Supplementary Notes

1 Preliminaries

The notations in this supplementary notes are same with main text if not defined particularly.

Definition 1 *Throughout our proof, we consider the parameter space*

$$\mathcal{U}_{s,m,M} = \left\{ \Theta \in \mathbb{R}_{p\times p} | \Theta \succ 0, \lambda_1(\Theta) > m, \|\Theta\|_2 \le \rho, \max_{j\in[p]} \|\Theta_j\|_0 \le s, \max_{j\in[p]} \|\Theta_j\|_1 \le M \right\},$$

where $0 < \lambda_1(\Theta) \le \lambda_2(\Theta) \le \ldots \le \lambda_p(\Theta)$ is the eigenvalues of Θ. A similar class of matrices were considered in the literature on inverse covariance matrix estimation[3][12]. We allow s to increase with n and p.

Proposition 1 *The right Epanechnikov kernel function* $K^+ = 1.5 \cdot (1-u^2) \cdot \mathbb{1}_{\{0\le u \le 1\}}$ *satisfies*

$$\int K^+(u)du = 1, \int u^l K^+(u)du < \infty, \int K^{+l} du < \infty,$$

for $l = 1, 2, 3, 4$. *Moreover,* $\|K^+\|_2^2 = \int K^{+2}(u)du = 6/5$, $\|K^+\|_\infty = \sup_u |K^+(u)| = 3.2$ *and* $\|K^+\|_{TV} = 1.5 \int_0^1 2udu = 3/2$. *The definition of left Epanechnikov kernel function is defined in the same way.*

Assumption 1 *Assume that there exists a constant* $\underline{f}_T$ *such that* $\inf_{t\in[0,1]} f_T(t) \ge \underline{f}_T > 0$. *Furthermore, assume that* f_T *is twice continuously differentiable and that there exists a constant* $\overline{f}_T < \infty$ *such that* $\max\{\|f_T\|_\infty, \|\dot{f}_T\|_\infty, \|\ddot{f}_T\|_\infty\} \le \overline{f}_T$.

Now, we give assumption 2. Different from assumption 2 in [12], to detect graph change, we only assume $\Sigma_{jk}(t)$ is right continuous in $[0, 1]$.

Assumption 2 *Assume* $\Sigma_{jk}(t)$ *is right continuous and* $\Sigma^{\pm}_{jk}(t)$, $\dot{\Sigma}^{\pm}_{jk}(t)$ *and* $\ddot{\Sigma}^{\pm}_{jk}(t)$ *exist for* $\forall t \in (0, 1)$, $j, k \in [p]$. $\Sigma^+_{jk}(0) = \Sigma_{jk}(0)$, $\Sigma^-_{jk}(1) = \Sigma_{jk}(1)$ *and there are finite number of discontinuities of* $\Sigma(t)$ *for* $t \in [0, 1]$. *There exists a constant* M_σ *such that*

$$\sup_{t\in[0,1]} \max_{j,k\in[p]} \max\{\Sigma^+_{jk}(t), \dot{\Sigma}^+_{jk}(t), \ddot{\Sigma}^+_{jk}(t), \Sigma^-_{jk}(t), \dot{\Sigma}^-_{jk}(t), \ddot{\Sigma}^-_{jk}(t)\} \le M_\sigma.$$

2 Notation

We define some notations. Let

$$\mathbb{P}_n[f] = \frac{1}{n} \sum_{i \in [n]} f(X_i)$$

and

$$\mathbb{G}_n[f] = \sqrt{n} \cdot (\mathbb{P}_n[f] - \mathbb{E}[f(X_i)]).$$

For notational convenience, for fixed $j, k \in [p]$, let

$$g_{t,jk}^{\pm}(T_i, X_{ij}, X_{ik}) = K_h^{\pm}(T_i - t) X_{ij} X_{ik}, \tag{1}$$

$$\omega_t^{\pm}(T_i) = K_h^{\pm}(T_i - t),$$

$$q_{t,jk}^{\pm}(T_i, X_{ij}, X_{ik}) = g_{t,jk}^{\pm}(T_i, X_{ij}, X_{ik}) - \mathbb{E}[g_{t,jk}^{\pm}(T, X_j, X_k)],$$

and let

$$k_t^{\pm}(T_i) = \omega_t^{\pm}(T_i) - \mathbb{E}[\omega_t^{\pm}(T)].$$

By the definition of $\hat{\Sigma}(t)^{\pm}$, we have

$$\hat{\Sigma}_{jk}^{\pm}(t) = \frac{\sum_{i \in [n]} g_{t,jk}^{\pm}(T_i, X_{ij}, X_{ik})}{\sum_{i \in [n]} \omega_t^{\pm}(T_i)} = \frac{\mathbb{P}_n[g_{t,jk}^{\pm}]}{\mathbb{P}_n[\omega_t^{\pm}]}.$$

Denote $\hat{\Theta}^{\pm}(t)$ as the CLIME estimator for $\Theta^{\pm}(t)$, we define

$$l_{t,jk}^{\pm}(T_i, X_i) = K_h^{\pm}(T_i - t) \left(X_i^T \hat{\Theta}_j^{\pm}(t) \right)^2 \left(X_i^T \hat{\Theta}_k^{\pm}(t) \right)^2$$

and

$$\hat{\Omega}_{jk}^{\pm}(t) = \frac{\sum_{i \in [n]} l_{t,jk}^{\pm}(T_i, X_i)}{\sum_{i \in [n]} \omega_t^{\pm}(T_i)} = \frac{\mathbb{P}_n[l_{t,jk}^{\pm}]}{\mathbb{P}_n[\omega_t^{\pm}]}.$$

In addition, let

$$J_{t,jk}^{\pm(1)}(T_i, X_i) = \sqrt{h} \cdot \left(\Theta_j^{\pm}(t) \right)^T \cdot \left[K_h^{\pm}(T_i - t) X_i X_i^T - \mathbb{E}\left[K_h^{\pm}(T - t) X X^T \right] \right] \cdot \Theta_k^{\pm}(t), \tag{2}$$

$$J_{t,jk}^{\pm(2)}(T_i) = \sqrt{h} \cdot \left(\Theta_j^{\pm}(t) \right)^T \cdot \left[K_h^{\pm}(T_i - t) - \mathbb{E}\left[K_h^{\pm}(T - t) \right] \right] \cdot \Sigma^{\pm}(t) \cdot \Theta_k^{\pm}(t), \tag{3}$$

$$J_{t,jk}^{\pm}(T_i, X_i) = J_{t,jk}^{\pm(1)}(T_i, X_i) - J_{t,jk}^{\pm(2)}(T_i), \tag{4}$$

and let

$$W_{t,jk}^{\pm}(T_i, X_{ij}, X_{ik}) = \sqrt{h} \cdot \left[K_h^{\pm}(T_i - t) X_{ij} X_{ik} - K_h^{\pm}(T_i - t) \Sigma_{jk}^{\pm}(t) \right]. \tag{5}$$

3 Theorem 1

Theorem 1 *Assume that* $h = o(1)$ *and that* $\log^2 np \cdot \log(np/\sqrt{h})/nh = o(1)$. *For any* $0 < a < b < 1$, *under assumptons 1-2, there exists a positive constant* C *such that*

$$\sup_{t\in[a,b]} \|\hat{\Sigma}^{\pm}(t) - \Sigma^{\pm}(t)\|_{max} \leq C \cdot \left(h + \sqrt{\frac{\log(np/\sqrt{h})}{nh}} \right)$$

with probability at least $1 - 3/np$, *for sufficiently large* n.

The proof of theorem 1 is inspired by the proof of theorem 1 in [12]. In [12], they upper bounded the quantity $\sup_{t\in[a,b]} \|\hat{\Sigma}^{\pm}(t) - \Sigma^{\pm}(t)\|_{max}$ by the summation of two terms: $\sup_{t\in[a,b]} \max_{j,k\in[p]} \left|\hat{\Sigma}^{+}_{jk}(t) - \mathbb{E}\left[\hat{\Sigma}^{+}_{jk}(t)\right]\right|$ and $\sup_{t\in[a,b]} \max_{j,k\in[p]} \left|\mathbb{E}\left[\hat{\Sigma}^{+}_{jk}(t)\right] - \Sigma^{+}_{jk}(t)\right|$, which are known as the variance and bias terms, respectively. To upper bound the variance term, they applied Talagrand's inequality[11] as well as the results of [9]. The bound they obtained for their estimator is $h^2 + \sqrt{\frac{\log(d/h)}{nh}}$. We generate their proof to fit in our model. To detect the change-point, we use right and left Epanechnikov kernel to estimate the covariance matrices so that the result in [12] is sharper due to the absence of first order term in 21 and the difference of the covering number of the function class. In corollary 2, we establish the consistency of $\Omega^{\pm}(t)$ with the same procedure to proof theorem 1. The upper bound of $\sup_{t\in[a,b]} \max_{j,k\in[p]} \left|\hat{\Omega}^{\pm}_{jk}(t) - \left(\Theta^{\pm}_{jj}(t)\Theta^{\pm}_{kk}(t) + 2\cdot\left(\Theta^{\pm}_{jk}(t)\right)^2\right)\right|$ is applied in lemma 5.1 to establish the consistency of nomalization term $\tilde{\sigma}^2_{jk}(t)$.

3.1 Thechnical Lemmas

Lemma 3.1 *Under the following conditions*

$$\left|\frac{\mathbb{G}_n[\omega^{\pm}_t]}{\sqrt{n}\cdot\mathbb{E}[\mathbb{P}_n[\omega^{\pm}_t]]}\right| < 1 \tag{6}$$

and

$$\mathbb{E}[\mathbb{P}_n[\omega^{\pm}_t]] \neq 0, \tag{7}$$

we have

$$\mathbb{E}[\hat{\Sigma}^{\pm}_{jk}(t)] = \frac{\mathbb{E}[\mathbb{P}_n[g^{\pm}_{t,jk}]]}{\mathbb{E}[\mathbb{P}_n[\omega^{\pm}_t]]} + \frac{1}{n}\mathcal{O}\left(\mathbb{E}[\mathbb{G}_n[\omega^{\pm}_t]\cdot\mathbb{G}_n[g^{\pm}_{t,jk}]] + \mathbb{E}[\mathbb{G}^2_n[g^{\pm}_{t,jk}]]\right).$$

Lemma 3.2 *Assume that* $h = o(1)$. *Under Assumptions 1-2, for any* $0 < a < b < 1$, *we have*

$$\sup_{t\in[a,b]} \max_{j,k\in[p]} \left|\mathbb{E}[\mathbb{P}_n[g^{\pm}_{t,jk}]] - f_T(t)\Sigma^{\pm}_{jk}(t)\right| = \mathcal{O}(h), \tag{8}$$

$$\sup_{t\in[a,b]} \left|\mathbb{E}[\mathbb{P}_n[\omega_t^{\pm}]] - f_T(t)\right| = \mathcal{O}(h), \tag{9}$$

$$\sup_{t\in[a,b]} \max_{j,k\in[p]} \frac{1}{n} \left|\mathbb{E}\left[\mathbb{G}_n[g_{t,jk}^{\pm}] \cdot \mathbb{G}_n[\omega_t^{\pm}]\right]\right| = \mathcal{O}\left(\frac{1}{nh}\right), \tag{10}$$

$$\sup_{t\in[a,b]} \frac{1}{n}\mathbb{E}[\mathbb{G}_n^2[\omega_t^{\pm}]] = \mathcal{O}\left(\frac{1}{nh}\right). \tag{11}$$

and

$$\sup_{t\in[a,b]} \max_{j,k\in[p]} \frac{1}{n}\mathbb{E}[\mathbb{G}_n^2[g_{t,jk}^{\pm}]] = \mathcal{O}\left(\frac{1}{nh}\right). \tag{12}$$

Lemma 3.3 *Assume that* $h = o(1)$ *and* $\log(np/\sqrt{h})/(nh) = o(1)$. *Under Assumptions 1-2, for sufficiently large* n *and* $0 < a < b < 1$*, there exists a universal constant* $C > 0$ *such that*

$$\sup_{t\in[a,b]} \left|\mathbb{G}_n[\omega_t^{\pm}]\right| \leq C \cdot \sqrt{\frac{\log(np/\sqrt{h})}{h}},$$

with probabiity at least $1 - 1/np$.

Lemma 3.4 *Assume that* $h = o(1)$ *and* $\log^2 np \cdot \log(np/\sqrt{h})/nh = o(1)$. *Under Assumptions 1-2, for sufficiently large* n *and* $0 < a < b < 1$*, there exists a universal constant* $C > 0$ *such that*

$$\sup_{t\in[a,b]} \max_{j,k\in[p]} \left|\mathbb{G}_n[g_{t,jk}^{\pm}]\right| \leq C \cdot \sqrt{\frac{\log(np/\sqrt{h})}{h}},$$

with probabiity at least $1 - 2/np$.

3.2 Proof of Theorem 1

For any a and b satisfying $0 < a < b < 1$, We have

$$\begin{aligned}\sup_{t\in[a,b]} \|\hat{\Sigma}^{+}(t) - \Sigma^{+}(t)\|_{max} &\leq \sup_{t\in[a,b]} \max_{j,k\in[p]} \left|\hat{\Sigma}_{jk}^{+}(t) - \mathbb{E}\left[\hat{\Sigma}_{jk}^{+}(t)\right]\right| + \sup_{t\in[a,b]} \max_{j,k\in[p]} \left|\mathbb{E}\left[\hat{\Sigma}_{jk}^{+}(t)\right] - \Sigma_{jk}^{+}(t)\right| \\ &= I_1 + I_2.\end{aligned}$$

It suffices to obtain upper bound for I_1 and I_2.

We first verify that the two conditions in 6 and 7 hold. By Lemma 3.2, we have

$$\left|\mathbb{E}\left[\mathbb{P}_n[\omega_t^{+}]\right]\right| = \mathcal{O}(h) + f_T(t) \geq \underline{f}_T(t) > 0,$$

where the last inequality follows from Assumption 1. Moreover,

$$\begin{aligned}\left|\frac{\mathbb{G}_n[\omega_t^+]}{\sqrt{n}\cdot\mathbb{E}\left[\mathbb{P}_n[\omega_t^+]\right]}\right| &\le C\cdot\frac{1}{\sqrt{n}}|\mathbb{G}_n[\omega_t^+]|\cdot\frac{1}{\underline{f}_T+\mathcal{O}(h)}\\ &\le C_1\cdot\sqrt{\frac{\log(np/\sqrt{h})}{nh}}\cdot\frac{1}{\underline{f}_T+\mathcal{O}(h)}\\ &< 1,\end{aligned}$$

for sufficiently lareg n, where the first inequality is obtained by an application of Lemma 3.2, the second inequality is obtained by an application of Lemma 3.3, and the last inequality is obtained by the scaling assumptions $h=o(1)$ and $\log(np/\sqrt{h})/(nh)=o(1)$.

Upper bound for I_1: By the proof of Lemma 3.1 in [12], we have

$$\hat{\Sigma}_{jk}^+(t)=\frac{\mathbb{G}_n[g_{t,jk}^+]}{\sqrt{n}\cdot\mathbb{E}[\mathbb{P}_n[\omega_t^+]]}+\frac{\mathbb{E}[\mathbb{P}_n[g_{t,jk}^+]]}{\mathbb{E}[\mathbb{P}_n[\omega_t^+]]}-\frac{\mathbb{G}_n[\omega_t^+]\cdot\mathbb{E}[\mathbb{P}_n[g_{t,jk}^+]]}{\sqrt{n}\cdot\mathbb{E}^2[\mathbb{P}_n[\omega_t^+]]}+\mathcal{O}\left(\left[\mathbb{G}_n[\omega_t^+]\cdot\mathbb{G}_n[g_{t,jk}^+]\right]+\mathbb{G}_n^2[g_{t,jk}^+]\right).$$

Thus, by Lemma 3.1, we have

$$\begin{aligned}I_1&=\sup_{t\in[a,b]}\max_{j,k\in[p]}\left|\frac{\mathbb{G}_n[g_{t,jk}^+]}{\sqrt{n}\cdot\mathbb{E}[\mathbb{P}_n[\omega_t^+]]}-\frac{\mathbb{G}_n[\omega_t^+]\cdot\mathbb{E}[\mathbb{P}_n[g_{t,jk}^+]]}{\sqrt{n}\cdot\mathbb{E}^2[\mathbb{P}_n[\omega_t^+]]}+I_{13}\right|\\ &\le\sup_{t\in[a,b]}\max_{j,k\in[p]}\{|I_{11}|+|I_{12}|+|I_{13}|\},\end{aligned}$$

where $I_{13}=\mathcal{O}\left(\left[\mathbb{G}_n[\omega_t^+]\cdot\mathbb{G}_n[g_{t,jk}^+]\right]+\mathbb{G}_n^2[g_{t,jk}^+]+\mathbb{E}[\mathbb{G}_n[\omega_t^+]\cdot\mathbb{G}_n[g_{t,jk}^+]]+\mathbb{E}[\mathbb{G}_n^2[g_{t,jk}^+]]\right)/n$.

We now provide upper bounds for I_{11}, I_{12}, and I_{13}. By an application of Lemmas 3.2, 3.3 and 3.4, we obtain

$$\sup_{t\in[a,b]}\max_{j,k\in[p]}|I_{11}|\le n^{-1/2}\cdot\sup_{t\in[a,b]}\max_{j,k\in[p]}\left|\frac{\mathbb{G}_n[g_{t,jk}^+]}{\underline{f}_T+\mathcal{O}(h)}\right|\le C\cdot\sqrt{\frac{\log(np/\sqrt{h})}{nh}}. \tag{13}$$

Similarly, we have

$$\sup_{t\in[a,b]}\max_{j,k\in[p]}|I_{12}|\le n^{-1/2}\cdot\sup_{t\in[a,b]}\max_{j,k\in[p]}\left|\frac{\mathbb{G}_n[g_{t,jk}^+](\overline{f}_T\Sigma_{jk}^+(t)+\mathcal{O}(h))}{(\underline{f}_T+\mathcal{O}(h))^2}\right|\le C\cdot\sqrt{\frac{\log(np/\sqrt{h})}{nh}}. \tag{14}$$

For I_{13}, we have

$$\begin{aligned}\sup_{t\in[a,b]}\max_{j,k\in[p]}|I_{13}|&\le\sup_{t\in[a,b]}\max_{j,k\in[p]}\left|\frac{1}{n}\mathcal{O}\left(\left[\mathbb{G}_n[\omega_t^+]\cdot\mathbb{G}_n[g_{t,jk}^+]\right]+\mathbb{G}_n^2[g_{t,jk}^+]\right)\right|+\mathcal{O}\left(\frac{1}{nh}\right)\\ &\le C\cdot\frac{\log(np/\sqrt{h})}{nh}+\mathcal{O}\left(\frac{1}{nh}\right)\\ &\le C\cdot\frac{\log(np/\sqrt{h})}{nh}.\end{aligned} \tag{15}$$

Combining 13,14 and 15, we have

$$I_1 \leq C \cdot \sqrt{\frac{\log(np/\sqrt{h})}{nh}}, \tag{16}$$

with probability at least $1 - 3/np$.

Upper bound for I_2: By Lemmas 3.1 and 3.2, we have

$$\begin{aligned}
I_2 &= \sup_{t\in[a,b]} \max_{j,k\in[p]} \left| \frac{\mathbb{E}[\mathbb{P}_n[g^+_{t,jk}]]}{\mathbb{E}[\mathbb{P}_n[\omega^+_t]]} - \Sigma^+_{jk}(t) + \frac{1}{n}\mathcal{O}\left(\mathbb{E}[\mathbb{G}_n[\omega^+_t]\cdot\mathbb{G}_n[g^+_{t,jk}]] + \mathbb{E}[\mathbb{G}^2_n[g^+_{t,jk}]]\right) \right| \\
&\leq \sup_{t\in[a,b]} \max_{j,k\in[p]} \left| \frac{f_T(t)\Sigma^+_{jk}(t) + \mathcal{O}(h)}{f_T(t) + \mathcal{O}(h)} - \Sigma^+_{jk}(t) + \frac{1}{n}\mathcal{O}\left(\mathbb{E}[\mathbb{G}_n[\omega^+_t]\cdot\mathbb{G}_n[g^+_{t,jk}]] + \mathbb{E}[\mathbb{G}^2_n[g^+_{t,jk}]]\right) \right| \\
&= \sup_{t\in[a,b]} \max_{j,k\in[p]} \left| \frac{f_T(t)\Sigma^+_{jk}(t) + \mathcal{O}(h)}{f_T(t) + \mathcal{O}(h)} - \Sigma^+_{jk}(t) + \mathcal{O}(\frac{1}{nh}) \right| \\
&= \sup_{t\in[a,b]} \max_{j,k\in[p]} \left| \frac{\mathcal{O}(h)(1-\Sigma^+_{jk}(t))}{f_T(t) + \mathcal{O}(h)} + \mathcal{O}\left(\frac{1}{nh}\right) \right| \\
&\leq C \cdot \left(h + \frac{1}{nh}\right),
\end{aligned} \tag{17}$$

where the first inequality follows from 8 and 9, the second equality follows from 10 and 11, and the last inequality follows from the assumption that $h = o(1)$.

Combining the upper bounds 16 and 17, we obtain

$$\sup_{t\in[a,b]} \|\hat{\Sigma}^+(t) - \Sigma^+(t)\|_{max} \leq C \cdot \left(h + \sqrt{\frac{\log(np/\sqrt{h})}{nh}}\right)$$

with probability at least $1 - 3/np$.

3.3 Corollary of Theorem 1

Given Theorem 1, the following corollary establishes the uniform rates of convergence for $\hat{\Theta}^\pm(t)$ using the CLIME estimator as defined in [2]. It follows directly from the proof of Theorem 6 in [2].

Corollary 1 *For any $0 < a < b < 1$, suppose that $\Theta^\pm(t) \in \mathcal{U}_{s,m,M}$ for all $t \in [a,b]$. Under the same conditions in Theorem 1, there exists a constant $C > 0$ such that if $\lambda \geq C \cdot M \cdot (h + \sqrt{\log(np/\sqrt{h})/nh})$, we have*

$$\sup_{t\in[a,b]} \|\hat{\Theta}^\pm(t) - \Theta^\pm(t)\|_{max} \leq C \cdot M^2 \cdot \left(h + \sqrt{\frac{\log(np/\sqrt{h})}{nh}}\right);$$

$$\sup_{t\in[a,b]}\max_{j\in[p]}\|\hat{\Theta}_j^{\pm}(t)-\Theta_j^{\pm}(t)\|_1 \le C\cdot M\cdot s\cdot\left(h+\sqrt{\frac{\log(np/\sqrt{h})}{nh}}\right);$$

$$\sup_{t\in[a,b]}\max_{j\in[p]}\|\hat{\Sigma}^{\pm}(t)\hat{\Theta}_j^{\pm}(t)-e_j\|_\infty \le C\cdot M\cdot\left(h+\sqrt{\frac{\log(np/\sqrt{h})}{nh}}\right)$$

with probability at least 1- 3/np for sufficiently large n. Please notice that λ here is the parameter of CLIME rather not the eigenvalues of Θ.

Following directly from the proof of Theorem 1, corollary 1 establishes the uniform rate of convergence for $\hat{\Omega}^{\pm}(t)$ under the maximum norm.

Corollary 2 *Assume that $h = o(1)$ and $\log^4 np\cdot\log(np/\sqrt{h})/nh = o(1)$. Under the same conditions in Corollary 1, there exists a constant $C > 0$ such that*

$$\sup_{t\in[a,b]}\max_{j,k\in[p]}\left|\hat{\Omega}_{jk}^{\pm}(t)-\left(\Theta_{jj}^{\pm}(t)\Theta_{kk}^{\pm}(t)+2\cdot\left(\Theta_{jk}^{\pm}(t)\right)^2\right)\right| \le C\cdot\left(h+\sqrt{\frac{\log(np/\sqrt{h})}{nh}}\right)$$

with probability at least $1-4/np$ for sufficiently large n.

4 Proof of Technical Lemmas in Theorem 1

4.1 Proof of Lemma 3.1

The proof of Lemma 3.1 directly follows from the proof of **Lemma 1** in [12].

4.2 Proof of Lemma 3.2

Proof of 8: Here, we only give the proof of the results for right Epanechnikov kernel function. The results for left Epanechnikov kernel function can be proved in the same way. We have

$$\begin{aligned}
\mathbb{E}[\mathbb{P}_n[g_{t,jk}^+]] &= \mathbb{E}\left[\frac{1}{h}K^+\left(\frac{T-t}{h}\right)X_jX_k\right] \\
&= \mathbb{E}\left[\frac{1}{h}K^+\left(\frac{T-t}{h}\right)\mathbb{E}[X_jX_k|T]\right] \\
&= \mathbb{E}\left[\frac{1}{h}K^+\left(\frac{T-t}{h}\right)\Sigma_{jk}(T)\right] \\
&= \int\frac{1}{h}K^+\left(\frac{T-t}{h}\right)\Sigma_{jk}(T)f_T(T)dT \\
&= \int K^+(u)\Sigma_{jk}(uh+t)f_T(uh+t)du. \qquad (18)
\end{aligned}$$

For $\forall t \in [a,b]$ and sufficiently small h, $\Sigma_{jk}(uh+t)$ is twice differentiable for $u \in (0,1]$. So, we can apply Taylor expansions to $\Sigma_{jk}(uh+t)$ and $f_T(uh+t)$. We have

$$\Sigma_{jk}(uh+t) = \Sigma^+_{jk}(t) + uh \cdot \dot{\Sigma}^+_{jk}(t) + \frac{1}{2}u^2h^2 \cdot \ddot{\Sigma}^+_{jk}(t') \tag{19}$$

and

$$f_T(uh+t) = f_T(t) + uh \cdot \dot{f}_T(t) + \frac{1}{2}u^2h^2 \cdot \ddot{f}_T(t''), \tag{20}$$

where t' and t'' are between t and $uh+t$. Substituting 19 and 20 into 18, we have

$$\int K^+(u)\left(\Sigma^+_{jk}(t) + uh \cdot \dot{\Sigma}^+_{jk}(t) + \frac{1}{2}u^2h^2 \cdot \ddot{\Sigma}^+_{jk}(t')\right)\left(f_T(t) + uh \cdot \dot{f}_T(t) + \frac{1}{2}u^2h^2 \cdot \ddot{f}_T(t'')\right) du \tag{21}$$

By assumptions 1 and 2, we have

$$h\Sigma^+_{jk}(t)\dot{f}_T(t)\int uK^+(u)du \leq hCM_\sigma \overline{f}_T = \mathcal{O}(h) \tag{22}$$

$$h\dot{\Sigma}^+_{jk}(t)f_T(t)\int uK^+(u)du \leq hCM_\sigma \overline{f}_T = \mathcal{O}(h) \tag{23}$$

Substituting 22 and 23 into 21 and bounding the other higher-order terms by $\mathcal{O}(h)$, we obtain

$$\mathbb{E}[\mathbb{P}_n[g^+_{t,jk}]] = f_T(t)\Sigma^+_{jk}(t) + \mathcal{O}(h),$$

for all $t \in [a,b]$ and $j,k \in [p]$. This implies that

$$\sup_{t\in[a,b]}\max_{j,k\in[p]}\left|\mathbb{E}[\mathbb{P}_n[g^+_{t,jk}]] - f_T(t)\Sigma^+_{jk}(t)\right| = \mathcal{O}(h).$$

The proof of 9 follows from the same set of argument.

Proof of 10: We have

$$\begin{aligned}
&\frac{1}{n}\mathbb{E}\left[\mathbb{G}_n[g^+_{t,jk}] \cdot \mathbb{G}_n[\omega^+_t]\right] \\
&= \mathbb{E}\left[\mathbb{P}_n[g^+_{t,jk}] \cdot \mathbb{P}_n[\omega^+_t]\right] - \mathbb{E}\left[\mathbb{P}_n[g^+_{t,jk}]\right] \cdot \mathbb{E}\left[\mathbb{P}_n[\omega^+_t]\right] \\
&= \mathbb{E}\left[\left(\frac{1}{n}\sum_{i\in[n]} K^+_h(T_i - t)X_{ij}X_{ik}\right) \cdot \left(\frac{1}{n}\sum_{i\in[n]} K^+_h(T_i - t)\right)\right] - \mathbb{E}\left[\mathbb{P}_n[g^+_{t,jk}]\right] \cdot \mathbb{E}\left[\mathbb{P}_n[\omega^+_t]\right] \\
&= \frac{1}{n}\mathbb{E}[K^{+2}_h(T-t)X_jX_k] + \frac{1}{n^2}\mathbb{E}\left[\sum_{i\in[n]}\sum_{i'\neq i} K^+_h(T_i - t)K^+_h(T_{i'} - t)X_{ij}X_{ik}\right] - \mathbb{E}\left[\mathbb{P}_n[g^+_{t,jk}]\right] \cdot \mathbb{E}\left[\mathbb{P}_n[\omega^+_t]\right] \\
&= \frac{1}{n}\mathbb{E}[K^{+2}_h(T-t)\Sigma_{jk}(T)] + \frac{n-1}{n}\left(\mathbb{E}\left[K^+_h(T-t)\right] \cdot \mathbb{E}\left[K^+_h(T-t)\Sigma_{jk}(T)\right]\right) - \mathbb{E}\left[\mathbb{P}_n[g^+_{t,jk}]\right] \cdot \mathbb{E}\left[\mathbb{P}_n[\omega^+_t]\right] \\
&= \frac{1}{n}\mathbb{E}[K^{+2}_h(T-t)\Sigma_{jk}(T)] - \frac{1}{n}\mathbb{E}\left[\mathbb{P}_n[g^+_{t,jk}]\right] \cdot \mathbb{E}\left[\mathbb{P}_n[\omega^+_t]\right],
\end{aligned}$$

where the second to the last equality follows from the fact that T_i and $T_{i'}$ are independent. By proposition 1 and Assumptions 1-2, we have

$$\frac{1}{n}\mathbb{E}[K_h^{+2}(T-t)\Sigma_{jk}(T)] = \frac{1}{nh}\int \frac{1}{h}K^{+2}\left(\frac{T-t}{h}\right)\Sigma_{jk}(T)f_T(t)dT$$
$$\leq \frac{1}{nh}M_\sigma \overline{f}_T \int \frac{1}{h}K^{+2}\left(\frac{T-t}{h}\right)dT = \mathcal{O}\left(\frac{1}{nh}\right),$$

where the last equality holds by a change of variable. Moreover, by 8 and 9, we have

$$\frac{1}{n}\mathbb{E}\left[\mathbb{P}_n[g_{t,jk}^+]\right]\cdot\mathbb{E}\left[\mathbb{P}_n[\omega_t^+]\right] = \frac{1}{n}\left(f_T(t)\Sigma_{jk}^+(t)+\mathcal{O}(h)\right)\cdot(f_T(t)+\mathcal{O}(h)) = \mathcal{O}\left(\frac{1}{n}\right).$$

Taking the supreme over $t \in [a,b]$ and $j,k \in [p]$ on both sides of the eqution, we obtain

$$\sup_{t\in[a,b]}\max_{j,k\in[p]}\left|\mathbb{E}\left[\mathbb{G}_n[g_{t,jk}^+]\cdot\mathbb{G}_n[\omega_t^+]\right]\right| = \mathcal{O}\left(\frac{1}{nh}\right)+\mathcal{O}\left(\frac{1}{n}\right) = \mathcal{O}\left(\frac{1}{nh}\right),$$

where the last equality holds by the scaling assumption of $h = o(1)$. The proof of 11 and 12 follows from the same set of argument.

4.3 Proof of Lemma 3.3

Lemma 3.3 and 3.4 provide upper bounds for the supreme of the empirical processes $\mathbb{G}_n[\omega_t^\pm]$ and $\mathbb{G}_n[g_{t,jk}^\pm]$, respectively. To this end, we apply the Talagrand's inequality [11] in Lemma 8.1. Let $\mathcal{F}$ be a function class. In order to apply Talagrand's inequality, we need to evaluate the quantities η and τ^2 such that

$$\sup_{f\in\mathcal{F}}\|f\|_\infty \leq \eta$$

and

$$\sup_{f\in\mathcal{F}} Var(f(X)) \leq \tau^2.$$

Talagrand's inequality in Lemma 8.1 provides an upper bound for the supreme of an empirical process in terms of its expectation. By Lemma 8.2, the expectation can then be upper bounded as a function of the covering number of the function class $\mathcal{F}$, denoted as $N(\mathcal{F}, L_2(Q), \epsilon)$.

The proof of Lemma 3.3 uses the set of arguments as detailed in appendix **E.3.1** of [12]. Recall the definition that $\omega_t^+(T_i) = K_h^+(T_i - t)$ and $k_t^+(T_i) = \omega_t^+(T_i) - \mathbb{E}[\omega_t^+(T)]$, respectively. We consider the class of functions

$$\mathcal{K}^+ = \{k_t^+ | t \in [a,b]\}.$$

First, note that

$$\begin{aligned}\sup_{t\in[a,b]} \|k_t^+\|_\infty &= \sup_{t\in[a,b]} \|\omega_t^+(T_i) - \mathbb{E}[\omega_t^+(T)]\|_\infty \\ &\le \frac{1}{h}\|K^+\|_\infty + \overline{f}_T + \mathcal{O}(h) \\ &\le \frac{2}{h}\|K^+\|_\infty \end{aligned} \tag{24}$$

where the first inequality holds by Proposition 1 and Lemma 3.2, and the last inequality holds by the scaling assumption $h = o(1)$ for sufficiently large n.

Next, we obtain an upper bound for the variance of $k_t(T_i)$. Note that

$$\begin{aligned}\sup_{t\in[a,b]} Var(k_t^+(T)) &= \sup_{t\in[a,b]} \mathbb{E}\left[(\omega_t^+(T) - \mathbb{E}[\omega_t^+(T)])^2\right] \\ &\le \sup_{t\in[a,b]} 2\mathbb{E}\left[\omega_t^{+2}(T)\right] + \sup_{t\in[a,b]} 2\mathbb{E}^2\left[\omega_t^+(T)\right],\end{aligned}$$

where we apply the inequality $(x-y)^2 \le 2x^2 + 2y^2$ for two scalars x, y. By Lemma 3.2, we have $\sup_{t\in[a,b]} 2\mathbb{E}^2\left[\omega_t^+(T)\right] \le 2(\overline{f}_T + \mathcal{O}(h))^2$. Also, by a change of variable and second-order Taylor expansion on the marginal density $f_T(\cdot)$, we have

$$\begin{aligned}\sup_{t\in[a,b]} 2\mathbb{E}\left[\omega_t^{+2}(T)\right] &= 2\sup_{t\in[a,b]} \int \frac{1}{h^2} K^{+2}\left(\frac{T-t}{h}\right) f_T(T)dT \\ &= 2\sup_{t\in[a,b]} \frac{1}{h}\int K^{+2}(u) f_T(uh+t)du \qquad (25) \\ &= 2\sup_{t\in[a,b]} \frac{1}{h}\int K^{+2}(u)\left(f_T(t) + uh\dot{f}_T(t) + \frac{1}{2}u^2h^2\ddot{f}_T(t')\right) \qquad (26) \\ &\le \frac{2}{h}\overline{f}_T\|K^+\|_2^2 + \mathcal{O}(1) + \mathcal{O}(h), \qquad (27)\end{aligned}$$

where $t' \in (t, t+uh)$. Thus for sufficiently large n and the assumption that $h = o(1)$, we have

$$\sup_{t\in[a,b]} Var(k_t^+(T)) \le \frac{3}{h} \cdot \overline{f}_T \cdot \|K^+\|_2^2. \tag{28}$$

By Lemma 7.4, the covering number for the function class $\mathcal{K}^+$ satisfies

$$\sup_Q N(\mathcal{K}^+, L_2(Q), \epsilon) \le \left(\frac{4\cdot\|K^+\|_{TV}\cdot C_{K^+}^{4/5}\cdot \overline{f}_T^{1/5}}{h\epsilon}\right)^5.$$

We are now ready to obtain an upper bound for the supreme of the empirial process, $\sup_{t\in[a,b]} |\mathbb{G}_n[\omega_t^+]|$. By Lemma 8.2 with $A = 2\cdot\|K^+\|_{TV}\cdot C_{K^+}^{4/5}\cdot\overline{f}_T^{1/5}/\|K^+\|_\infty$, $U = \|F\|_{L_2(\mathbb{P}_n)} = 2\cdot\|K^+\|_\infty/h$, $V = 5$,

$\sup_{t\in[a,b]} \mathbb{E}k_t^{+2} = \sup_{t\in[a,b]} Var(k_t^+(T)) \le \sigma_P^2 = 3\cdot \overline{f}_T \cdot \|K^+\|_2^2/h \le (2\cdot \|K^+\|_\infty/h)^2 = \|F\|^2_{L_2(\mathbb{P}_n)}$, for sufficiently large n, we obtain

$$\begin{aligned}
\mathbb{E}\left[\sup_{t\in[a,b]} \frac{1}{\sqrt{n}}\cdot|\mathbb{G}_n[\omega_t^+|\right] &= \mathbb{E}\left[\sup_{t\in[a,b]} \frac{1}{n}\left|\sum_{i\in[n]}(\omega_t^+(T_i) - \mathbb{E}[\omega_t^+(T)])\right|\right] \\
&\precsim \sqrt{\frac{\log(\sqrt{1/h})}{nh}} + \frac{\log(\sqrt{1/h})}{nh} \\
&\sim \sqrt{\frac{\log(\sqrt{1/h})}{nh}},
\end{aligned}$$

where $x \precsim y$ means there exists a constant $C > 0$ such that $x \le C\cdot y$ and $x \sim y$ means there exist $C_1, C_2 > 0$ such that $C_1\cdot y \ge x \ge C_2\cdot y$. The last expression holds by the assumption that $\log(p/\sqrt{h})/(nh) = o(1)$. By Lemma 8.1 with $\tau^2 = 3\overline{f}_T\cdot\|K^+\|_2^2/h$, $\eta = 2\cdot\|K^+\|_\infty/h$, $\mathbb{E}[Y] \le C\cdot\sqrt{\log(\sqrt{1/h})/nh}$ and picking $t = \sqrt{\log(np)/n}$, for sufficiently large n, we have

$$\begin{aligned}
\sup_{t\in[a,b]} \frac{1}{\sqrt{n}}\cdot|\mathbb{G}_n[\omega_t^+]| &= \sup_{t\in[a,b]} \frac{1}{n}\left|\sum_{i\in[n]}(\omega_t^+(T_i) - \mathbb{E}[\omega_t^+(T)])\right| \\
&\precsim \left[\sqrt{\frac{\log(\sqrt{1/h})}{nh}} + \sqrt{\frac{\log(np)}{nh}}\cdot\sqrt{1+\sqrt{\frac{\log(\sqrt{1/h})}{nh}} + \frac{\log(np)}{nh}}\right] \\
&\sim \sqrt{\frac{\log(np/\sqrt{h})}{nh}}
\end{aligned}$$

with probability $1-1/np$, where the last expression holds by the assumption that $\log(np/\sqrt{h})/(nh) = o(1)$ and $h = o(1)$. Multiplying both sides of the above equation by $\sqrt{n}$ completes the proof of Lemma 3.3.

4.4 Proof of Lemma 3.4

For convenince, we prove Lemma 3.4 by conditioning on the event

$$\mathcal{A} = \{\max_{i\in[n]}\max_{j\in[p]} |X_{ij}| \le M_X\cdot\sqrt{\log np}\}.$$

Since X_{ij} conditioned on T is a Gaussian random variable, the event $\mathcal{A}$ occurs with probability at least $1 - 1/np$ for sufficiently large constant $M_X > 0$. Recall the definition that

$g^+_{t,jk}(T_i, X_{ij}, X_{ik}) = K^+_h(T_i - t)X_{ij}X_{ik}$ and $q^+_{t,jk}(T_i, X_{ij}, X_{ik}) = g^+_{t,jk}(T_i, X_{ij}, X_{ik}) - \mathbb{E}[g^+_{t,jk}(T, X_j, X_k)]$, respectively. We consider the function class

$$\mathcal{Q}^+ = \{q^+_{t,jk} | t \in [a, b], j, k \in [p]\}.$$

We first obtain an upper bound for the function class

$$\begin{aligned}
\sup_{t\in[a,b]} \max_{j,k\in[p]} \|q^+_{t,jk}\|_\infty &= \sup_{t\in[a,b]} \max_{j,k\in[p]} \|g^+_{t,jk}(T_i, X_{ij}, X_{ik}) - \mathbb{E}[g^+_{t,jk}(T, X_j, X_k)]\|_\infty \\
&\leq \sup_{t\in[a,b]} \max_{j,k\in[p]} \|g^+_{t,jk}(T_i, X_{ij}, X_{ik})\|_\infty + \sup_{t\in[a,b]} \max_{j,k\in[p]} \|\mathbb{E}[g^+_{t,jk}(T, X_j, X_k)]\|_\infty \\
&\leq \sup_{t\in[a,b]} \max_{j,k\in[p]} \|K^+_h(T_i - t)X_{ij}X_{ik}\|_\infty + \overline{f}_T \cdot M_\sigma + \mathcal{O}(h) \\
&\leq \frac{1}{h} \cdot M_X^2 \cdot \|K^+\|_\infty \cdot \log np + \overline{f}_T \cdot M_\sigma + \mathcal{O}(h) \\
&\leq \frac{2}{h} \cdot M_X^2 \cdot \|K^+\|_\infty \cdot \log np, \qquad (29)
\end{aligned}$$

where the second inequality holds by Assumptions 1-2 and Lemma 3.2, the third inequality holds by Proposition 1 and by conditioning on the event $\mathcal{A}$, and the last inequality holds by the scaling assumption $h = o(1)$ for sufficiently large n.

Next, we obtain an upper bound for the variance of $q^+_{t,jk}(T_i, X_{ij}, X_{ik})$. Note that

$$\begin{aligned}
\sup_{t\in[a,b]} \max_{j,k\in[p]} Var(q^+_{t,jk}(T, X_j, X_k)) &= \sup_{t\in[a,b]} \max_{j,k\in[p]} \mathbb{E}\left[(g^+_{t,jk}(T_i, X_{ij}, X_{ik}) - \mathbb{E}[g^+_{t,jk}(T, X_j, X_k)])^2\right] \\
&\leq \sup_{t\in[a,b]} \max_{j,k\in[p]} 2\mathbb{E}\left[g^{+2}_{t,jk}(T, X_j, X_k)\right] + \sup_{t\in[a,b]} \max_{j,k\in[p]} 2\mathbb{E}^2\left[g^+_{t,jk}(T, X_j, X_k)\right],
\end{aligned}$$

where we apply the inequality $(x-y)^2 \leq 2x^2 + 2y^2$ for two scalars x, y. By Lemma 3.2, we have $\sup_{t\in[a,b]} \max_{j,k\in[p]} 2\mathbb{E}^2\left[g^+_{t,jk}(T, X_j, X_k)\right] \leq 2\left(\overline{f}_T \cdot M_\sigma + \mathcal{O}(h)\right)^2$. Also, by a change of variable and second-order Taylor expansion on the marginal density $f_T(\cdot)$ as in 25-26, we have

$$\begin{aligned}
\sup_{t\in[a,b]} \max_{j,k\in[p]} 2\mathbb{E}\left[g^{+2}_{t,jk}(T, X_j, X_k)\right] &= 2 \sup_{t\in[a,b]} \max_{j,k\in[p]} \mathbb{E}\left[K^{+2}_h(T - t) \cdot \mathbb{E}[X_j^2 X_k^2 | T]\right] \\
&\leq 2\kappa \sup_{t\in[a,b]} \mathbb{E}[K^{+2}_h(T - t)] \\
&\leq \frac{2\kappa}{h} \cdot \overline{f}_T \cdot \|K^+\|_2^2 + \mathcal{O}(1) + \mathcal{O}(h),
\end{aligned}$$

where the first inequality follows form the fact that $|\mathbb{E}[X_j^2 X_k^2 | T]| \leq \kappa$ for some $\kappa < \infty$, and the second inequality follows from 25-26. Thus, for sufficiently large n and the assumption that $h = o(1)$, we have

$$\sup_{t\in[a,b]} \max_{j,k\in[p]} Var(q^+_{t,jk}(T, X_j, X_k)) \leq \frac{3\kappa}{h} \cdot \overline{f}_T \cdot \|K^+\|_2^2.$$

By Lemma 7.5, the covering number for the function class $\mathcal{Q}$ satisfies

$$\sup_Q N(\mathcal{Q}^+, L_2(Q), \epsilon) \leq p^2 \cdot \left(\frac{4 \cdot \|K^+\|_{TV} \cdot C_{K^+}^{4/5} \cdot \overline{f}_T^{1/5} \cdot M_\sigma^{1/5} \cdot M_X^{8/5} \cdot \log^{4/5} np}{h\epsilon} \right)^5 ,$$

where we multiply p^2 on the right hand side since the function class $\mathcal{Q}$ is taken over all $j,k \in [p]$.We now obtain an upper bound for the supreme of the empirical process, $\sup_{t\in[a,b]} \max_{j,k\in[p]} |\mathbb{G}_n[g_{t,jk}^+]|$. By lemma 8.2 with $A = 2 \cdot \|K^+\|_{TV} \cdot C_{K^+}^{4/5} \cdot \overline{f}_T^{1/5} \cdot M_\sigma^{1/5} \cdot M_X^{-2/5} \cdot p^{2/5} \cdot \log^{-1/5} np/\|K^+\|_\infty$, $U = \|F\|_{L_2(\mathbb{P}_n)} = 2 \cdot \|K^+\|_\infty \cdot M_X^2 \cdot \log np/h$, $V = 5$, $\sup_{t\in[a,b]} \max_{j,k\in[p]} \mathbb{E} q_{t,jk}^{+2} \leq \sigma_P^2 = (3\kappa/h) \cdot \overline{f}_T \cdot \|K^+\|_2^2 \leq (2/h \cdot M_X^2 \cdot \|K^+\|_\infty \cdot \log np)^2 = \|F\|_{L_2(\mathbb{P}_n)}^2$, for sufficiently large n, we obtain

$$\begin{aligned}
\mathbb{E}\left[\sup_{t\in[a,b]} \max_{j,k\in[p]} \frac{1}{\sqrt{n}} \cdot |\mathbb{G}_n[g_{t,jk}^+]| \right] &= \mathbb{E}\left[\sup_{t\in[a,b]} \max_{j,k\in[p]} \frac{1}{n} \cdot \left| \sum_{i\in[n]} (g_{t,jk}^+(T_i, X_{ij}, X_{ik} - \mathbb{E}[g_{t,jk}^+(T, X_j, X_k)])) \right| \right] \\
&\lesssim \sqrt{\frac{\log(p^{2/5} \cdot \log^{4/5} np/\sqrt{h})}{nh}} + \frac{\log np \cdot \log(p^{2/5} \cdot \log^{4/5} np/\sqrt{h})}{nh} \\
&\sim \sqrt{\frac{\log(p^{2/5} \cdot \log^{4/5} np/\sqrt{h})}{nh}},
\end{aligned}$$

where the last inequality holds by the assumption $\log np \cdot \sqrt{\log(p^{2/5} \cdot \log^{4/5} np/\sqrt{h})/nh} = o(1)$. By Lemma 8.1 with $\tau^2 = (3\kappa/h) \cdot \overline{f}_T \cdot \|K^+\|_2^2$, $\eta = 2/h \cdot M_X^2 \cdot \|K^+\|_\infty \cdot \log np$, $\mathbb{E}[Y] \leq C \cdot \sqrt{\log(p^{2/5} \cdot \log^{4/5} np/\sqrt{h})/nh}$, and picking $t = \sqrt{\log np/n}$, for sufficiently large n, we have

$$\begin{aligned}
\sup_{t\in[a,b]} \max_{j,k\in[p]} \frac{1}{\sqrt{n}} \cdot |\mathbb{G}_n[g_{t,jk}^+]| &= \sup_{t\in[a,b]} \max_{j,k\in[p]} \frac{1}{n} \cdot \left| \sum_{i\in[n]} (g_{t,jk}^+(T_i, X_{ij}, X_{ik} - \mathbb{E}[g_{t,jk}^+(T, X_j, X_k)])) \right| \\
&\lesssim \sqrt{\frac{\log(p^{2/5} \cdot \log^{4/5} np/\sqrt{h})}{nh}} + \sqrt{\frac{\log np}{nh}} \\
&\quad \cdot \sqrt{1 + \log np \cdot \sqrt{\frac{\log(p^{2/5} \cdot \log^{4/5} np/\sqrt{h})}{nh}} + \frac{\log^2 np}{nh}} \\
&\lesssim \sqrt{\frac{\log(np/\sqrt{h})}{nh}}
\end{aligned}$$

with probability at least $1 - 2/np$. The second inequality holds by the assumption that $\log^2 np \cdot \log(np/\sqrt{h})/nh = o(1)$. Multiplying both sides of the equation by $\sqrt{n}$, we complete the proof of 3.4.

5 Theorem 2

Theorem 2 *Assume that* $\sqrt{nh^3} = o(1)$*. In addition, assume that* $\text{poly}(s)\cdot\sqrt{\log^4(np/\sqrt{h})/nh^2}+ \text{poly}(s)\cdot\log^8(p/h)\cdot\log^2(ns)/(nh) = o(1)$*, where* $\text{poly}(s)$ *is a polynomial of* s*. Under the same conditions in Corollary 2, we have*

$$\lim_{n\to\infty}\sup_{\Theta(\cdot)\in\mathcal{U}_{s,m,M}} P_{\Theta(\cdot)}\left(U_E \geq c\left(1-\alpha, E\right)\right) \leq \alpha,$$

where $c\left(1-\alpha,E\right)$ *follows the definition that*

$$c\left(1-\alpha,E\right) = \inf\{q\in\mathbb{R}|P\left(U_E^B \leq q|\{(t_i,X_i)\}_{i\in[n]}\right) \geq 1-\alpha\}.$$

To prove Theorem 2, we use a similar set of arguments in the series of work on Gaussian multiplier bootstrap of the supreme of empirical [6][5][4] and generate the proof of theorem 2 in [12]. Recall the definition that, for any $0<a<b<1$

$$U_E = \sup_{t\in[a,b]}\max_{(j,k)\in E(t)}\sqrt{nh}\cdot\left|\left(\hat{\Theta}^{d+}_{jk}(t) - \Theta^{+}_{jk}(t)\right) - \left(\hat{\Theta}^{d-}_{jk}(t) - \Theta^{-}_{jk}(t)\right)\right|/\tilde{\sigma}_{jk}(t),$$

where $\tilde{\sigma}_{jk}(t)$ is the normalization term. We also have the bootstrap statistic

$$M^{B+}_{ijk}(t) = \left(\hat{\Theta}^{+}_{j}(t)\right)^T K_h^+(t_i-t)\left(X_iX_i^T\hat{\Theta}^{+}_{k}(t) - e_k\right),$$
$$M^{B-}_{ijk}(t) = \left(\hat{\Theta}^{-}_{j}(t)\right)^T K_h^-(t_i-t)\left(X_iX_i^T\hat{\Theta}^{-}_{k}(t) - e_k\right),$$

and

$$U_E^B = \sup_{t\in[a,b]}\max_{(j,k)\in E(t)}\sqrt{nh}\cdot\left|\frac{\sum_{i\in[n]}M^{B+}_{ijk}(t)\xi_i}{\sum_{i\in[n]}K_h^+(t_i-t)} - \frac{\sum_{i\in[n]}M^{B-}_{ijk}(t)\xi_i}{\sum_{i\in[n]}K_h^-(t_i-t)}\right|/\tilde{\sigma}_{jk}(t)$$

in which $\xi_1,\ldots,\xi_n \overset{i.i.d.}{\sim} N(0,1)$.

We aim to show that U_E^B is a good approximation of U_E. However, U_E and U_E^B are not exact averages. To apply the results in [5], we define four intermediate processes:

$$U_0 = \sup_{t\in[a,b]}\max_{(j,k)\in E(t)}\sqrt{nh}\cdot\left|\frac{\sum_{i\in[n]}M^{+}_{ijk}(t)}{\sum_{i\in[n]}K_h^+(t_i-t)} - \frac{\sum_{i\in[n]}M^{-}_{ijk}(t)}{\sum_{i\in[n]}K_h^-(t_i-t)}\right|/\tilde{\sigma}_{jk}(t);$$

$$U_{00} = \sup_{t\in[a,b]}\max_{(j,k)\in E(t)}\sqrt{nh}\cdot \left|\left(\frac{\sum_{i\in[n]}M^{+}_{ijk}(t)}{\sum_{i\in[n]}K_h^+(t_i-t)} - \frac{\sum_{i\in[n]}M^{-}_{ijk}(t)}{\sum_{i\in[n]}K_h^-(t_i-t)}\right) - \left(\frac{n\cdot\Lambda^{+}_{jk}(t)}{\sum_{i\in[n]}K_h^+(t_i-t)} - \frac{n\cdot\Lambda^{-}_{jk}(t)}{\sum_{i\in[n]}K_h^-(t_i-t)}\right)\right|/\tilde{\sigma}_{jk}(t); \tag{30}$$

$$U_0^B = \sup_{t\in[a,b]} \max_{(j,k)\in E(t)} \sqrt{nh}\cdot \left| \frac{\sum_{i\in[n]} M_{ijk}^+(t)\xi_i}{\sum_{i\in[n]} K_h^+(t_i-t)} - \frac{\sum_{i\in[n]} M_{ijk}^-(t)\xi_i}{\sum_{i\in[n]} K_h^-(t_i-t)} \right| / \tilde{\sigma}_{jk}(t);$$

$$U_{00}^B = \sup_{t\in[a,b]} \max_{(j,k)\in E(t)} \sqrt{nh}\cdot \left| \left(\frac{\sum_{i\in[n]} M_{ijk}^+(t)\xi_i}{\sum_{i\in[n]} K_h^+(t_i-t)} - \frac{\sum_{i\in[n]} M_{ijk}^-(t)\xi_i}{\sum_{i\in[n]} K_h^-(t_i-t)} \right) - \left(\frac{\sum_{i\in[n]} \Lambda_{jk}^+(t)\xi_i}{\sum_{i\in[n]} K_h^+(t_i-t)} - \frac{\sum_{i\in[n]} \Lambda_{jk}^-(t)\xi_i}{\sum_{i\in[n]} K_h^-(t_i-t)} \right) \right| / \tilde{\sigma}_{jk}(t),$$

where

$$\begin{aligned}
M_{ijk}^+(t) &= \left(\Theta_j^+(t)\right)^T K_h^+(t_i-t)\left(X_iX_i^T - \Sigma^+(t)\right)\Theta_k^+(t),\\
M_{ijk}^-(t) &= \left(\Theta_j^-(t)\right)^T K_h^-(t_i-t)\left(X_iX_i^T - \Sigma^-(t)\right)\Theta_k^-(t),\\
\Lambda_{jk}^+(t) &= \left(\Theta_j^+(t)\right)^T \left(\mathbb{E}\left[K_h^+(T-t)XX^T\right] - \mathbb{E}\left[K_h^+(T-t)\right]\Sigma^+(t)\right)\Theta_k^+(t),\\
\Lambda_{jk}^-(t) &= \left(\Theta_j^-(t)\right)^T \left(\mathbb{E}\left[K_h^-(T-t)XX^T\right] - \mathbb{E}\left[K_h^-(T-t)\right]\Sigma^-(t)\right)\Theta_k^-(t)\\
\xi_i &\overset{i.i.d.}{\sim} N(0,1).
\end{aligned}$$

Similar to the proof of theorem 2 in [12], we show that U_{00} is a good approximation of U_E and that U_{00}^B is a good approximation of U_E^B. We then show that there exists a Gaussian process W such that both U_{00}^B and U_{00} can be accurately approximated by W. This is done by applications of Theorems A.1 and A.2 in [5]. The following summarizes the chain of empirical and Gaussian processes that we are going to study

$$U_E \longleftrightarrow U_0 \longleftrightarrow U_{00} \longleftrightarrow W \longleftrightarrow U_{00}^B \longleftrightarrow U_0^B \longleftrightarrow U_E^B.$$

5.1 Thechnical Lemmas

We first give the uniform rates of convergence for the normalization term $\tilde{\sigma}_{jk}(t)$ in lemma 5.1.

Lemma 5.1 *Under the same conditions in corollary 2, there exists a positive constant C such that*

$$\sup_{t\in[a,b]} \max_{j,k\in[p]} \left| \tilde{\sigma}_{jk}^2(t) - \left(\Theta_{jj}^+(t)\Theta_{kk}^+(t) + \left(\Theta_{jk}^+(t)\right)^2 + \Theta_{jj}^-(t)\Theta_{kk}^-(t) + \left(\Theta_{jk}^-(t)\right)^2 \right) \right| \le C\cdot \left(h + \sqrt{\frac{\log(np/\sqrt{h})}{nh}} \right)$$

with probability at least $1 - 8/np$ for sufficiently large n.

Lemma 5.2 *Assume that* $h + \sqrt{\log(np/\sqrt{h})/nh} = o(1)$ *and* $\sqrt{nh^3} + s \cdot \log\left(np/\sqrt{h}\right)/\sqrt{nh} = o(1)$. *Under Assumption 1-2, for sufficiently large* n*, there exists a universal constant* $C > 0$ *such that*

$$|U_E - U_{00}| \le C \cdot \left(\sqrt{nh^3} + s \cdot \frac{\log(np/\sqrt{h})}{\sqrt{nh}}\right),$$

with probability at least $1 - 8/np$.

We now apply Theorems A.1 and A.2 in [5] to show that there exists a Gaussian process W such that the quantities $|U_{00} - W|$ and $\left|T_{00}^B - W\right|$ can be controled, respectively. The results are stated in the following Lemmas.

Lemma 5.3 *Assume that* $\log^2(ns)\log^4(s)\log^4(p/h)/(nh) = o(1)$.*Under Assumptions 1-2, for sufficiently large* n*, there exist universal constants* $C, C' > 0$ *such that*

$$P\left(|U_{00} - W| \ge C \cdot \left(\frac{\log^2(ns)\log^4(s)\log^4(p/h)}{nh}\right)^{1/8}\right) \le C' \cdot \left(\frac{\log^2(ns)\log^4(s)\log^4(p/h)}{nh}\right)^{1/8}.$$

Lemma 5.4 *Assume that* $\log^2(ns) \cdot \log^2(s) \cdot \log^3(p/h)/(nh) = o(1)$.*Under Assumptions 1-2, for sufficiently large* n*, there exist universal constants* $C, C'' > 0$ *such that*

$$P\left(\left|U_{00}^B - W\right| > C \cdot \left(\frac{\log^2(ns) \cdot \log^2(s) \cdot \log^3(p/h)}{nh}\right)^{1/8} \middle| \{(T_i, X_i)\}_{i\in[n]}\right)$$
$$\le C'' \cdot \left(\frac{\log^2(ns) \cdot \log^2(s) \cdot \log^3(p/h)}{nh}\right)^{1/8},$$

with probability at least $1 - 3/n$.

Finally, the following lemma provides an upper bound on the difference between U_E^B and U_{00}^B, conditioned on the data $\{(T_i, X_i)\}_{i\in[n]}$.

Lemma 5.5 *Assume that* $s \cdot \sqrt{\log^4(np/\sqrt{h})/nh^2} = o(1)$ *and* $nh^3 = o(1)$. *Under Assumptions 1-2, for sufficiently large* n*, there exist universal constants* $C, C'' > 0$ *such that, with probability at least* $1 - 10/np$,

$$P\left(|U_E^B - U_{00}^B| \ge C \cdot s \cdot \sqrt{\frac{\log^4(np/\sqrt{h})}{nh^2}} \middle| \{(T_i, X_i)\}_{i\in[n]}\right) \le \frac{2}{n}.$$

5.2 Proof of Theorem 2

With Lemma 5.1 - 5.5, we are now ready to prove Theorem 2.

Now we show that U_E can be well-approximated by the $(1-\alpha)$ conditional quantile of U_E^B, i.e. $P(U_E \geq c(1-\alpha)) \leq \alpha$. For notation convenience, we let $r = r_1 + r_2 + r_3 + r_4$, where

$$r_1 = \sqrt{nh^3} + s \cdot \frac{\log(np/\sqrt{h})}{\sqrt{nh}}$$
$$r_2 = \left(\frac{\log^2(ns)\log^4(s)\log^4(p/h)}{nh}\right)^{1/8}$$
$$r_3 = \left(\frac{\log^2(ns)\cdot\log^2(s)\cdot\log^3(p/h)}{nh}\right)^{1/8}$$
$$r_4 = s \cdot \sqrt{\frac{\log^4(np/\sqrt{h})}{nh^2}}.$$

These are the scaling that appears in Lemmas 5.2-5.5. By Lemmas 5.2 and 5.3, it can be shown that

$$P(|U_E - W| \geq r_1 + r_2) \leq P(|U_E - U_{00}| + |U_{00} - W| \geq r_1 + r_2) \leq 2r_2 \tag{31}$$

since $r_2 \geq 1/np$. With some abuse of notation, throughout the proof, we write $P_\xi(U_E^B \geq t)$ to indicate $P(U_E^B \geq t|\{(T_i, X_i)\}_{i\in[n]})$. By Lemmas 5.4 and 5.5, we have

$$P_\xi(|U_E^B - W| \geq r_2 + r_4) \leq P_\xi(|U_E^B - U_{00}^B| + |U_{00}^B - W| \geq r_2 + r_4) \leq 2r_2 \tag{32}$$

since $r_2 \geq r_3$ and $r_2 \geq 1/n$. Define the event

$$\mathcal{E} = \{P_\xi(|U_E^B - W| \geq r_2 + r_4) \leq r_2\},$$

and note that $P(\mathcal{E}) \geq 1 - 3/n$ by Lemma 5.4 and 5.5. Throughout the proof, we condition on the event $\mathcal{E}$.

By the triangle inequality, we obtain

$$\begin{aligned} P(U_E \leq c(1-\alpha)) &\geq 1 - P\left(U_E - W + W + r \geq c(1-\alpha) + r\right) \\ &\geq 1 - P\left(|U_E - W| \geq r\right) - P\left(W \geq c(1-\alpha) - r\right) \\ &\geq P\left(|W| \leq c(1-\alpha) - r\right) - 2r_2, \end{aligned} \tag{33}$$

where the last inequality follows from 31. By a similar argument and by 32, we have

$$\begin{aligned} P\left(|W| \leq c(1-\alpha) - r\right) &\geq P_\xi\left(U_E^B \leq c(1-\alpha) - 2r\right) - 2r_2 \\ &\geq P_\xi\left(U_E^B \leq c(1-\alpha)\right) - 2r_2 - P_\xi\left(|U_E^B - c(1-\alpha)| \leq r\right), \end{aligned} \tag{34}$$

where the last inequality follows from the fact that $P(X \leq t-\epsilon) - P(X \leq t) \geq -P(|X-t| \leq \epsilon)$ for any $\epsilon > 0$. Thus, combining 33 and 34, we obtain

$$P(U_E \leq c(1-\alpha)) \geq 1 - \alpha - 4r_2 - P_\xi \left(|U_E^B - c(1-\alpha)| \leq r \right). \tag{35}$$

It remains to show that the quantity $P_\xi \left(|U_E^B - c(1-\alpha)| \leq r \right)$ converges to zero as we increase n.

By the defintion of $\tilde{U}_{00}$ and $\tilde{U}_{00}^B$, we have

$$\tilde{U}_{00} = \frac{\sup_{t \in [a,b]} \max_{(j,k) \in E(t)} \left| \sum_{i \in [n]} J_{t,jk}^+(T_i, X_i) - \sum_{i \in [n]} J_{t,jk}^-(T_i, X_i) \right|}{\sqrt{n} \cdot f_T(t) \cdot \bar{\sigma}_{jk}(t)}$$

and

$$\tilde{U}_{00}^B = \frac{\sup_{t \in [a,b]} \max_{(j,k) \in E(t)} \left| \left(\sum_{i \in [n]} J_{t,jk}^+(T_i, X_i) - \sum_{i \in [n]} J_{t,jk}^-(T_i, X_i) \right) \cdot \xi_i \right|}{\sqrt{n} \cdot f_T(t) \cdot \bar{\sigma}_{jk}(t)}.$$

Let $\hat{\sigma}_{t,jk}^2 = \sum_{i=1}^n \left(\sum_{i \in [n]} J_{t,jk}^+(T_i, X_i) - \sum_{i \in [n]} J_{t,jk}^-(T_i, X_i) \right)^2 / \left(n \cdot f_T^2(t) \cdot \bar{\sigma}_{jk}^2(t) \right)$ be the conditional variance, and let $\underline{\sigma} = \inf_{t,jk} \hat{\sigma}_{t,jk}$ and $\overline{\sigma} = \sup_{t,jk} \hat{\sigma}_{t,jk}$. By Lemma A.1 of [6] and Theorem 3 of [4], we obtain

$$\begin{aligned} & P_\xi \left(|T_E^B - c(1-\alpha)| \leq r \right) \\ & \leq C \cdot \overline{\sigma} / \underline{\sigma} \cdot r \cdot \left(\mathbb{E} \left[U_E^B | \{(T_i, X_i)\}_{i \in [n]} \right] + \sqrt{1 \vee \log(\underline{\sigma}/r)} \right) \\ & \leq C \cdot \overline{\sigma} / \underline{\sigma} \cdot r \cdot \left(\mathbb{E} \left[\tilde{U}_{00}^B | \{(T_i, X_i)\}_{i \in [n]} \right] + \mathbb{E} \left[|U_E^B - \tilde{U}_{00}^B| | \{(T_i, X_i)\}_{i \in [n]} \right] + \sqrt{1 \vee \log(\underline{\sigma}/r)} \right). \end{aligned} \tag{36}$$

We first calculate the quantity $\overline{\sigma}$. By 47, we have

$$\sup_{t \in [a,b]} \max_{j,k \in [p]} \| \left(J_{t,jk}^+(T_i, X_i) - J_{t,jk}^-(T_i, X_i) \right)^2 \Big/ \left(f_T^2(t) \cdot \bar{\sigma}_{jk}^2(t) \right) \|_\infty \leq C \cdot \frac{\log^2(2ns)}{h}$$

Moreover, by 47, we have

$$\sup_{t \in [a,b]} \max_{j,k \in [p]} \mathbb{E} \left[\left(J_{t,jk}^+(T_i, X_i) - J_{t,jk}^-(T_i, X_i) \right)^4 \Big/ \left(f_T^4(t) \cdot \bar{\sigma}_{jk}^4(t) \right) \right] \leq C \cdot \frac{\log^4(2ns)}{h^2}.$$

Define the function class

$$\mathcal{J}^* = \left\{ \left(J_{t,jk}^+(T_i, X_i) - J_{t,jk}^-(T_i, X_i) \right)^2 \Big/ \left(f_T^2(t) \cdot \bar{\sigma}_{jk}^2(t) \right) \Big| z \in [0,1], j, k \in [p] \right\}.$$

By Lemma 7.3, 7.6 and 7.7, we have

$$\sup_Q N(\mathcal{J}^*, L_2(Q), \epsilon) \le C \cdot p^2 \cdot \left(\frac{p^{11/6} \log^{5/6} np}{h^{11/12} \epsilon} \right)^{48}.$$

Thus, applying Lemma 8.2 with $\sigma_P^2 = C \cdot \log^4(2ns)/h^2$, $V = 48$, $\|F\|_{L_2(\mathbb{P}_n)} \le C \cdot p^2 \cdot \log^2(2ns)/h$ and $A\|F\|_{L_2(\mathbb{P}_n)} \le p^{15/8} \log^{5/6}(np)/h^{11/12}$, we have

$$\mathbb{E}\left[\sup_{t\in[a,b]} \max_{j,k\in[p]} \left| \frac{1}{n} \sum_{i\in[n]} \phi_{t,jk}(T_i, X_i) - \mathbb{E}\left[\phi_{t,jk}(T, X)\right] \right| \right]$$
$$\le C \cdot \sqrt{\frac{\log^5(2np)}{nh^2}},$$

where we denote

$$\phi_{t,jk}(T_i, X_i) = \frac{\left(J^+_{t,jk}(T_i, X_i) - J^-_{t,jk}(T_i, X_i)\right)^2}{f_T^2(t) \cdot \bar{\sigma}^2_{jk}(t)}$$

for notation simplicity.

By an application of the Marknov's inequality, we obtain

$$P\left(\sup_{t\in[a,b]} \max_{j,k\in[p]} \left| \frac{1}{n} \sum_{i\in[n]} \phi_{t,jk}(T_i, X_i) - \mathbb{E}\left[\phi_{t,jk}(T, X)\right] \right| \ge C \cdot \left(\frac{\log^5(2np)}{nh^2} \right)^{1/4} \right)$$
$$\le C \cdot \left(\frac{\log^5(2np)}{nh^2} \right)^{1/4}.$$

Thus, we have with probability at least $1 - C \cdot \left(\log^5(2np)/\left(nh^2\right)\right)^{1/4}$,

$$\overline{\sigma}^2 = \sup_{t\in[a,b]} \max_{j,k\in[p]} \frac{1}{n} \sum_{i=1}^n \phi_{t,jk}(T_i, X_i) \le \sup_{t\in[a,b]} \max_{j,k\in[p]} \mathbb{E}\left[\phi_{t,jk}(T, X)\right] + C \cdot \left(\frac{\log^5(2np)}{nh^2} \right)^{1/4}$$
$$\le C \log^2 s \tag{37}$$

where the last inequality follows from 49 for sufficiently large n. By Lemma 10 in [12], we have $\inf_{t,jk} \mathbb{E}\left[\phi_{t,jk}(T, X)\right] \ge c > 0$. Therefore, we have

$$\underline{\sigma}^2 = \inf_{t,jk} \frac{1}{n} \sum_{i=1}^n \phi_{t,jk}(T_i, X_i) \ge c - \sup_{t\in[a,b]} \max_{j,k\in[p]} \left| \frac{1}{n} \sum_{i\in[n]} \phi_{t,jk}(T_i, X_i) - \mathbb{E}\left[\phi_{t,jk}(T, X)\right] \right| \ge c/2 > 0$$

with probability at least $1 - C \cdot \left(\log^5(2np)/\left(nh^2\right)\right)^{1/4}$.

Next, we calculate the quantity $\mathbb{E}\left[\tilde{U}_{00}^B|\{(T_i,X_i)\}_{i\in[n]}\right]$. By Dudley's inequality (Corollary 2.2.8 in [14]), 50 and 37, we obtain

$$\mathbb{E}\left[\tilde{U}_{00}^B|\{(T_i,X_i)\}_{i\in[n]}\right] \le C\cdot \log s\cdot \sqrt{\log(p/h)}. \tag{38}$$

Moreover, by Lemma 5.5 and 5.4, we have

$$\begin{aligned}
&\mathbb{E}\left[|U_E^B-\tilde{U}_{00}^B||\{(T_i,X_i)\}_{i\in[n]}\right]\\
&\le C\cdot s\cdot\sqrt{\frac{\log^4(np/\sqrt{h})}{nh^2}}+C\cdot\sqrt{\log(np/\sqrt{h})}\cdot\left(\sqrt{h}+\left(\frac{\log(np/\sqrt{h})}{nh}\right)^{1/4}\right)\\
&\le C\cdot s\cdot\sqrt{\frac{\log^4(np/\sqrt{h})}{nh^2}}\le r, \qquad (39)
\end{aligned}$$

with probability at least 1- 3/n. Substituting 37, 38 and 39 into 36, we obtain

$$P_\xi\left(|T_E^B-c(1-\alpha)|\le r\right)\le C\cdot\left(\frac{\log^{20}(s)\cdot\log^8(p/h)\cdot\log^2(np)}{nh}\right)^{1/8}. \tag{40}$$

Thus, substituting 40 into 35, we have

$$P(U_E\le c(1-\alpha))\ge 1-\alpha-4r_2-\left(\frac{\log^{20}(s)\cdot\log^8(p/h)\cdot\log^2(np)}{nh}\right)^{1/8}.$$

By the scaling assumptions, $r_2=o(1)$ and $\log^{20}(s)\cdot\log^8(p/h)\cdot\log^2(np)/(nh)=o(1)$. Thus, this implies that

$$\lim_{n\to\infty}P(U_E\le c(1-\alpha))\ge 1-\alpha,$$

which implies that

$$\lim_{n\to\infty}P(U_E\ge c(1-\alpha))\le \alpha,$$

as desired

6 Proof of Technical Lemmas in Theorem 2

6.1 Proof of Lemma 5.1

Recall the definition that

$$\tilde{\sigma}^2\left(\hat{\Theta}_{jk}^{d+}(t)\right)=\sum_{i\in[n]}K_h^+(t_i-t)\left(\left(\hat{\Theta}_j^+(t)\right)^T\left(X_iX_i^T\hat{\Theta}_k^+(t)-e_k\right)\right)^2/\sum_{i\in[n]}K_h^+(t_i-t),$$

we have

$$
\begin{aligned}
\tilde{\sigma}^2\left(\hat{\Theta}^{d+}_{jk}(t)\right) &= \frac{\sum_{i\in[n]} K_h^+(t_i-t)\left(\left(\hat{\Theta}^+_j(t)\right)^T X_i\right)^2\left(\left(\hat{\Theta}^+_k(t)\right)^T X_i\right)^2}{\sum_{i\in[n]} K_h^+(t_i-t)} \\
&- \frac{2\cdot\sum_{i\in[n]} K_h^+(t_i-t)\hat{\Theta}^+_{jk}(t)\left(\left(\hat{\Theta}^+_j(t)\right)^T X_i\right)\left(\left(\hat{\Theta}^+_k(t)\right)^T X_i\right)}{\sum_{i\in[n]} K_h^+(t_i-t)} \\
&+ \frac{\sum_{i\in[n]} K_h^+(t_i-t)\left(\hat{\Theta}^+_{jk}(t)\right)^2}{\sum_{i\in[n]} K_h^+(t_i-t)} \\
&= \hat{\Omega}^+_{jk}(t) - 2\cdot\hat{\Theta}^+_{jk}(t)\left(\hat{\Theta}^+_j(t)\right)^T\hat{\Sigma}^+(t)\hat{\Theta}^+_k(t) + \left(\hat{\Theta}^+_{jk}(t)\right)^2.
\end{aligned}
$$

So,

$$
\begin{aligned}
&\sup_{t\in[a,b]}\max_{j,k\in[p]}\left|\tilde{\sigma}^2\left(\hat{\Theta}^{d+}_{jk}(t)\right) - \left(\Theta^+_{jj}(t)\Theta^+_{kk}(t) + \left(\Theta^+_{jk}(t)\right)^2\right)\right| \\
&\le \sup_{t\in[a,b]}\max_{j,k\in[p]}\left|\hat{\Omega}^{\pm}_{jk}(t) - \left(\Theta^{\pm}_{jj}(t)\Theta^{\pm}_{kk}(t) + 2\cdot\left(\Theta^+_{jk}(t)\right)^2\right)\right| \\
&+ 2\cdot\sup_{t\in[a,b]}\max_{j,k\in[p]}\left|\hat{\Theta}^+_{jk}(t)\left(\hat{\Theta}^+_j(t)\right)^T\hat{\Sigma}^+(t)\hat{\Theta}^+_k(t) - \left(\Theta^+_{jk}(t)\right)^2\right| \\
&+ \sup_{t\in[a,b]}\max_{j,k\in[p]}\left|\left(\hat{\Theta}^+_{jk}(t)\right)^2 - \left(\Theta^+_{jk}(t)\right)^2\right|.
\end{aligned}
$$

By Corollary 1-2,

$$
\sup_{t\in[a,b]}\max_{j,k\in[p]}\left|\tilde{\sigma}^2\left(\hat{\Theta}^{d+}_{jk}(t)\right) - \left(\Theta^+_{jj}(t)\Theta^+_{kk}(t) + \left(\Theta^+_{jk}(t)\right)^2\right)\right| \le C\cdot\left(h + \sqrt{\frac{\log(np/\sqrt{h})}{nh}}\right) \tag{41}
$$

with probability at least $1-4/np$. Similarly,

$$
\sup_{t\in[a,b]}\max_{j,k\in[p]}\left|\tilde{\sigma}^2\left(\hat{\Theta}^{d-}_{jk}(t)\right) - \left(\Theta^-_{jj}(t)\Theta^-_{kk}(t) + \left(\Theta^-_{jk}(t)\right)^2\right)\right| \le C\cdot\left(h + \sqrt{\frac{\log(np/\sqrt{h})}{nh}}\right) \tag{42}
$$

with probability at least $1-4/np$. Finally, combining 41 and 42, we complete the proof of Lemma 5.1.

6.2 Proof of Lemma 5.2

Lemmas 5.2-5.5 can be proved by the same set of argument of the proof of Lemma 6-9 in [12]. Lemma 5.2 provides an approximation error between the statistic U_E and the intermediate empirical process U_{00}.

By the triangle inequality, we have $|U_E - U_{00}| \le |U_E - U_0| + |U_0 - U_{00}|$. Thus, it is suffices to obtain upper bounds for the terms $|U_E - U_0|$ and $|U_0 - U_{00}|$. With the fact that for any $j \neq k \in [p]$, $t \in [a,b]$, $\Theta^{\pm}_{jj}(t)\Theta^{\pm}_{kk}(t) + \left(\Theta^{\pm}_{jk}(t)\right)^2 \ge m^2$ given $\lambda_1(\Theta^{\pm}(t)) \ge m$, we have

$$\|\Theta^{\pm}_{j}(t)\|_1 \|\Theta^{\pm}_{k}(t)\|_1 / \tilde{\sigma}_{jk}(t) \le M^2/(\sqrt{2}m) \tag{43}$$

with probability at least $1 - 8/np$. Following the results of 43 and the proof of Lemma 6 in [12], we give the uppper bounds for $|U_E - U_0|$ and $|U_0 - U_{00}|$ directly.

Upper Bound for $|U_E - U_0|$:

$$\begin{aligned} |U_E - U_0| &\le C \cdot \sqrt{nh} \cdot s \cdot \left(h + \sqrt{\frac{\log(np/\sqrt{h})}{nh}} \right)^2 \\ &= C \cdot \left(s \cdot \sqrt{nh^5} + \frac{s \cdot \log(np/\sqrt{h})}{\sqrt{nh}} + s \cdot h \sqrt{\log(np/\sqrt{h})} \right) \end{aligned} \tag{44}$$

with probability at least $1 - 8/np$, for sufficiently large n.

Upper Bound for $|U_0 - U_{00}|$:

$$|U_0 - U_{00}| \le C \cdot \sqrt{nh^3} \tag{45}$$

Thus, combining 44 and 45, there exists a constant $C > 0$ such that

$$|U_E - U_{00}| \le C \cdot \left(\sqrt{nh^3} + s \cdot \sqrt{nh^5} + s \cdot \frac{\log(np/\sqrt{h})}{\sqrt{nh}} + s \cdot h \sqrt{\log(np/\sqrt{h})} \right),$$

with probability at least $1 - 8/np$. By the assumption that $nh^3 = o(1)$, we obtain

$$|U_E - U_{00}| \le C \cdot \left(\sqrt{nh^3} + s \cdot \frac{\log(np/\sqrt{h})}{\sqrt{nh}} \right).$$

6.3 Proof of Lemma 5.3

Recall from 30 the definition

$$U_{00} = \sup_{t \in [a,b]} \max_{(j,k) \in E(t)} \sqrt{nh} \cdot \left| \left(\frac{\sum_{i \in [n]} M^{+}_{ijk}(t)}{\sum_{i \in [n]} K^{+}_{h}(t_i - t)} - \frac{\sum_{i \in [n]} M^{-}_{ijk}(t)}{\sum_{i \in [n]} K^{-}_{h}(t_i - t)} \right) - \left(\frac{n \cdot \Lambda^{+}_{jk}(t)}{\sum_{i \in [n]} K^{+}_{h}(t_i - t)} - \frac{n \cdot \Lambda^{-}_{jk}(t)}{\sum_{i \in [n]} K^{-}_{h}(t_i - t)} \right) \right| / \tilde{\sigma}_{jk}(t).$$

Recall from 4 that $J^{\pm}_{t,jk}(T_i, X_i) = J^{\pm(1)}_{t,jk}(T_i, X_i) - J^{\pm(2)}_{t,jk}(T_i)$, where $J^{\pm(1)}_{t,jk}(T_i, X_i)$ and $J^{\pm(2)}_{t,jk}(T_i)$ are as defined in 2 and 3, respectively. Let $\mathcal{J} = \{J^{+}_{t,jk} - J^{-}_{t,jk} | t \in [0,1], j,k \in [p]\}$ and $\mathcal{J}^{\pm} = \{J^{\pm}_{t,jk} | t \in [0,1], j,k \in [p]\}$. Then the intermediate empirical average U_{00} can be written as

$$U_{00} = \sup_{t \in [a,b]} \max_{(j,k) \in E(t)} \sqrt{n} \left| \left(\frac{\sum_{i \in [n]} J^{+}_{t,jk}(T_i, X_i)}{\sum_{i \in [n]} K^{+}_{h}(t_i - t)} - \frac{\sum_{i \in [n]} J^{-}_{t,jk}(T_i, X_i)}{\sum_{i \in [n]} K^{-}_{h}(t_i - t)} \right) \right| / \tilde{\sigma}_{jk}(t).$$

By Lemma 3.2 and Lemma 3.3, with probability $1 - 1/np$, we have

$$\mathbb{P}_n \left[K^{+}_{h}(T_i - t) \right] = f_T(t) + \mathcal{O}(h) + \mathcal{O}\left(\frac{\log(np/\sqrt{h})}{nh} \right),$$

for sufficiently large n.

By Lemma 5.1, denote $\bar{\sigma}_{jk}(t) = \sqrt{\Theta^{+}_{jj}(t)\Theta^{+}_{kk}(t) + \left(\Theta^{+}_{jk}(t)\right)^2 + \Theta^{-}_{jj}(t)\Theta^{-}_{kk}(t) + \left(\Theta^{-}_{jk}(t)\right)^2}$, with probability at least $1 - 8/np$, we have

$$\tilde{\sigma}_{jk}(t) = \bar{\sigma}_{jk}(t) + \mathcal{O}\left(\sqrt{h + \sqrt{\frac{\log(np/\sqrt{h})}{nh}}} \right).$$

Denote $\tilde{U}_{00} = \sup_{t \in [a,b]} \max_{(j,k) \in E(t)} \left| \sum_{i \in [n]} J^{+}_{t,jk}(T_i, X_i) - \sum_{i \in [n]} J^{-}_{t,jk}(T_i, X_i) \right| / (\sqrt{n} f_T(t) \cdot \bar{\sigma}_{jk}(t))$. Thus, with probability at least 1-10/np, there exists a positive constant C, such that

$$\left| U_{00} - \tilde{U}_{00} \right| \leq C \cdot \sqrt{\log(np/\sqrt{h})} \cdot \left(\sqrt{h} + \left(\frac{\log(np/\sqrt{h})}{nh} \right)^{1/4} \right). \tag{46}$$

We will show that there exists a Gaussian process W such that

$$\left| \tilde{U}_{00} - W \right| \leq C \cdot \left(\frac{\log^2(ns)\log^4(s)\log^4(p/h)}{nh} \right)^{1/8}$$

with high probability. To this end, we apply Theorem A.1 in [5], which involves the following quantiites

- upper bound for $\sup_{t \in [a,b]} \max_{j,k \in [p]} \| \left| J^{+}_{t,jk}(T_i, X_i) - J^{-}_{t,jk}(T_i, X_i) \right| / (f_T(t) \cdot \bar{\sigma}_{jk}(t)) \|_{\infty}$;
- upper bound for $\sup_{t \in [a,b]} \max_{j,k \in [p]} \mathbb{E}\left[\left(J^{+}_{t,jk}(T, X) - J^{-}_{t,jk}(T, X) \right)^2 / (f_T(t) \cdot \bar{\sigma}_{jk}(t))^2 \right]$;
- covering number for the function class $\mathcal{J}$.

Let $\mathcal{S}_j(t)$ and $\mathcal{S}_k(t)$ be the support of $\Theta_j(t)$ and $\Theta_k(t)$, respectively. Note that the cardinality for each set is less than s. We now obtain the above quantities.

Upper bound for $\sup_{t\in[a,b]}\max_{j,k\in[p]}\|\left|J^+_{t,jk}(T_i,X_i)-J^-_{t,jk}(T_i,X_i)\right|/(f_T(t)\cdot\bar{\sigma}_{jk}(t))\|_\infty$**:** We have, with probability at least $1-1/(2ns)$,

$$
\begin{aligned}
&\sup_{t\in[a,b]}\max_{j,k\in[p]}\|J^+_{t,jk}(T_i,X_i)\|_\infty/\left(f_T(t)\cdot\bar{\sigma}_{jk}(t)\right)\\
&\le\sqrt{h}\cdot\sup_{t\in[a,b]}\max_{j,k\in[p]}\|\Theta^+_j(t)\|_1\|\Theta^+_k(t)\|_1\cdot\left(\max_{j\in\mathcal{S}_j(t),k\in\mathcal{S}_k(t)}\|q^+_{t,jk}\|_\infty+M_\sigma\cdot\|k^+_t\|_\infty\right)/\left(f_T(t)\cdot\bar{\sigma}_{jk}(t)\right)\\
&\le\sqrt{h}\cdot M\cdot\frac{1}{\underline{f}_T}\left(\frac{2}{h}\cdot M_X^2\cdot\|K^+\|_\infty\log(2ns)+M_\sigma\cdot\frac{2}{h}\cdot\|K^+\|_\infty\right)\\
&\le\frac{4}{\sqrt{h}}\cdot M\cdot\frac{1}{\underline{f}_T}\cdot M_X^2\cdot M_\sigma\cdot\|K^+\|_\infty\cdot\log(2ns)\\
&=C_1\cdot\frac{\log(2ns)}{\sqrt{h}},
\end{aligned}
$$

where the first inequality follows by Holder's inequality and definition of $q^+_{t,jk}$ and k^+_t and the second inequality follows from 24 and 29. Note that since we are only taking max over the set $\mathcal{S}_j(t)$ and $\mathcal{S}_k(t)$, instead of a $\log np$ factor from 29, we obtain a $\log(2ns)$ factor. Thus,

$$
\begin{aligned}
&\sup_{t\in[a,b]}\max_{j,k\in[p]}\|J^+_{t,jk}(T_i,X_i)-J^-_{t,jk}(T_i,X_i)\|_\infty/\left(f_T(t)\cdot\bar{\sigma}_{jk}(t)\right)\\
&\le\sup_{t\in[a,b]}\max_{j,k\in[p]}\|J^+_{t,jk}(T_i,X_i)\|_\infty/\left(f_T(t)\cdot\bar{\sigma}_{jk}(t)\right)+\sup_{t\in[a,b]}\max_{j,k\in[p]}\|J^-_{t,jk}(T_i,X_i)\|_\infty/\left(f_T(t)\cdot\bar{\sigma}_{jk}(t)\right)\\
&\le 2C_1\cdot\frac{\log(2ns)}{\sqrt{h}}.
\end{aligned}
\tag{47}
$$

Upper bound for $\sup_{t\in[a,b]}\max_{j,k\in[p]}\mathbb{E}\left[\left(J^+_{t,jk}(T,X)-J^-_{t,jk}(T,X)\right)^2/\left(f_T(t)\cdot\bar{\sigma}_{jk}(t)\right)^2\right]$**:** By an application of the inequality $(x-y)^2\le 2x^2+2y^2$, we have

$$
\begin{aligned}
&\sup_{t\in[a,b]}\max_{j,k\in[p]}\mathbb{E}\left[J^{+2}_{t,jk}(T,X)\right]/\left(f_T(t)\cdot\bar{\sigma}_{jk}(t)\right)\\
&=\sup_{t\in[a,b]}\max_{j,k\in[p]}\mathbb{E}\left[\left(J^{+(1)}_{t,jk}(T,X)-J^{+(2)}_{t,jk}(T)\right)^2\right]/\left(f_T(t)\cdot\bar{\sigma}_{jk}(t)\right)^2\\
&\le\underbrace{2\sup_{t\in[a,b]}\max_{j,k\in[p]}\mathbb{E}\left[\left(J^{+(1)}_{t,jk}(T,X)\right)^2\right]/\left(f_T(t)\cdot\bar{\sigma}_{jk}(t)\right)^2}_{I_1}\\
&+\underbrace{2\sup_{t\in[a,b]}\max_{j,k\in[p]}\mathbb{E}\left[\left(J^{+(2)}_{t,jk}(T)\right)^2\right]/\left(f_T(t)\cdot\bar{\sigma}_{jk}(t)\right)^2}_{I_2}.
\end{aligned}
$$

To obtain an upper bound for I_1, we need an upper bound for $\sup_{t\in[a,b]}\max_{j,k\in[p]}\mathbb{E}\left[\max_{j\in\mathcal{S}_j(t),k\in\mathcal{S}_k(t)} q_{z,jk}^{+2}\right]$. Recall from 1 the definition of $g_{t,jk}^{+}(T_i, X_{ij}, X_{ik}) = K_h^{+}(T_i - t)X_{ij}X_{ik}$ and that $q_{t,jk}^{+}(T_i, X_{ij}, X_{ik}) = g_{t,jk}^{+}(T_i, X_{ij}, X_{ik}) - \mathbb{E}[g_{t,jk}^{+}(T, X_j, X_k)]$. Thus, we have

$$\sup_{t\in[a,b]}\max_{j,k\in[p]}\mathbb{E}\left[\max_{j\in\mathcal{S}_j(t),k\in\mathcal{S}_k(t)} q_{z,jk}^{+2}\right] = \sup_{t\in[a,b]}\max_{j,k\in[p]}\mathbb{E}\left[\max_{j\in\mathcal{S}_j(t),k\in\mathcal{S}_k(t)}\left(g_{t,jk}^{+} - \mathbb{E}[g_{t,jk}^{+}]\right)^2\right]$$
$$\leq 2\sup_{t\in[a,b]}\max_{j,k\in[p]}\mathbb{E}\left[\max_{j\in\mathcal{S}_j(t),k\in\mathcal{S}_k(t)} g_{t,jk}^{+2}\right] + 2\sup_{t\in[a,b]}\max_{j,k\in[p]}\mathbb{E}^2\left[g_{t,jk}^{+}\right],$$

where we apply the fact that $(x-y)^2 \leq 2x^2 + 2y^2$ to obtain the last inequality. By Lemma 3.2, we have $2\sup_{t\in[a,b]}\max_{j,k\in[p]}\mathbb{E}^2[g_{t,jk}] \leq 2\left(\overline{f}_T \cdot M_\sigma + \mathcal{O}(h)\right)^2$. Moreover, we have

$$2\sup_{t\in[a,b]}\max_{j,k\in[p]}\mathbb{E}\left[\max_{j\in\mathcal{S}_j(t),k\in\mathcal{S}_k(t)} g_{t,jk}^{+2}\right] = 2\sup_{t\in[a,b]}\max_{j,k\in[p]}\mathbb{E}\left[\max_{j\in\mathcal{S}_j(t),k\in\mathcal{S}_k(t)} K_h^{+2}(T-t)X_j^2X_k^2\right]$$
$$= 2\sup_{t\in[a,b]}\max_{j,k\in[p]}\mathbb{E}\left[K_h^{+2}(T-t)\mathbb{E}\left[\max_{j\in\mathcal{S}_j(t),k\in\mathcal{S}_k(t)} X_j^2X_k^2|Z\right]\right]$$
$$\leq 2\cdot M_X^4\cdot\log^2(2s)\sup_{t\in[a,b]}\max_{j,k\in[p]}\mathbb{E}\left[K_h^{+2}(T-t)\right]$$
$$\leq 2\cdot M_X^4\cdot\log^2(2s)\left(\frac{1}{h}\overline{f}_T\|K^+\|_2^2 + \mathcal{O}(1) + \mathcal{O}(h)\right)$$
$$\leq 3\cdot\overline{f}_T\cdot\|K^+\|_2^2\cdot M_X^4\cdot\frac{\log^2(2s)}{h},$$

where the second inequality follows from an application of 27.

Thus, by Holder's inequality, we have

$$I_1 \leq 2\cdot h\cdot\sup_{t\in[a,b]}\max_{j,k\in[p]}\mathbb{E}\left[\left(\|\Theta_j(t)\|_1\cdot\|\Theta_k(t)\|_1\cdot\max_{j\in\mathcal{S}_j(t),k\in\mathcal{S}_k(t)}|q_{t,jk}^{+}|\right)^2\right]/\left(f_T(t)\cdot\bar{\sigma}_{jk}(t)\right)^2$$
$$\leq 2\cdot h\cdot M^2\cdot\frac{1}{\underline{f}_T}\cdot\left(3\cdot\overline{f}_T\cdot\|K^+\|_2^2\cdot M_X^4\cdot\frac{\log^2(2s)}{h} + 2\left(\overline{f}_T\cdot M_\sigma + \mathcal{O}(h)\right)^2\right)$$
$$\leq 8\cdot M^2\cdot\frac{1}{\underline{f}_T}\cdot\overline{f}_T\cdot M_X^4\cdot\|K^+\|_2^2\cdot\log^2(2s),$$

where the second inequality holds by the fact that $\Theta(t)\in\mathcal{U}_{s,m,M}$.

Similarly, to obtain an upper bound for I_2, we use the fact from 28 that

$$\sup_{t\in[a,b]}\mathbb{E}\left[k_t^{+2}\right] \leq \frac{3}{h}\cdot\overline{f}_T\cdot\|K^+\|_2^2. \tag{48}$$

By Holder's inequality, we have

$$
\begin{aligned}
I_2 &\leq 2 \cdot h \sup_{t\in[a,b]} \max_{j,k\in[p]} \mathbb{E}\left[\left(\|\Theta_j(t)\|_1 \cdot \|\Theta_k(t)\|_1 \cdot \max_{(j,k)\in E(t)} |\Sigma^+_{jk}(t)| \cdot |k^+_t|\right)^2\right] / \left(f_T(t)\cdot \bar{\sigma}_{jk}(t)\right)^2 \\
&\leq 2 \cdot h \cdot M^2 \cdot \frac{1}{\underline{f}_T} \cdot M_\sigma^2 \sup_{t\in[a,b]} \mathbb{E}\left[k_t^{+2}\right] \\
&\leq 6 \cdot M_\sigma^2 \cdot M^2 \cdot \frac{1}{\underline{f}_T} \cdot \overline{f}_T \cdot \|K^+\|_2^2,
\end{aligned}
$$

where the second inequality holds by Assumption 2 and by the fact that $\Theta(t) \in \mathcal{U}_{s,m,M}$, and the last inequality holds by 48.

Combing the upper bounds for I_1 and I_2, we have

$$
\sup_{t\in[a,b]} \max_{j,k\in[p]} \mathbb{E}\left[J^{+2}_{t,jk}(T,X)\right] \leq 8 \cdot M^2 \cdot \frac{1}{\underline{f}_T} \cdot \overline{f}_T \|K^+\|_2^2 \cdot \left(M_\sigma^2 + M_X^4 \cdot \log^2(2s)\right) \leq C \cdot \log^2 s,
$$

for sufficiently large $C > 0$. Thus, we have

$$
\begin{aligned}
&\sup_{t\in[a,b]} \max_{j,k\in[p]} \mathbb{E}\left[\left(J^+_{t,jk}(T,X) - J^-_{t,jk}(T,X)\right)^2\right] / \left(f_T(t)\cdot\bar{\sigma}_{jk}(t)\right)^2 \\
&\leq 2 \sup_{t\in[a,b]} \max_{j,k\in[p]} \mathbb{E}\left[J^{+2}_{t,jk}(T,X)\right] / \left(f_T(t)\cdot\bar{\sigma}_{jk}(t)\right)^2 + 2 \sup_{t\in[a,b]} \max_{j,k\in[p]} \mathbb{E}\left[J^{-2}_{t,jk}(T,X)\right] / \left(f_T(t)\cdot\bar{\sigma}_{jk}(t)\right)^2 \\
&\leq 4C \cdot \log^2 s = \sigma_J^2 \qquad (49)
\end{aligned}
$$

Covering number for the function class $\mathcal{J}$: First, we note that the function class $\mathcal{J}^+$ is generated from the addition of two function classes

$$
\mathcal{J}^{+(1)}_{jk} = \{J^{+(1)}_{t,jk} | t \in [a,b]\}
$$

and

$$
\mathcal{J}^{+(2)}_{jk} = \{J^{+(2)}_{t,jk} | t \in [a,b]\}.
$$

Thus, to obtain the covering number of $\mathcal{J}$, we first obtain the covering numbers for the function classes $\mathcal{J}^{+(1)}_{jk}$ and $\mathcal{J}^{+(2)}_{jk}$. Then, we apply Lemma 7.3 to obtain the covering number of the function class $\mathcal{J}$. From Lemma 7.6, we have with probability at least $1 - 1/np$,

$$
N\left(\mathcal{J}^{+(1)}_{t,jk}, L_2(Q), \epsilon\right) \leq C \cdot \left(\frac{p^{3/2}\log^{5/3} np}{\sqrt{h}\cdot\epsilon}\right)^6.
$$

Moreover, from Lemma 7.7, we have

$$
N\left(\mathcal{J}^{+(2)}_{t,jk}, L_2(Q), \epsilon\right) \leq C \cdot \left(\frac{p^{1/6}}{h^{4/3}\cdot\epsilon}\right)^6.
$$

Applying Lemma 7.3 with $a_1 = p^{3/2}\log^{5/3} np/h^{1/2}$, $v_1 = 6$, $a_2 = p^{1/6}/h^{4/3}$, and $v_2 = 6$, we have

$$N(\mathcal{J}^+, L_2(Q), \epsilon) \leq C \cdot p^2 \cdot \left(\frac{p^{5/6}\log^{5/6} np}{h^{11/12}\epsilon}\right)^{12},$$

where we multiply p^2 on the right hand side since the function class $\mathcal{J}^+$ is taken over all $j, k \in [p]$. In the end, since $\mathcal{J}$ is generated by addition of $\mathcal{J}^+$ and $\mathcal{J}^-$, we have

$$N(\mathcal{J}, L_2(Q), \epsilon) \leq C \cdot p^2 \cdot \left(\frac{p^{5/6}\log^{5/6} np}{h^{11/12}\epsilon}\right)^{24}.$$

Thus, for function class $\mathcal{J}' = \{\left(J^+_{t,jk} - J^-_{t,jk}\right) / \left(f_T(t) \cdot \tilde{\sigma}_{jk}(t)\right) | t \in [0,1], j, k \in [p]\}$, we have

$$N(\mathcal{J}', L_2(Q), p\epsilon) \leq C \cdot p^2 \cdot \left(\frac{p^{5/6}\log^{5/6} np}{h^{11/12}\epsilon}\right)^{24}.$$

So,

$$N(\mathcal{J}', L_2(Q), p\epsilon) \leq C \cdot p^2 \cdot \left(\frac{p^{11/6}\log^{5/6} np}{h^{11/12}\epsilon}\right)^{24}. \tag{50}$$

Application of Theorem A.1 in [5]: Applying Theorem A.1 in [5] with $a = p^{11/12} \cdot h^{-5/12} \cdot log^{-1/6} np$, $b = C \cdot \log(ns)/\sqrt{h}$, $\sigma_J = C \cdot \log s$, $\nu = 24$, and

$$K_n = A\nu \cdot (\log n \vee \log(ab/\sigma_J)) = C \cdot \log(p/h).$$

For sufficiently large constant $A, C > 0$, there exists a random process W such that for any $\gamma \in (0, 1)$,

$$P\left(|\tilde{U}_{00} - W| \geq C \cdot \left[\frac{bK_n}{(\gamma n)^{1/2}} + \frac{(b\sigma_J)^{1/2}K_n^{3/4}}{\gamma^{1/2}n^{1/4}} + \frac{b^{1/3}\sigma_J^{2/3}K_n^{2/3}}{\gamma^{1/3}n^{1/6}}\right]\right) \leq C' \cdot \left(\gamma + \frac{\log n}{n}\right)$$

for some absolute constant C'. Picking $\gamma = \left(\log^2(ns)\log^4(s)\log^4(p/h)/(nh)\right)^{1/8}$, we have

$$P\left(|\tilde{U}_{00} - W| \geq C \cdot \left(\frac{\log^2(ns)\log^4(s)\log^4(p/h)}{nh}\right)^{1/8}\right) \leq C' \cdot \left(\frac{\log^2(ns)\log^4(s)\log^4(p/h)}{nh}\right)^{1/8}. \tag{51}$$

Combine 46 and 51, we get

$$P\left(|U_{00}-W| \geq C \cdot \left(\frac{\log^2(ns)\log^4(s)\log^4(p/h)}{nh}\right)^{1/8}\right)$$
$$\leq C' \cdot \left(\frac{\log^2(ns)\log^4(s)\log^4(p/h)}{nh}\right)^{1/8}.$$

6.4 Proof of Lemma 5.4

Recall from the proof of Lemma 5.3 that

$$U_{00} = \sup_{t\in[a,b]} \max_{(j,k)\in E(t)} \sqrt{n}\left|\left(\frac{\sum_{i\in[n]} J^+_{t,jk}(T_i,X_i)}{\sum_{i\in[n]} K^+_h(t_i-t)} - \frac{\sum_{i\in[n]} J^-_{t,jk}(T_i,X_i)}{\sum_{i\in[n]} K^-_h(t_i-t)}\right)\right| / \tilde{\sigma}_{jk}(t).$$

We note that

$$U^B_{00} = \sup_{t\in[a,b]} \max_{(j,k)\in E(t)} \sqrt{n}\left|\left(\frac{\sum_{i\in[n]} J^+_{t,jk}(T_i,X_i)\cdot\xi_i}{\sum_{i\in[n]} K^+_h(t_i-t)} - \frac{\sum_{i\in[n]} J^-_{t,jk}(T_i,X_i)\cdot\xi_i}{\sum_{i\in[n]} K^-_h(t_i-t)}\right)\right| / \tilde{\sigma}_{jk}(t),$$

where $\xi_i \overset{i.i.d.}{\sim} N(0,1)$. To show that the term $|W - U^B_{00}|$ can be controlled, similar to proof of Lemma 5.3, we use

$$\tilde{U}^B_{00} = \sup_{t\in[a,b]} \max_{(j,k)\in E(t)} \frac{1}{\sqrt{n}}\left|\left(\sum_{i\in[n]} J^+_{t,jk}(T_i,X_i) - \sum_{i\in[n]} J^-_{t,jk}(T_i,X_i)\right)\cdot\xi_i\right| / \left(f_T(t)\bar{\sigma}_{jk}(t)\right)$$

to approximate U^B_{00} and we apply Theorem A.2 in [5] to controll $|\tilde{U}^B_{00} - W|$.

Let

$$\psi_n = \sqrt{\frac{\sigma_J^2 K_n}{n}} + \left(\frac{b^2\sigma_J^2 K_n^3}{n}\right)^{1/4}$$

and

$$\gamma_n(\delta) = \frac{1}{\delta}\left(\frac{b^2\sigma_J^2 K_n^3}{n}\right)^{1/4} + \frac{1}{n},$$

as defined in Theorem A.2 in [5]. From the proof of Lemma 5.3, we have $b = C\cdot\log(ns)/\sqrt{h}$, $\sigma_J = C\cdot\log s$, and $K_n = C\cdot\log(p/h)$. Since $b^2K_n = C\log^2(ns)\cdot\log(p/h)/h \leq nC^2\cdot\log^2 s = n\sigma_J^2$ for sufficiently large n, there exists a constant $C'' > 0$ such that

$$P\left(|U^B_{00} - W| > \psi_n + \delta \,\middle|\, \{(T_i,X_i)\}_{i\in[n]}\right) \leq C''\cdot\gamma_n(\delta),$$

with probability at least $1-3/n$. Choosing $\delta = \left(\log^2(ns)\cdot\log^2(s)\cdot\log^3(p/h)/\left(nh\right)\right)^{1/8}$, we have

$$P\left(\left|U_{00}^B - W\right| > C\cdot\left(\frac{\log^2(ns)\cdot\log^2(s)\cdot\log^3(p/h)}{nh}\right)^{1/8}\left|\{(T_i,X_i)\}_{i\in[n]}\right.\right)$$
$$\leq C''\cdot\left(\frac{\log^2(ns)\cdot\log^2(s)\cdot\log^3(p/h)}{nh}\right)^{1/8},$$

with probability at least $1-3/n$.

6.5 Proof of Lemma 5.5

Similar to the proof of Lemma 5.2, it suffices to obtain upper bounds for the terms $|U_E^B - U_0^B|$ and $|U_0^B - U_{00}^B|$. Throungh the proof of this lemma, it conditions on the data $\{(T_i,X_i)\}_{i\in[n]}$. We will show that $|U_E^B - U_{00}^B|$ is upper bounded by the quantity

$$C\cdot\left[s\cdot\sqrt{\frac{\log^4(np/\sqrt{h})}{nh^2}}\right]$$

with high probability for sufficiently large constant $C>0$. Following the results of 43 and the proof of Lemma 9 in [12], we directly give the upper bounds

Upper bound for $|U_E^B - U_0^B|$: Applying the Dudley's inequality(Corollary 2.2.8 in[14]) and the Borell's inequality(Proposition A.2.1. in [14]), we have

$$\begin{aligned}|U_E^B - U_0^B| &\leq C\cdot s\cdot\sqrt{h\log^3(np/\sqrt{h})} + C\cdot s\cdot\sqrt{\frac{\log^4(np/\sqrt{h})}{nh^2}}\\ &\leq C\cdot s\cdot\sqrt{\frac{\log^4(np/\sqrt{h})}{nh^2}}\end{aligned}\tag{52}$$

with probability at least $1-4\sqrt{h}/np$, for some positive constant C. The last inequality holds by the assumption that $nh^3 = o(1)$.

Upper bound for $|U_0^B - U_{00}^B|$: By Lemma 3.2 and the results that if $\xi_i \overset{i.i.d.}{\sim} N(0,1)$, then

$$P\left(\left|\frac{1}{n}\sum_{i\in[n]}\xi_i\right| > \sqrt{\frac{2\log n}{n}}\right) \leq \frac{1}{n},$$

we obtain the Upper bound for $|U_0^B - U_{00}^B|$

$$|U_0^B - U_{00}^B| \leq C\cdot\sqrt{h^3\log n}\tag{53}$$

with probability at least $1-1/n$, for some positive constant C.

Combining the upper bounds 52 and 53, and applying the union bound, we have

$$\begin{aligned}
&P\left(|U_E^B - U_{00}^B| \geq C\cdot s\cdot\sqrt{\frac{\log^4(np/\sqrt{h})}{nh^2}} + C\cdot\sqrt{h^3\log n}\middle|\{(T_i,X_i)\}_{i\in[n]}\right)\\
&\leq P\left(|U_E^B - U_0^B| \geq C\cdot s\cdot\sqrt{\frac{\log^4(np/\sqrt{h})}{nh^2}}\middle|\{(T_i,X_i)\}_{i\in[n]}\right)\\
&+P\left(|U_0^B - U_{00}^B| \geq C\cdot\sqrt{h^3\log n}\middle|\{(T_i,X_i)\}_{i\in[n]}\right)\\
&\leq 4\sqrt{h}/np + \frac{1}{n} \leq \frac{2}{n}.
\end{aligned}$$

By the assumption that $nh^3 = o(1)$, we conclude

$$P\left(|U_E^B - U_{00}^B| \geq C\cdot s\cdot\sqrt{\frac{\log^4(np/\sqrt{h})}{nh^2}}\middle|\{(T_i,X_i)\}_{i\in[n]}\right) \leq \frac{2}{n}.$$

7 Technical Lemmas on Covering Number

In this section, we present some technical lemmas on the covering number of some function classes. Lemma 7.1(Lemma 3 in [7]) provides an upper bound on the covering number for the class of functions of bounded variation. Lemma 7.2(Lemma 14 in [12]) provides an upper bound on the covering number of a class of Lipschitz functions. Lemma 7.3(Lemma 15 in [12]) provides an upper bound on the covering numbers for function classes generated from the product and addition of two function classes.

Lemma 7.1 *(Lemma 3 in [7]) Let $K:\mathbb{R}\to\mathbb{R}$ be a function of bounded variation. Define the function class $\mathcal{F}_h = \{K((t-\cdot)/h)|t\in\mathbb{R}\}$. Then, there exists $C_K<\infty$ independent of h and K such that for all $0<\epsilon<1$,*

$$\sup_Q N(\mathcal{F}_h, L_2(Q), \epsilon) \leq \left(\frac{2\cdot C_K\cdot\|K\|_{TV}}{\epsilon}\right)^4,$$

where $\|K\|_{TV}$ is the total variation norm of the function K and Q is any probability mersure.

Lemma 7.2 *(Lemma 14 in [12]) For any $0<a<b<1$, let $f(l)$ be a Lipschitz function defined on $[a,b]$such that $|f(l)-f(l')|\leq L_f\cdot|l-l'|$ for any $l,l'\in[a,b]$. We define the constant* *function class* $\mathcal{F} = \{g_l : f(l) | l \in [a,b]\}$*. For any probability measure* Q*, the covering number of the function class* $\mathcal{F}$ *satisfies*

$$N(\mathcal{F}, L_2(Q), \epsilon) \leq \frac{L_f}{\epsilon},$$

where $\epsilon \in (0,1)$*.*

Lemma 7.3 *(Lemma 15 in [12]) Let* $\mathcal{F}_1$ *and* $\mathcal{F}_2$ *be two function classes satifying*

$$N(\mathcal{F}_1, L_2(Q), a_1\epsilon) \leq C_1\epsilon^{-v_1}$$

and

$$N(\mathcal{F}_2, L_2(Q), a_2\epsilon) \leq C_2\epsilon^{-v_2}$$

for some $C_1, C_2, a_1, a_2, v_1, v_2 > 0$ *and any* $0 < \epsilon < 1$*. Define* $\|F_l\|_\infty = \sup_{f \in \mathcal{F}_l} \|f\|_\infty$ *for* $l = 1, 2$ *and* $U = \|\mathcal{F}_1\|_\infty \vee \|\mathcal{F}_2\|_\infty$*. For the function classes* $F_\times = \{f_1 f_2 | f_1 \in \mathcal{F}_1, f_2 \in \mathcal{F}_2\}$ *and* $\mathcal{F}_+ = \{f_1 + f_2 | f_1 \in \mathcal{F}_1, f_2 \in \mathcal{F}_2\}$*, we have for any* $\epsilon \in (0,1)$*,*

$$N(\mathcal{F}_\times, L_2(Q), \epsilon) \leq C_1 \cdot C_2 \cdot \left(\frac{2a_1U}{\epsilon}\right)^{v_1} \cdot \left(\frac{2a_2U}{\epsilon}\right)^{v_2}$$

and

$$N(\mathcal{F}_+, L_2(Q), \epsilon) \leq C_1 \cdot C_2 \cdot \left(\frac{2a_1}{\epsilon}\right)^{v_1} \cdot \left(\frac{2a_2}{\epsilon}\right)^{v_2}.$$

Now, we introduce Lemma 7.4 and 7.5. The proof of Lemma 7.4 and 7.5 is a direct application of Lemma 7.1, 7.2 and 7.3.

Lemma 7.4 *(Lemma 16 in [12]) Let* $\omega_t^\pm(u) = K_h^\pm(u - t)$*. For any* $0 < a < b < 1$*, we define the function classes*

$$\mathcal{K}_1^\pm = \{\omega_t^\pm(\cdot) | t \in [a,b]\}$$

and

$$\mathcal{K}_2^\pm = \{\mathbb{E}[\omega_t^\pm(T)] | t \in [a,b]\}.$$

Given Assumptions 1-2, we have for any $\epsilon \in (0,1)$*,*

$$\sup_Q N(\mathcal{K}_1^\pm, L_2(Q), \epsilon) \leq \left(\frac{2 \cdot C_{K^\pm} \cdot \|K^\pm\|_{TV}}{h\epsilon}\right)^4$$

and

$$\sup_Q N(\mathcal{K}_2^\pm, L_2(Q), \epsilon) \leq \frac{2}{h\epsilon} \cdot \|K^\pm\|_{TV} \cdot \overline{f}_T.$$

Moreover, let $k_t^\pm(u) = \omega_t^\pm(u) - \mathbb{E}[\omega_t^\pm(T)]$ *and let* $\mathcal{K}^\pm = \{k_t^\pm(\cdot) | t \in [a,b]\}$*. We have*

$$\sup_Q N(\mathcal{K}^\pm, L_2(Q), \epsilon) \leq \left(\frac{4 \cdot \|K^\pm\|_{TV} \cdot C_{K^\pm}^{4/5} \cdot \overline{f}_T^{1/5}}{h\epsilon}\right)^5.$$

Lemma 7.5 *(Lemma 17 in [12]) Let $g^{\pm}_{t,jk}(u, X_{ij}, X_{ik}) = K^{\pm}_h(u-t)X_{ij}X_{ik}$. For any $0 < a < b < 1$, we define the function classes*

$$\mathcal{G}^{\pm}_{1,jk} = \{g^{\pm}_{t,jk}(\cdot) | t \in [a,b]\}$$

and

$$\mathcal{G}^{\pm}_{2,jk} = \{\mathbb{E}[g^{\pm}_{t,jk}(T, X_j, X_k)] | t \in [a,b]\}.$$

Given assumptions 1-2, for all $\epsilon \in c(0,1)$ and sufficiently large M_X,

$$\sup_Q N(\mathcal{G}^{\pm}_{1,jk}, L_2(Q), \epsilon) \leq \left(\frac{2 \cdot M_X^2 \cdot \log np \cdot C_{K^{\pm}} \cdot \|K^{\pm}\|_{TV}}{h\epsilon} \right)^4$$

and

$$\sup_Q N(\mathcal{G}^{\pm}_{2,jk}, L_2(Q), \epsilon) \leq \frac{2}{h\epsilon} \cdot \|K^{\pm}\|_{TV} \cdot \overline{f}_T \cdot M_\sigma,$$

with probability at least $1 - 1/np$. Moreover, let $q^{\pm}_{t,jk}(u, X_{ij}, X_{ik}) = g^{\pm}_{t,jk}(u, X_{ij}, X_{ik}) - \mathbb{E}[g^{\pm}_{t,jk}(T, X_j, X_k)]$ and let $\mathcal{G}^{\pm}_{jk} = \{q^{\pm}_{t,jk}(\cdot) | t \in [a,b]\}$. We have

$$\sup_Q N(\mathcal{G}^{\pm}_{jk}, L_2(Q), \epsilon) \leq \left(\frac{4 \cdot \|K^{\pm}\|_{TV} \cdot C_{K^{\pm}}^{4/5} \cdot \overline{f}_T^{1/5} \cdot M_\sigma^{1/5} \cdot M_X^{8/5} \cdot \log^{4/5} np}{h\epsilon} \right)^5$$

with probability at least $1 - 1/np$.

Lemma 7.6 *(Lemma 18 in [12]) For any $0 < a < b < 1$, let $\mathcal{J}^{\pm(1)}_{jk} = \{J^{\pm(1)}_{t,jk} | t \in [a,b]\}$. Given Assumption 1-2, for all probability measure Q on $\mathbb{R}$ and all $\epsilon \in (0,1)$*

$$N(\mathcal{J}^{\pm(1)}_{jk}, L_2(Q), \epsilon) \leq C \cdot \left(\frac{p^{3/2} \log^{5/3} np}{\sqrt{h} \cdot \epsilon} \right)^6,$$

with probability at least $1 - 1/np$, where $C > 0$ is a generic constant that does not depend on p, h, and n.

Lemma 7.7 *(Lemma 19 in [12])For any $0 < a < b < 1$, let $\mathcal{J}^{\pm(2)}_{jk} = \{J^{\pm(2)}_{t,jk} | t \in [a,b]\}$. Given Assumption 1-2, for all probability measure Q on $\mathbb{R}$ and all $\epsilon \in (0,1)$*

$$N(\mathcal{J}^{\pm(2)}_{jk}, L_2(Q), \epsilon) \leq C \cdot \left(\frac{p^{1/6}}{h^{4/3} \cdot \epsilon} \right)^6,$$

where $C > 0$ is a generic constant that does not depend on p, h, and n.

8 Technical Lemmas on Empirical Process

In this section, we present some existing tools on empirical process. The following Lemma claims that the supreme of any empirical process is concentrated near its mean. It directly follows from Theorem 2.3 in [1], an improvement on Rio's version[10] of Talagrand's inequality[11].

Lemma 8.1 *(Theorem A.1 in [13]) Let $X_1, X_2, \ldots, X_n$ be independent random variables and let $\mathcal{F}$ be a function class such that there exists η and τ^2 satisfying*

$$\sup_{f\in\mathcal{F}} \|f\|_\infty \le \eta$$

and

$$\sup_{f\in\mathcal{F}} \frac{1}{n}\sum_{i\in n} Var(f(X_i)) \le \tau^2.$$

Define

$$Y = \sup_{f\in\mathcal{F}} \left| \frac{1}{n}\sum_{i\in[n]} (f(X_i) - \mathbb{E}[f(X_i)]) \right|.$$

Then, for any $t > 0$,

$$P\left(Y \ge \mathbb{E}[Y] + t\sqrt{2(\tau^2 + 2\eta\mathbb{E}[Y])} + 2t^2\eta/3\right) \le \exp\left(-nt^2\right).$$

The above inequality involves evaluating the expectation of the supreme of the empirical process. The following Lemma directly follows from Theorem 3.12 in [8] and Corollary 5.1 in [6]. It provides an upper bound on the expectation of the supreme of the empirical process as function of its covering number.

Lemma 8.2 *(Lemma F.2 in [9]) Assume that the functions in $\mathcal{F}$ defined on $\mathcal{X}$ are uniformly bounded by a constant U and $F(\cdot)$ is the envelope of $\mathcal{F}$ such that $|f(x)| \le F(x)$ for all $x \in \mathcal{X}$ and $f \in \mathcal{F}$. Let $\sup_{f\in\mathcal{F}} \mathbb{E}[f^2] \le \sigma_P^2 \le \|F\|^2_{L_2(\mathbb{P}_n)}$. Let $X_1, \ldots, X_n$ be i.i.d. copies of the random variables X. We denote the empirical measure as $\mathbb{P}_n = \frac{1}{n}\sum_{i\in[n]} \delta_{X_i}$. If for some $A, V > 0$ and for all $\epsilon > 0$ and $n \ge 1$, the covering entropy satisfies*

$$N(\mathcal{F}, L_2(\mathbb{P}_n), \epsilon) \le \left(\frac{A\|F\|_{L_2(\mathbb{P}_n)}}{\epsilon}\right)^V,$$

then there exists a universal constant C such that

$$\mathbb{E}\left[\sup_{f\in\mathcal{F}} \frac{1}{n}\left|\sum_{i\in[n]} (f(X_i) - \mathbb{E}[f(X_i)])\right|\right] \le C\left[\sqrt{\frac{V}{n}}\sigma_P\sqrt{\log\left(\frac{A\|F\|_{L_2(\mathbb{P}_n)}}{\sigma_P}\right)} + \frac{VU}{n}\log\left(\frac{A\|F\|_{L_2(\mathbb{P}_n)}}{\sigma_P}\right)\right].$$

References

[1] O. Bousquet. A bennett concentration inequality and its application to suprema of empirical processes. *Comptes Rendus Mathematique*, 334(6):495–500, 2002.

[2] T. Cai, W. Liu, and X. Luo. A constrained l_1 minimization approach to sparse precision matrix estimation. *Journal of the American Statistical Association*, 106(494):594–607, 2011.

[3] T. T. Cai, W. Liu, H. H. Zhou, et al. Estimating sparse precision matrix: Optimal rates of convergence and adaptive estimation. *The Annals of Statistics*, 44(2):455–488, 2016.

[4] V. Chernozhukov, D. Chetverikov, K. Kato, et al. Gaussian approximations and multiplier bootstrap for maxima of sums of high-dimensional random vectors. *The Annals of Statistics*, 41(6):2786–2819, 2013.

[5] V. Chernozhukov, D. Chetverikov, K. Kato, et al. Anti-concentration and honest, adaptive confidence bands. *The Annals of Statistics*, 42(5):1787–1818, 2014.

[6] V. Chernozhukov, D. Chetverikov, K. Kato, et al. Gaussian approximation of suprema of empirical processes. *The Annals of Statistics*, 42(4):1564–1597, 2014.

[7] E. Giné and R. Nickl. An exponential inequality for the distribution function of the kernel density estimator, with applications to adaptive estimation. *Probability Theory and Related Fields*, 143(3-4):569–596, 2009.

[8] V. Koltchinskii. *Oracle Inequalities in Empirical Risk Minimization and Sparse Recovery Problems: Ecole d'Eté de Probabilités de Saint-Flour XXXVIII-2008*, volume 2033. Springer Science & Business Media, 2011.

[9] J. Lu, M. Kolar, and H. Liu. Post-regularization confidence bands for high dimensional nonparametric models with local sparsity. *arXiv preprint arXiv:1503.02978*, 2015.

[10] E. Rio. Une inégalité de bennett pour les maxima de processus empiriques. In *Annales de l'IHP Probabilités et statistiques*, volume 38, pages 1053–1057, 2002.

[11] M. Talagrand. New concentration inequalities in product spaces. *Inventiones mathematicae*, 126(3):505–563, 1996.

[12] K. M. Tan, J. Lu, T. Zhang, and H. Liu. Estimating and inferring the maximum degree of stimulus-locked time-varying brain connectivity networks. *arXiv preprint arXiv:1905.11588*, 2019.

[13] S. A. Van de Geer et al. High-dimensional generalized linear models and the lasso. *The Annals of Statistics*, 36(2):614–645, 2008.

[14] A. W. Van Der Vaart and J. A. Wellner. Weak convergence. In *Weak convergence and empirical processes*, pages 16–28. Springer, 1996.