WAVELET ESTIMATIONS FOR LONGITUDINAL DATA

Trung Tu Nguyen

University of Paris VII

Supplementary Material

We present in detail the proofs of the main theorems of our paper. This supplement contains two sections. The first section presents the proofs of Theorem 3.1 and Theorem 3.2 (the upper and the lower bounds of the quadratic risk of the mean estimation). The second section proves consecutively Theorem 4.3 and Theorem 4.4 which correspond to the known/unknown-mean-function cases for the covariance estimations.

S1. Mean estimation

S1.1. Upper bound of the mean estimation

Proof of Theorem 3.1:

For each fixed integer i, using Parseval's equality we have:

$$E\|\hat{\mu}_{Mi}(\mathcal{E}) - P_M(\mathcal{E})\mu\|_2^2 = \sum_{k=1}^{M} E\left(\{\hat{\alpha}_{ki} - \alpha_k\}^2\right), \tag{1.1}$$

with

$$E\left(\{\hat{\alpha}_{ki} - \alpha_k\}^2\right) =$$
$$E\left[\left(\frac{1}{m}\sum_{j=1}^{m}\mu(T_{ij})e_k(T_{ij}) - \alpha_k\right) + \left(\frac{1}{m}\sum_{j=1}^{m}\delta_i(T_{ij})e_k(T_{ij})\right) + \left(\frac{1}{m}\sum_{j=1}^{m}\epsilon_{ij}e_k(T_{ij})\right)\right]^2.$$

Always with i fixed, conditionally on $T_{ij}, j = 1, \ldots, m$, we have:

- $A_{ki1} = \frac{1}{m}\sum_{j=1}^{m}\mu(T_{ij})e_k(T_{ij}) - \alpha_k$ is a constant,
- $A_{ki2} = \frac{1}{m}\sum_{j=1}^{m}\delta_i(T_{ij})e_k(T_{ij})$, $A_{ki3} = \frac{1}{m}\sum_{j=1}^{m}\epsilon_{ij}e_k(T_{ij})$ are independent and $E[A_{kil}|T_{ij}, j = 1, \ldots, m] = 0$ with $l = 2, 3$.

So:

$$\begin{aligned} E\{\hat{\alpha}_{ki} - \alpha_k\}^2 &= E\left[E\left(\{\hat{\alpha}_{ki} - \alpha_k\}^2|T_{ij}, j = 1, \ldots, m\right)\right] \\ &= E\left[E\left(A_{ki1}^2 + A_{ki2}^2 + A_{ki3}^2|T_{ij}, j = 1, \ldots, m\right)\right] \\ &= E[A_{ki1}^2] + E[A_{ki2}^2] + E[A_{ki3}^2]. \end{aligned} \tag{1.2}$$

In the next paragraph, we will prove three following inequalities:

- $$E[A_{ki1}^2] \le \frac{\|\mu\|_\infty^2}{m}, \tag{1.3}$$

- $$E[A_{ki3}^2] \le \frac{\sigma^2}{m} \tag{1.4}$$

- $$E[A_{ki2}^2] \le \frac{\|\omega\|_\infty}{m} + \sum_{l\ge 1} \lambda_l^2 < \eta_l, e_k >^2, \tag{1.5}$$

  with the functional principal component expansion of $\delta(u) = \sum_{l\ge1} \zeta_l \eta_l(u)$ and $E(\zeta_l^2) = \lambda_l^2$.

Remark:

$$\begin{aligned} E\|P_M(\mathcal{E})\delta\|_2^2 &= \sum_{k=1}^M E(< \delta, e_k >^2) = \sum_{k=1}^M E(< \sum_{l\ge1} \zeta_l \eta_l(.), e_k(.) >^2) \\ &= \sum_{k=1}^M E\left(\left[\sum_{l\ge1} \zeta_l < \eta_l, e_k > \right]^2 \right) = \sum_{k=1}^M \sum_{l\ge1} \lambda_l^2 < \eta_l, e_k >^2 . \end{aligned}$$

So,

$$\begin{aligned} E\|\hat{\mu}_{Mi}(\mathcal{E}) - P_M(\mathcal{E})\mu\|_2^2 &= \sum_{k=1}^M \big(E[A_{ki1}^2] + E[A_{ki2}^2] + E[A_{ki3}^2] \big) \\ &\le \sum_{k=1}^M \left(\frac{\|\mu\|_\infty^2}{m} + \frac{\sigma^2}{m} + \frac{2\|\omega\|_\infty}{m} + 2\sum_{l\ge1} \lambda_l^2 < \eta_l, e_k >^2 \right) \\ &= \frac{M}{m} \big(\|\mu\|_\infty^2 + \sigma^2 + 2\|\omega\|_\infty \big) + 2\sum_{k=1}^M \sum_{l\ge1} \lambda_l^2 < \eta_l, e_k >^2 \\ &= \frac{M}{m} \big(\|\mu\|_\infty^2 + \sigma^2 + 2\|\omega\|_\infty \big) + 2E\|P_M(\mathcal{E})\delta\|_2^2. \end{aligned}$$

With the construction of $\hat{\alpha}_{ki}$, under conditions on ϵ, δ, X_{ij} and T_{ij} we have:

$$\begin{aligned} E(\hat{\alpha}_{ki}) &= E\big(\frac{1}{m} \sum_{j=1}^m Y_{ij} e_k(T_{ij}) \big) \\ &= \frac{1}{m} \sum_{j=1}^m E\{\mu(T_{ij}) e_k(T_{ij})\} + \frac{1}{m} \sum_{j=1}^m E\{\delta_i(T_{ij}) e_k(T_{ij})\} + \frac{1}{m} \sum_{j=1}^m E(\epsilon_{ij}) E(e_k(T_{ij})) \\ &= \int_0^1 \mu(t) e_k(t) dt = \alpha_k. \end{aligned}$$

So

$$E(\hat{\alpha}_k) = \frac{1}{n} \sum_{i=1}^n E(\hat{\alpha}_{ki}) = \alpha_k \quad \Rightarrow \quad E\{\hat{\mu}_M(\mathcal{E})\} = \sum_{k=1}^M E\{\hat{\alpha}_k\} e_k = \sum_{k=1}^M \alpha_k e_k = P_M(\mathcal{E})\mu$$

Because n models $(Y_{i.}, T_{i.}), i = 1, \dots, n$ are independent and $E\{\hat{\mu}_{Mi}(\mathcal{E}) - P_M(\mathcal{E})\mu\} = 0$, so we have:

$$\begin{aligned} E\|\hat{\mu}_M(\mathcal{E}) - P_M(\mathcal{E})\mu\|_2^2 &= E\|\frac{1}{n}\sum_{i=1}^{n}\{\hat{\mu}_{Mi}(\mathcal{E}) - P_M(\mathcal{E})\mu\}\|_2^2 = \frac{1}{n^2}\sum_{i=1}^{n} E\|\hat{\mu}_{Mi}(\mathcal{E}) - P_M(\mathcal{E})\mu\|_2^2 \\ &\leq \frac{1}{n}\left(\frac{M}{m}\left(\|\mu\|_\infty^2 + \sigma^2 + 2\|\omega\|_\infty\right) + 2E\|P_M(\mathcal{E})\delta\|_2^2\right). \end{aligned} \tag{1.6}$$

Now, we only need to prove (1.3), (1.4) and (1.5)

- **First part:** Expected value of A_{ki1}^2

  $R_{ijk} = (\mu(T_{ij})e_k(T_{ij}) - \alpha_k)$ are independent and identically distributed with:

$$E[R_{ijk}] = E[\mu(T_{ij})e_k(T_{ij}) - \alpha_k] = \int_0^1 \mu(u)e_k(u)du - \alpha_k = 0,$$

  and

$$E[R_{ijk}^2] = E[\{\mu(T_{ij})e_k(T_{ij})\}^2] - \alpha_k^2 \leq E[\{\mu(T_{ij})e_k(T_{ij})\}^2] = \int_0^1 \mu^2(t)e_k^2(t)dt.$$

  So

$$E[A_{ki1}^2] = \frac{1}{m^2}\sum_{j=1}^{m} E[R_{ijk}^2] \leq \frac{1}{m}\int_0^1 \mu^2(t)e_k^2(t)dt \leq \frac{\|\mu\|_\infty^2}{m}.$$

- **Third part:** Expected value of A_{ki3}^2

$$\begin{aligned} E[A_{ki3}^2] = E\big[(\frac{\sum_{j=1}^m \epsilon_{ij}e_k(T_{ij})}{m})^2\big] &= E\big(E\{\big[(\frac{\sum_{j=1}^m \epsilon_{ij}e_k(T_{ij})}{m})^2\big]|T_{ij} : j = 1, \dots, m\}\big) \\ &= \frac{1}{m^2}\sigma^2 E[\sum_{j=1}^{m} e_k^2(T_{ij}))] = \frac{\sigma^2}{m}\int_0^1 e_k^2(t)dt = \frac{\sigma^2}{m}. \end{aligned}$$

- **Second part:** Expected value of A_{ki2}^2

  Using the spectral decomposition, we have :

$$\delta_i(T_{ij})e_k(T_{ij}) = e_k(T_{ij})\sum_{l\geq 1}\zeta_{il}\eta_l(T_{ij}) \Rightarrow \sum_{j=1}^{m}\delta_i(T_{ij})e_k(T_{ij}) = \sum_{l\geq 1}\zeta_{il}[\sum_{j=1}^{m} e_k(T_{ij})\eta_l(T_{ij})].$$

  We denote $< \eta_l, e_k >= \int \eta_l(t)e_k(t)dt = E\left(e_k(T_{ij})\eta_l(T_{ij})\right)$, and

$$\begin{aligned} \frac{1}{m}\sum_{j=1}^{m}\delta_i(T_{ij})e_k(T_{ij}) &= \sum_{l\geq 1}\zeta_{il}\frac{[\sum_{j=1}^m e_k(T_{ij})\eta_l(T_{ij})]}{m} \\ &= \sum_{l\geq 1}\zeta_{il}\sum_{j=1}^{m}\frac{[e_k(T_{ij})\eta_l(T_{ij}) - < \eta_l, e_k >]}{m} + \sum_{l\geq 1}\zeta_{il} < \eta_l, e_k > . \end{aligned}$$

So

$$\begin{aligned}
E[A_{ki2}^2] &= E\left[\left(\frac{1}{m}\sum_{j=1}^{m}\delta_i(T_{ij})e_k(T_{ij})\right)^2\right] \\
&= E\left[\left(\sum_{l\geq 1}\zeta_{il}\sum_{j=1}^{m}\frac{[e_k(T_{ij})\eta_l(T_{ij})- < \eta_l, e_k >]}{m} + \sum_{l\geq 1}\zeta_{il} < \eta_l, e_k >\right)^2\right] \\
&\leq 2\left[E\left(\sum_{l\geq 1}\zeta_{il}\sum_{j=1}^{m}\frac{[e_k(T_{ij})\eta_l(T_{ij})- < \eta_l, e_k >]}{m}\right)^2 + E\left(\sum_{l\geq 1}\zeta_{il} < \eta_l, e_k >\right)^2\right] \\
&= 2(E[B_{ik1}^2] + E[B_{ik2}^2]).
\end{aligned}$$

We have:

$$E[B_{ik2}^2] = E\left(\left[\sum_{l\geq 1}\zeta_l < \eta_l, e_k >\right]^2\right) = \sum_{l\geq 1}\lambda_l^2 < \eta_l, e_k >^2 .$$

So we only need to prove:

$$E[B_{ik1}^2] \leq \frac{\|\omega\|_\infty}{m} \Rightarrow (1.5).$$

Conditionally on $T_{ij}, j = 1, \ldots, m$,

$$\theta_{il} = \zeta_{il}\sum_{j=1}^{m}\frac{[e_k(T_{ij})\eta_l(T_{ij})- < \eta_l, e_k >]}{m}$$

are uncorrelated random variables. It means that for i fixed,

$$E[\theta_{il}\theta_{ih}|T_{ij}; j = 1, \ldots, m] = \lambda_l^2\left(\sum_{j=1}^{m}\frac{[e_k(T_{ij})\eta_l(T_{ij})- < \eta_l, e_k >]}{m}\right)^2\delta_{lh} \text{ with } \delta_{lh} = 1_{\{h=l\}}.$$

So

$$\begin{aligned}
E[B_{ik1}^2] &= E\left[(\sum_{l\geq 1}\theta_{il})^2\right] = E\left(E\left[(\sum_{l\geq 1}\theta_{il})^2|T_{ij} : j = 1, \ldots, m\right]\right) \\
&= \sum_{l\geq 1}\lambda_l^2 E\left((\sum_{j=1}^{m}\frac{[e_k(T_{ij})\eta_l(T_{ij})- < \eta_l, e_k >]}{m})^2\right) \\
&\leq \sum_{l\geq 1}\lambda_l^2\frac{1}{m^2}\sum_{j=1}^{m}E[e_k^2(T)\eta_l^2(T))] = \frac{1}{m}\sum_{l\geq 1}\lambda_l^2\int_0^1 e_k^2(t)\eta_l^2(t)dt \\
&= \frac{1}{m}\int_0^1 e_k^2(t)\left(\sum_{l\geq 1}\lambda_l^2\eta_l^2(t)\right)dt = \frac{1}{m}\int_0^1 e_k^2(t)E(\delta^2(t))dt \leq \frac{1}{m}\|\omega\|_\infty.
\end{aligned}$$

S1.2. Lower bound

Proof of Theorem 3.2:

In order to prove the lower bound, using the standard schema (see Tsybakov (2004)), the most important part is to bound the Kullback divergences. We have $M+1$ functions $\{\mu_0, \mu_1, \mu_2, \dots, \mu_M\}$ such that:

$$\|\mu_i - \mu_j\|_2^2 \asymp n^{-\frac{2s}{2s+1}}.$$

Conditionally on $\{T_{ij}\}_{\{i,j\}}$, we denote P_k as the joined probability of $(Y_{ij})_{i,j}$ for $\mu = \mu_k$. We should prove:

$$\forall k \leq M, \quad \mathcal{K}(P_k, P_0) \leq C \sum_{i=1}^{n} \sum_{j=1}^{m} \mu_k^2(T_{ij}).$$

Conditionally on $\{T_{ij}\}_{\{i,j\}}$, under the Gaussian assumption on process X, for each i we have a zero mean gaussian vector $(\delta_i(T_{i1}), \delta_i(T_{i2}), \dots, \delta_i(T_{im}))^t$. The independence between ϵ_{ij} and T_{ij} implies

$$(\delta_i(T_{i1}) + \epsilon_{i1}, \delta_i(T_{i2}) + \epsilon_{i2}, \dots, \delta_i(T_{im}) + \epsilon_{im})^t$$

is also a zero mean gaussian vector with the denoted density $p_{i,\epsilon,\delta}(u)$ and the denoted covariance Γ_{im}.

We remark that

$$\Gamma_{im}(j,k) = E\left(\left(\delta(T_{ij}) + \epsilon_{ij}\right)\left(\delta(T_{ik}) + \epsilon_{ik}\right)|T_{il} l = 1, \dots, m\right) = \omega(T_{ij}, T_{ik}) + \sigma^2 \mathbf{1}_{\{j=k\}}.$$

Under the assumption $\|\omega\|_\infty = \|cov(X(u), X(v))\|_\infty \leq C < \infty$, we have:

$$|\Gamma_{im}(j,k)| \leq L_1 = C + \sigma^2.$$

For all vector $v \in \mathbf{R}^m$, we have:

$$|v^t \Gamma_{im} v| = |\sum_{k=1}^{m} v_k \sum_{j=1}^{m} \Gamma_{im}(j,k) v_j| \leq L_1 \sum_{k=1}^{m} \sum_{j=1}^{m} |v_j v_k| \leq L_1 m \sum_{j=1}^{m} v_j^2.$$

By the properties of gaussian vector, we remark :

$$\begin{aligned}
\int \log \frac{p_{i,\epsilon,\delta}(u)}{p_{i,\epsilon,\delta}(u+v)} p_{i,\epsilon,\delta}(u) du &= \int \log \frac{p_{i,\epsilon,\delta}(u_1, \dots, u_m)}{p_{i,\epsilon,\delta}(u_1 + v_1, \dots, u_m + v_m)} p_{i,\epsilon,\delta}(u_1, \dots, u_m) du_1..du_m \\
&= \int \left\{(u+v)^t \Gamma_{im}(u+v) - u^t \Gamma_{im} u\right\} p_{i,\epsilon,\delta}(u) du \\
&= \int (v^t \Gamma_{im}) \left[u p_{i,\epsilon,\delta}(u)\right] du + \int \left[u p_{i,\epsilon,\delta}(u)\right]^t (\Gamma_{im} v) du \\
&\quad + (v^t \Gamma_{im} v) \int p_{i,\epsilon,\delta}(u) du \\
&= v^t \Gamma_{im} v \leq L_1 m \sum_{k=1}^{m} v_k^2,
\end{aligned}$$

where m is a finite integer, so

$$\int \log \frac{p_{i,\epsilon,\delta}(u)}{p_{i,\epsilon,\delta}(u+v)} p_{i,\epsilon,\delta}(u) du \leq C \sum_{j=1}^{m} v_j^2. \tag{1.7}$$

We remark that P_k admits the following density on $\mathcal{R}^{nm}$:

$$p_k(u_{ij} : i = 1, \ldots, n; j = 1, \ldots, m) = \prod_{i=1}^{n} p_{i,\epsilon,\delta}(u_{i1} - \mu_k(T_{i1}), u_{i2} - \mu_k(T_{i2}), \ldots, u_{im} - \mu_k(T_{im})).$$

The Kullback divergence could be written as:

$$\begin{aligned}
\mathcal{K}(P_k, P_0) &= \int \log \frac{dP_k}{dP_0} dP_k \\
&= \int .. \int \log \prod_{i=1}^{n} \frac{p_{i,\epsilon,\delta}(U - \mu_k(T_i))}{p_{i,\epsilon,\delta}(U)} \prod_{i=1}^{n} [p_{i,\epsilon,\delta}(U - \mu_k(T_i)) du_{i1}..du_{im}] \\
&= \sum_{i=1}^{n} \int \log \frac{p_{i,\epsilon,\delta}(U - \mu_k(T_i))}{p_{i,\epsilon,\delta}(U)} [p_{i,\epsilon,\delta}(U - \mu_k(T_i)) du_{i1}..du_{im}],
\end{aligned}$$

with

$$U = (u_{i1}, u_{i2}, \ldots, u_{im})^t, \quad U - \mu_k(T_i) = (u_{i1} - \mu_k(T_{i1}), \ldots, u_{im} - \mu_k(T_{im}))^t.$$

Using (1.7), we have:

$$\mathcal{K}(P_k, P_0) \quad \leq \quad C \sum_{i=1}^{n} [\sum_{j=1}^{m} \mu_k^2(T_{ij})],$$

which gives the lower bound of the quadratic risk.

S2. Covariance estimations

S2.1. Known-mean-function case

Proof of Theorem 4.3 :

For $0 < i \leq n$, we have:

$$\begin{aligned}
E\left((\hat{c}_{k,i} - c_k)^2\right) &= E\left((\tfrac{1}{l}\textstyle\sum_{q=1}^{l} U_{iq} V_{iq} \pi_{kiq} - c_k)^2\right) \\
&= E\left[\left(\tfrac{1}{l}\textstyle\sum_{q=1}^{l} \{\delta_i(R_{iq})\delta_i(S_{iq}) + \epsilon_{iq1}\delta_i(S_{iq}) + \delta_i(R_{iq})\epsilon_{iq2} + \epsilon_{iq1}\epsilon_{iq2}\} \pi_{kiq} - c_k\right)^2\right] \\
&\leq 4\left(E_{ik1} + E_{ik2} + E_{ik3} + E_{ik4}\right),
\end{aligned}$$

where

- $E_{ik1} = E\left((\frac{1}{l}\sum_{q=1}^{l} \delta_i(R_{iq})\delta_i(S_{iq})\pi_{kiq} - c_k)^2\right)$,
- $E_{ik2} = E\left((\frac{1}{l}\sum_{q=1}^{l} \epsilon_{iq1}\delta_i(S_{iq})\pi_{kiq})^2\right)$,
- $E_{ik3} = E\left((\frac{1}{l}\sum_{q=1}^{l} \delta_i(R_{iq})\epsilon_{iq2}\pi_{kiq})^2\right)$,
- $E_{ik4} = E\left((\frac{1}{l}\sum_{q=1}^{l} \epsilon_{iq1}\epsilon_{iq2}\pi_{kiq})^2\right)$

and $\pi_{kiq} = \pi_k(R_{iq}, S_{iq})$.

We have the following inequalities, which will be proved later:

1.
$$E_{ik4} = \frac{\sigma^4}{l}, \tag{2.1}$$

2.
$$E_{ik2} \leq \frac{\sigma^2\|\omega\|_\infty}{l} \ \textit{and}\ E_{ik3} \leq \frac{\sigma^2\|\omega\|_\infty}{l}, \tag{2.2}$$

3.
$$E_{ik1} \leq \frac{\|h\|_\infty}{l} + C_2\|h\|_\infty \left(\int_{[0,1]^2} |\pi_k(u,v)|dudv\right)^2. \tag{2.3}$$

Because the processes $\delta_{i_1}(.), \delta_{i_2}(.)$ are independent for $i_1 \neq i_2$ and $E(\hat{c}_{k,i}) = c_k$ so:

$$E\left((\hat{c}_{k,i_1} - c_k)(\hat{c}_{k,i_2} - c_k)\right) \quad = \quad E\left((\hat{c}_{k,i_1} - c_k)\right) E\left((\hat{c}_{k,i_2} - c_k)\right) = 0,$$

which implies

$$\Rightarrow E\left((\hat{c}_k - c_k)^2\right) = E\left([\frac{1}{n}\sum_{i=1}^{n}(\hat{c}_{k,i} - c_k)]^2\right) = \frac{1}{n^2}\sum_{i=1}^{n} E\left((\hat{c}_{k,i} - c_k)^2\right). \tag{2.4}$$

From (2.1), (2.3), (2.2) and (2.4), using Parseval's equality we have:

$$\begin{aligned} E\|\hat{\omega}_M(\Pi) - P_M(\Pi)\omega\|_2^2 &= \sum_{k=1}^{M} E\left((\hat{c}_k - c_k)^2\right) = \sum_{k=1}^{M}\frac{1}{n^2}\sum_{i=1}^{n} E\left((\hat{c}_{k,i} - c_k)^2\right) \\ &\leq C\left\{\frac{M}{nl}\left[\sigma^4 + 2\sigma^2\|\omega\|_\infty + \|h\|_\infty\right] + \frac{\|h\|_\infty}{n}\sum_{k=1}^{M}\left(\int_{[0,1]^2}|\pi_k(u,v)|dudv\right)^2\right\}. \end{aligned}$$

Now we only need to prove(2.1), (2.3) and (2.2):

Because of the independence between the errors ϵ_{iq}, the design R_{iq}, S_{iq} and the processes $\delta_i()$, we have:

$$E\left((\frac{1}{l}\sum_{q=1}^{l}\epsilon_{iq1}\epsilon_{iq2}\pi_{kiq})^2\right) \quad = \quad \frac{1}{l^2}\sum_{q=1}^{l}\sigma^4\int \pi_k^2(u,v)dudv = \frac{\sigma^4}{l} \quad \Rightarrow (2.1)$$

and

$$\begin{aligned} E\left((\frac{1}{l}\sum_{q=1}^{l}\delta_i(R_{iq})\epsilon_{iq2}\pi_{kiq})^2\right) &= \frac{1}{l^2}\sum_{q=1}^{l}E(\epsilon_{iq2}^2)E\left(\delta_i^2(R_{iq})\pi_k^2(R_{iq},S_{iq})\right) \\ &= \frac{1}{l^2}\sum_{q=1}^{l}\sigma^2\int_{[0,1]^2}\omega(u,u)\pi_k^2(u,v)dudv \ \leq \ \frac{\sigma^2\|\omega\|_\infty}{l} \quad \Rightarrow (2.2). \end{aligned}$$

For (2.3):

$$E\left((\frac{1}{l}\sum_{q=1}^{l}\delta_i(R_{iq})\delta_i(S_{iq})\pi_{kiq} - c_k)^2\right) = E\left((\frac{1}{l}\sum_{q=1}^{l}\delta_i(R_{iq})\delta_i(S_{iq})\pi_{kiq})^2\right) - c_k^2$$

$$\begin{aligned}
&\leq E\left((\frac{1}{l}\sum_{q=1}^{l}\delta_i(R_{iq})\delta_i(S_{iq})\pi_{kiq})^2 \right) \\
&= \frac{1}{l^2}\sum_{q\neq p} E\left(\delta_i(R_{iq})\delta_i(S_{iq})\delta_i(R_{ip})\delta_i(S_{ip})\pi_{kiq}\pi_{kip}\right) + \frac{1}{l^2}\sum_{q=1}^{l} E\left(\delta^2(R_{iq})\delta^2(S_{iq})\pi_{kiq}^2\right).
\end{aligned}$$

First term,

$$\begin{aligned}
& E\left[\delta_i(R_{iq})\delta_i(S_{iq})\delta_i(R_{ip})\delta_i(S_{ip})\pi_{kiq}\pi_{kip}\right] \\
=\ & E\left\{E[\delta_i(R_{iq})\delta_i(S_{iq})\delta_i(R_{ip})\delta_i(S_{ip})|R_{iq},R_{ip},S_{iq},S_{ip}]\pi_{kiq}\pi_{kip}\right\} \\
=\ & E\left\{h(R_{iq},S_{iq},R_{ip},S_{ip})\pi_{kiq}\pi_{kip}\right\} \leq \|h\|_\infty E(|\pi_{kiq}\pi_{kip}|) \\
=\ & \|h\|_\infty \left(\int_{[0,1]^2}|\pi_k(u,v)|dudv\right)^2 \quad \forall p\neq q,
\end{aligned}$$

with $|h(s,t,u,v)| = |E(\delta(s)\delta(t)\delta(u)\delta(v))| \leq \| h\|_\infty < \infty, \forall(s,t,u,v)\in[0,1]^4$.

Second term,

$$\begin{aligned}
E\left(\delta^2(R_{iq})\delta^2(S_{iq})\pi_{kiq}^2\right) &= E\left(E[\delta^2(R_{iq})\delta^2(S_{iq})|R_{iq},S_{iq}]\pi_{kiq}^2\right) \\
&= E\left(h(R_{iq},S_{iq},R_{iq},S_{iq})\pi_{kiq}^2\right) \\
&= \int_{[0,1]^2} h(u,v,u,v)\pi_k^2(u,v)dudv \leq \|h\|_\infty\int_{[0,1]^2}\pi_k^2(u,v)dudv = \|h\|_\infty.
\end{aligned}$$

So

$$\begin{aligned}
E\left[\left(\frac{1}{l}\sum_{q=1}^{l}\delta_i(R_{iq})\delta_i(S_{iq})\pi_{kiq} - c_k\right)^2\right] &\leq \frac{\|h\|_\infty}{l} + C_2\|h\|_\infty\left(\int_{[0,1]^2}|\pi_k(u,v)|dudv\right)^2 \\
&\Rightarrow (2.3).
\end{aligned}$$

S2.2. Unknown-mean-function case *Proof of Theorem 4.4*:

We recall the auxiliary estimator $\hat{\omega}_{M_1}(\Pi)$, presented in section 4.3:

$$\hat{\omega}_{M_1}(\Pi) = \sum_{k=1}^{M_1}\hat{c}_k\pi_k(u,v),$$

with

$$\hat{c}_k = \frac{1}{n}\sum_{i=1}^{n}\frac{1}{l}\sum_{q=1}^{l}U_{iq}V_{iq}\pi_k(R_{iq},S_{iq}) = \frac{1}{nl}\sum_{q=1}^{l}U_{iq}V_{iq}\pi_{kiq}, \quad \hat{c}_{k,i} = \frac{1}{l}\sum_{q=1}^{l}U_{iq}V_{iq}\pi_{kiq}.$$

Using Parseval's equality, we have:

$$E\|\hat{\omega}^*_{M_1}(\Pi) - \hat{\omega}_{M_1}(\Pi)\|_2^2 = \sum_{k=1}^{M_1}E(\hat{c}^*_k - \hat{c}_k)^2 = \sum_{k=1}^{M_1}E\left(\frac{1}{n}\sum_{i=1}^{n}[\hat{c}^*_{k,i} - \hat{c}_{k,i}]\right)^2$$

So we need to find a upper bound of $(\hat{c}_k^* - \hat{c}_k)^2$:

$$\begin{aligned}
\hat{c}_k^* - \hat{c}_k &= \frac{1}{n}\sum_{i=1}^{n}(\hat{c}_{k,i}^* - \hat{c}_{k,i}) = \frac{1}{n}\sum_{i=1}^{n}\frac{1}{l}\sum_{q=1}^{l}\left(U_{iq}^* V_{iq}^* - U_{iq}V_{iq}\right)\pi_{kiq} \\
&= \frac{1}{n}\sum_{i=1}^{n}\frac{1}{l}\sum_{q=1}^{l}\left[\left(U_{iq} + \Delta_i(R_{iq})\right)\left(V_{iq} + \Delta_i(S_{iq})\right) - U_{iq}V_{iq}\right]\pi_{kiq} \\
&= \frac{1}{nl}\sum_{i=1}^{n}\sum_{q=1}^{l}\left(F_{1,kiq} + F_{2,kiq} + F_{3,kiq}\right).
\end{aligned}$$

with

$$F_{1,kiq} = \Delta_i(R_{iq})V_{iq}\pi_{kiq}, \quad F_{2,kiq} = U_{iq}\Delta_i(S_{iq})\pi_{kiq} \textit{ and } F_{3,kiq} = \Delta_i(R_{iq})\Delta_i(S_{iq})\pi_{kiq}.$$

We have the following inequalities which will be proved later:

- $$E\left[\left(\frac{1}{nl}\sum_{i,q}F_{1,kiq}\right)^2\right] \le CW(\pi_{k1}, \pi_{k2})v_{nm}^2(M) \tag{2.5}$$
- $$E\left[\left(\frac{1}{nl}\sum_{i,q}F_{2,kiq}\right)^2\right] \le CW(\pi_{k2}, \pi_{k1})v_{nm}^2(M) \tag{2.6}$$
- $$E\left[\left(\frac{1}{nl}\sum_{i,q}F_{3,kiq}\right)^2\right] \le \left[M^2 C\left(\frac{d_4(e_k)}{n^3} + \frac{1}{n^2}\right) + \|P_M(\mathcal{E})\mu - \mu\|_2^4\right]\left(\frac{2\|\pi_k^2\|_\infty}{nl} + 1\right) \tag{2.7}$$

From (2.5),(2.6) and (2.7), we have Theorem 4.4.

Now we prove the above inequalities:

For the first and the second terms

$$E\left[\left(\frac{1}{nl}\sum_{i=1}^{n}\sum_{q=1}^{l}F_{1,kiq}\right)^2\right] = \frac{1}{(nl)^2}\sum_{i,q}E\left(F_{1,kiq}^2\right) + \frac{1}{(nl)^2}\sum_{(i_1,q_1)\neq(i_2,q_2)}E\left(F_{1,ki_1q_1}F_{1,ki_2q_2}\right).$$

Because in the construction of the cross-estimator $\hat{\mu}_{(i_0),M}$, the data of the i_0^{th} subject $\{(Y_{i_0j}, T_{i_0j})\}_{j=1,\ldots,m}$ are not used, we have the independence between $\hat{\mu}_{(i_0),M}$ and $\{T_{i_0j}\}_{j=1,\ldots,m}$, so

$$\begin{aligned}
E\left[\left(\hat{\mu}_{(i_0),M}(T_{i_0j}) - \mu(T_{i_0j})\right)^2\right] &= E\left(\Delta_{i_0}^2(T_{i_0j})\right) = E\left(E\left[\left(\hat{\mu}_{(i_0),M}(T_{i_0j}) - \mu(T_{i_0j})\right)^2 | T_{i_0j}\right]\right) \\
&= \int_0^1 E\left[\left(\hat{\mu}_{(i_0),M}(u) - \mu(u)\right)^2\right]du = E\|\hat{\mu}_{(i_0),M} - \mu\|_2^2 \\
&= E\|\Delta_{i_0}\|_2^2 \le v_{nm}^2(M),
\end{aligned}$$

which implies:

$$E\left\{\Delta_i^2(u)\left(\delta_i(v)+\epsilon_{iq1}\right)^2\right\}=E(\Delta_i^2(u))\left[E(\delta_i^2(v))+\sigma^2\right]\leq(\|\omega\|_\infty+\sigma^2)E(\Delta_i^2(u)).$$

So, by denoting $\xi_i(u)=E(\Delta_i^2(u))$, we have:

$$\begin{aligned}
E\left(F_{1,kiq}^2\right) &= E\left[(\Delta_i(R_{iq})V_{iq}\pi_{kiq})^2\right]=E\left[E\left\{\Delta_i^2(R_{iq})\left(\delta_i(S_{iq})+\epsilon_{iq1}\right)^2|R_{iq},S_{iq}\right\}\pi_k^2(R_{iq},S_{iq})\right]\\
&\leq (\|\omega\|_\infty+\sigma^2)E\left[\xi_i(R_{iq})\pi_k^2(R_{iq},S_{iq})\right]=(\|\omega\|_\infty+\sigma^2)\int_0^1\left(\int_0^1\xi_i(u)\pi_k^2(u,v)du\right)dv\\
&= (\|\omega\|_\infty+\sigma^2)\int_0^1\pi_{k2}^2(v)dv\left(\int_0^1E(\Delta_i^2(u))\pi_{k1}^2(u)du\right)\leq(\|\omega\|_\infty+\sigma^2)\|\pi_{k1}^2\|_\infty E\|\Delta_i\|_2^2\\
&\leq (\|\omega\|_\infty+\sigma^2)\|\pi_{k1}^2\|_\infty v_{nm}^2(M).
\end{aligned}$$

Using the same arguments, by denoting $E_{R,S}=E\{.|R_{i_1q_1},S_{i_1q_1},R_{i_2q_2},S_{i_2q_2}\}$ we have:

$$\begin{aligned}
& E\left(F_{1,ki_1q_1}F_{1,ki_2q_2}\right)\\
= & E\left[E_{R,S}\left\{\Delta_{i_1}(R_{i_1q_1})\Delta_{i_2}(R_{i_2q_2})V_{i_1q_1}V_{i_2q_2}\right\}\pi_{ki_1q_1}\pi_{ki_2q_2}\right]\\
\leq & \frac{1}{2}E\left[E_{R,S}\left\{[\Delta_{i_1}(R_{i_1q_1})V_{i_1q_1}]^2+[\Delta_{i_2}(R_{i_2q_2})V_{i_2q_2}]^2\right\}|\pi_{ki_1q_1}\pi_{ki_2q_2}|\right]\\
\leq & \frac{(\|\omega\|_\infty+\sigma^2)}{2}\left\{E[\xi_i(R_{i_1q_1})|\pi_{ki_1q_1}|]E\left[|\pi_{ki_2q_2}|\right]+E\left[\xi_i(R_{i_2q_2})|\pi_{ki_2q_2}|\right]E\left[|\pi_{ki_1q_1}|\right]\right\}\\
\leq & \frac{(\|\omega\|_\infty+\sigma^2)}{2}(\int_{[0,1]^2}|\pi_k(u,v)|dudv)\left(\int_0^1\int_0^1\xi_i(u)|\pi_k(u,v)|dudv+\int_0^1\int_0^1\xi_i(u)|\pi_k(u,v)|dudv\right)\\
= & 2(\|\omega\|_\infty+\sigma^2)(\int_{[0,1]^2}|\pi_k(u,v)|dudv)\int_0^1|\pi_{k2}(v)|dv\int_0^1E(\Delta_i^2(u))|\pi_{k1}(u)|du\\
\leq & 2(\|\omega\|_\infty+\sigma^2)(\int_{[0,1]^2}|\pi_k(u,v)|dudv)\|\pi_{k1}\|_\infty\int_0^1|\pi_{k2}(v)|dv\int_0^1E(\Delta_i^2(u))du\\
\leq & 2(\|\omega\|_\infty+\sigma^2)(\int_{[0,1]^2}|\pi_k(u,v)|dudv)\|\pi_{k1}\|_\infty(\int_0^1|\pi_{k2}(v)|dv)v_{nm}^2(M).
\end{aligned}$$

So,

$$\begin{aligned}
E\left[\left(\frac{1}{nl}\sum_{i=1}^n\sum_{q=1}^lF_{1,kiq}\right)^2\right] &\leq C\left(\frac{\|\pi_{k1}^2\|_\infty}{nl}+(\int_{[0,1]^2}|\pi_k(u,v)|dudv)\|\pi_{k1}\|_\infty(\int_0^1|\pi_{k2}(v)|dv)\right)v_{nm}^2(M)\\
&= CW(\pi_{k1},\pi_{k2})v_{nm}^2(M) \qquad \Rightarrow (2.5).
\end{aligned}$$

The second term could be bounded by the same way:

$$\begin{aligned}
E\left[\left(\frac{1}{nl}\sum_{i=1}^n\sum_{q=1}^lF_{2,kiq}\right)^2\right] &\leq \left(\frac{\|\pi_{k2}^2\|_\infty}{nl}+(\int_{[0,1]^2}|\pi_k(u,v)|dudv)\|\pi_{k2}\|_\infty(\int_0^1|\pi_{k1}(v)|dv)\right)v_{nm}^2(M)\\
&= CW(\pi_{k2},\pi_{k1})v_{nm}^2(M) \qquad \Rightarrow (2.6).
\end{aligned}$$

For the third term

Lemma S2.1. For all random process δ such : $|E(\delta^{2p}(u))| \leq C < \infty, \forall u \in [0,1]$, $p = 1, 2 \ldots,$, we have:

$$E\left[(\hat{\alpha}_{t,i} - \alpha_t)^{2p}\right] \leq C_p \left(\int_0^1 e_t^{2p}(u)du + (\int_0^1 |e_t(u)|du)^{2p} \right) \leq C d_{2p}(e_t), \tag{2.8}$$

with $d_{2p}(e_t) = \int_0^1 e_t^{2p}(u)du; \hat{\alpha}_{t,i} = \frac{1}{m}\sum_{j=1}^m Y_{ij} e_t(T_{ij}), t = 1, \ldots, M.$

Using Rosenthal's inequality, under the condition $E(\delta^{2p}(u)) \leq C < \infty$ we have

$$E\left((\hat{\alpha}_t - \alpha)^{2p}\right) \leq C\left(\frac{d_{2p}(e_t)}{n^{2p-1}} + \frac{1}{n^p}\right).$$

Using the above lemma, we have:

$$\begin{aligned}
E\left((\int_0^1 \Delta_i^2(u)du)^2\right) &= E\left((\int_0^1 [\hat{\mu}_{(i),M}(\mathcal{E})(u) - \mu(u)]^2 du)^2\right) \\
&= E\left((\sum_{t=1}^M (\hat{\alpha}_t - \alpha_t)^2 + \|P_M(\mathcal{E})\mu - \mu\|_2^2)^2\right) \\
&\leq 2\left[E\left(\sum_{t=1}^M (\hat{\alpha}_t - \alpha_t)^2\right)^2 + \|P_M(\mathcal{E})\mu - \mu\|_2^4\right] \\
&\leq 2\left[M\sum_{t=1}^M E\left((\hat{\alpha}_t - \alpha_t)^4\right) + \|P_M(\mathcal{E})\mu - \mu\|_2^4\right] \\
&\leq 2\left[MC\sum_{t=1}^M \left(\frac{d_4(e_t)}{n^3} + \frac{1}{n^2}\right) + \|P_M(\mathcal{E})\mu - \mu\|_2^4\right]
\end{aligned}$$

So

- $$\begin{aligned}
E\left(F_{3,kiq}^2\right) &= E\left[(\Delta_i(R_{iq})\Delta_i(S_{iq})\pi_k(R_{iq}, S_{iq}))^2\right] \\
&= E\left(\int_0^1 \Delta_i^2(u)\pi_{k1}^2(u)du \int_0^1 \Delta_i^2(v)\pi_{k2}^2(v)dv\right) \leq \|\pi_{k1}^2\|_\infty \|\pi_{k2}^2\|_\infty E\left((\int_0^1 \Delta_i^2(u)du)^2\right) \\
&= \|\pi_k^2\|_\infty E\left((\int_0^1 \Delta_i^2(u)du)^2\right) \\
&\leq 2\|\pi_k^2\|_\infty \left[MC\sum_{t=1}^M \left(\frac{d_4(e_t)}{n^3} + \frac{1}{n^2}\right) + \|P_M(\mathcal{E})\mu - \mu\|_2^4\right]
\end{aligned}$$

- $$\begin{aligned}
E\left(F_{3,ki_1q_1}F_{3,ki_2q_2}\right) &= E\left[\Delta_i(R_{i_1q_1})\Delta_i(S_{i_1q_1})\pi_k(R_{i_1q_1}, S_{i_1q_1})\Delta_i(R_{i_2q_2})\Delta_i(S_{i_2q_2})\pi_k(R_{i_2q_2}, S_{i_2q_2})\right] \\
&= E\left[\left(\int_0^1 \Delta_i(u)\pi_{k1}(u)du \int_0^1 \Delta_i(v)\pi_{k2}(v)dv\right)^2\right] \\
&\leq E\left([\int_0^1 \Delta_i^2(u)du]^2 \int_0^1 \pi_{k1}^2(u)du \int_0^1 \pi_{k2}^2(v)dv\right) = E\left([\int_0^1 \Delta_i^2(u)du]^2\right) \\
&\leq \left[MC\sum_{t=1}^M \left(\frac{d_4(e_t)}{n^3} + \frac{1}{n^2}\right) + \|P_M(\mathcal{E})\mu - \mu\|_2^4\right]
\end{aligned}$$

The third term (2.7):

$$E\left[\left(\frac{1}{nl}\sum_{i,q}F_{3,kiq}\right)^2\right] = \frac{1}{(nl)^2}\left(\sum_{i,q}E\left(F_{3,kiq}^2\right) + \sum_{(i_1,q_1)\neq(i_2,q_2)}E(F_{3,ki_1q_1}F_{3,ki_2q_2})\right)$$

$$\leq \left[MC\sum_{t=1}^{M}\left(\frac{d_4e_t}{n^3}+\frac{1}{n^2}\right) + \|P_M(\mathcal{E})\mu-\mu\|_2^4\right]\left(\frac{2\|\pi_k^2\|_\infty}{nl}+1\right) \quad \Rightarrow (2.7)$$

Proof of Lemma S2.1

We have the following inequalities :

- $$E\left((\frac{1}{m}\sum_{j=1}^{m}\delta_i(T_{ij})e_t(T_{ij}))^{2p}\right) \leq \frac{1}{m}\sum_{j=1}^{m}E\{E(\delta_i^{2p}(T_{ij})|T_{ij})e_t^{2p}(T_{ij})\} \leq C\int_0^1 e_t^{2p}(u)du,$$

- $$E\left((\frac{1}{m}\sum_{j=1}^{m}\epsilon_{ij}e_t(T_{ij}))^{2p}\right) \leq \frac{1}{m}\sum_{j=1}^{m}E(\epsilon_{ij}^{2p})E(e_t^{2p}(T_{ij})) = E(\epsilon^{2p})\int_0^1 e_t^{2p}(u)du,$$

- $$\begin{aligned} E\left((\frac{1}{m}\sum_{j=1}^{m}\mu(T_{ij})e_t(T_{ij})-\alpha_t)^{2p}\right) &\leq \frac{1}{m}\sum_{j=1}^{m}E\left[(\mu(T_{ij})e_t(T_{ij})-\alpha_t)^{2p}\right] \\ &\leq 2^{2p-1}\left(\int_0^1\mu^{2p}(u)e_t^{2p}(T_{ij})du + (\int_0^1\mu(u)e_t(u)du)^{2p}\right) \\ &\leq 2^{2p-1}\left(\|\mu\|_\infty^{2p}\left[\int_0^1 e_t^{2p}(u)du + (\int_0^1|e_t(u)|du)^{2p}\right]\right) \\ &\leq 2^{2p}\|\mu\|_\infty^{2p}\int_0^1 e_t^{2p}(u)du. \end{aligned}$$

So, we have:

$$\begin{aligned} & E\left((\hat{\alpha}_{t,i}-\alpha_t)^{2p}\right) = E\left((\frac{1}{m}\sum_{j=1}^{m}Y_{ij}e_t(T_{ij})-\alpha_t)^{2p}\right) \\ \leq\ & 3^{2p-1}\left[E\left((\frac{1}{m}\sum_{j=1}^{m}\mu(T_{ij})e_t(T_{ij})-\alpha_t)^{2p} + (\frac{1}{m}\sum_{j=1}^{m}\delta_i(T_{ij})e_t(T_{ij}))^{2p} + (\frac{1}{m}\sum_{j=1}^{m}\epsilon_{ij}e_t(T_{ij}))^{2p}\right)\right] \\ \leq\ & C_p\int_0^1 e_t^{2p}(u)du \quad \Rightarrow (2.8). \end{aligned}$$

The proof of Theorem 4.4 is complete.