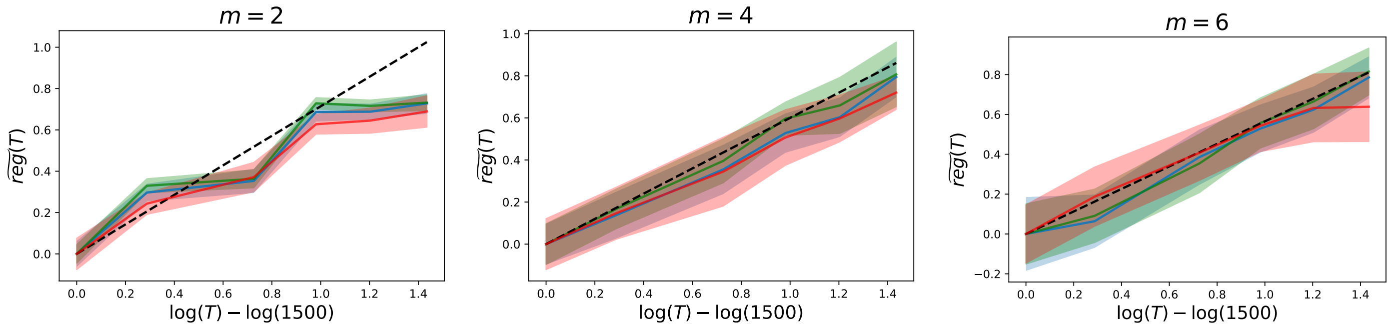
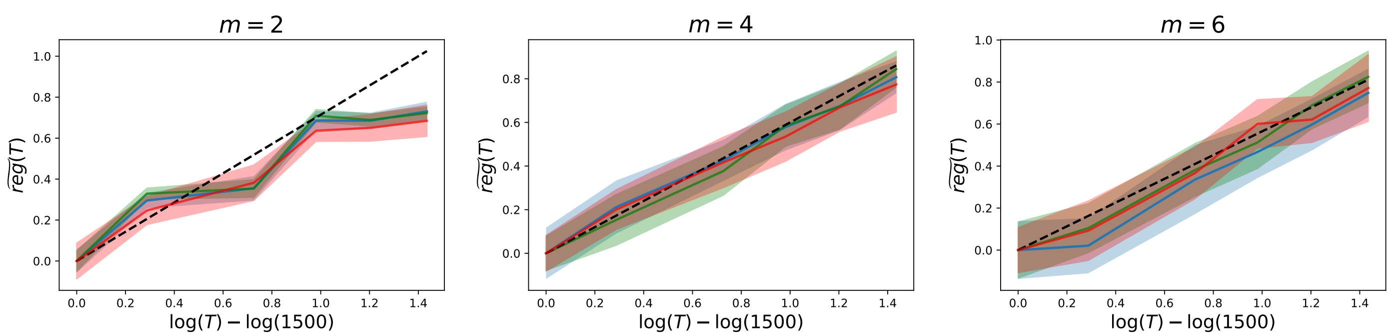
Supplement material for “ Policy Optimization Using Semi-parametric Models for Dynamic Pricing”

A Addition Figures



(a) (b) (c)

Figure 7: Regret log-log plot in the setting with i.i.d. covariates with dependent entries. The remaining caption is the same as Figure 1.



(a) (b) (c)

Figure 8: Regret log-log plot in the setting with strong mixing covariates. The remaining caption is the same as Figure 1.

B Discussion

1. [Minimax Lower Bound] Our work shares a similar setting with Broder and Rusmevichientong [2012], in which they study a general choice model with parametric structure and binary

response, but without any covariates. A lower bound of order $\Omega(\sqrt{T})$ is established by constructing an 'uninformative price' in their work. To be more precise, an uninformative price is a price that all demand curves (probability of successful sales) as offered price indexed by unknown parameters intersect. Namely, the demands at this uninformative price are the same for all unknown parameters. In addition, such price is also the optimal price with some parameters. In this case, the price is uninformative because it doesn't reveal any information on the true parameter. Intuitively, if one tries to learn model parameters, the only way is to offer prices that are sufficiently far from the uninformative price (optimal price) which leads to a larger regret.

Borrowing the idea from Broder and Rusmevichientong [2012] and Javanmard and Nazerzadeh [2019], we deduce that there exists an 'uninformative price' in the following class of models: Consider a class of distributions $\mathcal{F}$ which satisfies Assumption 1:

$$\mathcal{F} := \{F_\sigma : \sigma > 0, F_\sigma = F(x/\sigma)\}.$$

Here, F is the c.d.f. of a known distribution with mean zero. Moreover, we assume the support of F'_σ is contained in $[-a, a]$ (For instance, the class of distributions with density $f_\sigma(x) = 4/(3\sigma^3)(\sigma - x)^k(\sigma + x)^k \cdot \mathbb{I}_{\{|x|\leqslant\sigma\}}, k \geqslant 1$ or $f_\sigma(x) = C_\sigma \exp\left(-\frac{\sigma^2}{\sigma^2 - x^2}\right) \cdot \mathbb{I}_{\{|x|\leqslant\sigma\}}$ with $\sigma \leqslant a$ etc.)

Let $\beta = 1/\sigma$ and multiply β on both sides of (2), which leads to

$$\widetilde{v}(\mathbf{x}_t) = \widetilde{\boldsymbol{\beta}}_0^\top \mathbf{x}_t + \widetilde{\alpha}_0 + \widetilde{z}_t.$$

Here, $\widetilde{v}_t = \beta v_t, \widetilde{\boldsymbol{\beta}}_0 = \beta\boldsymbol{\beta}_0, \widetilde{\alpha}_0 = \beta\alpha_0$ and $\widetilde{z}_t = \beta z_t$. The distribution of $\widetilde{z}_t$ is F_1, which is denoted as F here for convenience. Next, in our sub-parameter class, we first let $\boldsymbol{\beta}_0 = 0$ and fix a number ξ with $F'(\xi) \neq 0$. Then we choose a collection of $\{(\sigma, \alpha_0)\}$ which satisfies $\beta = 1/\sigma = (\xi + \widetilde{\alpha}_0)$. Following the same arguments as in Javanmard and Nazerzadeh [2019], one can prove that $p = 1$ is indeed an uninformative price. Since in the sub-parametric class given above, all demand curves intersect at a point $1 - F(\xi)$ when $p = 1$, and for a special $(\sigma, \alpha_0) = (1/(\xi - \phi(\xi)), -\phi(\xi)/(\xi - \phi(\xi))$, $p = 1$ is the optimal price. Thus the $\Omega(\sqrt{T})$ lower bound applies.

Remark 1. When we only consider explore-then-commit algorithms and offer price as $p_t = \widehat{\phi}_k^{-1}(-\mathbf{x}_t^\top\widehat{\theta}) + \mathbf{x}_t^\top\widehat{\theta}$, with $\widehat{\phi}_k(u) = u - \frac{1-\widehat{F}_k(u)}{\widehat{F}^{(1)}(u)}$, the optimality of p_t reduces to the optimality of estimating $F(\cdot), f(\cdot)$ and $\boldsymbol{\theta}$. According to Stone [1980, 1982], Tsybakov [2008], the statistical rates of our estimators on $\widehat{F}$, $\widehat{F}^{(1)}$ and $\widehat{\boldsymbol{\theta}}$ are minimax optimal in every episode. Thus, our posted price is optimal constrained on this type of policies. However, if we consider a general policy class, there is currently no lower bound for feature-based pricing given unknown noise distribution with finite smoothness degree besides the general $\sqrt{T}$ lower bound mentioned above. It remains an open problem whether our upper bound is tight for finite m.

2. [The adversarial setting] We note that in some real applications with potentially adversarial contexts, the covariance of the feature vectors might be singular or ill-conditioned (e.g. due to repeated buyers recorded in $\mathbf{x}_t$). However, our algorithm can be adjusted to cope with such situations. The key observation here is that this assumption is *only* required in our exploration phase: For any k, we allow arbitrary $\mathbf{x}_t$ in the k-th exploitation phase, since we have already obtained accurate estimators $\widehat{\boldsymbol{\theta}}_k$ and $\widehat{g}_k(\cdot)$ for $\boldsymbol{\theta}_0$ and $g(\cdot)$. Therefore, whenever there is a sign of a repeated buyer, we can modify our algorithm slightly by using the $\widehat{g}_{k-1}(\cdot)$ in the last episode to offer a price, and then move this buyer to the corresponding exploitation phase. If the number of similar buyers in the k-th episode is ℓ_k^r with any $r < 1$ and we assume the remaining buyers are sampled i.i.d. from a distribution, we are still able to proceed by only arranging some contexts with similar buyers into the exploitation phase directly. This matches with some real situation in online shopping where personal preference features will be recorded by the seller in order to make recommendation in the future.

3. [Online inference of the demand] Recently, Wang et al. [2020] use a de-biased approach to quantify the uncertainty of the demand function in a parametric class which offers new insight to the field of statistical decision making.
In our work, we combine the non-parametric statistical estimation and online decision making to derive a policy that maximize the seller's revenue. We next also briefly discuss our intuition on depicting the uncertainty of the demand curve in a non-parametric class. Recall the demand curve given in (4). For given $p, \mathbf{x}$, and estimators $\widehat{F}_k, \widehat{\boldsymbol{\theta}}_k$, in the k-th exploitation phase, deriving asymptotic behavior of the demand curve reduces to deriving the asymptotic behavior of our estimator on $\widehat{F}_k(\cdot)$. This is due to the statistical rate of $\widehat{F}_k(\cdot)$ dominates that of $\widehat{\boldsymbol{\theta}}_k$. According to asymptotic behavior of the kernel regression [Fan and Gijbels, 1996, Carroll et al., 1997, Fan et al., 1998], we have the following pointwise confidence interval for $\widehat{F}$:

$$\sqrt{|I_k|h_k}(\widehat{F}_k(u) - F(u) - h_k^m \kappa_m B(u)) \to N\Big(0, \int K^2(x)\mathrm{d}x\sigma^2(u)/f(u)\Big),$$

where $f(\cdot)$ is the density of $p_t - \mathbf{x}_t^\top \theta_0$ with $p_t \sim \mathrm{Unif}(0, B)$ and we recall that $|I_k|$ is the length of our k-th exploration phase. In addition, $\kappa_m = \int K(x)x^m \mathrm{d}x$, $B(u) = F^{(m)}(u)f(u)/m! + F^{(m-1)}(u)f^{(1)}(u)/(m-1)! + \cdots + F^{(1)}(u)f^{(m-1)}(u)/(m-1)!$, and $\sigma^2(u) = \mathrm{Var}(y_t \,|\, p_t - \mathbf{x}_t^\top \boldsymbol{\theta}_0 = u)$. Thus, for any given $p, \mathbf{x}$, and an $\widehat{\boldsymbol{\theta}}_k$, we are able to derive the pointwise asymptotic behavior of our demand curve as follows:

$$\begin{aligned}&\sqrt{|I_k|h_k}(p\widehat{F}_k(p - \mathbf{x}^\top \widehat{\boldsymbol{\theta}}_k) - pF(p - \mathbf{x}^\top \boldsymbol{\theta}_0) - ph_k^m \kappa_m B(p - \mathbf{x}^\top \boldsymbol{\theta}_0))\\ &\to N\Big(0, p^2 \int K^2(s)\mathrm{d}s\sigma^2(p - \mathbf{x}^\top \boldsymbol{\theta}_0)/f(p - \mathbf{x}^\top \boldsymbol{\theta}_0)\Big).\end{aligned}$$

The data-driven confidence interval for our demand curve given in (4) can be established via bootstrap and the undersmoothing technique (to remove the bias), see e.g. Hall [1992], Horowitz [2001] for more details. Similarly, uniform statistical inference results can also be established by using similar non-parametric tools, see e.g. Eubank and Speckman [1993], Neumann and Polzehl [1998], Hall and Horowitz [2013] for more details. We will leave the detailed proof for future work.

4. In some situations, it might be difficult for retailers to adopt a uniform pricing strategy even during a short period of time. An alternative strategy might be the following: As in Algorithm 1, we divide the time horizon into episodes according to the doubling strategy. However, now we no longer divide an episode into explore-then-exploitation phases. Instead, at the beginning of each episode $k > 1$, we leverage all the data $\{p_t, \mathbf{x}_t, y_t\}$ collected from the previous episode to estimate $\boldsymbol{\theta}_0$ and F. Then, we compute $\widehat{g}_k$ from the estimates $\widehat{F}_k$ and $\widehat{F}_k^{(1)}$, and perform exploitation directly throughout this episode. This procedure can help us to get rid of uniform exploration in practice. We leave the theoretical guarantees for this refined algorithm as our future work.

B.1 Extension: High-dimensional Feature-based Dynamic Pricing

Algorithm 1 can be naturally extended to the high-dimensional setting, where $\boldsymbol{\theta}_0 \in \mathbb{R}^d$, d can be large compared to T, while $\|\boldsymbol{\theta}_0\|_0 \leqslant s$ for a relatively small sparsity s. This happens in applications when a large amount of covariate information is available, and the actual market value only depends on some essential factors. One way of extension is the following: at each episode, we can replace estimation of $\widehat{\boldsymbol{\theta}}_k$ in (7) with the two steps below.

Step 1. Let

$$\widetilde{\boldsymbol{\theta}}_k = \underset{\boldsymbol{\theta}}{\operatorname{argmin}}\, L_k(\boldsymbol{\theta}) + \lambda p(\boldsymbol{\theta}), \tag{1}$$

where

$$L_k(\boldsymbol{\theta}) := \frac{1}{|I_k|} \sum_{t \in I_k} (By_t - \boldsymbol{\theta}^\top \widetilde{\mathbf{x}}_t)^2, \quad p(\boldsymbol{\theta}) = \sum_{j=1}^{p} p(|\boldsymbol{\theta}^{(j)}|)$$

for some penalty function $p(\cdot)$. As in Zhao and Yu [2006], Fan and Li [2001], Zhang [2010], by choosing different $p(\cdot)$ such as in the ℓ_1, SCAD or MCP penalty, under suitable conditions such as irrepresentable condition, variable selection consistency is achieved with high probability.

Step 2. Let $\widehat{S}_k = \text{supp}(\widetilde{\boldsymbol{\theta}}_k)$, we then refit the least squares (7) on $\widehat{S}_k$:

$$\widehat{\boldsymbol{\theta}}_k = \underset{\text{supp}(\boldsymbol{\theta}) \subseteq \widehat{S}_k}{\operatorname{argmin}}\, L_k(\boldsymbol{\theta}). \tag{2}$$

Then the conclusions of Lemma 2 hold with high probability.

After Step 2, we continue the remaining steps of Algorithm 1 in the episode. In fact, if we can learn the support of $\boldsymbol{\theta}_0$, we essentially translate the problem into a low-dimensional one, and we can prove that Algorithm 1 achieves a regret upper bound of $\widetilde{\mathcal{O}}((Ts)^{\frac{4m+1}{2m-1}})$ if $F \in \mathcal{C}^{(m)}$ (or $\mathcal{O}((Ts)^{1/2})$ if F is super smooth).

C Proof under the time-independent feature setting

C.1 Proof of Lemma 1

First, recall that $R_{\mathcal{X}} := \sup_{\mathbf{x}\in\mathcal{X}} \|\mathbf{x}\|_2$, we deduce that $\mathbf{x}_t$ is also subgaussian with norm upper bounded by $\psi_x = R_{\mathcal{X}}$. This fact is useful in later proofs as well. Now according to (7), for the k-th episode, our loss function $L_k(\boldsymbol{\theta})$ is defined as

$$L_k(\boldsymbol{\theta}) = \frac{1}{|I_k|}\sum_{t\in I_k}(By_t - \boldsymbol{\theta}^\top \widetilde{\mathbf{x}}_t)^2. \tag{3}$$

For notational convenience, denote $n = |I_k|$. Then the gradient and Hessian of $L_k(\boldsymbol{\theta})$ is given by

$$\nabla_{\boldsymbol{\theta}} L_k(\boldsymbol{\theta}) = \frac{1}{n}\sum_{t\in I_k} 2(\boldsymbol{\theta}^\top \widetilde{\mathbf{x}}_t - By_t)\widetilde{\mathbf{x}}_t, \tag{4}$$

$$\nabla^2_{\boldsymbol{\theta}} L_k(\boldsymbol{\theta}) = \frac{1}{n}\sum_{t\in I_k} 2\widetilde{\mathbf{x}}_t\widetilde{\mathbf{x}}_t^\top. \tag{5}$$

Let $\widehat{\boldsymbol{\theta}}_k$ be the global minimizer of $L_k(\boldsymbol{\theta})$. We do a Taylor expansion of $L_k(\widehat{\boldsymbol{\theta}}_k)$ at $\boldsymbol{\theta}_0$:

$$L_k(\widehat{\boldsymbol{\theta}}_k) - L_k(\boldsymbol{\theta}_0) = \langle \nabla L_k(\boldsymbol{\theta}_0), \widehat{\boldsymbol{\theta}}_k - \boldsymbol{\theta}_0\rangle + \frac{1}{2}\langle \widehat{\boldsymbol{\theta}}_k - \boldsymbol{\theta}_0, \nabla^2_{\boldsymbol{\theta}} L_k(\widetilde{\boldsymbol{\theta}})(\widehat{\boldsymbol{\theta}}_k - \boldsymbol{\theta}_0)\rangle. \tag{6}$$

Here $\widetilde{\boldsymbol{\theta}}$ is a point lying between $\widehat{\boldsymbol{\theta}}_k$ and $\boldsymbol{\theta}_0$. As $\widehat{\boldsymbol{\theta}}_k$ is the global minimizer of loss (3), we have

$$\langle \nabla L_k(\boldsymbol{\theta}_0), \widehat{\boldsymbol{\theta}}_k - \boldsymbol{\theta}_0\rangle + \frac{1}{2}\langle \widehat{\boldsymbol{\theta}}_k - \boldsymbol{\theta}_0, \nabla^2_{\boldsymbol{\theta}} L_k(\widetilde{\boldsymbol{\theta}})(\widehat{\boldsymbol{\theta}}_k - \boldsymbol{\theta}_0)\rangle \leqslant 0$$

which implies

$$\langle \widehat{\boldsymbol{\theta}}_k - \boldsymbol{\theta}_0, \frac{1}{n}\sum_{t\in I_k}\widetilde{\mathbf{x}}_t\widetilde{\mathbf{x}}_t^\top(\widehat{\boldsymbol{\theta}}_k - \boldsymbol{\theta}_0)\rangle \leqslant \langle \nabla L_k(\boldsymbol{\theta}_0), \boldsymbol{\theta}_0 - \widehat{\boldsymbol{\theta}}_k\rangle \leqslant \sqrt{d}\|\nabla L_k(\boldsymbol{\theta}_0)\|_\infty \cdot \|\boldsymbol{\theta}_0 - \widehat{\boldsymbol{\theta}}_k\|_2. \tag{7}$$

In order to achieve ℓ_2-convergence rate of $\widehat{\boldsymbol{\theta}}_k$, we separate our following analysis into two steps.

Step I: In this step, we lower bound the minimum eigenvalue of

$$\boldsymbol{\Sigma}_k := \frac{1}{n}\sum_{t\in I_k}\widetilde{\mathbf{x}}_t\widetilde{\mathbf{x}}_t^\top. \tag{8}$$

using concentration inequalities.

Since $\boldsymbol{\Sigma}_k$ is an average of n i.i.d. random matrices with mean $\Sigma = \mathbb{E}[\widetilde{\mathbf{x}}_t\widetilde{\mathbf{x}}_t^\top]$ and that $\{\widetilde{\mathbf{x}}_t\}$ are sub-Gaussian random vectors, according to Remark 5.40 in Vershynin [2012], there exist c_1 and $C > c_{\min}$ such that with probability at least $1 - 2e^{-c_1 t^2}$,

$$\|\boldsymbol{\Sigma}_k - \boldsymbol{\Sigma}\| \leqslant \max\{\delta, \delta^2\}, \quad \text{where } \delta := C\sqrt{\frac{d+1}{n}} + \frac{t}{\sqrt{n}}. \tag{9}$$

Here c_1, C are both constants that are only related to sub-Gaussian norm of $\widetilde{\mathbf{x}}_t$. Now we plug in $t = c_{\min}\sqrt{n}/4$ and $c_0 = 16C^2/c_{\min}^2$, then as long as $n \geqslant c_0(d+1)$, with probability at least $1-2e^{-c_1 c_{\min}^2 n/16}$,

$$(c_{\min}/2) \cdot \mathbb{I} \preccurlyeq \mathbf{\Sigma}_k. \tag{10}$$

Step II: In this step, we provide an upper bound of $\|\nabla_{\boldsymbol{\theta}} L_k(\boldsymbol{\theta}_0)\|_\infty$.

First, we prove $\mathbb{E}[\nabla_\theta L_k(\boldsymbol{\theta}_0)] = 0$. By definition we have

$$\nabla_{\boldsymbol{\theta}} L_k(\boldsymbol{\theta}_0) = \frac{1}{n}\sum_{t\in I_k} 2(\boldsymbol{\theta}_0^\top \widetilde{\mathbf{x}}_t - By_t)\widetilde{\mathbf{x}}_t$$

We take the conditional expectation of $\nabla_{\boldsymbol{\theta}} L_k(\boldsymbol{\theta}_0)$ and obtain

$$\mathbb{E}[\nabla_{\boldsymbol{\theta}} L_k(\boldsymbol{\theta}_0) \,|\, \widetilde{\mathbf{x}}_t] = \frac{1}{n}\sum_{t\in I_k} 2\mathbb{E}[(\boldsymbol{\theta}_0^\top \widetilde{\mathbf{x}}_t - By_t) \,|\, \widetilde{\mathbf{x}}_t]\widetilde{\mathbf{x}}_t.$$

By our definition on y_t,

$$\begin{aligned}
\mathbb{E}[\boldsymbol{\theta}_0^\top \widetilde{\mathbf{x}}_t - By_t \,|\, \widetilde{\mathbf{x}}_t] &= \boldsymbol{\theta}_0^\top \widetilde{\mathbf{x}}_t - \mathbb{E}[B\mathbb{I}_{\{p_t \leqslant v_t\}} \,|\, \widetilde{\mathbf{x}}_t] \\
&= \boldsymbol{\theta}_0^\top \widetilde{\mathbf{x}}_t - \mathbb{E}[\mathbb{E}[B\mathbb{I}_{\{p_t \leqslant v_t\}} \,|\, v_t] \,|\, \widetilde{\mathbf{x}}_t] \\
&= \boldsymbol{\theta}_0^\top \widetilde{\mathbf{x}}_t - B \cdot \mathbb{E}[v_t/B \,|\, \widetilde{\mathbf{x}}_t] = 0,
\end{aligned}$$

where the third equality follows from $p_t \sim \text{Uniform}(0, B)$. After finally taking expectation with respective to $\widetilde{\mathbf{x}}_t$ we deduce that $\mathbb{E}[\nabla_{\boldsymbol{\theta}} L_k(\boldsymbol{\theta}_0)] = 0$.

Next, we get an upper bound of $\|\nabla_{\boldsymbol{\theta}} L_k(\boldsymbol{\theta})\|_\infty$. By (4), we have every entry of $\nabla_{\boldsymbol{\theta}} L_k(\boldsymbol{\theta}_0)$ is mean zero. In addition, according to our Assumption 2, we have $\mathbf{x}_t$ are i.i.d. sub-Gaussian random vectors with sub-Gaussian norm ψ_x. Thus, we have $\max_{i\in[d]} \|\mathbf{x}_{t,i}\|_{\psi_2} \leqslant \psi_x$. On the other hand, $\widetilde{\mathbf{x}}_t^\top \boldsymbol{\theta}_0 - By_t$ is bounded by the constant $R_{\mathcal{X}} R_\Theta + B$. Therefore,

$$\mathbb{P}\big(|2(\boldsymbol{\theta}_0^\top \widetilde{\mathbf{x}}_t - By_t)\widetilde{\mathbf{x}}_{t,i}| \geqslant u\big) \leqslant \mathbb{P}\big(2(R_{\mathcal{X}} R_\Theta + B)|\widetilde{\mathbf{x}}_{t,i}| \geqslant u\big) \leqslant 2\exp\Big(\frac{-u^2}{8\psi_x^2(R_{\mathcal{X}} R_\Theta + B)^2}\Big)$$

for $i \in [2:(d+1)]$, which implies that $2(\boldsymbol{\theta}_0^\top \widetilde{\mathbf{x}}_t - By_t)\widetilde{\mathbf{x}}_{t,i}$, $i \in [2:(d+1)]$ are sub-Gaussian random variables with variance proxy $2\psi_x(R_{\mathcal{X}} R_\Theta + B)$. Moreover, We can also obtain $\|2(\boldsymbol{\theta}_0^\top \widetilde{\mathbf{x}}_t - By_t)\widetilde{\mathbf{x}}_{t,1}\|_{\psi_2} \leqslant 2(R_{\mathcal{X}} R_\Theta + B)$ by Hoeffding's inequality.

We now take the union bound of all entries of $\nabla_{\boldsymbol{\theta}} L_k(\boldsymbol{\theta}_0)$:

$$\begin{aligned}
\mathbb{P}\big(\|\nabla_{\boldsymbol{\theta}} L_k(\boldsymbol{\theta}_0)\|_\infty \geqslant t\big) &\leqslant 2(d+1)\exp\Big(\frac{-t^2}{8\max\{\psi_x^2, 1\}(R_{\mathcal{X}} R_\Theta + B)^2}\Big) && (11) \\
&= 2\exp\Big(\frac{-nt^2}{8\max\{\psi_x^2, 1\}(R_{\mathcal{X}} R_\Theta + B)^2} + \log(d+1)\Big). && (12)
\end{aligned}$$

As we assume $n \geqslant d+1$, by taking $t = 4\max\{\psi_x, 1\}(R_{\mathcal{X}}R_\Theta + B)\sqrt{\log n/n}$ in (12), then with probability $1 - 2/n$, we have

$$\|\nabla_{\boldsymbol{\theta}} L_k(\boldsymbol{\theta}_0)\|_\infty \leqslant 4\max\{\psi_x, 1\}(R_{\mathcal{X}}R_\Theta + B)\sqrt{\frac{\log n}{n}}. \tag{13}$$

Finally, combining (7), (10) and (13), we obtain that with probability at least $1 - 2e^{-c_1 c_{\min}^2 |I_k|/16} - 2/|I_k|$,

$$\|\widehat{\boldsymbol{\theta}}_k - \boldsymbol{\theta}_0\|_2 \leqslant \frac{8\max\{\psi_x, 1\}(R_{\mathcal{X}}R_\Theta + B)}{c_{\min}}\sqrt{\frac{(d+1)\log|I_k|}{|I_k|}}.$$

C.2 Proof of Lemma 2

For the following analysis, we fix any episode index k satisfying the conditions of Lemma 2. It's easy to verify that for any $k \geqslant (\log(\sqrt{T} - \log \ell_0))/\log 2$, $\Theta_k \subset \Theta_0$. Therefore, all the assumptions hold for $\boldsymbol{\theta} \in \Theta_k$. Our goal is to prove (18) holds with high probability on the k-th episode.

Now we have the i.i.d. samples $\{w_t(\boldsymbol{\theta}) := p_t - \widetilde{\mathrm{x}}_t^\top \boldsymbol{\theta}, y_t\}_{t\in I_k}$ from some distribution $P_{w(\boldsymbol{\theta}),y}$. According to the previous notations, the marginal distribution $P_{w(\boldsymbol{\theta})}$ has density $f_{\boldsymbol{\theta}}(u)$. Moreover, $r_{\boldsymbol{\theta}}(u) := \mathbb{E}[y_t \,|\, w_t(\boldsymbol{\theta}) = u]$. We're interested in bounding the quantity $\sup_{u\in I, \boldsymbol{\theta}\in\Theta_k} |\widehat{r}_k(u, \boldsymbol{\theta}) - r_{\boldsymbol{\theta}_0}(u)|$, which leads to the desired conclusion of the lemma.

For notational simplicity, let $n = |I_k|$ be the length of the exploration phase. Recall that $\widehat{r}_k(u, \boldsymbol{\theta}) = h_k(u, \boldsymbol{\theta})/f_k(u, \boldsymbol{\theta})$, where

$$h_k(u, \boldsymbol{\theta}) = \frac{1}{nb_k}\sum_{t\in I_k} K(\frac{w_t(\boldsymbol{\theta}) - u}{b_k})Y_t, \quad f_k(u, \boldsymbol{\theta}) = \frac{1}{nb_k}\sum_{t\in I_k} K(\frac{w_t(\boldsymbol{\theta}) - u}{b_k}).$$

Here, $b_k > 0$ is the bandwidth (to be chosen), and $K(\cdot)$ is some kernel function.

Note that $r_{\boldsymbol{\theta}}(u) = \frac{h_{\boldsymbol{\theta}}(u)}{f_{\boldsymbol{\theta}}(u)}$, we can write the difference between $\widehat{r}_k$ and r as

$$\widehat{r}_k(u, \boldsymbol{\theta}) - r_{\boldsymbol{\theta}}(u) = \frac{h_k(u, \boldsymbol{\theta})}{f_k(u, \boldsymbol{\theta})} - \frac{h_{\boldsymbol{\theta}}(u)}{f_{\boldsymbol{\theta}}(u)} = \frac{h_k(u, \boldsymbol{\theta}) - h_{\boldsymbol{\theta}}(u)}{f_k(u, \boldsymbol{\theta})} + h_{\boldsymbol{\theta}}(u)\cdot[\frac{1}{f_k(u, \boldsymbol{\theta})} - \frac{1}{f_{\boldsymbol{\theta}}(u)}]. \tag{14}$$

The following lemmas are used as tools to control the right hand side of the above equation. The proof of the lemmas can be found in §F.1 and F.2.

Lemma 1. *Under Assumptions 3 – 5, for any $b_k \leqslant 1$,*

$$\sup_{u\in I, \boldsymbol{\theta}\in\Theta_k} |\mathbb{E}h_k(u, \boldsymbol{\theta}) - h_{\boldsymbol{\theta}}(u)| \leqslant C_{x,K}^{(1)} b_k^m, \tag{15}$$

$$\sup_{u\in I, \boldsymbol{\theta}\in\Theta_k} |\mathbb{E}f_k(u, \boldsymbol{\theta}) - f_{\boldsymbol{\theta}}(u)| \leqslant C_{x,K}^{(1)} b_k^m. \tag{16}$$

Here, $C_{x,K}^{(1)} = l_f \frac{\int |s^m K(s)|\mathbf{d}s}{(m-1)!}$.

Lemma 2. *Under Assumptions 3 – 5, $\forall b_k \leqslant 1$, $\delta \in [4e^{-nb_k/3}, \frac{1}{2})$, as long as $nb_k \geqslant \max\{132d(\log \frac{1}{b_k} + 1), 3\log n\}$, either of the following inequalities holds with probability at least $1-\delta$:*

$$\sup_{u\in I,\boldsymbol{\theta}\in\Theta_k} |h_k(u,\boldsymbol{\theta}) - \mathbb{E}h_k(u,\boldsymbol{\theta})| \leqslant C^{(2)}_{x,K}\sqrt{\frac{\log n}{nb_k}}\left(\sqrt{d}+\sqrt{\log 1/\delta}\right), \tag{17}$$

$$\sup_{u\in I,\boldsymbol{\theta}\in\Theta_k} |f_k(u,\boldsymbol{\theta}) - \mathbb{E}f_k(u,\boldsymbol{\theta})| \leqslant C^{(2)}_{x,K}\sqrt{\frac{\log n}{nb_k}}\left(\sqrt{d}+\sqrt{\log 1/\delta}\right). \tag{18}$$

Here $C^{(2)}_{x,K} = l_K\Bigg(8\sqrt{22}\max\{2\bar f\int K^2\mathbf{d}s, 2\bar f\int K'^2\mathbf{d}s, \frac{2}{3}\bar K, 1\}+ \frac{60(6\sqrt{\log 2}+\sqrt{c_0})}{c_0}\sqrt{1+R^2_{\mathcal{X}}}\max\{\delta_z, \frac{\max\{1,\psi_x\}(B+R_{\mathcal{X}}R_\Theta)}{c_{\min}}\}\Bigg)$ *(Numerical constants are not optimized).*

Now according to (14), we have

$$\begin{aligned}
\sup_{u\in I,\boldsymbol{\theta}\in\Theta_k} |\widehat{r}_k(u,\boldsymbol{\theta}) - r(u)| &\leqslant \sup_{u\in I,\boldsymbol{\theta}\in\Theta_k} \frac{|h_k(u,\boldsymbol{\theta}) - h_{\boldsymbol{\theta}}(u)|}{|f_{\boldsymbol{\theta}}(u) - |f_k(u,\boldsymbol{\theta}) - f_{\boldsymbol{\theta}}(u)||} \\
&\quad + \sup_{u\in I,\boldsymbol{\theta}\in\Theta_k} \frac{h_{\boldsymbol{\theta}}(u)}{f_{\boldsymbol{\theta}}(u)}\cdot\frac{|f_k(u,\boldsymbol{\theta}) - f_{\boldsymbol{\theta}}(u)|}{|f_{\boldsymbol{\theta}}(u) - |f_k(u,\boldsymbol{\theta}) - f_{\boldsymbol{\theta}}(u)||} \\
&\leqslant \frac{\sup_{u\in I,\boldsymbol{\theta}\in\Theta_k}|h_k(u,\boldsymbol{\theta}) - h_{\boldsymbol{\theta}}(u)|}{c - \sup_{u\in I,\boldsymbol{\theta}\in\Theta_k}|f_k(u,\boldsymbol{\theta}) - f_{\boldsymbol{\theta}}(u)|} \\
&\quad + \sup_{u\in I,\boldsymbol{\theta}\in\Theta_k} r_{\boldsymbol{\theta}}(u)\cdot\frac{\sup_{u\in I,\boldsymbol{\theta}\in\Theta_k}|f_k(u,\boldsymbol{\theta}) - f_{\boldsymbol{\theta}}(u)|}{c - \sup_{u\in I,\boldsymbol{\theta}\in\Theta_k}|f_k(u,\boldsymbol{\theta}) - f_{\boldsymbol{\theta}}(u)|} \qquad (19)\\
&\leqslant \frac{\sup_{u\in I,\boldsymbol{\theta}\in\Theta_k}|h_k(u,\boldsymbol{\theta}) - h_{\boldsymbol{\theta}}(u)|}{c - \sup_{u\in I,\boldsymbol{\theta}\in\Theta_k}|f_k(u,\boldsymbol{\theta}) - f_{\boldsymbol{\theta}}(u)|} + \frac{\sup_{u\in I,\boldsymbol{\theta}\in\Theta_k}|f_k(u,\boldsymbol{\theta}) - f_{\boldsymbol{\theta}}(u)|}{c - \sup_{u\in I,\boldsymbol{\theta}\in\Theta_k}|f_k(u,\boldsymbol{\theta}) - f_{\boldsymbol{\theta}}(u)|} \qquad (20)
\end{aligned}$$

as long as we ensure that $\sup_{u\in I,\boldsymbol{\theta}\in\Theta_k}|f_k(u,\boldsymbol{\theta}) - f_{\boldsymbol{\theta}}(u)| \leqslant \frac{c}{2}$.

Let $b_k = n^{-\frac{1}{2m+1}}$. By letting $B_{x,K} = \max\{4{C^{(3)}_{x,K}}^8/c^8, (2c_0)^4, (2C_b)^4\}$, we can verify that for any qualifying episode k, $nb_k \geqslant \max\{C_b d(\log\frac{1}{b_k}+1), 3\log n\}$. Combining (15) and (17), we have that $\forall\delta \in [4\exp(-n^{\frac{2m}{2m+1}}/3), \frac{1}{2})$, with probability at least $1-\delta$,

$$\begin{aligned}
\sup_{u\in I,\boldsymbol{\theta}\in\Theta_k}|h_k(u,\boldsymbol{\theta}) - h_{\boldsymbol{\theta}}(u)| &\leqslant \sup_{u\in I,\boldsymbol{\theta}\in\Theta_k}|h_k(u,\boldsymbol{\theta}) - \mathbb{E}h_k(u,\boldsymbol{\theta})| + \sup_{u\in I,\boldsymbol{\theta}\in\Theta_k}|\mathbb{E}h_k(u,\boldsymbol{\theta}) - h_{\boldsymbol{\theta}}(u)| \\
&\leqslant C^{(1)}_{x,K}n^{-\frac{m}{2m+1}} + C^{(2)}_{x,K}\sqrt{\frac{\log n}{nb_k}}\left(\sqrt{d}+\sqrt{\log 1/\delta}\right) \\
&\leqslant C^{(3)}_{x,K}n^{-\frac{m}{2m+1}}\sqrt{\log n}\left(\sqrt{d}+\sqrt{\log 1/\delta}\right).
\end{aligned}$$

Here, $C^{(3)}_{x,K} = C^{(1)}_{x,K} + C^{(2)}_{x,K}$. Similarly, with probability at least $1-\delta$,

$$\sup_{u\in I,\boldsymbol{\theta}\in\Theta_k} |f_k(u,\boldsymbol{\theta}) - f_{\boldsymbol{\theta}}(u)| \leqslant C^{(3)}_{x,K} n^{-\frac{m}{2m+1}}\sqrt{\log n}\left(\sqrt{d}+\sqrt{\log 1/\delta}\right).$$

It's easily seen that as long as $n^{\frac{m}{2m+1}}/\sqrt{\log n} \geqslant \frac{2C^{(3)}_{x,K}}{c}(\sqrt{d}+\sqrt{\log 1/\delta})$, The right hand side of the above inequality is upper bounded by $c/2$, which guarantees that

$$\sup_{u\in I,\boldsymbol{\theta}\in\Theta_k} |f_k(u,\boldsymbol{\theta}) - f_{\boldsymbol{\theta}}(u)| \leqslant \frac{c}{2}.$$

(*Remark*: From the conditions in the lemma, by letting $B_{x,K} = \max\{4{C^{(3)}_{x,K}}^8/c^8, (2c_0)^4, (2C_b)^4\}$ and $B'_{x,K} = \min\{\big(\frac{c}{4C^{(3)}_{x,K}}\big)^2, 1/3\}$, we have

$$n^{\frac{m}{2m+1}}/\sqrt{\log n} \geqslant \frac{4C^{(3)}_{x,K}}{c}\sqrt{d}, \quad n^{\frac{m}{2m+1}}/\sqrt{\log n} \geqslant \frac{4C^{(3)}_{x,K}}{c}\sqrt{\log\frac{1}{\delta}},$$

which lead to $n^{\frac{m}{2m+1}}/\sqrt{\log n} \geqslant \frac{2C^{(3)}_{x,K}}{c}(\sqrt{d}+\sqrt{\log 1/\delta})$.)

Plugging the above results into inequality (20) gives

$$\sup_{u\in I,\boldsymbol{\theta}\in\Theta_k} |\widehat{r}_k(u,\boldsymbol{\theta}) - r_{\boldsymbol{\theta}}(u)| \leqslant \frac{4C^{(3)}_{x,K}}{c} n^{-\frac{m}{2m+1}}\sqrt{\log n}\left(\sqrt{d}+\sqrt{\log 1/\delta}\right). \tag{21}$$

Next, we proceed to upper bound the quantity $\sup_{t\in I,\boldsymbol{\theta}\in\Theta_k}|r_{\boldsymbol{\theta}}(u) - r_{\boldsymbol{\theta}_0}(u)|$. We know that for any $\boldsymbol{\theta}\in\Theta_k$,

$$r_{\boldsymbol{\theta}}(u) = \mathbb{E}[Y_t\,|\,p_t - \widetilde{\mathbf{x}}_t^\top\boldsymbol{\theta} = u] = \mathbb{E}[\mathbb{E}[Y_t\,|\,\widetilde{\mathbf{x}}_t, p_t]\,|\,p_t - \widetilde{\mathbf{x}}_t^\top\boldsymbol{\theta} = u] = \mathbb{E}[r_{\boldsymbol{\theta}_0}(p_t - \widetilde{\mathbf{x}}_t^\top\boldsymbol{\theta}_0)\,|\,p_t - \widetilde{\mathbf{x}}_t^\top\boldsymbol{\theta} = u].$$

Moreover from the Lipchitz property of $r_{\boldsymbol{\theta}_0}$,

$$\sup_{\mathbf{x}\in\mathcal{X},\boldsymbol{\theta}\in\Theta_k} |r_{\boldsymbol{\theta}_0}(p_t - \widetilde{\mathbf{x}}^\top\boldsymbol{\theta}_0) - r_{\boldsymbol{\theta}_0}(p_t - \widetilde{\mathbf{x}}^\top\boldsymbol{\theta})| \leqslant l_r R_{\mathcal{X}} R_k = l_r R_{\mathcal{X}}\cdot\frac{10\max\{\psi_x,1\}(B+R_{\mathcal{X}}R_\Theta)}{c_{\min}}\sqrt{\frac{(d+1)\log n}{n}}.$$

Therefore,

$$\sup_{u\in I,\boldsymbol{\theta}\in\Theta_k} |r_{\boldsymbol{\theta}}(u) - r_{\boldsymbol{\theta}_0}(u)| \leqslant \mathbb{E}\Big[\sup_{\mathbf{x}\in\mathcal{X},\boldsymbol{\theta}\in\Theta_k} |r_{\boldsymbol{\theta}_0}(p_t - \widetilde{\mathbf{x}}^\top\boldsymbol{\theta}_0) - r_{\boldsymbol{\theta}_0}(p_t - \widetilde{\mathbf{x}}^\top\boldsymbol{\theta})|\,|\,p_t - \widetilde{\mathbf{x}}_t^\top\boldsymbol{\theta} = u\Big] \leqslant C^{(4)}_{x,K}\sqrt{\frac{d\log n}{n}}, \tag{22}$$

where $C^{(4)}_{x,K} = l_r R_{\mathcal{X}}\cdot\frac{10\max\{\psi_x,1\}(B+R_{\mathcal{X}}R_\Theta)}{c_{\min}}$.

Finally, after combing our results in (21)-(22), we claim our conclusion for Lemma 2.

C.3 Proof of Lemma 3

Following the same settings as in the proof of Lemma 2, we now aim at bounding the quantity $\sup_{u\in I,\boldsymbol{\theta}\in\Theta_k}|\widehat{r}_k^{(1)}(u,\boldsymbol{\theta})-r'_{\boldsymbol{\theta}_0}(u)|$, where

$$r_k^{(1)}(u,\boldsymbol{\theta}) = \frac{h_k^{(1)}(u,\boldsymbol{\theta})f_k(u,\boldsymbol{\theta}) - h_k(u,\boldsymbol{\theta})f_k^{(1)}(u,\boldsymbol{\theta})}{f_k^2(u,\boldsymbol{\theta})},$$

$$h_k(u,\boldsymbol{\theta}) = \frac{1}{nb_k}\sum_{u\in I_k}K(\frac{w_t(\boldsymbol{\theta})-u}{b_k})Y_t, \quad f_k(u,\boldsymbol{\theta}) = \frac{1}{nb_k}\sum_{t\in I_k}K(\frac{w_t(\boldsymbol{\theta})-u}{b_k}),$$

$$h_k^{(1)}(u,\boldsymbol{\theta}) = \frac{-1}{nb_k^2}\sum_{t\in I_k}K'(\frac{w_t(\boldsymbol{\theta})-u}{b_k})Y_t, \quad f_k^{(1)}(u,\boldsymbol{\theta}) = \frac{-1}{nb_k^2}\sum_{t\in I_k}K'(\frac{w_t(\boldsymbol{\theta})-u}{b_k}).$$

Similar to the proof of Lemma 2, we will bound $\sup_{u\in I,\boldsymbol{\theta}\in\Theta_k}|\widehat{r}_k^{(1)}(u,\boldsymbol{\theta})-r'_{\boldsymbol{\theta}}(u)|$ and $\sup_{u\in I,\boldsymbol{\theta}\in\Theta_k}|r'_{\boldsymbol{\theta}}(u)-r'_{\boldsymbol{\theta}_0}(u)|$ separately. First, notice that

$$r'_{\boldsymbol{\theta}}(u) = \frac{h'_{\boldsymbol{\theta}}(u)f_{\boldsymbol{\theta}}(u) - f'_{\boldsymbol{\theta}}(u)h_{\boldsymbol{\theta}}(u)}{f_{\boldsymbol{\theta}}^2(u)},$$

we can bound $\sup_{u\in I,\boldsymbol{\theta}\in\Theta_k}|\widehat{r}_k^{(1)}(u,\boldsymbol{\theta})-r'_{\boldsymbol{\theta}}(u)|$ from the following four terms: $\sup_{u\in I,\boldsymbol{\theta}\in\Theta_k}|f_k(u,\boldsymbol{\theta})-f_{\boldsymbol{\theta}}(u)|$, $\sup_{u\in I,\boldsymbol{\theta}\in\Theta_k}|h_k(u,\boldsymbol{\theta})-h_{\boldsymbol{\theta}}(u)|$, $\sup_{u\in I,\boldsymbol{\theta}\in\Theta_k}|f_k^{(1)}(u,\boldsymbol{\theta})-f'_{\boldsymbol{\theta}}(u)|$ and $\sup_{u\in I,\boldsymbol{\theta}\in\Theta_k}|h_k^{(1)}(u,\boldsymbol{\theta})-h'_{\boldsymbol{\theta}}(u)|$. In fact, we can upper bound the first two terms from Lemma 1 and 2. The lemmas below help us bound the last two terms. The proof can be found in §F.3 and F.4.

Lemma 3. *Given Assumptions 3-5, for any $b_k\leqslant 1$,*

$$\sup_{u\in I,\boldsymbol{\theta}\in\Theta_k}|\mathbb{E}h_k^{(1)}(u,\boldsymbol{\theta})-h'_{\boldsymbol{\theta}}(u)| \leqslant C_{x,K}^{(5)}b_k^{m-1}, \tag{23}$$

$$\sup_{u\in I,\boldsymbol{\theta}\in\Theta_k}|\mathbb{E}f_k^{(1)}(u,\boldsymbol{\theta})-f'_{\boldsymbol{\theta}}(u)| \leqslant C_{x,K}^{(5)}b_k^{m-1}. \tag{24}$$

Here, $C_{x,K}^{(5)} = \frac{l_f}{(m-2)!}\int|K(s)s^{m-1}|\mathrm{d}s$.

Lemma 4. *Given assumptions 3, 4 and 5, $\forall b_k\in[\frac{1}{n},1]$, $\delta\in[4e^{-nb_k/3},\frac{1}{2})$, either of the following inequalities holds with probability at least $1-\delta$:*

$$\sup_{u\in I,\boldsymbol{\theta}\in\Theta_k}|h_k^{(1)}(u,\boldsymbol{\theta})-\mathbb{E}h_k^{(1)}(u,\boldsymbol{\theta})| \leqslant C_{x,K}^{(2)}\sqrt{\frac{\log n}{nb_k^3}}\left(\sqrt{d}+\sqrt{\log 1/\delta}\right), \tag{25}$$

$$\sup_{u\in I,\boldsymbol{\theta}\in\Theta_k}|f_k^{(1)}(u,\boldsymbol{\theta})-\mathbb{E}f_k^{(1)}(u,\boldsymbol{\theta})| \leqslant C_{x,K}^{(2)}\sqrt{\frac{\log n}{nb_k^3}}\left(\sqrt{d}+\sqrt{\log 1/\delta}\right). \tag{26}$$

Here $C^{(2)}_{x,K} = l_K\bigg(8\sqrt{22}\max\{2\bar f\int K^2\mathbf{d}s, 2\bar f\int K'^2\mathbf{d}s, \frac{2}{3}\bar K, 1\}+$
$\frac{60(6\sqrt{\log 2}+\sqrt{c_0})}{c_0}\sqrt{1+R_{\mathcal{X}}^2}\max\{\delta_z, \frac{\max\{1,\psi_x\}(B+R_{\mathcal{X}}R_\Theta)}{c_{\min}}\}\bigg)$ *(Numerical constants are not optimized).*

Now let $b_k = n^{-\frac{1}{2m+1}}$. Combining (23) and (25), we obtain that $\forall\delta \in [4\exp(-n^{\frac{2m}{2m+1}}/3), \frac{1}{2})$, with probability at least $1-\delta$,

$$\begin{aligned}
\sup_{u\in I,\boldsymbol{\theta}\in\Theta_k}|h^{(1)}_k(u,\boldsymbol{\theta}) - h'_{\boldsymbol{\theta}}(u)| &\leqslant \sup_{u\in I,\boldsymbol{\theta}\in\Theta_k}|h^{(1)}_k(u,\boldsymbol{\theta}) - \mathbb{E}h^{(1)}_k(u,\boldsymbol{\theta})| + \sup_{u\in I,\boldsymbol{\theta}\in\Theta_k}|\mathbb{E}h^{(1)}_k(u,\boldsymbol{\theta}) - h'_{\boldsymbol{\theta}}(u)| \\
&\leqslant C^{(5)}_{x,K}n^{-\frac{m-1}{2m+1}} + C^{(2)}_{x,K}\sqrt{\frac{\log n}{nb_k^3}}\left(\sqrt{d}+\sqrt{\log 1/\delta}\right) \\
&\leqslant C^{(6)}_{x,K}n^{-\frac{m-1}{2m+1}}\sqrt{\log n}\left(\sqrt{d}+\sqrt{\log 1/\delta}\right)
\end{aligned}$$

Here, $C^{(6)}_{x,K} = C^{(5)}_{x,K} + C^{(2)}_{x,K}$. Similarly, with probability at least $1-\delta$,

$$\sup_{u\in I,\boldsymbol{\theta}\in\Theta_k}|f^{(1)}_k(u,\boldsymbol{\theta}) - f'_{\boldsymbol{\theta}}(u)| \leqslant C^{(6)}_{x,K}n^{-\frac{m-1}{2m+1}}\sqrt{\log n}\left(\sqrt{d}+\sqrt{\log 1/\delta}\right).$$

Recall that when $n^{\frac{m}{2m+1}}/\sqrt{\log n} \geqslant \frac{2C^{(3)}_{x,K}}{c}(\sqrt{d}+\sqrt{\log 1/\delta})$, we have

$$\sup_{u\in I,\boldsymbol{\theta}\in\Theta_k}|f_k(u,\boldsymbol{\theta}) - f_{\boldsymbol{\theta}}(u)| \leqslant \frac{c}{2}, \qquad \sup_{u\in I,\boldsymbol{\theta}\in\Theta_k}|h_k(u,\boldsymbol{\theta}) - h_{\boldsymbol{\theta}}(u)| \leqslant \frac{c}{2}.$$

Moreover, we have

$$\sup_{u\in I,\boldsymbol{\theta}\in\Theta_k}\max\{|h_{\boldsymbol{\theta}}(u)|, |f_{\boldsymbol{\theta}}(u)|, |f'_{\boldsymbol{\theta}}(u)|\} \leqslant \bar f, \qquad \sup_{u\in I,\boldsymbol{\theta}\in\Theta_k}|h'_{\boldsymbol{\theta}}(u)| = \sup_{u\in I,\boldsymbol{\theta}\in\Theta_k}|f'_{\boldsymbol{\theta}}(u)r_{\boldsymbol{\theta}}(u)+f_{\boldsymbol{\theta}}(u)r'_{\boldsymbol{\theta}}(u)| \leqslant l_f + l_r\bar f.$$

Therefore, from the definition of $r^{(1)}_k(u,\boldsymbol{\theta})$ and $r'_{\boldsymbol{\theta}}(u)$, we have

$$\begin{aligned}
&\sup_{u\in I,\boldsymbol{\theta}\in\Theta_k}|r^{(1)}_k(\boldsymbol{\theta},u) - r'_{\boldsymbol{\theta}}(u)| \\
\leqslant &\sup_{u\in I,\boldsymbol{\theta}\in\Theta_k}\left|[h'_{\boldsymbol{\theta}}(u)f_{\boldsymbol{\theta}}(u) - h_{\boldsymbol{\theta}}(u)f'_{\boldsymbol{\theta}}(u)]\left[\frac{1}{f_k(u,\boldsymbol{\theta})^2} - \frac{1}{f_{\boldsymbol{\theta}}(u)^2}\right]\right| \\
&+\sup_{u\in I,\boldsymbol{\theta}\in\Theta_k}\left|\frac{1}{f_k(u,\boldsymbol{\theta})^2}\{[h^{(1)}_k(u,\boldsymbol{\theta})f_k(u,\boldsymbol{\theta}) - h_k(u,\boldsymbol{\theta})f^{(1)}_k(u,\boldsymbol{\theta})] - [h'_{\boldsymbol{\theta}}(u)f_{\boldsymbol{\theta}}(u) - h_{\boldsymbol{\theta}}(u)f'_{\boldsymbol{\theta}}(u)]\right| \\
\leqslant &[l_f\bar f + (l_r+1)\bar f^2]\cdot\sup_{u\in I,\boldsymbol{\theta}\in\Theta_k}\left|\frac{f_k(u,\boldsymbol{\theta})^2 - f_{\boldsymbol{\theta}}(u)^2}{f_k(u,\boldsymbol{\theta})^2f_{\boldsymbol{\theta}}(u)^2}\right| \\
&+\sup_{u\in I,\boldsymbol{\theta}\in\Theta_k}\frac{1}{f_k(u,\boldsymbol{\theta})^2}|(h^{(1)}_k(u,\boldsymbol{\theta}) - h'_{\boldsymbol{\theta}}(u))f_k(u,\boldsymbol{\theta}) + h'_{\boldsymbol{\theta}}(u)(f_k(u,\boldsymbol{\theta}) - f_{\boldsymbol{\theta}}(u))
\end{aligned}$$

$$
\begin{aligned}
&\quad - (f_k^{(1)}(u,\boldsymbol{\theta}) - f'_{\boldsymbol{\theta}}(u))h_k(u,\boldsymbol{\theta}) - f'_{\boldsymbol{\theta}}(u)(h_k(u,\boldsymbol{\theta}) - h_{\boldsymbol{\theta}}(u))| \\
\leqslant & [l_f\bar{f} + (l_r+1)\bar{f}^2] \cdot \sup_{u\in I,\boldsymbol{\theta}\in\Theta_k} \frac{\frac{5}{2}f_{\boldsymbol{\theta}}(u)|f_k(u,\boldsymbol{\theta}) - f_{\boldsymbol{\theta}}(u)|}{f_k(u,\boldsymbol{\theta})^2 f_{\boldsymbol{\theta}}(u)^2} \\
&+ \sup_{u\in I,\boldsymbol{\theta}\in\Theta_k} \frac{1}{f_k(u,\boldsymbol{\theta})^2}\Big[\sup_{u\in I,\boldsymbol{\theta}\in\Theta_k} |f_k(u,\boldsymbol{\theta})| \cdot |h_k^{(1)}(u,\boldsymbol{\theta}) - h'_{\boldsymbol{\theta}}(u))| + \sup_{u\in I,\boldsymbol{\theta}\in\Theta_k} |h'_{\boldsymbol{\theta}}(u)| \cdot |f_k(u,\boldsymbol{\theta}) - f_{\boldsymbol{\theta}}(u)| \\
&\quad + \sup_{u\in I,\boldsymbol{\theta}\in\Theta_k} |h_k(u,\boldsymbol{\theta})| \cdot |f_k^{(1)}(u,\boldsymbol{\theta}) - f'_{\boldsymbol{\theta}}(u)| + \sup_{u\in I,\boldsymbol{\theta}\in\Theta_k} |f'_{\boldsymbol{\theta}}(u)| \cdot |h_k(u,\boldsymbol{\theta}) - h_{\boldsymbol{\theta}}(u)|\Big]. \\
\leqslant & C_{x,K}^{(7)} n^{-\frac{m-1}{2m+1}} \sqrt{\log n}\left(\sqrt{d} + \sqrt{\log\frac{1}{\delta}}\right). \qquad (27)
\end{aligned}
$$

when $n^{\frac{m}{2m+1}}/\sqrt{\log n} \geqslant \frac{2C_{x,K}^{(3)}}{c}(\sqrt{d} + \sqrt{\log 1/\delta})$. Here

$$
C_{x,K}^{(7)} = (\frac{10}{c^3} + \frac{4}{c^2})[l_f(\bar{f}+1) + (l_r+1)\bar{f}^2]C_{x,K}^{(3)} + (\frac{8\bar{f}}{c^2} + \frac{4}{c})C_{x,K}^{(6)}.
$$

Next, we bound the term $\sup_{u\in I,\boldsymbol{\theta}\in\Theta_k} |r'_{\boldsymbol{\theta}}(u) - r'_{\boldsymbol{\theta}_0}(u)|$. In fact, according to our assumptions,

$$
\sup_{u\in I,\boldsymbol{\theta}\in\Theta_k} |r'_{\boldsymbol{\theta}}(u) - r'_{\boldsymbol{\theta}_0}(u)| \leqslant C_{x,K}^{(4)}\sqrt{\frac{d\log n}{n}}, \qquad (28)
$$

where $C_{x,K}^{(4)} = l_r R_{\mathcal{X}} \cdot \frac{8\max\{\psi_x,1\}(B+R_{\mathcal{X}}R_\Theta)}{c_{\min}}$. Finally, after combing our results in (27)-(28), we claim our conclusion for Lemma 3.

C.4 Proof of Lemma 4

We'll need the following auxiliary result in order to prove the lemma. The proof of Lemma 5 can be found in section F.5.

Lemma 5. *Given conditions of Lemma 4, for any $\widetilde{\mathbf{x}}_t \in \mathcal{X}$ and $\boldsymbol{\theta} \in \Theta_0$, $\boldsymbol{\theta}^\top\widetilde{\mathbf{x}}_t \in [\delta_z, B-\delta_z]$.*

Now we proceed to the proof. First, we seek an uniform upper bound for $|\widehat{\phi}_k(u) - \phi(u)|$ from lemma 2 and 3. Recall that $\phi(u) = u - \frac{1-F(u)}{F'(u)}$ and $\widehat{\phi}_k(u) = u - \frac{1-\widehat{F}_k(u)}{\widehat{F}_k^{(1)}(u)}$. It's easy to see that the desired uniform bound can be achieved on an interval where F' is bounded below from 0. For this reason, we choose some positive constant $c_{F'}$ and some interval $[l_{F'}, r_{F'}]$ (we'll specify how to choose them later) such that

$$
\inf_{u\in[l_{F'},r_{F'}]} F'(u) \geqslant c_{F'}. \qquad (29)
$$

From Lemma 3 we know that if in addition $|I_k|^{\frac{m-1}{2m+1}} \geqslant \frac{2\widetilde{C}_{x,K}}{c_{F'}}\sqrt{\log|I_k|}(\sqrt{d} + \sqrt{\log 1/\delta})$, then $\sup_{u\in[l_{F'},r_{F'}]} |\widehat{F}_k^{(1)}(u) - F'(u)| \leqslant \frac{c_{F'}}{2}$ with probability at least $1-4\delta$. In fact, the above condition is ensured by

$$T \geqslant \left(\frac{4\widetilde{C}_{x,K}}{c_{F'}}\right)^8 (\log T + 2\log d)^{\frac{4m-1}{m-1}} d^{\frac{2m+1}{m-1}}.$$

Combining (29), Lemma 2 and Lemma 3, we deduce that with probability at least $1-6\delta$,

$$\begin{aligned}
\sup_{u\in[l_{F'},r_{F'}]} |\widehat{\phi}_k(u) - \phi(u)| &\leqslant \sup_{u\in[l_{F'},r_{F'}]} \left|\frac{(1-\widehat{F}_k(u))(F'(u)-\widehat{F}_k^{(1)}(u))}{\widehat{F}_k^{(1)}(u)F'(u)}\right| \\
&\quad + \sup_{v\in[l_{F'},r_{F'}]} \left|\frac{\widehat{F}_k(u)-F(u)}{F'(u)}\right| \\
&\leqslant \frac{2\widetilde{C}_{x,K} + C_{x,K}c_{F'}}{c_{F'}^2} |I_k|^{-\frac{m-1}{2m+1}} \sqrt{\log|I_k|} \left(\sqrt{d} + \sqrt{\log\frac{1}{\delta}}\right) \qquad (30)
\end{aligned}$$

Next, we proceed to bound $\sup_{u\in[\delta_z,B-\delta_z]} |\widehat{g}_k(u) - g(u)|$ from $\sup_{u\in[\delta_z,B-\delta_z]} |\widehat{\phi}_k^{-1}(-u) - \phi^{-1}(-u)|$ for some properly defined $\widehat{\phi}_k^{-1}$. To be more specific, we will also let

$$[\delta_z - B, -\delta_z] \subseteq \phi([l_{F'}, r_{F'}]) \cap \widehat{\phi}_k([l_{F'}, r_{F'}]). \qquad (31)$$

The way we ensure the above is the following: First, according to the assumptions, we know $\phi'(u) \geqslant c_\phi > 0$, and that $\lim_{u\to\delta_z-0}\phi(u) = \delta_z$, $\lim_{u\to l_F^{(1)}+0}\phi(u) = -\infty$ with $l_F^{(1)} = \inf\{u : F'(u) > 0\} > -\delta_z$. We can deduce that

$$m_{F'} = \inf_{u\in[\phi^{-1}(\delta_z-B),\phi^{-1}(-\delta_z)]} F'(u) > 0.$$

Therefore, there exists some $\delta_{F'} > 0$ such that

$$\inf_{u\in[\phi^{-1}(\delta_z-B)-\delta_{F'},\phi^{-1}(-\delta_z)+\delta_{F'}]} F'(u) > \frac{m_{F'}}{2}.$$

Now let $l_{F'} = \phi^{-1}(\delta_z - B) - \delta_{F'}$, $r_{F'} = \phi^{-1}(-\delta_z) + \delta_{F'}$, $c_{F'} = \frac{m_{F'}}{2}$. From the assumptions on ϕ, we have

$$\phi(l_{F'}) \leqslant \delta_z - B - c_\phi\delta_{F'}, \quad \phi(r_{F'}) \geqslant -\delta_z + c_\phi\delta_{F'}.$$

Combining (30), we obtain that as long as

$$\frac{2\widetilde{C}_{x,K} + C_{x,K}c_{F'}}{c_{F'}^2} |I_k|^{-\frac{m-1}{2m+1}} \sqrt{\log|I_k|} \left(\sqrt{d} + \sqrt{\log\frac{1}{\delta}}\right) \leqslant c_\phi\delta_{F'},$$

we can ensure (31). The above condition can be obtained from the fact that

$$T \geqslant \left(\frac{4\widetilde{C}_{x,K} + 2C_{x,K}c_{F'}}{c_{F'}^2 c_{\phi\delta_{F'}}}\right)^8 (\log T + 2\log d)^{\frac{4m-1}{m-1}} d^{\frac{2m+1}{m-1}}.$$

Define

$$\widehat{\phi}_k^{-1}(u) := \inf\{v \in [l_{F'}, r_{F'}] : \widehat{\phi}_k(v) = u\}. \tag{32}$$

We proceed to upper bound $\sup_{u\in[\delta_z-B,-\delta_z]} |\widehat{\phi}_k^{-1}(u) - \phi^{-1}(u)|$. In fact, for any u, let $v_1 = \phi^{-1}(u)$, $v_2 = \widehat{\phi}_k^{-1}(u)$. Then

$$\begin{aligned}|v_1 - v_2| &\leqslant 1/c_\phi \cdot |\phi(v_1) - \phi(v_2)| = 1/c_\phi \cdot |\widehat{\phi}_k(v_2) - \phi(v_2)| \\ &\leqslant 1/c_\phi \cdot \sup_{v\in[l_{F'},r_{F'}]} |\widehat{\phi}_k(v) - \phi(v)| \\ &\leqslant \frac{2\widetilde{C}_{x,K} + C_{x,K}c_{F'}}{c_\phi c_{F'}^2} |I_k|^{-\frac{m-1}{2m+1}} \sqrt{\log|I_k|} \left(\sqrt{d} + \sqrt{\log\frac{1}{\delta}}\right)\end{aligned}$$

with probability at least $1 - 6\delta$.

Finally, since $g(u) = u + \phi^{-1}(-u)$ and $\widehat{g}_k(u) = u + \widehat{\phi}_k^{-1}(-u)$, we conclude Lemma 4 by choosing

$$\bar{B}_{x,K} = \max\left\{B_{x,K}, (\frac{4\widetilde{C}_{x,K}}{c_{F'}})^8, \left(\frac{4\widetilde{C}_{x,K} + 2C_{x,K}c_{F'}}{c_{F'}^2 c_\phi \delta_{F'}}\right)^8, \left[\frac{C_{\boldsymbol{\theta}}^2}{\delta_v^2}(1 + R_{\mathcal{X}}^2)\right]^{\frac{2(4m-1)}{2m+1}}\right\},$$

$$\bar{B}'_{x,K} = \min\left\{B'_{x,K}, (\frac{c_{F'}}{4\widetilde{C}_{x,K}})^2, (\frac{c_{F'}^2 c_\phi \delta_{F'}}{4\widetilde{C}_{x,K} + 2C_{x,K}c_{F'}})^2\right\},$$

and

$$\bar{C}_{x,K} = \frac{2\widetilde{C}_{x,K} + C_{x,K}c_{F'}}{c_\phi c_{F'}^2}.$$

C.5 Proof of Theorem 1

In order to bound the total regret, we first try to bound the regret at each episode k. First, for all $k \leqslant \lfloor(\log(\sqrt{T} + \ell_0) - \log \ell_0) \log 2\rfloor + 1$, we bound the total regret during episode k by $B\ell_k$. It can be easily verified that

$$\sum_{k\leqslant\lfloor(\log(\sqrt{T}+\ell_0)-\log \ell_0)\log 2\rfloor+1} \text{Regret}_k \leqslant 2B\sqrt{T}.$$

We now turn to the case where $k > \lfloor(\log(\sqrt{T} + \ell_0) - \log \ell_0) \log 2\rfloor + 1$. Recall that the conditional expectation of regret at time t given previous information and $\widetilde{\mathbf{x}}_t$ is

$$\mathbb{E}[R_t \,|\, \bar{\mathcal{H}}_{t-1}] = \mathbb{E}[p_t^* \mathbb{I}_{(v_t \geqslant p_t^*)} - p_t \mathbb{I}_{(v_t \geqslant p_t)} \,|\, \bar{\mathcal{H}}_t] = \rho_t(p_t^*) - \rho_t(p_t),$$

where $\bar{\mathcal{H}}_t = \sigma(\mathbf{x}_1, \mathbf{x}_2, \cdots, \mathbf{x}_{t+1}; z_1, \cdots, z_t)$, and we denote $\rho_t(p) := p(1 - F(p - \boldsymbol{\theta}_0^\top \widetilde{\mathbf{x}}_t))$. Using Taylor expansion and the first order condition induced by the optimality of p_t^*, we have

$$\rho_t(p_t) = \rho_t(p_t^*) + \frac{1}{2}\rho_t''(\xi_t)(p_t - p_t^*)^2,$$

where ξ_t is some value lying between p_t and p_t^*. Note that for any $p \in [0, B]$, $|\rho_t''(p)| = |2F'(p - \boldsymbol{\theta}_0^\top \widetilde{\mathbf{x}}_t) - pF''(p - \boldsymbol{\theta}_0^\top \widetilde{\mathbf{x}}_t)| \leqslant 2l_r + Bl_r'$. Thus we deduce that

$$\mathbb{E}[R_t \mid \bar{\mathcal{H}}_{t-1}] = \rho_t(p_t^*) - \rho_t(p_t) \leqslant (2l_r + Bl_r')(p_t - p_t^*)^2,$$

which further implies that the expected regret at time t is bounded by

$$\mathbb{E}R_t \leqslant \frac{1}{2}(2l_r + Bl_r')\mathbb{E}(p_t - p_t^*)^2 \tag{33}$$

On the other hand,

$$\begin{aligned}(p_t - p_t^*)^2 &\leqslant (\widehat{g}_k(\widetilde{\mathbf{x}}_t^\top \widehat{\boldsymbol{\theta}}_k) - g(\widetilde{\mathbf{x}}_t^\top \boldsymbol{\theta}_0))^2 \\ &\leqslant 2(\widehat{g}_k(\widetilde{\mathbf{x}}_t^\top \widehat{\boldsymbol{\theta}}_k) - g(\widetilde{\mathbf{x}}_t^\top \widehat{\boldsymbol{\theta}}_k))^2 + 2(g(\widetilde{\mathbf{x}}_t^\top \widehat{\boldsymbol{\theta}}_k) - g(\widetilde{\mathbf{x}}_t^\top \boldsymbol{\theta}_0))^2 \\ &:= \boldsymbol{J}_1 + \boldsymbol{J}_2.\end{aligned}$$

We first analyze $\boldsymbol{J}_2$. In fact, define the event

$$\mathcal{E}_k := \{\|\widehat{\boldsymbol{\theta}}_k - \boldsymbol{\theta}_0\| \leqslant R_k\},$$

then according to Lemma 1, $\mathbb{P}(\mathcal{E}_k) \leqslant 1 - 2e^{-c_1 c_{\min}^2 |I_k|/16} - 2/|I_k|$. On $\mathcal{E}_k$ we have

$$\boldsymbol{J}_2 \leqslant \frac{2}{c_{\phi^2}}(\widetilde{\mathbf{x}}_t^\top \widehat{\boldsymbol{\theta}}_k - \widetilde{\mathbf{x}}_t^\top \boldsymbol{\theta}_0)^2 \leqslant \frac{2}{c_{\phi^2}} R_{\mathcal{X}}^2 \|\widehat{\boldsymbol{\theta}}_k - \boldsymbol{\theta}_0\|^2 \leqslant \frac{2}{c_{\phi^2}} R_{\mathcal{X}}^2 R_k^2.$$

Therefore,

$$\mathbb{E}\boldsymbol{J}_2 \leqslant \frac{2}{c_{\phi^2}} R_{\mathcal{X}}^2 R_k^2 + 2B^2(2e^{-c_1 c_{\min}^2 |I_k|/16} + 2/|I_k|). \tag{34}$$

As for $\boldsymbol{J}_1$, on the event $\mathcal{E}_k$, we deduce from Lemma 4 that for any

$$\delta \in [4\exp(-\bar{B}_{x,K}|I_k|^{\frac{2m-2}{2m+1}}/\log|I_k|, \frac{1}{2}]$$

, with probability at least $1 - 6\delta$,

$$\boldsymbol{J}_1 \leqslant 2\bigg[\sup_{u \in [\delta_z, B-\delta_z]} (\widehat{g}_k(u) - g(u))\bigg]^2 \leqslant 2\bar{C}_{x,K}^2 |I_K|^{-\frac{2(m-1)}{2m+1}} \log|I_K| \bigg(\sqrt{d} + \sqrt{\log\frac{1}{\delta}}\bigg)^2.$$

By choosing $\delta = 1/|I_k|$, we have

$$\begin{aligned}\mathbb{E}\boldsymbol{J}_1 &\leqslant 2\bar{C}_{x,K}^2 |I_K|^{-\frac{2(m-1)}{2m+1}} \log|I_K| \bigg(\sqrt{d} + \sqrt{\log\frac{1}{\delta}}\bigg)^2 + 2B^2 \cdot 6\delta \\ &\leqslant 4\bar{C}_{x,K}^2 |I_K|^{-\frac{2(m-1)}{2m+1}} \log|I_K| \bigg(d + \log|I_K|\bigg) + \frac{12B^2}{|I_k|}\end{aligned} \tag{35}$$

Combining (33), (34) and (35), we obtain an upper bound for the expected regret at any time t during episode k:

$$\mathbb{E}R_t \leqslant \bar{C}_{x,K}^{(1)} |I_K|^{-\frac{2(m-1)}{2m+1}} \log |I_K| \Big(d + \log |I_K| \Big),$$

where $\bar{C}_{x,K}^{(1)} = \frac{1}{2}(2l_r + Bl'_r) \cdot [\frac{4}{c_\phi^2} R_{\mathcal{X}}^2 (\frac{10 \max\{\psi_x, 1\}(R_{\mathcal{X}} R_\Theta + B)}{c_{\min}})^2 + 20B^2 + 4C'^2_{x,K}]$. We choose $|I_k| = \lceil (l_k d)^{\frac{2m+1}{4m-1}} \rceil$. The total regret during the k-th episode is

$$\begin{aligned}
\text{Regret}_k &= \sum_{t \in I_k} \mathbb{E}R_t + \sum_{t \in I'_k} \mathbb{E}R_t \\
&\leqslant B|I_k| + l_k \cdot \mathbb{E}R_t \\
&\leqslant B(l_k d)^{(2m+1)/(4m-1)} + B + l_k \cdot \bar{C}_{x,K}^{(1)} (l_k d)^{-(2m-2)/(4m-1)} \log T (d + \log T) \\
&\leqslant (2B + \bar{C}_{x,K}^{(1)}) l_k^{\frac{2m+1}{4m-1}} d^{\frac{2m+1}{4m-1}} \log T (1 + \log T / d).
\end{aligned}$$

Finally, the total regret defined in (6) can be bounded by

$$\begin{aligned}
&\text{Regret}_\pi(T) = \sum_{k=1}^K \text{Regret}_k \leqslant 2B\sqrt{T} + (2B + \bar{C}_{x,K}^{(1)}) d^{\frac{2m+1}{4m-1}} \log T (1 + \log T/d) \sum_{k=1}^K l_k^{(2m+1)/(4m-1)} \\
&\leqslant \left[2B + \frac{2 l_0^{(2m+1)/(4m-1)} (2B + \bar{C}_{x,K}^{(1)})}{2^{(2m+1)/(4m-1)} - 1} \right] (Td)^{\frac{2m+1}{4m-1}} \log T \left(1 + \frac{\log T}{d} \right). \qquad (36)
\end{aligned}$$

Here $K = \lceil \log_2 T \rceil$. The proof is then finished by letting $C_{x,K}^* = 2B + \frac{2 l_0^{(2m+1)/(4m-1)} (2B + \bar{C}_{x,K}^{(1)})}{2^{(2m+1)/(4m-1)} - 1}$.

D Proof under the strong-mixing feature setting

In this section, we mainly present the proof of Theorem 2. The proof will be decomposed to the following lemmas, and their proof is also attached.

Before stating the lemmas, we introduce the α-mixing condition.

Definition 1. *[α-mixing] For a sequence of random variables x_i defined on a probability space $(\Omega, \mathcal{X}, \mathbb{P})$, define*

$$\alpha_k = \sup_{l \geqslant 0} \alpha(\sigma(x_t, t \leqslant l), \sigma(x_t, t \geqslant l + k))$$

in which

$$\alpha(\mathcal{A}, \mathcal{B}) = \sup_{A \in \mathcal{A}, B \in \mathcal{B}} \{ |\mathbb{P}(A \cap B) - \mathbb{P}(A)\mathbb{P}(B)| \}$$

From the definition of strong β-mixing, we see that it can infer strong α-mixing conditions. So in this case, our sequence $\mathbf{x}_t$ also follows strong α-mixing conditions, with $\alpha_k \leqslant e^{-ck}$.

Lemma 6. *[Parametric estimation under dependence] Under Assumption 2 and 6, there exist positive constants c_1 and c_2 (only depend on constants given in Assumptions) such that when $|I_k| \geqslant \max\{c_1(d+1), c_2 \log^2 |I_k| \log\log |I_k|\}$, for any episode k within the horizon, with probability $1-4/|I_k|^2$, we obtain*

$$\|\widehat{\boldsymbol{\theta}}_k - \boldsymbol{\theta}_0\|_2 \leqslant \frac{2}{c_{\min}} \sqrt{\frac{(d+1)(6W_x^2 \log |I_k| + 6W_x \log^2 |I_k| \log\log |I_k|)}{C_w |I_k|}},$$

where $W_x = 2R_{\mathcal{X}}(R_{\mathcal{X}} R_\Theta + B)$.

The proof of Lemma 6 can be found in §F.6. Next, we present the following results on estimation error of $F(\cdot)$ and $F'(\cdot)$:

Lemma 7. *Suppose that Assumptions 3, 4, 5, 6 and 7 hold. Then there exist constants $B_{mx,K}$, $B'_{mx,K}$, $C_{mx,K}$ only depending on $R_{\mathcal{X}} := \sup_{\mathbf{x}\in\mathcal{X}} \|\mathbf{x}\|_2$ and constants within assumptions, such that as long as*

$$T \geqslant B_{mx,K}(\log T + 2\log d)^{\frac{12m-3}{m}} [(d+1)\log(d+1)]^{\frac{4m-1}{m}}/d^2,$$

we have for any $k \geqslant \lfloor(\log(\sqrt{T}+\ell_0) - \log \ell_0)/\log 2\rfloor + 2$, and $\delta \in [8\exp(-|I_k|^{\frac{2m}{2m+1}}/(B'_{mx,K}\log^2 |I_k|)), 1/2]$ with probability at least $1-2\delta$,

$$\sup_{u\in I, \boldsymbol{\theta}\in\Theta_k} |\widehat{F}_k(u,\boldsymbol{\theta}) - F(u)| \leqslant C_{mx,K} |I_k|^{-\frac{m}{2m+1}} \log |I_k| \Big(\sqrt{(d+1)\log(d+1)\log |I_k|} + \sqrt{2\log\frac{8}{\delta}}\Big). \tag{37}$$

Here $I = [-\delta_z, \delta_z]$ and we choose the bandwidth $b_k = |I_k|^{-\frac{1}{2m+1}}$.

The proof of Lemma 7 can be found in §F.7.

Lemma 8. *Suppose that Assumptions 3, 4, 5, 6 and 7 hold. Then there exist constants $\bar{B}_{mx,K}$, $\bar{B}'_{mx,K}$, $\bar{C}_{mx,K}$ that depending only on $R_{\mathcal{X}} := \sup_{\mathbf{x}\in\mathcal{X}} \|\mathbf{x}\|_2$ and the constants within the assmptions such that as long as*

$$T \geqslant \bar{B}_{mx,K}(\log T + 2\log d)^{\frac{12m-3}{m}} [(d+1)\log(d+1)]^{\frac{4m-1}{m}}/d^2,$$

for any $k \geqslant \lfloor(\log(\sqrt{T}+\ell_0) - \log \ell_0)/\log 2\rfloor + 2$ and $\delta \in [\{8\exp(-|I_k|^{\frac{2m}{2m+1}}/(\bar{B}'_{mx,K}\log^2 |I_k|)), 1/2]$ we have with probability at least $1-4\delta$,

$$\sup_{u\in I, \boldsymbol{\theta}\in\Theta_k} |\widehat{F}_k^{(1)}(u,\boldsymbol{\theta}) - F'(u)| \leqslant \bar{C}_{mx,K} |I_k|^{-\frac{m-1}{2m+1}} \log |I_k| \Big(\sqrt{(d+1)\log(d+1)\log |I_k|} + \sqrt{2\log\frac{8}{\delta}}\Big). \tag{38}$$

Here $I = [-\delta_z, \delta_z]$ and we choose the bandwidth $b_k = |I_k|^{-\frac{1}{2m+1}}$.

The proof of this lemma can be found in §F.8.

By combining these two lemmas and following our conclusions from Lemma 4, we are able to achieve the regret bound at the same order with Theorem 1 in Theorem 2.

E Proof under the super smooth noise distribution setting

Proof of Theorem 3 can be followed directly from the proof of Theorem 2 by substituting the Lemma 5 with Lemma 6. Below we'll only present the proof of Lemma 5.

Proof. We only bound $\sup_{u\in I,\boldsymbol{\theta}\in\Theta_k}|\mathbb{E}[f_k(u,\boldsymbol{\theta})]-f_{\boldsymbol{\theta}}(u)|$ and $\sup_{u\in I,\boldsymbol{\theta}\in\Theta_k}|\mathbb{E}[f_k^{(1)}(u,\boldsymbol{\theta})]-f'_{\boldsymbol{\theta}}(u)|$, since the analysis for $f_k(u,\boldsymbol{\theta})$ and $h_k(u,\boldsymbol{\theta})$ are the same. In fact, under the settings of Lemma 5, for any $u\in I,\boldsymbol{\theta}\in\Theta_k$,

$$\begin{aligned}\mathbb{E}[f_k(u,\boldsymbol{\theta})]-f_{\boldsymbol{\theta}}(u) &= \int_{\mathbb{R}}\frac{1}{b_k}K\Big(\frac{s-u}{b_k}\Big)f_{\boldsymbol{\theta}}(s)\mathrm{d}s - f_{\boldsymbol{\theta}}(u)\\ &= \mathcal{F}\bigg(\mathcal{F}^{-1}\Big(\int_{\mathbb{R}}\frac{1}{b_k}K\Big(\frac{s-u}{b_k}\Big)f_{\boldsymbol{\theta}}(s)\mathrm{d}s\Big)-\mathcal{F}^{-1}\circ f_{\boldsymbol{\theta}}(u)\bigg)\\ &= \mathcal{F}\bigg(\phi_{\boldsymbol{\theta}}(u)\Big[\mathcal{F}^{-1}\Big(\frac{1}{b_k}K\Big(\frac{-u}{b_k}\Big)\Big)-1\Big]\bigg)\\ &= \mathcal{F}(\phi_{\boldsymbol{\theta}}(u)[\kappa(-b_k u)-1]).\end{aligned}$$

Here $\mathcal{F}$ is the Fourier transform operator defined by

$$g\to\mathcal{F}\circ g(u)=\frac{1}{2\pi}\int_{\mathbb{R}}g(x)e^{-iux}\mathrm{d}x,$$

and we've utilized the fact that $K=\mathcal{F}\circ\kappa$, $\phi_{\boldsymbol{\theta}}(u)=\mathcal{F}^{-1}\circ f_{\boldsymbol{\theta}}$. Since $|\kappa(x)|\leqslant 1$ for all $x\in\mathbb{R}$ and that $\kappa(x)=1$ for $|x|\leqslant c_\kappa$,

$$\begin{aligned}\sup_{u\in I,\boldsymbol{\theta}\in\Theta_k}|\mathbb{E}[f_k(u,\boldsymbol{\theta})]-f_{\boldsymbol{\theta}}(u)| &\leqslant \sup_{u\in I,\boldsymbol{\theta}\in\Theta_k}|\mathcal{F}(\phi_{\boldsymbol{\theta}}(u)[\kappa(-b_k u)-1])|\\ &\leqslant \sup_{\boldsymbol{\theta}\in\Theta_k}\frac{1}{2\pi}\int|\phi_{\boldsymbol{\theta}}(s)|\cdot|\kappa(-b_k s)-1|\mathrm{d}s\\ &\leqslant \sup_{\boldsymbol{\theta}\in\Theta_0}\frac{1}{\pi}\int_{|s|>c_\kappa/b_k}|\phi_{\boldsymbol{\theta}}(s)|\mathrm{d}s\\ &\leqslant \frac{2}{\pi}\int_{s>0}D_\phi e^{-d_\phi(s+c_\kappa/b_k)^\alpha}\mathrm{d}s\\ &\leqslant \frac{2}{\pi}\int_{s>0}D_\phi e^{-d_\phi/2\cdot[s^\alpha+(c_\kappa/b_k)^\alpha]}\mathrm{d}s.\end{aligned}$$

Here, the last inequality is due to the fact that for $x,y\in\mathbb{R}$, $(x+y)^\alpha\geqslant\min\{2^{\alpha-1},1\}(x^\alpha+y^\alpha)\geqslant\frac{1}{2}(x^\alpha+y^\alpha)$. Thus, by choosing $b_k=c_\kappa(d_\phi/\log|I_k|)^{1/\alpha}$, we obtain that

$$\sup_{u\in I,\boldsymbol{\theta}\in\Theta_k}|\mathbb{E}[f_k(u,\boldsymbol{\theta})]-f_{\boldsymbol{\theta}}(u)|\leqslant C_{\inf}/\sqrt{n},$$

where $C_{\inf}=2D_\phi/\pi\cdot\int_{s>0}\exp(-d_\phi s^\alpha/2)\mathrm{d}s$.

The analysis for $\sup_{u\in I,\boldsymbol{\theta}\in\Theta_k}|\mathbb{E}[f_k^{(1)}(u,\boldsymbol{\theta})]-f'_{\boldsymbol{\theta}}(u)|$ is similar as above. In fact, for any $u\in I,\boldsymbol{\theta}\in\Theta_k$,

$$\begin{aligned}
\mathbb{E}[f_k^{(1)}(u,\boldsymbol{\theta})]-f'_{\boldsymbol{\theta}}(u) &= -\int_{\mathbb{R}}\frac{1}{b_k^2}K'\Big(\frac{s-u}{b_k}\Big)f_{\boldsymbol{\theta}}(s)\mathrm{d}s-f'_{\boldsymbol{\theta}}(u)\\
&= \int_{\mathbb{R}}\frac{1}{b_k}K\Big(\frac{s-u}{b_k}\Big)f'_{\boldsymbol{\theta}}(s)\mathrm{d}s-f'_{\boldsymbol{\theta}}(u)\\
&= \mathcal{F}\Big(\mathcal{F}^{-1}\Big(\int_{\mathbb{R}}\frac{1}{b_k}K\Big(\frac{s-u}{b_k}\Big)f'_{\boldsymbol{\theta}}(s)\mathrm{d}s\Big)-\mathcal{F}^{-1}\circ f'_{\boldsymbol{\theta}}(u)\Big)\\
&= \mathcal{F}\Big(\phi_{\boldsymbol{\theta}}^{(1)}(u)\Big[\mathcal{F}^{-1}\Big(\frac{1}{b_k}K\Big(\frac{-u}{b_k}\Big)\Big)-1\Big]\Big)\\
&= \mathcal{F}(\phi_{\boldsymbol{\theta}}^{(1)}(u)[\kappa(-b_k u)-1]).
\end{aligned}$$

Following the same arguments as above, we deduce that

$$\sup_{u\in I,\boldsymbol{\theta}\in\Theta_k}|\mathbb{E}[f_k^{(1)}(u,\boldsymbol{\theta})]-f'_{\boldsymbol{\theta}}(u)|\leqslant C_{\inf}/\sqrt{n}.$$

□

F Proof of technical lemmas

F.1 Proof of Lemma 1

We only prove (15), since (16) can be proved in the same way.

Recall that $h_k(u,\boldsymbol{\theta})=\frac{1}{nb_k}\sum_{t=1}^n K(\frac{w_t(\boldsymbol{\theta})-u}{b_k})y_t$, and $\mathbb{E}[y_t|w_t(\boldsymbol{\theta})=u]=r_{\boldsymbol{\theta}}(u)=\frac{h_{\boldsymbol{\theta}}(u)}{f_{\boldsymbol{\theta}}(u)}$. We have

$$\mathbb{E}h_k(u,\boldsymbol{\theta})=\frac{1}{b_k}\mathbb{E}K(\frac{w_t(\boldsymbol{\theta})-u}{b_k})y_t=\frac{1}{b_k}\mathbb{E}K(\frac{w_t(\boldsymbol{\theta})-u}{b_k})r(w_t(\boldsymbol{\theta})).$$

Thus,

$$\begin{aligned}
\mathbb{E}h_k(u,\boldsymbol{\theta})-h_{\boldsymbol{\theta}}(u) &= \int\frac{1}{b_k}K(\frac{w(\boldsymbol{\theta})-u}{b_k})r_{\boldsymbol{\theta}}(w(\boldsymbol{\theta}))f_{\boldsymbol{\theta}}(w(\boldsymbol{\theta}))\mathrm{d}w(\boldsymbol{\theta})-h_{\boldsymbol{\theta}}(u)\\
&= \int K(s)h_{\boldsymbol{\theta}}(u+b_k s)\mathrm{d}s-h_{\boldsymbol{\theta}}(u). \qquad (39)
\end{aligned}$$

Using Taylor's expansion, $\forall s\in\mathbb{R}$, there exists some $\xi(s,u)$ lying between the points u and $u+b_k s$ such that

$$h_{\boldsymbol{\theta}}(u+b_k s)=h_{\boldsymbol{\theta}}(u)+\sum_{i=1}^{m-2}\frac{h_{\boldsymbol{\theta}}^{(i)}(u)}{i!}(b_k s)^i+\frac{h_{\boldsymbol{\theta}}^{(m-1)}(\xi(s,u))}{(m-1)!}(b_k s)^{m-1}.$$

Plugging this into (39) gives

$$
\begin{aligned}
\mathbb{E}h_k(u,\boldsymbol{\theta}) - h_{\boldsymbol{\theta}}(u) &= \int K(s)\left[h_{\boldsymbol{\theta}}(u) + \sum_{i=1}^{m-2}\frac{h_{\boldsymbol{\theta}}^{(i)}(u)}{i!}(b_k s)^i + \frac{h_{\boldsymbol{\theta}}^{(m-1)}(\xi(s,u))}{(m-1)!}(b_k s)^{m-1}\right]\mathrm{d}s - h_{\boldsymbol{\theta}}(u) \\
&= \int K(s)\frac{h_{\boldsymbol{\theta}}^{(m-1)}(\xi(s,u))}{(m-1)!}(b_k s)^{m-1}\mathrm{d}s \\
&= \int K(s)\frac{h_{\boldsymbol{\theta}}^{(m-1)}(u)}{(m-1)!}(b_k s)^{m-1}\mathrm{d}s + \int K(s)\frac{[h_{\boldsymbol{\theta}}^{(m-1)}(\xi(s,u)) - h_{\boldsymbol{\theta}}^{(m-1)}(u)]}{(m-1)!}(b_k s)^{m-1}\mathrm{d}s \\
&= \int K(s)\frac{[h_{\boldsymbol{\theta}}^{(m-1)}(\xi(s,u)) - h_{\boldsymbol{\theta}}^{(m-1)}(u)]}{(m-1)!}(b_k s)^{m-1}\mathrm{d}s.
\end{aligned}
$$

Thus we have that

$$
\begin{aligned}
|\mathbb{E}h_k(u,\boldsymbol{\theta}) - h_{\boldsymbol{\theta}}(u)| &\leqslant \int |K(s)|\frac{|h_{\boldsymbol{\theta}}^{(m-1)}(\xi(s,u)) - h_{\boldsymbol{\theta}}^{(m-1)}(u)|}{(m-1)!}|b_k s|^{m-1}\mathrm{d}s \\
&\leqslant \int |K(s)|\frac{l_f|b_k s|}{(m-1)!}|b_k s|^{m-1}\mathrm{d}s \\
&\leqslant C_1 b_k^m,
\end{aligned}
$$

where $C_1 = l_f \cdot \int |s^m K(s)|\mathrm{d}s/(m-1)!$. Moreover, since the inequality holds for any $u \in I$ and $\boldsymbol{\theta} \in \Theta_k$, we finish the proof.

F.2 Proof of Lemma 2

We only prove (17), since (18) can be proved in the same way.

For any $u \in I$, $\boldsymbol{\theta} \in \Theta_k$, denote $Z(u,\boldsymbol{\theta}) := h_k(u,\boldsymbol{\theta}) - \mathbb{E}h_k(u,\boldsymbol{\theta}) = \frac{1}{nb_k}\sum_{t\in I_k}[K(\frac{w_t(\boldsymbol{\theta})-u}{b_k})y_t - \mathbb{E}K(\frac{w_t(\boldsymbol{\theta})-u}{b_k})y_t]$. Then

$$
\sup_{u\in I,\boldsymbol{\theta}\in\Theta_k} |h_k(u,\boldsymbol{\theta}) - \mathbb{E}h_k(u,\boldsymbol{\theta})| = \sup_{u\in I,\boldsymbol{\theta}\in\Theta_k} |Z(u,\boldsymbol{\theta})| = \max\left\{\sup_{u\in I,\boldsymbol{\theta}\in\Theta_k} Z(u,\boldsymbol{\theta}), \sup_{u\in I,\boldsymbol{\theta}\in\Theta_k}(-Z(u,\boldsymbol{\theta}))\right\}.
$$

We can then bound $\sup_{u\in I,\boldsymbol{\theta}\in\Theta_k} |h_k(u,\boldsymbol{\theta}) - \mathbb{E}h_k(u,\boldsymbol{\theta})|$ by upper bounding both $\sup_{u\in I,\boldsymbol{\theta}\in\Theta_k} Z(u,\boldsymbol{\theta})$ and $\sup_{u\in I,\theta\in\Theta_k}(-Z(u,\boldsymbol{\theta}))$. We now give upper bound for $\sup_{u\in I,\boldsymbol{\theta}\in\Theta_k} Z(u,\boldsymbol{\theta})$ with high probability (Bounding $\sup_{u\in I,\boldsymbol{\theta}\in\Theta_k}(-Z(u,\boldsymbol{\theta}))$ is essentially the same).

We use the chaining method to obtain the desired bound. First, we construct a sequence of ε-nets with decreasing scale. Denote the left and right endpoints of the interval I as L_I and R_I respectively. For any $i \in \mathbb{N}^+$, construct set $S_1^{(i)} \subseteq I$ as

$$
S_1^{(i)} \triangleq \left\{L_I + \frac{j}{2^i\sqrt{n}}(R_I - L_I) : j \in \{1, 2, \cdots, (2^i - 1)\lceil\sqrt{n}\rceil\}\right\}.
$$

For any $u \in I$, $i \in \mathbb{N}^+$, let $\pi_1^{(i)}(u) = \arg\min_{s\in S_1^{(i)}} |s - u|$. Moreover, let $\pi_1^{(0)}(u) = u$. Then we can easily verify that $|S_1^{(i)}| \leqslant 2^i(\sqrt{n}+1)$, and $\forall u \in I$, $|\pi_i(u) - \pi_{i+1}(u)| \leqslant \frac{2\delta_z}{2^{i-1}\sqrt{n}}$. At the same time, denote $S_2^{(i)}$ as a $R_k/2^i$-net with respective to l_2-distance of Θ_k, where R_k denotes the radius of Θ_k. Similar to $\pi_1^{(i)}$, define $\pi_2^{(i)}(\boldsymbol{u}) = \arg\min_{\boldsymbol{s}\in S_2^{(i)}} |\boldsymbol{u} - \boldsymbol{s}|$. By Corollary 4.2.13 in Vershynin [2018], $|S_2^{(i)}| \leqslant (2^{i+1}+1)^d$.

Combining the above two nets, we have $S^{(i)} := S_1^{(i)} \times S_2^{(i)}$ is a $2^{-i}\sqrt{4\delta_z^2/n + R_k^2}$-net of $U_k := I \times \Theta_k$ with cardinality $|S^{(i)}| \leqslant 2^i(\sqrt{n}+1)\cdot(2^{i+1}+1)^d$. In fact, for any $\boldsymbol{u} := (u, \boldsymbol{\theta}) \in I \times \Theta_k$ with $i \geqslant 1$, denote $\pi_i(\boldsymbol{u}) := (\pi_1^{(i)}(u), \pi_2^{(i)}(\boldsymbol{\theta}))$, then $\|\pi_i(\boldsymbol{u}) - \boldsymbol{u}\|_2 \leqslant 2^{-i}\sqrt{4\delta_z^2/n + R_k^2}$.

Now, since $Z(u, \boldsymbol{\theta})$ is continuous a.s., we have for any $M \in \mathbb{N}^+$

$$Z(\boldsymbol{u}) - Z(\pi_M(\boldsymbol{u})) = \sum_{i=M}^{\infty}[Z(\pi_{i+1}(\boldsymbol{u})) - Z(\pi_i(\boldsymbol{u}))],$$

and thus

$$\sup_{\boldsymbol{u}\in U_k} Z(\boldsymbol{u}) \leqslant \sup_{\boldsymbol{u}\in U_k} Z(\pi_M(\boldsymbol{u})) + \sum_{i=M}^{\infty}\sup_{\boldsymbol{u}\in U_k}[Z(\pi_{i+1}(\boldsymbol{u})) - Z(\pi_i(\boldsymbol{u}))] \tag{40}$$

almost surely. Our goal is to choose a suitable M such that both terms on the right hand side of (40) can be controlled in a reasonable manner.

For this reason, Let $M = \lceil \frac{3}{\log 2}\log\frac{1}{b_k}\rceil + 10$. We first upper bound $\sup_{\boldsymbol{u}\in U_k} Z(\pi_M(\boldsymbol{u}))$. Note that

$$Z(\boldsymbol{u}) = \frac{1}{nb_k}\sum_{t\in I_k} A_t(\boldsymbol{u}),$$

where $A_t(\boldsymbol{u}) = K(\frac{w_t(\boldsymbol{\theta})-u}{b_k})Y_t - \mathbb{E}K(\frac{w_t(\boldsymbol{\theta})-u}{b_k})Y_t$. We have $\mathbb{E}A_t(\boldsymbol{u}) = 0$ and $|A_t(\boldsymbol{u})| \leqslant \bar{K}$ almost surely. Moreover,

$$\begin{aligned}\mathrm{Var}(A_t(\boldsymbol{u})) &\leqslant \mathbb{E}\left[K(\frac{w_t(\boldsymbol{\theta})-u}{b_k})y_t\right]^2 \leqslant \mathbb{E}\left[K(\frac{w_t(\boldsymbol{\theta})-u}{b_k})\right]^2 \\ &\leqslant \int K(\frac{w_t(\boldsymbol{\theta})-u}{b_k})^2 f_{\boldsymbol{\theta}}(w_t(\boldsymbol{\theta}))\mathrm{d}w_t(\boldsymbol{\theta}) = b_k\int K(s)^2 f_{\theta}(u+b_k s)\mathrm{d}s \leqslant C_4 b_k,\end{aligned}$$

where $C_4 = \max\{\bar{f}\cdot\int K(s)^2\mathrm{d}s, \bar{f}\cdot\int K(s)'^2\mathrm{d}s\}$. Thus according to Bernstein's Inequality, for any $\epsilon > 0$,

$$\mathbb{P}(|Z(\boldsymbol{u})| \geqslant \epsilon) = \mathbb{P}(|\sum_{t\in I_k} A_t(\boldsymbol{u})| \geqslant nb_k\epsilon) \leqslant 2e^{-\frac{n^2 b_k^2\epsilon^2}{2C_4 nb_k + \frac{2}{3}\bar{K}nb_k\epsilon}} \leqslant 2e^{-C_5\frac{nb_k\epsilon^2}{1+\epsilon}},$$

where $C_5 = 1/\max\{2C_4, \frac{2}{3}\bar{K}, 1\}$. A union bound then gives

$$
\begin{aligned}
\mathbb{P}(\sup_{\boldsymbol{u}\in U_k} |Z(\pi_M(\boldsymbol{u}))| \geqslant \epsilon) &\leqslant |S^{(M)}| \cdot \mathbb{P}(|Z(\boldsymbol{u})| \geqslant \epsilon) \\
&\leqslant 2^M(\sqrt{n}+1)\cdot(2^{M+1}+1)^d \cdot 2e^{-C_5\frac{nb_k\epsilon^2}{1+\epsilon}} \\
&\leqslant \exp\left(4dM\log 2 + \log n - \frac{C_5}{2} nb_k \min\{\epsilon, \epsilon^2\}\right).
\end{aligned}
$$

When $\delta \geqslant 4e^{-nb_k/3}$ and $nb_k \geqslant \max\{C_b d(\log\frac{1}{b_k}+1), 3\log n\}$ for some absolute constant $C_b > 0$, by choosing
$\epsilon = \epsilon(k) = \frac{2}{C_5}\frac{1}{\sqrt{nb_k}}\sqrt{4dM\log 2 + \log n + \log\frac{4}{\delta}}$, we can verify that the last term above is upper bounded by $\frac{\delta}{4}$, and thus we have

$$
\mathbb{P}\left(\sup_{\boldsymbol{u}\in \boldsymbol{U}_k} |Z(\pi_M(\boldsymbol{u}))| \geqslant \epsilon(k)\right) \leqslant \frac{\delta}{4}. \tag{41}
$$

Now we proceed to bound the latter term on the right hand side of (40). For any $\boldsymbol{u}_1 := (u, \boldsymbol{\theta}_1), \boldsymbol{u}_2 := (s, \boldsymbol{\theta}_2) \in I \times \Theta_k$, we have

$$
Z(\boldsymbol{u}_1) - Z(\boldsymbol{u}_2) = Z(u, \boldsymbol{\theta}_1) - Z(s, \boldsymbol{\theta}_2) = \frac{1}{nb_k}\sum_{t\in I_k} B_t(u, \boldsymbol{\theta}_1, s, \boldsymbol{\theta}_2),
$$

where

$$
B_t(u, \boldsymbol{\theta}_1, s, \boldsymbol{\theta}_2) = y_t\left(K(\frac{w_t(\boldsymbol{\theta}_1)-u}{b_k}) - K(\frac{w_t(\boldsymbol{\theta}_2)-s}{b_k})\right) - \mathbb{E}y_t\left(K(\frac{w_t(\boldsymbol{\theta}_1)-u}{b_k}) - K(\frac{w_t(\boldsymbol{\theta}_2)-s}{b_k})\right).
$$

Then $\mathbb{E}B_j(u, \boldsymbol{\theta}_1, s, \boldsymbol{\theta}_2) = 0$, and

$$
\begin{aligned}
|Z(\boldsymbol{u}_1) - Z(\boldsymbol{u}_2)| = |B_t(u, \boldsymbol{\theta}_1, s, \boldsymbol{\theta}_2)| &\leqslant 2\left|y_t(K(\frac{w_t(\boldsymbol{\theta}_1)-u}{b_k}) - K(\frac{w_t(\boldsymbol{\theta}_2)-s}{b_k}))\right| \\
&\leqslant \frac{2l_K\sqrt{(\max_{\mathbf{x}\in\mathcal{X}}\|\mathbf{x}\|_2^2+1)}}{b_k}\cdot\|\boldsymbol{u}_1 - \boldsymbol{u}_2\|_2.
\end{aligned}
$$

Using Hoeffding's Inequality, for any $\epsilon > 0$,

$$
\mathbb{P}(|\sum_{t\in I_k} B_t(\boldsymbol{u}_1, \boldsymbol{u}_2)| \geqslant \epsilon) \leqslant 2e^{-\frac{2\epsilon^2}{4l_K^2(R_{\mathcal{X}}^2+1)/b_k^2\cdot n\|\boldsymbol{u}_1-\boldsymbol{u}_2\|_2^2}} = 2e^{-\frac{b_k^2\epsilon^2}{2l_K^2 n(R_{\mathcal{X}}^2+1)\|\boldsymbol{u}_1-\boldsymbol{u}_2\|_2^2}}
$$

Therefore,

$$
\mathbb{P}(|Z(\boldsymbol{u}_1) - Z(\boldsymbol{u}_2)| \geqslant \epsilon) = \mathbb{P}(|\sum_{t\in I_k} B_t(\boldsymbol{u}_1, \boldsymbol{u}_2)| \geqslant nb_k\epsilon) \leqslant 2e^{-\frac{nb_k^4\epsilon^2}{2l_K^2(R_{\mathcal{X}}^2+1)\|\boldsymbol{u}_1-\boldsymbol{u}_2\|_2^2}}.
$$

Recall that $\forall \boldsymbol{u}$, $\|\pi_i(\boldsymbol{u}) - \pi_{i+1}(\boldsymbol{u})\|_2 \leqslant 2^{-i}\sqrt{4\delta_z^2/n + R_k^2}$. We use union bound to obtain

$$\begin{aligned}
&\mathbb{P}(\sup_{\boldsymbol{u}\in \boldsymbol{U}_k} |Z(\pi_{i+1}(\boldsymbol{u})) - Z(\pi_i(\boldsymbol{u}))| \geqslant \epsilon) \\
&\leqslant 2^i(\sqrt{n}+1)\cdot(2^{i+1}+1)^d \cdot 2e^{-\frac{2^{2i-2}n^2 b_k^4 \epsilon^2}{2l_K^2(R_{\mathcal{X}}^2+1)(4\delta_z^2+nR_k^2)}}.
\end{aligned}$$

Let $\epsilon = \frac{l_K\sqrt{(R_{\mathcal{X}}^2+1)(4\delta_z^2+nR_k^2)}\epsilon_i}{2^{i-1}nb_k^2}$. The above inequality reduces to

$$\begin{aligned}
&\mathbb{P}\Big(\sup_{\boldsymbol{u}\in \boldsymbol{U}_k} |Z(\pi_{i+1}(\boldsymbol{u})) - Z(\pi_i(\boldsymbol{u}))| \geqslant \frac{l_K\sqrt{(R_{\mathcal{X}}^2+1)(4\delta_z^2+nR_k^2)}\epsilon_i}{2^{i-1}nb_k^2}\Big) \\
&\leqslant 2^i(\sqrt{n}+1)\cdot(2^{i+1}+1)^d \cdot 2e^{-\frac{\epsilon_i^2}{2}}. \qquad (42)
\end{aligned}$$

Now we choose $\epsilon_i = \sqrt{2\log\frac{8}{\delta} + \log n + (2i+4)(d+2)\log 2}$ and define $W^* := \sqrt{(R_{\mathcal{X}}^2+1)(4\delta_z^2+nR_k^2)}$. Notice that

$$\begin{aligned}
\sum_{i=M}^{\infty} \frac{l_K W^*}{nb_k^2}\frac{\epsilon_i}{2^{i-1}} &\leqslant \frac{l_K W^*}{nb_k^2}\sum_{i=M}^{\infty}\frac{\sqrt{2id\log 2} + \sqrt{(4d+8)\log 2 + \log n + 2\log\frac{8}{\delta}}}{2^{i-1}} \\
&\leqslant \frac{l_K W^*}{nb_k^2}\left[\sqrt{2d\log 2}\sum_{i=M}^{\infty}\frac{i}{2^{i-1}} + \frac{1}{2^{M-2}}\sqrt{(4d+8)\log 2 + \log n + 2\log\frac{8}{\delta}}\right] \\
&\leqslant \frac{l_K W^*}{nb_k^2}\left[\sqrt{2d\log 2}\frac{M+1}{2^{M-2}} + \frac{1}{2^{M-2}}\sqrt{(4d+8)\log 2 + \log n + 2\log\frac{8}{\delta}}\right] \\
&\leqslant \frac{l_K W^*}{\sqrt{n b_k}}\cdot\frac{1}{n^{1/2}b_k^{3/2}}\frac{M+2}{2^{M-2}}\left[\sqrt{2\log\frac{8}{\delta} + \log n} + 4\sqrt{d\log 2}\right] \\
&\leqslant \frac{l_K W^*}{\sqrt{n b_k}}\left[\sqrt{\frac{2}{n}\log\frac{8}{\delta} + 1} + \frac{6\sqrt{\log 2}}{\sqrt{c_0}}\right]
\end{aligned}$$

Here we use the fact that when $B_{x,K} \geqslant (2c_0)^4$, combining the assumptions in the lemma, we have $n \geqslant c_0 d$. Combining this fact and a union bound on (42), we get

$$\begin{aligned}
&\mathbb{P}\left(\sup_{\boldsymbol{u}\in \boldsymbol{U}_k} |Z(\boldsymbol{u}) - Z(\pi_M(\boldsymbol{u}))| \geqslant \frac{l_K W^*}{\sqrt{n b_k}}\left[\sqrt{\frac{2}{n}\log\frac{8}{\delta} + 1} + \frac{6\sqrt{\log 2}}{\sqrt{c_0}}\right]\right) \\
\leqslant &\mathbb{P}\left(\sup_{\boldsymbol{u}\in \boldsymbol{U}_k} |Z(\boldsymbol{u}) - Z(\pi_M(\boldsymbol{u}))| \geqslant \sum_{i=M}^{\infty}\frac{l_K W^*}{nb_k^2}\frac{\epsilon_i}{2^{i-1}}\right) \\
\leqslant &\mathbb{P}\left(\sum_{i=M}^{\infty}\sup_{\boldsymbol{u}\in \boldsymbol{U}_k} |Z(\pi_{i+1}(\boldsymbol{u})) - Z(\pi_i(\boldsymbol{u}))| \geqslant \sum_{i=M}^{\infty}\frac{l_K W^*}{nb_k^2}\frac{\epsilon_i}{2^{i-1}}\right)
\end{aligned}$$

$$
\begin{aligned}
&\leqslant \sum_{i=M}^{\infty} \mathbb{P}\left(\sup_{\boldsymbol{u}\in\boldsymbol{U}_k} |Z(\pi_{i+1}(\boldsymbol{u})) - Z(\pi_i(\boldsymbol{u}))| \geqslant \frac{l_K W^*}{nb_k^2}\frac{\epsilon_i}{2^{i-1}}\right)\\
&\leqslant \sum_{i=M}^{\infty} 2^i(\sqrt{n}+1)\cdot(2^{i+1}+1)^d \cdot 2e^{-\frac{\epsilon_i^2}{2}} \leqslant \sum_{i=M}^{\infty}\frac{\delta}{4}\cdot\frac{1}{2^{i+1}} \leqslant \frac{\delta}{4\cdot 2^M} \leqslant \frac{\delta}{4}.
\end{aligned} \tag{43}
$$

Finally, combining (40), (41) and (43), we obtain that

$$
\begin{aligned}
\frac{\delta}{2} &\geqslant \mathbb{P}\Big(\sup_{\boldsymbol{u}\in\boldsymbol{U}_k} |Z(\pi_M(\boldsymbol{u}))| \geqslant \epsilon(k)\Big) + \mathbb{P}\Big(\sup_{\boldsymbol{u}\in\boldsymbol{U}_k} |Z(\boldsymbol{u}) - Z(\pi_M(\boldsymbol{u}))| \geqslant \frac{l_K W^*}{\sqrt{nb_k}}\Big[\sqrt{\frac{2}{n}\log\frac{8}{\delta}} + 1 + \frac{6\sqrt{\log 2}}{\sqrt{c_0}}\Big]\Big)\\
&\geqslant \mathbb{P}\Big(\sup_{\boldsymbol{u}\in\boldsymbol{U}_k} Z(\boldsymbol{u}) \geqslant \epsilon(k) + \frac{l_K W^*}{\sqrt{nb_k}}\Big[\sqrt{\frac{2}{n}\log\frac{8}{\delta}} + 1 + \frac{6\sqrt{\log 2}}{\sqrt{c_0}}\Big]\Big)\\
&\geqslant \mathbb{P}\Big(\sup_{\boldsymbol{u}\in\boldsymbol{U}_k} Z(\boldsymbol{u}) \geqslant \frac{4\sqrt{11}/C_5}{\sqrt{nb_k}}\sqrt{d\left(1+\log\frac{1}{b_k}\right) + \log n + \log\frac{4}{\delta}} + \\
&\quad 16\sqrt{2}\left(1+\frac{6\sqrt{\log 2}}{c_0}\right)\frac{l_K\sqrt{1+R_{\mathcal{X}}^2}}{\sqrt{nb_k}}\max\left\{\delta_z, \frac{\max\{1,\psi_x\}(B+R_{\mathcal{X}}R_\Theta)}{c_{\min}}\right\}\Big(\sqrt{d\log n} + \sqrt{\frac{d\log n}{n}\log\frac{8}{\delta}}\Big)\Big)\\
&\geqslant \mathbb{P}\Big(\sup_{\boldsymbol{u}\in\boldsymbol{U}_k} Z(\boldsymbol{u}) \geqslant C_x l_K\sqrt{\frac{\log n}{nb_k}}\left(\sqrt{d} + \sqrt{\log 1/\delta}\right)\Big).
\end{aligned}
$$

Here we let $C_x = 8\sqrt{22}/C_5 + \frac{60(6\sqrt{\log 2}+\sqrt{c_0})}{c_0}\sqrt{1+R_{\mathcal{X}}^2}\max\{\delta_z, \frac{\max\{1,\psi_x\}(B+R_{\mathcal{X}}R_\Theta)}{c_{\min}}\}$.

For the same reason, we have that

$$
\mathbb{P}\left(\sup_{\boldsymbol{u}\in\boldsymbol{U}_k}(-Z(\boldsymbol{u})) \geqslant C_x l_K\sqrt{\frac{\log n}{nb_k}}\left(\sqrt{d}+\sqrt{\log 1/\delta}\right)\right) \leqslant \frac{\delta}{2}.
$$

Combining the above two inequalities, we finish the proof.

F.3 Proof of Lemma 3

We only prove (23), since (24) can be proved in a similar way. Recall $h_k^{(1)}(u,\boldsymbol{\theta}) = \frac{-1}{nb_k^2}\sum_{t\in I_k} K'(\frac{w_t(\boldsymbol{\theta})-u}{b_k})y_t$, we have

$$
\mathbb{E}h_k^{(1)}(\boldsymbol{\theta},u) = \frac{-1}{b_k^2}\mathbb{E}K'(\frac{w_t(\boldsymbol{\theta})-u}{b_k})y_u = \frac{-1}{b_k^2}\mathbb{E}K'(\frac{w_t(\boldsymbol{\theta})-u}{b_k})r(w_t(\boldsymbol{\theta})).
$$

Then

$$
\begin{aligned}
\mathbb{E}h_k^{(1)}(u,\boldsymbol{\theta}) - h'_{\boldsymbol{\theta}}(u) &= \int \frac{-1}{b_k^2}K'(\frac{w_t(\boldsymbol{\theta})-u}{b_k})h_{\boldsymbol{\theta}}(w_t(\boldsymbol{\theta}))\mathrm{d}w_t(\boldsymbol{\theta}) - h'_{\boldsymbol{\theta}}(u)\\
&= \int K(s)h'_{\boldsymbol{\theta}}(u+b_k s)\mathrm{d}s - h'_{\boldsymbol{\theta}}(u),
\end{aligned} \tag{44}
$$

where (44) follows from integration by parts. By Taylor's expansion, we have

$$h'_{\boldsymbol{\theta}}(u+b_ks) = h'_{\boldsymbol{\theta}}(u) + \sum_{i=2}^{m-2} \frac{h_{\boldsymbol{\theta}}^{(i)}(u)}{(i-1)!}(b_ks)^{i-1} + \frac{h_{\boldsymbol{\theta}}^{(m-1)}(\xi(s,u))}{(m-1)!}(b_ks)^{m-2}.$$

Similar to our proof procedure of Lemma 1, under Assumption 5, we get

$$\mathbb{E}h_k^{(1)}(u,\boldsymbol{\theta}) - h'_{\boldsymbol{\theta}}(u) = \int K(s)\frac{h_{\theta}^{(m-1)}(\xi(s,u)) - h_{\boldsymbol{\theta}}^{(m-1)}(u)}{(m-2)!}(b_ks)^{m-2}\mathrm{d}s.$$

Thus

$$\begin{aligned}
|\mathbb{E}h_k^{(1)}(u,\boldsymbol{\theta}) - h'_{\boldsymbol{\theta}}(u)| &\leqslant \int |K(s)\frac{[h_{\boldsymbol{\theta}}^{(m-1)}(\xi(s,u)) - h_{\boldsymbol{\theta}}^{(m-1)}(u)]}{(m-2)!}(b_ks)^{m-2}|\mathrm{d}s \\
&\leqslant |K(s)|\frac{l_f|b_ks|}{(m-2)!}|b_ks|^{m-2}\mathrm{d}s \\
&\leqslant C_{x,K}^{(5)}b_k^{m-1}, \qquad (45)
\end{aligned}$$

in which $C_{x,K}^{(5)} = \frac{l_f}{(m-2)!}\int |K(s)s^{m-1}|\mathrm{d}s$. Because (45) holds for any $t \in I$ and $\boldsymbol{\theta} \in \Theta_k$, we have

$$\sup_{u\in I,\boldsymbol{\theta}\in\Theta_k} |\mathbb{E}h_k^{(1)}(u,\boldsymbol{\theta}) - h'_{\boldsymbol{\theta}}(u)| \leqslant C_{x,K}^{(5)}b_k^{m-1},$$

which claims inequality 23 of Lemma 3. On the other hand, (24) follows directly from our proof procedure above, so we omit the details.

F.4 Proof of Lemma 4

For any $u \in I, \boldsymbol{\theta} \in \Theta_k$, write

$$Z^{(1)}(u,\boldsymbol{\theta}) = h_k^{(1)}(u,\boldsymbol{\theta}) - \mathbb{E}h_k^{(1)}(u,\boldsymbol{\theta}) = \frac{-1}{b_k}\cdot\frac{1}{nb_k}\sum_{t\in I_k}\Big[K'(\frac{w_t(\boldsymbol{\theta})-u}{b_k})y_t - \mathbb{E}K'(\frac{w_t(\boldsymbol{\theta})-u}{b_k})y_t\Big]$$

Under Assumption 4 and Assumption 5, by following a similar proof procedure with Lemma 2, for $\delta \in [4e^{-nb_k/3}, \frac{1}{2})$, with probability at least $1-\delta$,

$$\sup_{u\in I,\boldsymbol{\theta}\in\Theta_k}\Big|\frac{1}{nb_k}\sum_{t\in I_k}[K'(\frac{w_t(\boldsymbol{\theta})-u}{b_k})y_t - \mathbb{E}K'(\frac{w_t(\boldsymbol{\theta})-u}{b_k})y_t]\Big| \leqslant C_{x,K}^{(2)}\sqrt{\frac{\log n}{nb_k}}\left(\sqrt{d}+\sqrt{\log 1/\delta}\right),$$

where $C_{x,K}^{(2)} = l_K\Big(8\sqrt{22}\max\{2\bar{f}\int K^2\mathbf{d}s, 2\bar{f}\int K'^2\mathbf{d}s, \frac{2}{3}\bar{K}, 1\}+$

$\frac{60(6\sqrt{\log 2}+\sqrt{c_0})}{c_0}\sqrt{1+R_{\mathcal{X}}^2}\max\{\delta_z, \frac{\max\{1,\psi_x\}(B+R_{\mathcal{X}}R_{\Theta})}{c_{\min}}\}\Big)$. Thus, with probability at least $1-\delta$,

$$\sup_{u\in I,\boldsymbol{\theta}\in\Theta_k} |h_k^{(1)}(u,\boldsymbol{\theta}) - \mathbb{E}h_k^{(1)}(u,\boldsymbol{\theta})| \leqslant C_{x,K}^{(2)}\sqrt{\frac{\log n}{nb_k^3}}\left(\sqrt{d}+\sqrt{\log 1/\delta}\right),$$

which claims the inequality (25) in Lemma 4. Moreover, (26) also follows directly from our procedure given above. Thus, we claim our our conclusion of Lemma 4.

F.5 Proof of Lemma 5

First, we argue that for any $\widetilde{\mathbf{x}}_t$,

$$\boldsymbol{\theta}_0^\top \widetilde{\mathbf{x}}_t \in [\delta_z+\delta_v, B-\delta_z-\delta_v]. \tag{46}$$

In fact, we have $v_t = \boldsymbol{\theta}_0^\top \widetilde{\mathbf{x}}_t + z_t$, where $z_t \in [-\delta_z, \delta_z]$ and that $\boldsymbol{\theta}_0^\top \widetilde{\mathbf{x}}_t$ is independent from z_t. Therefore, in order to satisfy the condition $v_t \in [\delta_v, B-\delta_v]$, it ought to be true that $\boldsymbol{\theta}_0^\top \widetilde{\mathbf{x}}_t \in [\delta_z+\delta_v, B-\delta_z-\delta_v]$.

On the other hand,

$$\begin{aligned}
\sup_{\widetilde{\mathbf{x}}_t\in\mathcal{X},\boldsymbol{\theta}\in\Theta_0} |\boldsymbol{\theta}^\top \widetilde{\mathbf{x}}_t - \boldsymbol{\theta}_0^\top \widetilde{\mathbf{x}}_t| &\leqslant \sup_{\boldsymbol{\theta}\in\Theta_0} \|\boldsymbol{\theta}-\boldsymbol{\theta}_0\| \cdot \sup_{\widetilde{\mathbf{x}}_t\in\mathcal{X}} \|\mathbf{x}_t\| \\
&\leqslant C_{\boldsymbol{\theta}} T^{-\frac{2m+1}{4(4m-1)}} d^{\frac{m-1}{4m-1}} \sqrt{\log T + 2\log d}\cdot R_{\mathcal{X}} \\
&\leqslant \delta_v.
\end{aligned} \tag{47}$$

The last inequality is due to the condition on T. The lemma is proved by combining (46) and (47).

F.6 Proof of Lemma 6

The proof of Lemma 6 is similar with our proof of Lemma 1, the major difference between them is that here we assume our covaraites $\widetilde{\mathbf{x}}_t, t \geqslant 0$ follow β-mixing condition instead of of i.i.d. assumption. After following similar proof procedures of (3)-(7), we obtain the same inequality with (7) and we also divide the following proofs into two steps.

Step I: In this step, we prove under β-mixing conditions given in Assumption 6, with high-probability, there exists a constant $c > 0$ such that $\lambda_{\min}(\frac{1}{|I_k|}\sum_{t\in I_k}\widetilde{\mathbf{x}}_t\widetilde{\mathbf{x}}_t^\top) \geqslant c$. In order to prove this, we first use the following matrix Bernstein inequality under β-mixing conditions to prove the concentration between $\Sigma_k := \frac{1}{|I_k|}\sum_{t\in I_k}\widetilde{\mathbf{x}}_t\widetilde{\mathbf{x}}_t^\top$ and $\Sigma := \mathbb{E}[\widetilde{\mathbf{x}}_t\widetilde{\mathbf{x}}_t^\top]$. Similar to §C.1, here for notational convenience, we also denote $n = |I_k|$ for any $k \geqslant 1$ respectively.

Lemma 9 (Matrix Bernstein Inequality under Mixing). *We assume $\widetilde{\mathbf{x}}_t, t \geqslant 0$ satisfy Assumption 6, and we also assume there exists a positive constant M_x such that $\|\widetilde{\mathbf{x}}_t\|_2 \leqslant M_x$. Then for any x and integer $n \geqslant 2$ we have*

$$\mathbb{P}\Big(\|\sum_{t\in I_k}\widetilde{\mathbf{x}}_t\widetilde{\mathbf{x}}_t^\top - n\Sigma\| \geqslant nx\Big) \leqslant 2(d+1)\exp\left(-\frac{C_u n^2 x^2}{v^2 n + M_x^4 + nxM_x^2\log n}\right) \tag{48}$$

where C is a universal constant and

$$v^2 = \sup_{K \in \{1,\dots,n\}} \frac{1}{Card(K)} \lambda_{\max}\Big\{\mathbb{E}\big[\sum_{i\in K}(\widetilde{\mathbf{x}}_i\widetilde{\mathbf{x}}_i^\top - \Sigma)\big]^2\Big\}$$

and v^2 is at the order of M_x^4.

Proof. (48) is a direct consequence of Theorem 1 in Banna et al. [2016], so here we just need to prove the order of v^2.

$$\begin{aligned}\lambda_{\max}\Big\{\mathbb{E}\big[\sum_{i\in K}(\widetilde{\mathbf{x}}_i\widetilde{\mathbf{x}}_i^\top - \Sigma)\big]^2\Big\} &= \lambda_{\max}\Big\{\sum_{i,j\in K}\mathrm{Cov}\big(\widetilde{\mathbf{x}}_i\widetilde{\mathbf{x}}_i^\top, \widetilde{\mathbf{x}}_j\widetilde{\mathbf{x}}_j^\top\big)\Big\}\\ &= \lambda_{\max}\Big\{\sum_{i\in K}\mathrm{Var}(\widetilde{\mathbf{x}}_i\widetilde{\mathbf{x}}_i^\top) + 2\sum_{j>i,\, i,j\in K}\mathrm{Cov}(\widetilde{\mathbf{x}}_i\widetilde{\mathbf{x}}_i^\top, \widetilde{\mathbf{x}}_j\widetilde{\mathbf{x}}_j^\top)\Big\}\end{aligned}$$

Then we get

$$v^2 \leqslant \max_{i\in K}\lambda_{\max}\Big\{\mathrm{Var}(\widetilde{\mathbf{x}}_i\widetilde{\mathbf{x}}_i^\top) + 2\sum_{j>i,\, i,j\in K}\mathrm{Cov}(\widetilde{\mathbf{x}}_i\widetilde{\mathbf{x}}_i^\top, \widetilde{\mathbf{x}}_j\widetilde{\mathbf{x}}_j^\top)\Big\}$$

We know $\|\widetilde{\mathbf{x}}_i\|_2 \leqslant M_x$, so we have

$$\lambda_{\max}\{\mathrm{Var}(\widetilde{\mathbf{x}}_i\widetilde{\mathbf{x}}_i^\top)\} \leqslant \|\mathbb{E}[\widetilde{\mathbf{x}}_i\widetilde{\mathbf{x}}_i^\top\widetilde{\mathbf{x}}_i\widetilde{\mathbf{x}}_i^\top]\| \leqslant M_x^4$$

In addition, we obtain

$$\|\mathrm{Cov}(\widetilde{\mathbf{x}}_i\widetilde{\mathbf{x}}_i^\top, \widetilde{\mathbf{x}}_j\widetilde{\mathbf{x}}_j^\top)\| = \|\mathbb{E}[\widetilde{\mathbf{x}}_i\widetilde{\mathbf{x}}_i^\top\widetilde{\mathbf{x}}_j\widetilde{\mathbf{x}}_j^\top] - \mathbb{E}[\widetilde{\mathbf{x}}_i\widetilde{\mathbf{x}}_i^\top]\mathbb{E}[\widetilde{\mathbf{x}}_j\widetilde{\mathbf{x}}_j^\top]\| \tag{49}$$

By Lemma 1.1 (Berbee's Lemma) given in Bosq, we are able to construct a $\widetilde{\mathbf{x}}_j^*$ such that the distribution of $\widetilde{\mathbf{x}}_j^*$ is the same with $\widetilde{\mathbf{x}}_j$ but is independent with $\widetilde{\mathbf{x}}_i$. At the same time, we also have $\mathbb{P}(\widetilde{\mathbf{x}}_j^* \neq \widetilde{\mathbf{x}}_j) = \beta_{j-i}$ according to Berbee's Lemma. We then proceed to bound (49).

$$\begin{aligned}(49) &= \|\mathbb{E}[\widetilde{\mathbf{x}}_i\widetilde{\mathbf{x}}_i^\top\widetilde{\mathbf{x}}_j\widetilde{\mathbf{x}}_j^\top] - \mathbb{E}[\widetilde{\mathbf{x}}_i\widetilde{\mathbf{x}}_i^\top]\mathbb{E}[\widetilde{\mathbf{x}}_j^*\widetilde{\mathbf{x}}_j^{*\top}]\|\\ &= \|\mathbb{E}[\widetilde{\mathbf{x}}_i\widetilde{\mathbf{x}}_i^\top(\widetilde{\mathbf{x}}_j\widetilde{\mathbf{x}}_j^\top - \widetilde{\mathbf{x}}_j^*\widetilde{\mathbf{x}}_j^{*\top})]\|\\ &\leqslant \|\mathbb{E}[\widetilde{\mathbf{x}}_i\widetilde{\mathbf{x}}_i^\top(\widetilde{\mathbf{x}}_j\widetilde{\mathbf{x}}_j^\top - \widetilde{\mathbf{x}}_j^*\widetilde{\mathbf{x}}_j^{*\top}) \mid \widetilde{\mathbf{x}}_j \neq \widetilde{\mathbf{x}}_j^*]\|\beta_{j-i} \leqslant M_x^4\beta_{j-i}\end{aligned}$$

Then we obtain that there exists a constant $C_v \geqslant 1 + \sum_{j>i}\beta_{j-i}$ s.t.

$$v^2 \leqslant C_v M_x^4,$$

holds, since the term $1 + \sum_{j>i}\beta_{j-i}$ is finite by our Assumption 6 on $\beta_j,\ j \geqslant 0$. Then we conclude our proof of Lemma 9 □

By using conclusions from this Lemma 9, according to Assumption 2 we have $\lambda_{\min}(\Sigma) = c_{\min}$ and $\|\widetilde{\mathbf{x}}_t\|_2 \leqslant M_x := \sqrt{R_{\mathcal{X}}^2 + 1}$, so when

$$n \geqslant \max\{(12C_v(R_{\mathcal{X}}^2 + 1)^2 \log n + 6(R_{\mathcal{X}}^2 + 1)\log^2 n)/(C_u \min\{c_{\min}^2/4, 1\}), d + 1\},$$

$$\lambda_{\min}(\Sigma_k) \geqslant c_{\min}/2. \tag{50}$$

holds with probability $1 - 2/n^2$.

Step II: The next step is to prove the upper bound of $\|\nabla_{\boldsymbol{\theta}} L_k(\boldsymbol{\theta}_0)\|_\infty$. By definition we know

$$\nabla_{\boldsymbol{\theta}} L_k(\boldsymbol{\theta}_0) = \frac{1}{n}\sum_{t\in I_k} 2(\boldsymbol{\theta}_0^\top \widetilde{\mathbf{x}}_t - By_t)\widetilde{\mathbf{x}}_t.$$

Since the expression of $\nabla_{\boldsymbol{\theta}} L_k(\boldsymbol{\theta}_0)$ involves both $\widetilde{\mathbf{x}}_t$ and y_t, $t \in [n]$, next we show the sequence $(\widetilde{\mathbf{x}}_t, y_t), t \geqslant 0$ satisfy α-mixing condition with $\alpha_k \leqslant \exp(-ck)$ under Assumption 6.

Lemma 10 (strong α-mixing of both $\widetilde{\mathbf{x}}$ and y)**.** *Here we denote $\mathcal{A} = \sigma((\widetilde{\mathbf{x}}_t, y_t)_{t\leqslant l})$ and $\mathcal{B} = \sigma((\widetilde{\mathbf{x}}_t, y_t)_{t\leqslant l+k})$. In addition, we also denote $\mathcal{A}_x = \sigma(\widetilde{\mathbf{x}}_{t,\, t\leqslant l})$ and $\mathcal{B}_x = \sigma(\widetilde{\mathbf{x}}_{t,\, t\geqslant l+k})$. Then under Assumption 6, we have for any $l, k \geqslant 0$,*

$$\sup_{l\geqslant 0}\ \sup_{A\in\mathcal{A}, B\in\mathcal{B}} |\mathbb{P}(A, B) - \mathbb{P}(A)\cdot\mathbb{P}(B)| \leqslant \alpha_k$$

where the definition of α_k is given in Definition 1.

Proof.

$$\begin{aligned}\sup_{l\geqslant 0}\ \sup_{A\in\mathcal{A}, B\in\mathcal{B}} |\mathbb{P}(A, B) - \mathbb{P}(A)\cdot\mathbb{P}(B)| &= \sup_{l\geqslant 0}\ \sup_{A\in\mathcal{A}, B\in\mathcal{B}} \Big|\mathbb{E}[\mathbb{I}_{A,B}] - \mathbb{E}[\mathbb{I}_A]\mathbb{E}[\mathbb{I}_B]\Big| \\ &= \sup_{l\geqslant 0}\ \sup_{A\in\mathcal{A}, B\in\mathcal{B}} \Big|\mathbb{E}[\mathbb{E}[\mathbb{I}_{A,B}\,|\,\mathcal{A}_x, \mathcal{B}_x]] - \mathbb{E}[\mathbb{E}[\mathbb{I}_A\,|\,\mathcal{A}_x]]\mathbb{E}[\mathbb{E}[\mathbb{I}_B\,|\,\mathcal{B}_x]]\Big|\end{aligned}$$

After conditioning on $\widetilde{\mathbf{x}}_i, \widetilde{\mathbf{x}}_j$, we observe that y_i, y_j are independent with each other, then we get $\mathbb{E}[\mathbb{I}_{A,B}\,|\,\mathcal{A}_x, \mathcal{B}_x] = \mathbb{E}[\mathbb{I}_A\,|\,\mathcal{A}_x]\cdot\mathbb{E}[\mathbb{I}_B\,|\,\mathcal{B}_x]$. Thus, we have for any $k \geqslant 0$,

$$\begin{aligned}\sup_{l\geqslant 0}\ \sup_{A\in\mathcal{A}, B\in\mathcal{B}} \Big|\mathbb{E}[\mathbb{I}_{A,B}] - \mathbb{E}[\mathbb{I}_A]\mathbb{E}[\mathbb{I}_B]\Big| &= \sup_{l\geqslant 0}\ \sup_{A\in\mathcal{A}, B\in\mathcal{B}} \Big|\mathbb{E}[\mathbb{E}[\mathbb{I}_A\,|\,\mathcal{A}_x]\cdot\mathbb{E}[\mathbb{I}_B\,|\,\mathcal{B}_x]] - \mathbb{E}[\mathbb{E}[\mathbb{I}_A\,|\,\mathcal{A}_x]]\mathbb{E}[\mathbb{E}[\mathbb{I}_B\,|\,\mathcal{B}_x]]\Big| \\ &\leqslant \alpha_k\|\mathbb{I}_A\|_\infty\cdot\|\mathbb{I}_B\|_\infty = \alpha_k\end{aligned}$$

The last inequality follows directly from Corollary 1.1 in Bosq, since $\mathbb{E}[\mathbb{I}_A\,|\,\mathcal{A}_x]$ lies in $\mathcal{A}_x$ and $\mathbb{E}[\mathbb{I}_B\,|\,\mathcal{B}_x]$ lies in $\mathcal{B}_x$. $\square$

By using the same proof given in §C.1, we have $\mathbb{E}[\nabla_{\boldsymbol{\theta}} L_k(\boldsymbol{\theta}_0)] = 0$. In addition, we obtain an upper bound of every entry of $\nabla_{\boldsymbol{\theta}} L_k(\boldsymbol{\theta}_0)$ in a way that there exists a upper bound $W_x = 2R_{\mathcal{X}}(R_{\mathcal{X}}R_\Theta + B)$ of $|2(\boldsymbol{\theta}_0^\top\widetilde{\mathbf{x}}_t - By_t)\widetilde{\mathbf{x}}_{t,i}|$, for every $i \in [d]$. Then using the following vector Bernstein inequality under α-mixing conditions, we obtain an upper bound for $\|\nabla_{\boldsymbol{\theta}} L_k(\boldsymbol{\theta}_0)\|_\infty$.

Lemma 11. *(Vector Bernstein under α-Mixing Conditions, Theorem 1 in Merlevède et al. [2009]) Let $X_j, j \geqslant 0$ be a sequence of centered real-valued random variables. Suppose there exists a positive W_x such that $\sup_i \|X_i\|_\infty \leqslant W_x$, then when $n \geqslant 4$ and $x \geqslant 0$, we obtain*

$$\mathbb{P}\Big(\Big|\frac{1}{n}\sum_{i=1}^{n} X_i\Big| \geqslant x\Big) \leqslant \exp\Big(-\frac{C_w n^2 x^2}{nW_x^2 + W_x nx \log n \log\log n}\Big)$$

where C_w is a universal constant.

By leveraging conclusions from Lemma 11, we have

$$\mathbb{P}(\|\nabla_{\boldsymbol{\theta}} L_k(\boldsymbol{\theta}_0)\|_\infty \geqslant x) \leqslant 2(d+1)\exp\Big(-\frac{C_w n^2 x^2}{nW_x^2 + W_x nx \log n \log\log n}\Big).$$

Thus, when $n \geqslant \max\{(6W_x^2 \log n + 6W_x \log^2 n \log\log n)/C_w, d+1\}$ we obtain, with probability $1 - 2/n^2$, we have

$$\|\nabla_{\boldsymbol{\theta}} L_k(\boldsymbol{\theta}_0)\|_\infty \leqslant \sqrt{(6W_x^2 \log n + 6W_x \log^2 n \log\log n)/(C_w n)}. \tag{51}$$

Then combining our results given in (7), (50) and (51), with probability $1 - 4/|I_k|^2$ we obtain

$$\|\widehat{\boldsymbol{\theta}}_k - \boldsymbol{\theta}_0\|_2 \leqslant \frac{2}{c_{\min}}\sqrt{\frac{(d+1)(6W_x^2 \log|I_k| + 6W_x \log^2|I_k| \log\log|I_k|)}{C_w |I_k|}}$$

for any $k \geqslant 1$.

F.7 Proof of Lemma 7

Proof. Similar with our proof given in §C.2, we suppose $\{w_t(\boldsymbol{\theta}) := p_t - \widetilde{\mathbf{x}}_t^\top \boldsymbol{\theta}, y_t\}_{t\in[n]}$ are observations from the stationary distribution $P_{w(\boldsymbol{\theta}),y}$. We assume that the marginal distribution $P_{w(\boldsymbol{\theta})}$ has density $f_{\boldsymbol{\theta}}(u)$ and let $r_{\boldsymbol{\theta}}(u) = \mathbb{E}[y_t \,|\, w_t(\boldsymbol{\theta}) = u]$ be the regression function to be estimated by estimator

$$\widehat{r}_k(u,\theta) = \frac{h_k(u,\theta)}{f_k(u,\theta)},$$

where

$$h_k(u,\theta) = \frac{1}{nb_k}\sum_{t\in I_k}^{n} K(\frac{w_t(\boldsymbol{\theta}) - u}{b_k})Y_t, \quad f_k(u,\boldsymbol{\theta}) = \frac{1}{nb_k}\sum_{t\in I_k}^{n} K(\frac{w_t(\boldsymbol{\theta}) - u}{b_k}).$$

Here, $b_k > 0$ is the bandwidth (to be chosen) in episode k, $|I_k|$ is denoted as n for simplicity and $K(\cdot)$ is some kernel function. For the true signal $\boldsymbol{\theta}_0$, we denote the true regression function as $r_{\boldsymbol{\theta}_0}(u) = \mathbb{E}[y_t \,|\, w_t(\boldsymbol{\theta}_0) = u]$. The following proof procedures are similar with that given in §C.2, where their major differences are related to control the biases of $|\mathbb{E}[h_k(u,\boldsymbol{\theta})] - h_{\boldsymbol{\theta}}(u)|$ and $|\mathbb{E}[f_k(u,\boldsymbol{\theta})] - f_{\boldsymbol{\theta}}(u)|$ given in Lemma 12 and the variances of $h_k(u,\boldsymbol{\theta})$ and $f_k(u,\boldsymbol{\theta})$ given in Lemma 13 under strong-mixing settings respectively.

Lemma 12. *Under Assumptions 3-5 and 6, with any choice of $b_k \leqslant 1$, we obtain*

$$\sup_{u\in I,\boldsymbol{\theta}\in\Theta_k} |\mathbb{E}h_k(u,\boldsymbol{\theta}) - h_{\boldsymbol{\theta}}(u)| \leqslant C^{(1)}_{mx,K} b_k^m$$
$$\sup_{u\in I,\boldsymbol{\theta}\in\Theta_k} |\mathbb{E}f_k(u,\boldsymbol{\theta}) - f_{\boldsymbol{\theta}}(u)| \leqslant C^{(1)}_{mx,K} b_k^m$$

where $C_{mx,K} = l_f \frac{\int |s^m K(s)\mathrm{d}s}{(m-1)!}$.

Proof. The proof of Lemma 12 is the same with the proof of Lemma 1. So we omit the details. $\square$

Lemma 13. *Under Assumption 3-5 and 6, there exists a constant C'_{17} only depending on constants given in assumptions, such that for $I = [-\delta_z, \delta_z]$, if $b_k \in [1/n, 1]$, $nb_k \geqslant 4C'^2_{17}\log^3 n[(d+1)\log(d+1)]$ and $\delta \in [8\exp(-nb_k/(8C'^2_{17}\log^2 n)), 1/2]$, the following inequalities hold simultaneously with probability $1-\delta$:*

$$\sup_{u\in I,\boldsymbol{\theta}\in\Theta_k} |h_k(u,\boldsymbol{\theta}) - \mathbb{E}[h_k(u,\boldsymbol{\theta})]| \leqslant \frac{C'_{17}\log n}{\sqrt{nb_k}}\left(\sqrt{(d+1)\log(d+1)\log n} + \sqrt{2\log\frac{8}{\delta}}\right) \tag{52}$$

$$\sup_{u\in I,\boldsymbol{\theta}\in\Theta_k} |f_k(u,\boldsymbol{\theta}) - \mathbb{E}[f_k(u,\boldsymbol{\theta})]| \leqslant \frac{C'_{17}\log n}{\sqrt{nb_k}}\left(\sqrt{(d+1)\log(d+1)\log n} + \sqrt{2\log\frac{8}{\delta}}\right) \tag{53}$$

Proof. We only prove (52), since (53) can be proved in the same way. For any $u \in I$ and $\boldsymbol{\theta} \in \Theta_k$, we denote $Z(u,\boldsymbol{\theta}) := h_k(u,\boldsymbol{\theta}) - \mathbb{E}h_k(u,\boldsymbol{\theta}) = \frac{1}{nb_k}\sum_{t\in I_k}[K(\frac{w_t(\boldsymbol{\theta})-u}{b_k})y_t - \mathbb{E}K(\frac{w_t(\boldsymbol{\theta})-u}{b_k})y_t]$. Then we have that

$$\sup_{u\in I,\boldsymbol{\theta}\in\Theta_k} |h_k(u,\boldsymbol{\theta}) - \mathbb{E}h_k(u,\boldsymbol{\theta})| = \sup_{u\in I,\boldsymbol{\theta}\in\Theta_k} |Z(u,\theta)| = \max\left\{\sup_{u\in I,\boldsymbol{\theta}\in\Theta_k} Z(u,\boldsymbol{\theta}), \sup_{u\in I,\boldsymbol{\theta}\in\Theta_k}(-Z(u,\boldsymbol{\theta}))\right\}.$$

Similar with our proof procedure of Lemma 2, we then bound $\sup_{u\in I,\boldsymbol{\theta}\in\Theta_k}|h_k(u,\boldsymbol{\theta}) - \mathbb{E}h_k(u,\boldsymbol{\theta})|$ by upper bounding both $\sup_{u\in I,\boldsymbol{\theta}\in\Theta_k} Z(u,\boldsymbol{\theta})$ and $\sup_{u\in I,\boldsymbol{\theta}\in\Theta_k}(-Z(u,\boldsymbol{\theta}))$. We next also use chaining method to achieve desired bound. We also construct a sequence of ϵ-nets with decreasing scale.

As a reminder, here we also denote the left and right endpoints of the interval I as L_I and R_I respectively. For any $i \in \mathbb{N}^+$, construct set $S_1^{(i)} \subseteq I$ as

$$S_1^{(i)} \triangleq \left\{L_I + \frac{j}{2^i\sqrt{n}}(R_I - L_I) : j \in \{1, 2, \cdots, (2^i-1)\lceil\sqrt{n}\rceil\}\right\}.$$

For any $u \in I$, $i \in \mathbb{N}^+$, let $\pi_i(u) = \arg\min_{s\in S_1^{(i)}}|s-u|$. Moreover, let $\pi_0(u) = u$. Then we can easily verify that $|S_1^{(i)}| \leqslant 2^i(\sqrt{n}+1)$, and that $\forall u \in I$, $|\pi_i(u) - \pi_{i+1}(u)| \leqslant \frac{2\delta_z}{2^{i-1}\sqrt{n}}$.
As for the ϵ-net of Θ_k, we let S_2^i be a $R_m/(2^i\sqrt{n})$-net with respective to l_2-distance of Θ_k, where $R_m = 2/c_{\min}\sqrt{6W_x/C_w}$ (constants are specified in the Lemma 6). By Proposition 4.2.12 in Vershynin [2018], we have $|S_2^{(i)}| \leqslant (2^{i+1}C(d,n)+1)^d$, where $C(d,n) = \sqrt{(d+1)(W_x\log n + \log^2 n\log\log n)}$.

Then we have for any $\mathbf{u} := (u, \boldsymbol{\theta}) \in I \times \Theta_k$ with $i \geqslant 1$, there exist $\pi_i(u) \in S_1^{(i)}$ and $\pi_i(\boldsymbol{\theta}) \in S_2^{(i)}$ such that $\|\pi_i(\mathbf{u}) := (\pi_i(u), \pi_i(\boldsymbol{\theta})) - \mathbf{u}\|_2 \leqslant \sqrt{4\delta_z^2 + R_m^2}/(2^i\sqrt{n})$. So $S^{(i)} := S_1^{(i)} \times S_2^{(i)}$ is a $\sqrt{4\delta_z^2 + R_m^2}/(2^i\sqrt{n})$-net of $U_k := I \times \Theta_k$ with size $|S^{(i)}| \leqslant 2^i(\sqrt{n}+1) \cdot (2^{i+1}C(d,n)+1)^d$ and $C(d,n) = \sqrt{(d+1)(W_x \log n + \log^2 n \log\log n)}$.
Because $Z(u, \boldsymbol{\theta})$ is continuous almost surely, we have that for any $M \in \mathbb{N}^+$

$$Z(\mathbf{u}) - Z(\pi_M(\mathbf{u})) = \sum_{i=M}^{\infty} [Z(\pi_{i+1}(\mathbf{u})) - Z(\pi_i(\mathbf{u}))],$$

and thus

$$\sup_{\mathbf{u}\in U_k} Z(\mathbf{u}) \leqslant \sup_{\mathbf{u}\in U_k} Z(\pi_M(\mathbf{u})) + \sum_{i=M}^{\infty} \sup_{\mathbf{u}\in U_k} [Z(\pi_{i+1}(\mathbf{u})) - Z(\pi_i(\mathbf{u}))] \tag{54}$$

almost surely. If we can choose a M properly then the two terms at the right hand side of (54) can be both well controlled. For this reason, we let $M = \lceil \frac{4}{\log 2} \log \frac{1}{b_n} \rceil$. We then first bound $\sup_{\mathbf{u}\in U_k} Z(\pi_M(\mathbf{u}))$ by using a union bound. By our definition on $Z(\mathbf{u})$, we can write

$$Z(\mathbf{u}) = \frac{1}{nb_k} \sum_{t\in I_k} A_j(\mathbf{u}).$$

in which $A_t(\mathbf{u}) = K(\frac{w_t(\boldsymbol{\theta})-u}{b_k})y_t - \mathbb{E}K(\frac{w_t(\boldsymbol{\theta})-u}{b_k})y_t$. Similar with our case in proving Lemma 2, we have that $\mathbb{E}[A_t(\mathbf{u})] = 0$ and $|A_t(\mathbf{u})| \leqslant \bar{K}$ almost surely. We next prove the bound of variance of $A_t(\mathbf{u})$ and the covariance between $A_j(\mathbf{u})$ and $A_i(\mathbf{u})$ with $j > i$. Following similar procedures with Lemma 2, we first conclude that

$$\mathrm{Var}(A_t(\mathbf{u})) \leqslant C_4' b_k,$$

where $C_4' = C_4 = \max\{\bar{f} \cdot \int K(s)^2 \mathrm{d}s, \bar{f} \cdot \int K'(s)^2 \mathrm{d}s\}$ is defined in the same way with our proof of Lemma 2. We next control the covariance of $A_j(\mathbf{u})$ and $A_i(\mathbf{u})$ with $j > i$.

$$\begin{aligned}
\mathrm{Cov}(A_j(\mathbf{u}), A_i(\mathbf{u})) &= \mathbb{E}\Big[K(\frac{w_j(\boldsymbol{\theta})-u}{b_k})y_j K(\frac{w_i(\boldsymbol{\theta})-u}{b_k})y_i\Big] - \mathbb{E}\Big[K(\frac{w_j(\boldsymbol{\theta})-u}{b_k})y_j\Big]\mathbb{E}\Big[K(\frac{w_i(\boldsymbol{\theta})-u}{b_k})y_i\Big] \\
&= \mathbb{E}\Big[K(\frac{w_j(\boldsymbol{\theta})-u}{b_k})K(\frac{w_i(\boldsymbol{\theta})-u}{b_k})\mathbb{E}[y_j y_i \,|\, w_j(\boldsymbol{\theta}), w_i(\boldsymbol{\theta})]\Big] \\
&\quad - \mathbb{E}\Big[K(\frac{w_j(\boldsymbol{\theta})-u}{b_k})y_j\Big]\mathbb{E}\Big[K(\frac{w_i(\boldsymbol{\theta})-u}{b_k})y_i\Big]
\end{aligned}$$

For simplicity, for any $\boldsymbol{\theta} \in \Theta_0$, we define $r(u_i, u_j) := \mathbb{E}[y_i y_j \,|\, w_j(\boldsymbol{\theta}) = u_j, w_i(\boldsymbol{\theta}) = u_i]$ and $r(u_j) = \mathbb{E}[y_j \,|\, w_j(\boldsymbol{\theta}) = u_j]$. Then after some simple calculation, we further obtain

$$\mathrm{Cov}(A_j(\mathbf{u}), A_i(\mathbf{u})) = \int\int K(\frac{w_j(\boldsymbol{\theta})-u}{b_k})K(\frac{w_i(\boldsymbol{\theta})-u}{b_k})r(w_i(\boldsymbol{\theta}), w_j(\boldsymbol{\theta}))f(w_i(\boldsymbol{\theta}), w_j(\boldsymbol{\theta}))\mathrm{d}w_i(\boldsymbol{\theta})\mathrm{d}w_j(\boldsymbol{\theta})$$

$$
\begin{aligned}
&-\int\int K(\frac{w_j(\boldsymbol{\theta})-u}{b_k})K(\frac{w_i(\boldsymbol{\theta})-u}{b_k})r(w_i(\boldsymbol{\theta}))r(w_j(\boldsymbol{\theta}))f(w_i(\boldsymbol{\theta}))f(w_j(\boldsymbol{\theta}))\mathrm{d}w_i(\boldsymbol{\theta})\mathrm{d}w_j(\boldsymbol{\theta})\\
=\ & b_k^2\int\int K(s_1)K(s_2)[r(b_ks_1+u,b_ks_2+u)f(b_ks_1+u,b_ks_2+u)\\
&-r(b_ks_1+u)r(b_ks_2+u)f(b_ks_1+u)f(b_ks_2+u)]\mathrm{d}s_1\mathrm{d}s_2
\end{aligned}
$$

We next prove that $h(u_i,u_i):=r(u_i,u_j)f(u_i,u_j)$ stays close to $h(u_i)h(u_j):=r(u_i)f(u_i)r(u_j)f(u_j)$ for all $(u_i,\ u_j)$ in the following Lemma 14.

Lemma 14. *Under Assumptions given in Lemma 13. We let $g^*(u_i,u_j):=h(u_i,u_j)-h(u_i)h(u_j)$, if we further assume $g^*(u_i,u_j)$ is Lipschitz continuous w.r.t. (u_i,u_j) with Lipschitz constant l, then we have*

$$
\sup_{u_i,u_j}|g^*(u_i,u_j)|\leqslant(1/4+\sqrt{2}l)\beta_{j-i}^{1/3}
$$

Proof. For any x we define

$$
B(x,\epsilon);=\{x':\|x'-x\|\leqslant\epsilon\},\ \ \epsilon>0,x\in\mathbb{R}
$$

First, we prove $|\mathbb{E}[y_iy_j\mathbb{I}_{\{w_i(\boldsymbol{\theta})\in B(x,\epsilon),w_j(\boldsymbol{\theta})\in B(y,\epsilon)\}}]-\mathbb{E}[y_i\mathbb{I}_{\{w_i(\boldsymbol{\theta})\in B(x,\epsilon)\}}]\mathbb{E}[y_j\mathbb{I}_{\{w_j(\boldsymbol{\theta})\in B(y,\epsilon)\}}]|\leqslant\beta_{j-i}$. We have

$$
\begin{aligned}
&\big|\mathbb{E}[y_iy_j\mathbb{I}_{\{w_i(\boldsymbol{\theta})\in B(x,\epsilon),v_j(\boldsymbol{\theta})\in B(y,\epsilon)\}}]-\mathbb{E}[y_i\mathbb{I}_{\{w_i(\boldsymbol{\theta})\in B(x,\epsilon)\}}]\mathbb{E}[y_j\mathbb{I}_{\{v_j(\boldsymbol{\theta})\in B(y,\epsilon)\}}]\big|\\
&\quad=\big|\mathbb{E}[\mathbb{I}_{\{w_i(\boldsymbol{\theta})\in B(x,\epsilon),v_j(\boldsymbol{\theta})\in B(y,\epsilon)\}}\mathbb{E}[y_iy_j\,|\,\widetilde{\mathbf{x}}_i,\widetilde{\mathbf{x}}_j,p_i,p_j]]\\
&\quad\quad-\mathbb{E}[\mathbb{I}_{\{w_i(\boldsymbol{\theta})\in B(x,\epsilon)\}}\mathbb{E}[y_i\,|\,\widetilde{\mathbf{x}}_i,p_i]]\mathbb{E}[\mathbb{I}_{\{w_j(\boldsymbol{\theta})\in B(y,\epsilon)\}}\mathbb{E}[y_j\,|\,\widetilde{\mathbf{x}}_j,p_j]]\big|\\
&\quad=\big|\mathbb{E}[\mathbb{E}[y_i\mathbb{I}_{\{w_i(\boldsymbol{\theta})\in B(x,\epsilon)\}}\,|\,\widetilde{\mathbf{x}}_i,p_i]\mathbb{E}[y_j\mathbb{I}_{\{w_j(\boldsymbol{\theta})\in B(y,\epsilon)\}}\,|\,\widetilde{\mathbf{x}}_j,p_j]]\\
&\quad\quad-\mathbb{E}[\mathbb{E}[y_i\mathbb{I}_{\{w_i(\boldsymbol{\theta})\in B(x,\epsilon)\}}\,|\,\widetilde{\mathbf{x}}_i,p_i]]\mathbb{E}[\mathbb{E}[y_j\mathbb{I}_{\{w_j(\boldsymbol{\theta})\in B(y,\epsilon)\}}\,|\,\widetilde{\mathbf{x}}_j,p_j]]\big|
\end{aligned}
$$

As $p_i,i\in|I_k|,k\geqslant0$ are independent, so the σ-algebra generated by the joint distribution of $\widetilde{\mathbf{x}}_i,p_i$ still follows strong-β and -α conditions given in our Assumption 6. Moreover, we have $\mathbb{E}[y_i\mathbb{I}_{\{w_i(\boldsymbol{\theta})\in B(x,\epsilon)\}}\,|\,\widetilde{\mathbf{x}}_i,p_i]$ lies in $\sigma(\widetilde{\mathbf{x}}_i,p_i)$ and $\mathbb{E}[y_j\mathbb{I}_{\{w_j(\boldsymbol{\theta})\in B(y,\epsilon)\}}\,|\,\widetilde{\mathbf{x}}_j,p_j]$ lies in $\sigma(\widetilde{\mathbf{x}}_j,p_j)$ with $j>i$. So we are able to obtain the upper bound:

$$
\big|\mathbb{E}[y_iy_j\mathbb{I}_{\{w_i(\boldsymbol{\theta})\in B(x,\epsilon),w_j(\boldsymbol{\theta})\in B(y,\epsilon)\}}]-\mathbb{E}[y_i\mathbb{I}_{\{w_i(\boldsymbol{\theta})\in B(x,\epsilon)\}}]\mathbb{E}[y_j\mathbb{I}_{\{w_j(\boldsymbol{\theta})\in B(y,\epsilon)\}}]\big|\leqslant\beta_{j-i} \tag{55}
$$

by using Corollary 1.1 in Bosq.

Next, we get an upper bound of $\sup_{(u_i,u_j)}|g^*(u_i,u_j)|$. From (55) and our definition on g^*, we obtain

$$
\beta_{j-i}\geqslant\Big|\int_{B(x,\epsilon)\times B(y,\epsilon)}g^*(u_i,u_j)\mathrm{d}u_i\mathrm{d}u_j\Big|:=\mathcal{I}
$$

Then by the mean value property we have $\mathcal{I} = 4\epsilon^2|g^*(x', y')|$ for some $(x', y') \in B(x, \epsilon) \times B(y, \epsilon)$. Moreover, as we assume g is Lipschitz, then we get

$$|g^*(x, y)| \leqslant |g^*(x', y')| + \sqrt{2}l\epsilon$$

Hence, we finally achieve

$$|g^*(x, y)| \leqslant \beta_{j-i}/(4\epsilon^2) + \sqrt{2}l\epsilon.$$

for any fixed (x, y). As this inequality holds for all $\epsilon > 0$, we choose $\epsilon = \beta_{j-i}^{1/3}$ and we conclude the proof of our Lemma 14. $\square$

By our conclusion from Lemma 14, we are able to find a constant C_5' such that $|\sum_{j>i} \mathrm{Cov}(A_j(\boldsymbol{u}), A_i(\boldsymbol{u}))| \leqslant C_5' b_n$ holds according to our assumptions on β_{j-i}, $j > i$, where we set $C_5' = (1/4 + \sqrt{2}l) \sum_{j>0} \beta_j^{1/3}$. Next we introduce the following Bernstein inequality under strong-mixing conditions, in order to achieve an upper bound of $Z(\mathbf{u})$.

Lemma 15. *[Theorem 2 in Merlevède et al. [2009]] Under conditions of Lemma 13, for all $n \geqslant 2$, we have*

$$\mathbb{P}(|Z(\mathbf{u})| \geqslant nb_k x) = \mathbb{P}(|\sum_{j \in I_k} A_j(\mathbf{u})| \geqslant nb_k x) \leqslant 2 \exp\Big(- \frac{C_b b_k^2 n^2 x^2}{v^2 n + \bar{K}^2 + nb_k x \log^2 n}\Big)$$

Here

$$v^2 = \sup_{i>0}(\mathit{Var}(A_i(\mathbf{u})) + 2\sum_{j>i} |\mathit{Cov}(A_i(\mathbf{u}), A_j(\mathbf{u}))|),$$

C_b is a pure constant and $\bar{K}$ is defined as the upper bound of $|A_j(\mathbf{u})|$ with any $j \in [n]$.

By our conclusions from Lemma 14 and Lemma 15, we conclude there exists a constant $C_6' = (C_4' + 2C_5')$ such that $v^2 \leqslant C_6' b_n$, so we obtain

$$\begin{aligned}
\mathbb{P}(|Z(\mathbf{u})| \geqslant x) &\leqslant 2 \exp\Big(- \frac{C_b b_k^2 n^2 x^2}{C_6' nb_k + \bar{K}^2 + nb_k x \log^2 n}\Big) \\
&\leqslant 2 \exp\Big(- \frac{C_b nb_k x^2}{(C_6' + \bar{K}^2 + \log^2 n)(1 + x)}\Big)
\end{aligned}$$

The last inequality follows from our assumption that $b_k \geqslant 1/n = 1/|I_k|$ for any $k \geqslant 1$ in given Lemma 13. Further, we set $C_7' = C_b/(2C_6' + 2\bar{K}^2 + 2)$. Then we take the union bound over U_k, which gives

$$\begin{aligned}
\mathbb{P}(\sup_{\mathbf{u} \in U_k} |Z(\pi_M(\mathbf{u}))| \geqslant x) &\leqslant |S^{(M)}| \cdot \mathbb{P}(|Z(\mathbf{u})| \geqslant x) \\
&\leqslant 2 \cdot 2^M(\sqrt{n} + 1) \cdot (2^{M+1}\sqrt{d} + 1)^d \cdot e^{-\frac{C_7' nb_k}{\log^2 n} \min\{x, x^2\}}
\end{aligned}$$

$$\leqslant 2e^{(d+1)M\log 2+\log(\sqrt{n}+1)+d\log(2C(n,d)+2)-\frac{C_7' nb_k}{\log^2 n}\min\{x,x^2\}}.$$

Since we define $M = \lceil \frac{4}{\log 2}\log\frac{1}{b_k}\rceil$, then we choose

$$x(n,d) := \frac{\log n}{\sqrt{nb_k}}\sqrt{\Big[(d+1)4\log\frac{1}{b_k} + 2(d+1)\log 2 + \log(\sqrt{n}+1) + d\log(2C(n,d)+2) + \log\frac{8}{\delta}\Big]/C_7'}, \tag{56}$$

where $C(n,d) = \sqrt{(d+1)(W_x\log n + \log^2 n\log\log n)}$. We then have

$$\mathbb{P}(\sup_{\mathbf{u}\in U_k}|Z(\pi_M(\mathbf{u}))| \geqslant x(n,d)) \leqslant \frac{\delta}{4}.$$

when $\delta > 8\exp(-nb_k/(C_7'\log^2 n))$ and $nb_k \geqslant 2\log^2 n[(d+1)4\log\frac{1}{b_k} + 2(d+1)\log 2 + \log(\sqrt{n}+1) + d\log(2C(n,d)+2)]/C_7'$ (because under such conditions, we have $x(n,d)\leqslant 1$). Now, we proceed to bound the later term at the right hand side of (54). Similar with our cases stated in the proof of Lemma 2, for any $\mathbf{u}_1 := (u,\boldsymbol{\theta}_1), \mathbf{u}_2 := (s,\boldsymbol{\theta}_2) \in I\times\Theta_k$, we have that

$$Z(\mathbf{u}_1) - Z(\mathbf{u}_2) = Z(u,\boldsymbol{\theta}_1) - Z(s,\boldsymbol{\theta}_2) = \frac{1}{nb_k}\sum_{t\in I_k} B_t(u,\boldsymbol{\theta}_1,s,\boldsymbol{\theta}_2),$$

where

$$B_t(u,\boldsymbol{\theta}_1,s,\boldsymbol{\theta}_2) = y_t\left(K(\frac{w_t(\boldsymbol{\theta}_1)-u}{b_k}) - K(\frac{w_t(\boldsymbol{\theta}_2)-s}{b_k})\right) - \mathbb{E}y_t\left(K(\frac{w_t(\boldsymbol{\theta}_1)-t}{b_k}) - K(\frac{w_t(\boldsymbol{\theta}_2)-s}{b_k})\right).$$

We have $\mathbb{E}B_t(u,\boldsymbol{\theta}_1,s,\boldsymbol{\theta}_2) = 0$, and that

$$\begin{aligned}|Z(\mathbf{u}_1) - Z(\mathbf{u}_2)| = |B_t(u,\boldsymbol{\theta}_1,s,\boldsymbol{\theta}_2)| &\leqslant 2\left|y_j(K(\frac{w_t(\boldsymbol{\theta}_1)-u}{b_k}) - K(\frac{w_t(\boldsymbol{\theta}_2)-s}{b_k}))\right| \\ &\leqslant \frac{2l_K\sqrt{1+\max_{\mathbf{x}\in\mathcal{X}}\|\mathbf{x}\|_2^2+1}}{b_n}\cdot\|\mathbf{u}_1-\mathbf{u}_2\|_2 := \frac{C^*}{b_n}\|\mathbf{u}_1-\mathbf{u}_2\|_2.\end{aligned}$$

The last inequality follows from the Lipschitz property of $K(\cdot)$ and for simplicity we use C^* to denote the constant $2l_K\sqrt{\max_{\mathbf{x}\in\mathcal{X}}\|\mathbf{x}\|_2^2+2} = 2l_K\sqrt{R_{\mathcal{X}}^2+2}$. Then according to the Bernstein inequality given in Lemma 11, we have

$$\mathbb{P}(|\sum_{t=1}^n B_t(\mathbf{u}_1,\mathbf{u}_2)| \geqslant nb_k x) \leqslant 2\exp\left(-\frac{C_w n^2 b_k^2 x^2}{n\frac{C^{*2}\|\mathbf{u}_1-\mathbf{u}_2\|_2^2}{b_k^2} + nb_k x\frac{C^*\|\mathbf{u}_1-\mathbf{u}_2\|_2}{b_k}\log^2 n}\right).$$

Recall that $\forall\mathbf{u}\in U_n$, we have $\|\pi_i(\mathbf{u}) - \pi_{i+1}(\mathbf{u})\|_2 \leqslant \frac{\sqrt{4\delta_z^2+R_m^2}}{2^{i-1}\sqrt{n}}$. We then use the union bound to get

$$\mathbb{P}(\sup_{\mathbf{u}\in U_k} |Z(\pi_{i+1}(\mathbf{u})) - Z(\pi_i(\mathbf{u}))| \geqslant x)$$
$$\leqslant 2^{2i+2}(\sqrt{n}+1)^2(2^{i+2}C(n,d)+1)^{2d} \cdot 2e^{-\frac{C_8' 2^{i-1} n^{3/2} b_k^4 x^2}{(\frac{4\delta_z^2+R_m^2}{2^{i-1}\sqrt{n}} + b_k^2\sqrt{4\delta_z^2+R_m^2}\log^2 n)(1+x)}}$$

in which $C_8' = C_w/\max\{C^{*2}, C^*\}$. We let $x = \frac{\sqrt{(4\delta_z^2+R_m^2)/(2^{i-1}\sqrt{n}) + b_k^2\sqrt{4\delta_z^2+R_m^2}\log^2 n}}{2^{(i-1)/2}n^{3/4}b_k^2} \cdot \epsilon_i$. Then we have

$$\mathbb{P}\Bigg(\sup_{\mathbf{u}\in U_k} |Z(\pi_{i+1}(\mathbf{u})) - Z(\pi_i(\mathbf{u}))| \geqslant \sqrt{(4\delta_z^2+R_m^2)/(2^{i-1}\sqrt{n}) + b_k^2\sqrt{4\delta_z^2+R_m^2}\log^2 n}/(2^{(i-1)/2}n^{3/4}b_k^2)\cdot\epsilon_i\Bigg) \tag{57}$$

$$\leqslant 2^{2i+2}(\sqrt{n}+1)^2(2^{i+2}C(n,d)+1)^{2d} \cdot 2e^{-\frac{C_8'\epsilon_i^2}{1+\frac{\sqrt{(4\delta_z^2+R_m^2)/(2^{i-1}\sqrt{n})+b_k^2\sqrt{4\delta_z^2+R_m^2}\log^2 n}}{2^{(i-1)/2}n^{3/4}b_k^2}\cdot\epsilon_i}} \tag{58}$$

We observe that if we could choose ϵ_i such that

$$\frac{\sqrt{(4\delta_z^2+R_m^2)/(2^{i-1}\sqrt{n}) + b_k^2\sqrt{4\delta_z^2+R_m^2}\log^2 n}}{2^{(i-1)/2}n^{3/4}b_k^2} \cdot \epsilon_i < 1,$$

holds, then the right hand side of (58) satisfies

$$(58) \leqslant 2^{2i+2}(\sqrt{n}+1)^2(2^{i+2}C(n,d)+1)^{2d} \cdot 2e^{-\frac{C_8'\epsilon_i^2}{2}}. \tag{59}$$

Now we choose $\epsilon_i = \sqrt{[(4d+6)(i+1)\log 2 + 4\log(\sqrt{n}+1) + 4d\log(2C(n,d)+2) + 2\log(8/\delta)]/C_8'}$. Then we have

$$\frac{\sqrt{(4\delta_z^2+R_m^2)/(2^{i-1}\sqrt{n}) + b_k^2\sqrt{4\delta_z^2+R_m^2}\log^2 n}}{2^{(i-1)/2}n^{3/4}b_k^2} \cdot \epsilon_i$$
$$\leqslant \frac{1}{2^{(i-1)/2}n^{3/4}b_k}\Big[\frac{\sqrt{4\delta_z^2+R_m^2}}{2^{(i-1)/2}b_k n^{1/4}} + (4\delta_z^2+R_m^2)^{1/4}\log n\Big]\cdot\epsilon_i.$$

Here we only consider $i \geqslant M = \lceil \frac{4}{\log 2}\log\frac{1}{b_k}\rceil$, and we have $2^{M/4}\cdot b_k = 1$. In addition, we also get $\max_i (i+1)/2^{(i-2)/2} \leqslant 3$. Hence, we have

$$\frac{1}{2^{(i-1)/2}n^{3/4}b_k}\Big[\frac{\sqrt{4\delta_z^2+R_m^2}}{2^{(i-1)/2}b_k n^{1/4}} + (4\delta_z^2+R_m^2)^{1/4}\log n\Big]\cdot\epsilon_i < 1,$$

if $\delta \geqslant 8\exp(-C_8' n^{3/2}/(16(4\delta_z^2 + R_m^2)\log^2 n))$ and $n \geqslant \{8(4\delta_z^2 + R_m^2)\log^2 n \cdot [(12d + 18)\log 2 + 4\log(\sqrt{n} + 1) + 4d\log(2C(n,d) + 2)]/C_8'\}^{2/3}$. Then after plugging our setting of ϵ_i into (59), we obtain

$$\mathbb{P}\left(\sup_{\mathbf{u}\in U_k} |Z(\pi_{i+1}(\mathbf{u})) - Z(\pi_i(\mathbf{u}))| \geqslant \sqrt{(4\delta_z^2 + R_m^2)/(2^{i-1}\sqrt{n}) + b_k^2\sqrt{4\delta_z^2 + R_m^2}\log^2 n}/(2^{(i-1)/2}n^{3/4}b_k^2)\cdot\epsilon_i\right)$$
$$\leqslant \frac{1}{2^{i+1}}\cdot\frac{\delta}{4}.$$

And we notice

$$\begin{aligned}
&\sum_{i=M}^{\infty}\frac{\sqrt{(4\delta_z^2 + R_m^2)/(2^{i-1}\sqrt{n}) + b_k^2\sqrt{4\delta_z^2 + R_m^2}\log^2 n}}{2^{(i-1)/2}n^{3/4}b_k^2}\cdot\epsilon_i \\
&\quad\leqslant \sum_{i=M}^{\infty}\frac{\sqrt{4\delta_z^2 + R_m^2}}{2^{i-1}nb_k^2}\cdot\epsilon_i + \frac{\sqrt{(4\delta_z^2 + R_m^2)\log^2 n}}{2^{(i-1)/2}n^{3/4}b_k}\cdot\epsilon_i := \mathbf{I} + \mathbf{II}.
\end{aligned}$$

For term $\mathbf{I}$, we have

$$\begin{aligned}
\mathbf{I} &= \sum_{i=M}^{\infty}\frac{\sqrt{4\delta_z^2 + R_m^2}}{2^{i-1}nb_k^2}\cdot\sqrt{[(4d+6)(i+1)\log 2 + 4\log(\sqrt{n}+1) + 4d\log(2C(n,d)+2) + 2\log(8/\delta)]/C_8'} \\
&\leqslant \frac{\sqrt{(4\delta_z^2 + R_m^2)/C_8'}}{nb_k^2}\Bigg[\sqrt{(4d+6)\log 2}\sum_{i=M}^{\infty}\frac{i+1}{2^{i-1}} + \frac{\sqrt{4\log(\sqrt{n}+1) + 4d\log(2C(n,d)+2) + 2\log(8/\delta)}}{2^{M-2}}\Bigg] \\
&\leqslant \frac{\sqrt{(4\delta_z^2 + R_m^2)/C_8'}}{nb_k^2}\frac{2M}{2^{M-2}}\Big[\sqrt{(4d+6)\log 2} + \sqrt{4\log(\sqrt{n}+1)} + \sqrt{4d\log(2C(n,d)+2)} + \sqrt{2\log(8/\delta)}\Big] \\
&\leqslant \frac{\sqrt{(4\delta_z^2 + R_m^2)/C_8'}}{n}\frac{8M}{2^{M/2}}\frac{1}{2^{M/2}b_k^2}\Big[\sqrt{(4d+6)\log 2} + \sqrt{4\log(\sqrt{n}+1)} \\
&\qquad + \sqrt{4d\log(2C(n,d)+2)} + \sqrt{2\log(8/\delta)}\Big] \\
&\leqslant \frac{C_9'}{n}\Big[\sqrt{(4d+6)\log 2} + \sqrt{4\log(\sqrt{n}+1)} + \sqrt{4d\log(2C(n,d)+2)} + \sqrt{2\log(8/\delta)}\Big],
\end{aligned}$$

in which C_9' is a pure constant such that $C_9' = \sqrt{(4\delta_z^2 + R_m^2)/C_8'}\cdot\max_i(8i/2^{i/2}) = 16\sqrt{(4\delta_z^2 + R_m^2)/C_8'}$ and $C(n,d) \leqslant \sqrt{(d+1)(W_x\log n + \log^3 n)}$. Then we obtain

$$\begin{aligned}
\sqrt{4d\log(2C(n,d)+2)} &\leqslant \sqrt{4d\log\left(4\sqrt{(d+1)(W_x\log n + \log^3 n)}\right)} \\
&\leqslant \sqrt{4d\log\left(4\sqrt{2}\sqrt{(d+1)\max\{1, W_x\}\log^3 n}\right)}
\end{aligned}$$

$$\leqslant \sqrt{4d\log(4\sqrt{2})} + \sqrt{2d\log(\max\{W_x,1\}(d+1))} + \sqrt{6d\log n}. \tag{60}$$

Next, we are able to find a pure constant $C'_{10} = 6\sqrt{6}$ such that $\sqrt{(4d+6)\log 2} + \sqrt{4\log(\sqrt{n}+1)} + \sqrt{4d\log(2C(n,d)+2)} \leqslant 6\sqrt{6}\sqrt{(d+1)\log(\max\{W_x,1\}(d+1))\log n}$ as long as $n \geqslant 3$ according to (60). Thus, we finally achieve

$$\mathbf{I} \leqslant \frac{C'_{11}}{n}\Big(\sqrt{(d+1)\log(\max\{W_x,1\}(d+1))\log n} + \sqrt{2\log(8/\delta)}\Big),$$

where $C'_{11} = C'_{10} \cdot C'_9$. For term $\mathbf{II}$, we obtain

$$\begin{aligned}
\mathbf{II} &= \sum_{i=M}^{\infty} \frac{\sqrt{(4\delta_z^2 + R_m^2)\log^2 n/C'_8}}{2^{(i-1)/2} n^{3/4} b_k}[\sqrt{(4d+6)(i+1)\log 2} \\
&\quad + \sqrt{4\log(\sqrt{n}+1)} + \sqrt{4d\log(2C(n,d)+2)} + \sqrt{2\log(8/\delta)}] \\
&\leqslant \frac{\sqrt{(4\delta_z^2 + R_m^2)/C'_8}\log n}{n^{3/4} b_k}\Big[\sqrt{(4d+6)\log 2}\sum_{i=M}^{\infty}\frac{i+1}{2^{(i-1)/2}} \\
&\quad + \frac{\sqrt{4\log(\sqrt{n}+1) + 4d\log(2C(n,d)+2) + 2\log(8/\delta)}}{2^{(M-2)/2}}\Big] \\
&\leqslant \frac{\sqrt{(4\delta_z^2 + R_m^2)/C'_8}\log n}{n^{3/4}} \frac{8\sqrt{2}M}{2^{M/4}} \frac{1}{2^{M/4} b_k}\Big[\sqrt{(4d+6)\log 2} + \sqrt{4\log(\sqrt{n}+1)} \\
&\quad + \sqrt{4d\log(2C(n,d)+2)} + \sqrt{2\log(8/\delta)}\Big].
\end{aligned}$$

We are also able to find a pure constant C'_{12} such that $C'_{12} = \sqrt{(4\delta_z^2 + R_m^2)/C'_8}\max_i(8\sqrt{2}i/2^{i/4}) = 24\sqrt{(4\delta_z^2 + R_m^2)/C'_8}$ and $C'_{13} = C'_{10} \cdot C'_{12}$. Then we obtain

$$\mathbf{II} \leqslant \frac{C'_{13}\log n}{n^{3/4}}\Big(\sqrt{(d+1)\log(\max\{W_x,1\}(d+1))\log n} + \sqrt{2\log 8/\delta}\Big).$$

After combining our inequalities of $\mathbf{I}$ and $\mathbf{II}$, we obtain a union bound:

$$\begin{aligned}
\mathbb{P}\Big(\sup_{\mathbf{u}\in U_k}|Z(\mathbf{u}) - Z(\pi_M(\mathbf{u}))| \geqslant x_2(n,d) :&= \frac{C'_{14}\log n}{n^{3/4}}\Big(\sqrt{(d+1)\log(\max\{W_x,1\}(d+1))\log n} \\
&\quad + \sqrt{2\log(8/\delta)}\Big)\Big) \\
&\leqslant \sum_{i=M}^{\infty}\frac{1}{2^{i+1}}\frac{\delta}{4} \leqslant \frac{\delta}{4},
\end{aligned}$$

in which we choose $C'_{14} = 2\max\{C'_{11}, C'_{13}\}$. Then we get

$$\mathbb{P}\Big(\sup_{\mathbf{u}\in U_k} Z(\mathbf{u}) \geqslant x(n,d) + x_2(n,d)\Big) \leqslant \frac{\delta}{4} + \frac{\delta}{4} = \frac{\delta}{2}.$$

where the expression of $x(n,d)$ is given in (56). As a reminder, we have

$$x(n,d) := \frac{\log n}{\sqrt{nb_k}}\sqrt{\Big[(d+1)4\log\frac{1}{b_k} + 2(d+1)\log 2 + \log(\sqrt{n}+1) + d\log(2C(n,d)+2) + \log\frac{8}{\delta}\Big]/C_7'}.$$

We obtain there exist a universal constant $C_{15}' = 8/\sqrt{C_7'}$ such that

$$x(n,d) \leqslant \frac{C_{15}'\log n}{\sqrt{nb_k}}\bigg(\sqrt{(d+1)\log(\max\{W_x,1\}(d+1))\log n} + \sqrt{2\log\frac{8}{\delta}}\bigg).$$

Then we finally achieve

$$\mathbb{P}\bigg(\sup_{\mathbf{u}\in U_k} Z(\mathbf{u}) \geqslant \frac{C_{16}'\log n}{\sqrt{nb_k}}\bigg(\sqrt{(d+1)\log(\max\{W_x,1\}(d+1))\log n} + \sqrt{2\log\frac{8}{\delta}}\bigg)\bigg) = \frac{\delta}{2},$$

where we let $C_{16}' = 2\max\{C_{14}', C_{15}'\}$ and $C_{17}' = C_{16}\log(\max\{W_x, e\})$. Thus, $nb_k \geqslant 4C_{17}'^2\log^3 n[(d+1)\log(d+1)]$ and $\delta \geqslant 8\exp(-nb_k/(8C_{17}'^2\log^2 n))$ becomes a sufficient condition to make $x(n,d) + x_2(n,d)$ be smaller than 1. Following similar procedure, we are able to prove the same inequality for f_n, so we conclude our proof of Lemma 13. □

The remaining part of Lemma 7 only involves getting a uniform upper bound for $|r_{\boldsymbol{\theta}}(u) - r_{\boldsymbol{\theta}_0}(u)|$ and thus $|\widehat{r}_k(u,\boldsymbol{\theta}) - r_{\boldsymbol{\theta}_0}(u)|$ for any $\boldsymbol{\theta} \in \Theta_k$ and $u \in I$. Similar with the corresponding proof of Lemma 2, we have

$$\sup_{u\in I,\boldsymbol{\theta}\in\Theta_k} |r_{\boldsymbol{\theta}}(u) - r_{\boldsymbol{\theta}_0}(u)| \leqslant l_r R_{\mathcal{X}} \cdot \frac{2}{c_{\min}}\sqrt{\frac{(d+1)(6W_x^2\log n + 6W_x\log^2 n\log\log n)}{C_w n}}.$$

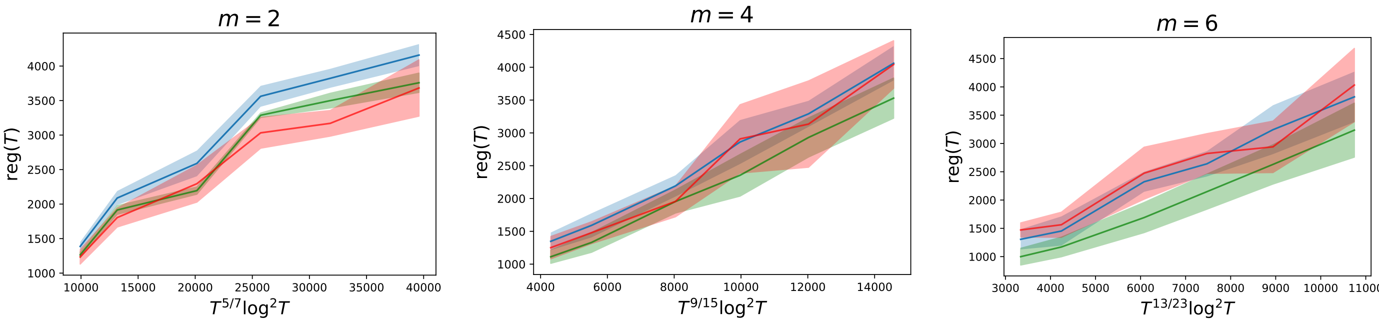
Finally, by setting $b_k = n^{-1/(2m+1)}$ and combining our results obtained in Lemma 12 and Lemma 13, we conclude our results for Lemma 7. In addition, our way of deriving constants $B_{mx,K}, B_{mx,K}'$ and $C_{mx,K}$ is similar with that in Lemma 7, so we omit the details here. □

F.8 Proof of Lemma 8 and Theorem 2

The proof of Lemma 8 and Theorem 2 are straight forward by combining the proof of Lemma 3 and Lemma 7, so we omit the details here.

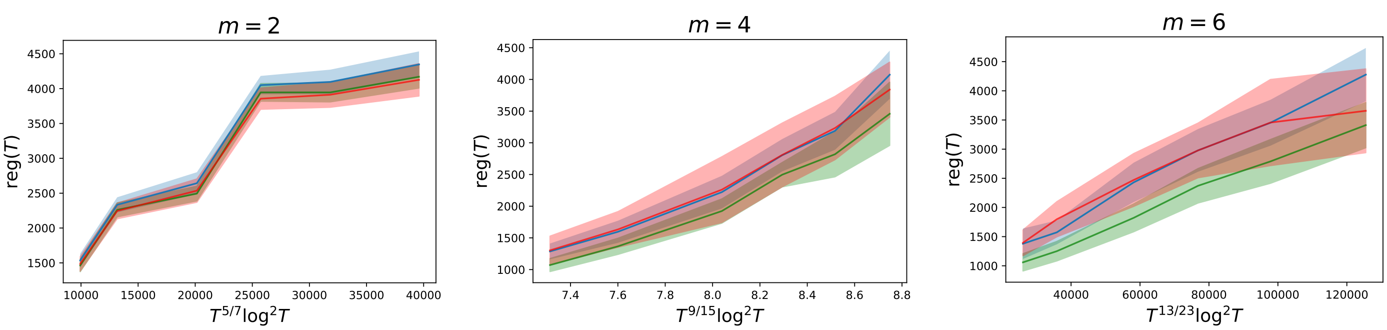
G Additional Plots

In this section, we directly plot reg(T) for all the settings discussed in the main paper. From Figure 9 - Figure 11, we see that the blue solid lines depicted in every figure are close to the other two lines that depict regrets with either known $\boldsymbol{\theta}_0$ or $g(\cdot)$ in Algorithm 1. This fact reflects the robustness of our estimators on $\boldsymbol{\theta}_0$ and $g(\cdot)$ in every episode.



(a) (b) (c)

Figure 9: From left to right, we plot empirical regret reg(T) against $T^{(2m+1)/(4m-1)}\log^2 T$ with $m \in [2, 4, 6]$ in the setting with i.i.d. covariates with independent entries. Solid blue, green, red lines, represent the mean regret collected by implementing the Algorithm 1 for 30 times with unknown $g(\cdot)$, $\boldsymbol{\theta}_0$, unknown $g(\cdot)$ but known $\boldsymbol{\theta}_0$ and known $g(\cdot)$ but unknown $\boldsymbol{\theta}_0$ in the exploitation phase respectively. Light color areas around those solid lines depict the standard error of our estimation of reg(T).



(a) (b) (c)

Figure 10: From left to right, we plot empirical regret reg(T) against $T^{(2m+1)/(4m-1)}\log^2 T$ with $m \in [2, 4, 6]$ in the setting with i.i.d. covariates but dependent entries. The rest caption is the same as in Figure 9.

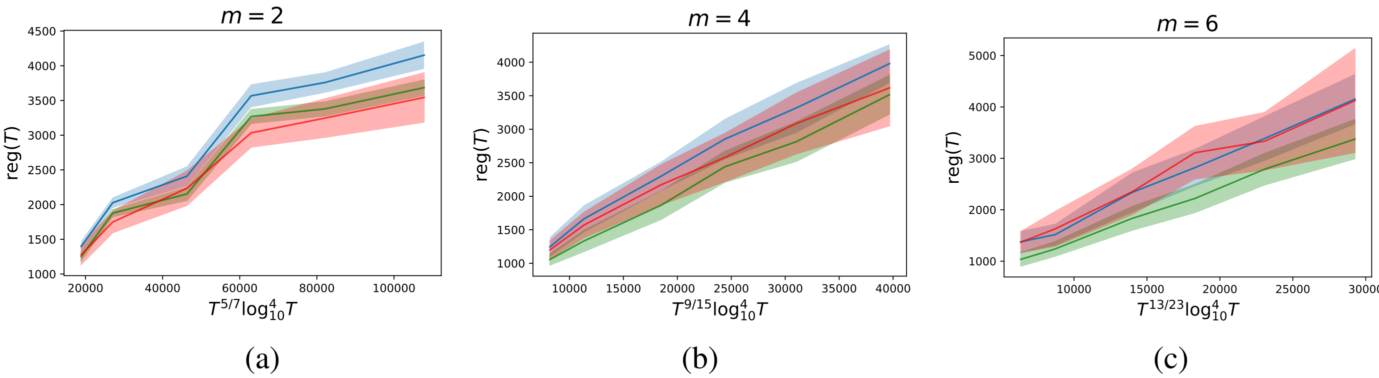


Figure 11: From left to right, we plot empirical regret $\mathrm{reg}(T)$ against $T^{(2m+1)/(4m-1)} \log_{10}^4 T$ with $m \in [2, 4, 6]$ in the setting with strong-mixing covariates. The rest caption is the same as in Figure 9.

H Regret bounds when $F(\cdot)$ is Lipschitz

All our main results require bounded second derivatives of F. This allows the pricing strategy $p_t = \widehat{\phi}_k^{-1}(-\mathbf{x}_t^\top \widehat{\boldsymbol{\theta}}) + \mathbf{x}_t^\top \widehat{\boldsymbol{\theta}}$ to achieve low regret if the revenue function has bounded second derivative. When $F(\cdot)$ is only Lipschitz continuous, the above method is no longer applicable. Fortunately, we can directly define the offered price based on the substitution of $\widehat{\boldsymbol{\theta}}$ and $\widehat{F}$ into (5) We summarize these in the following Algorithm 3.

Algorithm 3: Feature based dynamic pricing with unknown noise distribution when $F(\cdot)$ is ℓ-Lipschitz

1: **Input:** Upper bound of market value ($\{v_t\}_{t\geqslant 1}$): $B > 0$, minimum episode length: ℓ_0, degree of smoothness: $m = 0$.
2: **Initialization:** $p_1 = 0$, $\widehat{\boldsymbol{\theta}}_1 = 0$.
3: **for** each episode $k = 1, 2, \dots,$ **do**
4: Set length of the k-th episode $\ell_k = 2^{k-1}\ell_0$; Length of the exploration phase $a_k = \lceil (\ell_k d)^{\frac{3}{4}} \rceil$.
5: **Exploration Phase** ($t \in I_k := \{\ell_k, \cdots, \ell_k + a_k - 1\}$):
6: Offer price $p_t \sim \text{Unif}(0, B)$.
7: **Updating Estimates (at the end of the exploration phase with data** $\{(\widetilde{\mathbf{x}}_t, y_t)\}_{t\in I_k}$**):**
8: Update estimate of $\boldsymbol{\theta}_0$ by $\widehat{\boldsymbol{\theta}}_k = \widehat{\boldsymbol{\theta}}_k(\{(\widetilde{\mathbf{x}}_t, y_t)\}_{t\in I_k})$;

$$\widehat{\boldsymbol{\theta}}_k = \underset{\boldsymbol{\theta}}{\text{argmin}}\, L_k(\boldsymbol{\theta}) := \frac{1}{|I_k|}\sum_{t\in I_k}(By_t - \boldsymbol{\theta}^\top \boldsymbol{x}_t)^2 \tag{61}$$

9: Update estimates of F, by $F_k(u, \widehat{\boldsymbol{\theta}}_k) = F_k(u; \widehat{\boldsymbol{\theta}}_k, \{(\widetilde{\mathbf{x}}_t, y_t, p_t)\}_{t\in I_k})$ given by (14).
10: **Exploitation Phase** ($t \in I'_k := \{\ell_k + a_k, \cdots, \ell_{k+1} - 1\}$):
11: Offer p_t as

$$p_t = \text{argmax}_{p\geqslant 0}\{p(1 - \widehat{F}_k(p - \boldsymbol{x}_t^\top \widehat{\boldsymbol{\theta}}_k))\} \tag{62}$$

12: **end for**

Theorem 1. *Let Assumptions 1, 3, 4 and 5 hold. Then there exist constants C (depending only on the absolute constants within the assumptions) such that for all T satisfying $T \geqslant Cd$, the regret of Algorithm 3 over time T is no more than $C^*_{x,K}(Td)^{\frac{3}{4}}\log T(1 + \log T/d)$.*

Proof. We write

$$\mathbb{E}[R_t|\bar{\mathcal{H}}_{t-1}] = p_t^*(1 - F(p_t^* - \boldsymbol{x}_t^\top \boldsymbol{\theta}_0)) - p_t(1 - F(p_t - \boldsymbol{x}_t^\top \boldsymbol{\theta}_0)) \tag{63}$$
$$= \text{rev}_t(p_t^*, \boldsymbol{\theta}_0, F) - \text{rev}_t(p_t, \boldsymbol{\theta}_0, F). \tag{64}$$

The last inequality follows from our definition of (4). When $t \in I'_k$ (the k-th exploitation phase) we can then further expand (64) into

$$(64) = \text{rev}_t(p_t^*, \boldsymbol{\theta}_0, F) - \text{rev}_t(p_t^*, \widehat{\boldsymbol{\theta}}_k, F) \tag{65}$$
$$+ \text{rev}_t(p_t^*, \widehat{\boldsymbol{\theta}}_k, F) - \text{rev}_t(p_t^*, \widehat{\boldsymbol{\theta}}_k, \widehat{F}_k) \tag{66}$$
$$+ \text{rev}_t(p_t^*, \widehat{\boldsymbol{\theta}}_k, \widehat{F}_k) - \text{rev}_t(p_t, \widehat{\boldsymbol{\theta}}_k, \widehat{F}_k) \tag{67}$$
$$+ \text{rev}_t(p_t, \widehat{\boldsymbol{\theta}}_k, \widehat{F}_k) - \text{rev}_t(p_t, \widehat{\boldsymbol{\theta}}_k, F) \tag{68}$$

$$+ \text{rev}_t(p_t, \widehat{\boldsymbol{\theta}}_k, F) - \text{rev}_t(p_t, \boldsymbol{\theta}_0, F). \tag{69}$$

For (67), by our definition on p_t in (8) , we have

$$(67) = \text{rev}_t(p_t^*, \widehat{\boldsymbol{\theta}}_k, \widehat{F}_k) - \text{rev}_t(p_t, \widehat{\boldsymbol{\theta}}_k, \widehat{F}_k) \leqslant 0 \tag{70}$$

For terms (65) and (69), we can control both of them by difference between $\widehat{\boldsymbol{\theta}}_k$ and $\boldsymbol{\theta}_0$ in a sense that

$$(65), (69) \lesssim |\langle \boldsymbol{x}_t, \widehat{\boldsymbol{\theta}}_k - \boldsymbol{\theta}_0 \rangle| \lesssim \frac{1}{\sqrt{a_k}} \tag{71}$$

holds with high probability by our Lemma 4.1, since we assume F is Lipschitz continuous. Recall $a_k = |I_k|$, which is the length of the k-th exploration phase

For the rest two parts (66) and (68), as we are able to control $\widehat{F}(x)$ to $F(x)$ uniformly with rate $a_k^{-1/3}$ using data in the exploration phase. we can then bound $\mathbb{E}[R_t]$ by

$$\mathbb{E}[R_t] = \mathbb{E}[\mathbb{E}[R_t | \bar{\mathcal{H}}_{t-1}]] \lesssim \frac{1}{a_k^{1/3}} \tag{72}$$

Then for the regret in k-th episode we can bound it as

$$\text{Regret}_k = \sum_{t \in I_k} (\text{rev}_t^* - \text{rev}_t) + \sum_{t \in E_k \setminus I_k} (\text{rev}_t^* - \text{rev}_t) \tag{73}$$

$$\leqslant B a_k + a_k / a_k^{1/3} = a_k^{3/4} + a_k / a_k^{1/4} = \mathcal{O}(a_k^{3/4}) \tag{74}$$

let $K = \lfloor \log_2 T \rfloor + 1$, we have our total regret can be bounded by

$$\text{Regret}_\pi(T) = \sum_{k=1}^{K} 2^{3(k-1)/4} = \mathcal{O}(T^{3/4}). \tag{75}$$

□

I A Data Driven Way to Determine m

As mentioned in Remark 12, we are able to adopt the cross-validation method [Hall and Racine, 2015] to determine the order of smoothness m using data from the previous exploration phase. In the below, we briefly introduce how the order of smoothness can be determined in local polynomial regression in the context of nonparametric regression.

Given training data $\{x_i, y_i\}_{i=1}^n$ and we assume they are generated following model

$$Y = g^*(X) + \epsilon,$$

with $\mathbb{E}[\epsilon \mid X] = 0$. Define

$$\text{CV}(h, m) = \frac{1}{n} \sum_{i=1}^{n} (Y_i - \widehat{g}_{-i}(X_i))^2,$$

where $\widehat{g}_{-i}(\cdot)$ is fitted using all samples except the i-th pair (x_i, y_i). Here we use bandwidth h and m-th order local polynomial to fit the regression function. According to Theorem 3.2 given in Hall and Racine [2015], optimizing $\text{CV}(h, m)$ is equivalent to optimizing over (h, m) with respective to the averaged summed squared errors defined in (76) up to some small order terms.

$$\frac{1}{n} \sum_{i=1}^{n} (\widehat{g}(x_i) - g^*(x_i))^2. \tag{76}$$

Thus, this method is a valid way to determine the order of smoothness. We summarize the combined procedure in Algorithm 4.

Algorithm 4: Feature based dynamic pricing with unknown m

Input: Upper bound of market value ($\{v_t\}_{t\geqslant 1}$): $B>0$, minimum episode length: ℓ_0.

Initialization: $p_1=0$, $\widehat{\boldsymbol{\theta}}_1=0$.

for each episode $k=1,2,\ldots,$ **do**

  If $k\geqslant 2$, use $\{\mathbf{x}_t^\top\widehat{\boldsymbol{\theta}}_{k-1},p_t,y_t\}_{t\in I_{k-1}}$ and Algorithm 4 to determine m. If $k=1$, set $\widehat{m}=2$.

  Set length of the k-th episode $\ell_k=2^{k-1}\ell_0$; Length of the exploration phase $a_k=\lceil(\ell_k d)^{\frac{2\widehat{m}+1}{4\widehat{m}-1}}\rceil$.

  Exploration Phase ($t\in I_k:=\{\ell_k,\cdots,\ell_k+a_k-1\}$)**:**

    Offer price $p_t\sim \mathrm{Unif}(0,B)$.

  Updating Estimates (at the end of the exploration phase with data $\{(\widetilde{\mathbf{x}}_t,y_t)\}_{t\in I_k}$)**:**

    Update estimate of $\boldsymbol{\theta}_0$ by $\widehat{\boldsymbol{\theta}}_k=\widehat{\boldsymbol{\theta}}_k(\{(\widetilde{\mathbf{x}}_t,y_t)\}_{t\in I_k})$;

$$\widehat{\boldsymbol{\theta}}_k=\underset{\boldsymbol{\theta}}{\operatorname{argmin}}\, L_k(\boldsymbol{\theta}):=\frac{1}{|I_k|}\sum_{t\in I_k}(By_t-\boldsymbol{\theta}^\top\widetilde{\mathbf{x}}_t)^2. \tag{77}$$

    If $\widehat{m}\geqslant 1$, update estimates of F, F' by $F_k(u,\widehat{\boldsymbol{\theta}}_k)=F_k(u;\widehat{\boldsymbol{\theta}}_k,\{(\widetilde{\mathbf{x}}_t,y_t,p_t)\}_{t\in I_k},\widehat{h}_k)$, $F_k^{(1)}(u,\widehat{\boldsymbol{\theta}}_k)=F_k^{(1)}(u,\widehat{\boldsymbol{\theta}}_k,\{(\widetilde{\mathbf{x}}_t,y_t,p_t)\}_{t\in I_k},\widehat{h}_k)$. The detailed formulas are given by (14) and (16).

    Update estimate of ϕ by $\widehat{\phi}_k(u)=u-\frac{1-\widehat{F}_k(u)}{\widehat{F}^{(1)}(u)}$ and estimate of g by $\widehat{g}_k(u)=u+\widehat{\phi}_k^{-1}(-u)$.

    If $\widehat{m}=0$, update estimates of F, by $F_k(u,\widehat{\boldsymbol{\theta}}_k)=F_k(u;\widehat{\boldsymbol{\theta}}_k,\{(\widetilde{\mathbf{x}}_t,y_t,p_t)\}_{t\in I_k})$, The detailed formulas are given by (14).

  Exploitation Phase ($t\in I_k':=\{\ell_k+a_k,\cdots,\ell_{k+1}-1\}$)**:**

    If $\widehat{m}\geqslant 1$, offer p_t as

$$p_t=\min\{\max\{\widehat{g}_k(\widetilde{\mathbf{x}}_t^\top\widehat{\boldsymbol{\theta}}_k),0\},B\}. \tag{78}$$

    If $\widehat{m}=0$, offer p_t as

$$p_t=\operatorname{argmax}_{p\geqslant 0}\{p(1-\widehat{F}_k(p-\boldsymbol{x}_t^\top\widehat{\boldsymbol{\theta}}_k))\}.$$

end for

Algorithm 5: Selection of m.

Input: Data $\{\mathbf{x}_t^\top\widehat{\boldsymbol{\theta}}_{t-1},p_t,y_t\}_{t\in I_{k-1}}$

For $(m,h)\in\mathcal{M}\times\mathcal{H}$, compute:

$$(\widehat{m},\widehat{h})=\underset{(m,h)}{\operatorname{argmin}}\, L(m,h)=\frac{1}{|I_{k-1}|}\sum_{i=1}^{|I_{k-1}|}(Y_i-\widehat{g}_{-i}^{(m,h)}(X_i))^2$$

Output: $\widehat{m}$

References

M. Banna, F. Merlevède, and P. Youssef. Bernstein-type inequality for a class of dependent random matrices. *Random Matrices: Theory and Applications*, 05(02):1650006, 2016.

D. Bosq. *Nonparametric Statistics for Stochastic Processes: Estimation and Prediction*, volume 110. Springer Science & Business Media.

J. Broder and P. Rusmevichientong. Dynamic pricing under a general parametric choice model. *Operations Research*, 60(4):965–980, 2012.

R. J. Carroll, J. Fan, I. Gijbels, and M. P. Wand. Generalized partially linear single-index models. *Journal of the American Statistical Association*, 92(438):477–489, 1997. doi: 10.1080/01621459.1997.10474001.

R. Eubank and P. L. Speckman. Confidence bands in nonparametric regression. *Journal of the American Statistical Association*, 88(424):1287–1301, 1993.

J. Fan and I. Gijbels. *Local polynomial modelling and its applications*. Chapman and Hall, 1996.

J. Fan and R. Li. Variable selection via nonconcave penalized likelihood and its oracle properties. *Journal of the American statistical Association*, 96(456):1348–1360, 2001.

J. Fan, M. Farmen, and I. Gijbels. Local maximum likelihood estimation and inference. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 60(3):591–608, 1998.

P. Hall. *The bootstrap and Edgeworth expansion*. Springer-Verlag New York, 1992.

P. Hall and J. Horowitz. A simple bootstrap method for constructing nonparametric confidence bands for functions. *The Annals of Statistics*, 41(4):1892 – 1921, 2013.

P. G. Hall and J. S. Racine. Infinite order cross-validated local polynomial regression. *Journal of Econometrics*, 185(2):510–525, 2015. ISSN 0304-4076.

J. L. Horowitz. *The Bootstrap*, volume 5 of *Handbook of Econometrics*. Elsevier, 2001.

A. Javanmard and H. Nazerzadeh. Dynamic pricing in high-dimensions. *The Journal of Machine Learning Research*, 20(1):315–363, 2019.

F. Merlevède, M. Peligrad, and E. Rio. *Bernstein inequality and moderate deviations under strong mixing conditions*, volume Volume 5 of *Collections*. Institute of Mathematical Statistics, Beachwood, Ohio, USA, 2009. doi: 10.1214/09-IMSCOLL518.

M. H. Neumann and J. Polzehl. Simultaneous bootstrap confidence bands in nonparametric regression. *Journal of Nonparametric Statistics*, 9(4):307–333, 1998.

C. Stone. Optimal rates of convergence for nonparametric estimators. *The Annals of Statistics*, 8(6): 1348–1360, 1980.

C. J. Stone. Optimal Global Rates of Convergence for Nonparametric Regression. *The Annals of Statistics*, 10(4):1040 – 1053, 1982. doi: 10.1214/aos/1176345969.

A. B. Tsybakov. *Introduction to Nonparametric Estimation*. Springer Publishing Company, Incorporated, 2008. ISBN 0387790519.

R. Vershynin. *Introduction to the non-asymptotic analysis of random matrices*. Cambridge University Press, 2012.

R. Vershynin. *High-Dimensional Probability: An Introduction with Applications in Data Science*. Cambridge Series in Statistical and Probabilistic Mathematics. Cambridge University Press, 2018.

Y. Wang, X. Chen, X. Chang, and D. Ge. Uncertainty quantification for demand prediction in contextual dynamic pricing. *Production and Operations Management*, 30, 12 2020. doi: 10.1111/poms.13337.

C.-H. Zhang. Nearly unbiased variable selection under minimax concave penalty. *The Annals of Statistics*, 38(2):894– 942, 2010.

P. Zhao and B. Yu. On model selection consistency of lasso. *The Journal of Machine Learning Research*, 7:2541–2563, 2006.