Supplementary Material for Regularized Contextual Bandits

A PROOFS OF SLOW RATES

We prove in this section the propositions and theorem of Subsection 4.1.

We begin by a lemma on the concentration of T_b, the number of context samples falling in a bin b.

Lemma 3. *For all $b \in \mathcal{B}$, let T_b the number of context samples falling in the bin b. We have*

$$\mathbb{P}\left(\exists b \in \mathcal{B},\ \left|T_b - \frac{T}{B^d}\right| \geq \frac{1}{2}\frac{T}{B^d}\right) \leq 2B^d \exp\left(-\frac{T}{12B^d}\right).$$

Proof. For a bin $b \in \mathcal{B}$ and $t \in \{1, \dots, T\}$, let $Z_t^{(b)} = \mathbb{1}_{\{X_t \in \mathcal{B}\}}$ which is a random Bernoulli variable of parameter $1/B^d$.

We have $T_b = \sum_{t=1}^T Z_t$ and $\mathbb{E}[T_b] = T/B^d$.

Using a multiplicative Chernoff's bound (Vershynin, 2018) we obtain:

$$\mathbb{P}\left(|T_b - \mathbb{E}[T_b]| \geq \frac{1}{2}\mathbb{E}[T_b]\right) \leq 2\exp\left(-\frac{1}{3}\left(\frac{1}{2}\right)^2 \frac{T}{B^d}\right) = 2\exp\left(-\frac{T}{12B^d}\right).$$

We conclude with an union bound on all the bins. □

Proof of Proposition 1. We have

$$E(p_T) = \mathbb{E}L(p_T) - L(\tilde{p}^\star) = \frac{1}{B^d}\sum_{b\in\mathcal{B}} \mathbb{E}L_b(p_T(b)) - L_b(p_b^\star)$$

Let us now consider a single bin $b \in \mathcal{B}$. We have run the UCB Frank-Wolfe (Berthet and Perchet, 2017) algorithm for the function L_b on the bin b with T_b iterations.

For all $p \in \Delta^K$, $L_b(p) = \langle \bar{\mu}(b), p\rangle + \lambda\rho(p)$, then for all $p \in \Delta^K$, $\nabla L_b(p) = \bar{\mu}(b) + \lambda\nabla\rho(p)$ and $\nabla^2 L_b(p) = \lambda\nabla^2\rho(p)$. Since ρ is a S-smooth convex function, L_b is a λS-smooth convex function.

We consider the event A:

$$A \doteq \left\{\forall b \in \mathcal{B},\ T_b \in \left[\frac{T}{2B^d}, \frac{3T}{2B^d}\right]\right\}.$$

Lemma 3 shows that $\mathbb{P}(A^\complement) \leq 2B^d \exp\left(-\frac{T}{12B^d}\right)$.

Theorem 3 of Berthet and Perchet (2017) shows that, on event A:

$$\begin{aligned}\mathbb{E}L_b(p_T(b)) - L_b(p_b^\star) &\leq 4\sqrt{\frac{3K\log(T_b)}{T_b}} + \frac{S\log(eT_b)}{T_b} + \left(\frac{\pi^2}{6} + K\right)\frac{2\|\nabla L_b\|_\infty + \|L_b\|_\infty}{T_b} \\ &\leq 4\sqrt{\frac{6K\log(T)}{T/B^d}} + \frac{2S\log(eT)}{T/B^d} + 2\left(\frac{\pi^2}{6} + K\right)\frac{2\|\nabla L_b\|_\infty + \|L_b\|_\infty}{T/B^d}.\end{aligned}$$

Since ρ is of class $\mathcal{C}^1$, ρ and $\nabla\rho$ are bounded on the compact set Δ^K. It is also the case for L_b and consequently $\|L_b\|_\infty$ and $\|\nabla L_b\|_\infty$ exist and are finite and can be expressed in function of $\|\rho\|_\infty$, $\|\nabla\rho\|_\infty$ and $\|\lambda\|_\infty$. On event $A^\complement$, $\mathbb{E}L_b(p_T(b)) - L_b(p_b^\star) \leq 2\|L_b\|_\infty \leq 2 + 2\|\lambda\rho\|_\infty$.

Summing over all the bins in $\mathcal{B}$ we obtain:

$$\mathbb{E}L(p_T)-L(p^\star) \le 4B^{d/2}\sqrt{\frac{6K\log(T)}{T}}+B^d\frac{2S\log(eT)}{T}+4KB^d\frac{4+2\left\|\lambda\nabla\rho\right\|_\infty+\left\|\lambda\rho\right\|_\infty}{T}+4B^d(1+\left\|\lambda\rho\right\|_\infty)e^{-\frac{T}{12B^d}}. \tag{4}$$

The first term of Equation (4) dominates the others and we can therefore write that

$$\mathbb{E}L(p_T)-L(p^\star)=\mathcal{O}\left(\sqrt{K}B^{d/2}\sqrt{\frac{\log(T)}{T}}\right)$$

where the $\mathcal{O}$ is valid for $T\to\infty$.

□

Proof of Proposition 2. We consider a bin $b\in\mathcal{B}$ containing t samples.

Let $\mathcal{S}\doteq\left\{p\in\Delta^K \mid \forall i\in[K],\ p_i\ge\frac{\lambda}{\sqrt{t}}\right\}$. In order to force all the successive estimations of $p_b^\star$ to be in $\mathcal{S}$ we sample each arm $\lambda\sqrt{t}$ times. Thus we have $\forall i\in[K],\ p_i\ge\lambda/\sqrt{t}$. Then we apply the UCB-Frank Wolfe algorithm on the bin b. Let

$$\hat{p}_b\doteq\min_{p\in\mathcal{S}}L_b(p)\quad\text{and}\quad p_b^\star\doteq\min_{p\in\Delta^K}L_b(p).$$

- **Case 1:** $\hat{p}_b=p_b^\star$, *i.e.* the minimum of L_b is in $\mathcal{S}$.

  For all $p\in\Delta^K$, $L_b(p)=\langle\bar{\mu}(b),p\rangle+\lambda\rho(p)$, then for all $p\in\Delta^K$, $\nabla L_b(p)=\bar{\mu}(b)+\lambda(1+\log(p))$ and $\nabla^2_{ii}L_b(p)=\lambda/p_i$. Therefore on $\mathcal{S}$ we have

  $$\nabla^2_{ii}L_b(p)\le\sqrt{t}.$$

  And consequently L_b is $\sqrt{t}$-smooth. And since $\nabla_i L_b(p)=1+\lambda\log(p_i)$, $\left\|\nabla L_b(p)\right\|_\infty\lesssim\log(t)$. We can apply the same steps as in the proof of Proposition 1 to find that

  $$\mathbb{E}L_b(p_t(b))-L_b(p_b^\star)\le4\sqrt{\frac{3K\log(t)}{t}}+\frac{\sqrt{t}\log(et)}{t}+\left(\frac{\pi^2}{6}+K\right)\frac{2\log(t)+\log(K)}{t}=\mathcal{O}\left(\frac{\log(t)}{\sqrt{t}}\right).$$

- **Case 2:** $\hat{p}_b\ne p_b^\star$. By strong convexity of L_b, $\hat{p}_b$ cannot be a local minimum of L_b and therefore $\hat{p}_b\in\partial\Delta^K$. The Case 1 shows that

  $$\mathbb{E}L_b(p_t(b))-L_b(\hat{p}_b)\le\mathcal{O}\left(\frac{\log(t)}{\sqrt{t}}\right).$$

  Let $\pi=(\pi_1,\dots,\pi_K)$ with $\pi_i\doteq\max(\lambda/\sqrt{t},\hat{p}_{b,i})$. We have $\left\|\pi-\hat{p}_b\right\|_2\le\sqrt{K}\lambda/\sqrt{t}$.

  Let us derive an explicit formula for $p_b^\star$ knowing the explicit expression of ρ. In order to find the optimal $\rho^\star$ value let us minimize $(p\mapsto L_b(p))$ under the constraint that p lies in the simplex Δ^K. The KKT equations give the existence of $\xi\in\mathbb{R}$ such that for each $i\in[K]$, $\bar{\mu}_i(b)+\lambda\log(p_i)+\lambda+\xi=0$ which leads to $p_{b,i}^\star=e^{-\bar{\mu}_i(b)/\lambda}/Z$ where Z is a normalization factor. Since $Z=\sum_{i=1}^K e^{-\bar{\mu}_i(b)/\lambda}$ we have $Z\le K$ and $p_{b,i}^\star\ge e^{-1/\lambda}/K$. Consequently for all p on the segment between π and $p_b^\star$ we have $p_i\ge e^{-1/\lambda}/K$ and therefore $\lambda(1+\log(p_i))\ge\lambda(1-\log K)-1$ and finally $|\nabla_i L_b(p)|\le4\left\|\lambda\right\|_\infty\log(K)$.

  Therefore L_b is $4\sqrt{K}\log(K)$-Lipschitz and

  $$\left\|L_b(p_b^\star)-L_b(\pi)\right\|_2\le4\left\|\lambda\right\|_\infty\sqrt{K}\log(K)\left\|\pi-\hat{p}_b\right\|_2\le4K\log(K)\left\|\lambda\right\|_\infty^2/\sqrt{t}=\mathcal{O}(1/\sqrt{t}).$$

  Finally, since $L_b(\pi)\ge L_b(\hat{p}_b)$ (because $\pi\in\mathcal{S}$), we have

  $$\mathbb{E}L_b(p_t(b))-L_b(p_b^\star)\le\mathbb{E}L_b(p_t(b))-L_b(\hat{p}_b)+L_b(\hat{p}_b)-L_b(p_b^\star)\le\mathcal{O}\left(\frac{\log(t)}{\sqrt{t}}\right)+L(\pi)-L(p_b^\star)=\mathcal{O}\left(\frac{\log(t)}{\sqrt{t}}\right).$$

  We conclude by summing on the bins and using that $t\in[T/2B^d,3T/2B^d]$ with high probability, as in the proof of Proposition 1.

□

Proof of Proposition 3. We have to bound the quantity

$$L(\tilde{p}^\star) - L(p^\star) = \lambda \sum_{b \in \mathcal{B}} \int_b \rho^*(-\mu(x)/\lambda) - \rho^*(-\bar{\mu}(b)/\lambda)\, \mathrm{d}x.$$

Classical results on convex conjugates (Hiriart-Urruty and Lemaréchal, 2013a) give that $\nabla \rho^*(y) = \operatorname{argmin}_{x \in \Delta^K} \rho(x) - \langle x, y \rangle$ for all $y \in \mathbb{R}^K$. Consequently, $\nabla \rho^*(y) \in \Delta^K$ and for all $y \in \mathbb{R}^K$, $\|\nabla \rho^*(y)\| \leq 1$ showing that ρ^* is 1-Lipschitz continuous. This leads to

$$\begin{aligned} L(\tilde{p}^\star) - L(p^\star) &\leq \lambda \sum_{b \in \mathcal{B}} \int_b \left\| \frac{\mu(x) - \bar{\mu}(b)}{\lambda} \right\| \mathrm{d}x \\ &\leq \sum_{b \in \mathcal{B}} \int_b \sqrt{L_\beta K} \left(\frac{\sqrt{d}}{B} \right)^\beta \mathrm{d}x \\ &\leq \sqrt{L_\beta K d^\beta} B^{-\beta} \end{aligned}$$

because all the μ_k are (L_β, β)-Hölder. □

Proof of Theorem 1. We will denote by C_k with increasing values of k the constants. Since the regret is the sum of the approximation error and the estimation error we obtain

$$R(T) \leq \sqrt{L_\beta d^\beta K} B^{-\beta} + C_1 \sqrt{K} B^{d/2} \sqrt{\frac{\log(T)}{T}} + B^d \frac{2S \log(eT)}{T} + C_2 K \frac{B^d}{T} + 4B^d (1 + \|\lambda \rho\|_\infty) \exp\left(-\frac{T}{12 B^d} \right).$$

With the choice of

$$B = \left(C_2 \beta \sqrt{L_\beta} d^{\beta/2 - 1} \right)^{1/(\beta + d/2)} \left(\frac{T}{\log(T)} \right)^{1/(2\beta + d)},$$

we find that the three last terms of the regret are negligible with respect to the first two. This gives

$$R(T) \leq \mathcal{O}\left(\left(3\sqrt{K} L_\beta^{d/(4\beta + 2d)} d^{\beta(4+d)/(4\beta + 2d)} (C_2 \beta)^{-\beta/(2\beta + d)} \right) \left(\frac{T}{\log(T)} \right)^{-\beta/(2\beta + d)} \right).$$

□

B PROOFS OF FAST RATES

We prove now the propositions and theorem of Subsection 4.2.

Proof of Proposition 4. The proof is very similar to the one of Proposition 1. We decompose the estimation error on the bins:

$$\mathbb{E} L(p_T) - L(\tilde{p}^\star) = \frac{1}{B^d} \sum_{b \in \mathcal{B}} \mathbb{E} L_b(p_T(b)) - L_b(p_b^\star).$$

Let us now consider a single bin $b \in \mathcal{B}$. We have run the UCB Frank-Wolfe algorithm for the function L_b on the bin b with T_b samples.

As in the proof of Proposition 1 we consider the event A.

Theorem 7 of Berthet and Perchet (2017), applied to L_b which is a λS-smooth $\lambda\zeta$-strongly convex function, shows that on event A:

$$\mathbb{E} L(p_T) - L(p^\star) \leq 2\tilde{c}_1 \frac{\log^2(T)}{T/B^d} + 2\tilde{c}_2 \frac{\log(T)}{T/B^d} + \tilde{c}_3 \frac{2}{T/B^d}$$

with $\tilde{c}_1 = \dfrac{96K}{\zeta\lambda\eta^2}$, $\tilde{c}_2 = \dfrac{24}{\zeta\lambda\eta^3} + \lambda S$ and $\tilde{c}_3 = 24\left(\dfrac{20}{\zeta\lambda\eta^2}\right)^2 K + \dfrac{\lambda\zeta\eta^2}{2} + \lambda S$. Consequently

$$\mathbb{E}L(p_T) - L(p^\star) \leq 2\tilde{c}_1 \frac{\log^2(T)}{T/B^d} + 2\tilde{c}_2 \frac{\log(T)}{T/B^d} + \tilde{c}_3 \frac{2}{T/B^d} + 4B^d(1 + \|\lambda\rho\|_\infty)\exp\left(-\frac{T}{12B^d}\right).$$

In order to have a simpler expression we can use the fact that λ and η are constants that can be small while S can be large. Consequently $\tilde{c}_3$ is the largest constant among $\tilde{c}_1$, $\tilde{c}_2$ and $\tilde{c}_3$ and we obtain

$$\mathbb{E}L(p_T) - L(p^\star) \leq \mathcal{O}\left(\left(\frac{K}{\lambda^2\zeta^2\eta^4} + S\lambda\right) B^d \frac{\log^2(T)}{T}\right),$$

because the other terms are negligible. □

Proof of Lemma 1. We consider a single bin $b \in \mathcal{B}$. Let us consider the function

$$\hat{L}_b : p \mapsto L_b(\alpha p^o + (1-\alpha)p).$$

Since for all i, $p^\star_{b,i} \geq \alpha p^o_i$ and since Δ^K is convex we know that $\min_{p\in\Delta^K} \hat{L}_b(p) = L_b(p^\star_b)$.

If p is the frequency vector obtained by running the UCB-Frank Wolfe algorithm for function $\hat{L}_b$ with $(1-\alpha)T$ samples then minimizing $\hat{L}_b$ is equivalent to minimizing L with a presampling stage.

Consequently the whole analysis on the regret still holds with T replaced by $(1-\alpha)T$. Thus fast rates are kept with a constant factor $1/(1-\alpha) \leq 2$. □

Proof of Proposition 5. For the entropy regularization, we have

$$p^\star_{b,i} = \frac{\exp(-\bar{\mu}(b)_i/\lambda)}{\sum_{j=1}^K \exp(-\bar{\mu}(b)_j/\lambda)} \leq \frac{\exp(-1/\lambda)}{K}.$$

We apply Lemma 1 with $p^o = \left(\dfrac{1}{K}, \dots, \dfrac{1}{K}\right)$ and $\alpha = \exp(-1/\lambda)$. Consequently each arm is presampled $T\exp(-1/\lambda)/K$ times and finally we have

$$\forall i \in [K], p_i \geq \frac{\exp(-1/\lambda)}{K}.$$

Therefore we have

$$\forall i \in [K],\ \nabla_{ii}\rho(p) = \frac{1}{p_i} \leq K\exp(1/\lambda),$$

showing that ρ is $K\exp(1/\lambda)$-smooth. □

In order to prove the Proposition 6 we will need the following lemma which is a direct consequence of a result on smooth convex functions.

Lemma 4. *Let $f : \mathbb{R}^d \to \mathbb{R}$ be a convex function of class $\mathcal{C}^1$ and $L > 0$. Let $g : \mathbb{R}^d \ni x \mapsto \dfrac{L}{2}\|x\|^2 - f(x)$. Then g is convex if and only if ∇f is L-Lipschitz continuous.*

Proof. Since g is continuously differentiable we can write

$$\begin{aligned}
g \text{ convex } &\Leftrightarrow \forall x, y \in \mathbb{R}^d,\ g(y) \geq g(x) + \langle \nabla g(x), y - x\rangle \\
&\Leftrightarrow \forall x, y \in \mathbb{R}^d, \frac{L}{2}\|y\|^2 - f(y) \geq \frac{L}{2}\|x\|^2 - f(x) + \langle Lx - \nabla f(x), y - x\rangle \\
&\Leftrightarrow \forall x, y \in \mathbb{R}^d, f(y) \leq f(x) + \langle \nabla f(x), y - x\rangle + \frac{L}{2}\left(\|y\|^2 + \|x\|^2 - 2\langle x, y\rangle\right) \\
&\Leftrightarrow \forall x, y \in \mathbb{R}^d, f(y) \leq f(x) + \langle \nabla f(x), y - x\rangle + \frac{L}{2}\|x - y\|^2 \\
&\Leftrightarrow \nabla f \text{ is } L\text{-Lipschitz continuous.}
\end{aligned}$$

where the last equivalence comes from Theorem 2.1.5 of Nesterov (2013). □

Proof of Proposition 6. Since ρ is ζ-strongly convex then $\nabla\rho^*$ is $1/\zeta$-Lipschitz continuous (see for example Theorem 4.2.1 at page 82 in Hiriart-Urruty and Lemaréchal (2013b)). Since ρ^* is also convex, Lemma 4 shows that $g : x \mapsto \frac{1}{2\zeta}\left\|x\right\|^2 - \rho^*(x)$ is convex.

Let us now consider the bin b and the function $\mu = (\mu_1, \dots, \mu_k)$. Jensen's inequality gives:

$$\frac{1}{|b|}\int_b g(-\mu(x)/\lambda)\,\mathrm{d}x \geq g\left(\frac{1}{|b|}\int_b -\frac{\mu(x)}{\lambda}\,\mathrm{d}x\right).$$

This leads to

$$\begin{aligned}\int_b g(-\mu(x)/\lambda)\,\mathrm{d}x &\geq \int_b g(-\bar{\mu}(b)/\lambda)\,\mathrm{d}x\\ \int_b \frac{1}{2\zeta}\left\|-\mu(x)\right\|^2/\lambda^2 - \rho^*(-\mu(x)/\lambda)\,\mathrm{d}x &\geq \int_b \frac{1}{2\zeta}\left\|-\bar{\mu}(b)\right\|^2/\lambda^2 - \rho^*(-\bar{\mu}(b)/\lambda)\,\mathrm{d}x\\ \int_b \rho^*(-\mu(x)/\lambda) - \rho^*(-\bar{\mu}(b)/\lambda)\,\mathrm{d}x &\leq \frac{1}{2\zeta\lambda^2}\int_b \left\|\mu(x)\right\|^2 - \left\|\bar{\mu}(b)\right\|^2\,\mathrm{d}x.\end{aligned}$$

We use the fact that $\int_b \left\|\mu(x) - \bar{\mu}(b)\right\|^2\,\mathrm{d}x = \int_b \left\|\mu(x)\right\|^2 + \left\|\bar{\mu}(b)\right\|^2 - 2\langle\mu(x), \bar{\mu}(b)\rangle\,\mathrm{d}x = \int_b \left\|\mu(x)\right\|^2 + \left\|\bar{\mu}(b)\right\|^2\,\mathrm{d}x - 2\langle\bar{\mu}(b), \int_b \mu(x)\,\mathrm{d}x\rangle = \int_b \left\|\mu(x)\right\|^2 + \left\|\bar{\mu}(b)\right\|^2\,\mathrm{d}x - 2\langle\bar{\mu}(b), |b|\bar{\mu}(b)\rangle = \int_b \left\|\mu(x)^2\right\| - \left\|\bar{\mu}(b)\right\|^2\,\mathrm{d}x$ and we get finally

$$\int_b \rho^*(-\mu(x)/\lambda) - \rho^*(-\bar{\mu}(b)/\lambda)\,\mathrm{d}x \leq \frac{1}{2\zeta\lambda^2}\int_b \left\|\mu(x) - \bar{\mu}(b)\right\|^2\,\mathrm{d}x.$$

Equation (2) shows that

$$\begin{aligned}L(\tilde{p}^\star) - L(p^\star) &\leq \frac{1}{2\zeta\lambda}\sum_{b\in\mathcal{B}}\int_b \left\|\bar{\mu}(b) - \mu(x)\right\|^2\,\mathrm{d}x\\ &\leq \sum_{b\in\mathcal{B}}\int_b \frac{L_\beta K}{2\zeta\lambda}\left(\frac{\sqrt{d}}{B}\right)^{2\beta}\,\mathrm{d}x\\ &\leq \frac{L_\beta K d^\beta}{2\zeta\lambda}\left(\frac{1}{B}\right)^{2\beta}\end{aligned}$$

because each μ_k is (L_β, β)-Hölder.

□

Proof of Theorem 2. We denote again by C_k the constants. We sum the approximation and the estimation errors (given in Propositions 6 and 4) to obtain the following bound on the regret:

$$R(T) \leq C_1\frac{L_\beta K d^\beta}{\zeta\lambda}B^{-2\beta} + C_2\frac{\log^2(T)}{T}B^d\left(\frac{1}{\zeta\lambda\eta^3} + \frac{K}{\zeta^2\lambda^2\eta^4} + \lambda\zeta\eta^2 + \lambda S\right) + 4B^d(1 + \left\|\lambda\rho\right\|_\infty)\exp\left(-\frac{T}{12B^d}\right).$$

For the sake of clarity let us note $\xi_1 \doteq C_1\dfrac{L_\beta K d^\beta}{\zeta\lambda}$ and $\xi_2 \doteq C_2\left(\dfrac{1}{\zeta\lambda\eta^3} + \dfrac{K}{\zeta^2\lambda^2\eta^4} + \lambda\zeta\eta^2 + \lambda S\right)$.

We have

$$R(T) \leq \xi_1 B^{-2\beta} + \xi_2 B^d\frac{\log^2(T)}{T} + 4B^d(1 + \left\|\lambda\rho\right\|_\infty)\exp\left(-\frac{T}{12B^d}\right).$$

Taking

$$B = \left(\frac{2\xi_1\beta}{\xi_2}\right)^{1/(2\beta+d)}\left(\frac{T}{\log^2(T)}\right)^{1/(d+2\beta)},$$

we notice that the third term is negligible and we conclude that

$$R(T) \leq \mathcal{O}\left(2\xi_1\left(\frac{2\xi_1\beta}{\xi_2}\right)^{-2\beta/(2\beta+d)}\left(\frac{T}{\log^2(T)}\right)^{-2\beta/(2\beta+d)}\right).$$

□

C PROOFS OF INTERMEDIATE RATES

We begin with a lemma on convex conjugates.

Lemma 5. *Let $\lambda, \mu > 0$ and let $y \in \mathbb{R}^n$ and ρ a non-negative bounded convex function. Then*

$$(\lambda\rho)^*(y) - (\mu\rho)^*(y) \leq |\lambda - \mu| \, \|\rho\|_\infty .$$

Proof. $(\lambda\rho)^*(y) = \sup_x \langle x, y\rangle - \lambda\rho(x) = \langle x_\lambda, y\rangle - \lambda\rho(x_\lambda)$.

And $(\mu\rho)^*(y) = \sup_x \langle x, y\rangle - \mu\rho(x) = \langle x_\mu, y\rangle - \mu\rho(x_\mu) \geq \langle x_\lambda, y\rangle - \mu\rho(x_\lambda)$.

Then, $(\lambda\rho)^*(y) - (\mu\rho)^*(y) \leq \langle x_\lambda, y\rangle - \lambda\rho(x_\lambda) - (\langle x_\lambda, y\rangle - \mu\rho(x_\lambda)) = (\mu - \lambda)\rho(x_\lambda)$.

Finally $(\lambda\rho)^*(y) - (\mu\rho)^*(y) \leq |\lambda - \mu| \, \|\rho\|_\infty$. □

Proof of Proposition 7. There exists $x_0 \in b$ such that $\bar{\lambda}(b) = \lambda(x_0)$ and $x_1 \in b$ such that $\bar{\mu}(b) = \mu(x_1)$. We use Lemma 5 to derive a bound for the approximation error.

$$\begin{aligned}
&\int_b (\lambda(x)\rho)^* \left(-\mu(x)\right) - (\bar{\lambda}(b)\rho)^* \left(-\bar{\mu}(b)\right) \mathrm{d}x \\
&= \int_b (\lambda(x)\rho)^* \left(-\mu(x)\right) - (\lambda(x)\rho)^* \left(-\bar{\mu}(b)\right) \mathrm{d}x + \int_b (\lambda(x)\rho)^* \left(-\bar{\mu}(b)\right) - (\bar{\lambda}(b)\rho)^* \left(-\bar{\mu}(b)\right) \mathrm{d}x \\
&\leq \int_b \lambda(x) \left(\rho^* \left(-\frac{\mu(x)}{\lambda(x)}\right) - \rho^* \left(-\frac{\bar{\mu}(b)}{\lambda(x)}\right)\right) \mathrm{d}x + \int_b |\lambda(x) - \bar{\lambda}(b)| \, \|\rho\|_\infty \, \mathrm{d}x \\
&\leq \int_b \lambda(x) \left|\frac{\mu(x)}{\lambda(x)} - \frac{\bar{\mu}(b)}{\lambda(x)}\right| \mathrm{d}x + \|\rho\|_\infty \int_b |\lambda(x) - \lambda(x_0)| \, \mathrm{d}x \\
&\leq \int_b L_\beta |x - x_1|^\beta \, \mathrm{d}x + \|\rho\|_\infty \int_b \|\lambda'\|_\infty \, |x - x_0| \, \mathrm{d}x \\
&\leq B^{-d} \left(L_\beta d^{\beta/2} B^{-\beta} + \|\rho\|_\infty \, \|\lambda'\|_\infty \sqrt{d} B^{-1}\right) = \mathcal{O}(B^{-\beta-d}).
\end{aligned}$$

□

Proof of Proposition 8. As in the proof of Proposition 6 we consider a bin $b \in \mathcal{B}$ and the goal is to bound

$$\int_b \lambda(x)\rho^* \left(-\frac{\mu(x)}{\lambda(x)}\right) - \bar{\lambda}(b)\rho^* \left(-\frac{\bar{\mu}(b)}{\bar{\lambda}(b)}\right) \mathrm{d}x.$$

We use a similar method and we apply Jensen inequality with density $\dfrac{\lambda(x)}{|b|\bar{\lambda}(b)}$ to the function $g : x \mapsto \frac{1}{2\zeta} \|x\|^2 - \rho^*(x)$ which is convex.

$$\begin{aligned}
g\left(\int_b -\frac{\mu(x)}{\lambda(x)} \frac{\lambda(x)}{|b|\bar{\lambda}(b)} \mathrm{d}x\right) &\leq \int_b g\left(-\frac{\mu(x)}{\lambda(x)}\right) \frac{\lambda(x)}{|b|\bar{\lambda}(b)} \mathrm{d}x \\
g\left(-\frac{\bar{\mu}(b)}{\bar{\lambda}(b)}\right) &\leq \int_b g\left(-\frac{\mu(x)}{\lambda(x)}\right) \frac{\lambda(x)}{|b|\bar{\lambda}(b)} \mathrm{d}x \\
\frac{1}{2\zeta} \left\|-\frac{\bar{\mu}(b)}{\bar{\lambda}(b)}\right\|^2 - \rho^* \left(-\frac{\bar{\mu}(b)}{\bar{\lambda}(b)}\right) &\leq \frac{1}{|b|\bar{\lambda}(b)} \int_b \left[\frac{1}{2\zeta} \left\|-\frac{\mu(x)}{\lambda(x)}\right\|^2 - \rho^* \left(-\frac{\mu(x)}{\lambda(x)}\right)\right] \lambda(x) \, \mathrm{d}x \\
\int_b \lambda(x)\rho^* \left(-\frac{\mu(x)}{\lambda(x)}\right) - \bar{\lambda}(b)\rho^* \left(-\frac{\bar{\mu}(b)}{\bar{\lambda}(b)}\right) \mathrm{d}x &\leq \frac{1}{2\zeta} \int_b \frac{\|\mu(x)\|^2}{\lambda(x)} - \frac{\|\bar{\mu}(b)\|^2}{\bar{\lambda}(b)} \mathrm{d}x.
\end{aligned}$$

Consequently we have proven that

$$\int_b \lambda(x)\rho^*\left(-\frac{\mu(x)}{\lambda(x)}\right) - \bar{\lambda}(b)\rho^*\left(-\frac{\bar{\mu}(b)}{\bar{\lambda}(b)}\right) \mathrm{d}x \leq \frac{1}{2\zeta}\int_b \frac{\|\mu(x)\|^2}{\lambda(x)} - \frac{\|\bar{\mu}(b)\|^2}{\bar{\lambda}(b)} \mathrm{d}x$$
$$\leq \frac{1}{2\zeta}\sum_{k=1}^{K}\int_b \frac{\mu_k(x)^2}{\lambda(x)} - \frac{\bar{\mu}_k(b)^2}{\bar{\lambda}(b)} \mathrm{d}x.$$

Therefore we have to bound, for each k, $I = \int_b \frac{\mu_k(x)^2}{\lambda(x)} - \frac{\bar{\mu}_k(b)^2}{\bar{\lambda}(b)} \mathrm{d}x$.

Let us omit the subscript k and consider a β-Hölder function μ.

We have

$$\begin{aligned} I &= \int_b \frac{\mu(x)^2}{\lambda(x)} - \frac{\bar{\mu}(b)^2}{\bar{\lambda}(b)} \mathrm{d}x \\ &= \int_b \frac{\mu(x)^2}{\lambda(x)} - \frac{\mu(x)^2}{\bar{\lambda}(b)} + \frac{\mu(x)^2}{\bar{\lambda}(b)} - \frac{\bar{\mu}(b)^2}{\bar{\lambda}(b)} \mathrm{d}x \\ &= \underbrace{\int_b \left(\mu(x)^2 - \bar{\mu}(b)^2\right)\left(\frac{1}{\lambda(x)} - \frac{1}{\bar{\lambda}(b)}\right) \mathrm{d}x}_{I_1} + \underbrace{\int_b \bar{\mu}(b)^2\left(\frac{1}{\lambda(x)} - \frac{1}{\bar{\lambda}(b)}\right) \mathrm{d}x}_{I_2} + \underbrace{\int_b \frac{1}{\bar{\lambda}(b)}\left(\mu(x)^2 - \bar{\mu}(b)^2\right) \mathrm{d}x}_{I_3}. \end{aligned}$$

We now have to bound these three integrals.

Bounding I_1:

$$\begin{aligned} I_1 &= \int_b \left(\mu(x)^2 - \bar{\mu}(b)^2\right)\left(\frac{1}{\lambda(x)} - \frac{1}{\bar{\lambda}(b)}\right) \mathrm{d}x \\ &= \int_b \left(\mu(x) + \bar{\mu}(b)\right)\left(\mu(x) - \bar{\mu}(b)\right)\left(\frac{1}{\lambda(x)} - \frac{1}{\bar{\lambda}(b)}\right) \mathrm{d}x \\ &\leq \int_b 2|\mu(x) - \bar{\mu}(b)|\left|\frac{1}{\lambda(x)} - \frac{1}{\bar{\lambda}(b)}\right| \mathrm{d}x \\ &\leq 2L_\beta\left(\frac{\sqrt{d}}{B}\right)^\beta \int_b\left|\frac{1}{\lambda(x)} - \frac{1}{\bar{\lambda}(b)}\right| \mathrm{d}x. \end{aligned}$$

Since $1/\lambda$ is of class $\mathcal{C}^1$, Taylor-Lagrange inequality yields, using the fact that there exists $x_0 \in b$ such that $\bar{\lambda}(b) = \lambda(x_0)$:

$$\left|\frac{1}{\lambda(x)} - \frac{1}{\bar{\lambda}(b)}\right| \leq \left\|\left(\frac{1}{\lambda}\right)'\right\|_\infty |x - x_0| \leq \frac{\|\lambda'\|_\infty}{\lambda_{\min}^2}\frac{\sqrt{d}}{B}.$$

We obtain therefore

$$I_1 \leq 2L_\beta \|\lambda'\|_\infty \sqrt{d}^{\beta+1}\frac{1}{\lambda_{\min}^2}B^{-(1+\beta+d)} = \mathcal{O}\left(\frac{B^{-(1+\beta+d)}}{\lambda_{\min}^2}\right).$$

Bounding I_2:

We have

$$I_2 = \bar{\mu}(b)^2\int_b\left(\frac{1}{\lambda(x)} - \frac{1}{\bar{\lambda}(b)}\right) \mathrm{d}x \leq \int_b\left(\frac{1}{\lambda(x)} - \frac{1}{\bar{\lambda}(b)}\right) \mathrm{d}x$$

because $\int_b\left(\frac{1}{\lambda(x)} - \frac{1}{\bar{\lambda}(b)}\right) \mathrm{d}x \geq 0$ from Jensen's inequality.

Without loss of generality we can assume that the bin b is the closed cuboid $[0, 1/B]^d$. We suppose that for all $x \in b$, $\lambda(x) > 0$.

Since λ is of class $\mathcal{C}^\infty$, we have the following Taylor series expansion:

$$\lambda(x) = \lambda(0) + \sum_{i=1}^{d} \frac{\partial \lambda(0)}{\partial x_i} x_i + \frac{1}{2} \sum_{i,j} \frac{\partial^2 \lambda(0)}{\partial x_i \partial x_j} x_i x_j + o(\|x\|^2).$$

Integrating over the bin b we obtain

$$\bar{\lambda}(b) = \lambda(0) + \frac{1}{2}\frac{1}{B} \sum_{i=1}^{d} \frac{\partial \lambda(0)}{\partial x_i} + \frac{1}{8}\frac{1}{B^2} \sum_{i \neq j} \frac{\partial^2 \lambda(0)}{\partial x_i \partial x_j} + \frac{1}{6}\frac{1}{B^2} \sum_{i=1}^{d} \frac{\partial^2 \lambda(0)}{\partial x_i^2} + o\left(\frac{1}{B^2}\right).$$

Consequently

$$\begin{aligned}
\int_b \frac{\mathrm{d}x}{\bar{\lambda}(b)} &= \frac{1}{B^d \bar{\lambda}(b)} \\
&= \frac{1}{B^d \lambda(0)} \frac{1}{1 + \frac{1}{2\lambda(0)}\frac{1}{B} \sum_{i=1}^{d} \frac{\partial \lambda(0)}{\partial x_i} + \frac{1}{\lambda(0)}\frac{1}{B^2}\left(\frac{1}{8} \sum_{i \neq j} \frac{\partial^2 \lambda(0)}{\partial x_i \partial x_j} + \frac{1}{6} \sum_{i=1}^{d} \frac{\partial^2 \lambda(0)}{\partial x_i^2}\right) + o\left(\frac{1}{B^2}\right)} \\
&= \frac{1}{B^d \lambda(0)} \left(1 - \frac{1}{2\lambda(0)}\frac{1}{B} \sum_{i=1}^{d} \frac{\partial \lambda(0)}{\partial x_i} - \frac{1}{\lambda(0)}\frac{1}{B^2}\left(\frac{1}{8} \sum_{i \neq j} \frac{\partial^2 \lambda(0)}{\partial x_i \partial x_j} + \frac{1}{6} \sum_{i=1}^{d} \frac{\partial^2 \lambda(0)}{\partial x_i^2}\right)\right. \\
&\qquad \left. + \frac{1}{4\lambda(0)^2}\frac{1}{B^2}\left(\sum_{i=1}^{d} \frac{\partial \lambda(0)}{\partial x_i}\right)^2 + o\left(\frac{1}{B^2}\right)\right) \\
&= \frac{1}{B^d \lambda(0)} - \frac{1}{2\lambda(0)^2}\frac{1}{B^{d+1}} \sum_{i=1}^{d} \frac{\partial \lambda(0)}{\partial x_i} - \frac{1}{\lambda(0)^2}\frac{1}{B^{d+2}}\left(\frac{1}{8} \sum_{i \neq j} \frac{\partial^2 \lambda(0)}{\partial x_i \partial x_j} + \frac{1}{6} \sum_{i=1}^{d} \frac{\partial^2 \lambda(0)}{\partial x_i^2}\right) \\
&\qquad + \frac{1}{4\lambda(0)^3}\frac{1}{B^{d+2}}\left(\sum_{i=1}^{d} \frac{\partial \lambda(0)}{\partial x_i}\right)^2 + o\left(\frac{1}{B^2}\right).
\end{aligned}$$

Let us now compute the Taylor series development of $1/\lambda$. We have:

$$\frac{\partial}{\partial x_i}\frac{1}{\lambda(x)} = -\frac{1}{\lambda(x)^2}\frac{\partial \lambda(x)}{\partial x_i} \quad \text{and} \quad \frac{\partial^2}{\partial x_i \partial x_j}\frac{1}{\lambda(x)} = -\frac{1}{\lambda(x)^2}\frac{\partial^2 \lambda(x)}{\partial x_i \partial x_j} + \frac{2}{\lambda(x)^3}\frac{\partial \lambda(x)}{\partial x_i}\frac{\partial \lambda(x)}{\partial x_j}.$$

This lets us write

$$\begin{aligned}
\frac{1}{\lambda(x)} &= \frac{1}{\lambda(0)} - \frac{1}{\lambda(0)^2} \sum_{i=1}^{d} \frac{\partial \lambda(0)}{\partial x_i} x_i - \frac{1}{2}\frac{1}{\lambda(0)^2} \sum_{i,j} \frac{\partial^2 \lambda(0)}{\partial x_i \partial x_j} x_i x_j + \frac{1}{\lambda(0)^3} \sum_{i,j} \frac{\partial \lambda(0)}{\partial x_i}\frac{\partial \lambda(0)}{\partial x_j} x_i x_j + o(\|x\|^2) \\
\int_b \frac{\mathrm{d}x}{\lambda(x)} &= \frac{1}{\lambda(0)}\frac{1}{B^d} - \frac{1}{2\lambda(0)^2}\frac{1}{B^{d+1}} \sum_{i=1}^{d} \frac{\partial \lambda(0)}{\partial x_i} - \frac{1}{\lambda(0)^2}\frac{1}{B^{d+2}}\left(\frac{1}{8} \sum_{i \neq j} \frac{\partial^2 \lambda(0)}{\partial x_i \partial x_j} + \frac{1}{6} \sum_{i=1}^{d} \frac{\partial^2 \lambda(0)}{\partial x_i^2}\right) \\
&\qquad + \frac{1}{\lambda(0)^3}\frac{1}{B^{d+2}}\left(\frac{1}{4} \sum_{i \neq j} \frac{\partial \lambda(0)}{\partial x_i}\frac{\partial \lambda(0)}{\partial x_j} + \frac{1}{3} \sum_{i=1}^{d} \left(\frac{\partial \lambda(0)}{\partial x_i}\right)^2\right) + o\left(\frac{1}{B^{d+2}}\right).
\end{aligned}$$

And then

$$I_2 \leq \frac{1}{12}\frac{1}{\lambda(0)^3}\frac{1}{B^{d+2}}\sum_{i=1}^{d}\left(\frac{\partial\lambda(0)}{\partial x_i}\right)^2 + o\left(\frac{1}{B^{d+2}}\right).$$

Since the derivatives of λ are bounded we obtain that

$$I_2 = \mathcal{O}\left(\frac{B^{-2-d}}{\lambda_{\min}^3}\right).$$

Bounding I_3:

$$\begin{aligned} I_3 &= \int_b \frac{1}{\bar{\lambda}(b)}\left(\mu(x)^2 - \bar{\mu}(b)^2\right)\,\mathrm{d}x \\ &= \frac{1}{\bar{\lambda}(b)}\int_b \left(\mu(x) - \bar{\mu}(b)\right)^2\,\mathrm{d}x \\ &\leq \frac{1}{\lambda_{\min}} L_\beta^2 d^\beta B^{-(2\beta+d)} = \mathcal{O}\left(\frac{B^{-(2\beta+d)}}{\lambda_{\min}}\right). \end{aligned}$$

Putting this together we have $I = \mathcal{O}\left((dL_\beta^2\left\|\nabla\lambda\right\|_\infty^2)\frac{B^{-(2\beta+d)}}{\lambda_{\min}^3}\right)$. And finally

$$L(\tilde{p}^\star) - L(p^\star) \leq \mathcal{O}\left(KdL_\beta^2\left\|\nabla\lambda\right\|_\infty^2\frac{B^{-2\beta}}{\zeta\lambda_{\min}^3}\right).$$

□

Lemma 6 (Regularity of η)**.** *If η is the distance of the optimum $p^\star$ to the boundary of Δ^K as defined in Definition 5, and if the μ_k functions are all β-Hölder and λ of class $\mathcal{C}^1$, then η is β-Hölder. More precisely we have*

$$\forall x, y \in b, |\eta(x) - \eta(y)| \leq \sqrt{\frac{K}{K-1}}\frac{\|\lambda\|_\infty + \|\lambda'\|_\infty}{\zeta\lambda_{\min}(b)^2}|x-y|^\beta = \frac{C_L}{\lambda_{\min}(b)^2}|x-y|^\beta.$$

Proof. Let $x \in \mathcal{X}$. Since $\eta(x) = \mathrm{dist}(p_b^\star, \partial\Delta^K)$ we obtain

$$\eta(x) = \sqrt{\frac{K}{K-1}}\min_i p_i^\star(x).$$

And

$$\begin{aligned} p^\star(x) &= \operatorname{argmin}\langle\mu(x), p(x)\rangle + \lambda(x)\rho(p(x)) \\ &= \nabla(\lambda(x)\rho)^*(-\mu(x)) \\ &= \nabla\rho^*\left(-\frac{\mu(x)}{\lambda(x)}\right). \end{aligned}$$

Since ρ is ζ-strongly convex, $\nabla\rho^*$ is $1/\zeta$-Lipschitz continuous.

Therefore, for $x, y \in b$,

$$\begin{aligned} |p^\star(x) - p^\star(y)| &\leq \frac{1}{\zeta}\left|\frac{\mu(x)}{\lambda(x)} - \frac{\mu(y)}{\lambda(y)}\right| \\ &\leq \frac{1}{\zeta}\left|\frac{\mu(x)-\mu(y)}{\lambda(x)}\right| + \frac{1}{\zeta}|\mu(y)|\left|\frac{1}{\lambda(x)} - \frac{1}{\lambda(y)}\right| \\ &\leq \frac{1}{\zeta\lambda_{\min}(b)}|x-y|^\beta + \frac{1}{\zeta}\frac{\|\lambda'\|_\infty}{\lambda_{\min}(b)^2}|x-y| \end{aligned}$$

since all μ_k are bounded by 1 (the losses are bounded by 1). □

Proof of Lemma 2. We consider a well-behaved bin b. There exists $x_1 \in b$ such that $\lambda(x_1) \geq c_1 B^{-\beta/3}$. Since λ is $\mathcal{C}^\infty$ on $[0,1]^d$, it is in particular Lipschitz-continuous on b. And therefore

$$\forall x \in b,\ \lambda(x) \geq c_1 B^{-\beta/3} - \|\lambda'\|_\infty \operatorname{diam}(b) \geq c_1 B^{-\beta/3} - \|\lambda'\|_\infty \operatorname{diam}(b)^{\beta/3} = B^{-\beta/3}.$$

Lemma 6 shows that η is β-Hölder continuous (with constant denoted by $C_L/\lambda_{\min}^2$) and therefore we have

$$\forall x \in b,\ \eta(x) \geq c_2 B^{-\beta/3} - \frac{C_L}{\lambda_{\min}(b)^2} \operatorname{diam}(b)^\beta = B^{-\beta/3}.$$

□

Lemma 7. *If ρ is convex, η is an increasing function of λ.*

Proof. As in the proof of Proposition 2 we use the KKT conditions to find that on a bin b (without the index k for the arm):

$$\bar{\mu}(b) + \bar{\lambda}(b)\nabla\rho(p_b^\star) + \xi = 0.$$

Therefore

$$p_b^\star = (\nabla\rho)^{-1}\left(-\frac{\xi + \bar{\mu}(b)}{\bar{\lambda}(b)}\right).$$

Since ρ is convex, $\nabla\rho$ is an increasing function and its inverse as well. Consequently $p_b^\star$ is an increasing function of $\bar{\lambda}(b)$, and since $\eta(b) = \sqrt{K/(K-1)} \min_i p_{b,i}^\star$, η is also an increasing function of $\bar{\lambda}(b)$. □

Proof of Theorem 3. Since B will be chosen as an increasing function of T we only consider T sufficiently large in order to have $c_1 B^{-\beta/3} < \delta_1$ and $c_2 B^{-\beta/3} < \delta_2$. To ensure this we can also take smaller δ_1 and δ_2. Moreover we lower the value of δ_2 or δ_1 to be sure that $\frac{\delta_2}{c_2} = \eta(\frac{\delta_1}{c_1})$. These are technicalities needed to simplify the proof.

The proof will be divided into several steps. We will first obtain lower bounds on λ and η for the "well-behaved bins". Then we will derive bounds for the approximation error and the estimation error. And finally we will put that together to obtain the intermediate convergence rates.

As in the proofs on previous theorems we will denote the constants C_k with increasing values of k.

- **Lower bounds on η and λ:**

  Using a technique from Rigollet and Zeevi (2010) we notice that without loss of generality we can index the B^d bins with increasing values of $\bar{\lambda}(b)$. Let us note $\mathcal{IB} = \{1, \ldots, j_1\}$ and $\mathcal{WB} = \{j_1 + 1, \ldots, B^d\}$. Since η is an increasing function of λ (cf Lemma 7), the $\eta(b_j)$ are also increasingly ordered.

  Let $j_2 \geq j_1$ be the largest integer such that $\bar{\lambda}(b_j) \leq \dfrac{\delta_1}{c_1}$. Consequently we also have that j_2 is the largest integer such that $\eta(b_j) \leq \dfrac{\delta_2}{c_2}$.

  Let $j \in \{j_1+1, \ldots, j_2\}$. The bin b_j is a well-behaved bin and Lemma 2 shows that $\bar{\lambda}(b_j) \geq B^{-\beta/3}$. Then $\bar{\lambda}(b_j) + (c_1 - 1)B^{-\beta/3} \leq c_1\bar{\lambda}(b_j) \leq \delta_1$ and we can apply the margin condition (cf Assumption 3) which gives

  $$\mathbb{P}_X(\lambda(x) \leq \bar{\lambda}(b_j) + (c_1-1)B^{-\beta/3}) \leq C_m (c_1 \bar{\lambda}(b_j))^{6\alpha}.$$

  But since the context are uniformly distributed and since the $\bar{\lambda}(b_j)$ are increasingly ordered we also have that

  $$\mathbb{P}_X(\lambda(x) \leq \bar{\lambda}(b_j) + (c_1-1)B^{-\beta/3}) \geq \mathbb{P}_X(\lambda(x) \leq \bar{\lambda}(b_j)) \geq \frac{j}{B^d}.$$

  This gives $\bar{\lambda}(b_j) \geq \dfrac{1}{c_1 C_m^{1/6\alpha}}\left(\dfrac{j}{B^d}\right)^{1/6\alpha}$. The same computations give $\eta(b_j) \geq \dfrac{1}{c_2 C_m^{1/6\alpha}}\left(\dfrac{j}{B^d}\right)^{1/6\alpha}$. We note $C_\gamma \doteq \min((c_1 C_m^{1/6\alpha})^{-1}, (c_2 C_m^{1/6\alpha})^{-1}))$ and $\gamma_j \doteq C_\gamma \left(\dfrac{j}{B^d}\right)^{1/\alpha}$. Consequently $\bar{\lambda}(b_j) \geq \gamma_j$ and $\eta(b_j) \geq \gamma_j$.

Let us now compute the number of ill-behaved bins:

$$\begin{aligned}\#\{b\in\mathcal{B}, b\notin\mathcal{WB}\} &= B^d\ \mathbb{P}(b\notin\mathcal{WB})\\ &= B^d\ \mathbb{P}(\forall x\in\mathcal{B},\ \eta(x)\le c_2B^{-\beta/3}\text{ or }\forall x\in\mathcal{B},\ \lambda(x)\le c_1B^{-\beta/3})\\ &\le B^d\ \mathbb{P}(\eta(\bar{x})\le c_2B^{-\beta/3}\text{ or }\lambda(\bar{x})\le c_1B^{-\beta/3})\\ &\le C_m(c_1^{6\alpha}+c_2^{6\alpha})B^dB^{-2\alpha\beta}\doteq C_IB^dB^{-2\alpha\beta}\end{aligned}$$

where $\bar{x}$ is the mean context value in the bin b. Consequently if $j\ge j^\star\doteq C_IB^dB^{-2\alpha\beta}$, then $b_j\in\mathcal{WB}$. Let $\hat{j}\doteq C_IB^dB^{-\alpha\beta}\ge j^\star$. Consequently for all $j\ge j^\star$, $b_j\in\mathcal{WB}$.

We want to obtain an upper-bound on the constant $S\lambda(\bar{b}_j)+\dfrac{K}{\eta(b_j)^4\bar{\lambda}(b_j)^2}$ that arises in the fast rate for the estimation error. For the sake of clarity we will remove the dependency in b_j and denote this constant $C=S\lambda+\dfrac{K}{\lambda^2\eta^4}$.

In the case of the entropy regularization $S=1/\min_i p_i^\star$. Since $\eta=\sqrt{K/(K-1)}\min_i p_i^\star$, we have that $\min_i p_i^\star=\sqrt{(K-1)/K}\eta\ge\eta/2$. Consequently $S\le 2/\gamma_j$ and, on a well-behaved bin b_j, for $j\le j_2$,

$$C\le\frac{K+2\left\|\lambda\right\|_\infty}{\gamma_j^6}\doteq\frac{C_F}{\gamma_j^6},\tag{5}$$

where the subscript F stands for "Fast". When $j\ge j_2$, we have $\bar{\lambda}(b_j)\ge\delta_1/c_1$ and $\eta(b_j)\ge\delta_2/c_2$ and consequently

$$C\le\frac{K}{(\delta_1/c_1)^2(\delta_2/c_2)^4}+\frac{2\left\|\lambda\right\|_\infty}{\delta_2/c_2}\doteq C_{\max}.$$

Let us notice than λ being known by the agent, the agent knows the value of $\bar{\lambda}(b)$ on each bin b and can therefore order the bins. Consequently the agent can sample, on every well-behaved bin, each arm $T\gamma_j/2$ times and be sure that $\min_i p_i\ge\gamma_j/2$. On the first $\lfloor\hat{j}\rfloor$ bins the agent will sample each arm $\bar{\lambda}(b)\sqrt{T/B^d}$ times as in the proof of Proposition 2.

- **Approximation Error:**

  We now bound the approximation error. We separate the bins into two sets: $\{1,\dots,\lfloor j^\star\rfloor\}$ and $\{\lceil j^\star\rceil,\dots,B^d\}$. On the first set we use the slow rates of Proposition 7 and on the second set we use the fast rates of Proposition 8.

  We obtain that, for $\alpha<1/2$,

$$\begin{aligned}L(\tilde{p}^\star)-L(p^\star) &\le L_\beta d^{\beta/2}\sum_{j=1}^{\lfloor j^\star\rfloor}B^{-\beta-d}+\left\|\rho\right\|_\infty\left\|\nabla\lambda\right\|_\infty\sqrt{d}\sum_{j=1}^{\lfloor j^\star\rfloor}B^{-1-d}+(KdL_\beta^2\left\|\nabla\lambda\right\|_\infty^2)\sum_{j=\lceil j^\star\rceil}^{B^d}\frac{B^{-2\beta-d}}{\bar{\lambda}(b_j)^3}\\ &\le C_IL_\beta d^{\beta/2}B^{-\beta}B^{-2\alpha\beta}+(KdL_\beta^2\left\|\nabla\lambda\right\|_\infty^2)\left(\sum_{j=\lceil j^\star\rceil}^{j_2}\frac{B^{-2\beta-d}}{\gamma_j^3}+\sum_{j=j_2+1}^{B^d}\frac{B^{-2\beta-d}}{(c_1/\delta_1)^3}\right)+o(B^{-2\alpha\beta-\beta})\\ &\le C_IL_\beta d^{\beta/2}B^{-2\alpha\beta-\beta}+(KdL_\beta^2\left\|\nabla\lambda\right\|_\infty^2)\left(\frac{B^{-2\beta-d}}{C_\gamma^3}\sum_{j=\lceil j^\star\rceil}^{j_2}\left(\frac{j}{B^d}\right)^{-1/2\alpha}+B^{-2\beta}\left(\frac{\delta_1}{c_1}\right)^3\right)+o(B^{-2\alpha\beta-\beta})\\ &\le C_IL_\beta d^{\beta/2}B^{-2\alpha\beta-\beta}+(KdL_\beta^2\left\|\nabla\lambda\right\|_\infty^2)\frac{1}{C_\gamma^3}B^{-2\beta}\int_{C_IB^{-2\alpha\beta}}^1x^{-1/2\alpha}\,\mathrm{d}x+o(B^{-2\alpha\beta-\beta})\\ &\le\left(C_IL_\beta d^{\beta/2}+KdL_\beta^2\left\|\nabla\lambda\right\|_\infty^2\frac{2\alpha}{1-2\alpha}\frac{C_I^{(2\alpha-1)/2\alpha}}{C_\gamma^3}\right)B^{-\beta-2\alpha\beta}+o(B^{-2\alpha\beta-\beta})=\mathcal{O}\left(B^{-\beta-2\alpha\beta}\right)\end{aligned}$$

  since $\alpha<1/2$. We step from line 3 to 4 thanks to a series-integral comparison.

For $\alpha = 1/2$ we get

$$L(\tilde{p}^\star) - L(p^\star) \leq \left(C_I L_\beta d^{\beta/2} + \left(KdL_\beta^2 \left\|\nabla\lambda\right\|_\infty^2\right)(\delta_1^3 c_1^{-3} + 2\beta C_\gamma^{-3}\log(B))\right) B^{-2\beta} + o(B^{-2\beta}) = \mathcal{O}\left(B^{-2\beta}\log(B)\right).$$

And for $\alpha > 1/2$ we have

$$L(\tilde{p}^\star) - L(p^\star) \leq \left(KdL_\beta^2 \left\|\nabla\lambda\right\|_\infty^2\right)\left(\frac{1}{C_\gamma^3}\frac{2\alpha}{2\alpha-1} + \left(\frac{\delta_1}{c_1}\right)^3\right) B^{-2\beta} + o(B^{-2\beta}) = \mathcal{O}\left(B^{-2\beta}\right)$$

because $\beta + 2\alpha\beta > 2\beta$.

Let us note

$$\begin{aligned}
\xi_1 &\doteq \left(C_I L_\beta d^{\beta/2} + KdL_\beta^2 \left\|\nabla\lambda\right\|_\infty^2 \frac{2\alpha}{1-2\alpha}\frac{C_I^{(2\alpha-1)/2\alpha}}{C_\gamma^3}\right);\\
\xi_2 &\doteq \left(C_I L_\beta d^{\beta/2} + \left(KdL_\beta^2 \left\|\nabla\lambda\right\|_\infty^2\right)(\delta_1^3 c_1^{-3} + 2\beta C_\gamma^{-3}\log(B))\right);\\
\xi_3 &\doteq \left(KdL_\beta^2 \left\|\nabla\lambda\right\|_\infty^2\right)\left(\frac{1}{C_\gamma^3}\frac{2\alpha}{2\alpha-1} + \left(\frac{\delta_1}{c_1}\right)^3\right);\\
\xi_{app} &\doteq \max(\xi_1, \xi_2, \xi_3).
\end{aligned}$$

Finally we obtain that the approximation error is bounded by $\xi_{app} B^{-\min(\beta+2\alpha\beta, 2\beta)}\log(B)$ with $\alpha > 0$.

- **Estimation Error:**

  We proceed in a similar manner as for the approximation error, except that we do not split the bins around $j^\star$ but around $\hat{j}$.

  In a similar manner to the proofs of Theorems 1 and 2 we only need to consider the terms of dominating order from Propositions 1 and 4. As before we consider the same event A (cf the proof of Proposition 1) and we note $C_A \doteq 4B^d(1 + \left\|\lambda\rho\right\|_\infty)$. We obtain, for $\alpha < 1$, using (5):

$$\begin{aligned}
\mathbb{E}L(\tilde{p}_T) - L(\tilde{p}^\star) &= \frac{1}{B^d}\sum_{b\in\mathcal{B}} \mathbb{E}L_b(\tilde{p}_T) - L(p_b^\star)\\
&= \frac{1}{B^d}\sum_{j=\lceil\hat{j}\rceil}^{B^d} \mathbb{E}L_b(\tilde{p}_T) - L(p_b^\star) + \frac{1}{B^d}\sum_{j=1}^{\lfloor\hat{j}\rfloor} \mathbb{E}L_b(\tilde{p}_T) - L(p_b^\star)\\
&\leq \frac{1}{B^d}\sum_{j=\lceil\hat{j}\rceil}^{B^d} 2C\frac{\log^2(T)}{T/B^d} + \frac{1}{B^d}\sum_{j=1}^{\lfloor\hat{j}\rfloor} 4\sqrt{12K}\sqrt{\frac{\log(T)}{T/B^d}} + C_A e^{-\frac{T}{12B^d}}\\
&\leq 2C_F\sum_{j=\lceil\hat{j}\rceil}^{j_2} \frac{\log^2(T)}{T}\gamma_j^{-6} + \sum_{j=j_2+1}^{B^d} 2C_{\max}\frac{\log^2(T)}{T} + 6\sqrt{3K}\sqrt{\frac{\log(T)}{T}}B^{d/2}B^{-\alpha\beta} + C_A e^{-\frac{T}{12B^d}}\\
&\leq \frac{2C_F}{C_\gamma^6}\frac{\log^2(T)}{T}\sum_{j=\lceil\hat{j}\rceil}^{j_2}\left(\frac{j}{B^d}\right)^{-1/\alpha} + 2C_{\max}\frac{\log^2(T)}{T}B^d + 6\sqrt{3K}\sqrt{\frac{\log(T)}{T}}B^{d/2-\alpha\beta} + C_A e^{-\frac{T}{12B^d}}\\
&\leq \frac{2C_F}{C_\gamma^6}\frac{\log^2(T)}{T}B^d\int_{C_I B^{-\alpha\beta}}^1 x^{-1/\alpha}\,\mathrm{d}x + 2C_{\max}\frac{\log^2(T)}{T}B^d + 6\sqrt{3K}\sqrt{\frac{\log(T)}{T}}B^{d/2-\alpha\beta} + C_A e^{-\frac{T}{12B^d}}\\
&\leq \frac{2C_F}{C_\gamma^6}\frac{\log^2(T)}{T}B^d\frac{\alpha}{1-\alpha}B^{\beta(1-\alpha)} + 2C_{\max}\frac{\log^2(T)}{T}B^d + 6\sqrt{3K}\sqrt{\frac{\log(T)}{T}}B^{d/2-\alpha\beta} + C_A e^{-\frac{T}{12B^d}}\\
&\leq \frac{2C_F}{C_\gamma^6}\frac{\log^2(T)}{T}\frac{\alpha}{1-\alpha}B^{d+\beta-\alpha\beta} + 6\sqrt{3K}\sqrt{\frac{\log(T)}{T}}B^{d/2-\alpha\beta} + 2C_{\max}\frac{\log^2(T)}{T}B^d + C_A e^{-\frac{T}{12B^d}}.
\end{aligned}$$

- **Putting things together:**

  We note $C_\alpha \doteq \dfrac{2C_F}{C_\gamma^6}\dfrac{\alpha}{1-\alpha}$. This leads to the following bound on the regret:

$$R(T) \leq C_\alpha \frac{\log^2(T)}{T} B^{d+\beta-\alpha\beta} + 6\sqrt{3K}\sqrt{\frac{\log(T)}{T}} B^{d/2-\alpha\beta} + 2C_{\max}\frac{\log^2(T)}{T}B^d + C_A e^{-\frac{T}{12B^d}} + \xi_{app} B^{-\min(2\beta,\beta+2\alpha\beta)}\log(B).$$

  Choosing $B = \left(\dfrac{T}{\log^2(T)}\right)^{1/(2\beta+d)}$ we get

$$R(T) \leq (C_\alpha + 6\sqrt{3}K)\left(\frac{T}{\log^2(T)}\right)^{-\beta(1+\alpha)/(2\beta+d)} + o\left(\frac{T}{\log^2(T)}\right)^{-\beta(1+\alpha)/(2\beta+d)}$$

  which is valid for $\alpha \in (0,1)$.

  Finally we have

$$R(T) = \mathcal{O}\left(\left(\frac{T}{\log^2(T)}\right)^{-\beta(1+\alpha)/(2\beta+d)}\right).$$

□

D PROOFS OF LOWER BOUNDS

Proof of Theorem 4. We consider the model with $K = 2$ where $\mu(x) = (-\eta(x), \eta(x))^\top$, where η is a β-Hölder function on $\mathcal{X} = [0,1]^d$. We note that η is uniformly bounded over $\mathcal{X}$ as a consequence of smoothness, so one can take λ such that $|\eta(x)| < \lambda$. We denote by $e = (1/2, 1/2)$ the center of the simplex, and we consider the loss

$$L(p) = \int_{\mathcal{X}} \left(\langle \mu(x), p(x)\rangle + \lambda \|p(x) - e\|^2\right) \mathrm{d}x.$$

Denoting by $p_0(x)$ the vector $e + \mu(x)/(2\lambda)$, we have that $p_0(x) \in \Delta^2$ for all $x \in \mathcal{X}$. Further, we have that

$$\langle \mu(x), p(x)\rangle + \lambda\|p(x) - e\|^2 = \lambda\|p(x) - p_0(x)\|^2 + 1/(4\lambda)\|\mu(x)\|^2\,,$$

since $\langle \mu(x), e\rangle = 0$. As a consequence, L is minimized at p_0 and

$$L(p) - L(p_0) = \int_{\mathcal{X}} \lambda \|p(x) - p_0(x)\|^2\,\mathrm{d}x = 1/(2\lambda)\int_{\mathcal{X}} |\eta(x) - \eta_0(x)|^2\,\mathrm{d}x\,.$$

where η is such that $p(x) = \big(1/2 - \eta(x)/(2\lambda), 1/2 + \eta(x)/(2\lambda)\big)$. As a consequence, for any algorithm with final variable $\hat{p}_T$, we can construct an estimator $\hat{\eta}_T$ such that

$$\mathbb{E}[L(\hat{p}_T)] - L(p_0) = 1/(2\lambda)\mathbb{E}\int_{\mathcal{X}} |\hat{\eta}_T(x) - \eta_0(x)|^2\,\mathrm{d}x\,,$$

where the expectation is taken over the randomness of the observations Y_t, with expectation $\pm\eta(X_t)$, with sign depending on the known choice $\pi_t = 1$ or 2. As a consequence, any upper bound on the regret for a policy implies an upper bound on regression over β-Hölder functions in dimension d, with T observations. This yields that, in the special case where ρ is the 1-strongly convex function equal to the squared ℓ_2 norm

$$\inf_{\hat{p}} \sup_{\substack{\mu \in \mathcal{H}_\beta \\ \rho = \ell_2^2}} \mathbb{E}[L(\hat{p}_T)] - L(p_0) \geq \inf_{\hat{\eta}} \sup_{\eta \in \mathcal{H}_\beta} 1/(2\lambda)\mathbb{E}\int_{\mathcal{X}} |\hat{\eta}_T(x) - \eta_0(x)|^2\,\mathrm{d}x \geq CT^{-\frac{2\beta}{2\beta+d}}\,.$$

The final bound is a direct application of Theorem 3.2 in Györfi et al. (2006). □