SUPPLEMENTARY TO "TWO SAMPLE TESTING FOR HIGH-DIMENSIONAL FUNCTIONAL DATA: A MULTI-RESOLUTION PROJECTION METHOD"

Shouxia Wang, Jiguo Cao, Hua Liu, Jicai Liu and Jinhong You

Shanghai University of Finance and Economics, Simon Fraser University

Xi'an Jiaotong University, Shanghai Lixin University of Accounting and Finance

and Shanghai University of Finance and Economics

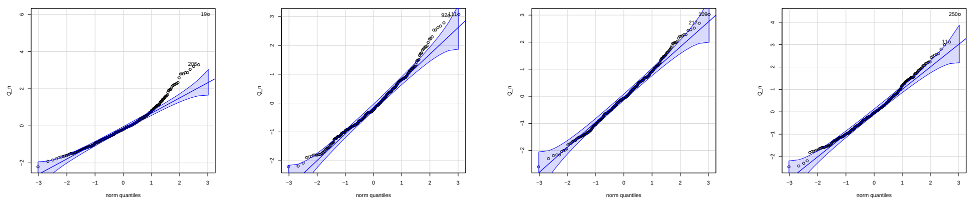
Supplementary Material

This supplementary material contains three sections. Section S1 presents additional simulation results. Additional real data analysis about the EEG data, which concerns the relationship between genetic predisposition and the tendency for alcoholism, is given in Section S2. Section S3 provides the technical details of the main manuscript.

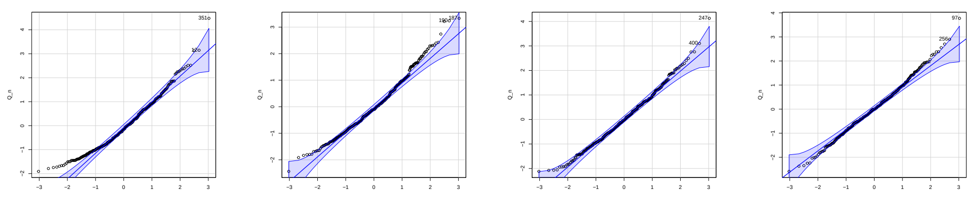
S1 Additional simulation results

Figure S.1 shows the QQ plot of $\widehat{Q}_n(\widehat{\mathbf{X}}, \widehat{\mathbf{Y}})$ with an Ornstein-Uhlenbeck process under $n = m = 25, 40$ and $p = 20, 50, 100, 200$ in Case I for model

(3.14). Similar to Figure 2, Figure S.1 reveals that the asymptotic distribution of $\widehat{Q}_n(\widehat{\mathbf{X}}, \widehat{\mathbf{Y}})$ is well approximated by $N(0,1)$ when the dimension and the sample sizes increase.



(a) $n = 25, p = 20$ (b) $n = 25, p = 50$ (c) $n = 25, p = 100$ (d) $n = 25, p = 200$



(e) $n = 40, p = 20$ (f) $n = 40, p = 50$ (g) $n = 40, p = 100$ (h) $n = 40, p = 200$

Figure S.1: Q-Q plots of $\widehat{Q}_n(\widehat{\mathbf{X}}, \widehat{\mathbf{Y}})$ under $n = m = 25, 40$ and $p = 20, 50, 100, 200$ for Case I for model (3.14) in Simulation II.

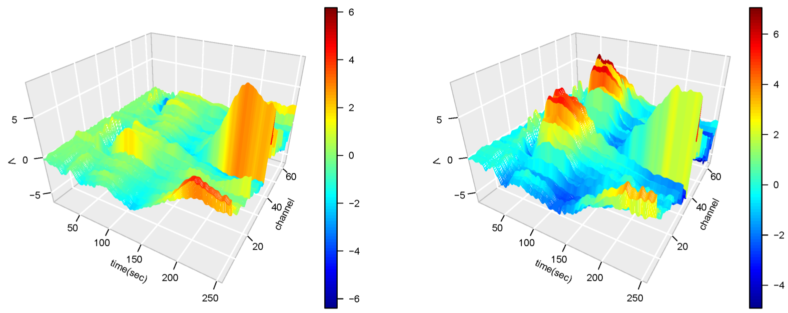
S2 Additional real data analysis: EEG Data

This dataset is from an electroencephalogram (EEG) study, which concerns the relationship between genetic predisposition and the tendency for alcoholism (`http://kdd.ics.uci.edu/databases/eeg/eeg.data.html`).

The EEG dataset includes two groups of subjects: an alcoholic group of 77 subjects and a control (non-alcoholic) group of 45 subjects. Each subject was exposed to one stimulus or two stimuli while voltage values were measured from 64 channels of electrodes placed on the subject's scalp for 256 time points. Li et al. (2010) applied the dataset to study the pattern of voltage over times and channels by treating voltage values for each subject as a 256×64 matrix.

It is noteworthy that, although the method proposed by Li et al. (2010) can preserve well the original structure of the data by the way of the matrix-valued objects, it ignores the smoothness of voltage values at the time points. In fact, it seems to be more natural that voltage values for each channel should be treated as a continuous function of time. Thus, it may be more reasonable to view the voltage values as a random functional curve. From this perspective, we represent the data as random functional data, denoted by $\{\mathbf{X}_i(t) = (X_{i1}(t), \dots, X_{i64}(t))^T, t \in [0,1]\}_{i=1}^{77}$ for the alcoholic group, and $\{\mathbf{Y}_j(t) = (Y_{j1}(t), \dots, Y_{j64}(t))^T, t \in [0,1]\}_{j=1}^{45}$ for the control group, where each $\mathbf{X}_{ik}(t)$ and $\mathbf{Y}_{jk}(t)$ has $N = 256$ observations.

We here are interested in testing whether the EEG patterns between the control and alcoholic groups are different or not. Using the proposed MRP test, we obtain that the test statistic is 4.680 with the upper-tail p-value approximately equal to 1.43×10^{-6}. Thus, there is a significant



(a) Non-alcoholic group

(b) Alcoholic group

Figure S.2: Sample means of the voltages measured from 64 channels of electrodes placed on the subject's scalp for 256 time points among non-alcoholic and alcoholic groups.

difference in the EEG pattern between the two groups. The reasonability of the result is supported by their sample mean functions shown in Figure S.2 (a) and (b), with two horizontal axes representing time and channel and the vertical axis denoting voltage.

S3 Proofs of the theoretical results

Proof of Lemma 1. Suppose that $\mathrm{E}\{\mathbf{X}\} = \mathrm{E}\{\mathbf{Y}\}$. For any $\boldsymbol{\alpha} \in \mathbb{R}^p$ and $\gamma \in L_2([0,1])$, $\langle \boldsymbol{\alpha}^{\mathrm{T}}\mathbf{X}, \gamma\rangle$ and $\langle \boldsymbol{\alpha}^{\mathrm{T}}\mathbf{Y}, \gamma\rangle$ are measurable functions of $\mathbf{X}$ and $\mathbf{Y}$ respectively. Then, we have that

$$\mathrm{E}\langle \boldsymbol{\alpha}^{\mathrm{T}}\mathbf{X}, \gamma\rangle = \boldsymbol{\alpha}^{\mathrm{T}} \int_0^1 \mathrm{E}\{\mathbf{X}(t)\}\gamma(t)dt = \boldsymbol{\alpha}^{\mathrm{T}} \int_0^1 \mathrm{E}\{\mathbf{Y}(t)\}\gamma(t)dt = \mathrm{E}\langle \boldsymbol{\alpha}^{\mathrm{T}}\mathbf{Y}, \gamma\rangle.$$

If $\mathrm{E}\langle \boldsymbol{\alpha}^{\mathrm{T}}\mathbf{X}, \gamma\rangle = \mathrm{E}\langle \boldsymbol{\alpha}^{\mathrm{T}}\mathbf{Y}, \gamma\rangle$ holds for any $\boldsymbol{\alpha} \in \mathbb{R}^p$ and $\gamma \in L_2([0,1])$, we have that

$$\boldsymbol{\alpha}^{\mathrm{T}} \int_0^1 \mathrm{E}\{[\mathbf{X}(t) - \mathbf{Y}(t)]\}\gamma(t)dt = \mathrm{E}\langle \boldsymbol{\alpha}^{\mathrm{T}}[\mathbf{X} - \mathbf{Y}], \gamma\rangle = 0.$$

By the arbitrariness of $\boldsymbol{\alpha} \in \mathbb{R}^p$ and $\gamma \in L_2([0,1])$, the above result implies that $\mathrm{E}\{\mathbf{X}\} = \mathrm{E}\{\mathbf{Y}\}$ in $L_2^p([0,1])$, and hence componentwise for almost every t.

Proof of Theorem 1. Suppose $\boldsymbol{\alpha} = (\alpha_1, \ldots, \alpha_p)^{\mathrm{T}} \sim G(\boldsymbol{\alpha})$, where $G(\cdot)$ is the CDF of $\boldsymbol{\alpha}$, $G(\boldsymbol{\alpha}) = \Pi_{k=1}^p G_k(\alpha_k)$, $G_k(\alpha_k)$ is a CDF of α_k with mean 0 and variance 1 for $k = 1, \ldots, p$, $\nu(\cdot)$ is the CDF of a centered process γ in $L_2([0,1])$ with a covariance function $v(s,t)$ satisfying the conditions in the definition of MRP, and $\boldsymbol{\alpha}$ and γ are independent of $\mathbf{X}$ and $\mathbf{Y}$. Then, for any fixed functions $\mathbf{Z}_1, \mathbf{Z}_2 \in L_2^p[0,1]$, we have

$$\begin{aligned}
&\int_{L_2}\int_{\mathbb{R}^p} \langle \boldsymbol{\alpha}^{\mathrm{T}}\mathbf{Z}_1, \gamma\rangle \langle \boldsymbol{\alpha}^{\mathrm{T}}\mathbf{Z}_2, \gamma\rangle dG(\boldsymbol{\alpha})\nu(\,\mathrm{d}\gamma) \\
&= \int_{L_2}\int_{\mathbb{R}^p} \Big[\sum_{k=1}^p \int_0^1 \alpha_k Z_{1k}(t)\gamma(t)\,\mathrm{d}t\Big]\Big[\sum_{k=1}^p \int_0^1 \alpha_k Z_{2k}(t)\gamma(t)\,\mathrm{d}t\Big] dG(\alpha_1, \ldots, \alpha_p)\nu(\,\mathrm{d}\gamma) \\
&= \int_{L_2}\int_{\mathbb{R}^p} \Big[\sum_{k=1}^p \int_0^1\int_0^1 \alpha_k^2 Z_{1k}(t) Z_{2k}(s)\gamma(s)\gamma(t)\,\mathrm{d}s\,\mathrm{d}t\Big] dG(\alpha_1, \ldots, \alpha_p)\nu(\,\mathrm{d}\gamma) \\
&= \int_{L_2}\int_0^1\int_0^1 \mathbf{Z}_1(t)^{\mathrm{T}}\mathbf{Z}_2(s)\gamma(s)\gamma(t)\,\mathrm{d}s\,\mathrm{d}t\nu(\,\mathrm{d}\gamma) \\
&= \int_0^1\int_0^1 \mathbf{Z}_1(t)^{\mathrm{T}}\mathbf{Z}_2(s)v(s,t)\,\mathrm{d}s\,\mathrm{d}t.
\end{aligned} \tag{S3.1}$$

Denote $\mathbf{Z}_1 = \mathbf{X}_1 - \mathbf{Y}_1$ and $\mathbf{Z}_2 = \mathbf{X}_2 - \mathbf{Y}_2$. Then, by the definition of

$\texttt{MRP}(\mathbf{X},\mathbf{Y})$ and (S3.1), we have

$$\begin{aligned}
\texttt{MRP}(\mathbf{X},\mathbf{Y}) &= \int_{L_2}\int_{\mathbb{R}^p}[\mathrm{E}\{\langle\boldsymbol{\alpha}^{\mathrm{T}}(\mathbf{X}-\mathbf{Y}),\gamma\rangle\}]^2\,\mathrm{d}G(\boldsymbol{\alpha})\nu(\,\mathrm{d}\gamma)\\
&= \int_{L_2}\int_{\mathbb{R}^p}\mathrm{E}\{\langle\boldsymbol{\alpha}^{\mathrm{T}}(\mathbf{X}_1-\mathbf{Y}_1),\gamma\rangle\}\mathrm{E}\{\langle\boldsymbol{\alpha}^{\mathrm{T}}(\mathbf{X}_2-\mathbf{Y}_2),\gamma\rangle\}\,\mathrm{d}G(\boldsymbol{\alpha})\nu(\,\mathrm{d}\gamma)\\
&= \mathrm{E}\Big\{\int_{L_2}\int_{\mathbb{R}^p}\langle\boldsymbol{\alpha}^{\mathrm{T}}\mathbf{Z}_1,\gamma\rangle\langle\boldsymbol{\alpha}^{\mathrm{T}}\mathbf{Z}_2,\gamma\rangle\,\mathrm{d}G(\boldsymbol{\alpha})\nu(\,\mathrm{d}\gamma)\Big\}\\
&= \mathrm{E}\Big\{\int_0^1\int_0^1\mathbf{Z}_1(t)^{\mathrm{T}}\mathbf{Z}_2(s)v(s,t)\,\mathrm{d}s\,\mathrm{d}t\Big\}\\
&= \mathrm{E}\Big\{\int_0^1\int_0^1(\mathbf{X}_1(t)-\mathbf{Y}_1(t))^{\mathrm{T}}(\mathbf{X}_2(s)-\mathbf{Y}_2(s))\,v(s,t)\,\mathrm{d}s\,\mathrm{d}t\Big\}\\
&= \int_0^1\int_0^1[\boldsymbol{\mu}_1(t)-\boldsymbol{\mu}_2(t)]^{\mathrm{T}}[\boldsymbol{\mu}_1(s)-\boldsymbol{\mu}_2(s)]v(s,t)\,\mathrm{d}s\,\mathrm{d}t.
\end{aligned}$$

This proves part (i). For part (ii), let $\delta_k(t)=\mu_{1k}(t)-\mu_{2k}(t)$. Then $\texttt{MRP}(\mathbf{X},\mathbf{Y})=\sum_{k=1}^p\langle\delta_k,\mathcal{V}\delta_k\rangle_{L_2}\ge 0$. By the injectivity (strict positive definiteness) of $\mathcal{V}$, this quantity is zero if and only if every δ_k is the zero element of $L_2([0,1])$, that is, if and only if $\boldsymbol{\mu}_1(t)=\boldsymbol{\mu}_2(t)$ for almost every t.

Proof of Theorem 2. We first calculate $\mathrm{E}\{\widehat{\texttt{MRP}}(\mathbf{X},\mathbf{Y})\}$ and $\mathrm{Var}\{\widehat{\texttt{MRP}}(\mathbf{X},\mathbf{Y})\}$. By the definition of $\widehat{\texttt{MRP}}(\mathbf{X},\mathbf{Y})$, it is easy to show that

$$\begin{aligned}
&\mathrm{E}\{\widehat{\texttt{MRP}}(\mathbf{X},\mathbf{Y})\}\\
=&\int_0^1\int_0^1[\mathrm{E}\{\mathbf{X}_1(t)^{\mathrm{T}}\mathbf{X}_2(s)\}+\mathrm{E}\{\mathbf{Y}_1(t)^{\mathrm{T}}\mathbf{Y}_2(s)\}-2\mathrm{E}\{\mathbf{X}_1(t)\}^{\mathrm{T}}\mathrm{E}\{\mathbf{Y}_1(s)\}]\\
&\quad\times v(s,t)\,\mathrm{d}s\,\mathrm{d}t\\
=&\int_0^1\int_0^1(\boldsymbol{\mu}_1(t)-\boldsymbol{\mu}_2(t))^{\mathrm{T}}(\boldsymbol{\mu}_1(s)-\boldsymbol{\mu}_2(s))\,v(s,t)\,\mathrm{d}s\,\mathrm{d}t=\texttt{MRP}(\mathbf{X},\mathbf{Y}).
\end{aligned}$$

For simplicity, let

$$
\begin{aligned}
D_1 &= \frac{1}{n(n-1)} \sum_{\substack{i,j=1\\ i\neq j}}^{n} \int_0^1 \int_0^1 \mathbf{X}_i(t)^{\mathrm{T}} \mathbf{X}_j(s) v(s,t)\, \mathrm{d}s\, \mathrm{d}t, \\
D_2 &= \frac{1}{m(m-1)} \sum_{\substack{i,j=1\\ i\neq j}}^{m} \int_0^1 \int_0^1 \mathbf{Y}_i(t)^{\mathrm{T}} \mathbf{Y}_j(s) v(s,t)\, \mathrm{d}s\, \mathrm{d}t, \\
D_3 &= -\frac{2}{nm} \sum_{i=1}^{n} \sum_{j=1}^{m} \int_0^1 \int_0^1 \mathbf{X}_i(t)^{\mathrm{T}} \mathbf{Y}_j(s) v(s,t)\, \mathrm{d}s\, \mathrm{d}t.
\end{aligned}
$$

Note that $\mathrm{Var}\{\widehat{\mathtt{MRP}}(\mathbf{X},\mathbf{Y})\}$ can be decomposed into

$$
\begin{aligned}
\mathrm{Var}\{\widehat{\mathtt{MRP}}(\mathbf{X},\mathbf{Y})\} &= \mathrm{Var}\{D_1\} + \mathrm{Var}\{D_2\} + \mathrm{Var}\{D_3\} \\
&+ 2\,\mathrm{Cov}\{D_1, D_2\} + 2\,\mathrm{Cov}\{D_1, D_3\} + 2\,\mathrm{Cov}\{D_2, D_3\}. \qquad \text{(S3.2)}
\end{aligned}
$$

To obtain $\mathrm{Var}\{\widehat{\mathtt{MRP}}(\mathbf{X},\mathbf{Y})\}$, we next calculate each term in the above decomposition.

By Lemma S.1 in Appendix B, we have that

$$
\begin{aligned}
\mathrm{Var}\{D_1\} &= \frac{1}{n^2(n-1)^2} \mathrm{Var}\Big\{ \sum_{\substack{i,j=1\\ i\neq j}}^{n} \iint \mathbf{X}_i(t)^{\mathrm{T}} \mathbf{X}_j(s) v(s,t)\, \mathrm{d}s\, \mathrm{d}t \Big\} \\
&= \frac{2}{n(n-1)} \iiiint \mathrm{tr}\{\mathbf{G}_1(t_1,t)\mathbf{G}_1(s_1,s)\} v(s,t) v(s_1,t_1)\, \mathrm{d}s\, \mathrm{d}t\, \mathrm{d}s_1\, \mathrm{d}t_1 \\
&+ \frac{4}{n} \iiiint \boldsymbol{\mu}_1(t)^{\mathrm{T}} \mathbf{G}_1(s,s_1) \boldsymbol{\mu}_1(t_1) v(s,t) v(s_1,t_1)\, \mathrm{d}s\, \mathrm{d}t\, \mathrm{d}s_1\, \mathrm{d}t_1. \qquad \text{(S3.3)}
\end{aligned}
$$

Similarly, it can be shown that

$$
\begin{aligned}
\mathrm{Var}\{D_2\} &= \frac{2}{m(m-1)} \iiiint \mathrm{tr}\{\mathbf{G}_2(t_1,t)\mathbf{G}_2(s_1,s)\} v(s,t) v(s_1,t_1)\, \mathrm{d}s\, \mathrm{d}t\, \mathrm{d}s_1\, \mathrm{d}t_1 \\
&+ \frac{4}{m} \iiiint \boldsymbol{\mu}_2(t)^{\mathrm{T}} \mathbf{G}_2(s,s_1) \boldsymbol{\mu}_2(t_1) v(s,t) v(s_1,t_1)\, \mathrm{d}s\, \mathrm{d}t\, \mathrm{d}s_1\, \mathrm{d}t_1, \qquad \text{(S3.4)}
\end{aligned}
$$

$$\begin{aligned}\mathrm{Var}\{D_3\} &= \frac{4}{nm}\iiiint \mathrm{tr}\{\mathbf{G}_1(s,s_1)\mathbf{G}_2(t,t_1)\}v(s,t)v(s_1,t_1)\,\mathrm{d}s\,\mathrm{d}t\,\mathrm{d}s_1\,\mathrm{d}t_1\\ &+\frac{4}{n}\iiiint \boldsymbol{\mu}_2(t)^{\mathrm{T}}\mathbf{G}_1(s,s_1)\boldsymbol{\mu}_2(t_1)v(s,t)v(s_1,t_1)\,\mathrm{d}s\,\mathrm{d}t\,\mathrm{d}s_1\,\mathrm{d}t_1\\ &+\frac{4}{m}\iiiint \boldsymbol{\mu}_1(t)^{\mathrm{T}}\mathbf{G}_2(s,s_1)\boldsymbol{\mu}_1(t_1)v(s,t)v(s_1,t_1)\,\mathrm{d}s\,\mathrm{d}t\,\mathrm{d}s_1\,\mathrm{d}t_1, \end{aligned}\tag{S3.5}$$

$$\mathrm{Cov}\{D_1,D_3\} = -\frac{4}{n}\iiiint \boldsymbol{\mu}_1(t)^{\mathrm{T}}\mathbf{G}_1(s,s_1)\boldsymbol{\mu}_2(t_1)v(s,t)v(s_1,t_1)\,\mathrm{d}s\,\mathrm{d}t\,\mathrm{d}s_1\,\mathrm{d}t_1,\tag{S3.6}$$

and

$$\mathrm{Cov}\{D_2,D_3\} = -\frac{4}{m}\iiiint \boldsymbol{\mu}_1(t)^{\mathrm{T}}\mathbf{G}_2(s,s_1)\boldsymbol{\mu}_2(t_1)v(s,t)v(s_1,t_1)\,\mathrm{d}s\,\mathrm{d}t\,\mathrm{d}s_1\,\mathrm{d}t_1.\tag{S3.7}$$

Additionally, we have that $\mathrm{Cov}\{D_1,D_2\}=0$ due to independence between $\{\mathbf{X}_i\}_{i=1}^n$ and $\{\mathbf{Y}_i\}_{i=1}^m$. This, together with (S3.2)-(S3.6), yields that

$$\begin{aligned}&\mathrm{Var}\{\widehat{\mathtt{MRP}}(\mathbf{X},\mathbf{Y})\}\\ &=\frac{2}{n(n-1)}\iiiint \mathrm{tr}\{\mathbf{G}_1(s,s_1)\mathbf{G}_1(t,t_1)\}v(s,t)v(s_1,t_1)\,\mathrm{d}s\,\mathrm{d}t\,\mathrm{d}s_1\,\mathrm{d}t_1\\ &+\frac{2}{m(m-1)}\iiiint \mathrm{tr}\{\mathbf{G}_2(s,s_1)\mathbf{G}_2(t,t_1)\}v(s,t)v(s_1,t_1)\,\mathrm{d}s\,\mathrm{d}t\,\mathrm{d}s_1\,\mathrm{d}t_1\\ &+\frac{4}{nm}\iiiint \mathrm{tr}\{\mathbf{G}_1(s,s_1)\mathbf{G}_2(t,t_1)\}v(s,t)v(s_1,t_1)\,\mathrm{d}s\,\mathrm{d}t\,\mathrm{d}s_1\,\mathrm{d}t_1\\ &+\frac{4}{n}\iiiint [\boldsymbol{\mu}_1(t)-\boldsymbol{\mu}_2(t)]^{\mathrm{T}}\mathbf{G}_1(s,s_1)[\boldsymbol{\mu}_1(t_1)-\boldsymbol{\mu}_2(t_1)]v(s,t)v(s_1,t_1)\,\mathrm{d}s\,\mathrm{d}t\,\mathrm{d}s_1\,\mathrm{d}t_1\\ &+\frac{4}{m}\iiiint [\boldsymbol{\mu}_1(t)-\boldsymbol{\mu}_2(t)]^{\mathrm{T}}\mathbf{G}_2(s,s_1)[\boldsymbol{\mu}_1(t_1)-\boldsymbol{\mu}_2(t_1)]v(s,t)v(s_1,t_1)\,\mathrm{d}s\,\mathrm{d}t\,\mathrm{d}s_1\,\mathrm{d}t_1.\end{aligned}$$

Let

$$
\begin{aligned}
&\sigma_{nm}^2(\mathbf{X},\mathbf{Y}) \\
&=: \frac{2}{n\,(n-1)} \iiiint \operatorname{tr}\left\{\mathbf{G}_1(s,s_1)\mathbf{G}_1(t,t_1)\right\} v(s,t)v(s_1,t_1)\,\mathrm{d}s\,\mathrm{d}t\,\mathrm{d}s_1\,\mathrm{d}t_1 \\
&+ \frac{2}{m\,(m-1)} \iiiint \operatorname{tr}\left\{\mathbf{G}_2(s,s_1)\mathbf{G}_2(t,t_1)\right\} v(s,t)v(s_1,t_1)\,\mathrm{d}s\,\mathrm{d}t\,\mathrm{d}s_1\,\mathrm{d}t_1 \\
&+ \frac{4}{nm} \iiiint \operatorname{tr}\left\{\mathbf{G}_1(s,s_1)\mathbf{G}_2(t,t_1)\right\} v(s,t)v(s_1,t_1)\,\mathrm{d}s\,\mathrm{d}t\,\mathrm{d}s_1\,\mathrm{d}t_1.
\end{aligned}
$$

Thus, under $H_0 : \boldsymbol{\mu}_1 = \boldsymbol{\mu}_2$, we can show that $\operatorname{Var}\{\widehat{\mathtt{MRP}}(\mathbf{X},\mathbf{Y})\} = \sigma_{nm}^2(\mathbf{X},\mathbf{Y})$. Under $H_1 : \boldsymbol{\mu}_1 \neq \boldsymbol{\mu}_2$, we have that $\operatorname{Var}\{\widehat{\mathtt{MRP}}(\mathbf{X},\mathbf{Y})\} = \sigma_{nm}^2(\mathbf{X},\mathbf{Y})\{1+o(1)\}$ by Condition (C3). Let

$$
\begin{aligned}
\sigma_{nm_2}^2(\mathbf{X},\mathbf{Y}) = \frac{4}{n} &\iiiint [\boldsymbol{\mu}_1(t) - \boldsymbol{\mu}_2(t)]^{\mathrm{T}} \mathbf{G}_1(s,s_1) [\boldsymbol{\mu}_1(t_1) - \boldsymbol{\mu}_2(t_1)] \\
&\qquad \times v(s,t)v(s_1,t_1)\,\mathrm{d}s\,\mathrm{d}t\,\mathrm{d}s_1\,\mathrm{d}t_1 \\
+ \frac{4}{m} &\iiiint [\boldsymbol{\mu}_1(t) - \boldsymbol{\mu}_2(t)]^{\mathrm{T}} \mathbf{G}_2(s,s_1) [\boldsymbol{\mu}_1(t_1) - \boldsymbol{\mu}_2(t_1)] \\
&\qquad \times v(s,t)v(s_1,t_1)\,\mathrm{d}s\,\mathrm{d}t\,\mathrm{d}s_1\,\mathrm{d}t_1.
\end{aligned}
$$

Then under Condition (C3'), $\operatorname{Var}\{\widehat{\mathtt{MRP}}(\mathbf{X},\mathbf{Y})\} = \sigma_{nm_2}^2(\mathbf{X},\mathbf{Y})\{1+o(1)\}$.

We next establish the asymptotic normality of $\widehat{\mathtt{MRP}}(\mathbf{X},\mathbf{Y})$. Note that

$$
\widehat{\mathtt{MRP}}(\mathbf{X},\mathbf{Y}) = \widehat{\mathtt{MRP}}_1(\mathbf{X},\mathbf{Y}) + \widehat{\mathtt{MRP}}_2(\mathbf{X},\mathbf{Y}),
$$

where

$$
\widehat{\mathtt{MRP}}_1(\mathbf{X},\mathbf{Y})
$$

$$
\begin{aligned}
= & \frac{1}{n(n-1)} \sum_{\substack{i,j=1 \\ i \neq j}}^{n} \int_0^1 \int_0^1 [\mathbf{X}_i(t) - \boldsymbol{\mu}_1(t)]^{\mathrm{T}} [\mathbf{X}_j(s) - \boldsymbol{\mu}_1(s)]\, v(s,t)\, \mathrm{d}s\, \mathrm{d}t \\
& + \frac{1}{m(m-1)} \sum_{\substack{i,j=1 \\ i \neq j}}^{m} \int_0^1 \int_0^1 [\mathbf{Y}_i(t) - \boldsymbol{\mu}_2(t)]^{\mathrm{T}} [\mathbf{Y}_j(s) - \boldsymbol{\mu}_2(s)]\, v(s,t)\, \mathrm{d}s\, \mathrm{d}t \\
& - \frac{2}{nm} \sum_{i=1}^{n} \sum_{j=1}^{m} \int_0^1 \int_0^1 [\mathbf{X}_i(t) - \boldsymbol{\mu}_1(t)]^{\mathrm{T}} [\mathbf{Y}_j(s) - \boldsymbol{\mu}_2(s)]\, v(s,t)\, \mathrm{d}s\, \mathrm{d}t
\end{aligned}
$$

and

$$
\begin{aligned}
\widehat{\mathtt{MRP}}_2(\mathbf{X},\mathbf{Y}) = & \frac{2}{n} \sum_{i=1}^{n} \int_0^1 \int_0^1 [\mathbf{X}_i(t) - \boldsymbol{\mu}_1(t)]^{\mathrm{T}} [\boldsymbol{\mu}_1(s) - \boldsymbol{\mu}_2(s)]\, v(s,t)\, \mathrm{d}s\, \mathrm{d}t \\
& + \frac{2}{m} \sum_{i=1}^{m} \int_0^1 \int_0^1 [\mathbf{Y}_i(t) - \boldsymbol{\mu}_2(t)]^{\mathrm{T}} [\boldsymbol{\mu}_2(s) - \boldsymbol{\mu}_1(s)]\, v(s,t)\, \mathrm{d}s\, \mathrm{d}t \\
& + \int_0^1 \int_0^1 [\boldsymbol{\mu}_1(t) - \boldsymbol{\mu}_2(t)]^{\mathrm{T}} [\boldsymbol{\mu}_1(s) - \boldsymbol{\mu}_2(s)]\, v(s,t)\, \mathrm{d}s\, \mathrm{d}t.
\end{aligned}
$$

It is easy to show that

$$
\mathrm{E}\{\widehat{\mathtt{MRP}}_1(\mathbf{X},\mathbf{Y})\} = 0, \quad \mathrm{E}\{\widehat{\mathtt{MRP}}_2(\mathbf{X},\mathbf{Y})\} = \mathtt{MRP}(\mathbf{X},\mathbf{Y}),
$$

and

$$
\begin{aligned}
\mathrm{Var}\{\widehat{\mathtt{MRP}}_2(\mathbf{X},\mathbf{Y})\} = & \frac{4}{n} \iiiint [\boldsymbol{\mu}_1(t) - \boldsymbol{\mu}_2(t)]^{\mathrm{T}} \mathbf{G}_1(s,s_1) [\boldsymbol{\mu}_1(t_1) - \boldsymbol{\mu}_2(t_1)] \\
& \times v(s,t) v(s_1,t_1)\, \mathrm{d}s\, \mathrm{d}t\, \mathrm{d}s_1\, \mathrm{d}t_1 \\
& + \frac{4}{m} \iiiint [\boldsymbol{\mu}_1(t) - \boldsymbol{\mu}_2(t)]^{\mathrm{T}} \mathbf{G}_2(s,s_1) [\boldsymbol{\mu}_1(t_1) - \boldsymbol{\mu}_2(t_1)] \\
& \times v(s,t) v(s_1,t_1)\, \mathrm{d}s\, \mathrm{d}t\, \mathrm{d}s_1\, \mathrm{d}t_1.
\end{aligned}
$$

Under Condition (C3), we obtain that

$$\mathrm{Var}\left\{\frac{\widehat{\mathtt{MRP}}_2(\mathbf{X},\mathbf{Y})-\mathtt{MRP}(\mathbf{X},\mathbf{Y})}{\sigma_{nm}(\mathbf{X},\mathbf{Y})}\right\}=o(1).$$

Thus, under Condition (C3) we have that

$$\begin{aligned}\frac{\widehat{\mathtt{MRP}}(\mathbf{X},\mathbf{Y})-\mathtt{MRP}(\mathbf{X},\mathbf{Y})}{\sqrt{\mathrm{Var}\{\widehat{\mathtt{MRP}}(\mathbf{X},\mathbf{Y})\}}}&=\frac{\widehat{\mathtt{MRP}}_1(\mathbf{X},\mathbf{Y})}{\sigma_{nm}(\mathbf{X},\mathbf{Y})\{1+o(1)\}}+\frac{\widehat{\mathtt{MRP}}_2(\mathbf{X},\mathbf{Y})-\mathtt{MRP}(\mathbf{X},\mathbf{Y})}{\sigma_{nm}(\mathbf{X},\mathbf{Y})\{1+o(1)\}}\\&=\frac{\widehat{\mathtt{MRP}}_1(\mathbf{X},\mathbf{Y})}{\sigma_{nm}(\mathbf{X},\mathbf{Y})}+o_p(1).\end{aligned}\tag{S3.8}$$

Under Condition (C3'), we have that

$$\frac{\widehat{\mathtt{MRP}}(\mathbf{X},\mathbf{Y})-\mathtt{MRP}(\mathbf{X},\mathbf{Y})}{\sqrt{\mathrm{Var}\{\widehat{\mathtt{MRP}}(\mathbf{X},\mathbf{Y})\}}}=\frac{\widehat{\mathtt{MRP}}_2(\mathbf{X},\mathbf{Y})-\mathtt{MRP}(\mathbf{X},\mathbf{Y})}{\sigma_{nm_2}(\mathbf{X},\mathbf{Y})}+o_p(1).\tag{S3.9}$$

To obtain the asymptotic normality of $\widehat{\mathtt{MRP}}_1(\mathbf{X},\mathbf{Y})$, we assume without loss of generality that $\boldsymbol{\mu}_1(t)=\boldsymbol{\mu}_2(t)=\mathbf{0}$. For simplicity, define $\mathbf{W}_i=\mathbf{X}_i$ for $i=1,\ldots,n$ and $\mathbf{W}_{j+n}=\mathbf{Y}_j$ for $j=1,\ldots,m$, and for $i\neq j$

$$\phi_{ij}=\begin{cases}\frac{1}{n(n-1)}\iint\mathbf{W}_i(t)^{\mathrm{T}}\mathbf{W}_j(s)v(s,t)\,\mathrm{d}s\,\mathrm{d}t,\text{ if } i,j\in\{1,\ldots,n\}\,,\\ -\frac{1}{nm}\iint\mathbf{W}_i(t)^{\mathrm{T}}\mathbf{W}_j(s)v(s,t)\,\mathrm{d}s\,\mathrm{d}t,\text{ if } i\in\{1,\ldots,n\},j\in\{n+1,\ldots,n+m\},\\ \frac{1}{m(m-1)}\iint\mathbf{W}_i(t)^{\mathrm{T}}\mathbf{W}_j(s)v(s,t)\,\mathrm{d}s\,\mathrm{d}t,\text{ if } i,j\in\{n+1,\ldots,n+m\}\,.\end{cases}$$

Define

$$V_j=\sum_{i=1}^{j-1}\phi_{ij},\ \text{for } j=2,3,\ldots,n+m;\quad S_k=\sum_{j=2}^{k}V_j;$$

$$\mathcal{F}_k=\sigma\{\mathbf{W}_1(t),\ldots,\mathbf{W}_k(t),t\in[0,1]\},$$

where $\sigma\{\mathbf{W}_1(t), \ldots, \mathbf{W}_k(t), t \in [0,1]\}$ is the σ-algebra generated by $\{\mathbf{W}_1(t), \ldots, \mathbf{W}_k(t)\}$, $t \in [0,1]$. Then, by the definition of $\widehat{\mathtt{MRP}}_1(\mathbf{X}, \mathbf{Y})$, we have that

$$\widehat{\mathtt{MRP}}_1(\mathbf{X}, \mathbf{Y}) = 2 \sum_{j=2}^{n+m} \sum_{i=1}^{j-1} \phi_{ij} = 2 \sum_{j=2}^{n+m} V_j.$$

Then, by Lemmas S.2-S.4 in Appendix B and Corollary 3.1 of Hall and Heyde (1980), we have that

$$\widehat{\mathtt{MRP}}_1(\mathbf{X}, \mathbf{Y})/\sigma_{nm}(\mathbf{X}, \mathbf{Y}) \xrightarrow{D} N(0,1), \quad \text{as } p \to \infty \text{ and } \min\{n, m\} \to \infty.$$

This, together with (S3.8), completes the proof of Theorem 2. Under condition (C3'), since $\widehat{\mathtt{MRP}}_2(\mathbf{X}, \mathbf{Y})$ is the sum of two independent averages, its asymptotic normality can be attained by following the standard means.

Proof of Theorem 3. Let $N = n + m$, $T_{ab} = \mathtt{ITR}(\mathbf{G}_a, \mathbf{G}_b)$, $\widehat{T}_{ab} = \widehat{\mathtt{ITR}}(\mathbf{G}_a, \mathbf{G}_b)$, and $\mathcal{T} = \mathtt{ITR}(\mathbf{G}_{12}, \mathbf{G}_{12})$. The positive semidefiniteness of v and of the covariance operators implies $T_{ab} \geq 0$, and Cauchy–Schwarz gives $T_{12}^2 \leq T_{11}T_{22}$. Since $T_{12} = T_{21}$, Condition (C2) gives $\sigma_{nm}^2(\mathbf{X}, \mathbf{Y}) \asymp N^{-2}\mathcal{T}$. We first derive the required expectation and variance bounds for

$$\begin{aligned}
\widehat{T}_{11} &= \iiiint \widehat{\mathtt{TR}}_{11}(s, s_1, t, t_1) v(s,t) v(s_1, t_1) \, \mathrm{d}s \, \mathrm{d}t \, \mathrm{d}s_1 \, \mathrm{d}t_1 \\
&= \iiiint \operatorname{tr}\{\mathbf{G}_1(\widehat{s, s_1)\mathbf{G}}_1(t, t_1)\} v(s,t) v(s_1, t_1) \, \mathrm{d}s \, \mathrm{d}t \, \mathrm{d}s_1 \, \mathrm{d}t_1
\end{aligned}$$

and then apply the same argument to $\widehat{T}_{22}$ and $\widehat{T}_{12}$. Consider the following decomposition

$$\operatorname{tr}\{\mathbf{G}_1(\widehat{s, s_1)\mathbf{G}}_1(t, t_1)\}$$

$$
\begin{aligned}
&= \frac{1}{n(n-1)} \operatorname{tr}\Big\{ \sum_{j\neq k}^{n} (\mathbf{X}_j(s) - \overline{\mathbf{X}}_{(j,k)}(s))\mathbf{X}_j^{\mathrm{T}}(s_1)(\mathbf{X}_k(t) - \overline{\mathbf{X}}_{(j,k)}(t))\mathbf{X}_k^{\mathrm{T}}(t_1)\Big\} \\
&= \frac{1}{n(n-1)} \operatorname{tr}\left\{ \sum_{j\neq k}^{n} \Big\{ \left(\mathbf{X}_j(s) - \boldsymbol{\mu}_1(s)\right)\left(\mathbf{X}_j(s_1) - \boldsymbol{\mu}_1(s_1)\right)^{\mathrm{T}}\left(\mathbf{X}_k(t) - \boldsymbol{\mu}_1(t)\right)\right. \\
&\quad \times \left(\mathbf{X}_k(t_1) - \boldsymbol{\mu}_1(t_1)\right)^{\mathrm{T}} - 2\left(\overline{\mathbf{X}}_{(j,k)}(s) - \boldsymbol{\mu}_1(s)\right)\left(\mathbf{X}_j(s_1) - \boldsymbol{\mu}_1(s_1)\right)^{\mathrm{T}} \\
&\quad \times \left(\mathbf{X}_k(t) - \boldsymbol{\mu}_1(t)\right)\left(\mathbf{X}_k(t_1) - \boldsymbol{\mu}_1(t_1)\right)^{\mathrm{T}} \Big\} \\
&\quad + \sum_{j\neq k}^{n} \Big\{ 2\left(\mathbf{X}_j(s) - \boldsymbol{\mu}_1(s)\right)\boldsymbol{\mu}_1(s_1)^{\mathrm{T}}\left(\mathbf{X}_k(t) - \boldsymbol{\mu}_1(t)\right)\left(\mathbf{X}_k(t_1) - \boldsymbol{\mu}_1(t_1)\right)^{\mathrm{T}} \\
&\quad -2\left(\overline{\mathbf{X}}_{(j,k)}(s) - \boldsymbol{\mu}_1(s)\right)\boldsymbol{\mu}_1(s_1)^{\mathrm{T}}\left(\mathbf{X}_k(t) - \boldsymbol{\mu}_1(t)\right)\left(\mathbf{X}_k(t_1) - \boldsymbol{\mu}_1(t_1)\right)^{\mathrm{T}} \Big\} \\
&\quad + \sum_{j\neq k}^{n} \Big\{ \left(\overline{\mathbf{X}}_{(j,k)}(s) - \boldsymbol{\mu}_1(s)\right)\left(\mathbf{X}_j(s_1) - \boldsymbol{\mu}_1(s_1)\right)^{\mathrm{T}}\left(\bar{\mathbf{X}}_{(j,k)}(t) - \boldsymbol{\mu}_1(t)\right) \\
&\quad \times \left(\mathbf{X}_k(t_1) - \boldsymbol{\mu}_1(t_1)\right)^{\mathrm{T}} \Big\} \\
&\quad - \sum_{j\neq k}^{n} \Big\{ 2\left(\mathbf{X}_j(s) - \boldsymbol{\mu}_1(s)\right)\boldsymbol{\mu}_1(s_1)^{\mathrm{T}}\left(\overline{\mathbf{X}}_{(j,k)}(t) - \boldsymbol{\mu}_1(t)\right)\left(\mathbf{X}_k(t_1) - \boldsymbol{\mu}_1(t_1)\right)^{\mathrm{T}} \\
&\quad -2\left(\overline{\mathbf{X}}_{(j,k)}(s) - \boldsymbol{\mu}_1(s)\right)\left(\mathbf{X}_j(s_1) - \boldsymbol{\mu}_1(s_1)\right)^{\mathrm{T}}\left(\overline{\mathbf{X}}_{(j,k)}(t) - \boldsymbol{\mu}_1(t)\right)\boldsymbol{\mu}_1(t_1)^{\mathrm{T}} \Big\} \\
&\quad + \sum_{j\neq k}^{n} \{ \left(\mathbf{X}_j(s) - \boldsymbol{\mu}_1(s)\right)\boldsymbol{\mu}_1(s_1)^{\mathrm{T}}\left(\mathbf{X}_k(t) - \boldsymbol{\mu}_1(t)\right)\boldsymbol{\mu}_1(t_1)^{\mathrm{T}} \\
&\quad -2\left(\overline{\mathbf{X}}_{(j,k)}(s) - \boldsymbol{\mu}_1(s)\right)\boldsymbol{\mu}_1(s_1)^{\mathrm{T}}\left(\mathbf{X}_k(t) - \boldsymbol{\mu}_1(t)\right)\boldsymbol{\mu}_1(t_1)^{\mathrm{T}} \} \\
&\quad \left. + \sum_{j\neq k}^{n} \left(\overline{\mathbf{X}}_{(j,k)}(s) - \boldsymbol{\mu}_1(s)\right)\boldsymbol{\mu}_1(s_1)^{\mathrm{T}}\left(\overline{\mathbf{X}}_{(j,k)}(t) - \boldsymbol{\mu}_1(t)\right)\boldsymbol{\mu}_1(t_1)^{\mathrm{T}} \right\} \\
&\triangleq \sum_{k=1}^{10} \operatorname{tr}\{\mathbf{A}_k\}. \qquad \text{(S3.10)}
\end{aligned}
$$

Simple calculations show that $\mathrm{E}\{\operatorname{tr}\{\mathbf{A}_1\}\} = \operatorname{tr}\{\mathbf{G}_1(s,s_1)\mathbf{G}_1(t,t_1)\}$, $\mathrm{E}\{\operatorname{tr}\{\mathbf{A}_k\}\} = 0, k = 2,\ldots,9$, $\mathrm{E}\{\operatorname{tr}\{\mathbf{A}_{10}\}\} = \boldsymbol{\mu}_1(s_1)^{\mathrm{T}}\mathbf{G}_1(s,t)\boldsymbol{\mu}_1(t_1)/(n-2)$. This, together with Condition (C3), yields

$$\begin{aligned}
&\mathrm{E}\left\{\iiiint \mathrm{tr}\,\{\mathbf{G}_1(\widehat{s,s_1)\mathbf{G}}_1(t,t_1)\}v(s,t)v(s_1,t_1)\,\mathrm{d}s\,\mathrm{d}t\,\mathrm{d}s_1\,\mathrm{d}t_1\right\}\\
&\quad=\iiiint \mathrm{E}\{\mathrm{tr}\{\mathbf{G}_1(\widehat{s,s_1)\mathbf{G}}_1(t,t_1)\}\}v(s,t)v(s_1,t_1)\,\mathrm{d}s\,\mathrm{d}t\,\mathrm{d}s_1\,\mathrm{d}t_1\\
&\quad=T_{11}+o(\mathcal{T}). \qquad\qquad \text{(S3.11)}
\end{aligned}$$

Note that

$$\begin{aligned}
&\mathrm{Var}\left\{\iiiint \mathrm{tr}\{\mathbf{G}_1(\widehat{s,s_1)\mathbf{G}}_1(t,t_1)\}v(s,t)v(s_1,t_1)\,\mathrm{d}s\,\mathrm{d}t\,\mathrm{d}s_1\,\mathrm{d}t_1\right\}\\
\leq\ & 10\sum_{k=1}^{10}\mathrm{Var}\left\{\iiiint \mathrm{tr}\{\mathbf{A}_k\}v(s,t)v(s_1,t_1)\,\mathrm{d}s\,\mathrm{d}t\,\mathrm{d}s_1\,\mathrm{d}t_1\right\}. \qquad \text{(S3.12)}
\end{aligned}$$

We next establish the orders of each term $\mathrm{Var}\left\{\iiiint \mathrm{tr}\{\mathbf{A}_k\}v(s,t)v(s_1,t_1)\,\mathrm{d}s\,\mathrm{d}t\,\mathrm{d}s_1\,\mathrm{d}t_1\right\}$ to obtain the order of $\mathrm{Var}\left\{\iiiint \mathrm{tr}\{\mathbf{G}_1(\widehat{s,s_1)\mathbf{G}}_1(t,t_1)\}v(s,t)v(s_1,t_1)\,\mathrm{d}s\,\mathrm{d}t\,\mathrm{d}s_1\,\mathrm{d}t_1\right\}$.

We first consider the term $\mathrm{Var}\left\{\iiiint \mathrm{tr}\{\mathbf{A}_1\}v(s,t)v(s_1,t_1)\,\mathrm{d}s\,\mathrm{d}t\,\mathrm{d}s_1\,\mathrm{d}t_1\right\}$.

Since

$$\begin{aligned}
&\mathrm{E}\left\{\Big[\iiiint \mathrm{tr}\,\{\mathbf{A}_1\}\,v(s,t)v(s_1,t_1)\,\mathrm{d}s\,\mathrm{d}t\,\mathrm{d}s_1\,\mathrm{d}t_1\Big]^2\right\}\\
&=\frac{1}{n^2(n-1)^2}\mathrm{E}\Big\{\iiiint \mathrm{tr}\{\sum_{j_1\neq k_1}^{n}(\mathbf{X}_{j_1}(s)-\boldsymbol{\mu}_1(s))\,(\mathbf{X}_{j_1}(s_1)-\boldsymbol{\mu}_1(s_1))^{\mathrm{T}}\\
&\quad\times(\mathbf{X}_{k_1}(t)-\boldsymbol{\mu}_1(t))\,(\mathbf{X}_{k_1}(t_1)-\boldsymbol{\mu}_1(t_1))^{\mathrm{T}}\}v(s,t)v(s_1,t_1)\,\mathrm{d}s\,\mathrm{d}t\,\mathrm{d}s_1\,\mathrm{d}t_1\\
&\quad\times\iiiint \mathrm{tr}\{\sum_{j_2\neq k_2}^{n}(\mathbf{X}_{j_2}(s_2)-\boldsymbol{\mu}_1(s_2))\,(\mathbf{X}_{j_2}(s_3)-\boldsymbol{\mu}_1(s_3))^{\mathrm{T}}\\
&\quad\times(\mathbf{X}_{k_2}(t_2)-\boldsymbol{\mu}_1(t_2))\,(\mathbf{X}_{k_2}(t_3)-\boldsymbol{\mu}_1(t_3))^{\mathrm{T}}\}v(s_2,t_2)v(s_3,t_3)\,\mathrm{d}s_2\,\mathrm{d}t_2\,\mathrm{d}s_3\,\mathrm{d}t_3\Big\},
\end{aligned}$$

we have that

$$
\begin{aligned}
& \operatorname{Var}\left\{\iiiint \operatorname{tr}\left\{\mathbf{A}_1\right\} v(s,t)v(s_1,t_1)\,\mathrm{d}s\,\mathrm{d}t\,\mathrm{d}s_1\,\mathrm{d}t_1\right\} \\
= \; & \frac{2}{n(n-1)}B_{11} + \frac{4(n-2)}{n(n-1)}B_{12} + o(\mathcal{T}^2),
\end{aligned}
$$

where

$$
\begin{aligned}
B_{11} &= \mathrm{E}\Big\{\Big[\iint (\mathbf{X}_1(s)-\boldsymbol{\mu}_1(s))^{\mathrm{T}}(\mathbf{X}_2(t)-\boldsymbol{\mu}_1(t))v(s,t)\,\mathrm{d}s\,\mathrm{d}t\Big]^4\Big\}, \\
B_{12} &= \mathrm{E}\Big\{\Big[\iiiint (\mathbf{X}_1(s)-\boldsymbol{\mu}_1(s))^{\mathrm{T}}\mathbf{G}_1(t,t_1)(\mathbf{X}_1(s_1)-\boldsymbol{\mu}_1(s_1)) \\
&\qquad \times v(s,t)v(s_1,t_1)\,\mathrm{d}s\,\mathrm{d}t\,\mathrm{d}s_1\,\mathrm{d}t_1\Big]^2\Big\}.
\end{aligned}
$$

Let J_{1111} denote the eightfold trace integral in Condition (C4) with all four covariance kernels equal to $\mathbf{G}_1$. The uniform moment inequality in Condition (C1)(3), applied conditionally and then integrated over the time arguments, gives $B_{11} \le CT_{11}^2, B_{12} \le C\{T_{11}^2 + J_{1111}\}$. Consequently, (S3.13), Condition (C4), and $n \to \infty$ imply

$$
\begin{aligned}
& \operatorname{Var}\left\{\iiiint \operatorname{tr}(\mathbf{A}_1)v(s,t)v(s_1,t_1)\,\mathrm{d}s\,\mathrm{d}t\,\mathrm{d}s_1\,\mathrm{d}t_1\right\} \\
& \le \frac{C}{n^2}T_{11}^2 + \frac{C}{n}\{T_{11}^2 + J_{1111}\} + o(\mathcal{T}^2) = o(\mathcal{T}^2).
\end{aligned}
$$

By similar procedures, we obtain

$$
\operatorname{Var}\left\{\iiiint \operatorname{tr}\{\mathbf{A}_k\}v(s,t)v(s_1,t_1)\,\mathrm{d}s\,\mathrm{d}t\,\mathrm{d}s_1\,\mathrm{d}t_1\right\} = o(\mathcal{T}^2), \qquad k = 2, \ldots, 10.
$$

These, together with (S3.12), yield that

$$
\operatorname{Var}(\widehat{T}_{11}) = o(\mathcal{T}^2). \tag{S3.13}
$$

The same expectation and variance calculations for the second sample and the cross-sample estimator give

$$\mathrm{E}(\widehat{T}_{ab} - T_{ab}) = o(\mathcal{T}), \mathrm{Var}(\widehat{T}_{ab}) = o(\mathcal{T}^2),\ (a,b) \in \{(1,1),(2,2),(1,2)\}. \tag{S3.14}$$

Using the definition of $\widehat{\sigma}^2_{nm}$ in (2.7), Condition (C2), and (S3.14), we obtain

$$\mathrm{E}\{\widehat{\sigma}^2_{nm} - \sigma^2_{nm}\} = o(N^{-2}\mathcal{T}) = o(\sigma^2_{nm}),$$

$$\mathrm{Var}(\widehat{\sigma}^2_{nm}) \le CN^{-4} \sum_{(a,b)\in\{(1,1),(2,2),(1,2)\}} \mathrm{Var}(\widehat{T}_{ab}) = o(N^{-4}\mathcal{T}^2) = o(\sigma^4_{nm}).$$

Chebyshev's inequality now implies $\widehat{\sigma}^2_{nm}(\mathbf{X},\mathbf{Y})/\sigma^2_{nm}(\mathbf{X},\mathbf{Y}) \to_p 1$, which completes the proof.

Proof of Theorem 4. By Theorem 3, Condition (C2), and the definition of σ^2_{nm}, we have

$$\begin{aligned}
&\widehat{\sigma}^2_{nm}(\mathbf{X},\mathbf{Y})\\
=\ & \sigma^2_{nm}(\mathbf{X},\mathbf{Y})\{1+o_p(1)\}\\
=\ & \frac{2(n+m)^2}{n^2m^2}\Big[\frac{m^2}{(n+m)^2}\iiiint \mathrm{tr}\{\mathbf{G}_1(s,s_1)\mathbf{G}_1(t,t_1)\}v(s,t)v(s_1,t_1)\,\mathrm{d}s\,\mathrm{d}t\,\mathrm{d}s_1\,\mathrm{d}t_1\\
&+\frac{n^2}{(n+m)^2}\iiiint \mathrm{tr}\{\mathbf{G}_2(s,s_1)\mathbf{G}_2(t,t_1)\}v(s,t)v(s_1,t_1)\,\mathrm{d}s\,\mathrm{d}t\,\mathrm{d}s_1\,\mathrm{d}t_1\\
&+\frac{2nm}{(n+m)^2}\iiiint \mathrm{tr}\{\mathbf{G}_1(s,s_1)\mathbf{G}_2(t,t_1)\}v(s,t)v(s_1,t_1)\,\mathrm{d}s\,\mathrm{d}t\,\mathrm{d}s_1\,\mathrm{d}t_1\Big]\{1+o_p(1)\}\\
=\ & \frac{2(n+m)^2}{n^2m^2}\iiiint \mathrm{tr}\{\mathbf{G}_\tau(s,s_1)\mathbf{G}_\tau(t,t_1)\}v(s,t)v(s_1,t_1)\,\mathrm{d}s\,\mathrm{d}t\,\mathrm{d}s_1\,\mathrm{d}t_1\{1+o_p(1)\},
\end{aligned}$$

where $\mathbf{G}_\tau(s,t) = (1-\tau)\mathbf{G}_1(s,t)+\tau\mathbf{G}_2(s,t)$. Denote $\Delta_{nm}(\boldsymbol{\mu}_1-\boldsymbol{\mu}_2,\mathbf{G}_1,\mathbf{G}_2) = (n+m)\tau(1-\tau)\mathtt{MRP}(\mathbf{X},\mathbf{Y})/\{\sqrt{2\mathtt{ITR}(\mathbf{G}_\tau,\mathbf{G}_\tau)}\}$, where $\mathbf{G}_\tau(s,t) = (1-\tau)\mathbf{G}_1(s,t)+$

$\tau\mathbf{G}_2(s,t)$. Thus, Theorems 2 and 3 and Slutsky's theorem imply, under $H_1 : \boldsymbol{\mu}_1 \neq \boldsymbol{\mu}_2$, that

$$\mathrm{P}\{\widehat{Q}_n(\mathbf{X},\mathbf{Y}) > z_\alpha\} = \Phi\{-z_\alpha + \Delta_{nm}(\boldsymbol{\mu}_1 - \boldsymbol{\mu}_2, \mathbf{G}_1, \mathbf{G}_2)\} + o(1).$$

Under Condition (C3'), suppose that $\sigma_{nm_2}(\mathbf{X},\mathbf{Y}) > 0$ eventually, $\mathrm{P}\{\widehat{\sigma}^2_{nm}(\mathbf{X},\mathbf{Y}) > 0\} \to 1$, and $\widehat{\sigma}_{nm}(\mathbf{X},\mathbf{Y})/\sigma_{nm_2}(\mathbf{X},\mathbf{Y}) = o_p(1)$. Let

$$Z_{nm_2} = \frac{\widehat{\mathtt{MRP}}(\mathbf{X},\mathbf{Y}) - \mathtt{MRP}(\mathbf{X},\mathbf{Y})}{\sigma_{nm_2}(\mathbf{X},\mathbf{Y})}.$$

The C3' version of Theorem 2 gives $Z_{nm_2} \xrightarrow{D} N(0,1)$. Hence, by Slutsky's theorem,

$$Z_{nm_2} - z_\alpha \frac{\widehat{\sigma}_{nm}(\mathbf{X},\mathbf{Y})}{\sigma_{nm_2}(\mathbf{X},\mathbf{Y})} \xrightarrow{D} N(0,1).$$

Since the limiting distribution is continuous, Pólya's theorem yields uniform convergence of the corresponding distribution functions. Therefore,

$$\begin{aligned}
\mathrm{P}\{\widehat{Q}_n(\mathbf{X},\mathbf{Y}) > z_\alpha\} &= \mathrm{P}\left\{Z_{nm_2} - z_\alpha \frac{\widehat{\sigma}_{nm}(\mathbf{X},\mathbf{Y})}{\sigma_{nm_2}(\mathbf{X},\mathbf{Y})} > -\frac{\mathtt{MRP}(\mathbf{X},\mathbf{Y})}{\sigma_{nm_2}(\mathbf{X},\mathbf{Y})}\right\} + o(1) \\
&= \Phi\left\{\frac{\mathtt{MRP}(\mathbf{X},\mathbf{Y})}{\sigma_{nm_2}(\mathbf{X},\mathbf{Y})}\right\} + o(1) \\
&= \Phi\left\{\Delta_{nm2}(\boldsymbol{\mu}_1 - \boldsymbol{\mu}_2, \mathbf{G}_1, \mathbf{G}_2)\right\} + o(1),
\end{aligned}$$

where the first equality uses $\mathrm{P}\{\widehat{\sigma}^2_{nm} > 0\} \to 1$. For $\tau_{nm} = n/(n+m)$, the definitions of σ_{nm_2} and $\mathbf{G}_{\tau_{nm}}$ give the finite-sample noncentrality parameter. Condition (C2) implies that its ratio to Δ_{nm2} tends to 1; because both quantities are nonnegative, the last equality follows.

Proof of Theorem 5. Let $N = n + m$ and let R_{nmp} be as defined in Condition (C5). By telescoping each bilinear kernel in the MRP statistic and applying Hölder's inequality, the uniform boundedness of v and of the fourth moments in Condition (C5) gives

$$\mathrm{E}\left|\widehat{\mathtt{MRP}}(\widehat{\mathbf{X}}, \widehat{\mathbf{Y}}) - \widehat{\mathtt{MRP}}(\mathbf{X}, \mathbf{Y})\right| \leq CR_{nmp} = o\{\sigma_{nm}(\mathbf{X}, \mathbf{Y})\}. \tag{S3.15}$$

Hence the difference on the left of (S3.15) is $o_p\{\sigma_{nm}(\mathbf{X}, \mathbf{Y})\}$.

More explicitly, put $\rho_{\max} = \max_k(r_{X,k} + r_{Y,k})$. For any pair of curves U, V from either sample, writing $e_U = \widehat{U} - U$ and $e_V = \widehat{V} - V$ gives

$$\begin{aligned}&\mathrm{E}\left|\iint \{\widehat{U}(s)^{\mathrm{T}}\widehat{V}(t) - U(s)^{\mathrm{T}}V(t)\}v(s,t)\,\mathrm{d}s\,\mathrm{d}t\right| \\ &\leq C\{R_{nmp} + \rho_{\max}R_{nmp}\} = O(R_{nmp}),\end{aligned}$$

where the two single-error terms follow from Hölder's inequality and the double-error term is absorbed using $\rho_{\max} = o(1)$. Leave-out sample means obey the same bounds by Jensen's inequality.

For $a, b \in \{1, 2\}$, denote the true-curve and reconstructed versions of the trace estimator by $\widehat{I}_{ab}$ and $\widehat{I}_{ab}^R$, respectively. A telescoping expansion of the four factors in each product, followed by Hölder's inequality and the same uniform fourth-moment bounds, yields

$$\mathrm{E}|\widehat{I}_{ab}^R - \widehat{I}_{ab}| \leq C\{pR_{nmp} + p\rho_{\max}R_{nmp} + R_{nmp}^2\} = O(pR_{nmp}),$$

because $R_{nmp} \leq p\rho_{\max}$. Because the coefficients of the three trace estimators in $\widehat{\sigma}_{nm}^2$ are $O(N^{-2})$ under Condition (C2), Condition (C5) implies

$$\widehat{\sigma}_{nm}^2(\widehat{\mathbf{X}}, \widehat{\mathbf{Y}}) - \widehat{\sigma}_{nm}^2(\mathbf{X}, \mathbf{Y}) = o_p\{\sigma_{nm}^2(\mathbf{X}, \mathbf{Y})\}.$$

Combining this result with Theorem 3 gives $\widehat{\sigma}_{nm}^2(\widehat{\mathbf{X}}, \widehat{\mathbf{Y}})/\sigma_{nm}^2(\mathbf{X}, \mathbf{Y}) \to_p 1$. Finally,

$$\begin{aligned}
&\frac{\widehat{\mathtt{MRP}}(\widehat{\mathbf{X}}, \widehat{\mathbf{Y}}) - \mathtt{MRP}(\mathbf{X}, \mathbf{Y})}{\widehat{\sigma}_{nm}(\widehat{\mathbf{X}}, \widehat{\mathbf{Y}})} \\
&= \frac{\widehat{\mathtt{MRP}}(\mathbf{X}, \mathbf{Y}) - \mathtt{MRP}(\mathbf{X}, \mathbf{Y})}{\sigma_{nm}(\mathbf{X}, \mathbf{Y})}\{1 + o_p(1)\} + o_p(1) \ \xrightarrow{D}\ N(0,1)
\end{aligned}$$

by Theorem 2 and Slutsky's theorem.

S3.1 Lemmas S.1-S.4

The following lemmas are needed to prove Theorem 2.

Lemma S.1 *We have that*

$$\begin{aligned}
&\mathrm{Var}\Big\{\sum_{\substack{i,j=1\\ i\neq j}}^{n} \iint \mathbf{X}_i(t)^{\mathrm{T}}\mathbf{X}_j(s)v(s,t)\,\mathrm{d}s\,\mathrm{d}t\Big\} \\
&= 2n(n-1)\iiiint \mathrm{tr}\left\{\mathbf{G}_1(t_1,t)\mathbf{G}_1(s_1,s)\right\} v(s,t)v(s_1,t_1)\,\mathrm{d}s\,\mathrm{d}t\,\mathrm{d}s_1\,\mathrm{d}t_1 \\
&\qquad + 4n(n-1)^2\iiiint \boldsymbol{\mu}_1(t)^{\mathrm{T}}\mathbf{G}_1(s,s_1)\boldsymbol{\mu}_1(t_1)v(s,t)v(s_1,t_1)\,\mathrm{d}s\,\mathrm{d}t\,\mathrm{d}s_1\,\mathrm{d}t_1.
\end{aligned}$$

Proof of Lemma S.1.

$$\mathrm{Var}\Big\{\sum_{\substack{i,j=1\\ i\neq j}}^{n} \iint \mathbf{X}_i(t)^{\mathrm{T}}\mathbf{X}_j(s)v(s,t)\,\mathrm{d}s\,\mathrm{d}t\Big\}$$

$$
\begin{aligned}
&= \mathrm{E}\Big\{\Big[\sum_{\substack{i,j=1\\ i\neq j}}^{n}\iint \mathbf{X}_i(t)^{\mathrm{T}}\mathbf{X}_j(s)v(s,t)\,\mathrm{d}s\,\mathrm{d}t\Big]^2\Big\}\\
&\quad -\Big[\mathrm{E}\Big\{\sum_{\substack{i,j=1\\ i\neq j}}^{n}\iint \mathbf{X}_i(t)^{\mathrm{T}}\mathbf{X}_j(s)v(s,t)\,\mathrm{d}s\,\mathrm{d}t\Big\}\Big]^2\\
&= \mathrm{E}\Big\{\sum_{i\neq j}^{n}\sum_{i'\neq j'}^{n}\iint \mathbf{X}_i(t)^{\mathrm{T}}\mathbf{X}_j(s)v(s,t)\,\mathrm{d}s\,\mathrm{d}t\iint \mathbf{X}_{i'}(t)^{\mathrm{T}}\mathbf{X}_{j'}(s)v(s,t)\,\mathrm{d}s\,\mathrm{d}t\Big\}\\
&\qquad -\Big[n(n-1)\iint \boldsymbol{\mu}_1(t)^{\mathrm{T}}\boldsymbol{\mu}_1(s)v(s,t)\,\mathrm{d}s\,\mathrm{d}t\Big]^2. \qquad \text{(S3.16)}
\end{aligned}
$$

We can divide the first term on the right hand side of (S3.16) into the following cases:

Case (1): For $\{i\neq j\neq i'\neq j'\}$, we have that

$$
\begin{aligned}
&\mathrm{E}\Big\{\sum_{i\neq j\neq i'\neq j'}^{n}\iint \mathbf{X}_i(t)^{\mathrm{T}}\mathbf{X}_j(s)v(s,t)\,\mathrm{d}s\,\mathrm{d}t\iint \mathbf{X}_{i'}(t)^{\mathrm{T}}\mathbf{X}_{j'}(s)v(s,t)\,\mathrm{d}s\,\mathrm{d}t\Big\}\\
&= n(n-1)(n-2)(n-3)\Big[\iint \boldsymbol{\mu}_1(t)^{\mathrm{T}}\boldsymbol{\mu}_1(s)v(s,t)\,\mathrm{d}s\,\mathrm{d}t\Big]^2. \qquad \text{(S3.17)}
\end{aligned}
$$

Case (2): For $\{i=i', j\neq j'\}\cup\{i=j', j\neq i'\}\cup\{j=i', i\neq j'\}\cup\{j=$

$j', i \neq i'\}$, we have that

$$
\begin{aligned}
&\mathrm{E}\Big\{\sum_{i\neq j}^{n}\sum_{i'\neq j'}^{n}\iint \mathbf{X}_i(t)^{\mathrm{T}}\mathbf{X}_j(s)v(s,t)\,\mathrm{d}s\,\mathrm{d}t\iint \mathbf{X}_{i'}(t)^{\mathrm{T}}\mathbf{X}_{j'}(s)v(s,t)\,\mathrm{d}s\,\mathrm{d}t\Big\}\\
&= 4\,\mathrm{E}\Big\{\sum_{i=i',j\neq j'}^{n}\iint \mathbf{X}_i(t)^{\mathrm{T}}\mathbf{X}_j(s)v(s,t)\,\mathrm{d}s\,\mathrm{d}t\iint \mathbf{X}_{i'}(t)^{\mathrm{T}}\mathbf{X}_{j'}(s)v(s,t)\,\mathrm{d}s\,\mathrm{d}t\Big\}\\
&= 4n(n-1)(n-2)\iiiint \mathrm{tr}\{\mathrm{E}\{\mathbf{X}_i(t)^{\mathrm{T}}\mathbf{X}_j(s)\mathbf{X}_i(t_1)^{\mathrm{T}}\mathbf{X}_{j'}(s_1)\}\}v(s,t)\\
&\qquad\qquad\times v(s_1,t_1)\,\mathrm{d}s\,\mathrm{d}t\,\mathrm{d}s_1\,\mathrm{d}t_1\\
&= 4n(n-1)(n-2)\iiiint \mathrm{tr}\{\mathrm{E}\{\mathbf{X}_i(t)\mathbf{X}_i(t_1)^{\mathrm{T}}\}\,\mathrm{E}\{\mathbf{X}_{j'}(s_1)\mathbf{X}_j(s)^{\mathrm{T}}\}\}v(s,t)\\
&\qquad\qquad\times v(s_1,t_1)\,\mathrm{d}s\,\mathrm{d}t\,\mathrm{d}s_1\,\mathrm{d}t_1\\
&= 4n(n-1)(n-2)\iiiint \mathrm{tr}\{[\mathbf{G}_1(t,t_1)+\boldsymbol{\mu}_1(t)\boldsymbol{\mu}_1(t_1)^{\mathrm{T}}]\boldsymbol{\mu}_1(s_1)\boldsymbol{\mu}_1(s)^{\mathrm{T}}\}\\
&\qquad\qquad\times v(s,t)v(s_1,t_1)\,\mathrm{d}s\,\mathrm{d}t\,\mathrm{d}s_1\,\mathrm{d}t_1\\
&= 4n(n-1)(n-2)\iiiint \boldsymbol{\mu}_1(s)^{\mathrm{T}}\mathbf{G}_1(t,t_1)\boldsymbol{\mu}_1(s_1)v(s,t)v(s_1,t_1)\,\mathrm{d}s\,\mathrm{d}t\,\mathrm{d}s_1\,\mathrm{d}t_1\\
&\quad + 4n(n-1)(n-2)\Big[\iint \boldsymbol{\mu}_1(t)^{\mathrm{T}}\boldsymbol{\mu}_1(s)v(s,t)\,\mathrm{d}s\,\mathrm{d}t\Big]^2.
\end{aligned}
$$

Case (3): For $\{i=i', j=j'\}\cup\{i=j', j=i'\}$, we have that

$$
\begin{aligned}
&\mathrm{E}\Big\{\sum_{i\neq j}^{n}\sum_{i'\neq j'}^{n}\iint \mathbf{X}_i(t)^{\mathrm{T}}\mathbf{X}_j(s)v(s,t)\,\mathrm{d}s\,\mathrm{d}t\iint \mathbf{X}_{i'}(t)^{\mathrm{T}}\mathbf{X}_{j'}(s)v(s,t)\,\mathrm{d}s\,\mathrm{d}t\Big\}\\
&= 2\,\mathrm{E}\Big\{\sum_{i=i',j=j'}^{n}\iint \mathbf{X}_i(t)^{\mathrm{T}}\mathbf{X}_j(s)v(s,t)\,\mathrm{d}s\,\mathrm{d}t\iint \mathbf{X}_i(t)^{\mathrm{T}}\mathbf{X}_j(s)v(s,t)\,\mathrm{d}s\,\mathrm{d}t\Big\}\\
&= 2n(n-1)\iiiint \mathrm{tr}\Big\{\mathrm{E}\{\mathbf{X}_i(t)\mathbf{X}_i(t_1)^{\mathrm{T}}\}\,\mathrm{E}\{\mathbf{X}_j(s_1)\mathbf{X}_j(s)^{\mathrm{T}}\}\Big\}v(s,t)\\
&\qquad\qquad\times v(s_1,t_1)\,\mathrm{d}s\,\mathrm{d}t\,\mathrm{d}s_1\,\mathrm{d}t_1
\end{aligned}
$$

$$
\begin{aligned}
&= 2n(n-1) \iiiint \operatorname{tr}\Big\{[\mathbf{G}_1(t,t_1) + \boldsymbol{\mu}_1(t)\boldsymbol{\mu}_1(t_1)^{\mathrm{T}}][\mathbf{G}_1(s,s_1) + \boldsymbol{\mu}_1(s)\boldsymbol{\mu}_1(s_1)^{\mathrm{T}}]\Big\} \\
&\qquad \times v(s,t)v(s_1,t_1)\,\mathrm{d}s\,\mathrm{d}t\,\mathrm{d}s_1\,\mathrm{d}t_1 \\
&= 2n(n-1) \iiiint \operatorname{tr}\{\mathbf{G}_1(t,t_1)\mathbf{G}_1(s,s_1)\} v(s,t)v(s_1,t_1)\,\mathrm{d}s\,\mathrm{d}t\,\mathrm{d}s_1\,\mathrm{d}t_1 \\
&\qquad + 4n(n-1) \iiiint \boldsymbol{\mu}_1(s)^{\mathrm{T}}\mathbf{G}_1(t,t_1)\boldsymbol{\mu}_1(s_1) v(s,t)v(s_1,t_1)\,\mathrm{d}s\,\mathrm{d}t\,\mathrm{d}s_1\,\mathrm{d}t_1 \\
&\qquad + 2n(n-1)\Big[\iint \boldsymbol{\mu}_1(t)^{\mathrm{T}}\boldsymbol{\mu}_1(s) v(s,t)\,\mathrm{d}s\,\mathrm{d}t\Big]^2. \qquad \text{(S3.18)}
\end{aligned}
$$

By (S3.16)-(S3.18), we have that

$$
\begin{aligned}
&\operatorname{Var}\Big\{\sum_{\substack{i,j=1\\ i\neq j}}^{n} \iint \mathbf{X}_i(t)^{\mathrm{T}}\mathbf{X}_j(s)v(s,t)\,\mathrm{d}s\,\mathrm{d}t\Big\} \\
= \;& n(n-1)(n-2)(n-3)\Big[\iint \boldsymbol{\mu}_1(t)^{\mathrm{T}}\boldsymbol{\mu}_1(s)v(s,t)\,\mathrm{d}s\,\mathrm{d}t\Big]^2 \\
& +4n(n-1)(n-2)\iint \boldsymbol{\mu}_1(s)^{\mathrm{T}}\mathbf{G}_1(t,t_1)\boldsymbol{\mu}_1(s_1)v(s,t)v(s_1,t_1)\,\mathrm{d}s\,\mathrm{d}t\,\mathrm{d}s_1\,\mathrm{d}t_1 \\
& +4n(n-1)(n-2)\Big[\iint \boldsymbol{\mu}_1(t)^{\mathrm{T}}\boldsymbol{\mu}_1(s)v(s,t)\,\mathrm{d}s\,\mathrm{d}t\Big]^2 \\
& +2n(n-1)\iiiint \operatorname{tr}\{\mathbf{G}_1(t,t_1)\mathbf{G}_1(s,s_1)\}v(s,t)v(s_1,t_1)\,\mathrm{d}s\,\mathrm{d}t\,\mathrm{d}s_1\,\mathrm{d}t_1 \\
& +4n(n-1)\iint \boldsymbol{\mu}_1(s)^{\mathrm{T}}\mathbf{G}_1(t,t_1)\boldsymbol{\mu}_1(s_1)v(s,t)v(s_1,t_1)\,\mathrm{d}s\,\mathrm{d}t\,\mathrm{d}s_1\,\mathrm{d}t_1 \\
& +2n(n-1)\Big[\iint \boldsymbol{\mu}_1(t)^{\mathrm{T}}\boldsymbol{\mu}_1(s)v(s,t)\,\mathrm{d}s\,\mathrm{d}t\Big]^2 \\
& -\Big[n(n-1)\iint \boldsymbol{\mu}_1(t)^{\mathrm{T}}\boldsymbol{\mu}_1(s)v(s,t)\,\mathrm{d}s\,\mathrm{d}t\Big]^2 \\
= \;& 2n(n-1)\iiiint \operatorname{tr}\{\mathbf{G}_1(t,t_1)\mathbf{G}_1(s,s_1)\}v(s,t)v(s_1,t_1)\,\mathrm{d}s\,\mathrm{d}t\,\mathrm{d}s_1\,\mathrm{d}t_1 \\
& +4n(n-1)^2\iint \boldsymbol{\mu}_1(s)^{\mathrm{T}}\mathbf{G}_1(t,t_1)\boldsymbol{\mu}_1(s_1)v(s,t)v(s_1,t_1)\,\mathrm{d}s\,\mathrm{d}t\,\mathrm{d}s_1\,\mathrm{d}t_1.
\end{aligned}
$$

Lemma S.2 *For any N, $\{S_j, \mathcal{F}_j\}_{j=1}^N$ is the sequence of zero mean and a* *square integrable martingale.*

Proof of Lemma S.2. By the definitions of S_j and $\mathcal{F}_j$, it is easy to show that $\mathcal{F}_{j-1} \subseteq \mathcal{F}_j$ and S_j is of zero mean and square integrable, for any $1 \le j \le n+m$.

We next show that $\{S_j\}_{j=1}^N$ is a martingale with respect to the filtration $\{\mathcal{F}_j\}_{j=1}^N$. We consider the following two cases:

Case (1): When $j \le k$, we have that

$$\mathrm{E}\{V_j \mid \mathcal{F}_k\} = \sum_{i=1}^{j-1} \mathrm{E}\{\phi_{ij} \mid \mathcal{F}_k\} = \sum_{i=1}^{j-1} \phi_{ij} = V_j.$$

Case (2): When $j > k$, we have that

$$\begin{aligned}
\mathrm{E}\{V_j \mid \mathcal{F}_k\} &= \sum_{i=1}^{k} \mathrm{E}\{\phi_{ij} \mid \mathcal{F}_k\} + \sum_{i=k+1}^{j-1} \mathrm{E}\{\phi_{ij} \mid \mathcal{F}_k\} \\
&= C_1 \sum_{i=1}^{k} \mathrm{E}\{\iint \mathbf{W}_i(t)^{\mathrm{T}} \mathbf{W}_j(s) v(s,t)\,\mathrm{d}s\,\mathrm{d}t \mid \mathcal{F}_k\} \\
&+ C_2 \sum_{i=k+1}^{j-1} \mathrm{E}\{\iint \mathbf{W}_i(t)^{\mathrm{T}} \mathbf{W}_j(s) v(s,t)\,\mathrm{d}s\,\mathrm{d}t \mid \mathcal{F}_k\} \\
&= C_1 \sum_{i=1}^{k} \iint \mathbf{W}_i(t)^{\mathrm{T}} \mathrm{E}\{\mathbf{W}_j(s) \mid \mathcal{F}_k\} v(s,t)\,\mathrm{d}s\,\mathrm{d}t \\
&+ C_2 \sum_{i=k+1}^{j-1} \iint \mathrm{E}\{\mathbf{W}_i(t)\}^{\mathrm{T}} \mathrm{E}\{\mathbf{W}_j(s)\} v(s,t)\,\mathrm{d}s\,\mathrm{d}t = 0,
\end{aligned}$$

where C_1 $C_2 \in \{\frac{1}{n(n-1)}, -\frac{1}{nm}, \frac{1}{m(m-1)}\}$.

Then, for any $q \geq k$, we have that

$$\mathrm{E}\{S_q \mid \mathcal{F}_k\} = \sum_{j=2}^{q} \mathrm{E}\{V_j \mid \mathcal{F}_k\} = \sum_{j=2}^{k} V_j = S_k.$$

Thus, $\{S_j\}_{j=1}^N$ is a martingale with respect to the filtration $\{\mathcal{F}_j\}_{j=1}^N$ for any N.

Lemma S.3 *Under Conditions (C1), (C2), and (C4), provided* $\sigma_{nm}^2(\mathbf{X}, \mathbf{Y}) > 0$*, we have that*

$$\sum_{j=2}^{n+m} \mathrm{E}\{V_j^2 \mid \mathcal{F}_{j-1}\}/\sigma_{nm}^2(\mathbf{X}, \mathbf{Y}) \xrightarrow{P} \frac{1}{4}.$$

Proof of Lemma S.3. For $a = 1, 2$, define

$$K_a(f, g) = \iiiint f(t)^{\mathrm{T}} \mathbf{G}_a(s, s_1) g(t_1) v(s, t) v(s_1, t_1) \,\mathrm{d}s \,\mathrm{d}t \,\mathrm{d}s_1 \,\mathrm{d}t_1.$$

By the definition of V_j and ϕ_{ij}, we have that

$$\mathrm{E}\{V_j^2 \mid \mathcal{F}_{j-1}\} = \mathrm{E}\{[\sum_{i=1}^{j-1} \phi_{ij}]^2 \mid \mathcal{F}_{j-1}\} = \sum_{i_1, i_2 = 1}^{j-1} \mathrm{E}\{\phi_{i_1 j} \phi_{i_2 j} \mid \mathcal{F}_{j-1}\},$$

which can be divided into the following three cases:

Case (1): When $j \leq n$, we have that

$$\begin{aligned}
&\mathrm{E}\{V_j^2 \mid \mathcal{F}_{j-1}\} \\
&= \frac{1}{n^2(n-1)^2} \mathrm{E}\Big\{ \sum_{i_1, i_2 = 1}^{j-1} \iiiint \mathbf{W}_{i_1}(t)^{\mathrm{T}} \mathbf{W}_j(s) \mathbf{W}_j(s_1)^{\mathrm{T}} \mathbf{W}_{i_2}(t_1) v(s, t) \\
&\qquad \times v(s_1, t_1) \,\mathrm{d}s \,\mathrm{d}t \,\mathrm{d}s_1 \,\mathrm{d}t_1 \mid \mathcal{F}_{j-1} \Big\}
\end{aligned}$$

$$= \frac{1}{n^2(n-1)^2}\sum_{i_1,i_2=1}^{j-1}\iiiint \mathbf{W}_{i_1}(t)^{\mathrm{T}}\mathbf{G}_1(s,s_1)\mathbf{W}_{i_2}(t_1)v(s,t)v(s_1,t_1)\,\mathrm{d}s\,\mathrm{d}t\,\mathrm{d}s_1\,\mathrm{d}t_1.$$

Case (2): When $j = n+1$, we have that

$$\begin{aligned}
&\mathrm{E}\{V_j^2 \mid \mathcal{F}_{j-1}\}\\
&= \frac{1}{n^2m^2}\mathrm{E}\Big\{\sum_{i_1,i_2=1}^{n}\iiiint \mathbf{W}_{i_1}(t)^{\mathrm{T}}\mathbf{W}_{n+1}(s)\mathbf{W}_{n+1}(s_1)^{\mathrm{T}}\mathbf{W}_{i_2}(t_1)v(s,t)\\
&\quad\times v(s_1,t_1)\,\mathrm{d}s\,\mathrm{d}t\,\mathrm{d}s_1\,\mathrm{d}t_1 \mid \mathcal{F}_{j-1}\Big\}\\
&= \frac{1}{n^2m^2}\sum_{i_1,i_2=1}^{n}\iiiint \mathbf{W}_{i_1}(t)^{\mathrm{T}}\mathbf{G}_2(s,s_1)\mathbf{W}_{i_2}(t_1)v(s,t)v(s_1,t_1)\,\mathrm{d}s\,\mathrm{d}t\,\mathrm{d}s_1\,\mathrm{d}t_1.
\end{aligned}$$

Case (3): When $j \ge n+2$, we have that

$$\begin{aligned}
&\mathrm{E}\{V_j^2 \mid \mathcal{F}_{j-1}\}\\
&= \frac{1}{n^2m^2}\sum_{i_1,i_2=1}^{n}\iiiint \mathbf{W}_{i_1}(t)^{\mathrm{T}}\mathbf{G}_2(s,s_1)\mathbf{W}_{i_2}(t_1)v(s,t)v(s_1,t_1)\,\mathrm{d}s\,\mathrm{d}t\,\mathrm{d}s_1\,\mathrm{d}t_1\\
&+ \frac{1}{m^2(m-1)^2}\sum_{i_1,i_2=n+1}^{j-1}\iiiint \mathbf{W}_{i_1}(t)^{\mathrm{T}}\mathbf{G}_2(s,s_1)\mathbf{W}_{i_2}(t_1)v(s,t)v(s_1,t_1)\,\mathrm{d}s\,\mathrm{d}t\,\mathrm{d}s_1\,\mathrm{d}t_1\\
&\quad - \frac{2}{nm^2(m-1)}\sum_{i=1}^{n}\sum_{r=n+1}^{j-1}K_2(\mathbf{W}_i,\mathbf{W}_r).
\end{aligned}$$

Denote $\eta_{n+m} = \sum_{j=2}^{n+m}\mathrm{E}\{V_j^2 \mid \mathcal{F}_{j-1}\}$. The added cross block has expectation zero because the two centered samples are mutually independent.

Based on the above Cases (1)–(3), we obtain that

$$\begin{aligned}
&\mathrm{E}\{\eta_{n+m}\}\\
= &\frac{1}{n^2(n-1)^2}\sum_{j=2}^{n}\sum_{i_1,i_2=1}^{j-1}\iiiint \mathrm{E}\{\mathbf{W}_{i_1}(t)^{\mathrm{T}}\mathbf{G}_1(s,s_1)\mathbf{W}_{i_2}(t_1)\}v(s,t)v(s_1,t_1)\,\mathrm{d}s\,\mathrm{d}t\,\mathrm{d}s_1\,\mathrm{d}t_1
\end{aligned}$$

$$
\begin{aligned}
&+\frac{1}{n^2m^2}\sum_{i_1,i_2=1}^{n}\iiiint \mathrm{E}\{\mathbf{W}_{i_1}(t)^{\mathrm{T}}\mathbf{G}_2(s,s_1)\mathbf{W}_{i_2}(t_1)\}v(s,t)v(s_1,t_1)\,\mathrm{d}s\,\mathrm{d}t\,\mathrm{d}s_1\,\mathrm{d}t_1\\
&+\sum_{j=n+2}^{n+m}\Big[\frac{1}{n^2m^2}\sum_{i_1,i_2=1}^{n}\iiiint \mathrm{E}\{\mathbf{W}_{i_1}(t)^{\mathrm{T}}\mathbf{G}_2(s,s_1)\mathbf{W}_{i_2}(t_1)\}v(s,t)v(s_1,t_1)\,\mathrm{d}s\,\mathrm{d}t\,\mathrm{d}s_1\,\mathrm{d}t_1\\
&+\frac{1}{m^2(m-1)^2}\sum_{i_1,i_2=n+1}^{j-1}\iiiint \mathrm{E}\{\mathbf{W}_{i_1}(t)^{\mathrm{T}}\mathbf{G}_2(s,s_1)\mathbf{W}_{i_2}(t_1)\}v(s,t)v(s_1,t_1)\,\mathrm{d}s\,\mathrm{d}t\,\mathrm{d}s_1\,\mathrm{d}t_1\Big]\\
=\ &\frac{1}{n^2(n-1)^2}\frac{n(n-1)}{2}\iiiint \mathrm{tr}\left\{\mathbf{G}_1(s,s_1)\mathbf{G}_1(t,t_1)\right\}v(s,t)v(s_1,t_1)\,\mathrm{d}s\,\mathrm{d}t\,\mathrm{d}s_1\,\mathrm{d}t_1\\
&+\frac{1}{m^2(m-1)^2}\frac{m(m-1)}{2}\iiiint \mathrm{tr}\left\{\mathbf{G}_2(s,s_1)\mathbf{G}_2(t,t_1)\right\}v(s,t)v(s_1,t_1)\,\mathrm{d}s\,\mathrm{d}t\,\mathrm{d}s_1\,\mathrm{d}t_1\\
&+\frac{1}{n^2m^2}mn\iiiint \mathrm{tr}\left\{\mathbf{G}_1(s,s_1)\mathbf{G}_2(t,t_1)\right\}v(s,t)v(s_1,t_1)\,\mathrm{d}s\,\mathrm{d}t\,\mathrm{d}s_1\,\mathrm{d}t_1\\
=\ &\frac{1}{4}\sigma_{nm}^2(\mathbf{X},\mathbf{Y}). \qquad\qquad (\mathrm{S3.19})
\end{aligned}
$$

Moreover, we have that

$$
\begin{aligned}
&\mathrm{E}\{\eta_{n+m}^2\}\\
=\ &\mathrm{E}\Big\{\frac{1}{n^2(n-1)^2}\sum_{j=2}^{n}\sum_{i_1,i_2=1}^{j-1}\iiiint \mathbf{W}_{i_1}(t)^{\mathrm{T}}\mathbf{G}_1(s,s_1)\mathbf{W}_{i_2}(t_1)v(s,t)\\
&\qquad\qquad\times v(s_1,t_1)\,\mathrm{d}s\,\mathrm{d}t\,\mathrm{d}s_1\,\mathrm{d}t_1\\
&+\frac{m}{n^2m^2}\sum_{i_1,i_2=1}^{n}\iiiint \mathbf{W}_{i_1}(t)^{\mathrm{T}}\mathbf{G}_2(s,s_1)\mathbf{W}_{i_2}(t_1)v(s,t)v(s_1,t_1)\,\mathrm{d}s\,\mathrm{d}t\,\mathrm{d}s_1\,\mathrm{d}t_1\\
&+\frac{1}{m^2(m-1)^2}\sum_{j=n+2}^{n+m}\sum_{i_1,i_2=n+1}^{j-1}\iiiint \mathbf{W}_{i_1}(t)^{\mathrm{T}}\mathbf{G}_2(s,s_1)\mathbf{W}_{i_2}(t_1)v(s,t)\\
&\qquad\qquad\times v(s_1,t_1)\,\mathrm{d}s\,\mathrm{d}t\,\mathrm{d}s_1\,\mathrm{d}t_1\\
&-\frac{2}{nm^2(m-1)}\sum_{j=n+2}^{n+m}\sum_{i=1}^{n}\sum_{r=n+1}^{j-1}K_2(\mathbf{W}_i,\mathbf{W}_r)\Big\}^2\\
=\ &\frac{1}{16}\sigma_{nm}^4(\mathbf{X},\mathbf{Y})+o\{\sigma_{nm}^4(\mathbf{X},\mathbf{Y})\}.
\end{aligned}
$$

Indeed, if $T = \mathtt{ITR}(\mathbf{G}_{12}, \mathbf{G}_{12})$ and J is the maximum absolute eightfold trace integral in Condition (C4), expanding the square, using independence and centering, and counting the diagonal and shared-index terms gives

$$\mathrm{Var}(\eta_{n+m}) \leq C\{N^{-5}T^2 + N^{-4}J\} = o\{\sigma_{nm}^4(\mathbf{X}, \mathbf{Y})\}, \qquad N = n + m.$$

Thus, we have

$$\mathrm{Var}\{\eta_{n+m}\} = \mathrm{E}\{\eta_{n+m}^2\} - (\mathrm{E}\{\eta_{n+m}\})^2 = o(\sigma_{nm}^4(\mathbf{X}, \mathbf{Y})). \tag{S3.20}$$

By (S3.19) and (S3.20), we have

$$\mathrm{E}\Big\{\sum_{j=2}^{n+m} \mathrm{E}\{V_j^2 \mid \mathcal{F}_{j-1}\}\Big\}/\sigma_{nm}^2(\mathbf{X}, \mathbf{Y}) = \mathrm{E}\{\eta_{n+m}\}/\sigma_{nm}^2(\mathbf{X}, \mathbf{Y}) = \frac{1}{4},$$

and

$$\mathrm{Var}\Big\{\sum_{j=2}^{n+m} \mathrm{E}\{V_j^2 \mid \mathcal{F}_{j-1}\}\Big\}/\sigma_{nm}^4(\mathbf{X}, \mathbf{Y}) = \mathrm{Var}\{\eta_{n+m}\}/\sigma_{nm}^4(\mathbf{X}, \mathbf{Y}) = o(1).$$

This completes the proof of Lemma S.3.

Lemma S.4 *Under Conditions (C1), (C2), and (C4), provided* $\sigma_{nm}^2(\mathbf{X}, \mathbf{Y}) > 0$, *for any* $\epsilon > 0$, *we have that*

$$\sum_{j=2}^{n+m} \mathrm{E}\{V_j^2 I(|V_j| > \epsilon\sigma_{nm}(\mathbf{X}, \mathbf{Y})) \mid \mathcal{F}_{j-1}\}/\sigma_{nm}^2(\mathbf{X}, \mathbf{Y}) \xrightarrow{P} 0.$$

Proof of Lemma S.4. Let $N = n + m$ and

$$H(U, V) = \iint U(s)^{\mathrm{T}} V(t) v(s, t)\, \mathrm{d}s\, \mathrm{d}t,\ T_{ab} = \mathtt{ITR}(\mathbf{G}_a, \mathbf{G}_b),\ \mathcal{T} = \sum_{a,b=1}^{2} T_{ab}.$$

Write $\phi_{ij} = \omega_{ij} H(\mathbf{W}_i, \mathbf{W}_j)$. By Condition (C2), $\max_{i<j} |\omega_{ij}| = O(N^{-2})$ and $\sigma_{nm}^2 \asymp N^{-2}\mathcal{T}$. To verify the required integrated moment bounds, let $\mathcal{G}_b$ denote the covariance operator of U_b, let $\mathcal{V}$ be the integral operator induced by v, and put $A_b = \mathcal{V}\mathcal{G}_b\mathcal{V}^*$. Conditional on U_a, Condition (C1)(3) gives

$$\mathrm{E}\{H(U_a, U_b)^4 \mid U_a\} \le C\langle U_a, A_b U_a\rangle_{L_2^p}^2.$$

Applying the same condition to finite-rank spectral approximations of A_b, followed by monotone approximation, yields

$$\mathrm{E}\langle U_a, A_b U_a\rangle_{L_2^p}^2 \le C\{\mathrm{tr}(A_b\mathcal{G}_a)\}^2 = CT_{ab}^2.$$

Consequently, for independent centered curves U_a, U_b, U_c from the indicated samples, and using Cauchy–Schwarz for the second bound,

$$\mathrm{E}\{H(U_a, U_b)^4\} \le CT_{ab}^2 \le C\mathcal{T}^2,$$
$$\mathrm{E}\{H(U_a, U_c)^2 H(U_b, U_c)^2\} \le C(T_{ac}T_{bc}) \le C\mathcal{T}^2.$$

The remaining eightfold covariance terms are controlled by Condition (C4).

In the expansion of $\sum_{j=2}^N \mathrm{E}(V_j^4)$, independence and centering eliminate every index pattern in which one of $\mathbf{W}_1, \ldots, \mathbf{W}_{j-1}$ occurs only once. There are $O(N^3)$ paired patterns and $O(N^2)$ diagonal patterns. Hence

$$\mathrm{E}\left\{\sum_{j=2}^N \mathrm{E}(V_j^4 \mid \mathcal{F}_{j-1})\right\} \le CN^{-8}\{N^3\mathcal{T}^2 + N^2\mathcal{T}^2\} = O(N^{-5}\mathcal{T}^2) = o(\sigma_{nm}^4).$$

Markov's inequality therefore gives $\sum_{j=2}^{N} \mathrm{E}(V_j^4 \mid \mathcal{F}_{j-1}) = o_p(\sigma_{nm}^4)$. Finally,

$$\frac{1}{\sigma_{nm}^2} \sum_{j=2}^{N} \mathrm{E}\{V_j^2 I(|V_j| > \epsilon\sigma_{nm}) \mid \mathcal{F}_{j-1}\} \\ \leq \epsilon^{-2}\sigma_{nm}^{-4} \sum_{j=2}^{N} \mathrm{E}(V_j^4 \mid \mathcal{F}_{j-1}) = o_p(1),$$

which proves the result. For completeness, the index-counting step used above can be written in the following expanded form. Note that

$$\sum_{j=2}^{n+m} \sigma_{nm}^{-2}(\mathbf{X},\mathbf{Y})\mathrm{E}\{V_j^2 I(|V_j| > \epsilon\sigma_{nm}(\mathbf{X},\mathbf{Y})) \mid \mathcal{F}_{j-1}\} \\ \leq \sigma_{nm}^{-4}(\mathbf{X},\mathbf{Y})\epsilon^{-2} \sum_{j=2}^{n+m} \mathrm{E}\{V_j^4 \mid \mathcal{F}_{j-1}\}$$

and

$$\mathrm{E}\{V_j^4 \mid \mathcal{F}_{j-1}\} = \mathrm{E}\Big\{\Big[\sum_{i=1}^{j-1} \phi_{ij}\Big]^4 \mid \mathcal{F}_{j-1}\Big\} = \sum_{i_1,i_2,i_3,i_4=1}^{j-1} \mathrm{E}\{\phi_{i_1 j}\phi_{i_2 j}\phi_{i_3 j}\phi_{i_4 j} \mid \mathcal{F}_{j-1}\}.$$

Therefore, we have

$$\mathrm{E}\Big\{\sum_{j=2}^{n+m} \mathrm{E}\{V_j^4 \mid \mathcal{F}_{j-1}\}\Big\} \leq C(n+m)^{-8}(3Q+P), \tag{S3.21}$$

where

$$Q = \sum_{j=2}^{n+m} \sum_{i_1 \neq i_2}^{j-1} \mathrm{E}\Big\{\int \cdots \int \mathbf{W}_j(t)^{\mathrm{T}}\mathbf{W}_{i_1}(s)\mathbf{W}_{i_1}(s_1)^{\mathrm{T}}\mathbf{W}_j(t_1)\mathbf{W}_j(t_2)^{\mathrm{T}}\mathbf{W}_{i_2}(s_2) \\ \times \mathbf{W}_{i_2}(s_3)^{\mathrm{T}}\mathbf{W}_j(t_3)v(s,t)v(s_1,t_1)v(s_2,t_2)v(s_3,t_3)\,\mathrm{d}s\,\mathrm{d}t\,\mathrm{d}s_1\,\mathrm{d}t_1\,\mathrm{d}s_2\,\mathrm{d}t_2\,\mathrm{d}s_3\,\mathrm{d}t_3\Big\}$$

and $P = \sum_{j=2}^{n+m}\sum_{i=1}^{j-1} \mathrm{E}\Big\{\iint \mathbf{W}_i(s)^{\mathrm{T}}\mathbf{W}_j(t)v(s,t)\,\mathrm{d}s\,\mathrm{d}t\Big\}^4$.

Note that

$$
\begin{aligned}
Q =& \sum_{j=2}^{n+m} \sum_{i_1 \neq i_2}^{j-1} \int \cdots \int \mathrm{E}\Big\{ \mathrm{tr}\{\mathbf{W}_j(t_3)\mathbf{W}_j(t)^{\mathrm{T}}\mathbf{W}_{i_1}(s)\mathbf{W}_{i_1}(s_1)^{\mathrm{T}}\mathbf{W}_j(t_1)\mathbf{W}_j(t_2)^{\mathrm{T}} \\
& \times \mathbf{W}_{i_2}(s_2)\mathbf{W}_{i_2}(s_3)^{\mathrm{T}}\}\Big\} v(s,t)v(s_1,t_1)v(s_2,t_2)v(s_3,t_3)\,\mathrm{d}s\,\mathrm{d}t\,\mathrm{d}s_1\,\mathrm{d}t_1\,\mathrm{d}s_2\,\mathrm{d}t_2\,\mathrm{d}s_3\,\mathrm{d}t_3 \\
=& \sum_{j=2}^{n} \sum_{i_1 \neq i_2}^{j-1} \int \cdots \int \mathrm{E}\Big\{ \mathbf{W}_j(t)^{\mathrm{T}}\mathbf{G}_1(s,s_1)\mathbf{W}_j(t_1)\mathbf{W}_j(t_2)^{\mathrm{T}}\mathbf{G}_1(s_2,s_3)\mathbf{W}_j(t_3)\Big\} \\
& \times v(s,t)v(s_1,t_1)v(s_2,t_2)v(s_3,t_3)\,\mathrm{d}s\,\mathrm{d}t\,\mathrm{d}s_1\,\mathrm{d}t_1\,\mathrm{d}s_2\,\mathrm{d}t_2\,\mathrm{d}s_3\,\mathrm{d}t_3 \\
& + \sum_{i_1 \neq i_2}^{n} \int \cdots \int \mathrm{E}\Big\{ \mathbf{W}_{n+1}(t)^{\mathrm{T}}\mathbf{G}_1(s,s_1)\mathbf{W}_{n+1}(t_1)\mathbf{W}_{n+1}(t_2)^{\mathrm{T}}\mathbf{G}_1(s_2,s_3)\mathbf{W}_{n+1}(t_3)\Big\} \\
& \times v(s,t)v(s_1,t_1)v(s_2,t_2)v(s_3,t_3)\,\mathrm{d}s\,\mathrm{d}t\,\mathrm{d}s_1\,\mathrm{d}t_1\,\mathrm{d}s_2\,\mathrm{d}t_2\,\mathrm{d}s_3\,\mathrm{d}t_3 \\
& + \sum_{j=n+2}^{n+m} \sum_{a,b=1}^{2} \sum_{\substack{i_1 \in \mathcal{I}_a(j),\ i_2 \in \mathcal{I}_b(j) \\ i_1 \neq i_2}} \int \cdots \int \mathrm{E}\Big\{ \mathbf{W}_j(t)^{\mathrm{T}}\mathbf{G}_a(s,s_1)\mathbf{W}_j(t_1)\mathbf{W}_j(t_2)^{\mathrm{T}}\mathbf{G}_b(s_2,s_3)\mathbf{W}_j(t_3)\Big\} \\
& \times v(s,t)v(s_1,t_1)v(s_2,t_2)v(s_3,t_3)\,\mathrm{d}s\,\mathrm{d}t\,\mathrm{d}s_1\,\mathrm{d}t_1\,\mathrm{d}s_2\,\mathrm{d}t_2\,\mathrm{d}s_3\,\mathrm{d}t_3 \\
=& O\{(n+m)^3\mathcal{T}^2\} = o\left\{(n+m)^8\sigma_{nm}^4(\mathbf{X},\mathbf{Y})\right\},
\end{aligned}
$$

where $\mathcal{I}_1(j) = \{1,\ldots,n\}$ and $\mathcal{I}_2(j) = \{n+1,\ldots,j-1\}$; the last equality follows from Condition (C1)(3), Condition (C4), and $\sigma_{nm}^2 \asymp (n+m)^{-2}\mathcal{T}$.

and

$$
\begin{aligned}
P &= \sum_{j=2}^{n+m} \sum_{i=1}^{j-1} \mathrm{E}\Big\{ \iint \mathbf{W}_i(s)^{\mathrm{T}}\mathbf{W}_j(t)v(s,t)\,\mathrm{d}s\,\mathrm{d}t\Big\}^4 \\
&= \sum_{j=2}^{n} \sum_{i=1}^{j-1} \mathrm{E}\Big\{ \iint \mathbf{W}_i(s)^{\mathrm{T}}\mathbf{W}_j(t)v(s,t)\,\mathrm{d}s\,\mathrm{d}t\Big\}^4
\end{aligned}
$$

$$+\sum_{j=n+1}^{n+m}\sum_{i=1}^{n}\mathrm{E}\Big\{\iint \mathbf{W}_i(s)^{\mathrm{T}}\mathbf{W}_j(t)v(s,t)\,\mathrm{d}s\,\mathrm{d}t\Big\}^4$$

$$+\sum_{j=n+1}^{n+m}\sum_{i=n+1}^{j-1}\mathrm{E}\Big\{\iint \mathbf{W}_i(s)^{\mathrm{T}}\mathbf{W}_j(t)v(s,t)\,\mathrm{d}s\,\mathrm{d}t\Big\}^4$$

$$=\sum_{j=2}^{n}\sum_{i=1}^{j-1}\mathrm{E}\Big\{\iint \mathbf{X}_i(s)^{\mathrm{T}}\mathbf{X}_j(t)v(s,t)\,\mathrm{d}s\,\mathrm{d}t\Big\}^4$$

$$+\sum_{j=n+1}^{n+m}\sum_{i=1}^{n}\mathrm{E}\Big\{\iint \mathbf{X}_i(s)^{\mathrm{T}}\mathbf{Y}_{j-n}(t)v(s,t)\,\mathrm{d}s\,\mathrm{d}t\Big\}^4$$

$$+\sum_{j=n+1}^{n+m}\sum_{i=n+1}^{j-1}\mathrm{E}\Big\{\iint \mathbf{Y}_{i-n}(s)^{\mathrm{T}}\mathbf{Y}_{j-n}(t)v(s,t)\,\mathrm{d}s\,\mathrm{d}t\Big\}^4$$

$$=P_1+P_2+P_3,$$

where

$$P_1=\sum_{j=2}^{n}\sum_{i=1}^{j-1}\mathrm{E}\Big\{\iint \mathbf{X}_i(s)^{\mathrm{T}}\mathbf{X}_j(t)v(s,t)\,\mathrm{d}s\,\mathrm{d}t\Big\}^4,$$

$$P_2=\sum_{j=n+1}^{n+m}\sum_{i=1}^{n}\mathrm{E}\Big\{\iint \mathbf{X}_i(s)^{\mathrm{T}}\mathbf{Y}_{j-n}(t)v(s,t)\,\mathrm{d}s\,\mathrm{d}t\Big\}^4,$$

$$P_3=\sum_{j=n+1}^{n+m}\sum_{i=n+1}^{j-1}\mathrm{E}\Big\{\iint \mathbf{Y}_{i-n}(s)^{\mathrm{T}}\mathbf{Y}_{j-n}(t)v(s,t)\,\mathrm{d}s\,\mathrm{d}t\Big\}^4.$$

We first consider the term P_2. With H and $A_2=\mathcal{V}\mathcal{G}_2\mathcal{V}^*$ defined above, conditioning first on $\mathbf{X}_i$, the integrated moment inequality in Condition (C1)(3), applied first to truncations of the latent coefficients and then extended by continuity in the covariance norm, gives

$$\mathrm{E}\{H(\mathbf{X}_i,\mathbf{Y}_{j-n})^4\mid \mathbf{X}_i\}\le C\langle \mathbf{X}_i, A_2\mathbf{X}_i\rangle_{L_2^p}^2.$$

Since A_2 is positive and trace class, let $A_{2,L} = \sum_{\ell=1}^{L} \lambda_\ell e_\ell \otimes e_\ell$ be its finite-rank spectral approximation. Condition (C1)(3) and the Cauchy–Schwarz inequality give

$$\mathrm{E}\langle \mathbf{X}_i, A_{2,L}\mathbf{X}_i\rangle_{L_2^p}^2 \leq C\left\{\sum_{\ell=1}^{L} \lambda_\ell \langle e_\ell, \mathcal{G}_1 e_\ell\rangle_{L_2^p}\right\}^2 = C\{\mathrm{tr}(A_{2,L}\mathcal{G}_1)\}^2.$$

Letting $L \to \infty$ by monotone convergence, we obtain

$$\mathrm{E}\{H(\mathbf{X}_i, \mathbf{Y}_{j-n})^4\} \leq C\{\mathrm{tr}(A_2\mathcal{G}_1)\}^2 = CT_{12}^2 \leq C\mathcal{T}^2.$$

Here $\mathcal{G}_1$ and $\mathcal{G}_2$ are the covariance operators induced by the full covariance kernels $\mathbf{G}_1$ and $\mathbf{G}_2$, so the temporal covariance factors c_1 and c_2 in Condition (C1)(2) are included. Since P_2 contains $nm = O(N^2)$ summands under Condition (C2),

$$P_2 \leq Cnm\, T_{12}^2 = O(N^2\mathcal{T}^2),$$

$$N^{-8}P_2 = O(N^{-6}\mathcal{T}^2) = o\{\sigma_{nm}^4(\mathbf{X}, \mathbf{Y})\},$$

where $N = n + m$ and $\sigma_{nm}^4(\mathbf{X}, \mathbf{Y}) \asymp N^{-4}\mathcal{T}^2$. Similarly, we have

$$O((n+m)^{-8})P_1 = o(\sigma_{nm}^4(\mathbf{X}, \mathbf{Y})), \quad O((n+m)^{-8})P_3 = o(\sigma_{nm}^4(\mathbf{X}, \mathbf{Y})).$$

By Markov's inequality and (S3.21), we have

$$\sum_{j=2}^{n+m} \mathrm{E}\{V_j^4 \mid \mathcal{F}_{j-1}\} = o_p(\sigma_{nm}^4(\mathbf{X}, \mathbf{Y})).$$

Therefore, we have

$$\sum_{j=2}^{n+m} \mathrm{E}\{V_j^2 I(|V_j| > \epsilon\sigma_{nm}(\mathbf{X}, \mathbf{Y})) \mid \mathcal{F}_{j-1}\}/\sigma_{nm}^2(\mathbf{X}, \mathbf{Y}) \xrightarrow{P} 0.$$

References

Hall, P. and C. C. Heyde (1980). *Martingale Limit Theory and its Application*. Probability and Mathematical Statistics. New York: Academic Press.

Li, B., M. K. Kim, and N. Altman (2010). On dimension folding of matrix- or array-valued statistical objects. *The Annals of Statistics 38*(2), 1094 – 1121.