Supplemental File: Second–Order Comparison of Gaussian Random Functions and the Geometry of DNA Minicircles

This supplementary note contains additional plots and tables in Section 1. In addition, Section 2 contains a more detailed study of the problem of comparing the complete spectrum, extending the discussion in the last part of Section 3.2 in the main body of the paper.

1 Supplementary Figures and Tables

This section contains figures and a table not presented in the main body of the paper. The first two figures contain plots of the projected aligned curves onto their principal axes of inertia, including their superimposition. The third figure contains scree plots with respect to the mixed eigenbasis for the two groups separately, as well as jointly. The last figure depicts the Normal QQ plots of the Karhunen-Loève residuals, as described in the discussion section of the paper.

Finally, a complete table containing the results of the simulations for level and power corresponding to Section 4 is also given. In addition to the main test statistic proposed in the paper, the complete table also presents simulations for the diagonal form of the statistic (which compares only the eigenvalues). It is observed that when the difference lies only in the eigenvalues, this test statistic performs more powerfully, as would be expected. However, in the cases where differences also lie in the eigenfunctions, it is outperformed by the full version of the test statistic.

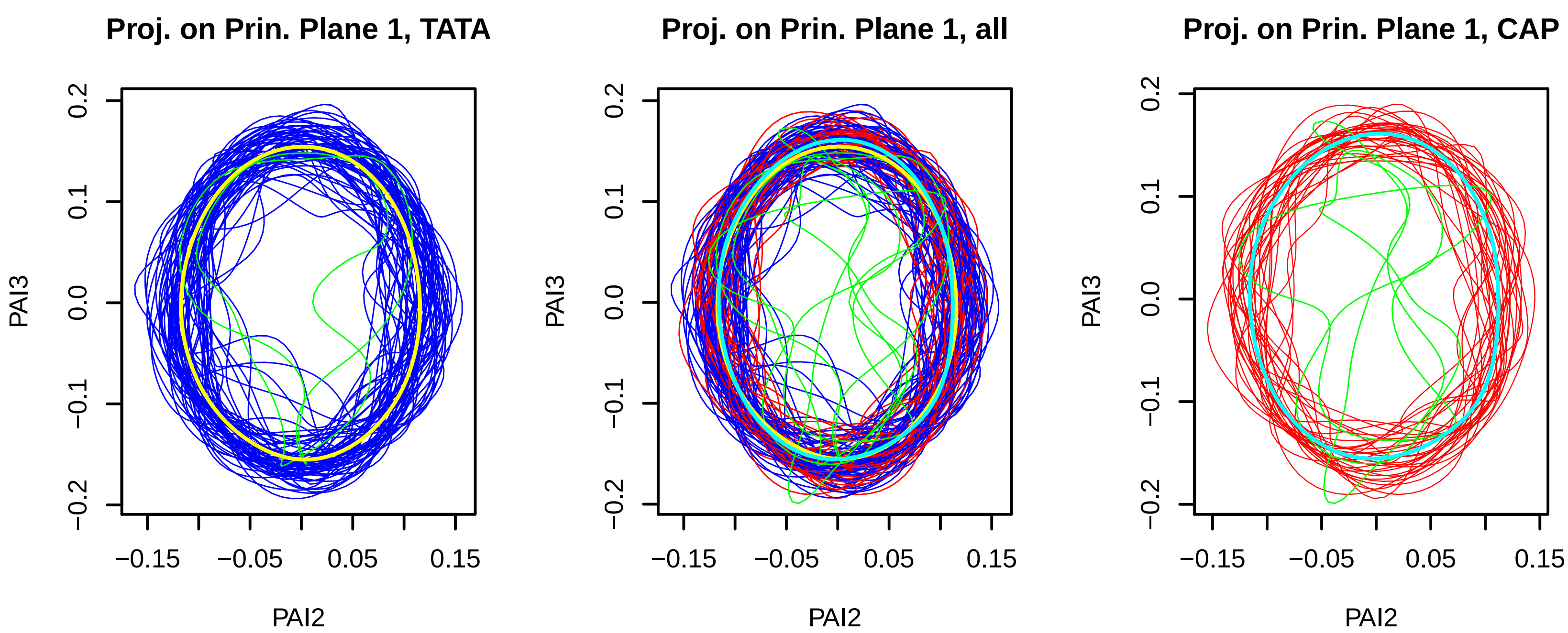

Figure 1: Projection of DNA curves on the first principal plane. Five removed outlying observations plotted in green. Mean curves (yellow and cyan) computed without outlying observations.

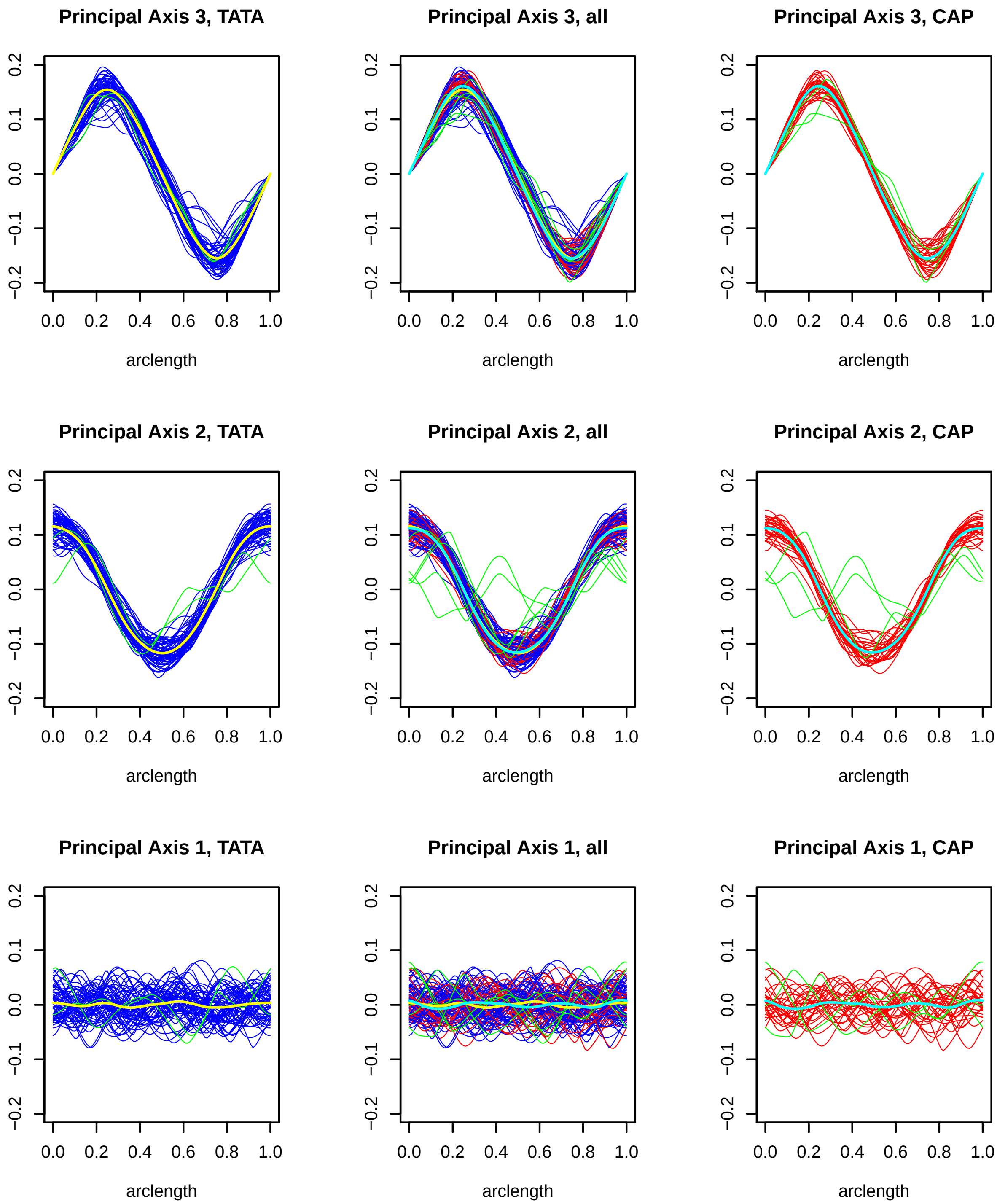

Figure 2: Coordinates of DNA curves on the principal axes of inertia. Five removed outlying observations plotted in green. Mean curves (yellow and cyan) computed without outlying observations.

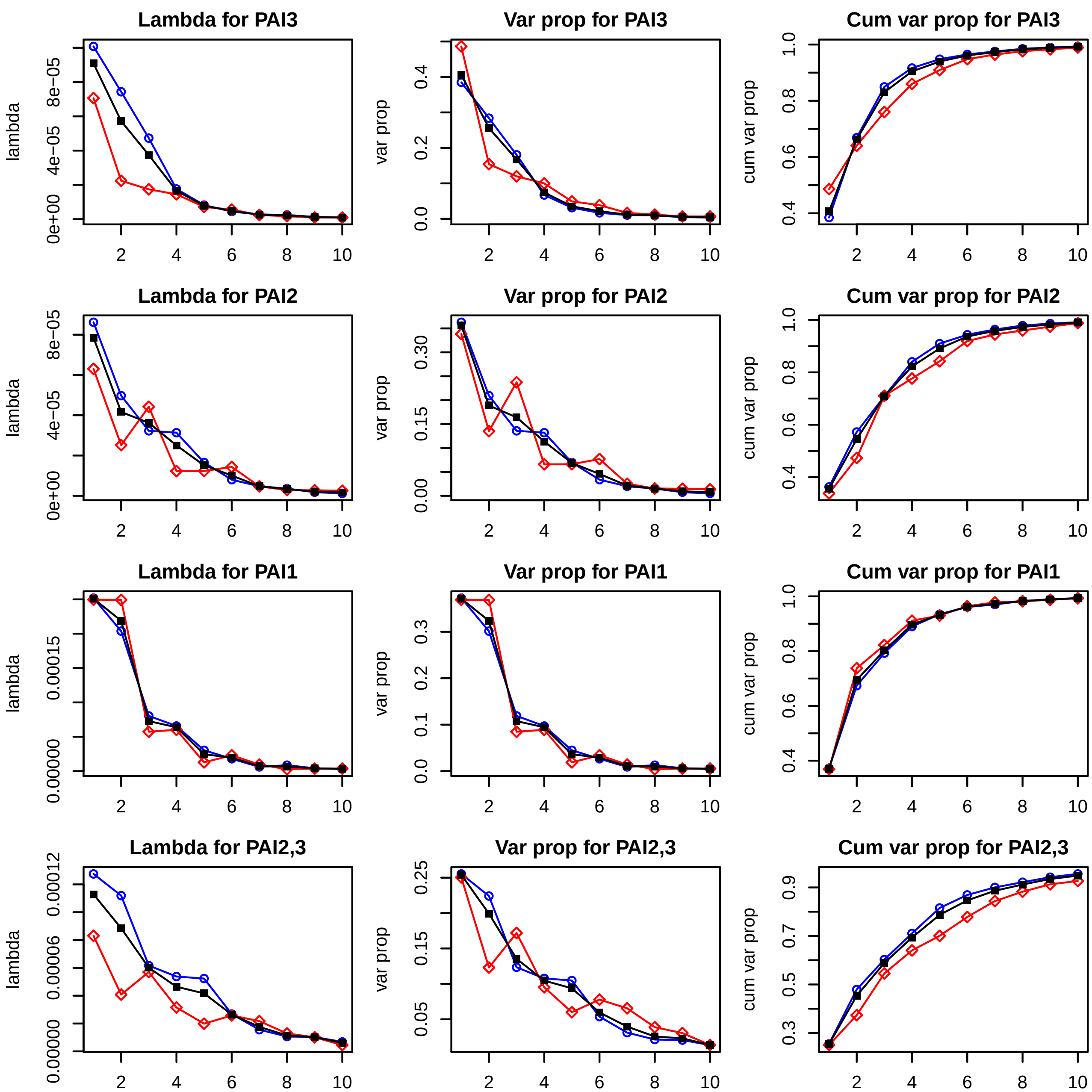

Figure 3: Empirical variances (scree plot), proportions and cumulative proportions of variance explained by components for the TATA (blue lines with circles) and CAP (red with diamonds) group and for both groups together (black with squares).

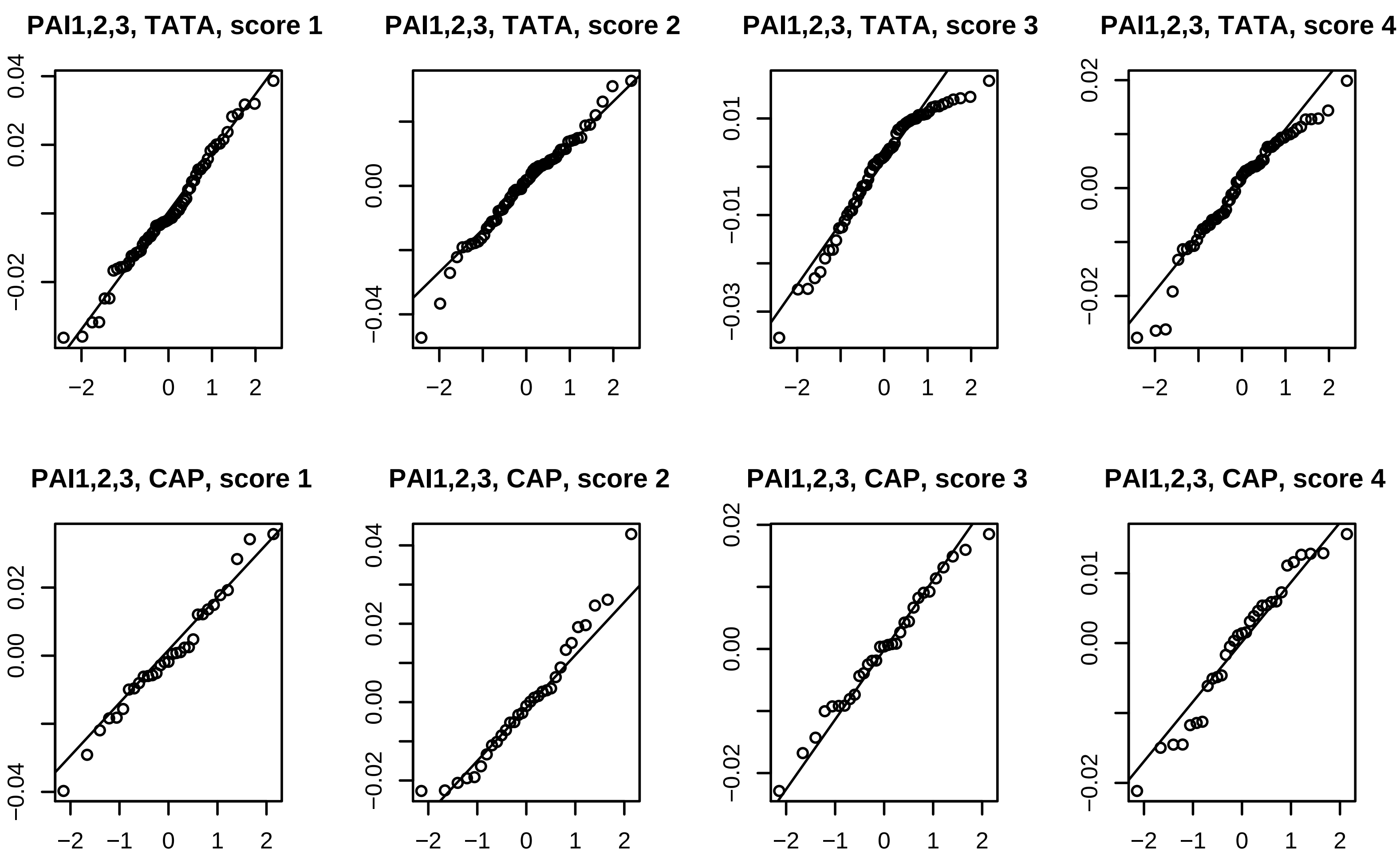

Figure 4: QQ plots corresponding to the centred Fourier coefficients when projecting onto the first four empirical eigenfunctions for each sample of curves, respectively. The exact distribution of these quantities will not be Gaussian, even if the processes are Gaussian. However, asymptotically, their distribution will be Gaussian. There do not appear systematic deviations, except for the plot corresponding to the third Fourier coefficient in the TATA group, which seems to suggest lighter upper tails as compared to the Gaussian.

Table 1: Empirical rejection probabilities on the nominal level 5 %, sample size $n_1 = n_2 = 50$, number of replications 5000 for A, 1000 for B–I. Here, $\boldsymbol{u}^X = (\boldsymbol{v}^X, \boldsymbol{w}^X)$ (resp. $\boldsymbol{u}^Y$) and K^* is the automatic truncation choice given by the penalised fit criterion.

			K				
	Parameters	Test	1	2	3	4	K^*
A	$\boldsymbol{u}^X = (12, 7, 0.5, 9, 5, 0.3)$	T	0.045	0.049	0.044	0.044	0.047
	$\boldsymbol{u}^Y = (12, 7, 0.5, 9, 5, 0.3)$	T^*	0.051	0.056	0.057	0.056	0.059
		T_1	0.045	0.046	0.045	0.047	0.047
		T_1^*	0.051	0.054	0.056	0.061	0.061
B	$\boldsymbol{u}^X = (14, 7, 0.5, 6, 5, 0.3)$	T	0.422	0.264	0.185	0.150	0.148
	$\boldsymbol{u}^Y = (8, 7, 0.5, 6, 5, 0.3)$	T^*	0.443	0.315	0.223	0.174	0.175
		T_1	0.422	0.317	0.265	0.219	0.222
		T_1^*	0.443	0.350	0.306	0.267	0.267
C	$\boldsymbol{u}^X = (15, 10, 0.5, 4, 3, 0.3)$	T	0.186	0.331	0.218	0.169	0.167
	$\boldsymbol{u}^Y = (11, 6, 0.5, 4, 3, 0.3)$	T^*	0.201	0.366	0.269	0.207	0.208
		T_1	0.186	0.380	0.312	0.279	0.273
		T_1^*	0.201	0.420	0.358	0.317	0.314
D	$\boldsymbol{u}^X = (12, 7, 0.5, 9, 3, 0.3)$	T	0.040	0.204	0.836	0.973	0.962
	$\boldsymbol{u}^Y = (12, 7, 0.5, 2, 5, 0.3)$	T^*	0.047	0.221	0.848	0.984	0.980
		T_1	0.040	0.202	0.766	0.803	0.799
		T_1^*	0.047	0.217	0.783	0.822	0.820
E	$\boldsymbol{u}^X = (12, 7, 0.5, 9, 3, 0.3)$	T	0.047	0.246	0.644	0.964	0.962
	$\boldsymbol{u}^Y = (12, 7, 0.5, 3, 9, 0.3)$	T^*	0.055	0.267	0.686	0.976	0.975
		T_1	0.047	0.227	0.477	0.597	0.594
		T_1^*	0.055	0.250	0.509	0.620	0.617
F	$\boldsymbol{u}^X = \boldsymbol{u}^Y = (12, 7, 4, 0.5, 0.3, 0.1)$	T	0.257	0.693	0.909	1.000	1.000
	$\boldsymbol{\delta}^X = (0.15, 0.15, 0.15)$	T^*	0.273	0.706	0.916	1.000	1.000
		T_1	0.257	0.474	0.521	0.567	0.637
		T_1^*	0.273	0.496	0.544	0.594	0.655
G	$\boldsymbol{u}^X = (12, 7, 0.5, 8, 6, 0.3)$	T	0.042	0.040	0.054	1.000	1.000
	$\boldsymbol{u}^Y = (12, 7, 0.5, 8, 0, 0.3)$	T^*	0.047	0.048	0.068	1.000	1.000
		T_1	0.042	0.047	0.051	1.000	1.000
		T_1^*	0.047	0.061	0.062	1.000	1.000
H	$\boldsymbol{u}^X = (12, 7, 0.5, 9, 5, 0.3)$	T	0.044	0.140	0.500	1.000	1.000
	$\boldsymbol{u}^Y = (12, 7, 0.5, 0, 5, 0.3)$	T^*	0.049	0.154	0.520	1.000	1.000
		T_1	0.044	0.139	0.478	0.992	0.992
		T_1^*	0.049	0.155	0.497	0.993	0.993
I	Brownian motion versus	T	0.719	0.608	0.483	0.377	0.493
	Ornstein–Uhlenbeck process	T^*	0.731	0.644	0.532	0.443	0.546
		T_1	0.719	0.627	0.547	0.476	0.551
		T_1^*	0.731	0.666	0.596	0.542	0.595

2 Comparing the Full Spectrum

The test procedure developed in the paper employs an optimal finite dimensional reduction in order to regularise the problem of testing. This is motivated by a Parseval decomposition of the Hilbert-Schmidt distance between the two operators,

$$\|\mathscr{R}_X - \mathscr{R}_Y\|_{HS}^2 = \sum_{k=1}^{K} \| (\mathscr{R}_X - \mathscr{R}_Y) \boldsymbol{\varphi}_{XY}^k \|_{\mathcal{L}^2}^2 + \epsilon,$$

where ϵ can be made arbitrarily small by appropriate choice of K. By making such a choice, the statistic will be (eventually) able to detect departures from the null hypothesis unless one operator is contained within a ball of small radius centred at the other operator; in this latter case, the test will still be able to detect the difference (eventually), except if this small difference lies completely at the high frequency end of the spectrum (in which case, for all practical purposes, the difference is irrelevant).

We are willing to tolerate this small level of "bias", in order to control the overall type II error of the problem. Comparison of the higher order terms of the operator spectrum on the basis of a finite sample is an ill-defined estimation problem: the fast decay of the spectrum means that we are attempting to compare extremely small quantities that have variance roughly proportional to their magnitude. In addition, the estimators of higher order eigenfunction will be characterised by very large integrated mean squared errors (available bounds grow for fixed N depending inversely on the rate of decay of the spectrum). Therefore, by trying to increase K in order to eliminate the small type II error introduced by the truncation, we are in effect causing an overall blow-up of the type II error.

If one nevertheless wishes to compare even the finest differences in the spectrum, then one needs to let K grow to infinity along with N, $K = K_N$ and modify the test statistic so as to obtain a Gaussian limit. Regularisation now manifests itself by the imposition of an allowed rate of growth of K_N. That is, a rate of growth of K relative to N that does not allow overwhelming instabilities due to the growing K. As one might expect, this growth will depend inversely on the rate of decay of the true eigenvalues (a lot of data is required to compare the finest details of the two procsses). Inevitably, in fact, this rate will be rather slow due to the following:

(a) Although the truncation level will grow as K_N, the number of terms being compared is K_N^2.

(b) While these K^2 summation terms do become independent as N grows (allowing for a CLT phenomenon) no mixing concept applies. In effect, this means that one has to look at the convergence in distribution to independence of a random vector of increasing dimension ($= K_N^2$). For any fixed dimension, the weak convergence will be at a rate of $N^{-1/2}$. Therefore, if one wishes to use L^p norms in order to use the Hilbert structure of the problem, K_N must grow slow enough to allow the $N^{-1/2}$ rate to compensate for the K_N^2 rate of increase in dimension.

(c) This required "global convergence" to independence is regulated by the convergence of the empirical eigenfunctions to the true ones; this in turn depends on the spacings between the true eigenvalues: the rate of convergence of the Kth empirical eigenfunction behaves like $N^{-1/2}\lambda_K^{-1}$. Therefore, when we let K grow, it has to be at rate slow enough, to allow $N^{-1/2}$ to annihilate the blow-up of the inverse eigenvalues.

The above heuristics are made precise in the proof of the next theorem, which provides a sufficient *regularisation rate* for asymptotically comparing the whole spectrum of infinite rank processes.

Theorem 1. *Let $\{\boldsymbol{X}_n\}_{n=1}^{n_1}$ and $\{\mathbf{Y}_n\}_{n=1}^{n_2}$ be two collections of zero mean iid continuous Gaussian random functions indexed by the interval $[0,1]$ and taking values in $\mathbb{R}^d$, possessing covariance operators $\mathscr{R}_X$ and $\mathscr{R}_Y$. Suppose that both operators are of infinite rank and have distinct eigenvalues. Let $\widehat{\mathscr{R}}_X^{n_1}$ and $\widehat{\mathscr{R}}_Y^{n_2}$ denote the empirical covariance operators based on*

$\{\boldsymbol{X}_n\}_{n=1}^{n_1}$ and $\{\boldsymbol{Y}_n\}_{n=1}^{n_2}$. For $N = n_1 + n_2$, let $\widehat{\mathscr{R}}_{XY}^{N}$ denote the empirical covariance operator of the pooled collection, and $\{\hat{\boldsymbol{\varphi}}_{XY}^{k,N}\}_{k=1}^{N}$ the corresponding eigenfunctions. Finally, let $\hat{\lambda}_{X,XY}^{k,n_1}$, $\hat{\lambda}_{Y,XY}^{k,n_2}$ denote the empirical variance of the kth Fourier coefficient of $\{\boldsymbol{X}_n\}_{n=1}^{n_1}$ and $\{\boldsymbol{Y}_n\}_{n=1}^{n_2}$, respectively, with respect to the eigenfunctions $\{\hat{\boldsymbol{\varphi}}_{XY}^{n,K}\}_{n=1}^{N}$. Assuming that $\mathbb{E}[\|\boldsymbol{X}_1\|_{L^2}^4] < \infty$, $\mathbb{E}[\|\boldsymbol{Y}_1\|_{L^2}^4] < \infty$, and $n_1/N \to \theta \in (0,1)$ as $N = n_1 + n_2 \to \infty$, it follows that, under the hypothesis $H_0 : \mathscr{R}_X = \mathscr{R}_Y$,

$$S_N := \frac{n_1 n_2}{2N\sqrt{K_N(K_N+1)/2}} \sum_{i=1}^{K_N}\sum_{j=1}^{K_N} \left\langle (\widehat{\mathscr{R}}_X^{n_1} - \widehat{\mathscr{R}}_Y^{n_2})\check{\boldsymbol{\varphi}}_{XY}^{i,N}, \check{\boldsymbol{\varphi}}_{XY}^{j,N} \right\rangle^2 - \sqrt{\frac{K_N(K_N+1)}{2}} \xrightarrow{w} \mathcal{N}(0,1),$$

as $N \to \infty$, for any $K_N \uparrow \infty$ such that $K_N^7 \lambda_{3K_N/2}^{-3/2} = o(\sqrt{N})$, where

$$\check{\boldsymbol{\varphi}}_{XY}^{k,N} = \frac{\hat{\boldsymbol{\varphi}}_{XY}^{k,N}}{\sqrt{\frac{n_1}{N}\hat{\lambda}_{X,XY}^{k,n_1} + \frac{n_2}{N}\hat{\lambda}_{Y,XY}^{k,n_2}}}.$$

Proof of Theorem 2. Let $\{Z_{Nk}\}$ denote the triangular array of random variables defined as

$$Z_{Nk} := \frac{1}{\sqrt{K_N(K_N+1)/2}} \left(\frac{n_1 n_2}{N} \left\langle (\widehat{\mathscr{R}}_X^{n_1} - \widehat{\mathscr{R}}_Y^{n_2})\check{\boldsymbol{\varphi}}_{XY}^{i(k),N}, \check{\boldsymbol{\varphi}}_{XY}^{j(k),N} \right\rangle^2 - 1 \right), \qquad i(k) \neq j(k)$$

and

$$Z_{Nk} := \frac{1}{\sqrt{K_N(K_N+1)/2}} \left(\frac{n_1 n_2}{2N} \left\langle (\widehat{\mathscr{R}}_X^{n_1} - \widehat{\mathscr{R}}_Y^{n_2})\check{\boldsymbol{\varphi}}_{XY}^{i(k),N}, \check{\boldsymbol{\varphi}}_{XY}^{i(k),N} \right\rangle^2 - 1 \right), \qquad \text{otherwise},$$

where $(i(k), j(k))$ is the the kth element of the index array $\{(i,j) : i \leq j \leq K_N\}$, when enumerating row-wise. Clearly, for $\kappa_N = K_N(K_N+1)/2$,

$$S_N = \sum_{k=1}^{\kappa_N} Z_{Nk}.$$

Write $\mathscr{Z}_N := (n_1 n_2/N)^{1/2}(\widehat{\mathscr{R}}_X^{n_1} - \widehat{\mathscr{R}}_Y^{n_2})$ and define

$$\widetilde{Z}_{Nk} := \sqrt{\frac{n_1 n_2}{N}} \left\langle (\widehat{\mathscr{R}}_X^{n_1} - \widehat{\mathscr{R}}_Y^{n_2})\mathrm{sgn}[\langle \check{\boldsymbol{\varphi}}_{XY}^{i(k),N}, \check{\boldsymbol{\varphi}}_{i(k)}\rangle]\check{\boldsymbol{\varphi}}_{XY}^{i(k),N}, \mathrm{sgn}[\langle \check{\boldsymbol{\varphi}}_{XY}^{j(k),N}, \check{\boldsymbol{\varphi}}_{j(k)}\rangle]\check{\boldsymbol{\varphi}}_{XY}^{j(k),N} \right\rangle, \quad i(k) \neq j(k)$$

and

$$\widetilde{Z}_{Nk} := \sqrt{\frac{n_1 n_2}{2N}} \left\langle (\widehat{\mathscr{R}}_X^{n_1} - \widehat{\mathscr{R}}_Y^{n_2})\mathrm{sgn}[\langle \check{\boldsymbol{\varphi}}_{XY}^{i(k),N}, \check{\boldsymbol{\varphi}}_{i(k)}\rangle]\check{\boldsymbol{\varphi}}_{XY}^{i(k),N}, \mathrm{sgn}[\langle \check{\boldsymbol{\varphi}}_{XY}^{i(k),N}, \check{\boldsymbol{\varphi}}_{i(k)}\rangle]\check{\boldsymbol{\varphi}}_{XY}^{i(k),N} \right\rangle, \quad \text{otherwise},$$

where we use the notation $\check{\boldsymbol{\varphi}}_k := \lambda_k^{-\frac{1}{2}}\boldsymbol{\varphi}_k$. The corresponding natural filtration is denoted by $\mathcal{F}_{N,k} := \sigma(\widetilde{Z}_{Nm}; 1 \leq m \leq k)$, and notice that $\{Z_{Nk}\}$ is also adapted to the filtration $\{\mathcal{F}_{N,k}\}$. Finally, we will write $\boldsymbol{Z}_{Nj} := (Z_{N1}, \ldots, Z_{Nj})^\top$ (resp. $\widetilde{\boldsymbol{Z}}_{Nj}$). We will show that

(A) $\sum_{k=1}^{\kappa_N} \mathbb{E}\left[Z_{Nk}\mathbf{1}_{\{|Z_{Nk}|\leq 1\}}|\mathcal{F}_{N,k-1}\right] \xrightarrow{\mathbb{P}} 0.$

(B) $\sum_{k=1}^{\kappa_N} \mathrm{Var}\left[Z_{Nk}\mathbf{1}_{\{|Z_{Nk}|\leq 1\}}|\mathcal{F}_{N,k-1}\right] \xrightarrow{\mathbb{P}} 1.$

(C) $\sum_{k=1}^{\kappa_N} \mathbb{P}[|Z_{Nk}| > \epsilon|\mathcal{F}_{N,k-1}] \xrightarrow{\mathbb{P}} 0, \forall\, \epsilon > 0.$

The conclusion will then follow from an "almost-martingale" central limit theorem for triangular arrays, Shorack (5, Thm. 12.2). Fix some N, let $d = \kappa_N$, and let $\boldsymbol{\zeta} \sim \mathcal{N}_d(\mathbf{0}, I)$. Letting d_∞ denote the Kolmogorov metric, we obtain

$$\begin{aligned} d_\infty\left(\widetilde{\boldsymbol{Z}}_{Nd}, \boldsymbol{\zeta}\right) \leq\ & d_\infty\left(\widetilde{\boldsymbol{Z}}_{Nd}, \left\{\sqrt{\frac{n_1 n_2}{2N}}\left\langle(\widehat{\mathscr{R}}_X^{n_1} - \widehat{\mathscr{R}}_Y^{n_2})\check{\boldsymbol{\varphi}}_{i(m)}, \check{\boldsymbol{\varphi}}_{j(m)}\right\rangle\right\}_{m=1}^{d}\right) \\ & + d_\infty\left(\left\{\sqrt{\frac{n_1 n_2}{2N}}\left\langle(\widehat{\mathscr{R}}_X^{n_1} - \widehat{\mathscr{R}}_Y^{n_2})\check{\boldsymbol{\varphi}}_{i(m)}, \check{\boldsymbol{\varphi}}_{j(m)}\right\rangle\right\}_{m=1}^{d}, \boldsymbol{\zeta}\right) \end{aligned}$$

First we concentrate on the second term of the right hand side. From the proof of Theorem 1 and Pólya's theorem we know that this term converges to zero. In fact, recalling that $\widehat{\mathscr{R}}_X^{n_1} = n_1^{-1}\sum_{i=1}^{n_i} \mathscr{X}_i$ (resp. $\widehat{\mathscr{R}}_Y^{n_2}$) and that the $\boldsymbol{\varphi}_k$ are the eigenfunctions of the common covariance operator, the convergence can be seen to be due to the standard multidimensional central limit theorem. We therefore have the following Berry-Esseen upper bound (e.g. DasGupta (2, Cor. 11.1)),

$$d_\infty\left(\left\{\sqrt{\frac{n_1n_2}{2N}}\left\langle(\widehat{\mathscr{R}}_X^{n_1}-\widehat{\mathscr{R}}_Y^{n_2})\check{\boldsymbol{\varphi}}_{i(m)},\check{\boldsymbol{\varphi}}_{j(m)}\right\rangle\right\}_{m=1}^{d},\boldsymbol{\zeta}\right)\leq\frac{Cd^{\frac{1}{4}}}{\sqrt{N}}.$$

Turning our attention to the first term in our triangle inequality, and letting $\nu_{i(k)} := \mathrm{sgn}[\langle\check{\boldsymbol{\varphi}}_{XY}^{i(k),N},\check{\boldsymbol{\varphi}}_{i(k)}\rangle]$, we note that

$$\mathbb{E}\left\|\widetilde{\boldsymbol{Z}}_{Nd}-\left\{\sqrt{\frac{n_1n_2}{2N}}\left\langle(\widehat{\mathscr{R}}_X^{n_1}-\widehat{\mathscr{R}}_Y^{n_2})\check{\boldsymbol{\varphi}}_{i(m)},\check{\boldsymbol{\varphi}}_{j(m)}\right\rangle\right\}_{m=1}^{d}\right\|_1=$$

$$=\sum_{k=1}^{d}\mathbb{E}\left|\widetilde{Z}_{Nk}-\sqrt{\frac{n_1n_2}{2N}}\left\langle(\widehat{\mathscr{R}}_X^{n_1}-\widehat{\mathscr{R}}_Y^{n_2})\check{\boldsymbol{\varphi}}_{i(k)},\check{\boldsymbol{\varphi}}_{j(k)}\right\rangle\right|$$

where, for every $1\leq k\leq d$ we have

$$\begin{aligned}
&\left|\widetilde{Z}_{Nk}-\sqrt{\frac{n_1n_2}{2N}}\left\langle(\widehat{\mathscr{R}}_X^{n_1}-\widehat{\mathscr{R}}_Y^{n_2})\check{\boldsymbol{\varphi}}_{i(k)},\check{\boldsymbol{\varphi}}_{j(k)}\right\rangle\right|\\
&\qquad=\left|\left\langle\mathscr{Z}_N\nu_{i(k)}\check{\boldsymbol{\varphi}}_{XY}^{i(k),N},\nu_{j(k)}\check{\boldsymbol{\varphi}}_{XY}^{j(k),N}\right\rangle-\left\langle\mathscr{Z}_N\check{\boldsymbol{\varphi}}_{i(k)},\check{\boldsymbol{\varphi}}_{j(k)}\right\rangle\right|\\
&=\left|\left\langle\mathscr{Z}_N\nu_{i(k)}\check{\boldsymbol{\varphi}}_{XY}^{i(k),N},\nu_{j(k)}\check{\boldsymbol{\varphi}}_{XY}^{j(k),N}\right\rangle-\langle\mathscr{Z}_N\nu_{i(k)}\check{\boldsymbol{\varphi}}_{XY}^{i(k),N},\check{\boldsymbol{\varphi}}_{j(k)}\rangle+\langle\mathscr{Z}_N\nu_{i(k)}\check{\boldsymbol{\varphi}}_{XY}^{i(k),N},\check{\boldsymbol{\varphi}}_{j(k)}\rangle-\left\langle\mathscr{Z}_N\check{\boldsymbol{\varphi}}_{i(k)},\check{\boldsymbol{\varphi}}_{j(k)}\right\rangle\right|\\
&=\left|\left\langle\mathscr{Z}_N\nu_{i(k)}\check{\boldsymbol{\varphi}}_{XY}^{i(k),N},\nu_{j(k)}\check{\boldsymbol{\varphi}}_{XY}^{j(k),N}-\check{\boldsymbol{\varphi}}_{j(k)}\right\rangle+\left\langle\mathscr{Z}_N\left(\nu_{i(k)}\check{\boldsymbol{\varphi}}_{XY}^{i(k),N}-\check{\boldsymbol{\varphi}}_{i(k)}\right),\check{\boldsymbol{\varphi}}_{j(k)}\right\rangle\right|\\
&=\left|\left\langle\mathscr{Z}_N\nu_{i(k)}\check{\boldsymbol{\varphi}}_{XY}^{i(k),N},\nu_{j(k)}\check{\boldsymbol{\varphi}}_{XY}^{j(k),N}-\check{\boldsymbol{\varphi}}_{j(k)}\right\rangle+\left\langle\mathscr{Z}_N\check{\boldsymbol{\varphi}}_{j(k)},\nu_{i(k)}\check{\boldsymbol{\varphi}}_{XY}^{i(k),N}-\check{\boldsymbol{\varphi}}_{i(k)}\right\rangle\right|\\
&\leq\left\|\mathscr{Z}_N\nu_{i(k)}\check{\boldsymbol{\varphi}}_{XY}^{i(k),N}\right\|_{\mathcal{L}^2}\left\|\nu_{j(k)}\check{\boldsymbol{\varphi}}_{XY}^{j(k),N}-\check{\boldsymbol{\varphi}}_{j(k)}\right\|_{\mathcal{L}^2}+\left\|\mathscr{Z}_N\check{\boldsymbol{\varphi}}_{j(k)}\right\|_{\mathcal{L}^2}\left\|\nu_{i(k)}\check{\boldsymbol{\varphi}}_{XY}^{i(k),N}-\check{\boldsymbol{\varphi}}_{i(k)}\right\|_{\mathcal{L}^2}\\
&\leq\|\mathscr{Z}_N\|_{HS}\left\|\nu_{i(k)}\check{\boldsymbol{\varphi}}_{XY}^{i(k),N}\right\|_{\mathcal{L}^2}\left\|\nu_{j(k)}\check{\boldsymbol{\varphi}}_{XY}^{j(k),N}-\check{\boldsymbol{\varphi}}_{j(k)}\right\|_{\mathcal{L}^2}+\|\mathscr{Z}_N\|_{HS}\left\|\check{\boldsymbol{\varphi}}_{j(k)}\right\|_{\mathcal{L}^2}\left\|\nu_{i(k)}\check{\boldsymbol{\varphi}}_{XY}^{i(k),N}-\check{\boldsymbol{\varphi}}_{i(k)}\right\|_{\mathcal{L}^2}\\
&=\|\mathscr{Z}_N\|_{HS}\left(\left\|\nu_{j(k)}\check{\boldsymbol{\varphi}}_{XY}^{j(k),N}-\check{\boldsymbol{\varphi}}_{j(k)}\right\|_{\mathcal{L}^2}+\left\|\nu_{i(k)}\check{\boldsymbol{\varphi}}_{XY}^{i(k),N}-\check{\boldsymbol{\varphi}}_{i(k)}\right\|_{\mathcal{L}^2}\right)
\end{aligned}$$

Here we have used the Cauchy-Schwartz inequality and the fact that $\mathscr{Z}_N$ is a bounded operator. By the triangle inequality we now obtain

$$
\begin{aligned}
&\|\mathscr{Z}_N\|_{HS}\left(\left\|\nu_{j(k)}\check{\boldsymbol{\varphi}}_{XY}^{j(k),N}-\check{\boldsymbol{\varphi}}_{j(k)}\right\|_{\mathcal{L}^2}+\left\|\nu_{i(k)}\check{\boldsymbol{\varphi}}_{XY}^{i(k),N}-\check{\boldsymbol{\varphi}}_{i(k)}\right\|_{\mathcal{L}^2}\right)\\
&\quad\le\|\mathscr{Z}_N\|_{HS}\left(\left\|\nu_{j(k)}\check{\boldsymbol{\varphi}}_{XY}^{j(k),N}-\nu_{j(k)}\lambda_{j(k)}^{-1/2}\hat{\boldsymbol{\varphi}}_{XY}^{j(k),N}\right\|_{\mathcal{L}^2}+\left\|\nu_{j(k)}\lambda_{j(k)}^{-1/2}\hat{\boldsymbol{\varphi}}_{XY}^{j(k),N}-\check{\boldsymbol{\varphi}}_{j(k)}\right\|_{\mathcal{L}^2}\right.\\
&\qquad\qquad\left.+\left\|\nu_{i(k)}\check{\boldsymbol{\varphi}}_{XY}^{i(k),N}-\nu_{i(k)}\lambda_{i(k)}^{-1/2}\hat{\boldsymbol{\varphi}}_{XY}^{i(k),N}\right\|_{\mathcal{L}^2}+\left\|\nu_{i(k)}\lambda_{i(k)}^{-1/2}\hat{\boldsymbol{\varphi}}_{XY}^{i(k),N}-\check{\boldsymbol{\varphi}}_{i(k)}\right\|_{\mathcal{L}^2}\right)\\
&\quad=\|\mathscr{Z}_N\|_{HS}\left((\hat{\lambda}_{j(k)}^{-1/2}-\lambda_{j(k)}^{-1/2})+\lambda_{j(k)}^{-1/2}\left\|\nu_{j(k)}\hat{\boldsymbol{\varphi}}_{XY}^{j(k),N}-\boldsymbol{\varphi}_{j(k)}\right\|_{\mathcal{L}^2}\right.\\
&\qquad\qquad\left.+(\hat{\lambda}_{i(k)}^{-1/2}-\lambda_{i(k)}^{-1/2})+\lambda_{i(k)}^{-1/2}\left\|\nu_{i(k)}\hat{\boldsymbol{\varphi}}_{XY}^{i(k),N}-\boldsymbol{\varphi}_{i(k)}\right\|_{\mathcal{L}^2}\right)
\end{aligned}
$$

where we have used the simplified notation

$$
\hat{\lambda}_{i(k)}=\sqrt{\frac{n_1}{N}\hat{\lambda}_{X,XY}^{i(k),n_1}+\frac{n_2}{N}\hat{\lambda}_{Y,XY}^{i(k),n_2}}.
$$

We now apply the inequality given in Bosq (1, Lem. 4.3) and obtain

$$
\begin{aligned}
&\|\mathscr{Z}_N\|_{HS}\left((\hat{\lambda}_{j(k)}^{-1/2}-\lambda_{j(k)}^{-1/2})+(\hat{\lambda}_{i(k)}^{-1/2}-\lambda_{i(k)}^{-1/2})+\lambda_{j(k)}^{-1/2}\left\|\nu_{j(k)}\hat{\boldsymbol{\varphi}}_{XY}^{j(k),N}-\boldsymbol{\varphi}_{j(k)}\right\|_{\mathcal{L}^2}\right.\\
&\qquad\qquad\left.+\lambda_{i(k)}^{-1/2}\left\|\nu_{i(k)}\hat{\boldsymbol{\varphi}}_{XY}^{i(k),N}-\boldsymbol{\varphi}_{i(k)}\right\|_{\mathcal{L}^2}\right)\\
&\quad\le\|\mathscr{Z}_N\|_{HS}\left((\hat{\lambda}_{j(k)}^{-1/2}-\lambda_{j(k)}^{-1/2})+(\hat{\lambda}_{i(k)}^{-1/2}-\lambda_{i(k)}^{-1/2})\right.\\
&\qquad\qquad\lambda_{j(k)}^{-1/2}2\sqrt{2}\max\left\{(\lambda_{j(k)-1}-\lambda_{j(k)})^{-1},(\lambda_{j(k)}-\lambda_{j(k)+1})^{-1}\right\}\|\widehat{\mathscr{R}}_{XY}^N-\mathscr{R}_X\|_{HS}\\
&\qquad\qquad\left.+\lambda_{i(k)}^{-1/2}2\sqrt{2}\max\left\{(\lambda_{i(k)-1}-\lambda_{i(k)})^{-1},(\lambda_{i(k)}-\lambda_{i(k)+1})^{-1}\right\}\|\widehat{\mathscr{R}}_{XY}^N-\mathscr{R}_X\|_{HS}\right)
\end{aligned}
$$

Recapitulating, we have obtained

$$
\begin{aligned}
&\left|\widetilde{Z}_{Nk}-\sqrt{\frac{n_1n_2}{2N}}\left\langle(\widehat{\mathscr{R}}_X^{n_1}-\widehat{\mathscr{R}}_Y^{n_2})\check{\boldsymbol{\varphi}}_{i(k)},\check{\boldsymbol{\varphi}}_{j(k)}\right\rangle\right|\\
&\quad\le\|\mathscr{Z}_N\|_{HS}\left((\hat{\lambda}_{j(k)}^{-1/2}-\lambda_{j(k)}^{-1/2})+(\hat{\lambda}_{i(k)}^{-1/2}-\lambda_{i(k)}^{-1/2})\right.\\
&\qquad\qquad\lambda_{j(k)}^{-1/2}2\sqrt{2}\max\left\{(\lambda_{j(k)-1}-\lambda_{j(k)})^{-1},(\lambda_{j(k)}-\lambda_{j(k)+1})^{-1}\right\}\|\widehat{\mathscr{R}}_{XY}^N-\mathscr{R}_X\|_{HS}
\end{aligned}
$$

$$+ \lambda_{i(k)}^{-1/2} 2\sqrt{2} \max\left\{(\lambda_{i(k)-1} - \lambda_{i(k)})^{-1}, (\lambda_{i(k)} - \lambda_{i(k)+1})^{-1}\right\} \|\widehat{\mathscr{R}}_{XY}^N - \mathscr{R}_X\|_{HS}\Big)$$

Now we take expectations on both sides, expand the right hand side, and repeatedly apply the Cauchy-Schwartz inequality (with respect to the mean-square norm) to obtain

$$\begin{aligned}
&\mathbb{E}\left|\widetilde{Z}_{Nk} - \sqrt{\frac{n_1 n_2}{2N}} \left\langle (\widehat{\mathscr{R}}_X^{n_1} - \widehat{\mathscr{R}}_Y^{n_2})\check{\boldsymbol{\varphi}}_{i(k)}, \check{\boldsymbol{\varphi}}_{j(k)}\right\rangle\right| \\
&\quad \leq \sqrt{\mathbb{E}\,\|\mathscr{Z}_N\|_{HS}^2}\sqrt{\mathbb{E}(\hat{\lambda}_{j(k)}^{-1/2} - \lambda_{j(k)}^{-1/2})^2} + \sqrt{\mathbb{E}\,\|\mathscr{Z}_N\|_{HS}^2}\sqrt{\mathbb{E}(\hat{\lambda}_{i(k)}^{-1/2} - \lambda_{i(k)}^{-1/2})^2} \\
&+\lambda_{j(k)}^{-1/2} 2\sqrt{2} \max\left\{(\lambda_{j(k)-1} - \lambda_{j(k)})^{-1}, (\lambda_{j(k)} - \lambda_{j(k)+1})^{-1}\right\} \sqrt{\mathbb{E}\,\|\mathscr{Z}_N\|_{HS}^2}\sqrt{\mathbb{E}\|\widehat{\mathscr{R}}_{XY}^N - \mathscr{R}_X\|_{HS}^2} \\
&+\lambda_{i(k)}^{-1/2} 2\sqrt{2} \max\left\{(\lambda_{i(k)-1} - \lambda_{i(k)})^{-1}, (\lambda_{i(k)} - \lambda_{i(k)+1})^{-1}\right\} \sqrt{\mathbb{E}\,\|\mathscr{Z}_N\|_{HS}^2}\sqrt{\mathbb{E}\|\widehat{\mathscr{R}}_{XY}^N - \mathscr{R}_X\|_{HS}^2}
\end{aligned}$$

We note first that, by Minkowski's inequality, $\sqrt{\mathbb{E}\,\|\mathscr{Z}_N\|_{HS}^2}$ is bounded above for all N, by definition of the random operator $\mathscr{Z}_N$. Next, $\sqrt{\mathbb{E}(\hat{\lambda}_{i(k)}^{-1} - \lambda_{i(k)}^{-1})^2}$ and $\sqrt{\mathbb{E}(\hat{\lambda}_{i(k)}^{-1} - \lambda_{i(k)}^{-1})^2}$ are, asymptotically in N, of the order of $O(\lambda_{i(k)}^{-1/2} N^{-1/2})$ and so are also of the order of $O(\lambda_{i(d)}^{-1/2} N^{-1/2})$, when $k \leq d$. This can be seen by applying the Delta method to the CLT given in Dauxois et. al (3, Prop. 8). Finally, $\sqrt{\mathbb{E}\|\widehat{\mathscr{R}}_{XY}^N - \mathscr{R}_X\|_{HS}^2}$ is asymptotically of the order of $O(N^{-1/2})$ by the CLT in Hilbert Space (Bosq (1, Thm 2.7)).

Now by definition of $i(k)$ and $j(k)$, we have that $i(d)[i(d)+1]/2 = j(d)[j(d)+1]/2 = d$, so that it holds that

$$\lambda_{i(k)} = \lambda_{\frac{\sqrt{8d+1}-1}{2}} \geq \lambda_{\frac{3\sqrt{d}}{2}}.$$

Combining all the above, we arrive at

$$\mathbb{E}\left|\widetilde{Z}_{Nk} - \sqrt{\frac{n_1 n_2}{2N}} \left\langle (\widehat{\mathscr{R}}_X^{n_1} - \widehat{\mathscr{R}}_Y^{n_2})\check{\boldsymbol{\varphi}}_{i(k)}, \check{\boldsymbol{\varphi}}_{j(k)}\right\rangle\right| = O\left(\lambda_{3\sqrt{d}/2}^{-3/2} N^{-1/2}\right).$$

so that

$$\mathbb{E}\left\|\widetilde{\boldsymbol{Z}}_{Nd} - \left\{\sqrt{\frac{n_1 n_2}{2N}} \left\langle (\widehat{\mathscr{R}}_X^{n_1} - \widehat{\mathscr{R}}_Y^{n_2})\check{\boldsymbol{\varphi}}_{i(m)}, \check{\boldsymbol{\varphi}}_{j(m)}\right\rangle\right\}_{m=1}^{d}\right\|_1 = O\left(\lambda_{3\sqrt{d}/2}^{-3/2} N^{-1/2} d\right).$$

Letting d_W denote the L_1-Wasserstein distance between two probability measures, we have (e.g. Gibbs & Su (4)),

$$\begin{aligned}
d_\infty(G_{N,d}, H_{N,d}) &\leq (1+\|h_{N,d}\|_\infty)\sqrt{d_W(G_{N,d}, H_{N,d})} \\
&\leq (1+\|h_{N,d}\|_\infty)\sqrt{\mathbb{E}\left\|\widetilde{\boldsymbol{Z}}_{Nd} - \left\{\sqrt{\frac{n_1 n_2}{2N}}\left\langle(\widehat{\mathscr{R}}_X^{n_1} - \widehat{\mathscr{R}}_Y^{n_2})\check{\boldsymbol{\varphi}}_{i(m)}, \check{\boldsymbol{\varphi}}_{j(m)}\right\rangle\right\}_{m=1}^{d}\right\|_1} \\
&= (1+\|h_{N,d}\|_\infty) O\left(\lambda_{3\sqrt{d}/2}^{-3/4} N^{-1/4} d^{1/2}\right).
\end{aligned}$$

where $H_{N,d}$ is the distribution function of $\left\{\sqrt{\frac{n_1 n_2}{2N}}\left\langle(\widehat{\mathscr{R}}_X^{n_1} - \widehat{\mathscr{R}}_Y^{n_2})\check{\boldsymbol{\varphi}}_{i(m)}, \check{\boldsymbol{\varphi}}_{j(m)}\right\rangle\right\}_{m=1}^{d}$, $G_{N,d}$ is the distribution function of $\widetilde{\boldsymbol{Z}}_{Nd}$, and $h_{N,d}$ is the density function of $H_{N,d}$. But $h_{N,d}$ is the density of a difference of two independent random vectors, each of which is in turn the sum of n_1 and n_2 iid random vectors, respectively. Thus, letting $h_d^{[1]}$ and $h_d^{[2]}$ be the respective densities, and by symmetry, we have,

$$\begin{aligned}
\|h_{N,d}\|_\infty = \|\underbrace{h_{d,n_1}^{[1]} * \ldots * h_{d,n_1}^{[1]}}_{n_1\ times} * \underbrace{h_{d,n_2}^{[2]} * \ldots * h_{d,n_2}^{[2]}}_{n_2\ times}\|_\infty &\leq \|\underbrace{h_{d,n_1}^{[1]} * \ldots * h_{d,n_1}^{[1]}}_{n_1\ times}\|_1 \|\underbrace{h_{d,n_2}^{[2]} * \ldots * h_{d,n_2}^{[2]}}_{n_2\ times}\|_\infty \\
&= \|\underbrace{h_{d,n_2}^{[2]} * \ldots * h_{d,n_2}^{[2]}}_{n_2\ times}\|_\infty
\end{aligned}$$

Now it is immediate that

$$\|h_{d,n_2}^{[2]} * \ldots * h_{d,n_2}^{[2]}\|_\infty \leq \|h_{n_2}^{[2]} * \ldots * h_{n_2}^{[2]}\|_\infty,$$

where $h_{n_2}^{[2]}$ is the marginal density of $\sqrt{\frac{n_1 n_2}{2N}}\left\langle(\frac{1}{n_2}\mathscr{X}_1)\check{\boldsymbol{\varphi}}_{i(1)}, \check{\boldsymbol{\varphi}}_{j(1)}\right\rangle$. But it must the case that $\|h_{n_2}^{[2]} * \ldots * h_{n_2}^{[2]}\|_\infty$ be bounded above, since $\sum_{i=1}^{n_2}\sqrt{\frac{n_1 n_2}{2N}}\left\langle(\frac{1}{n_2}\mathscr{X}_i)\check{\boldsymbol{\varphi}}_{i(1)}, \check{\boldsymbol{\varphi}}_{j(1)}\right\rangle$ is a sequence of variables with diffuse laws converging weakly to a non-degenerate Gaussian.

We are thus in a position to conclude that

$$d_\infty\left(\widetilde{\boldsymbol{Z}}_{Nd}, \boldsymbol{\zeta}\right) = O\left(\lambda_{3\sqrt{d}/2}^{-3/4} N^{-1/4} d^{1/2}\right). \tag{1}$$

Now recall that, with probability one,

$$\mathbb{E}\left(Z_{Nk}\mathbf{1}_{\{|Z_{Nk}|\leq 1\}}|\mathcal{F}_{N,k-1}\right) = \int_{-\infty}^{+\infty} \frac{1}{\sqrt{\kappa_N}}\left(x^2-1\right)\mathbf{1}_{\left\{|x^2-1|\leq\sqrt{2\kappa_N}\right\}} F_{\widetilde{Z}_{Nk}|\widetilde{\boldsymbol{Z}}_{N,k-1}}(dx|\widetilde{\boldsymbol{Z}}_{N,k-1})$$

where he have used standard notation for conditional distribution functions. It follows that, given ζ a standard Gaussian random variable,

$$\begin{aligned}
&\mathbb{E}\left(Z_{Nk}\mathbf{1}_{\{|Z_{Nk}|\leq 1\}}|\mathcal{F}_{N,k-1}\right) - \mathbb{E}\left(\frac{1}{\sqrt{\kappa_N}}(\zeta^2-1)\mathbf{1}_{\left\{|\zeta^2-1|\leq\sqrt{\kappa_N}\right\}}\right)\\
&= \int_{-\infty}^{+\infty} \frac{1}{\sqrt{\kappa_N}}\left(x^2-1\right)\mathbf{1}_{\left\{|x^2-1|\leq\sqrt{2\kappa_N}\right\}} F_{\widetilde{Z}_{Nk}|\widetilde{\boldsymbol{Z}}_{N,k-1}}(dx|\widetilde{\boldsymbol{Z}}_{N,k-1})\\
&\quad - \int_{-\infty}^{+\infty} \frac{1}{\sqrt{\kappa_N}}\left(x^2-1\right)\mathbf{1}_{\left\{|x^2-1|\leq\sqrt{2\kappa_N}\right\}} F_{\zeta}(dx)\\
&\quad = \int_{-\infty}^{+\infty} \frac{1}{\sqrt{\kappa_N}}\left(x^2-1\right)\mathbf{1}_{\left\{|x^2-1|\leq\sqrt{2\kappa_N}\right\}} \left[F^{\widetilde{\boldsymbol{Z}}_{N,k-1}}_{\widetilde{Z}_{Nk}|\widetilde{\boldsymbol{Z}}_{N,k-1}} - F_\zeta\right](dx)
\end{aligned}$$

with the alternative notation $F^{\widetilde{\boldsymbol{Z}}_{N,k-1}}_{\widetilde{Z}_{Nk}|\widetilde{\boldsymbol{Z}}_{N,k-1}}(x) \equiv F_{\widetilde{Z}_{Nk}|\widetilde{\boldsymbol{Z}}_{N,k-1}}(x|\widetilde{\boldsymbol{Z}}_{N,k-1})$. From (1) we have that for $\boldsymbol{\zeta}\sim\mathcal{N}_k(0,I)$, $d_\infty(\widetilde{\boldsymbol{Z}}_{Nk},\boldsymbol{\zeta}) = O\left(\lambda^{-1/3}_{3\sqrt{d}/2}N^{-1/4}k^{1/2}\right)$, so by Lemma 1 (see below), given any $\boldsymbol{z}\in\mathbb{R}^{k-1}$,

$$\sup_{x\in\mathbb{R}}\left|F^{\boldsymbol{z}}_{\widetilde{Z}_{Nk}|\widetilde{\boldsymbol{Z}}_{N,k-1}}(x) - F_\zeta(x)\right| = O\left(\lambda^{-3/4}_{3\sqrt{d}/2}N^{-1/4}k^{1/2}\right)$$

and so given $z\in\mathbb{R}^{k-1}$

$$\int_{-\infty}^{+\infty} \frac{1}{\sqrt{\kappa_N}}\left(x^2-1\right)\mathbf{1}_{\left\{|x^2-1|\leq\sqrt{2\kappa_N}\right\}} \left[F^{\boldsymbol{z}}_{\widetilde{Z}_{Nk}|\widetilde{\boldsymbol{Z}}_{N,k-1}} - F_\zeta\right](dx) = O\left(\lambda^{-3/4}_{3\sqrt{\kappa_N}/2}N^{-1/4}k^{1/2}\kappa_N^{1/4}\right).$$

Consequently, for $\{\zeta_k\}$ an iid sequence of standard Gaussian variables, and for all $\omega\in\Omega$,

$$\sum_{k=1}^{\kappa_N}\left[\mathbb{E}\left[Z_{Nk}\mathbf{1}_{\{|Z_{Nk}|\leq 1\}}|\mathcal{F}_{N,k-1}\right] - \mathbb{E}\left[\frac{1}{\sqrt{\kappa_N}}(\zeta_k^2-1)\mathbf{1}_{\{|\zeta_k|\leq\sqrt{\kappa_N}\}}\right]\right] = O\left(\frac{\kappa_N^{7/4}}{N^{1/4}\lambda^{3/4}_{3\sqrt{\kappa_N}/2}}\right) = O\left(\frac{K_N^{7/2}}{N^{1/4}\lambda^{3/4}_{3\sqrt{\kappa_N}/2}}\right)$$

And, since

$$K_N^7 \lambda^{-3/2}_{\frac{3\sqrt{2K_N(K_N+1)}}{2}} \leq K_N^7 \lambda^{-3/2}_{\frac{3K_N}{2}} = o\left(\sqrt{N}\right),$$

it follows from our assumptions that the quantity above converges to zero almost certainly. But, on the other hand,

$$\begin{aligned}
&\left|\sum_{k=1}^{\kappa_N} \mathbb{E}\left[Z_{Nk}\mathbf{1}_{\{|Z_{Nk}|\leq 1\}}|\mathcal{F}_{N,k-1}\right]\right| \\
&\quad\leq \left|\sum_{k=1}^{\kappa_N}\left[\mathbb{E}\left[Z_{Nk}\mathbf{1}_{\{|Z_{Nk}|\leq 1\}}|\mathcal{F}_{N,k-1}\right] - \mathbb{E}\left[\frac{1}{\sqrt{\kappa_N}}(\zeta_k^2-1)\mathbf{1}_{\{|\zeta_k|\leq\sqrt{\kappa_N}\}}\right]\right]\right| \\
&\qquad + \left|\sum_{k=1}^{\kappa_N}\mathbb{E}\left[\frac{1}{\sqrt{\kappa_N}}(\zeta_k^2-1)\mathbf{1}_{\{|\zeta_k|\leq\sqrt{\kappa_N}\}}\right]\right|
\end{aligned}$$

with the last term obviously converging to zero as $N \to \infty$ so that condition (A) is fulfilled.

We now turn our attention to condition (B). By definition:

$$\sum_{k=1}^{\kappa_N} \mathrm{Var}\left[Z_{Nk}\mathbf{1}_{\{|Z_{Nk}|\leq 1\}}|\mathcal{F}_{N,k-1}\right] = \sum_{k=1}^{\kappa_N}\mathbb{E}\left[Z_{Nk}^2\mathbf{1}_{\{|Z_{Nk}|\leq 1\}}|\mathcal{F}_{N,k-1}\right] - \sum_{k=1}^{\kappa_N}\mathbb{E}^2\left[Z_{Nk}\mathbf{1}_{\{|Z_{Nk}|\leq 1\}}|\mathcal{F}_{N,k-1}\right]$$

That the second term converges to zero almost surely follows from our proof of condition (A). Hence, it suffices to concentrate on the first term. Following the same steps as with (A), we may write

$$\int_{-\infty}^{+\infty} \frac{\left(x^2-1\right)^2}{2\kappa_N}\mathbf{1}_{\left\{|x^2-1|\leq\sqrt{2\kappa_N}\right\}}\left[F^{\boldsymbol{z}}_{\widetilde{Z}_{Nk}|\widetilde{\boldsymbol{Z}}_{N,k-1}} - F_\zeta\right](dx) = O\left(\frac{K_N^{3/2}}{N^{1/4}\lambda^{3/4}_{3\sqrt{\kappa_N}/2}}\right)$$

This in turn imples that, with probability one,

$$\sum_{k=1}^{\kappa_N}\left[\mathbb{E}\left[Z_{Nk}\mathbf{1}_{\{|Z_{Nk}|\leq 1\}}|\mathcal{F}_{N,k-1}\right] - \mathbb{E}\left[\frac{1}{\sqrt{\kappa_N}}(\zeta_k^2-1)\mathbf{1}_{\{|\zeta_k|\leq\sqrt{\kappa_N}\}}\right]\right] \overset{N\to\infty}{\longrightarrow} 0.$$

Finally, we see that

$$\begin{aligned}
&\sum_{k=1}^{\kappa_N} \mathbb{E}\left[Z_{Nk}^2 \mathbf{1}_{\{|Z_{Nk}|\le 1\}} | \mathcal{F}_{N,k-1}\right] \\
&\quad = \sum_{k=1}^{\kappa_N} \left[\mathbb{E}\left[Z_{Nk}^2 \mathbf{1}_{\{|Z_{Nk}|\le 1\}} | \mathcal{F}_{N,k-1}\right] - \mathbb{E}\left[\frac{1}{2\kappa_N}(\zeta_k^2-1)^2 \mathbf{1}_{\{|\zeta_k|\le\sqrt{\kappa_N}\}}\right]\right] \\
&\qquad + \sum_{k=1}^{\kappa_N} \mathbb{E}\left[\frac{1}{2\kappa_N}(\zeta_k^2-1)^2 \mathbf{1}_{\{|\zeta_k|\le\sqrt{\kappa_N}\}}\right]
\end{aligned}$$

with the last term clearly converging to 1 almost certainly. This establishes condition (B).

Finally, we concentrate on condition (C). By definition,

$$\begin{aligned}
\mathbb{P}[|Z_{Nk}| > \epsilon | \mathcal{F}_{N,k-1}] &= 1 - \mathbb{E}\left[\mathbf{1}\left\{|Z_{Nk}| < \epsilon\right\} | \mathcal{F}_{N,k-1}\right] \\
&= 1 + \left(\mathbb{E}\left[\mathbf{1}\left\{|\zeta^2 - 1| < \epsilon\sqrt{\kappa_N}\right\}\right] - \mathbb{E}\left[\mathbf{1}\left\{|Z_{Nk}| < \epsilon\right\} | \mathcal{F}_{N,k-1}\right]\right) \\
&\quad - \mathbb{E}\left[\mathbf{1}\left\{|\zeta^2 - 1| < \epsilon\sqrt{\kappa_N}\right\}\right] \\
&= \left(\mathbb{E}\left[\mathbf{1}\left\{|\zeta^2 - 1| < \epsilon\sqrt{\kappa_N}\right\}\right] - \mathbb{E}\left[\mathbf{1}\left\{|Z_{Nk}| < \epsilon\right\} | \mathcal{F}_{N,k-1}\right]\right) + \mathbb{P}[|\zeta^2 - 1| > \epsilon\sqrt{\kappa_N}]
\end{aligned}$$

It is clear from our analysis of (A) and (B) that

$$\sum_{k=1}^{\kappa_N} \left(\mathbb{E}\left[\mathbf{1}\left\{|\zeta^2 - 1| < \epsilon\sqrt{\kappa_N}\right\}\right] - \mathbb{E}\left[\mathbf{1}\left\{|Z_{Nk}| < \epsilon\right\} | \mathcal{F}_{N,k-1}\right]\right) \xrightarrow{a.s.} 0.$$

Finally, we have

$$\sum_{k=1}^{\kappa_N} \mathbb{P}[|\zeta^2 - 1| > \epsilon\sqrt{\kappa_N}] = \kappa_N \mathbb{P}[|\zeta^2 - 1| > \epsilon\sqrt{\kappa_N}] = O\left(\frac{\kappa_N e^{-\left(1+\epsilon\sqrt{\kappa_N}\right)^{1/2}}}{\left(1+\epsilon\sqrt{\kappa_N}\right)^{1/4}}\right) \xrightarrow{N\to\infty} 0$$

by the tail decay properties of the Gaussian distribution. This completes the proof. □

Lemma 1. *Assume that F_n is a sequence of distribution functions on $\mathbb{R}^d$ converging weakly* *to a standard Gaussian distribution function Φ^d, at a rate ϵ_n in the Kolmogorov distance,*

$$\sup_{\boldsymbol{x}\in\mathbb{R}^d} |F_n(\boldsymbol{x}) - \Phi^d(\boldsymbol{x})| = O(\epsilon_n).$$

Letting $d = p + q$, and given $\boldsymbol{y} \in \mathbb{R}^q$, we have

$$\sup_{x\in\mathbb{R}^p} |F_n(\boldsymbol{x}|\boldsymbol{y}) - \Phi^q(\boldsymbol{x})| = O(\epsilon_n).$$

Proof. By definition, and by our uniform bound, given any $\boldsymbol{y} \in \mathbb{R}^q$ we have that

$$\sup_{\boldsymbol{x}\in\mathbb{R}^p} |F_n(\boldsymbol{x}|\boldsymbol{y})F_n(\boldsymbol{y}) - \Phi^p(\boldsymbol{x})\Phi^q(\boldsymbol{y})| = \sup_{\boldsymbol{x}\in\mathbb{R}^p} |F_n(\boldsymbol{x},\boldsymbol{y}) - \Phi^d(\boldsymbol{x},\boldsymbol{y})| = O(\epsilon_n).$$

Now divide across by $\Phi^q(\boldsymbol{y})$, and obtain

$$\sup_{\boldsymbol{x}\in\mathbb{R}^p} \left| F_n(\boldsymbol{x}|\boldsymbol{y})\frac{F_n(\boldsymbol{y})}{\Phi^q(\boldsymbol{y})} - \Phi^p(\boldsymbol{x}) \right| = O(\epsilon_n) \tag{2}$$

By assumption of the theorem, it must also be that

$$|F_n(\boldsymbol{y}) - \Phi^q(\boldsymbol{y})| = O(\epsilon_n).$$

In turn, this implies that

$$\left| \frac{F_n(\boldsymbol{y})}{\Phi^q(\boldsymbol{y})} - 1 \right| = O(\epsilon_n), \tag{3}$$

for if this were not the case, for every $\alpha > 0$ and $M \geq 1$, there would exist and $m \geq M$ such that

$$\left| \frac{F_m(\boldsymbol{y})}{\Phi^q(\boldsymbol{y})} - 1 \right| > \frac{\alpha}{\Phi^q(\boldsymbol{y})}|\epsilon_m|,$$

or equivalently, for every $\alpha > 0$ and $M \geq 1$, there would exist and $m \geq M$ such that

$$|F_m(\boldsymbol{y}) - \Phi^q(\boldsymbol{y})| > \alpha|\epsilon_m|,$$

which would contradict the fact that $\sup_{\boldsymbol{u}} |F_n(\boldsymbol{u}) - \Phi^q(\boldsymbol{u})| \in O(\epsilon_n)$.

Now conditions (2) and (3) allow us to complete the proof by applying the triangle inequality:

$$d_\infty \left(F_n(\cdot|\boldsymbol{y}), \Phi_p\right) \leq d_\infty \left(F_n(\cdot|\boldsymbol{y}), \frac{F_n(\boldsymbol{y})}{\Phi_q(\boldsymbol{y})} F_n(\cdot|\boldsymbol{y})\right) + d_\infty \left(\frac{F_n(\boldsymbol{y})}{\Phi_q(\boldsymbol{y})} F_n(\cdot|\boldsymbol{y}), \Phi_p\right)$$

since

$$\begin{aligned} d_\infty \left(F_n(\cdot|\boldsymbol{y}), \frac{F_n(\boldsymbol{y})}{\Phi_q(\boldsymbol{y})} F_n(\cdot|\boldsymbol{y})\right) &= \sup_{\boldsymbol{x}\in\mathbb{R}^p} \left|F_n(\boldsymbol{x}|\boldsymbol{y}) - \frac{F_n(\boldsymbol{y})}{\Phi_q(\boldsymbol{y})} F_n(\boldsymbol{x}|\boldsymbol{y})\right| \\ &= \left|1 - \frac{F_n(\boldsymbol{y})}{\Phi_q(\boldsymbol{y})}\right| \sup_{\boldsymbol{x}\in\mathbb{R}^p} |F_n(\boldsymbol{x}|\boldsymbol{y})| \\ &= \left|1 - \frac{F_n(\boldsymbol{y})}{\Phi_q(\boldsymbol{y})}\right| = O(\epsilon_n) \end{aligned}$$

□

References

[1] Bosq, D.(2000). *Linear processes in function spaces.* Springer.

[2] DasGupta, A. (2008). *Asymptotic Theory of Statistics and Probability.* Springer.

[3] Dauxois, J. Pousse, A. & Romain, Y. (1982). Asymptotic theory for the principal component analysis of a random vector function: some applications to statistical inference. *Journal of Multivariate Analysis, **12**: 136–154.*

[4] Gibbs, A.L. & Su, F.E. (2002). On choosing and bounding probability metrics. *International Statistical Review, **70**(3): 419–435.*

[5] Shorack, G. R. (2000). *Probability for Statisticians.* Springer.