

# Online Companion: Smoothness-Adaptive Contextual Bandits

Yonatan Gur  
Stanford University

Ahmadreza Momeni  
Stanford University

Stefan Wager\*  
Stanford University

October 14, 2021

## 1 Proofs of main results

### 1.1 Proof of Part 1 of Theorem 3.1

In this section, we describe the proof of Part 1 of Theorem 3.1. The proof follows the next steps. In Step 1, we discuss some notations and definitions including the definition of inferior sampling rate. In Step 2, we leverage a result from Rigollet and Zeevi (2010) to connect regret and inferior sampling rate, which enables one to simplify analysis by focusing on the inferior sampling rate throughout the proof.

In Step 3, which is a key step of the proof, we reduce the problem at hand to a hypothesis testing problem by introducing a novel construction of a set of problem instances. This set consists of a nominal problem instance with smoothness parameter  $\gamma$  and some other problem instances with smoothness parameter  $\beta$ , each of which differs from the nominal one only over a specific region of the covariate space. These problem instances are designed to connect between the amount of exploration and the ability to identify the correct smoothness parameter. This construction is designed for showing that if a policy achieves rate-optimal performance for smooth problems, it is likely to underexplore in “rougher” problems, and hence not be able to differentiate between the two.

In Step 4, we verify that the aforementioned problem instances satisfy the margin condition. In Step 5, we show that with high probability, the number of covariates that belong to the regions mentioned in Step 3 grow linearly with respect to the time horizon and the volume of the regions. In Steps 6 and 7, we show that since the policy is rate optimal for  $\gamma$ -smooth problems, it cannot distinguish between the nominal problem and at least one of the  $\beta$ -smooth problems. In Step 8, we lower bound the inferior sampling rate due to not being able to identify the correct smoothness parameter (some high-level ideas

---

\*Correspondence: ygur@stanford.edu, amomenis@stanford.edu, swager@stanford.edu.

in Steps 7 and 8 are adopted from the proof of Theorem 3 in Locatelli and Carpentier (2018)). In Step 9, we revert back the lower bound on inferior sampling rate to a lower bound on regret.

**Step 1 (Preliminaries).** For any policy  $\pi$  and decision horizon length  $T$ , let  $\mathcal{S}^\pi(\mathbf{P}; T)$  be the *inferior sampling rate* defined as

$$\mathcal{S}^\pi(\mathbf{P}; T) := \mathbb{E}^\pi \left[ \sum_{t=1}^T \mathbb{1} \{f_{\pi_t^*}(X_t) \neq f_{\pi_t}(X_t)\} \right]. \quad (1.1)$$

Fix a covariate distribution  $\mathbf{P}_X$ . For any policy  $\pi$  and function  $f : [0, 1]^d \rightarrow [0, 1]$ , denote by  $\mathcal{S}^\pi(f; T)$  the inferior sampling rate of  $\pi$  when  $\mathbf{P}_X$  is the covariate distribution,  $\mathbb{E}[Y_{1,t} | X_t] = f(X_t)$ , and  $\mathbb{E}[Y_{2,t} | X_t] = \frac{1}{2}$ . Notably, the oracle policy  $\pi_f^*$  is given by  $\pi_f^*(x) = 2 - \mathbb{1} \{f(x) \geq \frac{1}{2}\}$ . We further denote by  $\mathbb{P}_{\pi, f}$  and  $\mathbb{E}_{\pi, f}$  the corresponding probability and expectation. Finally, for any Hölder exponent  $\beta > 0$  and margin parameter  $\alpha > 0$ , define:

$$\mathcal{R}_{\beta, \alpha}^\pi(T) := \sup_{\mathbf{P} \in \mathcal{P}(\beta, \alpha, d)} \mathcal{R}^\pi(\mathbf{P}; T); \quad \mathcal{S}_{\beta, \alpha}^\pi(T) := \sup_{\mathbf{P} \in \mathcal{P}(\beta, \alpha, d)} \mathcal{S}^\pi(\mathbf{P}; T).$$

Fix  $T \geq 1$ , two Hölder exponents  $0 < \beta < \gamma \leq 1$ , a margin parameter  $0 \leq \alpha \leq \frac{1}{\gamma}$ , a positive Lipschitz constant  $L$ , and positive constants  $\underline{\rho}, \bar{\rho}$  such that  $\mathbf{P}_X$  satisfies Assumption 2 with parameters  $\underline{\rho}, \bar{\rho}$ .

**Step 2 (From regret to inferior sampling rate).** The following lemma implies that it suffices to first analyze inferior sampling rate and then, revert the result back to regret.

**Lemma 1.1** (Rigollet and Zeevi 2010, Lemma 3.1). *For any  $\alpha > 0$  under the margin condition in Assumption 3, one has*

$$\mathcal{S}^\pi(\mathbf{P}; T) \leq C_{sr} T^{\frac{1}{\alpha+1}} [\mathcal{R}^\pi(\mathbf{P}; T)]^{\frac{\alpha}{\alpha+1}},$$

for any policy  $\pi$  and some positive constant  $C_{sr}$ .

By Lemma 1.1, we have  $\mathcal{S}_{\beta, \alpha}^\pi(T) \leq C_{sr} T^{\frac{1}{\alpha+1}} [\mathcal{R}_{\beta, \alpha}^\pi(T)]^{\frac{\alpha}{\alpha+1}}$ . Note that when  $\pi$  is rate-optimal over  $\mathcal{P}(\gamma, \alpha, d)$ , Lemma 1.1 implies that for some constants  $C_r, C_s > 0$ , one has:

$$\mathcal{R}_{\gamma, \alpha}^\pi(T) \leq C_r T^{1 - \frac{\gamma(1+\alpha)}{2\gamma+d}} =: \mathcal{R}_{\gamma, \alpha}^*(T); \quad \mathcal{S}_{\gamma, \alpha}^\pi(T) \leq C_s T^{1 - \frac{\gamma\alpha}{2\gamma+d}} =: \mathcal{S}_{\gamma, \alpha}^*(T).$$

**Step 3 (Constructing problem instances).** In this step we reduce our problem to a hypothesis testing problem. In order to do so, we first construct some problem instances. Defining  $M := \lceil \Delta^{\alpha - \frac{d}{\beta}} \rceil$  and  $C_\phi := \frac{L}{2^{2+2\beta}}$ , fix the parameter  $\Delta > 0$  such that

$$\frac{64C_\phi^2 \Delta^2 \mathcal{S}_{\gamma, \alpha}^*(T)}{3M} = \frac{1}{2}.$$

This selection of  $\Delta$  implies that for large enough  $T$  one has  $C_\phi \Delta \leq \frac{1}{4}$ . For any  $0 < \kappa \leq 1$ , define the functions  $\tilde{\psi}_\kappa$  and  $\hat{\psi}_\kappa$  as follows:

$$\tilde{\psi}_\kappa(x) := \begin{cases} |1 - \|x\|_\infty|^\kappa & \text{if } 0 \leq \|x\|_\infty \leq 1; \\ 0 & \text{o.w.}; \end{cases} \quad \hat{\psi}_\kappa(x) := \begin{cases} |1 - \|x\|_\infty|^\kappa & \text{if } 0 \leq \|x\|_\infty \leq 1; \\ -\|\|x\|_\infty - 1\|^\kappa & \text{if } 1 \leq \|x\|_\infty \leq 2; \\ -1 & \text{o.w.} \end{cases}$$

Note that  $\tilde{\psi}_\kappa \in \mathcal{H}_{\mathbb{R}^d}(\kappa, 1)$  and  $\hat{\psi}_\kappa \in \mathcal{H}_{\mathbb{R}^d}(\kappa, 2)$ . The following two lemmas (proved in Appendix 5) are the main tools to analyze the smoothness of the payoff functions that we construct in this step.

**Lemma 1.2** (Scaling and smoothness). *Suppose  $f \in \mathcal{H}_{\mathbb{R}^d}(\beta, L)$  for some  $0 < \beta \leq 1$  and  $L > 0$ , and define the function  $g$  such that  $g(x) = C^{-\beta} f(Cx)$  for all  $x \in \mathbb{R}^d$  and some  $C > 0$ . Then,  $g \in \mathcal{H}_{\mathbb{R}^d}(\beta, L)$ .*

**Lemma 1.3** (Min/Max and smoothness). *Suppose  $f, g \in \mathcal{H}_{\mathcal{X}}(\beta, L)$  for some  $\mathcal{X} \subseteq \mathbb{R}^d$ ,  $0 < \beta \leq 1$  and  $L > 0$ , and define the functions  $h_1 := \max(f, g)$  and  $h_2 := \min(f, g)$ . Then,  $h_1, h_2 \in \mathcal{H}_{\mathcal{X}}(\beta, L)$ .*

Define a hypercube  $H_0 := [0, 2\Delta^{\frac{\alpha}{d}}]^d$  with a center  $a_0 := (\Delta^{\frac{\alpha}{d}}, \Delta^{\frac{\alpha}{d}}, \dots, \Delta^{\frac{\alpha}{d}}) \in \mathbb{R}^d$ . Define the function

$$\phi_0(x) := \frac{1}{2} - C_\phi \cdot \min \left\{ \Delta, \Delta^{\frac{\alpha\gamma}{d}} \cdot \tilde{\psi}_\gamma \left( \Delta^{-\frac{\alpha}{d}} [x - a_0] \right) \right\}.$$

By Lemmas 1.2 and 1.3,  $\phi_0 \in \mathcal{H}(\gamma, L)$  since  $C_\phi \leq L$ . Consider the grid  $G$  that partitions the hypercube  $H_0$  into  $M$  disjoint hypercubes  $(H_m)_{m \in \{1, \dots, M\}}$ . Let  $a_m \in \mathbb{R}^d, m \in \{1, \dots, M\}$ , be the center of the hypercube  $H_m$ . Let  $\tilde{H}_m$  be the hypercube of side-length  $l := \frac{\Delta^{\frac{\alpha}{d}}}{4M^{\frac{1}{d}}}$  centered around  $a_m$ . Note that  $\tilde{H}_m \subset H_m$  and that the side-length of  $H_m$  is  $4l$ . Define the functions  $\phi_m, m \in \{1, \dots, M\}$ , as follows:

$$\phi_m(x) := \max \left\{ \phi_0(x), \frac{1}{2} + C_\phi \cdot \Delta \cdot \hat{\psi}_\beta (2l^{-1} [x - a_m]) \right\}.$$

Since  $C_\phi = \frac{L}{2^{2+2\beta}}$  and  $\Delta \leq 2^{\beta+1} l^\beta$  for large enough  $T$ , by Lemmas 1.2 and 1.3 one has that  $\phi_m \in \mathcal{H}(\beta, L)$  for  $1 \leq m \leq M$ .

**Step 4 (Verifying the margin condition).** By examining different cases of parametric values, we verify that the margin condition is satisfied with parameters  $\alpha$  and  $C_0 := 2^{d+3} d \bar{\rho} C_\phi^{-\alpha}$  when  $f_1 = \phi_m$  and  $f_2 = \frac{1}{2}$  for all  $0 \leq m \leq M$ .

- For  $m = 0$  and  $\delta \leq C_\phi \Delta$ , one has

$$\begin{aligned}
\mathbf{P}_X \left\{ 0 < \left| \phi_0(X) - \frac{1}{2} \right| \leq \delta \right\} &\leq \bar{\rho} \int_{H_0} \mathbb{1} \left\{ C_\phi \Delta^{\frac{\alpha\gamma}{d}} \cdot \tilde{\psi}_\gamma \left( \Delta^{-\frac{\alpha}{d}} [x - a_0] \right) \leq \delta \right\} dx \\
&\leq 2^d \bar{\rho} \Delta^\alpha \int_{[0,1]^d} \mathbb{1} \left\{ \tilde{\psi}_\gamma(x) \leq \delta C_\phi^{-1} \Delta^{-\frac{\alpha\gamma}{d}} \right\} dx \\
&\leq 2^d \bar{\rho} \Delta^\alpha \left[ 1 - \int_{[0,1]^d} \mathbb{1} \left\{ \|x\|_\infty \leq 1 - \delta^{\frac{1}{\gamma}} C_\phi^{-\frac{1}{\gamma}} \Delta^{-\frac{\alpha}{d}} \right\} dx \right] \\
&\leq 2^d \bar{\rho} \Delta^\alpha \left[ 1 - \left( 1 - \delta^{\frac{1}{\gamma}} C_\phi^{-\frac{1}{\gamma}} \Delta^{-\frac{\alpha}{d}} \right)^d \right] \\
&\leq 2^d \bar{\rho} \Delta^\alpha \left[ d \delta^{\frac{1}{\gamma}} C_\phi^{-\frac{1}{\gamma}} \Delta^{-\frac{\alpha}{d}} \right] \stackrel{(a)}{\leq} 2^d d \bar{\rho} C_\phi^{-\alpha} \Delta^{\frac{1}{\gamma} - \frac{\alpha}{d}} \delta^\alpha \leq 2^d d \bar{\rho} C_\phi^{-\alpha} \delta^\alpha, \quad (1.2)
\end{aligned}$$

where (a) holds since  $\alpha \leq \frac{1}{\gamma}$ .

- For  $m = 0$  and  $\delta > C_\phi \Delta$ , one has

$$\mathbf{P}_X \left\{ 0 < \left| \phi_0(X) - \frac{1}{2} \right| \leq \delta \right\} \leq 2^d \bar{\rho} \Delta^\alpha \leq 2^d \bar{\rho} C_\phi^{-\alpha} \delta^\alpha.$$

- For  $1 \leq m \leq M$  and  $\delta \leq C_\phi \Delta$ , one has

$$\begin{aligned}
\mathbf{P}_X \left\{ 0 < \left| \phi_m(X) - \frac{1}{2} \right| \leq \delta \right\} &\leq \mathbf{P}_X \left\{ 0 < C_\phi \cdot \Delta \cdot |1 - 2l^{-1} \|x - a_m\|_\infty|^\beta \leq \delta, X \in \tilde{H}_m \right\} \\
&\quad + \mathbf{P}_X \left\{ 0 < C_\phi \cdot \Delta \cdot |2l^{-1} \|x - a_m\|_\infty - 1|^\beta \leq \delta, X \in H_m \setminus \tilde{H}_m \right\} \\
&\quad + \mathbf{P}_X \left\{ 0 < \left| \phi_0(X) - \frac{1}{2} \right| \leq \delta \right\}. \quad (1.3)
\end{aligned}$$

Next, we analyze each term separately. One has

$$\begin{aligned}
\mathbf{P}_X \left\{ 0 < C_\phi \cdot \Delta \cdot |1 - 2l^{-1} \|x - a_m\|_\infty|^\beta \leq \delta, X \in \tilde{H}_m \right\} \\
&\leq \bar{\rho} \int_{\tilde{H}_m} \mathbb{1} \left\{ C_\phi \cdot \Delta \cdot |1 - 2l^{-1} \|x - a_m\|_\infty|^\beta \leq \delta \right\} dx \\
&\leq \bar{\rho} \int_{\tilde{H}_m} \mathbb{1} \left\{ \|x - a_m\|_\infty \geq \frac{l}{2} \left( 1 - C_\phi^{-\frac{1}{\beta}} \Delta^{-\frac{1}{\beta}} \delta^{\frac{1}{\beta}} \right) \right\} dx \\
&= \bar{\rho} 2^{-d} l^d \left[ 1 - \left( 1 - C_\phi^{-\frac{1}{\beta}} \Delta^{-\frac{1}{\beta}} \delta^{\frac{1}{\beta}} \right)^d \right] \\
&\stackrel{(a)}{\leq} \bar{\rho} 2^{-d} l^d \left[ d \delta^{\frac{1}{\beta}} C_\phi^{-\frac{1}{\beta}} \Delta^{-\frac{1}{\beta}} \right] \stackrel{(b)}{\leq} d \bar{\rho} 2^{-2d} C_\phi^{-\alpha} \delta^\alpha, \quad (1.4)
\end{aligned}$$

where (a) follows from the inequality  $(1-x)^r \geq 1-rx$  for  $0 \leq x \leq 1, r \geq 1$ , and (b) holds by  $\alpha \leq \frac{1}{\gamma}$ . Similarly,

$$\begin{aligned}
\mathbf{P}_X \left\{ 0 < C_\phi \Delta |2l^{-1} \|x - a_m\|_\infty - 1|^\beta \leq \delta, X \in H_m \setminus \tilde{H}_m \right\} \\
&\leq \bar{\rho} \int_{H_m \setminus \tilde{H}_m} \mathbb{1} \left\{ C_\phi \Delta |2l^{-1} \|x - a_m\|_\infty - 1|^\beta \leq \delta \right\} dx \\
&\leq \bar{\rho} \int_{H_m \setminus \tilde{H}_m} \mathbb{1} \left\{ \|x - a_m\|_\infty \leq \frac{l}{2} \left( 1 + C_\phi^{-\frac{1}{\beta}} \Delta^{-\frac{1}{\beta}} \delta^{\frac{1}{\beta}} \right) \right\} dx \\
&= \bar{\rho} 2^{-d} l^d \left[ \left( 1 + C_\phi^{-\frac{1}{\beta}} \Delta^{-\frac{1}{\beta}} \delta^{\frac{1}{\beta}} \right)^d - 1 \right] \\
&\stackrel{(a)}{\leq} \bar{\rho} l^d C_\phi^{-\frac{1}{\beta}} \Delta^{-\frac{1}{\beta}} \delta^{\frac{1}{\beta}} \stackrel{(b)}{\leq} \bar{\rho} C_\phi^{-\alpha} \delta^\alpha, \tag{1.5}
\end{aligned}$$

where (a) follows from the inequality  $(1+x)^r \leq 2^r x + 1$  for  $0 \leq x \leq 1, r \geq 1$ , and (b) holds by  $\alpha \leq \frac{1}{\gamma}$ . Putting together (1.2), (1.3), (1.4), and (1.5), yields for  $\delta \leq C_\phi \Delta$ :

$$\mathbf{P}_X \left\{ 0 < |\phi_m(X) - \frac{1}{2}| \leq \delta \right\} \leq C_0 \delta^\alpha.$$

- The case  $1 \leq m \leq M$  and  $\delta > C_\phi \Delta$  can be analyzed similar to the case  $m = 0$  and  $\delta > C_\phi \Delta$ .

**Step 5 (Desirable event).** For  $m \in \{1, \dots, M\}$ , define  $Q_m := \sum_{t=1}^T \mathbb{1} \{X_t \in \tilde{H}_m\} =: \sum_{t=1}^T Z_{m,t}$  to be the number of times periods at which the realized covariates belong to the hypercube  $\tilde{H}_m$ . Define  $\mathcal{A} := \left\{ \exists m \in \{1, \dots, M\} : Q_m < \frac{\rho}{2} T l^d \text{ or } Q_m > \frac{3\bar{\rho}}{2} T l^d \right\}$  to be the event where  $Q_m$  is less than  $\frac{\rho}{2} T l^d$  or larger than  $\frac{3\bar{\rho}}{2} T l^d$  for at least one value of  $m \in \{1, \dots, M\}$ . Note that

$$\mathbb{P} \{ \mathcal{A} \} \leq \sum_{m=1}^M \mathbb{P} \left\{ Q_m < \frac{\rho}{2} T l^d \right\} + \mathbb{P} \left\{ Q_m > \frac{3\bar{\rho}}{2} T l^d \right\}.$$

In order to bound each of the summands on the right hand side of the above inequality, one may apply Bernstein's inequality in the following lemma 1.4 to  $Q_m$ :

**Lemma 1.4** (Bernstein inequality). *Let  $X_1, \dots, X_n$  be random variables with range  $|X_i| \leq M$  and  $\sum_{t=1}^n \text{Var} [X_t | X_{t-1}, \dots, X_1] = \sigma^2$ . Let  $S_n = X_1 + \dots + X_n$ . Then for all  $a \geq 0$*

$$\mathbb{P} \{ S_n \geq \mathbb{E}[S_n] + a \} \leq \exp \left( -\frac{a^2/2}{\sigma^2 + Ma/3} \right).$$

Note that since  $\underline{\rho} l^d T \leq \mathbb{E} Z_{m,t} \leq \bar{\rho} T l^d$ ,  $|Z_{m,t}| \leq 1$ , and  $\text{Var} Z_{m,t} \leq \mathbb{E} Z_{m,t}^2 \leq 2\bar{\rho} l^d$ , one obtains:

$$\begin{aligned}
\mathbb{P}\{\mathcal{A}\} &\leq M \exp\left(-c_1 T l^d / 5\right) \\
&\stackrel{(a)}{\leq} c_2 [\mathcal{S}_{\gamma,\alpha}^*]^{-\frac{\alpha\beta-d}{2\beta+d-\alpha\beta}} \exp\left(-c_3 T [\mathcal{S}_{\gamma,\alpha}^*]^{-\frac{d}{2\beta+d-\alpha\beta}}\right) \\
&\leq c_2 T^{-\frac{\alpha\beta-d}{2\beta+d-\alpha\beta}} \exp\left(-c_4 T^{-\frac{2\beta(2\gamma+d-\alpha\gamma)+\alpha d(\gamma-\beta)}{(2\gamma+1)(2\beta+d-\alpha\beta)}}\right) \stackrel{(b)}{\leq} c_5 T^{-3},
\end{aligned}$$

for large enough  $T$  and constants  $c_1, c_2, c_3, c_4, c_5 > 0$ , where (a) follows from the definition of  $M$  and  $l$ , and (b) holds by  $\frac{2\beta(2\gamma+d-\alpha\gamma)+\alpha d(\gamma-\beta)}{(2\gamma+1)(2\beta+d-\alpha\beta)} > 0$  for  $\alpha \leq \frac{1}{\gamma}$ . For any problem instance  $\mathbf{P}$  and horizon length  $T$ , denote the inferior sampling rate of  $\pi$  when the event  $\mathcal{A}$  does not occur by

$$\bar{\mathcal{S}}^\pi(\mathbf{P}; T) := \mathbb{E}^\pi \left[ \sum_{t=1}^T \mathbb{1}\{f_{\pi_t^*}(X_t) \neq f_{\pi_t}(X_t)\} \mid \bar{\mathcal{A}} \right].$$

Define  $\bar{\mathcal{S}}_{\gamma,\alpha}^\pi(T) := \sup_{\mathbf{P} \in \mathcal{P}(\gamma,\alpha,d)} \bar{\mathcal{S}}^\pi(\mathbf{P}; T)$ . Note that

$$(1 - \mathbb{P}\{\mathcal{A}\}) \bar{\mathcal{S}}^\pi(\mathbf{P}; T) \leq \mathcal{S}^\pi(\mathbf{P}; T) \leq \bar{\mathcal{S}}^\pi(\mathbf{P}; T) + T \mathbb{P}\{\mathcal{A}\},$$

which implies that

$$|\bar{\mathcal{S}}_{\gamma,\alpha}^\pi(T) - \mathcal{S}_{\gamma,\alpha}^\pi(T)| \leq c_4 T^{-2}. \quad (1.6)$$

For the rest of the proof, all probabilities and expectations will be computed conditional on  $\bar{\mathcal{A}}$ .

**Step 6 (Selecting a single problem with smoothness  $\beta$ ).** Let  $N_{m,T} := \sum_{t=1}^T \mathbb{1}\{\pi_t = 1, X_t \in H_m\}$  denote the number of times policy  $\pi$  selects arm 1 when realized covariates belong to the hypercube  $H_m$ . By definition,  $\mathbb{E}_{\pi,\phi_0}^T \left[ \sum_{m=1}^M N_{m,T} \mid \bar{\mathcal{A}} \right] \leq \bar{\mathcal{S}}_{\gamma,\alpha}^\pi(T)$ , implying that there exists some  $m^* \in \{1, \dots, M\}$  such that

$$\mathbb{E}_{\pi,\phi_0}^T [N_{m^*,T} \mid \bar{\mathcal{A}}] \leq \frac{\bar{\mathcal{S}}_{\gamma,\alpha}^\pi(T)}{M} \leq \frac{\mathcal{S}_{\gamma,\alpha}^\pi(T)}{M} + c_4 T^{-2},$$

where the last inequality holds by (1.6).

**Step 7 (Likelihood of distinguishing between different smoothness parameters).** We show that policy  $\pi$  cannot distinguish between  $\phi_0$  and  $\phi_{m^*}$  with a strictly positive probability. For any set of samples  $\{(\pi_t, X_t, Y_{\pi_t,t})\}_{t=1}^T$ , define the log-likelihood ratio  $L_{m,T} = L_{m,T} \left( \{(\pi_t, X_t, Y_{\pi_t,t})\}_{t=1}^T \right)$  for  $m \in \{1, \dots, M\}$  as:

$$\begin{aligned}
L_{m,T} &:= \sum_{t=1}^T \log \left( \frac{\mathbb{P}_{\pi, \phi_0} \{Y_{\pi_t, t} \mid \pi_t, X_t\}}{\mathbb{P}_{\pi, \phi_m} \{Y_{\pi_t, t} \mid \pi_t, X_t\}} \right) \\
&\leq \sum_{t=1}^T \mathbb{1} \{ \pi_t = 1, X_t \in H_m \} \cdot \left[ Y_{\pi_t, t} \log \left( \frac{\phi_0(X_t)}{\phi_m(X_t)} \right) + (1 - Y_{\pi_t, t}) \log \left( \frac{(1 - \phi_0(X_t))}{(1 - \phi_m(X_t))} \right) \right] \\
&\leq \sum_{t=1}^T \mathbb{1} \{ \pi_t = 1, X_t \in H_m \} \cdot \left[ Y_{\pi_t, t} \frac{(\phi_0(X_t) - \phi_m(X_t))}{\phi_m(X_t)} + (1 - Y_{\pi_t, t}) \frac{(\phi_m(X_t) - \phi_0(X_t))}{(1 - \phi_m(X_t))} \right] \\
&= \sum_{t=1}^T \mathbb{1} \{ \pi_t = 1, X_t \in H_m \} \cdot \frac{(Y_{\pi_t, t} - \phi_m(X_t))(\phi_0(X_t) - \phi_m(X_t))}{\phi_m(X_t)(1 - \phi_m(X_t))},
\end{aligned}$$

where the last inequality follows from  $\log(1+x) \leq x$  for all  $x > 0$ . By taking expectations of the above inequality and conditioning on the event  $\bar{\mathcal{A}}$  for  $m = m^*$ , one obtains:

$$\begin{aligned}
\mathbb{E}_{\pi, \phi_0} [L_{m^*, T} \mid \bar{\mathcal{A}}] &\leq \mathbb{E}_{\pi, \phi_0} \left[ \sum_{t=1}^T \mathbb{1} \{ \pi_t = 1, X_t \in H_{m^*} \} \cdot \frac{(Y_{\pi_t, t} - \phi_{m^*}(X_t))(\phi_0(X_t) - \phi_{m^*}(X_t))}{\phi_{m^*}(X_t)(1 - \phi_{m^*}(X_t))} \mid \bar{\mathcal{A}} \right] \\
&= \mathbb{E}_{\pi, \phi_0} \left[ \sum_{t=1}^T \mathbb{E} \left[ \mathbb{1} \{ \pi_t = 1, X_t \in H_{m^*} \} \cdot \frac{(Y_{\pi_t, t} - \phi_{m^*}(X_t))(\phi_0(X_t) - \phi_{m^*}(X_t))}{\phi_{m^*}(X_t)(1 - \phi_{m^*}(X_t))} \mid X_t \right] \mid \bar{\mathcal{A}} \right] \\
&= \mathbb{E}_{\pi, \phi_0} \left[ \sum_{t=1}^T \mathbb{1} \{ \pi_t = 1, X_t \in H_{m^*} \} \cdot \frac{(\phi_0(X_t) - \phi_{m^*}(X_t))^2}{\phi_{m^*}(X_t)(1 - \phi_{m^*}(X_t))} \mid \bar{\mathcal{A}} \right] \\
&\stackrel{(a)}{\leq} \frac{64C_\phi^2 \Delta^2}{3} \mathbb{E}_{\pi, \phi_0} \left[ \sum_{t=1}^T \mathbb{1} \{ \pi_t = 1, X_t \in H_{m^*} \} \mid \bar{\mathcal{A}} \right] \\
&= \frac{64C_\phi^2 \Delta^2}{3} \mathbb{E}_{\pi, \phi_0} [N_{m^*, T} \mid \bar{\mathcal{A}}] \stackrel{(b)}{\leq} \frac{64C_\phi^2 \Delta^2 \mathcal{S}_{\gamma, \alpha}^*(T)}{3M} + c_4 T^{-2} \stackrel{(c)}{\leq} 1, \tag{1.7}
\end{aligned}$$

for large enough  $T$ , where: (a) follows from  $C_\phi \Delta \leq \frac{1}{4}$ ; (b) follows from the definition of  $m^*$ ; and (c) holds by the definition of  $\Delta$ .

**Step 8 (Lower bound on regret as a function of inferior sampling rate).** Let  $\tilde{N}_{m, T} := \sum_{t=1}^T \mathbb{1} \{ \pi_t = 1, X_t \in \tilde{H}_m \}$  denote the number of times policy  $\pi$  selects arm 1 when realized covariates belong to the hypercube  $\tilde{H}_m$ . We next use two lemmas in order to show that with a strictly positive probability one has  $\tilde{N}_{m^*, T} < \frac{\rho T l^d}{2}$  conditional on the event  $\bar{\mathcal{A}}$  under problem  $m^*$ , implying that  $\pi$  selects an inferior arm at least  $\frac{\rho T l^d}{2}$  times. The first lemma is a simple variation of Lemma 2.6 in Tsybakov (2008) and is proved for completeness in Appendix 5; the second lemma is a straightforward extension of Lemma 19 in Kaufmann et al. (2016).

**Lemma 1.5** (Hypothesis testing error probability). *Let  $\rho_0, \rho_1$  be two probability distributions supported on  $\mathcal{X}$ , with  $\rho_0$  absolutely continuous with respect to  $\rho_1$ . Then, for any measurable function  $\Psi : \mathcal{X} \rightarrow \{0, 1\}$ :*

$$\mathbb{P}_{\rho_0}\{\Psi(X) = 1\} + \mathbb{P}_{\rho_1}\{\Psi(X) = 0\} \geq \frac{1}{2} \exp(-\text{KL}(\rho_0, \rho_1)).$$

**Lemma 1.6** (Log-likelihood ratio and historical events). *For any event*

$\mathcal{E} \in \mathcal{F}_t^- = \sigma(\pi_1, X_1, Y_{\pi_1,1}, \dots, \pi_T, X_T, Y_{\pi_T,T})$  *and an arbitrary event*  $\mathcal{A}$ , *one has*

$$\mathbb{E}_{\pi, \phi_0} [L_{m,T} | \mathcal{E}, \mathcal{A}] \geq \log \left( \frac{\mathbb{P}_{\pi, \phi_0} \{\mathcal{E} | \mathcal{A}\}}{\mathbb{P}_{\pi, \phi_m} \{\mathcal{E} | \mathcal{A}\}} \right).$$

Denote by  $\rho_0$  and  $\rho_m$  the distributions of  $\tilde{N}_{m,T}$  under the problems 0 and  $m$  conditional on the event  $\bar{\mathcal{A}}$ . Define the test function  $\Psi(x) = \mathbb{1} \left\{ x \geq \frac{\rho T l^d}{2} \right\}$ . With this selection of  $\rho_0$ ,  $\rho_m$ , and  $\Psi$ , Lemma 1.5 yields:

$$\mathbb{P}_{\pi, \phi_0} \left\{ \tilde{N}_{m^*,T} \geq \frac{\rho T l^d}{2} \mid \bar{\mathcal{A}} \right\} + \mathbb{P}_{\pi, \phi_{m^*}} \left\{ \tilde{N}_{m^*,T} < \frac{\rho T l^d}{2} \mid \bar{\mathcal{A}} \right\} \geq \frac{1}{2} \exp(-\text{KL}(\rho_0, \rho_{m^*})).$$

To establish a lower bound on the right hand side of the above inequality, we note that:

$$\begin{aligned} \mathbb{E}_{\pi, \phi_0} [L_{m^*,T} | \bar{\mathcal{A}}] &= \sum_{s=1}^T \mathbb{E}_{\pi, \phi_0} [L_{m^*,T} | \bar{\mathcal{A}}, \tilde{N}_{m^*,T} = s] \mathbb{P}_{\pi, \phi_0} \left\{ \tilde{N}_{m^*,T} = s \mid \bar{\mathcal{A}} \right\} \\ &\geq \sum_{s=1}^T \log \left( \frac{\mathbb{P}_{\pi, \phi_0} \left\{ \tilde{N}_{m^*,T} = s \mid \bar{\mathcal{A}} \right\}}{\mathbb{P}_{\pi, \phi_{m^*}} \left\{ \tilde{N}_{m^*,T} = s \mid \bar{\mathcal{A}} \right\}} \right) \mathbb{P}_{\pi, \phi_0} \left\{ \tilde{N}_{m^*,T} = s \mid \bar{\mathcal{A}} \right\} = \text{KL}(\rho_0, \rho_{m^*}), \end{aligned}$$

where the inequality follows from Lemma 1.6. The last two inequalities, along with (1.7), yield

$$\mathbb{P}_{\pi, \phi_0} \left\{ \tilde{N}_{m^*,T} \geq \frac{\rho T l^d}{2} \mid \bar{\mathcal{A}} \right\} + \mathbb{P}_{\pi, \phi_{m^*}} \left\{ \tilde{N}_{m^*,T} < \frac{\rho T l^d}{2} \mid \bar{\mathcal{A}} \right\} \geq \frac{1}{2} \exp(-1).$$

Next, we show that  $\mathbb{P}_{\pi, \phi_0} \left\{ \tilde{N}_{m^*,T} \geq \frac{\rho T l^d}{2} \mid \bar{\mathcal{A}} \right\}$  is small. We apply Markov's inequality to obtain:

$$\begin{aligned} \mathbb{P}_{\pi, \phi_0} \left\{ \tilde{N}_{m^*,T} \geq \frac{\rho T l^d}{2} \mid \bar{\mathcal{A}} \right\} &\leq \frac{\mathbb{E}_{\pi, \phi_0} [\tilde{N}_{m^*,T} | \bar{\mathcal{A}}]}{\frac{\rho T l^d}{2}} \stackrel{(a)}{\leq} \frac{\mathcal{S}_{\gamma, \alpha}^*(T)}{M} + c_4 T^{-2} \\ &\stackrel{(b)}{\leq} \frac{2^d [\mathcal{S}_{\gamma, \alpha}^*]^{1 + \frac{\alpha \beta}{2\beta + d - \alpha \beta}} + c_4 l^{-d} T^{-2}}{\frac{\rho T}{2}} \leq c_5 T^{\frac{\alpha d(\beta - \gamma)}{(2\beta + d - \alpha \beta)(2\gamma + d)}} \stackrel{(c)}{\leq} \frac{1}{4} \exp(-1), \end{aligned}$$

for large enough  $T$  and some constant  $c_5 > 0$ , where: (a) follows from the definition of  $m^*$  and (1.6);

(b) holds due to the definition of  $l$  and  $M$ ; and (c) holds due to the fact that  $\frac{\alpha d(\beta - \gamma)}{(2\beta + d - \alpha \beta)(2\gamma + d)} < 0$  since

$\alpha \leq \frac{1}{\gamma}$ . The last two displays yield that for large enough  $T$ , one has  $\mathbb{P}_{\pi, \phi_{m^*}} \left\{ \tilde{N}_{m^*,T} < \frac{\rho T l^d}{2} \mid \bar{\mathcal{A}} \right\} \geq \frac{1}{4e}$ .

By definition, when event  $\bar{\mathcal{A}}$  holds, at least  $\rho T l^d$  times realized covariates belong to the hypercube  $\tilde{H}_{m^*}$

where  $f_1(x) \geq f_2(x) + \frac{\Delta}{2}$  for problem  $m^*$ , that is, for some constant  $c_6 > 0$ , one has:

$$\mathcal{R}_{\beta,\alpha}^\pi(T) \geq \frac{\rho T l^d}{2} \cdot \frac{\Delta}{2} \cdot \mathbb{P}_{\pi, \phi_{m^*}} \left\{ \tilde{N}_{m^*,T} < \frac{\rho T l^d}{2} \mid \bar{\mathcal{A}} \right\} \mathbb{P} \{ \bar{\mathcal{A}} \} \geq \frac{\rho T l^d}{32e} \geq c_6 T [\mathcal{S}_{\gamma,\alpha}^*(T)]^{-\frac{\beta+d}{2\beta+d-\alpha\beta}}. \quad (1.8)$$

The final result follows by noting that  $\mathcal{S}_{\gamma,\alpha}^\pi(T) \leq \tilde{C}_{sr} T^{\frac{1}{\alpha+1}} [T^{\zeta(\gamma,\alpha,d)}]^{\frac{\alpha}{\alpha+1}}$ . This concludes the proof. ■

## 1.2 Proof of Part 2 of Theorem 3.1

The proof follows similar lines of argument as in the proof of Part 1 of Theorem 3.1.

**Step 1 (Preliminaries).** Fix time horizon length  $T \geq 1$ , and some Hölder exponent  $\gamma > 1$ , some margin parameter  $0 \leq \alpha \leq 1$ , some positive Lipschitz constants  $L$ , and some positive constants  $\underline{\rho}, \bar{\rho}$  such that  $\mathbf{P}_X$ , the covariate distribution, satisfies Assumption 2 with parameters  $\underline{\rho}, \bar{\rho}$ .

**Step 2 (From regret to inferior sampling rate).** By Lemma 1.1, we have  $\mathcal{S}_{\beta,\alpha}^\pi(T) \leq C_{sr} T^{\frac{1}{\alpha+1}} \left[ \mathcal{R}_{\beta,\alpha}^\pi(T) \right]^{\frac{\alpha}{\alpha+1}}$ .

Note that by the assumption that  $\pi$  is rate-optimal over  $\mathcal{P}(\gamma, \alpha, d)$  and Lemma 1.1, one has

$$\mathcal{R}_{\gamma,\alpha}^\pi(T) \leq C_r T^{1-\frac{\gamma(1+\alpha)}{2\beta+d}} =: \mathcal{R}_{\gamma,\alpha}^*(T), \quad \mathcal{S}_{\gamma,\alpha}^\pi(T) \leq C_s T^{1-\frac{\gamma\alpha}{2\gamma+d}} =: \mathcal{S}_{\gamma,\alpha}^*(T),$$

for some constants  $C_r, C_s > 0$ .

**Step 3 (Constructing problem instances).** We will reduce our problem to a hypothesis testing problem. To do so, we construct some problem instances first. Define the parameter  $\Delta > 0$  such that

$$\frac{64C_\phi^2\Delta^2\mathcal{S}_{\gamma,\alpha}^*(T)}{3} = \frac{1}{2},$$

where we define  $C_\phi := \frac{L}{2^{2\beta}}$ . Note that the definition of  $\Delta$  implies that for large enough  $T$ , one has  $C_\phi\Delta \leq \frac{1}{4}$ . Define the function:

$$\phi_0(x) := \frac{1}{2} - C_\phi \cdot \left( \frac{1}{2} - x_1 \right).$$

Note that  $\phi_0 \in \mathcal{H}(\gamma, L)$  since  $C_\phi \leq L$ . Define the hypercube  $H := [\frac{1}{2} - \Delta, \frac{1}{2}] \times [0, 1]^{d-1}$ , with a center  $a_0 := (\frac{1-\Delta}{2}, \frac{1}{2}, \dots, \frac{1}{2}) \in \mathbb{R}^d$ , and the function:

$$\phi_1(x) := \phi_0(x) + 2C_\phi \cdot \Delta \cdot \tilde{\psi} \left( 2\Delta^{-1}[x - a_0] \right).$$

For any  $0 < \kappa \leq 1$ , define the functions  $\tilde{\psi}$ :

$$\tilde{\psi}(x) := \begin{cases} |1 - |x_1|| & \text{if } |x_1| \leq 1 \\ 0 & \text{o.w.} \end{cases}.$$

Note that by Lemmas 1.2 and 1.3,  $\phi_0 \in \mathcal{H}(1, L)$  since  $C_\phi \leq \frac{L}{2^{2\beta}}$ .

**Step 4 (Verifying the margin condition).** We verify that the margin condition is satisfied with parameters  $\alpha$  and  $C_0 := \frac{5\bar{\rho}}{2C_\phi}$  when  $f_1 = \phi_m$  and  $f_2 = \frac{1}{2}$  for all  $0 \leq m \leq 1$ .

- For  $m = 0$  and  $0 < \delta \leq 1$ , one has

$$\mathbf{P}_X \left\{ 0 < |\phi_0(X) - \frac{1}{2}| \leq \delta \right\} \leq \frac{2\bar{\rho}\delta}{C_\phi} \leq \frac{2\bar{\rho}\delta^\alpha}{C_\phi}.$$

- For  $m = 1$  and  $0 < \delta \leq 1$ , one has

$$\mathbf{P}_X \left\{ 0 < |\phi_1(X) - \frac{1}{2}| \leq \delta \right\} \leq \frac{5\bar{\rho}\delta}{2C_\phi} \leq \frac{5\bar{\rho}\delta^\alpha}{2C_\phi}.$$

**Step 5 (Desirable event).** Note that for  $x \in \tilde{H} := [\frac{1}{2} - \frac{7\Delta}{10}, \frac{1}{2} - \frac{\Delta}{6}] \times [0, 1]^{d-1}$ , the first arm is optimal by a gap of at least  $\frac{\Delta}{2}$  under the problem  $m = 1$ . Define  $Q := \sum_{t=1}^T \mathbb{1} \{X_t \in \tilde{H}\} =: \sum_{t=1}^T Z_t$  to be the number of times covariates fall into the hypercube  $\tilde{H}_m$  during the entire time horizon. Define the event

$$\mathcal{A} := \left\{ Q < \frac{2}{15} \underline{\rho} T \Delta \right\}$$

to be the event on which the number of covariates that have fallen into the hypercube  $\tilde{H}$  is less than  $\frac{2}{15} \underline{\rho} T \Delta$ . In order to bound  $\mathbb{P}\{\mathcal{A}\}$ , one can apply Bernstein's inequality in Lemma 1.4 to  $Q$  by noting that  $\mathbb{E}Z_t \geq \frac{2}{15} \underline{\rho} \Delta$ ,  $|Z_t| \leq 1$ , and  $\mathbb{V}\text{ar}Z_t \leq \mathbb{E}Z_t^2 \leq \frac{2}{15} \underline{\rho} \Delta$  to obtain

$$\mathbb{P}\{\mathcal{A}\} \leq \exp\left(-\frac{2}{15} \underline{\rho} T \Delta / 5\right) \stackrel{(a)}{\leq} \exp\left(-c_1 T [\mathcal{S}_{\gamma, \alpha}^*(T)]^{-\frac{1}{2}}\right) \leq \exp\left(-c_1 T^{\frac{(1+\frac{\alpha}{2})\gamma + \frac{d}{2}}{(2\gamma+d)\frac{1}{2}}}\right) \leq c_2 T^{-3},$$

for large enough  $T$  and constants  $c_1, c_2 > 0$ , where (a) follows from the definition of  $\Delta$ .

For any problem instance  $\mathbf{P}$  and time horizon length  $T$ , denote by

$$\bar{\mathcal{S}}^\pi(\mathbf{P}; T) := \mathbb{E}^\pi \left[ \sum_{t=1}^T \mathbb{1} \{f_{\pi_t^*}(X_t) \neq f_{\pi_t}(X_t)\} \mid \bar{\mathcal{A}} \right]$$

the inferior sampling rate of  $\pi$  when the event  $\mathcal{A}$  fails, and let  $\bar{\mathcal{S}}_{\gamma, \alpha}^\pi(T) := \sup_{\mathbf{P} \in \mathcal{P}(\gamma, \alpha, d)} \bar{\mathcal{S}}^\pi(\mathbf{P}; T)$ . Note that

$$(1 - \mathbb{P}\{\mathcal{A}\}) \bar{\mathcal{S}}^\pi(\mathbf{P}; T) \leq \mathcal{S}^\pi(\mathbf{P}; T) \leq \bar{\mathcal{S}}^\pi(\mathbf{P}; T) + T \mathbb{P}\{\mathcal{A}\},$$

which implies

$$|\bar{\mathcal{S}}_{\gamma,\alpha}^\pi(T) - \mathcal{S}_{\gamma,\alpha}^\pi(T)| \leq c_2 T^{-2}. \quad (1.9)$$

For the rest of the proof probabilities and expectations will be computed conditional on the event  $\bar{\mathcal{A}}$ .

**Step 6 (Likelihood of distinguishing different smoothness parameters).** In this step, we will show that policy  $\pi$  cannot distinguish between  $\phi_0$  and  $\phi_1$  with a strictly positive probability. For any set of samples  $\{(\pi_t, X_t, Y_{\pi_t,t})\}_{t=1}^T$ , define the log-likelihood ratio  $L_T = L_T \left( \{(\pi_t, X_t, Y_{\pi_t,t})\}_{t=1}^T \right)$  as:

$$\begin{aligned} L_T &:= \sum_{t=1}^T \log \left( \frac{\mathbb{P}_{\pi,\phi_0} \{Y_{\pi_t,t} \mid \pi_t, X_t\}}{\mathbb{P}_{\pi,\phi_1} \{Y_{\pi_t,t} \mid \pi_t, X_t\}} \right) \\ &\leq \sum_{t=1}^T \mathbb{1} \{ \pi_t = 1, X_t \in H \} \cdot \left[ Y_{\pi_t,t} \log \left( \frac{\phi_0(X_t)}{\phi_1(X_t)} \right) + (1 - Y_{\pi_t,t}) \log \left( \frac{(1 - \phi_0(X_t))}{(1 - \phi_1(X_t))} \right) \right] \\ &\leq \sum_{t=1}^T \mathbb{1} \{ \pi_t = 1, X_t \in H \} \cdot \left[ Y_{\pi_t,t} \frac{(\phi_0(X_t) - \phi_1(X_t))}{\phi_m(X_t)} + (1 - Y_{\pi_t,t}) \frac{(\phi_1(X_t) - \phi_0(X_t))}{(1 - \phi_1(X_t))} \right] \\ &= \sum_{t=1}^T \mathbb{1} \{ \pi_t = 1, X_t \in H \} \cdot \frac{(Y_{\pi_t,t} - \phi_1(X_t))(\phi_0(X_t) - \phi_1(X_t))}{\phi_1(X_t)(1 - \phi_1(X_t))}, \end{aligned}$$

where the last inequality follows from  $\log(1+x) \leq x$  for all  $x > 0$ . Taking expectation conditional on the event  $\bar{\mathcal{A}}$ , one obtains:

$$\begin{aligned} \mathbb{E}_{\pi,\phi_0} [L_T \mid \bar{\mathcal{A}}] &\leq \mathbb{E}_{\pi,\phi_0} \left[ \sum_{t=1}^T \mathbb{1} \{ \pi_t = 1, X_t \in \tilde{H} \} \cdot \frac{(Y_{\pi_t,t} - \phi_1(X_t))(\phi_0(X_t) - \phi_1(X_t))}{\phi_1(X_t)(1 - \phi_1(X_t))} \mid \bar{\mathcal{A}} \right] \\ &\leq \mathbb{E}_{\pi,\phi_0} \left[ \sum_{t=1}^T \mathbb{E} \left[ \mathbb{1} \{ \pi_t = 1, X_t \in \tilde{H}_1 \} \cdot \frac{(Y_{\pi_t,t} - \phi_1(X_t))(\phi_0(X_t) - \phi_1(X_t))}{\phi_1(X_t)(1 - \phi_1(X_t))} \mid X_t \right] \mid \bar{\mathcal{A}} \right] \\ &= \mathbb{E}_{\pi,\phi_0} \left[ \sum_{t=1}^T \mathbb{1} \{ \pi_t = 1, X_t \in \tilde{H} \} \cdot \frac{(\phi_0(X_t) - \phi_1(X_t))^2}{\phi_1(X_t)(1 - \phi_1(X_t))} \mid \bar{\mathcal{A}} \right] \\ &\stackrel{(a)}{\leq} \frac{64C_\phi^2\Delta^2}{3} \mathbb{E}_{\pi,\phi_0} \left[ \sum_{t=1}^T \mathbb{1} \{ \pi_t = 1, X_t \in \tilde{H} \} \mid \bar{\mathcal{A}} \right] \\ &\leq \frac{64C_\phi^2\Delta^2\mathcal{S}_{\gamma,\alpha}^*(T)}{3} + c_2 T^{-2} \stackrel{(b)}{\leq} 1, \end{aligned} \quad (1.10)$$

for large enough  $T$ , where (a) follows from  $C_\phi\Delta \leq \frac{1}{4}$ , and (b) follows from the definition of  $\Delta$ .

**Step 7 (Lower bound on regret as a function of inferior sampling rate ).** Let

$\tilde{N}_T := \sum_{t=1}^T \mathbb{1} \{ \pi_t = 1, X_t \in \tilde{H} \}$  be the number of times policy  $\pi$  pulls arm 1 when covariates fall into the hypercube  $\tilde{H}$ . We will show that with a strictly positive probability one has  $\tilde{N}_T < \frac{\rho T \Delta}{15}$  conditional on the event  $\bar{\mathcal{A}}$ . This will imply that policy  $\pi$  makes at least  $\frac{\rho T \Delta}{15}$  number of mistakes under the problem  $m = 1$ , where each mistake is associated with at least  $\frac{\Delta}{2}$  instantaneous regret .

Denote by  $\rho_0$  and  $\rho_1$  the distribution of  $\tilde{N}_T$  under the problems  $m = 0$  and  $m = 1$  conditional on the event  $\bar{\mathcal{A}}$ . Define the test function  $\Psi(x) = \mathbb{1}\left\{x \geq \frac{\rho T \Delta}{15}\right\}$ . With this choice of  $\rho_0$ ,  $\rho_1$ , and  $\Psi$ , one can apply Lemma 1.5 to obtain

$$\mathbb{P}_{\pi, \phi_0} \left\{ \tilde{N}_T \geq \frac{\rho T \Delta}{15} \mid \bar{\mathcal{A}} \right\} + \mathbb{P}_{\pi, \phi_1} \left\{ \tilde{N}_T < \frac{\rho T \Delta}{15} \mid \bar{\mathcal{A}} \right\} \geq \frac{1}{2} \exp(-\text{KL}(\rho_0, \rho_1)).$$

In order to lower bound the right hand side of this inequality, we note that

$$\begin{aligned} \mathbb{E}_{\pi, \phi_0} [L_T \mid \bar{\mathcal{A}}] &= \sum_{s=1}^T \mathbb{E}_{\pi, \phi_0} [L_T \mid \bar{\mathcal{A}}, \tilde{N}_T = s] \mathbb{P}_{\pi, \phi_0} \left\{ \tilde{N}_T = s \mid \bar{\mathcal{A}} \right\} \\ &\geq \sum_{s=1}^T \log \left( \frac{\mathbb{P}_{\pi, \phi_0} \left\{ \tilde{N}_T = s \mid \bar{\mathcal{A}} \right\}}{\mathbb{P}_{\pi, \phi_1} \left\{ \tilde{N}_T = s \mid \bar{\mathcal{A}} \right\}} \right) \mathbb{P}_{\pi, \phi_0} \left\{ \tilde{N}_T = s \mid \bar{\mathcal{A}} \right\} = \text{KL}(\rho_0, \rho_1), \end{aligned}$$

where the inequality follows from Lemma 1.6. The last two displays along with (1.10) yield

$$\mathbb{P}_{\pi, \phi_0} \left\{ \tilde{N}_T \geq \frac{\rho T \Delta}{15} \mid \bar{\mathcal{A}} \right\} + \mathbb{P}_{\pi, \phi_1} \left\{ \tilde{N}_T < \frac{\rho T \Delta}{15} \mid \bar{\mathcal{A}} \right\} \geq \frac{1}{2} \exp(-1).$$

To show that  $\mathbb{P}_{\pi, \phi_0} \left\{ \tilde{N}_T \geq \frac{\rho T \Delta}{15} \mid \bar{\mathcal{A}} \right\}$  is small, we apply Markov's inequality:

$$\begin{aligned} \mathbb{P}_{\pi, \phi_0} \left\{ \tilde{N}_T \geq \frac{\rho T \Delta}{15} \mid \bar{\mathcal{A}} \right\} &\leq \frac{\mathbb{E}_{\pi, \phi_0} [\tilde{N}_T \mid \bar{\mathcal{A}}]}{\frac{\rho T \Delta}{15}} \stackrel{(a)}{\leq} \frac{\mathcal{S}_{\gamma, \alpha}^*(T) + c_4 T^{-2}}{\frac{\rho T \Delta}{15}} \\ &\leq \frac{c_3 [\mathcal{S}_{\gamma, \alpha}^*]^{\frac{1}{2}} + c_2 l^{-d} T^{-2}}{\rho T} \leq c_4 T^{-\frac{1}{2}} \leq \frac{1}{4} \exp(-1) \end{aligned}$$

for large enough  $T$  and some constant  $c_3, c_4 > 0$ , where (a) follows from (1.9). The last two displays yield that for large enough  $T$ , one has

$$\mathbb{P}_{\pi, \phi_1} \left\{ \tilde{N}_T < \frac{\rho T \Delta}{15} \mid \bar{\mathcal{A}} \right\} \geq \frac{1}{4e}.$$

Note that by definition, when the event  $\bar{\mathcal{A}}$  holds, at least  $\frac{2\rho T \Delta}{3}$  number of covariates fall into the hypercube  $\tilde{H}$ , that is,

$$\mathcal{R}_{\beta, \alpha}^\pi(T) \geq \frac{\rho T \Delta}{15} \cdot \frac{\Delta}{2} \cdot \mathbb{P}_{\pi, \phi_1} \left\{ \tilde{N}_T < \frac{\rho T \Delta}{15} \mid \bar{\mathcal{A}} \right\} \mathbb{P} \{ \bar{\mathcal{A}} \} \geq \frac{\rho T \Delta^2}{240e} \geq c_5 T [\mathcal{S}_{\gamma, \alpha}^*(T)]^{-1}, \quad (1.11)$$

for some constant  $c_5 > 0$ . The final result follows by noting that  $\mathcal{S}_{\gamma, \alpha}^\pi(T) \leq \tilde{C}_{sr} T^{\frac{1}{\alpha+1}} [T^{\zeta(\gamma, \alpha, d)}]^{\frac{\alpha}{\alpha+1}}$ . This concludes the proof.  $\blacksquare$

### 1.3 Proof of Theorem 4.3

The following lemma characterizes a general class of self-similar payoff functions for any non-integer smoothness parameter  $\beta \in [\underline{\beta}, \bar{\beta}]$ .

**Lemma 1.7.** *Fix dimension  $d$ , some positive non-integer  $\beta$  and some  $\bar{\beta} \geq \beta$ . Consider some set of payoff functions  $\{f_k\}_k$  such that  $f_k \in \mathcal{H}(\beta)$ ,  $k \in \mathcal{K}$ . Suppose  $f_1(x) = a + bx_1^\beta$  for  $x_1 \in [0, c]$  where  $a, b$  and  $0 \leq c \leq 1$  are some constants. Then, the set of payoff functions  $\{f_k\}_k$  is self-similar as in Definition 4.1 with some finite constants  $l_0 \geq 0$  and  $b > 0$ .*

*Proof.* It suffices to show that for any non-negative integer  $p$ , one has

$$\max_{\mathbf{B} \in \mathcal{B}_l^q} \max_{k \in \mathcal{K}} \sup_{x \in \mathbf{B}} \left| \mathbf{\Gamma}_{q^{-l}}^p f_k(x; \mathbf{B}) - f_k(x) \right| \geq b' q^{-l\beta},$$

for any  $l \geq l_0 = \lceil \log \frac{1}{c} \rceil$  and some  $b' > 0$ . Fix some  $l > l_0$ . Let  $\mathbf{B}_0 := [0, q^{-l}]^d$ . One has

$$\max_{\mathbf{B} \in \mathcal{B}_l^q} \max_{k \in \mathcal{K}} \sup_{x \in \mathbf{B}} \left| \mathbf{\Gamma}_{q^{-l}}^p f_k(x; \mathbf{B}) - f_k(x) \right| \geq \left| \mathbf{\Gamma}_{q^{-l}}^p f_1(0; \mathbf{B}_0) - f_1(0) \right| = b \left| \mathbf{\Gamma}_{q^{-l}}^p g(0; \mathbf{B}_0) \right|, \quad (1.12)$$

where  $g(x) = x^\beta$ . By Part 1 of Lemma 2.1, one has  $\mathbf{\Gamma}_{q^{-l}}^p g(0; \mathbf{B}_0) = e_1^\top B^{-1} W$ , where

$$e_1 = (\mathbb{1}\{s = 0\})_{s \in \{0, 1, \dots, p\}}, \quad B = \left( \frac{1}{s_1 + s_2 + 1} \right)_{s_1, s_2 \in \{0, 1, \dots, p\}}, \quad W = \left( \frac{q^{-l\beta}}{s + \beta + 1} \right)_{s \in \{0, 1, \dots, p\}}.$$

By Cramer's rule for linear matrix equations, one has

$$\mathbf{\Gamma}_{q^{-l}}^p g(0; \mathbf{B}_0) = \frac{\det(B_1)}{\det(B)} q^{-l\beta}, \quad (1.13)$$

where

$$B_1 = \begin{pmatrix} \frac{1}{\beta+1} & \frac{1}{2} & \frac{1}{3} & \cdots & \frac{1}{p+1} \\ \frac{1}{\beta+2} & \frac{1}{3} & \frac{1}{4} & \cdots & \frac{1}{p+2} \\ \frac{1}{\beta+3} & \frac{1}{4} & \frac{1}{5} & \cdots & \frac{1}{p+3} \\ \vdots & \vdots & \vdots & \cdots & \vdots \\ \frac{1}{\beta+p+1} & \frac{1}{p+2} & \frac{1}{p+3} & \cdots & \frac{1}{2p+1} \end{pmatrix}.$$

Note that one can rewrite both matrices  $B$  and  $B_1$  as follows

$$B = \left( \frac{1}{u_i + w_j} \right)_{1 \leq i, j \leq p+1}, \quad u_i = i, w_j = j - 1;$$

$$B_1 = \left( \frac{1}{u'_i + w'_j} \right)_{1 \leq i, j \leq p+1}, \quad u'_i = i, w'_j = \beta \mathbb{1}\{j = 1\} + (j - 1) \mathbb{1}\{j > 1\}.$$

The next theorem shows that the determinants of  $B$  and  $B_1$  are non-zero.

**Theorem 1.8** (Cauchy double alternant determinant). *For any set of indeterminates  $\{u_i\}_{1 \leq i \leq n}$  and  $\{v_j\}_{1 \leq j \leq n}$  such that  $u_i + v_j \neq 0, \forall i, j \in \{1, \dots, n\}$ , one has*

$$\det \left( \frac{1}{u_i + w_j} \right)_{1 \leq i, j \leq n} = \frac{\prod_{1 \leq i < j \leq n} (u_i - u_j)(w_i - w_j)}{\prod_{1 \leq i \neq j \leq n} (u_i + w_j)}.$$

Hence, putting together 1.12 and 1.13 yields that for any integer  $l > l_0$ ,

$$\max_{\mathbf{B} \in \mathcal{B}_l^q} \max_{k \in \mathcal{K}} \sup_{x \in \mathbf{B}} \left| \mathbf{\Gamma}_{q^{-l}}^{\mathbf{P}} f_k(x; \mathbf{B}) - f_k(x) \right| \geq b \frac{\det(B_1)}{\det(B)} q^{-l\beta}.$$

This concludes the proof. ■

Using Lemma 1.7, one can adjust the lower bound arguments in Rigollet and Zeevi (2010) and Hu et al. (2019) in order to establish the same lower bounds for optimal regret when payoff functions are self-similar. We provide here the proof of the second part of the theorem; the proof of the first part is very similar, except for using Theorem 4.1 in Rigollet and Zeevi (2010) instead of Theorem 3 in Hu et al. (2019). First, we define the class of problems of interest.

**Definition 1.9.** *For any  $\beta \geq 0$  and  $\alpha \geq 0$ , we denote by  $\tilde{\mathcal{P}}(\beta, \alpha, d) = \tilde{\mathcal{P}}(\beta, L, \alpha, C_0, \underline{\rho}, \bar{\rho})$  the class of problems  $\mathbf{P} = \left( \mathbf{P}_X, \mathbf{P}_{Y|X}^{(1)}, \mathbf{P}_{Y|X}^{(2)} \right)$  that satisfy Assumption 1 for  $\beta$  and  $L > 0$ , Assumption 3 for  $\alpha$  and some  $C_0 > 0$ , and the following assumption regarding covariate distribution: the covariate density  $p_X$  has a compact support  $\mathcal{X} \subseteq [0, 1]^d$  and  $\underline{\rho} \leq p_X(x) \leq \bar{\rho}$  for some  $\bar{\rho} \geq \underline{\rho} > 0$  and  $x \in \mathcal{X}$ . Furthermore, for any  $\beta \in [\underline{\beta}, \bar{\beta}]$ ,  $\alpha \geq 0$ ,  $l_0 \geq 0$ , and  $b > 0$ , we define by  $\mathcal{P}^{\text{ss}}(\beta, \alpha, d, b, l_0) \subseteq \{\mathbf{P} \in \tilde{\mathcal{P}}(\beta, \alpha, d) : \{f_k\}_{k \in \mathcal{K}} \in \mathcal{F}^{\text{ss}}(\beta, b, l_0)\}$  the corresponding class of problems with self-similar payoffs.*

In their Theorem 3, Hu et al. (2019) construct a problem instance  $\mathbf{P}^* \in \tilde{\mathcal{P}}(\beta, \alpha, d)$  such that  $\mathcal{R}^\pi(\mathbf{P}^*; T) \geq CT^{1 - \frac{\beta(\alpha+1)}{2\beta+d}}$  for some constant  $C > 0$ . Let  $\{f_k^*\}_k$  be the set of payoff functions of  $\mathbf{P}^*$ . Define the set of payoff func-

tions  $\{f_k^{**}\}_k$  such that

$$f_k^{**}(x) := \begin{cases} \frac{1+L_1x_1^\beta}{2} & \text{if } 0 \leq x_1 \leq \frac{1}{8}, \\ \frac{1+L_1u(x_1)x_1^\beta}{2} & \text{if } \frac{1}{8} \leq x_1 \leq \frac{1}{4}, \\ \frac{1}{2} & \text{if } \frac{1}{4} \leq x_1 \leq \frac{1}{2}, \\ L_1f_k^*(g(x)) & \text{if } \frac{1}{2} \leq x_1 \leq 1, \end{cases}$$

where  $L_1 > 0$  is some constant and we define

$$g(x) := \begin{pmatrix} 2x_1 - 1 \\ x_2 \\ x_3 \\ \vdots \\ x_d \end{pmatrix}, \quad u(x_1) := \frac{\int_{x_1}^{\frac{1}{4}} \exp\left(\frac{-1}{|s-\frac{1}{8}||s-\frac{1}{4}|}\right) ds}{\int_{\frac{1}{8}}^{\frac{1}{4}} \exp\left(\frac{-1}{|s-\frac{1}{8}||s-\frac{1}{4}|}\right) ds}.$$

Now, we show that  $f_k^{**} \in \mathcal{H}(\beta)$ ,  $k \in \mathcal{K}$ . Note that  $u(x_1)$  is infinitely differentiable over  $[\frac{1}{8}, \frac{1}{4}]$  and  $x_1^\beta \in \mathcal{H}(\beta)$ . Hence, by the following lemma,  $u(x_1)x_1^\beta \in \mathcal{H}_{[\frac{1}{8}, \frac{1}{4}]}(\beta)$ .

**Lemma 1.10.** *Suppose  $f, g \in \mathcal{H}_{\mathcal{X}}(\beta, L)$  for some  $\mathcal{X} \subseteq [0, 1]$ ,  $\beta > 0$ , and  $L > 0$ , and define the function  $h := f \cdot g$  as the product of  $f$  and  $g$ . Then,  $h \in \mathcal{H}(\beta, L')$  for some  $L' > 0$ .*

Furthermore, any derivative of  $f_k^{**}$  up to degree  $\lfloor \beta \rfloor$  exists for  $x_1 \in \{\frac{1}{8}, \frac{1}{4}, \frac{1}{2}\}$ . Hence,  $f_k^{**} \in \mathcal{H}(\beta)$ . One can also make  $L_1 > 0$  small enough so that  $f_k^{**} \in \mathcal{H}(\beta, L)$ . Finally, by Lemma 1.7, the set of payoff functions  $\{f_k^{**}\}_k$  is self-similar. Now, let  $\mathbf{P}^{**}$  be a problem instance that is the same as  $\mathbf{P}^*$  except for its payoff functions that are  $\{f_k^{**}\}_k$ . One can perform a similar analysis as in the proof of Theorem 3 in Hu et al. (2019) in order to show that  $\mathcal{R}^\pi(\mathbf{P}^{**}; T) \geq CT^{1-\frac{\beta(\alpha+1)}{2\beta+d}}$ . This concludes the proof. ■

## 1.4 Proof of Proposition 5.6

Let  $\tilde{r} := \lfloor 2 \log_q(\frac{\gamma}{2C_3}) + 2l\beta + (\frac{d}{\beta} + 1) \log_q \log T \rfloor$  where the constant  $C_3$  was introduced in Proposition 3.2. We will prove the result by bounding the following probability

$$\begin{aligned} & \mathbb{P} \left\{ \exists r \leq \tilde{r} : \sup_{k \in \mathcal{K}, x \in \mathcal{M}^{(\mathbf{B})}} \left| \hat{f}_k^{(\mathbf{B}, r)}(x; j_1^{(\mathbf{B})}) - \hat{f}_k^{(\mathbf{B}, r)}(x; j_2^{(\mathbf{B})}) \right| \geq \frac{\gamma (\log T)^{\frac{d}{2\beta} + \frac{1}{2}}}{q^{r/2}} \right\} \\ & \leq \sum_{r \in [\tilde{r}]} \sum_{k \in \mathcal{K}} \sum_{x \in \mathcal{M}^{(\mathbf{B})}} \mathbb{P} \left\{ \left| \hat{f}_k^{(\mathbf{B}, r)}(x; j_1^{(\mathbf{B})}) - \hat{f}_k^{(\mathbf{B}, r)}(x; j_2^{(\mathbf{B})}) \right| \geq \frac{\gamma (\log T)^{\frac{d}{2\beta} + \frac{1}{2}}}{q^{r/2}} \right\}. \end{aligned} \quad (1.14)$$

Note that by the triangle inequality,

$$\left| \hat{f}_k^{(\mathbf{B},r)}(x; j_1^{(\mathbf{B})}) - \hat{f}_k^{(\mathbf{B},r)}(x; j_2^{(\mathbf{B})}) \right| \leq \left| f_k(x) - \hat{f}_k^{(\mathbf{B},r)}(x; j_1^{(\mathbf{B})}) \right| + \left| f_k(x) - \hat{f}_k^{(\mathbf{B},r)}(x; j_2^{(\mathbf{B})}) \right|.$$

That is,

$$\begin{aligned} & \mathbb{P} \left\{ \left| \hat{f}_k^{(\mathbf{B},r)}(x; j_1^{(\mathbf{B})}) - \hat{f}_k^{(\mathbf{B},r)}(x; j_2^{(\mathbf{B})}) \right| \geq \frac{\gamma (\log T)^{\frac{d}{2\beta} + \frac{1}{2}}}{q^{r/2}} \right\} \\ & \leq \mathbb{P} \left\{ \left| f_k(x) - \hat{f}_k^{(\mathbf{B},r)}(x; j_1^{(\mathbf{B})}) \right| \geq \frac{\gamma (\log T)^{\frac{d}{2\beta} + \frac{1}{2}}}{q^{1+r/2}} \right\} \\ & + \mathbb{P} \left\{ \left| f_k(x) - \hat{f}_k^{(\mathbf{B},r)}(x; j_2^{(\mathbf{B})}) \right| \geq \frac{\gamma (\log T)^{\frac{d}{2\beta} + \frac{1}{2}}}{q^{1+r/2}} \right\}. \end{aligned} \quad (1.15)$$

Note that since when  $r \leq \tilde{r}$  one has  $\frac{\gamma (\log T)^{\frac{d}{2\beta} + \frac{1}{2}}}{q^{1+r/2}} \geq C_3 q^{-\beta j_1^{\mathbf{B}}} \geq C_3 q^{-\beta j_2^{\mathbf{B}}}$ , one can apply Proposition 3.2 to bound the two terms on the right hand side of above inequality. Namely, one can apply Proposition 3.2 with  $n = q^r$ ,  $\underline{\mu} = \frac{\underline{\rho}}{\rho q^{-d}}$ ,  $\bar{\mu} = \frac{\bar{\rho}}{\rho q^{-d}}$ ,  $\delta = \frac{\gamma (\log T)^{\frac{d}{2\beta} + \frac{1}{2}}}{q^{1+r/2}}$ , and  $h = q^{-j_1^{\mathbf{B}}}$  for the first term and  $h = q^{-j_2^{\mathbf{B}}}$  for the second term to obtain

$$\begin{aligned} & \mathbb{P} \left\{ \left| f_k(x) - \hat{f}_k^{(\mathbf{B},r)}(x; j_1^{(\mathbf{B})}) \right| \geq \frac{\gamma (\log T)^{\frac{d}{2\beta} + \frac{1}{2}}}{q^{1+r/2}} \right\} \leq \tilde{C}_1 T^{-\gamma^2 \tilde{C}_2}, \\ & \mathbb{P} \left\{ \left| f_k(x) - \hat{f}_k^{(\mathbf{B},r)}(x; j_2^{(\mathbf{B})}) \right| \geq \frac{\gamma (\log T)^{\frac{d}{2\beta} + \frac{1}{2}}}{q^{1+r/2}} \right\} \leq \tilde{C}_1 T^{-\gamma^2 \tilde{C}_2}, \end{aligned}$$

where the constants  $\tilde{C}_1, \tilde{C}_2$  depend only on  $L, \underline{\rho}, \bar{\rho}$ , and  $d$ . These two inequalities along with (1.14) and (1.15) imply

$$\begin{aligned} \mathbb{P} \left\{ \exists r \leq \tilde{r} : \sup_{k \in \mathcal{K}, x \in \mathcal{M}^{(\mathbf{B})}} \left| \hat{f}_k^{(\mathbf{B},r)}(x; j_1^{(\mathbf{B})}) - \hat{f}_k^{(\mathbf{B},r)}(x; j_2^{(\mathbf{B})}) \right| \geq \frac{\gamma (\log T)^{\frac{d}{2\beta} + \frac{1}{2}}}{q^{r/2}} \right\} & \leq q^{1-ld} |\mathcal{M}^{(\mathbf{B})}| \tilde{r} \tilde{C}_1 T^{-\gamma^2 \tilde{C}_2} \\ & \leq C_7 q^{-ld} (\log T)^{\frac{d}{\beta}} T^{-\gamma^2 C_8 + C_9}, \end{aligned}$$

where the constants  $C_7, C_8, C_9$  depend only on  $\underline{\beta}, \bar{\beta}, L, \underline{\rho}, \bar{\rho}$ , and  $d$ . The results follows by applying union bound over  $\mathbf{B} \in \mathcal{B}_l$ . This concludes the proof.  $\blacksquare$

## 1.5 Proof of Proposition 5.7

By Assumption 4, there exists at least one bin  $\tilde{\mathbf{B}} \in \mathcal{B}_l$ , an arm  $\tilde{k} \in \mathcal{K}$ , and a point  $\hat{x} \in \tilde{\mathbf{B}}$  such that

$$\left| \mathbf{\Gamma}_{j_1^{(\tilde{\mathbf{B}})}}^0 f_k(\hat{x}; \tilde{\mathbf{B}}) - f_k(\hat{x}) \right| = \left| \mathbf{\Gamma}_l^{[\underline{\beta}]} f_k(\hat{x}) - f_k(\hat{x}; \tilde{\mathbf{B}}) \right| \geq b q^{-l\beta}. \quad (1.16)$$

Let  $\tilde{x} = \arg \min_{x \in \mathcal{M}(\tilde{\mathbf{B}})} \|x - \hat{x}\|_\infty$  (if there is more than one minimizer we choose the one with the minimum  $L_1$ -norm). Note that  $\|\tilde{x} - \hat{x}\|_\infty \leq q^{-\tilde{l}}$ , which along with the assumption  $f_{\tilde{k}} \in \mathcal{H}(\underline{\beta}, L)$  implies that

$$\left| f_{\tilde{k}}(\tilde{x}) - f_{\tilde{k}}(\hat{x}) \right| \leq L \|\tilde{x} - \hat{x}\|_\infty^\beta \leq L q^{-\tilde{l}\beta} \leq \frac{L}{\log T} q^{-l\beta}. \quad (1.17)$$

In addition, by Lemma 2.1, one has

$$\left| \mathbf{\Gamma}_l^{[\underline{\beta}]} f_{\tilde{k}}(\hat{x}; \tilde{\mathbf{B}}) - \mathbf{\Gamma}_l^{[\underline{\beta}]} f_{\tilde{k}}(\tilde{x}; \tilde{\mathbf{B}}) \right| \leq \kappa_0 q^l \|\hat{x} - \tilde{x}\|_\infty \leq \kappa_0 q^{l-\tilde{l}} \leq \frac{\kappa_0}{\log T} q^{-l\beta}, \quad (1.18)$$

where  $\kappa_0$  was introduced in Lemma 2.1. Let  $\hat{r} := \lfloor 2 \log_q(\frac{4\gamma}{L \wedge \kappa_0}) + 2l\beta + (\frac{d}{\underline{\beta}} + 3) \log_q \log T \rfloor$ . One has

$$\mathbb{P} \left\{ r_{\text{last}}^{(\tilde{\mathbf{B}})} > \hat{r} \right\} \leq \mathbb{P} \left\{ \left| \hat{f}_{\tilde{k}}^{(\tilde{\mathbf{B}}, \hat{r})}(\tilde{x}; j_1^{(\tilde{\mathbf{B}})}) - \hat{f}_{\tilde{k}}^{(\tilde{\mathbf{B}}, \hat{r})}(\tilde{x}; j_2^{(\tilde{\mathbf{B}})}) \right| < \frac{\gamma (\log T)^{\frac{d}{2\underline{\beta}} + \frac{1}{2}}}{q^{\hat{r}/2}} \right\}. \quad (1.19)$$

Note that by the triangle inequality,

$$\left| \hat{f}_{\tilde{k}}^{(\tilde{\mathbf{B}}, \hat{r})}(\tilde{x}; j_1^{(\tilde{\mathbf{B}})}) - \hat{f}_{\tilde{k}}^{(\tilde{\mathbf{B}}, \hat{r})}(\tilde{x}; j_2^{(\tilde{\mathbf{B}})}) \right| \geq \left| f_{\tilde{k}}(\tilde{x}) - \hat{f}_{\tilde{k}}^{(\tilde{\mathbf{B}}, \hat{r})}(\tilde{x}; j_1^{(\tilde{\mathbf{B}})}) \right| - \left| f_{\tilde{k}}(\tilde{x}) - \hat{f}_{\tilde{k}}^{(\tilde{\mathbf{B}}, \hat{r})}(\tilde{x}; j_2^{(\tilde{\mathbf{B}})}) \right|. \quad (1.20)$$

Note that since one has  $\frac{\gamma (\log T)^{\frac{d}{2\underline{\beta}} + \frac{1}{2}}}{q^{1+\hat{r}/2}} \geq C_3 q^{-\beta j_2^{\mathbf{B}}}$ , one can apply Proposition 3.2 to show that second term on the right hand side of above inequality is ‘‘small’’ with high probability. Namely, one can apply Proposition 3.2 with  $n = q^{\hat{r}}$ ,  $\underline{\mu} = \frac{\rho}{\rho q^{-d\hat{r}}}$ ,  $\bar{\mu} = \frac{\bar{\rho}}{\rho q^{-d\hat{r}}}$ ,  $\delta = \frac{\gamma (\log T)^{\frac{d}{2\underline{\beta}} + \frac{1}{2}}}{q^{1+\hat{r}/2}}$ , and  $h = q^{-j_2^{\mathbf{B}}}$  to obtain

$$\mathbb{P} \left\{ \left| f_{\tilde{k}}(\tilde{x}) - \hat{f}_{\tilde{k}}^{(\tilde{\mathbf{B}}, \hat{r})}(\tilde{x}; j_2^{(\tilde{\mathbf{B}})}) \right| \geq \frac{\gamma (\log T)^{\frac{d}{2\underline{\beta}} + \frac{1}{2}}}{q^{1+\hat{r}/2}} \right\} \leq \tilde{C}_1 T^{-\gamma^2 \tilde{C}_2}, \quad (1.21)$$

where the constants  $\tilde{C}_1, \tilde{C}_2$  depend only on  $\underline{\beta}, L, \underline{\rho}, \bar{\rho}$ , and  $d$ . Now, we show that the first term on the right hand side of (1.19) cannot get ‘‘small’’ with high probability. One can write

$$\left| f_{\tilde{k}}(\tilde{x}) - \hat{f}_{\tilde{k}}^{(\tilde{\mathbf{B}}, \hat{r})}(\tilde{x}; j_1^{(\tilde{\mathbf{B}})}) \right| \geq \left| f_{\tilde{k}}(\tilde{x}) - \mathbf{\Gamma}_{j_1^{(\tilde{\mathbf{B}})}}^{[\underline{\beta}]} f_{\tilde{k}}(\tilde{x}; \tilde{\mathbf{B}}) \right| - \left| \mathbf{\Gamma}_{j_1^{(\tilde{\mathbf{B}})}}^{[\underline{\beta}]} f_{\tilde{k}}(\tilde{x}; \tilde{\mathbf{B}}) - \hat{f}_{\tilde{k}}^{(\tilde{\mathbf{B}}, \hat{r})}(\tilde{x}; j_1^{(\tilde{\mathbf{B}})}) \right|. \quad (1.22)$$

The first term corresponds to bias and the second term corresponds to stochastic error. Note that by (1.16), (1.17), and (1.18), one has

$$\begin{aligned} \left| f_{\tilde{k}}(\tilde{x}) - \mathbf{\Gamma}_{j_1^{(\tilde{\mathbf{B}})}}^{[\underline{\beta}]} f_{\tilde{k}}(\tilde{x}; \tilde{\mathbf{B}}) \right| &\geq \left| f_{\tilde{k}}(\hat{x}) - \mathbf{\Gamma}_{j_1^{(\tilde{\mathbf{B}})}}^{[\underline{\beta}]} f_{\tilde{k}}(\hat{x}; \tilde{\mathbf{B}}) \right| - |f_{\tilde{k}}(\tilde{x}) - f_{\tilde{k}}(\hat{x})| - \left| \mathbf{\Gamma}_{j_1^{(\tilde{\mathbf{B}})}}^{[\underline{\beta}]} f_{\tilde{k}}(\hat{x}; \tilde{\mathbf{B}}) - \mathbf{\Gamma}_{j_1^{(\tilde{\mathbf{B}})}}^{[\underline{\beta}]} f_{\tilde{k}}(\tilde{x}; \tilde{\mathbf{B}}) \right| \\ &\geq bq^{-l\beta} - \frac{L}{\log T} q^{-l\beta} - \frac{\kappa_0}{\log T} q^{-l\beta} \geq \frac{L \wedge \kappa_0}{2 \log T} q^{-l\beta} \geq \frac{2\gamma (\log T)^{\frac{d}{2\underline{\beta}} + \frac{1}{2}}}{q^{\hat{r}/2}} \end{aligned} \quad (1.23)$$

for large enough  $T \geq T_0(L, b, \underline{\rho}, \bar{\rho}, d)$ . In order to bound the second term on the right hand side of (1.22), we apply Proposition 3.3, with  $n = q^{\hat{r}}$ ,  $\delta = \frac{\gamma (\log T)^{\frac{d}{2\underline{\beta}} + \frac{1}{2}}}{q^{1+\hat{r}/2}}$ , and  $h = q^{-j_1^{\tilde{\mathbf{B}}}}$  to obtain

$$\mathbb{P} \left\{ \left| \mathbf{\Gamma}_{j_1^{(\tilde{\mathbf{B}})}}^{[\underline{\beta}]} f_{\tilde{k}}(\tilde{x}; \tilde{\mathbf{B}}) - \hat{f}_{\tilde{k}}^{(\tilde{\mathbf{B}}, \hat{r})}(\tilde{x}; j_1^{(\tilde{\mathbf{B}})}) \right| \geq \frac{\gamma (\log T)^{\frac{d}{2\underline{\beta}} + \frac{1}{2}}}{q^{1+\hat{r}/2}} \right\} \leq \tilde{C}_4 T^{-\gamma^2 \tilde{C}_5}, \quad (1.24)$$

where the constants  $\tilde{C}_4, \tilde{C}_5$  depend only on  $\underline{\beta}, L, \underline{\rho}, \bar{\rho}$ , and  $d$ . Putting together (1.19), (1.20), (1.22), and (1.23), one obtains

$$\begin{aligned} \mathbb{P} \left\{ r_{\text{last}}^{(\tilde{\mathbf{B}})} > \hat{r} \right\} &\leq \mathbb{P} \left\{ \left| \mathbf{\Gamma}_{j_1^{(\tilde{\mathbf{B}})}}^{[\underline{\beta}]} f_k(\tilde{x}; \tilde{\mathbf{B}}) - \hat{f}_k^{(\tilde{\mathbf{B}}, \hat{r})}(\tilde{x}; j_1^{(\tilde{\mathbf{B}})}) \right| \geq \frac{\gamma (\log T)^{\frac{d}{2\underline{\beta}} + \frac{1}{2}}}{q^{1+\hat{r}/2}} \right\} \\ &\quad + \mathbb{P} \left\{ \left| \mathbf{\Gamma}_{j_1^{(\tilde{\mathbf{B}})}}^{[\underline{\beta}]} f_k(\tilde{x}; \tilde{\mathbf{B}}) - \hat{f}_k^{(\tilde{\mathbf{B}}, \hat{r})}(\tilde{x}; j_1^{(\tilde{\mathbf{B}})}) \right| \geq \frac{\gamma (\log T)^{\frac{d}{2\underline{\beta}} + \frac{1}{2}}}{q^{1+\hat{r}/2}} \right\} \\ &\leq C_{10} T^{-\gamma^2 C_{11}}, \end{aligned}$$

where the last inequality follows from (1.21) and (1.24), and the constants  $C_{10}, C_{11}$  depend only on  $\underline{\beta}, L, \underline{\rho}, \bar{\rho}$ , and  $d$ . This concludes the proof.  $\blacksquare$

## 1.6 Proof of Theorem 5.1

Note that for large enough  $T$ , one has

$$\begin{aligned} &\mathbb{P} \left\{ \hat{\beta}_{\text{SACB}} \in \left[ \beta - \frac{3(2\underline{\beta} + d)^2 \log_q \log T}{(\underline{\beta} + d - 1) \log_q T}, \beta \right] \right\} \\ &\leq \mathbb{P} \left\{ 2l\beta + \left( \frac{d}{\underline{\beta}} + 1 \right) \log_q \log T \leq r_{\text{last}}^{(\tilde{\mathbf{B}})} \leq 2l\beta + \left( \frac{d}{\underline{\beta}} + 4 \right) \log_q \log T \right\} \\ &\leq 1 - C_7 q^{ld} (\log T)^{\frac{d}{\underline{\beta}}} T^{-\gamma^2 C_8 + C_9} - C_{10} T^{-\gamma^2 C_{11}}, \end{aligned}$$

where the last inequality follows from Propositions 5.6, and 5.7, and the constants  $C_7, C_8, C_9 > 0$  were introduced in Proposition 5.6, and the constants  $C_{10}, C_{11} > 0$  were introduced in Proposition 5.7.

Next, we show that with high probability,  $T_{\text{SACB}} \leq \frac{4}{\rho} (\log T)^{\frac{2d}{\beta}+4} T^{\frac{(\beta+d-1)}{(2\beta+d)}} =: \bar{T}_{\text{SACB}}$ . Note that the smoothness estimation sub-routine terminates when all the bins  $\mathbf{B} \in \mathcal{B}_l$  have reached round  $\bar{r} = \lceil 2l\bar{\beta} + (\frac{2d}{\beta} + 4) \log_q \log T \rceil$ . That is,  $T_{\text{SACB}}$  is less than the time step by which  $2 \sum_{r=\underline{r}}^{\bar{r}} q^r$  covariates have realized in each  $\mathbf{B} \in \mathcal{B}_l$ . Note that

$$\sum_{r=\underline{r}}^{\bar{r}} q^r \leq q^{\bar{r}+1} \leq 2 (\log T)^{\frac{2d}{\beta}+4} T^{\frac{2\bar{\beta}(\beta+d-1)}{(2\bar{\beta}+d)^2}}.$$

Let  $\bar{N}^{(\mathbf{B})} := \sum_{t=1}^{\bar{T}_{\text{SACB}}} Z_t$  be the number of covariates that have realized in  $\mathbf{B}$  by  $t = \bar{T}_{\text{SACB}}$ , where  $Z_t$ 's are i.i.d. Bernoulli random variables with  $\mathbb{E}[Z_t] \geq \rho T^{-\frac{d(\beta+d-1)}{(2\beta+d)^2}}$  and  $\text{Var}(Z_t) \leq \mathbb{E}[Z_t^2] \leq \bar{\rho} T^{-\frac{d(\beta+d-1)}{(2\beta+d)^2}}$ . Applying Bernstein's inequality in Lemma 1.4 to  $\bar{N}^{(\mathbf{B})}$  with  $a = 2 (\log T)^{\frac{2d}{\beta}+4} T^{\frac{2\bar{\beta}(\beta+d-1)}{(2\bar{\beta}+d)^2}}$  yields:

$$\begin{aligned} \mathbb{P} \left\{ \bar{N}^{(\mathbf{B})} < 2 (\log T)^{\frac{2d}{\beta}+4} T^{\frac{2\bar{\beta}(\beta+d-1)}{(2\bar{\beta}+d)^2}} \right\} &\leq \exp \left( -\frac{a^2}{2\bar{T}_{\text{SACB}} \text{Var}(Z_t) + a} \right) \\ &\leq \exp \left( -\frac{\rho}{4\bar{\rho} + 2\rho} (\log T)^{\frac{2d}{\beta}+4} T^{\frac{2\bar{\beta}(\beta+d-1)}{(2\bar{\beta}+d)^2}} \right), \end{aligned}$$

and, by the union bound:

$$\begin{aligned} \mathbb{P} \{ T_{\text{SACB}} > \bar{T}_{\text{SACB}} \} &\leq \sum_{\mathbf{B} \in \mathcal{B}_l} \mathbb{P} \left\{ \bar{N}^{(\mathbf{B})} < 2 (\log T)^{\frac{2d}{\beta}+4} T^{\frac{2\bar{\beta}(\beta+d-1)}{(2\bar{\beta}+d)^2}} \right\} \\ &\leq 2T^{\frac{d(\beta+d-1)}{(2\beta+d)^2}} \exp \left( -\frac{\rho}{4\bar{\rho} + 2\rho} (\log T)^{\frac{2d}{\beta}+4} T^{\frac{2\bar{\beta}(\beta+d-1)}{(2\bar{\beta}+d)^2}} \right). \end{aligned}$$

This concludes the proof. ■

## 1.7 Proof of Theorem 7.2

The regret incurred by the SACB policy up to  $t = \lfloor \bar{T}_{\text{SACB}} \rfloor$  is bounded by

$$\begin{aligned}
\mathbb{E}^\pi \left[ \sum_{t=1}^{\lfloor \bar{T}_{\text{SACB}} \rfloor} f_{\pi_t^*}(X_t) - f_{\pi_t}(X_t) \right] &\leq T \cdot \mathbb{P} \{ T_{\text{SACB}} > \bar{T}_{\text{SACB}} \} + \bar{T}_{\text{SACB}} \\
&\stackrel{(a)}{\leq} 2T^{1+\frac{d(\underline{\beta}+d-1)}{(2\underline{\beta}+d)^2}} \exp \left( -\frac{\underline{\rho}}{4\bar{\rho}+2\underline{\rho}} (\log T)^{\frac{2d}{\underline{\beta}}+4} T^{\frac{2\underline{\beta}(\underline{\beta}+d-1)}{(2\underline{\beta}+d)^2}} \right) \\
&\quad + \frac{4}{\underline{\rho}} (\log T)^{\frac{2d}{\underline{\beta}}+4} T^{\frac{(\underline{\beta}+d-1)}{(2\underline{\beta}+d)}} \\
&\stackrel{(b)}{=} o \left( T^{1-\frac{\beta(\alpha+1)}{2\beta+d}} \right), \tag{1.25}
\end{aligned}$$

where (a) follows from Theorem 5.1 and (b) holds by  $\frac{(\beta+d-1)}{(2\bar{\beta}+d)} \leq 1 - \frac{\beta(\alpha+1)}{2\beta+d}$  for any  $\underline{\beta} \leq \beta \leq \bar{\beta}$  and  $\alpha \leq \frac{1}{\min\{1,\beta\}}$ . Define  $\hat{\beta}_T := \beta - \frac{3(2\bar{\beta}+d)^2 \log_q \log T}{(\underline{\beta}+d-1) \log_q T}$ . The regret from  $t = \lfloor \bar{T}_{\text{SACB}} \rfloor + 1$  to  $t = T$  is bounded by

$$\begin{aligned}
\mathbb{E}^\pi \left[ \sum_{t=\lfloor \bar{T}_{\text{SACB}} \rfloor + 1}^T f_{\pi_t^*}(X_t) - f_{\pi_t}(X_t) \right] &\leq T \cdot \mathbb{P} \left\{ \hat{\beta}_{\text{SACB}} \notin [\hat{\beta}_T, \beta] \right\} + \bar{C}_0 (\log T)^{\iota_0(\hat{\beta}_T, \alpha, d)} T^{1-\frac{\hat{\beta}_T(\alpha+1)}{2\hat{\beta}_T+d}} \\
&\leq C_4 (\log T)^{\frac{d}{\underline{\beta}}} T^{-\gamma^2 C_5 + C_6 + 1 + \frac{d(\beta+d-1)}{(2\bar{\beta}+d)^2}} \\
&\quad + CT^{1-\frac{\beta(\alpha+1)}{2\beta+d}} (\log T)^{\frac{3d(\alpha+1)(2\bar{\beta}+d)^2}{(2\bar{\beta}+d)(\beta+d)(\underline{\beta}+d-1)} + \iota_0(\hat{\beta}_T, \alpha, d)}, \tag{1.26}
\end{aligned}$$

for some constant  $C > 0$ , where the last inequality follows from Corollary 5.1 and the constants  $C_4, C_5$ , and  $C_6$  were introduced in Theorem 5.1. Putting together (1.25) and (1.26) concludes the proof.  $\blacksquare$

## 1.8 Proof of Corollary 5.3

The result follows from Theorem 7.2 and the fact that for any  $\beta_0 \leq 1$ :

$$\sup_{\mathbf{P} \in \mathcal{P}(\beta_0, \alpha, d)} \mathcal{R}^{\text{ABSE}(\beta_0)}(\mathbf{P}; T) = \mathcal{O} \left( T^{\zeta(\beta_0, \alpha, d)} \right).$$

## 1.9 Proof of Corollary 5.4

The result follows from Theorem 7.2 and since for any problem instance  $\mathbf{P} \in \mathcal{P}(\beta_0, \alpha, d)$ , and decision regions which satisfy the regularity condition in Assumption 5, one has for any  $\beta_0 \geq 1$ :

$$\mathcal{R}^{\text{SmoothBandit}(\beta_0)}(\mathbf{P}; T) = \mathcal{O} \left( (\log T)^{\frac{2\beta_0+d}{2\beta_0}} T^{\zeta(\beta_0, \alpha, d)} \right).$$

## 1.10 Proof of Remark 2 ■

Note that

$$\pi_0(\beta_0) = \begin{cases} \text{ABSE}(\beta_0) & \text{if } \beta_0 \leq 1; \\ \text{SmoothBandit}(\beta_0) & \text{if } \beta_0 > 1. \end{cases}$$

Furthermore, for any  $\beta_0 \leq 1$

$$\sup_{\mathbf{P} \in \mathcal{P}(\beta_0, \alpha, d)} \mathcal{R}^{\text{ABSE}(\beta_0)}(\mathbf{P}; T) = \mathcal{O}\left(T^{\zeta(\beta_0, \alpha, d)}\right),$$

and for any  $\beta_0 > 1$  and any problem instance  $\mathbf{P} \in \mathcal{P}(\beta_0, \alpha, d)$ , and decision regions which satisfy the regularity condition in Assumption 5,

$$\mathcal{R}^{\text{SmoothBandit}(\beta_0)}(\mathbf{P}; T) = \mathcal{O}\left((\log T)^{\frac{2\beta_0+d}{2\beta_0}} T^{\zeta(\beta_0, \alpha, d)}\right).$$

The result follows from applying Theorem 7.2 with

$$\iota_0(\beta_0, \alpha, d) := \begin{cases} 0 & \text{if } \beta_0 \leq 1; \\ \frac{2\beta_0}{2\beta_0+d} & \text{o.w.} \end{cases}$$

## 2 Properties of the $L_2(P_X)$ -projection ■

**Lemma 2.1.** *Fix non-negative integers  $l$  and  $p$ , a hypercube  $U$  of side-length  $q^{-l}$ ,  $l' \in \mathbb{R}_+$ , and some point  $x \in U$  and let  $K(\cdot) = \mathbb{1}\{\|\cdot\|_\infty \leq 1\}$  and  $h = q^{-l}$ . Let  $\mu_0$ ,  $\kappa_0$ , and  $L_0$  be some constants that only depend on  $p, \rho, \bar{\rho}$  (introduced in Assumption 2), and  $d$ . The following statements hold:*

1.  $\mathbf{\Gamma}_h^p f(x; U) = R^\top(0)B^{-1}W$ , where we define the vector  $R(u) := (u^s)_{|s| \leq p}$ , the matrix  $B := (B_{s_1, s_2})_{|s_1|, |s_2| \leq p}$ , and the vector  $W := (W_s)_{|s| \leq p}$  with elements

$$B_{s_1, s_2} := \int_{\mathbb{R}^d} u^{s_1+s_2} K(u) p_X(x+hu | U) du, \quad W_s := \int_{\mathbb{R}^d} u^s f(x+hu) K(u) p_X(x+hu | U) du;$$

2.  $\lambda_{\min}(B) \geq \mu_0 q^{dl'}$ ;
3.  $|\mathbf{\Gamma}_h^p f(x; U) - \mathbf{\Gamma}_h^p f(\hat{x}; U)| \leq \kappa_0 h^{-1} \|\hat{x} - x\|_\infty$  for all  $x, \hat{x} \in U$ ;
4. If  $f \in \mathcal{H}(\beta, L)$  for  $0 < \beta \leq p+1$  then,  $|\mathbf{\Gamma}_h^p f(x; U) - f(x)| \leq L_0 h^\beta$  for all  $x \in U$ .

*Proof.* Fix some  $x \in U$ . Let  $\tilde{\theta}(u; p, l, U) := \sum_{|s| \leq p} \xi_s u^s$  be a polynomial of degree  $p$  on  $\mathbb{R}^d$  that minimizes

$$\begin{aligned} & \int_U \left| f(u) - \tilde{\theta} \left( \frac{u-x}{h}; p, l, U \right) \right|^2 K \left( \frac{u-x}{h} \right) p_X(u | U) du = \int_U f^2(u) K \left( \frac{u-x}{h} \right) p_X(u | U) du \\ & + \sum_{|s_1|, |s_2| \leq p} \xi_{s_1} \xi_{s_2} \int_U \left( \frac{u-x}{h} \right)^{s_1+s_2} K \left( \frac{u-x}{h} \right) p_X(u | U) du \\ & - 2 \sum_{|s| \leq p} \xi_s \int_U f(u) \left( \frac{u-x}{h} \right)^s K \left( \frac{u-x}{h} \right) p_X(u | U) du, \end{aligned}$$

where  $h = q^{-l}$ . Equivalently,  $\tilde{\theta}(u; p, l, U)$  can be characterized by its vector of coefficients  $\boldsymbol{\xi}$  that minimizes

$$\sum_{|s_1|, |s_2| \leq p} \xi_{s_1} \xi_{s_2} \int_{\mathbb{R}^d} u^{s_1+s_2} K(u) p_X(x+hu | U) du - 2 \sum_{|s| \leq p} \xi_s \int_{\mathbb{R}^d} f(u) u^s K(u) p_X(x+hu | U) du = \boldsymbol{\xi}^\top B \boldsymbol{\xi} - 2W^\top \boldsymbol{\xi}, \quad (2.1)$$

where we define the matrix  $B := (B_{s_1, s_2})_{|s_1|, |s_2| \leq p}$  and the vector  $W := (W_s)_{|s| \leq p}$  with elements

$$B_{s_1, s_2} := \int_{\mathbb{R}^d} u^{s_1+s_2} K(u) p_X(x+hu | U) du, \quad W_s := \int_{\mathbb{R}^d} f(u) u^s K(u) p_X(x+hu | U) du.$$

Note that if  $B$  is a positive definite matrix then, the minimizer of (2.1) is  $\boldsymbol{\xi} = B^{-1}W$ , which implies the desired result:  $\mathbf{\Gamma}_h^p f(x; U) = R^\top(0)B^{-1}W$ . In order to show that this is indeed the case, we note that

$$\lambda_{\min}(B) = \min_{\|Z\|=1} Z^\top B Z = \min_{\|Z\|=1} \int_{\mathbb{R}^d} \left( \sum_{|s| \leq p} Z_s u^s \right)^2 K(u) p_X(x+hu | U) du \geq \min_{\|Z\|=1} \frac{\rho 2^{dl'}}{\bar{\rho}} \int_A \left( \sum_{|s| \leq p} Z_s u^s \right)^2 du,$$

where  $A = \{u \in \mathbb{R}^d : \|u\|_\infty \leq 1; x+hu \in U\}$ . Note that

$$\lambda[A] \geq h^{-d} \lambda[\Xi(x, h) \cap U] \geq q^{-d} h^{-d} \lambda[\Xi(x, h)] = q^{-d} \lambda[\Xi(0, 1)].$$

Let  $\mathcal{A}$  denote the class of compact subsets of  $\Xi(0, 1)$  having the Lebesgue measure  $q^{-d} \lambda[\Xi(0, 1)]$ . Using the previous display, we obtain

$$\lambda_{\min}(B) \geq \frac{\rho q^{dl'}}{\bar{\rho}} \min_{\|Z\| \leq 1; S \in \mathcal{A}} \int_S \left( \sum_{|s| \leq p} Z_s u^s \right)^2 du =: \frac{\rho q^{dl'}}{\bar{\rho}} \tilde{\mu}_0. \quad (2.2)$$

By the compactness argument, the minimum in the above expression exists, and is strictly positive.

In order to prove the last claim in the lemma, note that for any  $\hat{x} \in U$ ,

$$\begin{aligned} |\mathbf{\Gamma}_h^p f(x; U) - \mathbf{\Gamma}_h^p f(\hat{x}; U)| &= \left| \tilde{\theta}(0; p, l, U) - \tilde{\theta}\left(\frac{\hat{x} - x}{h}; p, l, U\right) \right| \\ &= \left| \sum_{|s| \leq p, s \neq (0, \dots, 0)} \xi_s \left(\frac{\hat{x} - x}{h}\right)^s \right| \leq M h^{-1} \|\hat{x} - x\|_\infty \|\boldsymbol{\xi}\|. \end{aligned}$$

Also, by (2.2), one has

$$\|\boldsymbol{\xi}\| \leq \|B^{-1}W\| \leq \frac{q^{-dl'} \bar{\rho}}{\underline{\rho}} \tilde{\mu}_0^{-1} M^{\frac{1}{2}} \max_s |W_s|,$$

and

$$|W_s| = \left| \int_{\mathbb{R}^d} u^s f(x + hu) K(u) p_X(x + hu | U) du \right| \leq \int_{\mathbb{R}^d} K(u) p_X(x + hu | U) du \leq q^{dl'}.$$

Putting together the above three displays, one obtains

$$|\mathbf{\Gamma}_h^p f(x; U) - \mathbf{\Gamma}_h^p f(\hat{x}; U)| \leq \bar{\rho} \underline{\rho}^{-1} \tilde{\mu}_0^{-1} M^{3/2} h^{-1} \|\hat{x} - x\|_\infty.$$

To prove the last part, define the vector  $Z := (Z_s)_{|s| \leq p}$  with elements

$$Z_s := \frac{h^{|s|} f^{(s)}(x)}{s!} \cdot \mathbb{1}\{|s| \leq \lfloor \beta \rfloor\}.$$

Note that

$$f(x) = R^\top(0) B^{-1} BZ.$$

As a result, one has

$$|f(x) - \mathbf{\Gamma}_h^p f(x; U)| = \left| R^\top(0) B^{-1} (BZ - W) \right| \leq \|B^{-1}\| \cdot \|BZ - W\| \leq \frac{q^{-dl'} \bar{\rho}}{\underline{\rho} \tilde{\mu}_0} M^{\frac{1}{2}} \max_s |(BZ)_s - W_s|,$$

where the last inequality follows from (2.2). Furthermore, one has

$$\begin{aligned} |(BZ)_s - W_s| &= \left| \int_{\mathbb{R}^d} u^s \left( \sum_{|s'| \leq \lfloor \beta \rfloor} \frac{(hu)^{s'} f^{(s')}(x)}{s'!} - f(x + hu) \right) K(u) p_X(x + hu | U) du \right| \\ &\leq \int_{\mathbb{R}^d} |u^s| \left| \sum_{|s'| \leq \lfloor \beta \rfloor} \frac{(hu)^{s'} f^{(s')}(x)}{s'!} - f(x + hu) \right| K(u) p_X(x + hu | U) du \\ &\leq \int_{\mathbb{R}^d} L h^\beta p_X(x + hu | U) du = L h^\beta q^{dl'} \end{aligned}$$

where the last inequality follows from the assumption that  $f \in \mathcal{H}(\beta, L)$ . Putting the last two displays

together, the result follows. This concludes the proof. ■

### 3 Proofs and analysis for the review of local polynomial regression

In this section of the appendix, we provide the proofs for our review of the local polynomial regression estimation method. Fix a set of pairs  $\mathcal{D} = \{(X_i, Y_i)\}_{i=1}^n$ , a point  $x \in \mathbb{R}^d$ , a bandwidth  $h > 0$ , an integer  $p > 0$  and a kernel function  $K : \mathbb{R}^d \rightarrow \mathbb{R}_+$ . Define the matrix  $Q := (Q_{s_1, s_2})_{|s_1|, |s_2| \leq p}$  and the vector  $V := (V_s)_{|s| \leq p}$  with the elements

$$Q_{s_1, s_2} := \sum_{i=1}^n (X_i - x)^{s_1 + s_2} K\left(\frac{X_i - x}{h}\right), \quad V_s := \sum_{i=1}^n Y_i (X_i - x)^s K\left(\frac{X_i - x}{h}\right).$$

Also, define the matrix  $U := (u^s)_{|s| \leq p}$ . The next result from Audibert and Tsybakov (2007) provides a closed-form expression for local polynomial regression at any arbitrary point.

**Lemma 3.1** (Audibert and Tsybakov 2007, Proposition 2.1). *If the matrix  $Q$  is positive definite, there exists a polynomial on  $\mathbb{R}^d$  of degree  $p$  minimizing (5.3). Its vector of coefficients is given by  $\xi = Q^{-1}V$  and the corresponding local polynomial regression function at point  $x$  is given by*

$$\hat{\eta}^{\text{LP}}(x; \mathcal{D}, h, p) = U(0)^\top Q^{-1}V = \sum_{i=1}^n Y_i K\left(\frac{X_i - x}{h}\right) U(0)^\top Q^{-1}U(X_i - x).$$

The following simple extension of Theorem 3.2 in Audibert and Tsybakov (2007) will be one of the main tools to bound our estimation error in our proposed policy.

**Proposition 3.2.** *Let  $\mathcal{D} = \{(X_i, Y_i)\}_{i=1}^n$  be a set of  $n$  i.i.d pairs  $(X_i, Y_i) \in \mathcal{X} \times \mathbb{R}$ . If the marginal density  $\mu$  of  $X_i$ 's satisfies  $\underline{\mu} \leq \mu(x) \leq \bar{\mu}$  for some  $0 < \underline{\mu} \leq \bar{\mu}$  with a support  $\mathcal{X}$  that is a closed hypercube in  $\mathbb{R}^d$  of side-length  $q^{-l}$ ,  $l \geq 0$ , and the function  $\eta$  belongs to the Hölder class of functions  $\mathcal{H}_{\mathcal{X}}(\beta, L)$  for some  $\beta, L > 0$  then, there exist constants  $C_1, C_2, C_3 > 0$  such that for any  $0 < h < q^{-l}$ , any  $C_3 h^\beta < \delta$ , any  $n \geq 1$  and the kernel function  $K(\cdot) = \mathbb{1}\{\|\cdot\|_\infty \leq 1\}$ , the local polynomial estimator  $\hat{\eta}^{\text{LP}}(x; \mathcal{D}, h, p)$  satisfies*

$$|\hat{\eta}^{\text{LP}}(x; \mathcal{D}, h, p) - \eta(x)| \leq \delta$$

*with probability at least  $1 - C_1 \exp(-C_2 n h^d \underline{\mu}^2 \bar{\mu}^{-1} \delta^2)$  for all  $x \in \mathcal{X}$ . The constants  $C_1, C_2, C_3$  depend only on  $p, d, L$ .*

The next proposition states that local polynomial regression estimation of a function inside a hypercube cannot largely deviate from the  $L_2(P_X)$ -projection of that function with high probability.

**Proposition 3.3.** Fix a hypercube  $U \subseteq (0, 1)^d$  with side-length  $q^{-l'}$ ,  $l' \in \mathbb{R}_+$ . Let  $\mathcal{D} = \{(X_i, Y_i)\}_{i=1}^n$  be a set of  $n$  i.i.d pairs  $(X_i, Y_i) \in U \times \mathbb{R}$ . If the marginal density  $\mu$  of  $X_i$ 's satisfies  $\mu(\cdot) = p_X(\cdot|U)$ , where  $p_X$  is the density of a distribution  $\mathbf{P}_X$  that satisfies Assumption 2 then, there exist constants  $C_4, C_5, C_6 > 0$  such that for any  $\delta < C_6$ , any  $n \geq 1$ ,  $h = q^{-l}$ ,  $l \geq l'$ , and the kernel function  $K(\cdot) = \mathbb{1}\{\|\cdot\|_\infty \leq 1\}$ , the local polynomial estimator  $\hat{\eta}^{\text{LP}}(x; \mathcal{D}, h, p)$  satisfies

$$\left| \hat{\eta}^{\text{LP}}(x; \mathcal{D}, h, p) - \mathbf{\Gamma}_{q^{-l}}^p \eta(x; U) \right| \leq \delta$$

with probability at least  $1 - C_4 \exp(-C_5 n q^{d(l'-l)} \delta^2)$  for all  $x \in U$ . The constants  $C_4, C_5, C_6$  depend only on  $p, \bar{\rho}, \rho$ , and  $d$ .

### 3.1 Proof of Proposition 3.2

The proof is a simple extension of the proof of Theorem 3.2 in Audibert and Tsybakov (2007); however, we provide the proof for completeness. Fix  $x \in \mathcal{X}$  and  $\delta > 0$ . Consider the matrices  $B := (B_{s_1, s_2})_{|s_1|, |s_2| \leq p}$  and  $\bar{B} := (\bar{B}_{s_1, s_2})_{|s_1|, |s_2| \leq p}$  with the elements

$$B_{s_1, s_2} := \int_{\mathbb{R}^d} u^{s_1 + s_2} K(u) \mu(x + hu) du, \quad \bar{B}_{s_1, s_2} := \frac{1}{nh^d} \sum_{i=1}^n \left( \frac{X_i - x}{h} \right)^{s_1 + s_2} K\left( \frac{X_i - x}{h} \right).$$

The smallest eigenvalue of  $\bar{B}$  satisfies

$$\begin{aligned} \lambda_{\min}(\bar{B}) &= \min_{\|W\|=1} W^\top \bar{B} W \\ &\geq \min_{\|W\|=1} W^\top B W + \min_{\|W\|=1} W^\top (\bar{B} - B) W \\ &\geq \min_{\|W\|=1} W^\top B W - \sum_{|s_1|, |s_2| \leq p} |\bar{B}_{s_1, s_2} - B_{s_1, s_2}|. \end{aligned} \quad (3.1)$$

Define  $\mathcal{X}_n := \{u \in \mathbb{R}^d : \|u\| \leq 1; x + hu \in \mathcal{X}\}$ . For any vector  $W$  satisfying  $\|W\|=1$ , we obtain

$$W^\top B W = \int_{\mathbb{R}^d} \left( \sum_{|s| \leq p} W_s u_s \right)^2 K(u) \mu(x + hu) du \geq \underline{\mu} \int_{\mathcal{X}_n} \left( \sum_{|s| \leq p} W_s u_s \right)^2 du.$$

Since  $\mathcal{X}$  is a closed hypercube and we have assumed that  $h \leq l$ , we get

$$\lambda[\mathcal{X}_n] \geq h^{-d} \lambda[\text{Ball}_2(x, h) \cap \mathcal{X}] \geq q^{-d} h^{-d} \lambda[\text{Ball}_2(x, h)] \geq q^{-d} \lambda[\text{Ball}_2(0, 1)],$$

where  $\text{Ball}_2(x, h)$  is the Euclidean ball of radius  $h$  centered around  $x$ .

Let  $\mathcal{A}$  denote the class of all compact subsets of  $\text{Ball}_2(0, 1)$  having the Lebesgue measure  $q^{-d}\lambda[\text{Ball}_2(0, 1)]$ .

Using the previous display, we obtain

$$\min_{\|W\|=1} W^\top BW \geq \underline{\mu} \min_{\|W\|=1; S \in \mathcal{A}} \int_S \left( \sum_{|s| \leq p} W_s u_s \right)^2 du =: 2c\underline{\mu} \quad (3.2)$$

By the compactness argument, the above minimum exists and is strictly positive.

For  $i = 1, \dots, n$  and any multi-indices  $s_1, s_2$  such that  $|s_1|, |s_2| \leq p$ , define

$$T_i^{(s_1, s_2)} := \frac{1}{h^d} \left( \frac{X_i - x}{h} \right)^{s_1 + s_2} K \left( \frac{X_i - x}{h} \right) - \int_{\mathbb{R}^d} u^{s_1 + s_2} K(u) \mu(x + hu) du.$$

We have  $\mathbb{E}T_i^{(s_1, s_2)} = 0$ ,  $|T_i^{(s_1, s_2)}| \leq 2h^{-d}$ , and the following bound on the variance of  $T_i^{(s_1, s_2)}$ :

$$\begin{aligned} \text{Var}T_i^{(s_1, s_2)} &\leq \frac{1}{h^{2d}} \mathbb{E} \left[ \left( \frac{X_i - x}{h} \right)^{2s_1 + 2s_2} K^2 \left( \frac{X_i - x}{h} \right) \right] \\ &\leq \frac{1}{h^d} \int_{\mathbb{R}^d} u^{2s_1 + 2s_2} K^2(u) \mu(x + hu | \mathbf{B}) du \\ &\leq \frac{\bar{\mu}}{h^d} \max_{j \leq p} \int_{\mathbb{R}^d} (1 + |u^{4j}|) K^2(u) du =: \frac{\kappa \bar{\mu}}{h^d}. \end{aligned}$$

From Bernstein's inequality, we get

$$\mathbb{P} \left\{ |\bar{B}_{s_1, s_2} - B_{s_1, s_2}| > \epsilon \right\} = \mathbb{P} \left\{ \left| \frac{1}{n} \sum_{i=1}^n T_i^{(s_1, s_2)} \right| > \epsilon \right\} \leq 2 \exp \left( \frac{-nh^d \epsilon^2}{2\kappa \bar{\mu} + 4\epsilon/3} \right)$$

This inequality along with (3.1) and (3.2) imply that

$$\mathbb{P} \left\{ \lambda_{\min}(\bar{B}) \leq c\underline{\mu} \right\} \leq 2M^2 \exp \left( \frac{-nh^d M^{-4} c^2 \underline{\mu}^2}{2\kappa \bar{\mu} + 4M^{-2} c \underline{\mu} / 3} \right), \quad (3.3)$$

where  $M^2$  is the number of elements in the matrix  $\bar{B}$ . In what follows assume that  $\lambda_{\min}(\bar{B}) \geq c\underline{\mu}$ .

Therefore,

$$\mathbb{P} \left\{ |\hat{\eta}^{\text{LP}}(x; \mathcal{D}, h, p) - \eta(x)| \geq \delta \right\} \leq \mathbb{P} \left\{ \lambda_{\min}(\bar{B}) \leq c\underline{\mu} \right\} + \mathbb{P} \left\{ |\hat{\eta}^{\text{LP}}(x; \mathcal{D}, h, p) - \eta(x)| \geq \delta, \lambda_{\min}(\bar{B}) > c\underline{\mu} \right\}. \quad (3.4)$$

We now evaluate the second term on the right hand side of the above inequality. Define the matrix

$Z := (Z_{i,s})_{1 \leq i \leq n, |s| \leq p}$  with elements

$$Z_{i,s} := (X_i - x)^s \sqrt{K \left( \frac{X_i - x}{h} \right)}.$$

The  $s$ -th column of  $Z$  is denoted by  $Z_s$ , and we introduce  $Z^{(\eta)} := \sum_{|s| \leq \lfloor \beta \rfloor} \frac{\eta^{(s)}(x)}{s!} Z_s$ . Since  $Q = Z^\top Z$  we get

$$\forall |s| \leq \lfloor \beta \rfloor : U^\top(0) Q^{-1} Z^\top Z = \mathbb{1}_{\{s = (0, \dots, 0)\}},$$

hence  $R^\top(0) Q^{-1} Z^\top Z^{(\eta)} = \eta(x)$ . So we can write

$$\hat{\eta}^{\text{LP}}(x; \mathcal{D}, h, p) - \eta(x) = R^\top(0) Q^{-1} (V - Z^\top Z^{(\eta)}) = R^\top(0) \bar{B}^{-1} \mathbf{a},$$

where  $\mathbf{a} := \frac{1}{nh^d} H (V - Z^\top Z^{(\eta)}) \in \mathbb{R}^M$  and  $H$  is a diagonal matrix  $H := (H_{s_1, s_2})_{|s_1|, |s_2| \leq p}$  with elements  $H_{s_1, s_2} := h^{-s_1} \mathbb{1}_{\{s_1 = s_2\}}$ . For  $\lambda_{\min}(\bar{B}) > c\underline{\mu}$ , one has

$$|\hat{\eta}^{\text{LP}}(x; \mathcal{D}, h, p) - \eta(x)| \leq \|\bar{B}^{-1} \mathbf{a}\| \leq \lambda_{\min}^{-1}(\bar{B}) \|\mathbf{a}\| \leq c^{-1} \underline{\mu}^{-1} M \max_s \|a_s\|, \quad (3.5)$$

where  $a_s$  are the components of the vector  $\mathbf{a}$  given by

$$a_s = \frac{1}{nh^d} \sum_{i=1}^n [Y_i - \eta_x(X_i)] \left( \frac{X_i - x}{h} \right)^s K \left( \frac{X_i - x}{h} \right).$$

Note that  $\eta_x(X_i)$  is the Taylor expansion of  $\eta$  at  $x$  and of degree  $\lfloor \beta \rfloor$  (not necessarily  $p$ ) evaluated at  $X_i$ .

Define:

$$\begin{aligned} T_i^{(s,1)} &:= [Y_i - \eta(X_i)] \left( \frac{X_i - x}{h} \right)^s K \left( \frac{X_i - x}{h} \right), \\ T_i^{(s,2)} &:= [\eta(X_i) - \eta_x(X_i)] \left( \frac{X_i - x}{h} \right)^s K \left( \frac{X_i - x}{h} \right). \end{aligned}$$

One has

$$|a_s| \leq \left| \frac{1}{n} \sum_{i=1}^n T_i^{(s,1)} \right| + \left| \frac{1}{n} \sum_{i=1}^n [T_i^{(s,2)} - \mathbb{E}T_i^{(s,2)}] \right| + |\mathbb{E}T_i^{(s,2)}|. \quad (3.6)$$

Note that  $\mathbb{E}T_i^{(s,1)} = 0$ ,  $|T_i^{(s,1)}| \leq 2h^{-d}$ , and

$$\begin{aligned}
\text{Var}T_i^{(s,1)} &\leq \frac{1}{4h^d} \int_{\mathbb{R}^d} u^{2s} K^2(u) \mu(x+hu) du \leq \frac{\kappa \bar{\mu}}{4h^d}, \\
\left| T_i^{(s,2)} - \mathbb{E}T_i^{(s,2)} \right| &\leq Lh^{\beta-d} + L\kappa h^\beta \leq Ch^{\beta-d}, \\
\text{Var}T_i^{(s,2)} &\leq L^2 h^{2\beta-d} \int_{\mathbb{R}^d} |u^{2s}| K^2(u) \mu(x+hu) \leq L^2 \bar{\mu} \kappa h^{2\beta-d}.
\end{aligned}$$

From Bernstein's inequality, for  $\epsilon_1, \epsilon_2 > 0$ , we obtain

$$\mathbb{P} \left\{ \left| \frac{1}{n} \sum_{i=1}^n T_i^{(s,1)} \right| \geq \epsilon_1 \right\} \leq 2 \exp \left( \frac{-nh^d \epsilon_1^2}{\kappa \bar{\mu}/2 + 4\epsilon_1/3} \right)$$

and

$$\mathbb{P} \left\{ \left| \frac{1}{n} \sum_{i=1}^n \left[ T_i^{(s,2)} - \mathbb{E}T_i^{(s,2)} \right] \right| \geq \epsilon_2 \right\} \leq 2 \exp \left( \frac{-nh^d \epsilon_2^2}{2L^2 \kappa \bar{\mu} h^{2\beta} + 2Ch^\beta \epsilon_2/3} \right).$$

Since also

$$\left| \mathbb{E}T_i^{(s,2)} \right| \leq Lh^\beta \int_{\mathbb{R}^d} |u^s| K^2(u) \mu(x+hu) du \leq L\kappa \bar{\mu} h^\beta$$

we get, using (3.6), that if  $3L\kappa \bar{\mu} h^\beta c^{-1} \underline{\mu}^{-1} M \leq \delta \leq 1$  the following inequality holds

$$\begin{aligned}
\mathbb{P} \left\{ |a_s| \geq \frac{c\mu\delta}{M} \right\} &\leq \mathbb{P} \left\{ \left| \frac{1}{n} \sum_{i=1}^n T_i^{(s,1)} \right| > \frac{c\mu\delta}{3M} \right\} + \mathbb{P} \left\{ \left| \frac{1}{n} \sum_{i=1}^n \left[ T_i^{(s,2)} - \mathbb{E}T_i^{(s,2)} \right] \right| > \frac{c\mu\delta}{3M} \right\} \\
&\leq 4 \exp \left( -Cnh^d \underline{\mu}^2 \bar{\mu}^{-1} \delta^2 \right).
\end{aligned}$$

Combining this inequality with (3.3), (3.4), and (3.5), one has

$$\mathbb{P} \left\{ \left| \hat{\eta}^{\text{LP}}(x; \mathcal{D}, h, p) - \eta(x) \right| \geq \delta \right\} \leq C_1 \exp \left( -C_2 nh^d \underline{\mu}^2 \bar{\mu}^{-1} \delta^2 \right)$$

for  $3L\kappa \bar{\mu} h^\beta c^{-1} \underline{\mu}^{-1} M \leq \delta$  (for  $\delta > 1$ , this inequality is obvious since  $\eta, \hat{\eta}^{\text{LP}}$  take values in  $[0, 1]$ ). The constants  $C_1, C_2$  do not depend on the density  $\mu$ , on its support  $\mathcal{X}$  and the point  $x \in \mathcal{X}$ . This concludes the proof.  $\blacksquare$

### 3.2 Proof of Proposition 3.3

Fix a bin  $U \subseteq (0, 1)^d$  with side-length  $q^{-l'}$ ,  $l' \in \mathbb{R}_+$ . Consider the matrix  $B := (B_{s_1, s_2})_{|s_1|, |s_2| \leq p}$  and the vector  $W := (W_s)_{|s| \leq p}$  with elements

$$B_{s_1, s_2} := \int_{\mathbb{R}^d} u^{s_1 + s_2} K(u) \mu(x+hu) du, \quad W_s := \int_{\mathbb{R}^d} u^s \eta(x+hu) K(u) \mu(x+hu) du,$$

as well as the matrix  $\bar{B} := (\bar{B}_{s_1, s_2})_{|s_1|, |s_2| \leq p}$  and the vector  $\bar{W} := (\bar{W}_s)_{|s| \leq p}$  with elements

$$\bar{B}_{s_1, s_2} := \frac{1}{nh^d} \sum_{i=1}^n \left( \frac{X_i - x}{h} \right)^{s_1 + s_2} K \left( \frac{X_i - x}{h} \right), \quad \bar{W}_s := \frac{1}{nh^d} \sum_{i=1}^n Y_i \left( \frac{X_i - x}{h} \right)^s K \left( \frac{X_i - x}{h} \right).$$

By Lemmas 2.1 and 3.1, one has

$$\begin{aligned} \left| \hat{\eta}^{\text{LP}}(x; \mathcal{D}, h, p) - \mathbf{\Gamma}_{q^{-l}}^p \eta(x) \right| &= \left| U(0)^\top Q^{-1} V - U(0)^\top B^{-1} W \right| \\ &= \left| U(0)^\top \bar{B}^{-1} \bar{W} - U(0)^\top B^{-1} W \right| \\ &\leq \left| U(0)^\top B^{-1} (\bar{W} - W) \right| + \left| U(0)^\top (\bar{B}^{-1} - B^{-1}) \bar{W} \right| =: J_1 + J_2. \end{aligned}$$

That is,

$$\mathbb{P} \left\{ \left| \hat{\eta}^{\text{LP}}(x; \mathcal{D}, h, p) - \mathbf{\Gamma}_{q^{-l}}^p \eta(x; U) \right| \geq \delta \right\} \leq \mathbb{P} \{ J_1 \geq 3\delta/4 \} + \mathbb{P} \{ J_2 \geq \delta/4 \}. \quad (3.7)$$

First, we analyze  $J_1$ . Note that

$$J_1 \leq \|B^{-1} (\bar{W} - W)\| \leq \lambda_{\min}^{-1}(B) \|\bar{W} - W\| \leq \mu_0^{-1} q^{-dl'} \|\bar{W} - W\| \leq \mu_0^{-1} q^{-dl'} M \max_s |\bar{W}_s - W_s|, \quad (3.8)$$

where the third inequality follows from  $\lambda_{\min}(B) \geq \mu_0 q^{dl'}$  by Lemma 2.1, and  $M$  is the number of elements in the vector  $W$ . Define:

$$T_i^{(s)} := \frac{1}{h^d} Y_i \left( \frac{X_i - x}{h} \right)^s K \left( \frac{X_i - x}{h} \right) - \int_{\mathbb{R}^d} \eta(x + hu) u^s K(u) p_X(x + hu | U) du.$$

We have  $\mathbb{E} [T_i^{(s)}] = 0$ ,  $|T_i^{(s)}| \leq 2h^{-d}$ , and

$$\mathbb{V}\text{ar} [T_i^{(s)}] \leq \frac{1}{h^{2d}} \mathbb{E} \left[ \left( \frac{X_i - x}{h} \right)^{2s} K^2 \left( \frac{X_i - x}{h} \right) \right] \leq \frac{1}{h^d} \int_{\mathbb{R}^d} u^{2s} K^2(u) p_X(x + hu | U) du \leq \frac{q^{dl'}}{h^d}.$$

By Bernstein's inequality, we get

$$\mathbb{P} \left\{ |\bar{W}_s - W_s| \geq \epsilon \right\} = \mathbb{P} \left\{ \left| \frac{1}{n} \sum_{i=1}^n T_i^{(s)} \right| > \epsilon \right\} \leq 2 \exp \left( \frac{-nh^d \epsilon^2}{q^{1+dl'} + 4\epsilon/3} \right).$$

Combining this inequality with (3.8), one obtains

$$\begin{aligned} \mathbb{P} \{ J_1 \geq 3\delta/4 \} &\leq \sum_{|s| \leq p} \mathbb{P} \left\{ |\bar{W}_s - W_s| \geq 3\mu_0 q^{dl'} M^{-1} \delta/4 \right\} \\ &= \sum_{|s| \leq p} \mathbb{P} \left\{ \left| \frac{1}{n} \sum_{i=1}^n T_i^{(s)} \right| \geq 3\mu_0 q^{dl'} M^{-1} \delta/4 \right\} \leq 2M \exp \left( \frac{-9\mu_0^2 M^{-2} q^{d(l'-l)} n \delta^2 / 16}{2 + \mu_0 M^{-1} \delta} \right). \end{aligned} \quad (3.9)$$

Now, we analyze  $J_2$ . Note that

$$J_2 \leq \|(\bar{B}^{-1} - B^{-1}) \bar{W}\| \leq \|\bar{B}^{-1} - B^{-1}\| \|\bar{W}\| \leq M \|\bar{B}^{-1} - B^{-1}\| \cdot \max_s |\bar{W}_s| \leq M \|\bar{B}^{-1} - B^{-1}\| h^{-d}. \quad (3.10)$$

Define  $Z := \bar{B} - B$ . One has

$$\lambda_{\max}(Z) \leq \sum_{|s_1|, |s_2| \leq p} |Z_{s_1, s_2}|.$$

Define

$$T_i^{(s_1, s_2)} := \frac{1}{h^d} \left( \frac{X_i - x}{h} \right)^{s_1 + s_2} K \left( \frac{X_i - x}{h} \right) - \int_{\mathbb{R}^d} u^{s_1 + s_2} K(u) p_X(x + hu | U) du.$$

We have  $\mathbb{E} [T_i^{(s_1, s_2)}] = 0$ ,  $|T_i^{(s_1, s_2)}| \leq 2h^{-d}$ , and

$$\text{Var} [T_i^{(s_1, s_2)}] \leq \mathbb{E} \left[ \frac{1}{h^{2d}} \left( \frac{X_i - x}{h} \right)^{2s_1 + 2s_2} K^2 \left( \frac{X_i - x}{h} \right) \right] = \frac{1}{h^d} \int_{\mathbb{R}^d} u^{2s_1 + 2s_2} K^2(u) p_X(x + hu | U) du \leq \frac{q^{dl'}}{h^d}.$$

By Bernstein's inequality, one obtains

$$\begin{aligned} \mathbb{P} \left\{ \lambda_{\max}(Z) \geq q^{dl'} M^{-1} \mu_0^2 \delta / 8 \right\} &\leq \mathbb{P} \left\{ \sum_{|s_1|, |s_2| \leq p} |Z_{s_1, s_2}| \geq q^{dl'} M^{-1} \mu_0^2 \delta / 8 \right\} \\ &\leq \sum_{|s_1|, |s_2| \leq p} \mathbb{P} \left\{ |Z_{s_1, s_2}| \geq q^{dl'} M^{-3} \mu_0^2 \delta / 8 \right\} \\ &= \sum_{|s_1|, |s_2| \leq p} \mathbb{P} \left\{ \left| \frac{1}{n} \sum_{i=1}^n T_i^{(s_1, s_2)} \right| \geq h^{-d} M^{-3} \mu_0^2 \delta / 8 \right\} \\ &\leq 2M^2 \exp \left( \frac{-n q^{d(l'-l)} M^{-6} \mu_0^4 \delta^2 / 64}{2 + M^{-3} \mu_0^2 \delta / 6} \right). \end{aligned}$$

By Lemma 2.1,  $\|B^{-1}\| \leq q^{-dl'} \mu_0^{-1}$ . That is, on the event  $\{\lambda_{\max}(Z) \leq q^{dl'} M^{-1} \mu_0^2 \delta / 8\}$ , one has  $\|B^{-\frac{1}{2}} Z B^{-\frac{1}{2}}\| \leq M^{-1} \mu_0 \delta / 8$  in which case if  $M^{-1} \mu_0 \delta / 8 < \frac{1}{2}$ , one obtains

$$\begin{aligned} \|\bar{B}^{-1} - B^{-1}\| &= \left\| B^{-\frac{1}{2}} \left( \left( I + B^{-\frac{1}{2}} Z B^{-\frac{1}{2}} \right)^{-1} - I \right) B^{-\frac{1}{2}} \right\| \\ &\leq \|B^{-1}\| \left\| \left( I + B^{-\frac{1}{2}} Z B^{-\frac{1}{2}} \right)^{-1} - I \right\| \\ &\leq q^{-dl'} \mu_0^{-1} \sum_{j=1}^{\infty} \left\| B^{-\frac{1}{2}} Z B^{-\frac{1}{2}} \right\|^j \\ &\leq q^{-dl'} \mu_0^{-1} \sum_{j=1}^{\infty} (M^{-1} \mu_0 \delta / 8)^j \leq q^{-dl'} M^{-1} \delta / 4. \end{aligned}$$

This inequality along with (3.10) imply  $J_2 \leq \delta/4$ . In other words,

$$\mathbb{P}\{J_2 \geq \delta/4\} \leq \mathbb{P}\left\{\lambda_{\max}(Z) \geq q^{d'l} M^{-1} \mu_0^2 \delta/8\right\} \leq 2M^2 \exp\left(\frac{-nq^{d(l'-l)} M^{-6} \mu_0^4 \delta^2/64}{2 + M^{-3} \mu_0^2 \delta/6}\right).$$

Combining this inequality with (3.7) and (3.9) gives

$$\mathbb{P}\left\{\left|\hat{\eta}^{\text{LP}}(x; \mathcal{D}, h, p) - \Gamma_{q^{-l}}^p \eta(x; U)\right| \geq \delta\right\} \leq C_4 \exp\left(-C_5 n q^{d(l'-l)} \delta^2\right)$$

if  $M^{-1} \mu_0 \delta/8 < \frac{1}{2}$ . This concludes the proof.  $\blacksquare$

## 4 Auxiliary analysis for Section 2.1

### 4.1 Analysis of Part 1 of Example 1

**Step 1.** Following the proof of Theorem 4.1 in Rigollet and Zeevi (2010), we first construct a problem

instance in  $\mathcal{P}(\beta, \alpha, d)$ . Define  $M := \left\lfloor 2^{-1} c_0^{-1} \left(\frac{2 \log 2}{T}\right)^{\frac{-\beta}{2\beta+d}} \right\rfloor^{\frac{1}{\beta}}$  and let  $\mathcal{B} := \{\mathbf{B}_m, m = 1, \dots, M^d\}$  be a re-indexed collection of the hypercubes

$$\mathbf{B}_m = \mathbf{B}_{\mathbf{m}} := \left\{x \in [0, 1]^d : \frac{\mathbf{m}_i - 1}{2^l} \leq x_i \leq \frac{\mathbf{m}_i}{2^l}, i \in \{1, \dots, d\}\right\},$$

for  $\mathbf{m} = (\mathbf{m}_1, \dots, \mathbf{m}_d)$  with  $\mathbf{m}_i \in \{1, \dots, M\}$ . Consider the regular grid  $\mathcal{Q} = \{a_1, a_2, \dots, a_{M^d}\}$ , where  $a_k$  denotes the center of bin  $\mathbf{B}_k, k = 1, \dots, M^d$ . Define  $C := 2^{\beta-1} L \wedge \frac{1}{4}$  and let  $\phi$  be defined as follows:

$$\phi(x) = \begin{cases} (1 - \|x\|_{\infty})^{\beta} & \text{if } \|x\|_{\infty} \leq 1 \\ 0 & \text{o.w.} \end{cases}.$$

Define  $m := \lceil \mu M^{d-\alpha\beta} \rceil$ , where  $\mu \in (0, 1)$  is chosen small enough to ensure  $m \leq M^d$ . Define the payoff functions as follows:

$$f_1(x) = \frac{1}{2} + \sum_{j=1}^m M^{-\beta} C \phi(M[x - a_j]), \quad f_2(x) = \frac{1}{2},$$

and assume that covariates are distributed uniformly. Similar to the proof of Theorem 4.1 in Rigollet and Zeevi (2010), one can show that the margin condition and smoothness condition in Assumptions 3 and 1 are satisfied for the constructed problem instance.

**Step 2.** Next, we lower bound the regret of  $\text{ABSE}(\tilde{\beta})$  under the constructed problem instance. To do so, we use the same exact terminology and notation as in Perchet and Rigollet (2013); for the sake of brevity, we do not re-introduce the notation here. By construction, for all bins  $\mathbf{B}$  with  $|\mathbf{B}| = 2^{-k}$ ,  $k = 0, 1, \dots, k_0$ , we have  $\underline{\mathcal{I}}_{\mathbf{B}} = \mathcal{K} = \{1, 2\}$ . Define the event  $\mathcal{W}_{\mathbf{B},s} := \{\underline{\mathcal{I}}_{\mathbf{B}} \subseteq \mathbb{I}_{\mathbf{B},s}\} = \{\mathbb{I}_{\mathbf{B},s} = \mathcal{K}\}$  and  $\mathcal{V}_{\mathbf{B}} := \bigcap_{\mathbf{B}' \in \mathcal{P}(\mathbf{B})} \mathcal{W}_{\mathbf{B}',t_{\mathbf{B}}}$ . Let

$$\mathcal{A}_1 := \left\{ \exists t \leq T; \exists \mathbf{B} \in \mathcal{L}_t; \exists s \leq l_{\mathbf{B}} : \mathbb{I}_{\mathbf{B},s} \neq \mathcal{K} \text{ and } |\mathbf{B}| \geq 2^{-k_0+1} \right\}$$

denote the event where one of the arms is eliminated in at least one of the bins at depth less than  $k_0$ . One has:

$$\mathbb{P}\{\mathcal{A}_1\} \leq \sum_{k=1}^{k_0-1} \sum_{|\mathbf{B}|=2^{-k}} \mathbb{P}\{\mathcal{V}_{\mathbf{B}} \cap \bar{\mathcal{W}}_{\mathbf{B},t_{\mathbf{B}}}\}. \quad (4.1)$$

Note that for any bin  $\mathbf{B}$  with  $|\mathbf{B}| \geq 2^{-k_0+1}$ ,  $\left| \bar{f}_{\mathbf{B}}^{(1)} - \bar{f}_{\mathbf{B}}^{(2)} \right| < c_0 |\mathbf{B}|^{\beta} \leq \frac{\epsilon_{\mathbf{B},l_{\mathbf{B}}}}{2}$ . This implies that  $\mathcal{W}_{\mathbf{B}}$  can only happen if either  $\bar{f}_{\mathbf{B}}^{(1)}$  or  $\bar{f}_{\mathbf{B}}^{(2)}$  does not belong to its respective confidence interval  $[\bar{Y}_{\mathbf{B},s}^{(1)} \pm \epsilon_{\mathbf{B},s}]$  or  $[\bar{Y}_{\mathbf{B},s}^{(2)} \pm \epsilon_{\mathbf{B},s}]$  for some  $s \leq l_{\mathbf{B}}$ . Therefore, since  $-\bar{f}_{\mathbf{B}}^{(i)} \leq Y_s - \bar{f}_{\mathbf{B}}^{(i)} \leq 1 - \bar{f}_{\mathbf{B}}^{(i)}$ ,

$$\mathbb{P}\{\mathcal{V}_{\mathbf{B}} \cap \bar{\mathcal{W}}_{\mathbf{B},t_{\mathbf{B}}}\} \leq \mathbb{P}\left\{ \exists s \leq l_{\mathbf{B}}; \exists i \in \mathcal{K} : \left| \bar{Y}_{\mathbf{B},s}^{(i)} - \bar{f}_{\mathbf{B}}^{(i)} \right| \geq \frac{\epsilon_{\mathbf{B},s}}{4} \right\} \leq \frac{4l_{\mathbf{B}}}{T|\mathbf{B}|^d}. \quad (4.2)$$

Putting together (4.1) and (4.2), one obtains

$$\mathbb{P}\{\mathcal{A}_1\} \leq \sum_{k=1}^{k_0-1} \frac{4C_l 2^{-2\tilde{\beta}k} \log\left(T 2^{(2\tilde{\beta}+d)k}\right)}{T 2^{-kd}} \leq \frac{4C_l 2^{-(2\tilde{\beta}-d)k_0} \log(T^2)}{T} \leq 8C_l T^{\frac{-4\tilde{\beta}}{2\tilde{\beta}+d}} \log T. \quad (4.3)$$

**Step 3.** Let  $\tilde{c} := 2^{1-d-2\tilde{\beta}} c_0^{-2} \log 2$  and define

$$\mathcal{A}_2 := \left\{ \exists t \leq \tilde{c}T/2; \exists \mathbf{B} \in \mathcal{L}_t : |\mathbf{B}| \geq 2^{-k_0+1} \right\}$$

to be the event that for some  $t \leq \tilde{c}T/2$  some bin at depth  $k_0$  becomes live. Note that for a bin  $\mathbf{B}$  to become live by  $t = \lfloor \tilde{c}T/2 \rfloor$ , we need  $l_{\mathbf{p}(\mathbf{B})}$  number of covariates to fall into its parent  $\mathbf{p}(\mathbf{B})$  by  $t = \lfloor \tilde{c}T/2 \rfloor$ . Let  $Z_{\mathbf{B},t} = \mathbb{1}\{X_t \in \mathbf{p}(\mathbf{B})\}$ . Note that  $|Z_t| \leq 1$ ,  $\mathbb{E}Z_t = |\mathbf{B}|^d$ , and  $\text{Var}Z_{\mathbf{B},t} \leq \mathbb{E}Z_{\mathbf{B},t}^2 = |\mathbf{B}|^d$ . Hence, one can

apply the Bernstein's inequality in Lemma 1.4 to to  $\sum_{t=1}^{\lfloor \tilde{c}T/2 \rfloor} Z_{\mathbf{B},t}$  for  $|\mathbf{B}| = 2^{-k_0}$  to obtain

$$\begin{aligned}
\mathbb{P}\{\mathcal{A}_2\} &\leq \sum_{|\mathbf{B}|=2^{-k_0}} \mathbb{P}\left\{ \sum_{t=1}^{\lfloor \tilde{c}T/2 \rfloor} Z_{\mathbf{B},t} \geq l_{\mathbf{p}(\mathbf{B})} \right\} \\
&\stackrel{(a)}{\leq} \sum_{|\mathbf{B}|=2^{-k_0}} \mathbb{P}\left\{ \sum_{t=1}^{\lfloor \tilde{c}T/2 \rfloor} Z_{\mathbf{B},t} \geq c_0^{-2} |\mathbf{p}(\mathbf{B})|^{-2\tilde{\beta}} \right\} \\
&\leq 2^{k_0 d} \exp\left( -\frac{c_0^{-4} 2^{4(k_0-1)\tilde{\beta}-1}/2}{\tilde{c}T 2^{(k_0-1)d-1} + c_0^{-2} 2^{2(k_0-1)\tilde{\beta}-1}/3} \right) \\
&= 2^{k_0 d} \exp\left( -\frac{c_0^{-2} 2^{2(k_0-1)\tilde{\beta}-1}/2}{\frac{1}{2} + 1/3} \right) \\
&\leq c_1 T^{\frac{d}{2\beta+d}} \exp\left( -c_2 T^{\frac{2\tilde{\beta}}{2\beta+d}} \right) \leq c_3 T^{-1}, \tag{4.4}
\end{aligned}$$

for some constants  $c_1, c_2, c_3 > 0$ , where (a) follows from  $l_{\mathbf{B}} \geq c_0^{-2} |\mathbf{B}|^{-2\tilde{\beta}}$  by the definition of  $l_{\mathbf{B}}$ .

**Step 4.** Let  $S := \{x \in [0, 1]^d : f_1(x) \neq \frac{1}{2}\}$ . Define the event

$$\mathcal{A}_3 := \left\{ \sum_{t=1}^{\lfloor \tilde{c}T/2 \rfloor} \mathbb{1}\{X_t \in S\} < \tilde{c}mM^{-d}T/4 \right\}.$$

Define  $Z_t := \mathbb{1}\{X_t \in S\}$  and note that  $|Z_t| \leq 1$ ,  $\mathbb{E}Z_t = mM^{-d}$ , and  $\text{Var}Z_t \leq \mathbb{E}Z_t = mM^{-d}$ . As a result we can apply the Bernstein's inequality in Lemma 1.4 to obtain

$$\mathbb{P}\{\mathcal{A}_3\} \leq \exp\left( -\tilde{c}mM^{-d}T/20 \right) \leq \exp\left( -c_5 T^{\frac{2\tilde{\beta}+d-\alpha\tilde{\beta}}{2\beta+d}} \right) \stackrel{(a)}{\leq} \exp\left( -c_4 T^{\frac{2\tilde{\beta}}{2\beta+d}} \right) \leq c_5 T^{-1}, \tag{4.5}$$

for some constants  $c_4, c_5 > 0$ , where (a) follows from the assumption that  $\alpha \leq \frac{1}{\tilde{\beta}} \leq \frac{1}{\beta}$ .

**Step 5.** Note that on the event  $\bar{\mathcal{A}}_1 \cap \bar{\mathcal{A}}_2$ , the  $\text{ABSE}(\tilde{\beta})$  has not eliminated any arms over any region of the covariate space up to time  $t = \lfloor \tilde{c}T/2 \rfloor$ . On the other hand, on the event  $\bar{\mathcal{A}}_3$ , up to time  $t = \lfloor \tilde{c}T/2 \rfloor$ , at least  $\tilde{c}mM^{-d}T/4$  number of covariates have fallen into  $S$ , where the first arm is strictly optimal. Recall the definition of the inferior sampling rate in (1.1). One has:

$$\begin{aligned}
\mathcal{S}^{\text{ABSE}(\tilde{\beta})}(\mathbf{P}; T) &\geq \mathbb{E}^\pi \left[ \sum_{t=1}^T \mathbb{1} \{f_{\pi_t^*}(X_t) \neq f_{\pi_t}(X_t)\} \middle| \bar{\mathcal{A}}_1 \cap \bar{\mathcal{A}}_2 \cap \bar{\mathcal{A}}_3 \right] \mathbb{P} \{ \bar{\mathcal{A}}_1 \cap \bar{\mathcal{A}}_2 \cap \bar{\mathcal{A}}_3 \} \\
&\stackrel{(a)}{\geq} \tilde{c} m M^{-d} T / 8 \left( 1 - 8C_l T^{\frac{-4\tilde{\beta}}{2\beta+d}} \log T - c_3 T^{-1} - c_5 T^{-1} \right) \\
&\geq c_6 T^{\frac{\alpha\tilde{\beta}}{2\beta+d}},
\end{aligned}$$

for some constant  $c_6 > 0$ , where (a) follows from (4.3), (4.4), and (4.5). Using this inequality along with Lemma 1.1, the result follows.  $\blacksquare$

## 4.2 Analysis of Part 2 of Example 1

**Step 1.** let  $\tilde{k} := \left\lceil \frac{\log_2(T/2 \log 2)}{(2\tilde{\beta}+1)} \right\rceil$ ,  $M := 2^{\tilde{k}}$ , and  $\mathcal{B} := \{\mathbf{B}_m, m = 1, \dots, M^d\}$  be a re-indexed collection of the hypercubes

$$\mathbf{B}_m = \mathbf{B}_m := \left\{ x \in [0, 1]^d : \frac{\mathbf{m}_i - 1}{2^l} \leq x_i \leq \frac{\mathbf{m}_i}{2^l}, i \in \{1, \dots, d\} \right\},$$

for  $\mathbf{m} = (\mathbf{m}_1, \dots, \mathbf{m}_d)$  with  $\mathbf{m}_i \in \{1, \dots, M\}$ . Consider the regular grid  $\mathcal{Q} = \{a_1, a_2, \dots, a_{M^d}\}$ , where  $a_k$  denotes the center of bin  $\mathbf{B}_k, k = 1, \dots, M^d$ . Define  $C := 2^{\beta-1} L \wedge \frac{1}{4}$  and let  $\phi$  be defined as follows:

$$\phi(x) = \begin{cases} (1 - |x_1|)^\beta & \text{if } |x_1| \leq 1 \\ 0 & \text{o.w.} \end{cases}.$$

Define  $m := 2 \times \lceil \mu M^{1-\alpha\beta} \rceil$ , where  $\mu \in (0, 1)$  is chosen small enough to ensure  $m \leq M^d$ . Define the payoff functions as follows:

$$f_1(x) = \frac{1}{2} + \sum_{j=1}^m M^{-\beta} C (-1)^j \phi(M[x_1 - \tilde{a}_j]), \quad f_2(x) = \frac{1}{2},$$

where  $\tilde{a}_j = \frac{j+\frac{1}{2}}{M}$  for each  $j = 1, \dots, m$ , and assume that covariates are distributed uniformly. Similar to the proof of Theorem 4.1 in Rigollet and Zeevi (2010), one can show that the margin condition and smoothness condition in Assumptions 3 and 1 are satisfied for the constructed problem instance.

**Step 2.** Next, we lower bound the regret of  $\text{ABSE}(\tilde{\beta})$  under the constructed problem instance. To do so, we use the same exact terminology and notation as in Perchet and Rigollet (2013); for the sake of brevity, we do not re-introduce the notation here. By construction, for all bins  $\mathbf{B}$  with  $|\mathbf{B}| = 2^{-k}, k = 0, 1, \dots, k_0$ ,

we have

$$\bar{f}_{\mathbf{B}}^{(i)} = \frac{1}{P_X(\mathbf{B})} \int_{\mathbf{B}} f_k(x) ddP_X(x) = \frac{1}{2}, \quad i \in \mathcal{K}.$$

Define the event  $\mathcal{W}_{\mathbf{B},s} := \{\mathbb{I}_{\mathbf{B},s} = \mathcal{K}\}$  and  $\mathcal{V}_{\mathbf{B}} := \bigcap_{\mathbf{B}' \in \mathcal{P}(\mathbf{B})} \mathcal{W}_{\mathbf{B}',t_{\mathbf{B}}}$ . Let

$$\mathcal{A}_1 := \left\{ \exists t \leq T; \exists \mathbf{B} \in \mathcal{L}_t; \exists s \leq l_{\mathbf{B}} : \mathbb{I}_{\mathbf{B},s} \neq \mathcal{K} \text{ and } |\mathbf{B}| \geq 2^{-k_0+1} \right\}$$

denote the event where one of the arms is eliminated in at least one of the bins at depth less than  $k_0$ .

One has:

$$\mathbb{P}\{\mathcal{A}_1\} \leq \sum_{k=1}^{k_0-1} \sum_{|\mathbf{B}|=2^{-k}} \mathbb{P}\{\mathcal{V}_{\mathbf{B}} \cap \bar{\mathcal{W}}_{\mathbf{B},t_{\mathbf{B}}}\}. \quad (4.6)$$

Note that for any bin  $\mathbf{B}$  with  $|\mathbf{B}| \geq 2^{-k_0+1}$ ,  $|\bar{f}_{\mathbf{B}}^{(1)} - \bar{f}_{\mathbf{B}}^{(2)}| = 0 < \frac{\epsilon_{\mathbf{B},t_{\mathbf{B}}}}{2}$ . This implies that  $\mathcal{W}_{\mathbf{B}}$  can only happen if either  $\bar{f}_{\mathbf{B}}^{(1)}$  or  $\bar{f}_{\mathbf{B}}^{(2)}$  does not belong to its respective confidence interval  $[\bar{Y}_{\mathbf{B},s}^{(1)} \pm \epsilon_{\mathbf{B},s}]$  or  $[\bar{Y}_{\mathbf{B},s}^{(2)} \pm \epsilon_{\mathbf{B},s}]$  for some  $s \leq l_{\mathbf{B}}$ . Therefore, since  $-\bar{f}_{\mathbf{B}}^{(i)} \leq Y_s - \bar{f}_{\mathbf{B}}^{(i)} \leq 1 - \bar{f}_{\mathbf{B}}^{(i)}$ ,

$$\mathbb{P}\{\mathcal{V}_{\mathbf{B}} \cap \bar{\mathcal{W}}_{\mathbf{B},t_{\mathbf{B}}}\} \leq \mathbb{P}\left\{ \exists s \leq l_{\mathbf{B}}; \exists i \in \mathcal{K} : \left| \bar{Y}_{\mathbf{B},s}^{(i)} - \bar{f}_{\mathbf{B}}^{(i)} \right| \geq \frac{\epsilon_{\mathbf{B},s}}{4} \right\} \leq \frac{4l_{\mathbf{B}}}{T|\mathbf{B}|^d}. \quad (4.7)$$

Putting together (4.6) and (4.7), one obtains

$$\mathbb{P}\{\mathcal{A}_1\} \leq \sum_{k=1}^{k_0-1} \frac{4C_l 2^{-2\tilde{\beta}k} \log(T 2^{(2\tilde{\beta}+d)k})}{T 2^{-kd}} \leq \frac{4C_l 2^{-(2\tilde{\beta}-d)k_0} \log(T^2)}{T} \leq 8C_l T^{\frac{-4\tilde{\beta}}{2\tilde{\beta}+d}} \log T. \quad (4.8)$$

**Step 3.** Let  $\tilde{c} := 2^{1-d-2\tilde{\beta}} c_0^{-2} \log 2$  and define

$$\mathcal{A}_2 := \left\{ \exists t \leq \tilde{c}T/2; \exists \mathbf{B} \in \mathcal{L}_t : |\mathbf{B}| < 2^{-k_0+1} \right\}$$

to be the event that for some  $t \leq \tilde{c}T/2$  some bin at depth  $k_0$  becomes live. Note that for a bin  $\mathbf{B}$  to become live by  $t = \lfloor \tilde{c}T/2 \rfloor$ , we need  $l_{\mathbf{p}(\mathbf{B})}$  number of covariates to fall into its parent  $\mathbf{p}(\mathbf{B})$  by  $t = \lfloor \tilde{c}T/2 \rfloor$ .

Let  $Z_{\mathbf{B},t} = \mathbb{1}\{X_t \in \mathbf{p}(\mathbf{B})\}$ . Note that  $|Z_t| \leq 1$ ,  $\mathbb{E}Z_t = |\mathbf{B}|^d$ , and  $\text{Var}Z_{\mathbf{B},t} \leq \mathbb{E}Z_{\mathbf{B},t}^2 = |\mathbf{B}|^d$ . Hence, one can

apply the Bernstein's inequality in Lemma 1.4 to to  $\sum_{t=1}^{\lfloor \tilde{c}T/2 \rfloor} Z_{\mathbf{B},t}$  for  $|\mathbf{B}| = 2^{-k_0}$  to obtain

$$\begin{aligned}
\mathbb{P}\{\mathcal{A}_2\} &\leq \sum_{|\mathbf{B}|=2^{-k_0}} \mathbb{P}\left\{ \sum_{t=1}^{\lfloor \tilde{c}T/2 \rfloor} Z_{\mathbf{B},t} \geq l_{\mathbf{p}(\mathbf{B})} \right\} \\
&\stackrel{(a)}{\leq} \sum_{|\mathbf{B}|=2^{-k_0}} \mathbb{P}\left\{ \sum_{t=1}^{\lfloor \tilde{c}T/2 \rfloor} Z_{\mathbf{B},t} \geq c_0^{-2} |\mathbf{p}(\mathbf{B})|^{-2\tilde{\beta}} \right\} \\
&\leq 2^{k_0 d} \exp\left( -\frac{c_0^{-4} 2^{4(k_0-1)\tilde{\beta}-1}/2}{\tilde{c}T 2^{(k_0-1)d-1} + c_0^{-2} 2^{2(k_0-1)\tilde{\beta}-1}/3} \right) \\
&= 2^{k_0 d} \exp\left( -\frac{c_0^{-2} 2^{2(k_0-1)\tilde{\beta}-1}/2}{\frac{1}{2} + 1/3} \right) \\
&\leq c_1 T^{\frac{d}{2\beta+d}} \exp\left( -c_2 T^{\frac{2\tilde{\beta}}{2\beta+d}} \right) \leq c_3 T^{-1}, \tag{4.9}
\end{aligned}$$

for some constants  $c_1, c_2, c_3 > 0$ , where (a) follows from  $l_{\mathbf{B}} \geq c_0^{-2} |\mathbf{B}|^{-2\tilde{\beta}}$  by the definition of  $l_{\mathbf{B}}$ .

**Step 2.** Let  $S_1 := \{x \in [0, 1]^d : f_1(x) > \frac{1}{2}\}$  and  $S_2 := \{x \in [0, 1]^d : f_1(x) < \frac{1}{2}\}$ . Define the events

$$\mathcal{A}_{31} := \left\{ \sum_{t=1}^T \mathbb{1}\{X_t \in S_1\} < \tilde{c}mM^{-d}T/8 \right\}, \quad \mathcal{A}_{32} := \left\{ \sum_{t=1}^T \mathbb{1}\{X_t \in S_2\} < \tilde{c}mM^{-d}T/8 \right\},$$

and let  $\mathcal{A}_3 = \mathcal{A}_{31} \cap \mathcal{A}_{32}$ . Define  $Z_t := \mathbb{1}\{X_t \in S_1\}$  and note that  $|Z_t| \leq 1$ ,  $\mathbb{E}Z_t = \tilde{c}mM^{-1}/4$ , and  $\text{Var}Z_t \leq \mathbb{E}Z_t = \tilde{c}mM^{-1}/4$ . As a result we can apply the Bernstein's inequality in Lemma 1.4 to obtain

$$\mathbb{P}\{\mathcal{A}_{31}\} \leq \exp(-\tilde{c}mM^{-1}T/40) \leq \exp\left( -c_5 T^{\frac{2\tilde{\beta}+d-\alpha\beta}{2\beta+d}} \right) \stackrel{(a)}{\leq} \exp\left( -c_4 T^{\frac{2\tilde{\beta}}{2\beta+d}} \right) \leq c'_5 T^{-1},$$

for some constants  $c_4, c'_5 > 0$ , where (a) follows from the assumption that  $\alpha \leq \frac{1}{\beta}$ . A similar upper bound can be shown for the probability of the event  $\mathcal{A}_{32}$ , which implies

$$\mathbb{P}\{\mathcal{A}_3\} \leq c_5 T^{-1}, \tag{4.10}$$

for some constants  $c_5 > 0$ .

**Step 5.** Note that on the event  $\bar{\mathcal{A}}_1 \cap \bar{\mathcal{A}}_2$ , the ABSE( $\tilde{\beta}$ ) has not eliminated any arms over any region of the covariate space up to time  $t = \lfloor \tilde{c}T/2 \rfloor$ . On the other hand, on the event  $\bar{\mathcal{A}}_3$ , up to time  $t = \lfloor \tilde{c}T/2 \rfloor$ , at least  $\tilde{c}mM^{-d}T/8$  number of covariates have fallen into  $S_1$  and also  $S_2$ , where the first arm and the second arm are strictly optimal, respectively. Recall the definition of the inferior sampling rate in (1.1). One has:

$$\begin{aligned} \mathcal{S}^{\text{ABSE}(\tilde{\beta})}(\mathcal{P}; T) &\geq \mathbb{E}^\pi \left[ \sum_{t=1}^T \mathbb{1} \{f_{\pi_t^*}(X_t) \neq f_{\pi_t}(X_t)\} \middle| \bar{\mathcal{A}}_1 \cap \bar{\mathcal{A}}_2 \cap \bar{\mathcal{A}}_3 \right] \mathbb{P} \{ \bar{\mathcal{A}}_1 \cap \bar{\mathcal{A}}_2 \cap \bar{\mathcal{A}}_3 \} \\ &\stackrel{(a)}{\geq} \tilde{c}mM^{-1}T/8 \left( 1 - 8C_l T^{\frac{-4\tilde{\beta}}{2\beta+d}} \log T - c_3 T^{-1} - c_5 T^{-1} \right) \\ &\geq c_6 T^{\frac{-\alpha\tilde{\beta}}{2\beta+d}}, \end{aligned}$$

for some constant  $c_6 > 0$ , where (a) follows from (4.8), (4.9), and (4.10). Using this inequality along with Lemma 1.1, the result follows.  $\blacksquare$

## 5 Proof of auxiliary lemmas

### 5.1 Proof of Lemma 1.2

**Lemma.** Suppose  $f \in \mathcal{H}_{\mathbb{R}^d}(\beta, L)$  for some  $\mathcal{X} \in \mathbb{R}^d$ ,  $0 < \beta \leq 1$  and  $L > 0$ , and define the function  $g$  such that  $g(x) = C^{-\beta} f(Cx)$  for all  $x \in \mathbb{R}^d$  and some  $C > 0$ . Then,  $g \in \mathcal{H}_{\mathbb{R}^d}(\beta, L)$ .

*Proof.* For any  $x, y \in \mathbb{R}^d$ , one has

$$|g(x) - g(y)| = C^{-\beta} |f(Cx) - f(Cy)| \leq C^{-\beta} L \|Cx - Cy\|_\infty^\beta = L \|x - y\|_\infty^\beta.$$

This concludes the proof.  $\blacksquare$

### 5.2 Proof of Lemma 1.3

**Lemma.** Suppose  $f, g \in \mathcal{H}_{\mathcal{X}}(\beta, L)$  for some  $\mathcal{X} \subseteq \mathbb{R}^d$ ,  $0 < \beta \leq 1$  and  $L > 0$ , and define the functions  $h_1 := \max(f, g)$  and  $h_2 := \min(f, g)$ . Then,  $h_1, h_2 \in \mathcal{H}_{\mathcal{X}}(\beta, L)$ .

*Proof.* We only prove the result for the function  $h_1$ . A similar analysis can be used for  $h_2$ . Fix some

$x, y \in \mathcal{X}$ . If  $h_1(x) = f(x)$  and  $h_1(y) = f(y)$ , or  $h_1(x) = g(x)$  and  $h_1(y) = g(y)$  then, one has

$$|h_1(x) - h_1(y)| \leq L \|x - y\|_\infty^\beta.$$

Now suppose  $h_1(x) = f(x)$  and  $h_1(y) = g(y)$ . Without loss of generality, assume that  $f(x) \leq g(y)$  then, one has

$$|h_1(x) - h_1(y)| \leq |g(x) - g(y)| \leq L \|x - y\|_\infty^\beta.$$

The case  $h_1(x) = f(x)$  and  $h_1(y) = g(y)$  can be analyzed similarly. This concludes the proof.  $\blacksquare$

### 5.3 Proof of Lemma 1.5

**Lemma.** *Let  $\rho_0, \rho_1$  be two probability distributions supported on some set  $\mathcal{X}$ , with  $\rho_0$  absolutely continuous with respect to  $\rho_1$ . Then for any measurable function  $\Psi : \mathcal{X} \rightarrow \{0, 1\}$ , one has:*

$$\mathbb{P}_{\rho_0}\{\Psi(X) = 1\} + \mathbb{P}_{\rho_1}\{\Psi(X) = 0\} \geq \frac{1}{2} \exp(-\text{KL}(\rho_0, \rho_1)).$$

*Proof.* Define  $\mathcal{B}$  to be the event that  $\Psi(X) = 1$ . One has

$$\mathbb{P}_{\rho_0}\{\Psi(X) = 1\} + \mathbb{P}_{\rho_1}\{\Psi(X) = 0\} = \mathbb{P}_{\rho_0}\{\mathcal{B}\} + \mathbb{P}_{\rho_1}\{\bar{\mathcal{B}}\} \geq \int \min\{d\rho_0, d\rho_1\} \geq \frac{1}{2} \exp(-\text{KL}(\rho_0, \rho_1)),$$

where the last inequality follows from Tsybakov 2008, Lemma 2.6.  $\blacksquare$

### 5.4 Proof of Lemma 1.10

**Lemma.** *Suppose  $f, g \in \mathcal{H}_{\mathcal{X}}(\beta, L)$  for some  $\mathcal{X} \subseteq [0, 1]$ ,  $\beta > 0$ , and  $L > 0$ , and define the function  $h := f \cdot g$  as the product of  $f$  and  $g$ . Then,  $h \in \mathcal{H}(\beta, L')$  for some  $L' > 0$ .*

*Proof.* Note that  $h$  is  $\lfloor \beta \rfloor$  times continuously differentiable. Hence, we only need to show that there exists some  $L' > 0$  such that for any  $x, x' \in \mathcal{X}$ ,

$$|h(x') - h_x(x')| \leq L' \|x - x'\|_\infty^\beta.$$

By the triangle inequality, one has

$$\begin{aligned} |h(x') - h_x(x')| &\leq |(f \cdot g)_x(x') - f_x(x') \cdot g_x(x')| + |f_x(x') \cdot g(x') - f_x(x') \cdot g_x(x')| \\ &\quad + |f_x(x') \cdot g(x') - f(x') \cdot g(x')|. \end{aligned} \quad (5.1)$$

Since  $\mathcal{X} \subseteq [0, 1]$  and  $f, g \in \mathcal{H}_{\mathcal{X}}(\beta, L)$ , one has for some  $L_1, L_2 > 0$ :

$$|f_x(x') \cdot g(x') - f_x(x') \cdot g_x(x')| = |f_x(x')| \cdot |g(x') - g_x(x')| \leq L_1 \|x - x'\|_{\infty}^{\beta}; \quad (5.2)$$

$$|f_x(x') \cdot g(x') - f(x') \cdot g(x')| = |f_x(x') - f(x')| |g(x')| \leq L_2 \|x - x'\|_{\infty}^{\beta}. \quad (5.3)$$

Furthermore, let  $\{a_s\}_{0 \leq s \leq \lfloor \beta \rfloor}$ ,  $\{b_s\}_{0 \leq s \leq \lfloor \beta \rfloor}$ , and  $\{c_s\}_{0 \leq s \leq \lfloor \beta \rfloor}$  be the coefficients of the Taylor expansions  $f_x(x')$ ,  $g_x(x')$ , and  $h_x(x')$ , respectively. Notably,  $c_s = \sum_{s'=0}^s a_{s'} b_{s-s'}$ . This equality implies that

$$|(f \cdot g)_x(x') - f_x(x') \cdot g_x(x')| = \left| \sum_{s=\lfloor \beta \rfloor+1}^{2\lfloor \beta \rfloor} \sum_{s'=0}^s a_{s'} b_{s-s'} (x - x')^s \right| \leq L_3 \|x - x'\|_{\infty}^{\beta}, \quad (5.4)$$

for some  $L_3 > 0$ . Then, the result follows from putting together (5.1), (5.2), (5.3), and (5.4).  $\blacksquare$

## 6 Numerical Analysis

### 6.1 Different parameters

In this section we provide the auxiliary results of the numerical study in §6, which also includes experiments with different values for smoothness parameter  $\beta$  and horizon length  $T$ . Let  $\hat{\mathcal{R}}^{\pi}(\mathbf{P}; T)$  be empirical cumulative regret of policy  $\pi$  in the simulation then, the average Relative Loss,  $\text{RL}^{\pi}(\mathbf{P}; T)$ , with respect the  $\text{ABSE}(\beta)$  is defined as follows:

$$\text{RL}^{\pi}(\mathbf{P}; T) := \frac{\hat{\mathcal{R}}^{\pi}(\mathbf{P}; T) - \hat{\mathcal{R}}^{\text{ABSE}(\beta)}(\mathbf{P}; T)}{\hat{\mathcal{R}}^{\text{ABSE}(\beta)}(\mathbf{P}; T)}.$$

Table 1 (Table 2) provides the average cumulative regret and Relative Loss for Setting I (Setting II) for a fixed horizon length  $T = 2 \times 10^6$  and different smoothness parameters  $\beta \in \{0.85, 0.9, 0.95\}$  ( $\beta \in \{0.45, 0.5, 0.55\}$ ). Table 3 (Table 4) provides the average cumulative regret and Relative Loss for Setting I (Setting II) for a fixed smoothness parameter  $\beta = 0.9$  ( $\beta = 0.5$ ) and different horizon lengths  $T \in \{2 \times 10^6, 2.5 \times 10^6, 3 \times 10^6\}$ . These tables show that the results of numerical analysis in §6 are consistent across different smoothness values and larger horizon lengths.

| $\beta$ | ABSE( $\beta$ ) | SACB | ABSE( $\tilde{\beta}$ ) |      |      |      |      |      |      |      |      |      |      |      |      |
|---------|-----------------|------|-------------------------|------|------|------|------|------|------|------|------|------|------|------|------|
|         |                 |      | 0.4                     | 0.45 | 0.5  | 0.55 | 0.6  | 0.65 | 0.7  | 0.75 | 0.8  | 0.85 | 0.9  | 0.95 | 1    |
| 0.85    | 3.38            | 4.08 | 4.29                    | 4.30 | 4.30 | 4.30 | 4.30 | 4.30 | 4.29 | 3.92 | 4.13 | 3.38 | 3.10 | 2.98 | 2.73 |
| 0.9     | 1.88            | 2.30 | 2.72                    | 2.72 | 2.72 | 2.72 | 2.72 | 2.73 | 2.59 | 2.42 | 2.39 | 2.01 | 1.88 | 1.71 | 1.58 |
| 0.95    | 2.01            | 2.63 | 3.38                    | 3.38 | 3.39 | 3.39 | 3.38 | 3.39 | 3.06 | 2.92 | 2.75 | 2.35 | 2.24 | 2.01 | 1.82 |

| $\beta$ | SACB | ABSE( $\tilde{\beta}$ ) |      |     |      |     |      |     |      |     |      |     |      |      |  |
|---------|------|-------------------------|------|-----|------|-----|------|-----|------|-----|------|-----|------|------|--|
|         |      | 0.4                     | 0.45 | 0.5 | 0.55 | 0.6 | 0.65 | 0.7 | 0.75 | 0.8 | 0.85 | 0.9 | 0.95 | 1    |  |
| 0.85    | 20%  | 27%                     | 27%  | 27% | 27%  | 27% | 27%  | 26% | 16%  | 22% | 0%   | -8% | -11% | -19% |  |
| 0.9     | 22%  | 44%                     | 44%  | 45% | 45%  | 45% | 45%  | 37% | 28%  | 27% | 7%   | 0%  | -8%  | -15% |  |
| 0.95    | 31%  | 68%                     | 68%  | 68% | 68%  | 68% | 68%  | 52% | 45%  | 36% | 17%  | 11% | 0%   | -9%  |  |

Table 1: Results for Setting I with fixed horizon  $T = 2 \times 10^6$ , and different choices of  $\beta$ . *Above*: Average cumulative regret divided by  $10^4$ ; *Below*: Relative Loss.

| $\beta$ | ABSE( $\beta$ ) | SACB | ABSE( $\tilde{\beta}$ ) |      |      |      |      |      |      |      |      |      |      |      |      |
|---------|-----------------|------|-------------------------|------|------|------|------|------|------|------|------|------|------|------|------|
|         |                 |      | 0.4                     | 0.45 | 0.5  | 0.55 | 0.6  | 0.65 | 0.7  | 0.75 | 0.8  | 0.85 | 0.9  | 0.95 | 1    |
| 0.45    | 0.75            | 1.27 | 0.55                    | 0.75 | 1.04 | 1.46 | 2.08 | 2.91 | 4.04 | 5.55 | 5.92 | 5.92 | 5.92 | 5.91 | 5.91 |
| 0.5     | 0.81            | 1.16 | 0.48                    | 0.62 | 0.81 | 1.10 | 1.52 | 2.09 | 2.88 | 3.89 | 4.08 | 4.06 | 4.07 | 4.08 | 4.07 |
| 0.55    | 0.88            | 1.39 | 0.46                    | 0.54 | 0.66 | 0.88 | 1.16 | 1.56 | 2.09 | 2.75 | 2.77 | 2.77 | 2.77 | 2.77 | 2.76 |

| $\beta$ | SACB | ABSE( $\tilde{\beta}$ ) |      |      |      |      |      |      |      |      |      |      |      |      |  |
|---------|------|-------------------------|------|------|------|------|------|------|------|------|------|------|------|------|--|
|         |      | 0.4                     | 0.45 | 0.5  | 0.55 | 0.6  | 0.65 | 0.7  | 0.75 | 0.8  | 0.85 | 0.9  | 0.95 | 1    |  |
| 0.45    | 68%  | -25%                    | 0%   | 38%  | 94%  | 176% | 286% | 436% | 635% | 685% | 685% | 685% | 683% | 684% |  |
| 0.5     | 42%  | -41%                    | -24% | 0%   | 35%  | 86%  | 156% | 252% | 376% | 399% | 396% | 397% | 399% | 397% |  |
| 0.55    | 58%  | -47%                    | -38% | -24% | 0%   | 32%  | 77%  | 138% | 213% | 214% | 214% | 214% | 215% | 214% |  |

Table 2: Results for Setting II with fixed horizon  $T = 2 \times 10^6$ , and different choices of  $\beta$ . *Above*: Average cumulative regret divided by  $10^3t$ ; *Below*: Relative Loss.

| $T$               | ABSE( $\beta$ ) | SACB | ABSE( $\tilde{\beta}$ ) |      |      |      |      |      |      |      |      |      |      |      |      |      |
|-------------------|-----------------|------|-------------------------|------|------|------|------|------|------|------|------|------|------|------|------|------|
|                   |                 |      | 0.4                     | 0.45 | 0.5  | 0.55 | 0.6  | 0.65 | 0.7  | 0.75 | 0.8  | 0.85 | 0.9  | 0.95 | 1    |      |
| $2 \times 10^6$   | 1.88            | 2.30 | 2.72                    | 2.72 | 2.72 | 2.72 | 2.72 | 2.72 | 2.73 | 2.59 | 2.42 | 2.39 | 2.01 | 1.88 | 1.71 | 1.58 |
| $2.5 \times 10^6$ | 1.97            | 2.45 | 2.94                    | 2.95 | 2.95 | 2.95 | 2.95 | 2.95 | 2.95 | 2.94 | 2.55 | 2.64 | 2.23 | 1.97 | 1.90 | 1.66 |
| $3 \times 10^6$   | 2.09            | 2.60 | 3.14                    | 3.14 | 3.14 | 3.14 | 3.14 | 3.14 | 3.11 | 3.14 | 2.71 | 2.70 | 2.44 | 2.09 | 2.03 | 1.74 |

| $T$               | SACB | ABSE( $\tilde{\beta}$ ) |      |     |      |     |      |     |      |     |      |     |      |      |      |
|-------------------|------|-------------------------|------|-----|------|-----|------|-----|------|-----|------|-----|------|------|------|
|                   |      | 0.4                     | 0.45 | 0.5 | 0.55 | 0.6 | 0.65 | 0.7 | 0.75 | 0.8 | 0.85 | 0.9 | 0.95 | 1    |      |
| $2 \times 10^6$   | 22%  | 44%                     | 44%  | 45% | 45%  | 45% | 45%  | 45% | 37%  | 28% | 27%  | 7%  | 0%   | -8%  | -15% |
| $2.5 \times 10^6$ | 23%  | 49%                     | 49%  | 49% | 49%  | 49% | 49%  | 49% | 48%  | 29% | 33%  | 12% | 0%   | -3%  | -15% |
| $3 \times 10^6$   | 24%  | 49%                     | 49%  | 49% | 50%  | 50% | 48%  | 50% | 29%  | 28% | 16%  | 0%  | -3%  | -16% |      |

Table 3: Results for Setting I with fixed smoothness  $\beta = 0.9$  and different values of horizon length  $T$ . *Above*: Average cumulative regret divided by  $10^4$ ; *Below*: Relative Loss.

| $T$               | ABSE( $\beta$ ) | SACB | ABSE( $\tilde{\beta}$ ) |      |      |      |      |      |      |      |      |      |      |      |      |
|-------------------|-----------------|------|-------------------------|------|------|------|------|------|------|------|------|------|------|------|------|
|                   |                 |      | 0.4                     | 0.45 | 0.5  | 0.55 | 0.6  | 0.65 | 0.7  | 0.75 | 0.8  | 0.85 | 0.9  | 0.95 | 1    |
| $2 \times 10^6$   | 0.81            | 1.16 | 0.48                    | 0.62 | 0.81 | 1.10 | 1.52 | 2.09 | 2.88 | 3.89 | 4.08 | 4.06 | 4.07 | 4.08 | 4.07 |
| $2.5 \times 10^6$ | 0.85            | 1.42 | 0.50                    | 0.64 | 0.85 | 1.17 | 1.61 | 2.23 | 3.08 | 4.21 | 5.07 | 5.08 | 5.08 | 5.09 | 5.07 |
| $3 \times 10^6$   | 0.89            | 1.46 | 0.52                    | 0.67 | 0.89 | 1.20 | 1.69 | 2.36 | 3.28 | 4.53 | 6.09 | 6.08 | 6.09 | 6.11 | 6.08 |

  

| $T$               | SACB | ABSE( $\tilde{\beta}$ ) |      |     |      |     |      |      |      |      |      |      |      |      |
|-------------------|------|-------------------------|------|-----|------|-----|------|------|------|------|------|------|------|------|
|                   |      | 0.4                     | 0.45 | 0.5 | 0.55 | 0.6 | 0.65 | 0.7  | 0.75 | 0.8  | 0.85 | 0.9  | 0.95 | 1    |
| $2 \times 10^6$   | 42%  | -41%                    | -24% | 0%  | 35%  | 86% | 156% | 252% | 376% | 399% | 396% | 397% | 399% | 397% |
| $2.5 \times 10^6$ | 67%  | -40%                    | -24% | 0%  | 37%  | 88% | 162% | 261% | 394% | 495% | 496% | 496% | 498% | 495% |
| $3 \times 10^6$   | 64%  | -40%                    | -23% | 0%  | 35%  | 90% | 165% | 269% | 409% | 585% | 584% | 584% | 587% | 584% |

Table 4: Results for Setting II with fixed smoothness  $\beta = 0.5$  and different values of horizon length  $T$ . *Above*: Average cumulative regret divided by  $10^3$ ; *Below*: Relative Loss.

## 6.2 Different payoff functions

In addition, we consider a different Setting III with random payoffs:

$$f_k(x) = \begin{cases} \frac{1+L_1(\frac{1}{2})^\beta - L_1x^\beta}{2} & \text{if } 0 \leq x \leq \frac{1}{2}; \\ \frac{1}{2} + g(2x - 1) & \text{if } \frac{1}{2} < x \leq 1, \end{cases} \quad \forall k \in \mathcal{K},$$

where  $g$  is a random fractional Brownian motion (Mandelbrot and Van Ness 1968) with parameter  $H = \beta$  (which is guaranteed to generate a  $\beta$ -Hölder function almost surely), or a random Brownian Bridge (which is guaranteed to generate a function that is  $\beta$ -Hölder for any  $\beta < \frac{1}{2}$  almost surely), and scaled to the range  $[-\frac{1}{2}, \frac{1}{2}]$ . We note that the first case in the definition of the above payoffs is to make sure payoffs are self-similar.

Table 5 (Table 6) provides the average cumulative regret and Relative Loss when payoffs are random fractional Brownian motions with  $H = 0.5$  (Brownian bridge) for a fixed horizon length  $T = 2 \times 10^6$ . We note that one could draw the same high-level conclusions that we made based on the results of Settings I and II, also through the results of this new setting. However, since the first two settings are designed to demonstrate the cost of smoothness under-estimation for ABSE, they tend to result in a larger gap between the performance of ABSE( $\tilde{\beta}$ ) and ABSE( $\beta$ ). In addition, since randomly generated functions do not behave as abruptly as in Setting II, one needs less exploration to identify the sub-optimal arm in each region, and hence, over-estimation of smoothness, in the range  $\tilde{\beta} < 1$  that we have considered, results in better performance.

| $T$             | ABSE( $\beta$ ) | SACB  | ABSE( $\tilde{\beta}$ ) |       |       |       |      |      |      |      |      |      |
|-----------------|-----------------|-------|-------------------------|-------|-------|-------|------|------|------|------|------|------|
|                 |                 |       | 0.1                     | 0.2   | 0.3   | 0.4   | 0.5  | 0.6  | 0.7  | 0.8  | 0.9  | 1    |
| $2 \times 10^6$ | 8.19            | 14.39 | 19.69                   | 18.39 | 18.25 | 14.58 | 8.19 | 4.97 | 3.04 | 2.02 | 1.33 | 0.95 |

  

| $T$             | SACB | ABSE( $\tilde{\beta}$ ) |      |      |     |     |      |      |      |      |      |
|-----------------|------|-------------------------|------|------|-----|-----|------|------|------|------|------|
|                 |      | 0.1                     | 0.2  | 0.3  | 0.4 | 0.5 | 0.6  | 0.7  | 0.8  | 0.9  | 1    |
| $2 \times 10^6$ | 75%  | 140%                    | 124% | 122% | 77% | 0%  | -39% | -62% | -75% | -83% | -88% |

Table 5: Results when payoffs are random fractional Brownian motion with  $H = 0.5$  *Above*: Average cumulative regret divided by  $10^4$ ; *Below*: Relative Loss.

| $T$             | ABSE( $\beta$ ) | SACB | ABSE( $\tilde{\beta}$ ) |      |      |      |      |      |      |      |      |      |
|-----------------|-----------------|------|-------------------------|------|------|------|------|------|------|------|------|------|
|                 |                 |      | 0.1                     | 0.2  | 0.3  | 0.4  | 0.5  | 0.6  | 0.7  | 0.8  | 0.9  | 1    |
| $2 \times 10^6$ | 5.59            | 5.60 | 6.02                    | 5.66 | 5.62 | 5.60 | 5.59 | 5.24 | 4.26 | 3.54 | 2.67 | 2.22 |

  

| $T$             | SACB  | ABSE( $\tilde{\beta}$ ) |       |       |       |      |        |         |        |         |         |
|-----------------|-------|-------------------------|-------|-------|-------|------|--------|---------|--------|---------|---------|
|                 |       | 0.1                     | 0.2   | 0.3   | 0.4   | 0.5  | 0.6    | 0.7     | 0.8    | 0.9     | 1       |
| $2 \times 10^6$ | 0.12% | 7.74%                   | 1.26% | 0.54% | 0.18% | 0.0% | -6.25% | -23.79% | -36.7% | -52.27% | -60.16% |

Table 6: Results when payoffs are random Brownian bridges. *Above*: Average cumulative regret divided by  $10^4$ ; *Below*: Relative Loss.

## 7 Adaptivity to smoothness with full-feedback

In this section, we consider a setting where at every time step, the agent observes the reward of both arms as opposed to the reward of the selected arm. We refer to this setting as the full-feedback setting. We establish that even with full feedback the optimal regret rate is characterized by (2.1):

**Theorem 7.1.** *Fix some Hölder exponent  $0 < \beta \leq 1$ , some margin parameter  $0 < \alpha \leq \max\{1, \frac{1}{\beta}\}$  and covariate dimension  $d$ . Then, in the full-feedback setting, for any horizon length  $T \geq 1$  and admissible policy  $\pi$ , the worst-case regret is lower bounded as follows*

$$\sup_{\mathbf{P} \in \mathcal{P}(\beta, \alpha, d)} \mathcal{R}^\pi(\mathbf{P}; T) \geq CT^{\eta(\beta, \alpha, d)},$$

where  $\eta(\beta, \alpha, d)$  is given in (2.1) and  $C$  is independent of  $T$ .

*Proof.* The following proof adopts the lower bound proof in Rigollet and Zeevi (2010) and consists of the following steps:

**Step 1 (Preliminaries).** Recall the definition of inferior sampling rate  $\mathcal{S}^\pi(\mathbf{P}; T)$  in (1.1). By Lemma 1.1, it would suffice to show that there exists a problem instances  $\mathbf{P} \in \mathcal{P}(\beta, \alpha, d)$  such that

$$\mathcal{S}^\pi(\mathbf{P}; T) \geq C'T^{1 - \frac{\beta\alpha}{2\beta+d}},$$

for some  $C'$  independent of  $T$ . Fix a uniform covariate distribution  $\mathbf{P}_X$ . For any policy  $\pi$  and function  $f : [0, 1]^d \rightarrow [0, 1]$ , denote by  $\mathcal{S}^\pi(f; T)$  the inferior sampling rate of  $\pi$  when  $\mathbf{P}_X$  is the covariate distribution,  $\mathbb{E}[Y_{1,t} | X_t] = f(X_t)$ , and  $\mathbb{E}[Y_{2,t} | X_t] = \frac{1}{2}$ . Notably, the oracle policy  $\pi_f^*$  is given by  $\pi_f^*(x) = 2 - \mathbb{1}\{f(x) \geq \frac{1}{2}\}$ . We further denote by  $\mathbb{P}_{\pi,f}$  and  $\mathbb{E}_{\pi,f}$  the corresponding probability and expectation.

**Step 2 (Problem construction).** Let  $\Delta = T^{-\frac{\beta}{2\beta+d}}$ ,  $M = \mu\Delta^{\alpha-\frac{d}{\beta}}$  for some constant  $\mu > 0$ , and

$$\psi(x) := \begin{cases} |1 - \|x\|_\infty|^\beta & \text{if } 0 \leq \|x\|_\infty \leq 1; \\ 0 & \text{o.w.} \end{cases}$$

Note that  $\psi \in \mathcal{H}_{\mathbb{R}^d}(\beta, 1)$ . Define the hypercube  $H_0 := [0, \Delta^{\frac{\alpha}{d}}]^d$ , and let grid  $G$  partition this hypercube into  $M$  disjoint hypercubes  $(H_m)_{m \in \{1, \dots, M\}}$  of equal side-length. Let  $a_m \in \mathbb{R}^d$ ,  $m \in \{1, \dots, M\}$ , be the center of the hypercube  $H_m$ . Define the function

$$\phi_m(x) := \Delta \cdot \psi_\gamma\left(\Delta^{-\frac{1}{\beta}}[x - a_m]\right).$$

By Lemmas 1.2,  $\phi_m \in \mathcal{H}(\beta, L)$  since  $C_\phi \leq L$ . For a given vector  $\boldsymbol{\omega} \in \{-1, 1\}^M$  with elements  $\omega_m$ , define the first arm's payoff function as follows:

$$f^\boldsymbol{\omega}(x) := \frac{1}{2} + \sum_{m=1}^M C_\phi \omega_m \phi_m(x).$$

**Step 3 (Desirable event).** For  $m \in \{1, \dots, M\}$ , define  $Q_m := \sum_{t=1}^T \mathbb{1}\{X_t \in H_m\} =: \sum_{t=1}^T Z_{m,t}$  to be the number of times periods at which the realized covariates belong to the hypercube  $H_m$ . Define  $\mathcal{A} := \left\{ \exists m \in \{1, \dots, M\} : Q_m < \frac{1}{2}T\Delta^{\frac{d}{\beta}} \text{ or } Q_m > \frac{3}{2}T\Delta^{\frac{d}{\beta}} \right\}$  to be the event where  $Q_m$  is less than  $\frac{1}{2}T\Delta^{\frac{d}{\beta}}$  or larger than  $\frac{3}{2}T\Delta^{\frac{d}{\beta}}$  for at least one value of  $m \in \{1, \dots, M\}$ . Note that

$$\mathbb{P}\{\mathcal{A}\} \leq \sum_{m=1}^M \mathbb{P}\left\{Q_m < \frac{1}{2}T\Delta^{\frac{d}{\beta}}\right\} + \mathbb{P}\left\{Q_m > \frac{3}{2}T\Delta^{\frac{d}{\beta}}\right\}.$$

In order to bound each of the summands on the right hand side of the above inequality, one may apply Bernstein's inequality in Lemma 1.4 to  $Q_m$ : Note that since  $\Delta^{\frac{d}{\beta}}T \leq \mathbb{E}Z_{m,t} \leq T\Delta^{\frac{d}{\beta}}$ ,  $|Z_{m,t}| \leq 1$ , and  $\text{Var}Z_{m,t} \leq \mathbb{E}Z_{m,t}^2 \leq 2\bar{\rho}\Delta^{\frac{d}{\beta}}$ , one obtains:

$$\begin{aligned}\mathbb{P}\{\mathcal{A}\} &\stackrel{(a)}{\leq} M \exp\left(-c_1 T \Delta^{\frac{d}{\beta}}/5\right) \\ &\leq \mu T^{\frac{\alpha\beta-d}{2\beta+d}} \exp\left(-c_2 T^{\frac{2\beta}{2\beta+d}}\right) \leq c_3 T^{-3},\end{aligned}$$

for large enough  $T$  and constants  $c_1, c_2, c_3 > 0$ , where (a) follows from the definition of  $M$  and  $\Delta$ . For any problem instance  $\mathbf{P}$  and horizon length  $T$ , denote the inferior sampling rate of  $\pi$  when the event  $\mathcal{A}$  does not occur by

$$\bar{\mathcal{S}}^\pi(\mathbf{P}; T) := \mathbb{E}^\pi \left[ \sum_{t=1}^T \mathbb{1}\{f_{\pi_t^*}(X_t) \neq f_{\pi_t}(X_t)\} \mid \bar{\mathcal{A}} \right].$$

Note that

$$(1 - \mathbb{P}\{\mathcal{A}\}) \bar{\mathcal{S}}^\pi(\mathbf{P}; T) \leq \mathcal{S}^\pi(\mathbf{P}; T) \leq \bar{\mathcal{S}}^\pi(\mathbf{P}; T) + T \mathbb{P}\{\mathcal{A}\},$$

which implies that

$$|\bar{\mathcal{S}}^\pi(\mathbf{P}; T) - \mathcal{S}^\pi(\mathbf{P}; T)| \leq c_4 T^{-2}. \quad (7.1)$$

For the rest of the proof, all probabilities and expectations will be computed conditional on  $\bar{\mathcal{A}}$ .

**Step 4 (Lower bounding information sampling rate).** One has

$$\begin{aligned}\sup_{\omega \in \{-1,1\}^M} \bar{\mathcal{S}}^\pi(f^\omega; T) &= \sup_{\omega \in \{-1,1\}^M} \sum_{t=1}^T \mathbb{E}_{\pi, f^\omega} [\mathbb{1}\{2\pi_t(X_t) \neq 3 - \text{sign}(f^\omega(X_t))\} \mid \bar{\mathcal{A}}] \\ &= \sup_{\omega \in \{-1,1\}^M} \sum_{m=1}^M \sum_{t=1}^T \mathbb{E}_{\pi, f^\omega} [\mathbb{1}\{2\pi_t(X_t) \neq 3 - \omega_m, X_t \in H_m\} \mid \bar{\mathcal{A}}] \\ &\geq \frac{1}{2^M} \sum_{m=1}^M \sum_{t=1}^T \sum_{\omega \in \{-1,1\}^M} \mathbb{E}_{\pi, f^\omega} [\mathbb{1}\{2\pi_t(X_t) \neq 3 - \omega_m, X_t \in H_m\} \mid \bar{\mathcal{A}}] \quad (7.2)\end{aligned}$$

Now observe that the summation  $\sum_{\omega \in \{-1,1\}^M} [\dots]$  can be decomposed as

$$R_{m,t} := \sum_{\omega_{[-m]} \in \{-1,1\}^{M-1}} \sum_{i \in \{-1,1\}} \mathbb{E}_{\pi, f^{\omega_{[-m]}^i}} [\mathbb{1}\{2\pi_t(X_t) \neq 3 - i, X_t \in H_m\} \mid \bar{\mathcal{A}}],$$

where  $\omega_{[-m]} = (\omega_1, \dots, \omega_{m-1}, \omega_{m+1}, \dots, \omega_M)$  and  $\omega_{[-m]}^i = (\omega_1, \dots, \omega_{m-1}, i, \omega_{m+1}, \dots, \omega_M)$  for  $i \in \{-1, 1\}$ . Using Lemma 1.5, and denoting by  $\bar{P}_X^m$  the conditional probability  $\mathbb{P}\{\cdot \mid \bar{\mathcal{A}}, X_t \in H_m\}$ , one

has

$$\begin{aligned}
\sum_{i \in \{-1, 1\}} \mathbb{E}_{\pi, f, \omega_{[-m]}^i} [\mathbb{1} \{2\pi_t(X_t) \neq 3 - i, X_t \in H_m\} \mid \bar{\mathcal{A}}] &= \Delta^{\frac{d}{\beta}} \sum_{i \in \{-1, 1\}} \mathbb{E}_{\pi, f, \omega_{[-m]}^i} [\bar{P}_X^m \{2\pi_t(X_t) \neq 3 - i\} \mid \bar{\mathcal{A}}] \\
&\geq \Delta^{\frac{d}{\beta}} \exp \left[ \text{KL} \left( \bar{\mathbb{P}}_{\pi, f, \omega_{[-m]}^{-1}}^t \times \bar{P}_X^m, \bar{\mathbb{P}}_{\pi, f, \omega_{[-m]}^1}^t \times \bar{P}_X^m \right) \right] \\
&= \Delta^{\frac{d}{\beta}} \exp \left[ \text{KL} \left( \bar{\mathbb{P}}_{\pi, f, \omega_{[-m]}^{-1}}^t, \bar{\mathbb{P}}_{\pi, f, \omega_{[-m]}^1}^t \right) \right], \quad (7.3)
\end{aligned}$$

where the probability measures  $\bar{\mathbb{P}}_{\pi, f, \omega_{[-m]}^i}$ , for  $i \in \{-1, 1\}$  are conditional on the event  $\bar{\mathcal{A}}$ . For any  $t = 2, \dots, T$ , let  $\mathcal{F}_t$  denote the  $\sigma$ -algebra generated by the information available at time  $t$  immediately after observing  $X_t$ , i.e.,  $\mathcal{F}_t = \sigma \left( X_t, (\pi_s, X_s, Y_{\pi_s, s})_{s=1, \dots, t-1} \right)$ . Define the conditional distribution  $\bar{\mathbb{P}}_{\pi, f}^{\cdot | \mathcal{F}_t}$  of the random couple  $(X_t, Y_{\pi_t, t})$  conditioned on  $\mathcal{F}_t$  and the event  $\bar{\mathcal{A}}$ . Applying the chain rule for KL divergence, we find that for any  $t = 1, \dots, T$  and any functions  $f, g : \mathcal{X} \rightarrow [0, 1]$ , we have

$$\begin{aligned}
\text{KL} \left( \bar{\mathbb{P}}_{\pi, f}^t, \bar{\mathbb{P}}_{\pi, g}^t \right) &= \text{KL} \left( \bar{\mathbb{P}}_{\pi, f}^{t-1}, \bar{\mathbb{P}}_{\pi, g}^{t-1} \right) + \bar{\mathbb{E}}_{\pi, f}^{t-1} \left[ \text{KL} \left( \bar{\mathbb{P}}_{\pi, f}^{\cdot | \mathcal{F}_t}, \bar{\mathbb{P}}_{\pi, g}^{\cdot | \mathcal{F}_t} \right) \right] \\
&= \text{KL} \left( \bar{\mathbb{P}}_{\pi, f}^{t-1}, \bar{\mathbb{P}}_{\pi, g}^{t-1} \right) + \bar{\mathbb{E}}_{\pi, f}^{t-1} \left[ \text{KL} \left( \bar{\mathbb{P}}_{\pi, f}^{Y_{\pi_t, t} | \mathcal{F}_t}, \bar{\mathbb{P}}_{\pi, g}^{Y_{\pi_t, t} | \mathcal{F}_t} \right) \right],
\end{aligned}$$

where  $\bar{\mathbb{P}}_{\pi, f}^{Y_{\pi_t, t} | \mathcal{F}_t}$  denotes the conditional distribution of  $Y_{\pi_t, t}$  given  $\mathcal{F}_t$  and  $\bar{\mathcal{A}}$ . Since for any  $\omega$  we have  $\mathbb{E}[Y_{\pi_t, t}] \in (\frac{1}{2} - \tau, \frac{1}{2} + \tau)$  for some  $\tau \in (0, \frac{1}{2})$ , we can apply Lemma 4.1 in Rigollet and Zeevi (2010) to derive the following bound:

$$\begin{aligned}
\left[ \text{KL} \left( \bar{\mathbb{P}}_{\pi, f, \omega_{[-m]}^{-1}}^{Y_{\pi_t, t} | \mathcal{F}_t}, \bar{\mathbb{P}}_{\pi, f, \omega_{[-m]}^1}^{Y_{\pi_t, t} | \mathcal{F}_t} \right) \right] &\leq \frac{1}{\kappa^2} \left( f^{\omega_{[-m]}^{-1}} - f^{\omega_{[-m]}^1} \right)^2 \mathbb{1} \{ \pi_t(X_t) = 1, X_t \in H_m \} \\
&\leq \frac{4C_\phi^2 \Delta^2}{\kappa^2} \left( f^{\omega_{[-m]}^{-1}} - f^{\omega_{[-m]}^1} \right)^2 \mathbb{1} \{ X_t \in H_m \}.
\end{aligned}$$

By induction, the last two displays yield that for any  $t = 1, \dots, T$ ,

$$\text{KL} \left( \bar{\mathbb{P}}_{\pi, f, \omega_{[-m]}^{-1}}^t, \bar{\mathbb{P}}_{\pi, f, \omega_{[-m]}^1}^t \right) \leq \frac{4C_\phi^2 \Delta^2}{\kappa^2} \sum_{t=1}^T \mathbb{1} \{ X_t \in H_m \} \stackrel{(a)}{\leq} \frac{6\bar{\rho}C_\phi^2 \Delta^2 T \Delta^{\frac{d}{\beta}}}{\kappa^2} \stackrel{(b)}{=} \frac{6\bar{\rho}C_\phi^2}{\kappa^2},$$

where (a) follows since we assume event  $\bar{\mathcal{A}}$  holds and (b) follows from the definition of  $\Delta$ . Combining the above inequality and (7.3) one has

$$R_{m,t} \geq \frac{2^{M-1} \Delta^{\frac{d}{\beta}}}{4} \exp \left( -\frac{6\bar{\rho}C_\phi^2}{\kappa^2} \right).$$

This inequality along with (7.2), results in the desired lower bound for the inferior sampling rate:

$$\sup_{\omega \in \{-1,1\}^M} \bar{S}^\pi(f^\omega; T) \geq \frac{2^{M-1} M T \Delta^{\frac{d}{\beta}}}{4 \cdot 2^M} \exp\left(-\frac{6\bar{\rho} C_\phi^2}{\kappa^2}\right) \geq C' T^{1-\frac{\beta\alpha}{2\beta+d}},$$

for some constant  $C'$  independent of  $T$ . This inequality along with (7.1) gives the desired result for large enough  $T$ . This concludes the proof.  $\blacksquare$

We next detail a meta policy (see Algorithm 1) that is smoothness-adaptive under the full-feedback assumption. This policy integrates some collection of non-adaptive policies  $\{\pi_0(\beta_0)\}_{\beta_0 \in [\underline{\beta}, \bar{\beta}]}$  that are rate-optimal under accurate tuning of the smoothness parameter.

The key idea of this policy is to consider a collection of smoothness parameters  $\beta_j \leftarrow \underline{\beta} + j\delta\beta_T$  over interval  $[\underline{\beta}, \bar{\beta}]$  and to initialize a separate policy  $\pi_0(\beta_j)$  for each smoothness parameter. Our meta policy keeps track of the regret that each of these policies have incurred up to each time step  $t$  (which can be computed due to the full-feedback assumption) and selects the arm suggested by the policy that has incurred minimum regret so far. Let  $j^* := \max\{j : \beta_j \leq \beta\}$  be the largest  $j$  for which  $\beta_j$  is smaller than the true smoothness parameter  $\beta$ . We know that policy  $\pi_0(\beta_{j^*})$  will incur a regret that is near-optimal. As a result, by following the policy that has incurred minimum regret up to each time step, we will make sure that the cumulative regret of our meta policy will not exceed the regret of  $\pi_0(\beta_{j^*})$  up to a multiplicative factor which depends on the number of policies  $\pi_0(\beta_j)$ . We next formalize this policy.

---

**Algorithm 1:** Full-Feedback Adaptive Bandits

---

- 1 **Input:** Set of non-adaptive policies  $\{\pi_0(\beta_0)\}_{\beta_0 \in [\underline{\beta}, \bar{\beta}]}$ , horizon length  $T$ , minimum and maximum smoothness exponents  $\underline{\beta}$  and  $\bar{\beta}$ , a tuning parameter  $\gamma$ .
  - 2 **Initialize:**  $\delta\beta_T \leftarrow \frac{(2\bar{\beta}+d)^2}{(\bar{\beta}+d-1)\log T}$ ,  $\beta_j \leftarrow \underline{\beta} + j\delta\beta_T$  and  $R_{j,1} \leftarrow 0$  for  $j \in \left\{0, 1, \dots, \left\lfloor \frac{(\bar{\beta}+d-1)(\bar{\beta}-\underline{\beta})\log T}{(2\bar{\beta}+d)^2} \right\rfloor\right\}$
  - 3 **for**  $t = 1, \dots$  **do**
  - 4     Determine the policy with minimum regret so far:  $j_t \leftarrow \arg \min R_{j,t}$
  - 5     Pull the arm suggested by the policy  $\pi_0(\beta_{j_t})$ :  $\pi_t \leftarrow \pi_0(\beta_{j_t}, X_t)$
  - 6     Observe feedback  $Y_{k,t}$  for all  $k \in \mathcal{K}$
  - 7     Advance the policy  $\pi_0(\beta_{j_t})$  by feeding observation  $Y_{\pi_t,t}$  into it
  - 8     Update the regret of the policy  $\pi_0(\beta_{j_t})$ :  $R_{j_t,t} \leftarrow R_{j_t,t-1} + \mathbb{1}\{j_t = j\} \cdot (\max_{k \in \mathcal{K}} Y_{k,t} - Y_{\pi_t,t})$
- 

The next theorem establishes that, when coupled with appropriate off-the-shelf non-adaptive policies, the above meta policy guarantees optimal regret rate up to poly-logarithmic terms, and smoothness-adaptive performance as stated in Definition 2.4.

**Theorem 7.2** (Smoothness-adaptive policy under full-feedback). *Let  $\pi$  be the SACB policy detailed in Algorithm 1, and let  $\{\pi_0(\beta_0)\}_{\beta_0 \in [\underline{\beta}, \bar{\beta}]}$  be a set of non-adaptive policies such that if initialized with the true*

smoothness parameter, for any  $\underline{\beta} \leq \beta_0 \leq \bar{\beta}$ ,  $\alpha \leq \frac{1}{\min\{1, \beta_0\}}$ , and  $T \geq 1$ , it satisfies the following upper bound on the regret:

$$\sup_{\mathbf{P} \in \mathcal{P}(\beta_0, \alpha, d)} \mathcal{R}^{\pi_0(\beta_0)}(\mathbf{P}; T) \leq \bar{C}_0 (\log T)^{\iota_0(\beta_0, \alpha, d)} T^{\zeta(\beta_0, \alpha, d)},$$

for some  $\iota_0(\beta_0, \alpha, d)$  and a constant  $\bar{C}_0 > 0$  that is independent of  $T$ , where the function  $\zeta(\beta_0, \alpha, d)$  is given in (2.1). Then, there exists  $\bar{C} > 0$ , such that for any problem instance  $\mathbf{P} \in \mathcal{P}(\beta, \alpha, d)$  with  $\underline{\beta} \leq \beta \leq \bar{\beta}$ ,  $\alpha \leq \frac{1}{\min\{1, \beta\}}$  and any horizon length  $T$ :

$$\mathcal{R}^\pi(\mathbf{P}; T) \leq \bar{C} T^{\zeta(\beta, \alpha, d)} (\log T)^{1 + \iota_0(\beta - \delta\beta, \alpha, d)}.$$

*Proof.* For each  $j$ , let  $\bar{t}_j$  be the last time step the corresponding policy has been selected. Recall the definition of  $j^* := \max\{j : \beta_j \leq \beta\}$  which is the largest  $j$  for which  $\beta_j$  is smaller than the true smoothness parameter  $\beta$ . For each  $j$ , one has

$$\mathbb{E} \left[ R_{j, \bar{t}_j} \right] \stackrel{(a)}{\leq} \mathbb{E} \left[ R_{j, \bar{t}_j - 1} \right] + 1 \stackrel{(b)}{\leq} \mathbb{E} \left[ R_{j^*, \bar{t}_j - 1} \right] + 1 \leq \mathbb{E} \left[ R_{j^*, \bar{t}_{j^*} - 1} \right] + 1 \leq \bar{C}_0 (\log T)^{\iota_0(\beta_{j^*}, \alpha, d)} T^{\zeta(\beta_{j^*}, \alpha, d)} + 1.$$

Noting  $\beta_{j^*} \geq \beta - \delta\beta_T$  and that  $\mathcal{R}^\pi(\mathbf{P}; T) = \sum_j \mathbb{E} \left[ R_{j, \bar{t}_j} \right]$ , the result follows.  $\blacksquare$

## References

- Audibert, J.-Y. and A. B. Tsybakov (2007). Fast learning rates for plug-in classifiers. *The Annals of statistics* 35(2), 608–633.
- Hu, Y., N. Kallus, and X. Mao (2019). Smooth contextual bandits: Bridging the parametric and non-differentiable regret regimes. *arXiv preprint arXiv:1909.02553*.
- Kaufmann, E., O. Cappé, and A. Garivier (2016). On the complexity of best-arm identification in multi-armed bandit models. *The Journal of Machine Learning Research* 17(1), 1–42.
- Locatelli, A. and A. Carpentier (2018). Adaptivity to smoothness in x-armed bandits. In *Conference on Learning Theory*, pp. 1463–1492.
- Mandelbrot, B. B. and J. W. Van Ness (1968). Fractional brownian motions, fractional noises and applications. *SIAM review* 10(4), 422–437.
- Perchet, V. and P. Rigollet (2013). The multi-armed bandit problem with covariates. *The Annals of Statistics* 41(2), 693–721.
- Rigollet, P. and A. Zeevi (2010). Nonparametric bandits with covariates. In *Conference on Learning Theory*, pp. 54–66.
- Tsybakov, A. B. (2008). *Introduction to Nonparametric Estimation* (1st ed.). Springer Publishing Company, Incorporated.