbb{E}_{y \sim q}[y]$ and $\int_{y \in \mathcal{G}_t} K_t^* q(y) dy$, and $\mathbb{E}_{y \sim K_t^* q}[y]$ can differ at most by $2\epsilon$, i.e. $|\mathbb{E}_{y \sim q}[y] - \mathbb{E}_{y \sim K_t^* q}[y]| \leq 2\epsilon$ [Bubeck et al., 2017]. We write, $\mathbb{E}_{y \sim K_t^* q}[y] = \mathbb{E}_{y \sim q}[y] + \psi$, clearly $\psi \in [-2\epsilon, 2\epsilon]$. Hence:

$$
\ell(\bar{y}) = \ell(\mathbb{E}_{y \sim K_t^* q}[y]) - \psi
$$

$$
\leq \ell \left( \int_{y \in \mathcal{G}_t} y K_t^* q(y) dy \right) + \psi L \leq \int_{y \in \mathcal{G}_t} \ell(y) K_t^* q(y) dy + 2\epsilon L
$$
where the first inequality follows using the $L$-lipschitzness of $\ell$ and the second inequality follows using Jensen’s inequality (since $\ell$ is convex). Now consider the case $|y - \bar{y}| \geq \epsilon$ in (15):

$$K_t^* \ell(y) = \mathbb{E}_{U \sim \text{unif}[0,1]} [\ell(U \bar{y} + (1 - U)y)] \leq \frac{\ell(\bar{y}) + \ell(y) + 2 \epsilon L}{2} \tag{16}$$

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:

$$K_t^* \ell(y) \leq \ell(\bar{y}) + \epsilon L = \ell(\bar{y})/2 + \ell(\bar{y})/2 + \epsilon L \leq (K_t^* \ell(y))/2 + (\ell(y) + \epsilon L)/2 + 2\epsilon L$$

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

3. For this part, note that:

$$\int_{y \in G_t} \frac{K_t^2(y, y')}{K_t(y)} \, dy \leq \int_{\alpha}^{\beta} \frac{1}{\max(|y - \bar{y}|, \epsilon)} \, dy$$

where (a) follows noting $K_t^2(y, y') \leq \frac{1}{\max(|y - \bar{y}|, \epsilon)}$, $\forall y, y' \in G_t$ which implies $K_t^2(y, y') \leq \frac{K_t(y)}{\max(|y - \bar{y}|, \epsilon)}$. 

### A.3. Proof of Lemma 6

**Proof.** For any $\ell_t : \mathbb{R} \to [0, C], t \in [T]$, define $\hat{\ell}_t : \mathbb{R} \to [0, C]$ such that $\hat{\ell}_t(y) = \mathbb{E}_{u \sim \text{unif}([0,1])} \ell_t(y + \delta u)$, for any $y \in \mathbb{R}$. Let us also define $\hat{f}_t(w) = \hat{\ell}_t(g_t(w; x_t)), \forall w \in W$. Let $g_t = g_t(w_t; x_t), \forall t \in [T]$.

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

$$\mathbb{E}_{u \sim \text{unif}([0,1])} \left[ \frac{\partial}{\partial u} \ell_t(g_t(w; x_t) + \delta u | w_t) \right] \nabla_w (g_t(w; x_t))$$

$$= \frac{d\hat{f}_t(y_t)}{dy_t} \nabla_w (g_t(w_t; x_t)) = \nabla_w \hat{f}_t(w_t) = \nabla_w \mathbb{E}_u \left[ \ell_t(g_t(w_t; x_t) + \delta u) \right],$$

where the first equality is due to Lemma 1 of Flaxman et al., 2005 applied to the 1-dimensional ball $B_1(1)$.

### 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 $S \subset B_d(R) \subset \mathbb{R}^d$ be a convex set. $f_1, f_2, \ldots, f_T : S \rightarrow \mathbb{R}$ be a sequence of convex, differentiable functions. Let $w_1, w_2, \ldots, w_T \in S$ be a sequence of predictions defined as $w_1 = 0$ and $w_{t+1} = P_S(w_t - \eta_t h_t)$, where $\eta > 0$, and $h_1, h_2, \ldots, h_T$ are random variables such that $\mathbb{E}[h_t | w_t] = \nabla f_t(w_t)$, and $\|h_t\|_2 \leq G$, for some $G > 0$ then, for $\eta = \frac{R}{G \sqrt{T}}$, the expected regret incurred by above prediction sequence is:

$$\mathbb{E}\left[ \sum_{t=1}^{T} f_t(w_t) \right] - \min_{w \in S} \sum_{t=1}^{T} f_t(w) \leq RG \sqrt{T}.$$ 

Coming back to our problem setup, let us first denote $\hat{f}_t(w) = \ell_t(g_t(w; x_t))$, for all $w \in W$, $t \in [T]$ (recall from the proof of Lemma 6, we define $\hat{f}_t : \mathbb{R} \rightarrow [0, C]$ such that $\hat{f}_t(y) = \mathbb{E}_{u \sim U(1,1)} \ell_t(y + \delta u)$, for any $y \in \mathbb{R}$). We can now apply Lemma 12 in the setting of Algorithm 3 on the sequence of convex (by Lemma 6, we define $\hat{w}$), differentiable functions $f_1, f_2, \ldots, f_T : W_o \rightarrow [0, C]$, with $h_t = \frac{1}{\alpha} (\ell_t(a_t) u) \nabla g_t(w_t; x_t)$, with $u \sim B(1)$ (note that Lemma 6 implies $\mathbb{E}[h_t | w_t] = \nabla_w \hat{f}_t(w_t) = \nabla_w \mathbb{E}_u [\ell_t(g_t(w_t; x_t) + \delta u)]$). We get:

$$\mathbb{E}\left[ \sum_{t=1}^{T} \hat{f}_t(w_t) \right] - \min_{w \in W_o} \sum_{t=1}^{T} \hat{f}_t(w) \leq \frac{W \sqrt{T}}{C},$$

(17)

as in this case $R \leq (1 - \alpha)W < W$, and, by (A3) (ii), $\|h_t\|_2 = \frac{1}{2} (\ell_t(a_t) u) \nabla (g_t(w_t; x_t)) \leq \frac{DC}{\sqrt{\alpha}}$, so $G = \frac{DC}{\sqrt{\alpha}}$, and $\eta = \frac{W \sqrt{T}}{\sqrt{C}}$. Further, since $\ell_t(\cdot)$s are assumed to be $L$-Lipschitz, (17) yields:

$$\mathbb{E}\left[ \sum_{t=1}^{T} (f_t(w_t) - \delta L) \right] - \min_{w \in W_o} \sum_{t=1}^{T} (f_t(w) + \delta L) \leq \frac{W DC \sqrt{T}}{\delta},$$

$$\Rightarrow \mathbb{E}\left[ \sum_{t=1}^{T} f_t(w_t) \right] - \min_{w \in W_o} \sum_{t=1}^{T} f_t(w) \leq \frac{W DC \sqrt{T}}{\delta} + 2\delta LT,$$

$$\Rightarrow \mathbb{E}\left[ \sum_{t=1}^{T} f_t(w_t) \right] - \min_{w \in W_o} \sum_{t=1}^{T} f_t(w) \leq \frac{W DC \sqrt{T}}{\delta} + 2\delta LT + \alpha LT,$$

$$\Rightarrow \mathbb{E}\left[ \sum_{t=1}^{T} f_t(w_t) \right] - \min_{w \in W} \sum_{t=1}^{T} f_t(w) \leq \frac{W DC \sqrt{T}}{\delta} + 3\delta LT,$$

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

$$\mathbb{E}[R_T(A)] = \mathbb{E}\left[ \sum_{t=1}^{T} f_t(w_t) \right] - \min_{w \in W} \sum_{t=1}^{T} f_t(w) \leq 2\sqrt{3WLDCT^{3/4}},$$

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’ $W$; in particular, to maintain and update the probability distribution $p_t$ over $W$, and to sample from $p_t$ given $y_t$ at round $t$. Towards this we use an epsilon-net trick to discretize $W$ into finitely many points—specifically, since we choose $W = B_d(1)$, we discretize the $[0, 1]$ interval every $d$ direction with a grid size of $O(1/d)$, and consider only the points inside $B_d(1)$. This reduces the action space $W$ 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 $p_1 \leftarrow 1/N$ for all $N$ points in the epsilon net).