Supplementary Material for " Estimating the covariance of fragmented and other related types of functional data"

Aurore Delaigle[1], Peter Hall[1,3], Wei Huang[1] and Alois Kneip[2]
[1]ACEMS and School of Mathematics and Statistics, University of Melbourne, Australia
[2]Department of Economics and Hausdorff Center for Mathematics, Universität Bonn, Germany
[3]This work was initiated with Peter Hall during a visit of Kneip to the University Melbourne. Peter died on January 9, 2016.

A Discretely observed fragments

The data on the left panel of Figure A1 consist of discretely observed fragments where the observations are connected by a straight line.

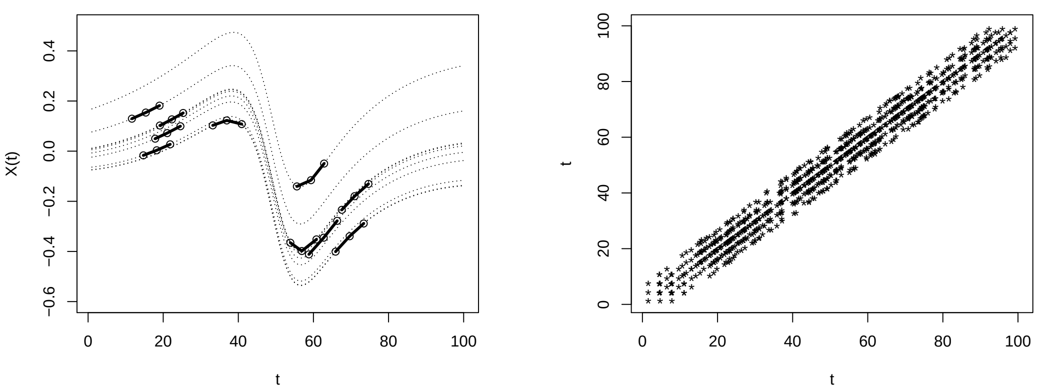

Figure A1: Left: subsample of 10 discretely observed curve fragments taken from a sample of size $n = 100$ of curve fragments $X_i(t)$, for $t \in \mathcal{I}_i \subset \mathcal{I} = [1, 100]$ and $i = 1, \ldots, n$; the observations (o) are connected by straight lines and the long dotted lines show the 10 corresponding unobserved full curves. Right: scatterplot of points $(s, t) \in \mathcal{I} \times \mathcal{I}$ where at least one pair $\big(X_i(s), X_i(t)\big)$ is observed, for $i = 1, \ldots, n$.

B More computational details

B.1 Spline estimator

In this section we show how to compute the spline estimator $\widehat{K}$ used on the diagonal band $\mathcal{S}$ in our simulations and in our real data analysis. Since we will need to refer both to this estimator and the empirical estimator at (2.2), in order to distinguish them, here we denote the spline estimator by $\widehat{K}_s$. As in sections 5.2 and 5.3, we assume that the data are observed on a discrete grid of points, denoted by $\mathcal{I}_D = \{T_1 < T_2 < \ldots < T_G\}$ such that $[T_1, T_G] = \mathcal{I}$. Let (s_i, t_i) for $i = 1, \ldots, N$ denote the points from $\mathcal{I}_D \times \mathcal{I}_D$ that are in $\mathcal{S}$. The spline estimator $\widehat{K}_s$ is obtained by minimising

$$\frac{1}{(G-1)^2} \sum_{i=1}^{N} \left[\widehat{K}(s_i, t_i) - \{\mathcal{B}(s_i) \otimes \mathcal{B}(t_i)\}^{\mathrm{T}} c\right]^2 m(s_i, t_i)$$
$$+\lambda \iint_{\mathcal{S}} \left[\{\mathcal{B}^{(2)}(s) \otimes \mathcal{B}(t) + \mathcal{B}(s) \otimes \mathcal{B}^{(2)}(t)\}^{\mathrm{T}} c\right]^2 ds\, dt \tag{B.1}$$

w.r.t. $c \in \mathbb{R}^{k+4}$, where $\otimes$ denotes the tensor product operator, $\mathcal{B}(t) = \big(B_1(t), \ldots, B_{k+4}(t)\big)^\top$ denotes a set of $k+4$ cubic B-splines defined on $\mathcal{I}$ with k knots, $\lambda > 0$ is a smoothing parameter and $m(s,t)$ was defined below (2.2). Finding a closed form solution to this minimisation problem is possible but not straightforward because the domain $\mathcal{S}$ has an irregular shape, so that the usual matrix formulations of the spline estimator do not work. Huang (2019) constructed an analytic solution which can be computed easily. The smoothing parameters k and λ are chosen by leave-one-curve-out cross validation.

B.2 Initial guess

As described in section 5.1, to find the matrix $\hat{A}_p = \hat{B}\hat{B}^{\mathrm{T}}$ that minimises (2.5), we use an iterative procedure that requires an initial guess. We take it to be the minimiser of (2.5) without the semi-positive definite constraint, that is, without imposing the form BB^{T}. Then we turn it into a covariance matrix by truncating its negative eigenvalues to zero. As in sections 5.2 and 5.3, we assume that the data are observed on a discrete grid of points, denoted by $\mathcal{I}_D = \{T_1 < T_2 < \ldots < T_G\}$ such that $[T_1, T_G] = \mathcal{I}$. Let (s_i, t_i) for $i = 1, \ldots, N$ denote the points from $\mathcal{I}_D \times \mathcal{I}_D$ that are in $\mathcal{S}$. In this setting, for a given p, the initial guess is obtained by first computing the matrix $\hat{A}_p^0$ that minimises, over $\mathbb{R}^{p\times p}$, the following discretised version of (2.5):

$$\begin{aligned} S_p(A_p) \;=\; & \frac{1}{(G-1)^2}\sum_{i=1}^{N}\left\{\widehat{K}(s_i,t_i) - M_p(s_i,t_i\,|\,A_p)\right\}^2 \\ & + \frac{1}{(G-1)}\sum_{i=1}^{G}\left\{\widehat{K}(t_i,t_i) - M_p(t_i,t_i\,|\,A_p)\right\}^2 . \end{aligned} \tag{B.2}$$

Recalling that $\otimes$ denotes the tensor product operator and letting $\widehat{\boldsymbol{K}} = \{\widehat{K}(s_1,t_1), \ldots, \widehat{K}(s_N,t_N)\}^{\mathrm{T}}$, we can express (B.2) as

$$S_p(A_p) \;=\; \frac{1}{(G-1)^2}\big(\widehat{\boldsymbol{K}} - \Psi_D \boldsymbol{a}_p\big)^{\mathrm{T}} W_D \big(\widehat{\boldsymbol{K}} - \Psi_D \boldsymbol{a}_p\big)\,,$$

where $W_D = \mathrm{diag}\{w_D(s_1,t_1), \ldots, w_D(s_N,t_N)\}$, with $w_D(s_i,t_i) = G$ if $s_i = t_i$ and $w_D(s_i,t_i) = 1$, otherwise, $\Psi_D = \{\Psi(s_1)\otimes\Psi(t_1), \ldots, \Psi(s_N)\otimes\Psi(t_N)\}^{\mathrm{T}} \in \mathbb{R}^{N\times p^2}$ with $\Psi(t) = \big(\psi_1(t), \ldots, \psi_p(t)\big)^{\mathrm{T}}$ and where $\boldsymbol{a}_p = (a_{11}, a_{12}, \ldots, a_{pp})^{\mathrm{T}}$ denotes the vectorised A_p. Denoting the vectorisation operator by $vec(\cdot)$, we have

$$vec(\hat{A}_p^0) = \big(\Psi_D^{\mathrm{T}} W_D \Psi_D\big)^{-1}\Psi_D^{\mathrm{T}} W_D \widehat{\boldsymbol{K}}\,. \tag{B.3}$$

Then, we compute the eigenvalues and eigenvectors of $\hat{A}_p^0$ and take our initial guess to be $\hat{B}_{p,0}\hat{B}_{p,0}^{\mathrm{T}}$, where $\hat{B}_{p,0} = \Gamma\Lambda^{1/2}$, with Γ the matrix whose columns are the p eigenvectors of $\hat{A}_p^0$ and Λ the diagonal matrix of eigenvalues of $\hat{A}_p^0$ arranged in the same order as those eigenvectors, but replacing the negative eigenvalues by zero.

Note that for a given p, $\Psi_D \in \mathbb{R}^{N\times p^2}$ and $W_D \in \mathbb{R}^{N\times N}$, so that the matrix $\Psi_D^{\mathrm{T}} W_D \Psi_D \in \mathbb{R}^{p^2\times p^2}$ is of rank $\min\{N, p^2\}$. Therefore, in order to ensure that $\Psi_D^{\mathrm{T}} W_D \Psi_D$ is invertible, we need to take $p \le \sqrt{N}$.

C Proofs of main theorems

C.1 Proof of Theorem 1

Proof of part a): We prove this result by induction. Consider an arbitrary continuous covariance function L defined on $\mathcal{S}_0$ satisfying $L(t,s) = K(t,s)$ for all $(t,s) \in \mathcal{S}$. Let $C_0^\epsilon := 0$, $C_1^\epsilon := 2\Theta_0\epsilon_1$, and $C_j^\epsilon := D_j C_{j-1}^\epsilon + 2\Theta_0\epsilon_j$ for $j = 2, \ldots, Q$.

Since $J_0 \times J_0 \subset \mathcal{S}$, the assertion holds for $j = 0$ since, for $j = 0$, we have

$$\sup_{(t,s)\in\bigcup_{k=0}^{j} J_k \times \bigcup_{k=0}^{j} J_k} |L(t,s) - K(t,s)| \leq \max_{k=0,\ldots,j} C_k^\epsilon \,. \tag{C.1}$$

Next, assuming that (C.1) holds for some $j \in \{0, 1, \ldots, Q-1\}$, we need to prove that it also holds when replacing j by $j+1$. Under the conditions of the theorem, for $j+1$ we assume that there exists $j^* \in \{0, \ldots, j\}$ with $(J_{j^*} \cup J_{j+1}) \times (J_{j^*} \cup J_{j+1}) \subset \mathcal{S}$ such that the values of $X(t)$, for $t \in J_{j+1}$, are linearly $(D_{j+1}, \epsilon_{j+1})$-predictable from $X^{J_{j^*}}$. Using the notation $X^*(t) = X(t) - \mu(t)$, this means that we assume that there exists a measurable function $\ell_t(s)$ such that $\sup_{t\in J_{j+1}} \int_{J_j^*} |\ell_t(s)|\, ds \leq D_{j+1}$,

$$X^*(t) = \int_{J_j^*} \ell_t(s) X^*(s)\, ds + Z(t)\,, \quad t \in J_{j+1}\,,$$

and $Var\{Z(t)\} = E[\{X^*(t) - \int_{J_{j^*}} \ell_t(s) X^*(s)\, ds\}^2] \leq \epsilon_{j+1}^2$. This can be expressed in terms of the covariance function as follows:

$$\left| K(t,t) - 2\int_{J_{j^*}} \ell_t(u) K(u,t)\, du + \int_{J_{j^*}} \int_{J_{j^*}} \ell_t(u)\ell_t(v) K(u,v)\, du\, dv \right| \leq \epsilon_{j+1}^2 \tag{C.2}$$

for all $t \in J_{j+1}$.

Now, let $\widetilde{X} \in L_2(\mathcal{I})$ denote a process with mean zero and covariance function equal to $L(t,s)$. Since $L(t,s) = K(t,s)$ for all $t, s \in J_{j^*} \cup J_{j^*}$, (C.2) implies that for all $t \in J_{j+1}$, we have $E[\{\widetilde{X}(t) - \int_{J_{j^*}} \ell_t(s)\widetilde{X}(s)\, ds\}^2] \leq \epsilon_{j+1}^2$, and hence

$$\widetilde{X}(t) = \int_{J_{j^*}} \ell_t(s)\widetilde{X}(s)\, ds + \widetilde{Z}(t)\,,$$

where $\widetilde{Z}(t)$ is a zero mean random variable with $Var\{\widetilde{Z}(t)\} = Var\{Z(t)\} \leq \epsilon_{j+1}^2$. Therefore, for all $s \in \bigcup_{k=0}^{j} J_k$ and for all $t \in J_{j+1}$ the Cauchy-Schwarz inequality together with (C.1) lead to

$$\begin{aligned}
&|K(t,s)-L(t,s)| = |E\{X^*(t)X^*(s)\} - E\{\widetilde{X}(t)\widetilde{X}(s)\}| \\
&= \left| E\left[\left\{\int_{J_{j^*}} \ell_t(u)X^*(u)\,du + Z(t)\right\}X^*(s)\right] - E\left[\left\{\int_{J_{j^*}} \ell_t(u)\widetilde{X}(u)\,du + \widetilde{Z}(t)\right\}\widetilde{X}(s)\right]\right| \\
&= \left|\int_{J_{j^*}} \ell_t(u)\{K(u,s)-L(u,s)\}\,du + E\{Z(t)X^*(s)\} - E\{\widetilde{Z}(t)\widetilde{X}(s)\}\right| \\
&\le C_j^\epsilon \left|\int_{J_{j^*}} \ell_t(u)\frac{K(u,s)-L(u,s)}{\sup_{(v,w)\in\bigcup_{k=0}^j J_k\times\bigcup_{k=0}^j J_k}|L(v,w)-K(v,w)|}\,du\right| + 2\Theta_0\epsilon_{j+1} \\
&\le D_{j+1}C_j^\epsilon + 2\Theta_0\epsilon_{j+1} = C_{j+1}^\epsilon\,,
\end{aligned}$$

where the last inequality follows from the fact that $\int_{J_{j^*}}|\ell_t(u)|\,du \le D_{j+1}$ and $J_{j^*} \subset \bigcup_{k=0}^j J_k$. Recall that $J_{j+1}\times J_{j+1} \subset \mathcal{S}$ and thus $K(t,s) = L(t,s)$ for all $(t,s) \in J_{j+1}\times J_{j+1}$. Combining these results with (C.1), we deduce that

$$\sup_{(t,s)\in\bigcup_{k=0}^{j+1} J_k\times\bigcup_{k=0}^{j+1} J_k} |L(t,s)-K(t,s)| \le \max_{k=0,\dots,j+1} C_k^\epsilon\,.$$

We conclude that (C.1) holds for $j = 0,\dots,Q$, which proves (3.4) because $C_0^\epsilon = 0$.

Proof of part b): Select an arbitrary $\delta > 0$. By assumption there exists $\epsilon_Q < \delta/(2\Theta_0)$ such that for some $0 < D_Q < \infty$ and some $j^* \in \{0,\dots,Q-1\}$ with $(J_{j^*}\cup J_Q)\times(J_{j^*}\cup J_Q) \subset \mathcal{S}$, the values of $X(t)$, for $t \in J_Q$, are linearly (D_Q,ϵ_Q)-predictable from $X^{J_{j^*}}$. Similarly, for any $j = Q-1,\dots,1$ there exists

$$\epsilon_j < \max\left\{\frac{\delta}{2\Theta_0}, \frac{\delta}{2\Theta_0 D_{j+1}}, \dots, \frac{\delta}{2\Theta_0 D_{j+1}\cdot\ldots\cdot D_Q}\right\}$$

such that for some $0 < D_j < \infty$ and some $j^* \in \{0,\dots,j-1\}$ with $(J_{j^*}\cup J_j)\times(J_{j^*}\cup J_j) \subset \mathcal{S}$, the values of $X(t)$, for $t \in J_j$, are linearly (D_j,ϵ_j)-predictable from $X^{J_{j^*}}$. It follows from (3.4) that $\sup_{(t,s)\in\mathcal{S}_0}|L(t,s)-K(t,s)| < Q\,\delta$. Since δ is arbitrary, this proves the result.

C.2 Proof of Theorem 2

The next proposition will be useful to prove the theorem. It states that, on the observed domain $\mathcal{S}$ and on its diagonal, as long as the number p of basis functions is sufficiently large, the rate of convergence of the empirical covariance estimator $\widehat{K}$ is preserved by our estimator $\widetilde{K}_p$ at (2.6), where $\hat{A}_p$ is the coefficient matrix that minimises $S_p(A_p)$ at (2.5). Its proof can be found in section D.4.

Proposition 3. *Assume that* (4.1) *holds. Then, as* $n \to \infty$*, there exists a sequence* $p(n)$ *with* $p(n) \to \infty$ *such that for* $p \geq p(n)$

$$\int_{\mathcal{S}} (\widetilde{K}_p - K)^2 = O_p(n^{-\kappa}) \tag{C.3}$$

and

$$\int_{\mathcal{I}} \left\{\widetilde{K}_p(t,t) - K(t,t)\right\}^2 dt = O_p(n^{-\kappa}) \tag{C.4}$$

Proof of Theorem 2. Proof of b): By the Cauchy-Schwarz inequality, using Proposition 3 we have $\int_{\mathcal{S}} \left|\widetilde{K}_p - K\right| = O_P(n^{-\kappa/2})$. Since $J_0 \times J_0 \subset \mathcal{S}$, for $j = 0$ we have

$|\widetilde{K}_p(t,s) - K(t,s)| \leq \max_{k=0,\dots,j} C_k^{\epsilon} + R_n^j(t,s)$ for all $(t,s) \in \bigcup_{k=0}^j J_k \times \bigcup_{k=0}^j J_k$, where the measurable random function R_n^j satisfies $\int_{\bigcup_{k=0}^j J_k \times \bigcup_{k=0}^j J_k} \left|R_n^j(t,s)\right| dt\, ds = O_P(n^{-\kappa/2})$. (C.5)

The proof now proceeds by induction. We will show that if (C.5) holds for some $j \in \{0, 1, \dots, Q-1\}$, then the inequality remains valid when replacing j by $j+1$.

Therefore, assume that (C.5) holds for some $j \in \{0, 1, \dots, Q-1\}$. By assumption, there exists $j^* \in \{0, \dots, j\}$ with $(J_{j^*} \cup J_{j+1}) \times (J_{j^*} \cup J_{j+1}) \subset \mathcal{S}$ such that the values of $X(t)$ for $t \in J_{j+1}$ are linearly $(D_{j+1}, \epsilon_{j+1})$-predictable from $X^{J_{j^*}}$. Using the notation $X^*(t) = X(t) - \mu(t)$ for all $t \in \mathcal{I}$, this means that there exists an integrable function $\ell_t(u)$ such that $L_{max} := \sup_{t \in J_{j+1}, u \in J_{j^*}} |\ell_t(u)| < \infty$,

$$X^*(t) = \int_{J_j^*} \ell_t(s) X^*(s)\, ds + Z(t)\,, \quad t \in J_{j+1}$$

and $Var\{Z(t)\} = E[\{X^*(t) - \int_{J_{j^*}} \ell_t(s) X^*(s)\, ds\}^2] \leq \epsilon_{j+1}^2$.

Using (C.2) we deduce that

$$\begin{aligned} z_{j,n} &:= \left|\widetilde{K}_p(t,t) - 2\int_{J_{j^*}} \ell_t(u) \widetilde{K}_p(u,t)\, du + \int_{J_{j^*}} \int_{J_{j^*}} \ell_t(u)\ell_t(v) \widetilde{K}_p(u,v)\, du\, dv\right| \\ &\leq \epsilon_{j+1}^2 + \widetilde{R}_{n,1}^j(t) \end{aligned} \tag{C.6}$$

for all $t \in J_{j+1}$, where

$$\begin{aligned}\widetilde{R}^j_{n,1}(t) =& |\widetilde{K}_p(t,t) - K(t,t)| + 2\Big|\int_{J_{j^*}} \ell_t(u)\big\{\widetilde{K}_p(u,t) - K(u,t)\big\}\,du\Big| \\ &+ \Big|\int_{J_{j^*}}\int_{J_{j^*}} \ell_t(u)\ell_t(v)\big\{\widetilde{K}_p(u,v) - K(u,v)\big\}\,du\,dv\Big|.\end{aligned}$$

By the Cauchy-Schwarz inequality, using Proposition 3 we also have $\int_{\mathcal{I}}\big|\widetilde{K}_p(t,t) - K(t,t)\big|\,dt = O_P(n^{-\kappa/2})$. Since $\ell_t(u)$ is uniformly bounded, which implies bounds such as $|\int_{J_{j^*}}\int_{J_{j^*}} \ell_t(u)\ell_t(v)\big\{\widetilde{K}_p(u,v) - K(u,v)\big\}\,du\,dv \le L_{max}\int_{J_{j^*}}\int_{J_{j^*}} |\widetilde{K}_p(u,v) - K(u,v)|\,du\,dv$, we deduce that

$$\int_{J_{j+1}} |\widetilde{R}^j_{n,1}(t)|\,dt = O_P(n^{-\kappa/2})\,. \tag{C.7}$$

By definition, $\widetilde{K}_p$ is a well-defined covariance function, and the corresponding covariance operator has at most p nonzero eigenvalues $\hat\theta_1,\ldots,\hat\theta_p$ and corresponding eigenfunctions $\hat\phi_1,\ldots,\hat\phi_p$. Given the estimate $\widetilde{K}_p$, let $\zeta_1,\ldots,\zeta_p$ be independent random variables with $\zeta_j \sim N(0,\hat\theta_j)$, $j = 1,\ldots,p$, and define the stochastic process $\widetilde{X}(t) = \sum_{j=1}^p \zeta_j\hat\phi_j(t)$ for $t \in \mathcal{I}$. We then obtain that $E\big\{\widetilde{X}(t)|\widetilde{K}_p\big\} = 0$ and $E\big\{\widetilde{X}(t)\widetilde{X}(s)|\widetilde{K}_p\big\} = \sum_{j=1}^p \hat\theta_j\hat\phi_j(t)\hat\phi_j(s) = \widetilde{K}_p(t,s)$ for all $t,s \in \mathcal{I}$.

Letting $\widetilde{Z}(t) = \widetilde{X}(t) - \int_{J_{j^*}} \ell_t(s)\widetilde{X}(s)\,ds$, $t \in \mathcal{I}$, we deduce from (C.6) that

$$Var\big\{\widetilde{Z}(t)|\widetilde{K}_p\big\} = z_{j,n} \le \epsilon^2_{j+1} + \widetilde{R}^j_{n,1}(t) \tag{C.8}$$

for all $t \in J_{j+1}$. Now, similarly as in the proof of Theorem 1 we have

$$\begin{aligned}&|K(t,s) - \widetilde{K}_p(t,s)| = |E\{X^*(t)X^*(s)\} - E\{\widetilde{X}(t)\widetilde{X}(s)|\widetilde{K}_p\}| \\ &= \Big|\int_{J_{j^*}} \ell_t(u)\{K(u,s) - \widetilde{K}_p(u,s)\}\,du + E\{Z(t)X^*(s)\} - E\{\widetilde{Z}(t)\widetilde{X}(s)|\widetilde{K}_p\}\Big|\end{aligned} \tag{C.9}$$

for all $t \in J_{j+1}$ and $s \in \bigcup_{k=0}^j J_k$. Next we bound each of those terms separately.

For the third one we have

$$\begin{aligned}|E\{\widetilde{Z}(t)\widetilde{X}(s)|\widetilde{K}_p\}| &\le \Big[Var\{\widetilde{X}(s)|\widetilde{K}_p\}\Big]^{1/2}\Big[Var\{\widetilde{Z}(t)|\widetilde{K}_p\}\Big]^{1/2} \\ &\le \Big\{\Theta_0^2 + |\widetilde{K}_p(s,s) - K(s,s)|\Big\}^{1/2}\Big\{\epsilon^2_{j+1} + \widetilde{R}^j_{n,1}(t)\Big\}^{1/2} \\ &\le \Theta_0\epsilon_{j+1} + \widetilde{R}^j_{n,2}(t,s)\,,\end{aligned}$$

where

$$\int_{J_{j+1}} \int_{\bigcup_{k=0}^{j} J_k} |\widetilde{R}_{n,2}^{j}(t,s)|\, ds\, dt = O_P(n^{-\kappa/2})\,. \tag{C.10}$$

Similarly, for the second term we have

$$|E\{Z(t)X^*(s)\}| \le \Theta_0 \epsilon_{j+1}, \quad t \in J_{j+1},\ s \in \bigcup_{k=0}^{j} J_k, \tag{C.11}$$

while $J_{j^*} \subset \bigcup_{k=0}^{j} J_k$. Finally, for the first term, using (C.5) we have

$$\begin{aligned} &\Big|\int_{J_{j^*}} \ell_t(u)\{K(u,s) - \widetilde{K}_p(u,s)\}\, du\Big| \le \int_{J_{j^*}} \ell_t(u)\{C_j^{\epsilon} + |R_n^j(u,s)|\}\, du \\ &\le D_{j+1} C_j^{\epsilon} + L_{max} \int_{J_{j^*}} |R_n^j(u,s)|\, du \end{aligned} \tag{C.12}$$

for all $t \in J_{j+1}$ and $s \in \bigcup_{k=0}^{j} J_k$. Combing the bounds for the three terms, we deduce that (C.5) holds for $j+1$ and with $t \in J_{j+1}$ and $s \in \bigcup_{k=0}^{j} J_k$.

Moreover, using Proposition 3 and the fact that $J_{j+1} \times J_{j+1} \subset \mathcal{S}$, we have $\int_{J_{j+1}\times J_{j+1}} |K - \widetilde{K}_p| = O_P(n^{-\kappa/2})$. We deduce from (C.9)–(C.12) that, with $C_{j+1}^{\epsilon} = D_{j+1}C_j^{\epsilon} + 2\Theta_0\epsilon_{j+1}$,

$$|\widetilde{K}_p(t,s) - K(t,s)| \le \max_{k=0,\dots,j+1} C_k^{\epsilon} + R_n^{j+1}(t,s)$$

holds for all $(t,s) \in \bigcup_{k=0}^{j+1} J_k \times \bigcup_{k=0}^{j+1} J_k$, where the measurable random functions R_n^{j+1} satisfies $\int_{\bigcup_{k=0}^{j+1} J_k \times \bigcup_{k=0}^{j+1} J_k} \left|R_n^{j+1}(t,s)\right| dt\, ds = O_P(n^{-\kappa/2})$. Since Q is fixed, this proves the result.

Proof of a): The proof proceeds by induction. First, since $J_0 \times J_0 \subset \mathcal{S}$, for $j = 0$ we have

$$\int_{\bigcup_{k=0}^{j} J_k \times \bigcup_{k=0}^{j} J_k} |\widetilde{K}_p(t,s) - K(t,s)|\, dt\, ds = O_P(n^{-\kappa/2})\,. \tag{C.13}$$

Then we show that if (C.13) holds for some $j \in \{0, 1, \dots, Q-1\}$, then the inequality remains valid when replacing j by $j+1$.

Therefore, assume that (C.13) holds for some $j \in \{0, 1, \ldots, Q-1\}$ and recall that, by assumption, there exists some $j^* \in \{0, \ldots, j\}$ such that $(J_{j^*} \cup J_{j+1}) \times (J_{j^*} \cup J_{j+1}) \subset \mathcal{S}$. Since we have assumed that, for $t \in \mathcal{I}$, $X(t)$ is $(D_{j^*}, 0)$-predictable from $X^{J_{j^*}}$, there exists a measurable function $\ell_t(u)$ such that $L_{max} := \sup_{t \in \mathcal{I}, u \in J_{j^*}} |\ell_t(u)| < \infty$ and for which, if we let $X^* = X - \mu$, we have

$$X^*(t) = \int_{J_{j^*}} \ell_t(s) X^*(s)\, ds\,, \quad t \in \mathcal{I}\,. \tag{C.14}$$

As in the proof of part b), define let $\widetilde{X}$ be a process satisfying $E\{\widetilde{X}(t)|\widetilde{K}_p\} = 0$ and $E\{\widetilde{X}(t)\widetilde{X}(s)|\widetilde{K}_p\} = \widetilde{K}_p(t, s)$, $t, s \in \mathcal{I}$, and for $t \in \mathcal{I}$, let $\widetilde{Z}(t) = \widetilde{X}(t) - \int_{J_{j^*}} \ell_t(s)\widetilde{X}(s)\, ds$. Using (C.8) with $\epsilon_{j+1} = 0$ we get, for $t \in J_{j+1}$,

$$Var\{\widetilde{Z}(t)|\widetilde{K}_p\} = \widetilde{R}^j_{n,1}(t) \text{ with } \int_{J_{j+1}} |\widetilde{R}^j_{n,1}(t)|\, dt = O_P(n^{-\kappa/2})\,. \tag{C.15}$$

Now by (C.14), (C.13), and the fact that $\int_{\bigcup_{k=0}^j J_k} \big|\widetilde{K}_p(s, s) - K(s, s)\big|\, ds = O_P(n^{-\kappa/2})$, the same type of arguments can be applied to show that also for $s \in \bigcup_{k=0}^j J_k$ there exists a measurable random function $\widetilde{R}^j_{n,3}(s)$ with

$$Var\{\widetilde{Z}(s)|\widetilde{K}_p\} = \widetilde{R}^j_{n,3}(t) \text{ with } \int_{\bigcup_{k=0}^j J_k} |\widetilde{R}^j_{n,3}(t)|\, dt = O_P(n^{-\kappa/2})\,. \tag{C.16}$$

Since, under our assumptions, $0 = Z(t) = X^*(t) - \int_{J_{j^*}} \ell_t(u)X^*(u)\, du$ for every $t \in \mathcal{I}$, then for all $t \in J_{j+1}$ and $s \in \bigcup_{k=0}^j J_k$ the equality at (C.9) becomes

$$\begin{aligned} |K(t, s) - \widetilde{K}_p(t, s)| = \Big| &\int_{J_{j^*}} \ell_t(u)\{K(u, s) - \widetilde{K}_p(u, s)\}\, du \\ &- E\Big\{\widetilde{Z}(t) \int_{J_{j^*}} \ell_s(u)\widetilde{X}(u)\, du|\widetilde{K}_p\Big\} - E\big\{\widetilde{Z}(t)\widetilde{Z}(s)|\widetilde{K}_p\big\}\Big|\,. \end{aligned} \tag{C.17}$$

Next we bound those three terms. For the second one, using again the fact that $0 = Z(t) = X^*(t) - \int_{J_{j^*}} \ell_t(u)X^*(u)\, du$, we have $0 = E\big\{Z(t) \int_{J_{j^*}} \ell_s(u)X^*(u)\, du\big\} = \int_{J_{j^*}} \ell_s(u)K(u, t)\, du - \int_{J_{j^*}} \int_{J_{j^*}} \ell_t(u)\ell_s(v)K(u, v)\, du\, dv$. Therefore

$$\begin{aligned} &E\Big\{\widetilde{Z}(t) \int_{J_{j^*}} \ell_s(u)\widetilde{X}(u)\, du|\widetilde{K}_p\Big\} \\ &= \int_{J_{j^*}} \ell_s(u)\{\widetilde{K}_p(u, t) - K(u, t)\}\, du - \int_{J_{j^*}} \int_{J_{j^*}} \ell_t(u)\ell_s(v)\{\widetilde{K}_p(u, v) - K(u, v)\}\, du\, dv\,. \end{aligned}$$

Consequently, using Proposition 3 and the fact that $(J_{j^*} \cup J_{j+1}) \times (J_{j^*} \cup J_{j+1}) \subset \mathcal{S}$, uniform boundedness of $\ell_t(u)$ and $\ell_s(u)$ implies that

$$\int_{J_{j+1}} \int_{\bigcup_{k=0}^{j} J_k} \Big| E\Big\{ \widetilde{Z}(t) \int_{J_{j^*}} \ell_s(u) \widetilde{X}(u)\, du \Big| \widetilde{K}_p \Big\} \Big|\, ds\, dt = O_P(n^{-\kappa/2}) \tag{C.18}$$

and

$$\int_{J_{j+1}} \int_{\bigcup_{k=0}^{j} J_k} \Big| \int_{J_{j^*}} \ell_t(u) \{K(u,s) - \widetilde{K}_p(u,s)\}\, du \Big|\, ds\, dt = O_P(n^{-\kappa/2})\,. \tag{C.19}$$

Moreover, (C.15) and (C.16) lead to

$$\int_{J_{j+1}} \int_{\bigcup_{k=0}^{j} J_k} \Big| E\{\widetilde{Z}(t) \widetilde{Z}(s) | \widetilde{K}_p\} \Big|\, ds\, dt = O_P(n^{-\kappa/2})\,. \tag{C.20}$$

Finally note that by Proposition 3 and $J_{j+1} \times J_{j+1} \subset \mathcal{S}$ we have $\int_{J_{j+1} \times J_{j+1}} |K - \widetilde{K}_p| = O_P(n^{-\kappa/2})$. Therefore, we can deduce from (C.17)–(C.20) that

$$\int_{\bigcup_{k=0}^{j+1} J_k \times \bigcup_{k=0}^{j+1}} |\widetilde{K}_p(t,s) - K(t,s)|\, dt\, ds = O_P(n^{-\kappa/2})\,.$$

Since Q is fixed, the desired result a) is an immediate consequence.

Proof of c): Under the additional condition made in c), the arguments used in the proof of Theorem 1.*b* imply that for any $\delta > 0$ there exist $\epsilon_j > 0$, $j = 1, \ldots, Q$, such that $\max_{j=1,\ldots,Q} C_j^{\epsilon} \le Q \cdot \delta$. By Theorem 2.*b* we deduce that $P\big\{ \int_{\mathcal{S}_0} |K - \widetilde{K}_p| \le (Q+1) \cdot \delta \int_{S_0} ds\, dt \big\} \to 1$ as $n \to \infty$. Since δ is arbitrary, the result follows. □

D Other technical details

D.1 Proof of identifiability for the setting in section 3.2.1

In this section we derive slightly more general identifiability results than those under which the $\Phi_q^{J_k}$ matrices of inner products at (3.8) are invertible. We include the case where the ϕ_j's vanish on some parts of their domain, where the zero parts arise in a hierarchical manner to the left and to the right of an interval

$J_0 \subset \mathcal{I} = [a,b]$. Specifically, we decompose $\mathcal{I}$ into the union of J_0 and disjoint intervals $J_{-1},\dots,J_{-Q_1} \subset \mathcal{I}$ located to the left of J_0 (if $a \notin J_0$) and $J_1,\dots,J_{Q_2} \subset \mathcal{I}$ located to the right of J_0 (if $b \notin J_0$). That is, we suppose that for two integers Q_1 and Q_2 such that $\max(Q_1,Q_2) > 0$, there is a sequence $J_k, k = -Q_1,\dots,Q_2$ of disjoint intervals such that $\bigcup_{k=-Q_1}^{Q_2} J_k = \mathcal{I}$, and such that for $j = -Q_1,\dots,Q_2 - 1$, $(J_j \cup J_{j+1}) \times (J_j \cup J_{j+1}) \subset \mathcal{S}$. Suppose too that the J_j's are organised in increasing order, that is, for $j = -Q_1,\dots,Q_2 - 1$, if $s \in J_j$ and $t \in J_{j+1}$ then $s \le t$. Then if we define $m_0 = 0$, the conditions of Theorem 1 are satisfied if the following conditions are satisfied:

(i) The q eigenfunctions $\phi_1,\dots,\phi_q$ are linearly independent on J_0.

(ii) For $j = 1,\dots,Q_2$, $m_j \in [m_{j-1}, q-1]$ of the q eigenfunctions $\phi_1,\dots,\phi_q$ vanish on $J_j \cup \ldots \cup J_{Q_2}$ and the $q - m_j$ remaining ones are linearly independent on $J_0 \cup \ldots \cup J_j$.

(iii) If $Q_1 \ge 1$, for $j = -1,\dots,-Q_1$, $m_j \in [m_{j+1}, q-1]$ of the q eigenfunctions $\phi_1,\dots,\phi_q$ vanish on $J_{-Q_1} \cup \ldots \cup J_j$ and the $q - m_j$ remaining ones are linearly independent on $J_j \cup \ldots \cup J_0$.

Note that in (i) and (ii) above, we allow for the case where, but do not impose that, some of the eigenfunctions vanish on some subintervals.

Next we prove this result in the case where $Q_1 = 0$ and $Q_2 \ge 1$. The proof is similar when Q_1 and Q_2 are both nonzero and when $Q_1 \ge 1$ and $Q_2 = 0$. Also, without loss of generality we assume that the m_j eigenfunctions that vanish in (i) are the last m_j ones. Finally, for $k = -Q_1,\dots,Q_2$ and $r,s \in \{1,\dots,q\}$, let $\Phi_{q;rs}^{J_k} = \int_{J_k} \phi_r(t)\phi_s(t)\,dt$ as in (3.8).

Using (i), for $j = 0,\dots,Q_2$, we have that $\phi_k = 0$ for $k > q - m_j$. Therefore, when $j = 1,\dots,Q_2$, for $s \in J_{j-1}$, we have

$$X(s) = \mu(s) + \sum_{r=1}^{q-m_{j-1}} \xi_r \phi_r(s)\,.$$

Now $\phi_1, \ldots, \phi_{q-m_{j-1}}$ are linearly independent on J_{j-1}, so that the $(q-m_{j-1}) \times (q-m_{j-1})$ matrix $\Phi^{J_{j-1}}_{q-m_{j-1}}$ of inner products $\Phi^{J_{j-1}}_{q;rs}$ at (3.8), for $r, s = 1, \ldots, q-m_{j-1}$, is invertible. Therefore we can write

$$(\xi_1, \ldots, \xi_{q-m_{j-1}})^{\mathrm{T}} = (\Phi^{J_{j-1}}_{q-m_{j-1}})^{-1}\Big(\int_{J_{j-1}} \phi_1(s)\{X(s)-\mu(s)\}\, ds, \ldots, \tag{D.1}$$

$$\int_{J_{j-1}} \phi_{q-m_{j-1}}(s)\{X(s)-\mu(s)\}\, ds\Big)^{\mathrm{T}}. \tag{D.2}$$

Likewise, for $t \in J_j$ we have $X(t) = \mu(t) + \sum_{r=1}^{q-m_j} \xi_r \phi_r(t)$, and using the above decomposition for $X(s)$, we deduce that

$$X(t) = \mu(t) + \int_{J_{j-1}} \ell^{J_{j-1}}_{t,q}(s)\{X(s)-\mu(s)\}\, ds\,,$$

where

$$\ell^{J_{j-1}}_{t,q}(s) = \big(\phi_1(t), \ldots, \phi_{q-m_j}(t)\big)\Omega\big(\phi_1(s), \ldots, \phi_{q-m_{j-1}}(s)\big)^{\mathrm{T}}, \tag{D.3}$$

and where Ω is the $(q-m_j) \times (q-m_{j-1})$ matrix obtained by keeping only the first $q-m_j$ rows of $(\Phi^{J_{j-1}}_{q-m_j})^{-1}$.

Thus we have just proved that, for $j = 1, \ldots, Q_2$, and for all $t \in J_j$, $X(t)$ is linearly $(\sup_{t\in\mathcal{I}} \int_{J_{j-1}} |\ell^{J_{j-1}}_{t,q}(s)|\, ds, 0)$-predictable from $X^{J_{j-1}}$. Therefore, the condition of Theorem 1.a is satisfied for $\epsilon_1 = \epsilon_2 = \ldots \epsilon_q = 0$, and thus the condition of Theorem 1.b is satisfied too. As a consequence, $L(t,s) = K(t,s)$ for all $(t,s) \in \mathcal{S}$ implies $\sup_{(t,s)\in\mathcal{S}_0} |L(t,s) - K(t,s)| = 0$.

D.2 Proof of Proposition 1.i

Let $J = J_k$ and $X^* = X - \mu$ and define $\chi_q(s) = \sum_{r=1}^q \xi_r \phi_r(s)$ for all $s \in \mathcal{I}$. Since Φ^J_q is invertible, we have

$$(\xi_1, \ldots, \xi_q)^{\mathrm{T}} = (\Phi^J_q)^{-1}\Big(\int_J \phi_1(s)\chi_q(s)\, ds, \ldots, \int_J \phi_q(s)\chi_q(s)\, ds\Big)^{\mathrm{T}}.$$

Therefore, for any $t \in \mathcal{I}$, if we take

$$\ell_t(s) = \Big(\phi_1(t), \ldots, \phi_q(t)\Big)(\Phi^J_q)^{-1}\Big(\phi_1(s), \ldots, \phi_q(s)\Big)^{\mathrm{T}}, \tag{D.4}$$

we can write $\chi_q(t) = \sum_{r=1}^q \xi_r \phi_r(t) = \int_J \ell_t(s)\chi_q(s)\,ds$.

Using the Karhunen-Loève decomposition, we have $X^*(t) = \chi_q(t) + \sum_{r=q+1}^\infty \xi_r\phi_r(t)$. We deduce that for $t \in \mathcal{I}$, we can write

$$X(t) = \mu(t) + \int_J \ell_t(s)X^*(s)\,ds + Z(t)\,,$$

where

$$Z(t) = X^*(t) - \int_J \ell_t(s)X^*(s)\,ds = \int_J \ell_t(s)\{X^*(s) - \chi_q(s)\}\,ds + \sum_{r=q+1}^\infty \xi_r\phi_r(t)\,.$$

Now since, $X^*(s) - \chi_q(s) = \sum_{r=q+1}^\infty \xi_r\phi_r(s)$ for all $s \in J$, the definition of Θ_q implies that for all $t \in \mathcal{I}$

$$Var\{Z(t)\} \le \Theta_q^2\left\{\int_J\int_J |\ell_t(s)||\ell_t(u)|\,ds\,du + 2\int_J |\ell_t(s)|\,ds + 1\right\} \le (D^{q,J}+1)^2\Theta_q^2\,.$$

This shows that $X(t)$ is $\big(D^{q,J_k},(D^{q,J_k}+1)\Theta_q\big)$-predictable from X^J. The bound on $|L(t,s) - K(t,s)|$ follows by combining this with Theorem 1.a.

D.3 Proof of Proposition 2

By definition of $\mathcal{S}$ at (2.1), there exists a sufficiently large integer Q and a sequence of disjoint subintervals $J_0, J_1, \ldots, J_Q$ of $\mathcal{I}$ with $\bigcup_{k=0}^Q J_k = \mathcal{I}$ and with the following property: For any $k \in \{1,\ldots,Q\}$ there exists a $k^* \in \{0,\ldots,k-1\}$ such that $(J_k \cup J_{k^*}) \times (J_k \cup J_{k^*}) \subset \mathcal{S}$ and $\sup_{t\in J_k, s\in J_{k^*}} |t-s| \le c$, where c is defined at (3.11) i).

Let $k \in \{1,\ldots,Q\}$. Next we show that for $t \in J_k$, $X(t)$ is linearly predictible from $X(s)$ with $s \in J_{k^*}$. For this, let $Z(t) = X(t) - \int_{J_{k^*}} \ell_{h,q,s;t}(u)X(u)\,du$, where

$$\ell_{h,q,s;t}(u) = \sum_{r=0}^q \frac{1}{h^{r+1}r!}W^{(r)}\left(\frac{s-u}{h}\right)(t-s)^r\,, \tag{D.5}$$

with $h > 0$ a real number, q an integer, and W a $q+2$-times continuously differentiable probability density, symmetric around 0, and with compact support $[-1,1]$.

Now, we have

$$
\begin{aligned}
Var\{Z(t)\} &= E\left[\left\{X(t) - \int_{J_{k^*}} \ell_{h,q,s;t}(u)X(u)\,du\right\}^2\right] \\
&\leq 2E\left[\left\{\int_{J_{k^*}} \ell_{h,q,s;t}(u)X(u)du - \sum_{r=0}^{q} \frac{1}{r!}X^{(r)}(s)(t-s)^r\right\}^2\right] \\
&\quad + 2E\left[\left\{X(t) - \sum_{r=0}^{q} \frac{1}{r!}X^{(r)}(s)(t-s)^r\right\}^2\right] \leq \epsilon_k^2 \qquad \text{(D.6)}
\end{aligned}
$$

To bound the second term of the last inequality, note that when s belongs to the interior of J_{k^*}, by (3.11) i) and iii), for all $\epsilon_k > 0$ there exists an integer $q \equiv q_{\epsilon_k}$ such that for any $t \in J_k$, if h is small enough,

$$
E\left[\left\{X(t) - \sum_{r=0}^{q} \frac{1}{r!}X^{(r)}(s)(t-s)^r\right\}^2\right] \leq \frac{\epsilon_k^2}{4}\,. \qquad \text{(D.7)}
$$

For the first inequality, we have

$$
\begin{aligned}
&\int_{J_{k^*}} \ell_{h,q,s;t}(u)X(u)\,du - \sum_{r=0}^{q} \frac{1}{r!}X^{(r)}(s)(t-s)^r \\
&= \sum_{r=0}^{q} \frac{(t-s)^r}{r!}\left\{\int_{J_{k^*}} \frac{1}{h^{r+1}}W^{(r)}\left(\frac{s-u}{h}\right)X(u)\,du - X^{(r)}(s)\right\}.
\end{aligned}
$$

Using repeated integration by parts followed by a first order Taylor expansion of $X^{(r)}$, it can be shown that there exist constants $0 < D_{W,r}$, $r = 1, \ldots, q$, such that

$$
\left|X^{(r)}(s) - \frac{1}{h^{r+1}}\int_{J_{k^*}} W^{(r)}\left(\frac{s-u}{h}\right)X(u)\,du\right| \leq h^2 D_{W,r} \sup_{v\in[s-h,s+h]} |X^{(r+2)}(v)|\,. \qquad \text{(D.8)}
$$

Therefore, using also (3.11) ii), we have

$$
\begin{aligned}
E\Big\{ \int_{J_{k^*}} \ell_{h,q,s;t}(u)X(u)\,du - \sum_{r=0}^{q} \frac{1}{r!} X^{(r)}(s)(t-s)^r \Big\}^2 \\
&\leq h^4 E\Big[\Big\{ \sum_{r=0}^{q} \frac{(t-s)^r}{r!} D_{W,r} \sup_{v\in[s-h,s+h]} |X^{(r+2)}(v)| \Big\}^2 \Big] \\
&\leq q h^4 \sum_{r=0}^{q} \frac{(t-s)^{2r}}{(r!)^2} D_{W,r}^2 E\Big\{ \sup_{v\in[s-h,s+h]} |X^{(r+2)}(v)|^2 \Big\} \\
&\leq q h^4 \sum_{r=0}^{q} \frac{c^{2r}}{(r!)^2} D_{W,r}^2 E\Big\{ \sup_{|v-s|\leq c} |X^{(r+2)}(v)|^2 \Big\} \\
&\leq \frac{\epsilon_k^2}{4}, \qquad \text{(D.9)}
\end{aligned}
$$

as long as we take $h \equiv h_{\epsilon_k}$ small enough.

Combining (D.6), (D.7) and (D.9), we deduce that $Var\{Z(t)\} \leq \epsilon_k^2$. Thus we have proved that for all $k \in \{1,\dots,Q\}$ and all $\epsilon_k > 0$ there exists $D_k > 0$ such that the random variables $X(t)$, $t \in J_k$, are linearly (D_k, ϵ_k)-predictable from $X^{J_{k^*}}$ for some interval J_{k^*}, where $k^* \in \{0,\dots,k-1\}$ and $(J_k \cup J_{k^*}) \times (J_k \cup J_{k^*}) \subset \mathcal{S}$. The proposition follows from Theorem 1.b.

D.4 Proof of Proposition 3

Recall that $A_p = BB^{\mathrm{T}}$ and that (2.5) is thus minimized with respect to all possible covariance functions which can be represented in the form $\sum_{k=1}^{p} v_{k,p}(s)v_{k,p}(t)$, where $v_{k,p} = \sum_{j=1}^{p} b_{j,k}\psi_j$ and the $b_{j,k}$'s are coefficients. The sequence $\psi_1, \psi_2, \dots$ forms a basis of $L_2(\mathcal{I})$. Hence, there exist coefficients $b^*_{j,k}$ such that if p is sufficiently large, any rescaled eigenfunction $\sqrt{\theta_r}\phi_r$ in (3.2) can be approximated with arbitrary accuracy, in the $L_2(\mathcal{I})$ sense, by a function of the form $v^*_{k,p} = \sum_{j=1}^{p} b^*_{j,k}\psi_j$.

This implies that for any $D > 0$ and all integers n and Q there exists an integer $p_{n,Q}$ with the following property: For all $p \geq p_{n,Q}$ there are coefficients $b^*_{j,k}$, $j,k = 1,\dots,p$, with $b^*_{j,k} = 0$ for $j > Q$ such that with $A^*_p = B^*(B^*)^{\mathrm{T}}$, $M_p(s,t\,|\,A^*_p) =$

$\sum_{k=1}^{Q} v_{k,p}^*(s)v_{k,p}^*(t)$, and $K_Q(t,s) = \sum_{r=1}^{Q} \theta_r\phi_r(s)\phi_r(t)$ we have

$$\int_{\mathcal{S}_0} \{K_Q - M_p(\,\cdot\,|\,A_p^*)\}^2 \le Dn^{-\kappa}, \quad \text{and} \int_{\mathcal{I}} \{K_Q(t,t) - M_p(t,t\,|\,A_p^*)\}^2\, dt \le Dn^{-\kappa}.$$

On the other hand, by (3.2) there necessarily exists an integer Q_n such that

$$\int_{\mathcal{S}_0} (K_{Q_n} - K)^2 \le Dn^{-\kappa}, \quad \text{and} \int_{\mathcal{I}} \{K_{Q_n}(t,t) - K(t,t)\}^2\, dt \le Dn^{-\kappa}.$$

By the Cauchy-Schwarz inequality we can thus conclude that for all $p \ge p_{n,Q_n} \equiv p(n)$

$$\int_{\mathcal{S}_0} \{K - M_p(\,\cdot\,|\,A_p^*)\}^2 \le 2\int_{\mathcal{S}_0} \{K_{Q_n} - M_p(\,\cdot\,|\,A_p^*)\}^2 + 2\int_{\mathcal{S}_0} (K_{Q_n} - K)^2 \le 4Dn^{-\kappa},$$

$$\begin{aligned}\int_{\mathcal{I}} \{K(t,t) - M_p(t,t\,|\,A_p^*)\}^2\, dt &\le 2\int_{\mathcal{I}} \{K_{Q_n}(t,t) - M_p(t,t\,|\,A_p^*)\}^2\, dt \\ &\quad + 2\int_{\mathcal{I}} \{K_{Q_n}(t,t) - K(t,t)\}^2\, dt \le 4Dn^{-\kappa}. \end{aligned} \tag{D.10}$$

Using again the Cauchy-Schwarz inequality, (D.10) and (4.1)(iii) imply that for all $p \ge p_{n,Q_n} \equiv p(n)$

$$\begin{aligned} S_p(\hat{A}_p) &= \int_{\mathcal{S}} \big\{\widehat{K} - M_p(\,\cdot\,|\,\hat{A}_p)\big\}^2 + \int_{\mathcal{I}} \{\widehat{K}(t,t) - M_p(t,t\,|\,\hat{A}_p)\}^2 dt \\ &\le \int_{\mathcal{S}} \{\widehat{K} - M_p(\,\cdot\,|\,A_p^*)\}^2 + \int_{\mathcal{I}} \{\widehat{K}(t,t) - M_p(t,t\,|\,A_p^*)\}^2\, dt \\ &\le 2\Big[\int_{\mathcal{S}} \{K - M_p(\,\cdot\,|\,A_p^*)\}^2 + \int_{\mathcal{I}} \{K(t,t) - M_p(t,t\,|\,A_p^*)\}^2\, dt\Big] \\ &\qquad + 2\Big[\int_{\mathcal{S}} (K - \widehat{K})^2 + \int_{\mathcal{I}} \{K(t,t) - \widehat{K}(t,t)\}^2\, dt\Big] \\ &\le 16Dn^{-\kappa} + 2\Big[\int_{\mathcal{S}} (K - \widehat{K})^2 + \int_{\mathcal{I}} \{K(t,t) - \widehat{K}(t,t)\}^2\, dt\Big] = O_P(n^{-\kappa})\,, \end{aligned}$$

which leads to

$$\int_{\mathcal{S}} \{\widehat{K} - M_p(\,\cdot\,|\,\hat{A}_p)\}^2 = O_P(n^{-\kappa})\,, \quad \int_{\mathcal{I}} \{\widehat{K}(t,t) - M_p(t,t\,|\,\hat{A}_p)\}^2\, dt = O_P(n^{-\kappa})\,. \tag{D.11}$$

Now recall from (2.6) that $\widetilde{K}_p(s,t)$ can be written as $M_p(s,t\,|\,\hat{A}_p)$. Since

$$\int_{\mathcal{S}} \{K - M_p(\,\cdot\,|\,\hat{A}_p)\}^2 \le 2\int_{\mathcal{S}} \big\{\widehat{K} - M_p(\,\cdot\,|\,\hat{A}_p)\big\}^2 + 2\int_{\mathcal{S}} (K - \widehat{K})^2,$$

$$\int_{\mathcal{I}} \{K(t,t) - M_p(t,t\,|\,\hat{A}_p)\}^2\, dt \le 2\int_{\mathcal{I}} \{\widehat{K}(t,t) - M_p(t,t\,|\,\hat{A}_p)\}^2\, dt + 2\int_{\mathcal{I}} \{K(t,t) - \widehat{K}(t,t)\}^2\, dt,$$

the assertions of the proposition follow (D.11) and (4.1)(iii).

References

Huang, W. (2019). Spline techniques for incomplete and complex data (Doctoral dissertation). Retrieved from The University of Melbourne Library.