Supplemental paper for:

On the Optimal Reconstruction of Partially Observed Functional Data

by Dominik Liebl and Alois Kneip

CONTENT

Appendix A of this supplemental material contains the detailed list of assumptions. The proofs of Theorems 2.1, 2.2, 4.1, and 4.2 are given in Appendix B. The main steps in our proofs of the Theorems 4.1 and 4.2 are as in Yao, Müller and Wang (2005a). Though, by contrast to Yao, Müller and Wang (2005a), we allow for a time series context (see Assumption A2), consider a different asymptotic setup (see Assumption A4), and exploit the assumption of kernel functions with compact support (see Assumption A7).

APPENDIX A: LIST OF ASSUMPTIONS

A2 (Weak Dependency) Let U_{it} i.i.d. as U_t for all $i \in \{1, \dots, n\}$. Let $(X_t)_t$, $(U_t)_t$, $(A_t)_t$, and $(B_t)_t$ be strictly stationary processes, where $(U_t)_t$, $(A_t)_t$, and $(B_t)_t$ are independent from $(X_t)_t$. Define the following autocovariance functions: $\gamma_h^U = \mathrm{Cov}_h(U_t, U_{t+h})$, $\gamma_h(u,v) = \mathrm{Cov}_h(X_t(u), X_{t+h}(v))$, and $\dot{\gamma}_h((u_1,v_1),(u_2,v_2)) = \mathrm{Cov}_h(\dot{X}_t(u_1,v_1), \dot{X}_{t+h}(u_2,v_2))$ for $h \in \{1,2,\dots\}$, where $\dot{X}_s(u,v) = (X_j(u) - \mu(u))(X_j(v) - \mu(v))$. We assume that there are (generic) constants c and r, with $0 < c < \infty$ and $0 < r < 1$, such that $|\gamma_h^U| \leq cr^h$, $\sup_{(u,v)\in I_0^2} |\gamma_h(u,v)| \leq cr^h$, and $\sup_{(u_1,v_1,u_2,v_2)\in I_0^4} |\dot{\gamma}_h((u_1,v_1),(u_2,v_2))| \leq cr^h$.

A3 (Random Design under Proper Subsets) Given a time point t, U_{it} is a continuous random variable, i.i.d. as $U_t \sim f_{U_t}$, where $f_{U_t}(u)$ is a random probability density function (pdf) defined as

$$f_{U_t}(u) := f_{U_0}(u)\Big(\mathbb{1}(u \in \mathcal{I}_t) / \textstyle\int_{\mathcal{I}_t} f_{U_0}(v)dv\Big), \tag{21}$$

where $\mathbb{1}(.)$ denotes the indicator function, $\mathcal{I}_t$ denotes a proper (random) subset as in Assumption RS, and the pdf $f_{U_0}(u) > 0$ for all $u \in I_0 = [a,b] \subset \mathbb{R}$ and zero else.

For estimating μ (i.e., pooled data): Unconditionally on time points t, the data (Y_{it}, U_{it}) is i.i.d. as $(Y,U) \sim g_{YU}$, where $g_{YU}(y,u) = f_{Y|U}(y|u)f_U(u)$, and where $f_U(u) = \mathbb{E}(f_{U_t}(u))$. It is assumed $f_U(u) > 0$ for all $u \in \mathcal{I}_0$ and zero else, that f_U' is continuous, and that there exist

(generic) constants $c > 0$ such that $\sup_{u \in I_0} |f_U''(u)| < c < \infty$ and $\sup_{(u,y) \in I_0 \times \mathbb{R}} |(\partial^2/\partial u^2) g_{YU}(y,u)| \leq c < \infty$.
For estimating γ (i.e., pooled data): Furthermore, $(Y_{it}, Y_{is}, U_{it}, U_{is})$, $t \neq s$, is i.i.d. as $(Y_1, Y_2, U_1, U_2) \sim g_{YYUU}$ and there exist a (generic) constant $c > 0$ such that $\sup_{(y_1,y_2,u_1,u_2) \in \mathbb{R}^2 \times I_0^2} |(\partial^4/(\partial u_1^2 \partial u_2^2)) g_{YYUU}(y_1,y_2,u_1,u_2)| \leq c < \infty$.

A4 (Asymptotic Scenario) $Tn \to \infty$, where $n = n(T) \geq 2$ such that $n(T) \sim T^\theta$ with $0 \leq \theta < \infty$. Here "$a_T \sim b_T$" denotes that two sequences a_T and b_T are asymptotically equivalent up to some positive constant $0 < c < \infty$, i.e., that $\lim_{T \to \infty}(a_T/b_T) = c$.

A5 (Smoothness) For estimating μ: The functions $\mu''(u)$, $\partial\gamma(u,v)/\partial u$, and $\partial\gamma(u,v)/\partial v$ are continuous for all $u, v \in I_0$.
For estimating γ: All second order derivatives of $\gamma(u,v)$ and all first order derivatives of $\dot{\gamma}((u_1,v_1),(u_2,v_2))$ are continuous for all points within its supports I_0^2 and I_0^4.

A6 (Bandwidths) For estimating μ: $h_\mu = h_{Tn,\mu} \to 0$ and $(Tn\, h_\mu)^{-1/2} \to \infty$ as $Tn \to \infty$. For estimating γ: $h_\gamma = h_{TN,\gamma} \to 0$ and $(TN\, h_\gamma)^{-1/2} \to \infty$ as $TN \to \infty$, where $N = n^2 - n$.

A7 (Kernel Function) The kernel function κ is assumed to be a univariate, symmetric, pdf with compact support $\text{supp}(\kappa) = [-1, 1]$, such as, e.g., the univariate Epanechnikov kernel.

A8 (Eigenvalues) For $k \in \{1, 2, \dots\}$, let $\lambda_k = \mathcal{O}(k^{-a})$, with $a > 1$, and $\lambda_k - \lambda_{k+1} \geq \text{const.} \times k^{-a-1}$. Correspondingly, for λ_k^{sm} and λ_k^{la}, but possibly with different constants and parameters a.

A9 (Additional Regularity Assumption on X_t)

$$\sup_{u \in I_0}\{\mathbb{E}(|X_t(u)|^C)\} < \infty \quad \text{for all} \quad C > 0,$$

$$\sup_{u,v \in I_0}(\mathbb{E}[\{|u-v|^{-\varepsilon}|X_t(u) - X_t(v)|\}^C]) < \infty \quad \text{for some} \quad \varepsilon > 0 \quad \text{and}$$

$$\sup_{k \geq 1} \lambda_k^{-r}\{\mathbb{E}[(\int_a^b (X_t(u) - \mathbb{E}(X_t(u)))\phi_k(u)\, du)^{2r}]\} < \infty$$

for each $r = 1, 2, \dots$.

APPENDIX B: PROOFS

Proof of Theorem 2.1: Note that for any K and every $u \in I_{\text{la}}$, and particularly for every $u \in I_{\text{la}} \setminus I_{\text{sm}}$, we have

$$0 \leq \mathbb{E}\left(\left(X_t^{\text{la}}(u) - \sum_{k=1}^K \xi_{tk}^{\text{sm}} \tilde{\phi}_k^{\text{la}}(u)\right)^2\right) = \gamma^{\text{la}}(u,u) - \sum_{k=1}^K \lambda_k^{\text{sm}} \tilde{\phi}_k^{\text{la}}(u)^2, \tag{22}$$

which implies that $\mathbb{V}(\sum_{k=1}^{K}\xi_{tk}^{\mathrm{sm}}\tilde{\phi}_k^{\mathrm{la}}(u)) = \sum_{k=1}^{K}\lambda_k^{\mathrm{sm}}\tilde{\phi}_k^{\mathrm{la}}(u)^2$ converges to a fixed limit $0 \leq \mathbb{V}(\widetilde{X}_t^{\mathrm{la}}(u)) < \infty$ as $K \to \infty$ for all $u \in I_{\mathrm{la}}$. Moreover, continuity of $\gamma^{\mathrm{la}}(u,v)$ implies continuity of $\mathbb{V}(\widetilde{X}_t^{\mathrm{la}}(u))$.

Proof of Theorem 2.2, part (a): For all $v \in I_{\mathrm{sm}}$ and $u \in I_{\mathrm{la}} \setminus I_{\mathrm{sm}}$ we have that

$$\begin{aligned}
\mathbb{E}\left(X_t^{\mathrm{sm}}(v)Z_t(u)\right) &= \mathbb{E}\left(X_t^{\mathrm{sm}}(v)\left(X_t^{\mathrm{la}}(u) - \widetilde{X}_t^{\mathrm{la}}(u)\right)\right) = \\
&= \mathbb{E}\left(\sum_{k=1}^{\infty}\xi_{tk}^{\mathrm{sm}}\phi_k^{\mathrm{sm}}(v)\Big(X_t^{\mathrm{la}}(u) - \sum_{k=1}^{\infty}\xi_{tk}^{\mathrm{sm}}\tilde{\phi}_k^{\mathrm{la}}(u)\Big)\right) = \\
&= \sum_{k=1}^{\infty}\phi_k^{\mathrm{sm}}(v)\Big(\,\mathbb{E}(\xi_{tk}^{\mathrm{sm}}X_t^{\mathrm{la}}(u)) - \lambda_r\tilde{\phi}_k^{\mathrm{la}}(u)\Big) = 0,
\end{aligned}$$

since $\mathbb{E}(\xi_{tk}^{\mathrm{sm}}X_t^{\mathrm{la}}(u)) = \lambda_k^{\mathrm{sm}}\,\tilde{\phi}_k^{\mathrm{la}}(u)$. This proves Eq. (5), while Eq. (6) directly follows from the definition of $Z_t(x)$.

Part (b): Note that the Riesz representation theorem implies that

$$\ell_u(X_t^{\mathrm{sm}}) = \int_{I_{\mathrm{sm}}} b(v)X_t^{\mathrm{sm}}(v)dv$$

for some $b \in L^2(I_{\mathrm{sm}})$. By Eq. (5) and the orthogonality property of the least squares projection we thus obtain

$$\begin{aligned}
&\mathbb{E}\left(\left(X_t^{\mathrm{la}}(u) - \ell_u(X_t^{\mathrm{sm}})\right)^2\right) = \\
&= E\left(\left(\widetilde{X}_t^{\mathrm{la}}(u) + Z_t(u) - \int_{I_{\mathrm{sm}}} b(v)X_t^{\mathrm{sm}}(v)dv\right)^2\right) = \\
&= \mathbb{E}\left(\left(\widetilde{X}_t^{\mathrm{la}}(u) - \int_{I_{\mathrm{sm}}} b(v)X_t^{\mathrm{sm}}(v)dv\right)^2\right) + \mathbb{E}(Z_t(u)^2) + \\
&+ 2\left(\mathbb{E}(\widetilde{X}_t^{\mathrm{la}}(u)\,Z_t(u)) - \int_{I_{\mathrm{sm}}} b(v)\,\mathbb{E}\left(X_t^{\mathrm{sm}}(v)Z_t(u)\right)dv\right) = \\
&= \mathbb{E}\left(\left(\widetilde{X}_t^{\mathrm{la}}(u) - \int_{I_{\mathrm{sm}}} b(v)X_t^{\mathrm{sm}}(v)dv\right)^2\right) + \mathbb{E}(Z_t^2(u)) \geq \mathbb{E}(Z_t^2(u)).
\end{aligned}$$

Part (c): From basic statistics we know that $\mathbb{V}(Z_t(u)-Z_s(u)) = \mathbb{V}(Z_t(u)) + \mathbb{V}(Z_s(u)) - 2\,Cov(Z_t(u), Z_s(u))$ for all $u \in I_{\mathrm{la}}$. Under our stationarity and weak dependency assumptions (Assumption A2) we have that $\mathbb{V}(Z_t(u) - Z_s(u)) = 2\ \mathbb{V}(Z_t(u)) + \mathcal{O}(r^{|t-s|})$, where we use the typical α-mixing covariance inequality; see, e.g., in Fan and Yao (2003) Ch. 2.6.2. Rearranging and using that $\mathbb{E}(Z_t(u)) = \mathbb{E}(Z_s(u)) = 0$ for all $u \in I_{\text{la}}$ and all $t, s \in \{1, \dots, T\}$ yields $\mathbb{V}(Z_t(u)) = \frac{1}{2}\,\mathbb{E}((Z_t(u) - Z_s(u))^2) + \mathcal{O}(r^{|t-s|})$. From result (a) we know that $Z_t(u)$ and $X_t^{\text{sm}}(v)$ are orthogonal and therefore uncorrelated for all $u \in I_{\text{la}} \setminus I_0$ and all $v \in I_{\text{sm}}$, i.e., $\mathbb{E}\left(X_t^{\text{sm}}(v) Z_t(u)\right) = \text{Cov}\left(X_t^{\text{sm}}(v), Z_t(u)\right) = 0$. Under the assumption of a Gaussian time series process $(X_t)_{t\in\mathbb{Z}}$, we have then independency between $Z_t(u)$ and $X_t^{\text{sm}}(v)$, such that

$$\mathbb{V}(Z_t(u)) = \frac{1}{2}\,\mathbb{E}\left(\mathbb{E}\left((Z_t(u) - Z_s(u))^2\right) | X_t^{\text{sm}} = X_s^{\text{sm}}\right) + \mathcal{O}(r^{|t-s|}),$$

where $X_t^{\text{sm}} = X_s^{\text{sm}}$ means that $X_t^{\text{sm}}(u) = X_s^{\text{sm}}(u)$ for all $u \in I_{\text{sm}}$.

$\square$

Proof of Theorem 4.1, part (a): Note that the estimator $\hat{\mu}(u; h_\mu)$ can be written as

$$\hat{\mu}(u; h_\mu) = e_1^\top S_{1,Tn,u}^{-1} S_{2,Tn,u}, \tag{23}$$

with 2×2 matrix

$$\begin{aligned} S_{1,Tn,u} &= (Tn)^{-1}[\mathbf{1}, \mathbf{U}_u]^\top \mathbf{W}_{\mu,u}[\mathbf{1}, \mathbf{U}_u] \\ &= \begin{pmatrix} \frac{1}{Tnh_\mu}\sum_{it}\kappa\left(\frac{U_{it}-u}{h_\mu}\right) & \frac{1}{Tnh_\mu}\sum_{it}\kappa\left(\frac{U_{it}-u}{h_\mu}\right)(U_{it}-u) \\ \frac{1}{Tnh_\mu}\sum_{it}\kappa\left(\frac{U_{it}-u}{h_\mu}\right)(U_{it}-u) & \frac{1}{Tnh_\mu}\sum_{it}\kappa\left(\frac{U_{it}-u}{h_\mu}\right)(U_{it}-u)^2 \end{pmatrix}, \end{aligned}$$

and 2×1 vector

$$S_{2,Tn,u} = (Tn)^{-1}[\mathbf{1}, \mathbf{U}_u]^\top \mathbf{W}_{\mu,u}\mathbf{Y} = \begin{pmatrix} \frac{1}{Tnh_\mu}\sum_{it}\kappa\left(\frac{U_{it}-u}{h_\mu}\right) Y_{it} \\ \frac{1}{Tnh_\mu}\sum_{it}\kappa\left(\frac{U_{it}-u}{h_\mu}\right)(U_{it}-u)\, Y_{it} \end{pmatrix}.$$

Using the notation and the results from Lemma B.1 we have that

$$\begin{aligned} S_{1,Tn,u} &= \begin{pmatrix} \Psi_{0,Tn}(u; h_\mu) & \Psi_{1,Tn}(u; h_\mu) \\ \Psi_{1,Tn}(u; h_\mu) & \Psi_{2,Tn}(u; h_\mu) \end{pmatrix} \\ &= \begin{pmatrix} f_U(u) & 0 \\ 0 & f_U(u)\nu_2(\kappa) \end{pmatrix} + \mathcal{O}_p^{\text{Unif}}\left(h_\mu^2 + \frac{1}{\sqrt{Tn\,h_\mu}} + \frac{1}{\sqrt{T}}\right) \quad \text{and} \end{aligned} \tag{24}$$

$$S_{2,Tn,u} = \begin{pmatrix} \Psi_{3,Tn}(u; h_\mu) \\ \Psi_{4,Tn}(u; h_\mu) \end{pmatrix} = \begin{pmatrix} \mu(u) f_U(u) \\ 0 \end{pmatrix} + \mathcal{O}_p^{\text{Unif}}\left(h_\mu^2 + \frac{1}{\sqrt{Tn\,h_\mu}} + \frac{1}{\sqrt{T}}\right), \tag{25}$$

where we write $\Psi_{q,Tn}(u;h_\mu)-m_q(u)=\mathcal{O}_p^{\text{Unif}}(\texttt{rate}_n)$ in order to denote that $\sup_{u\in I_0}|\Psi_{q,Tn}(u;h_\mu)-m_q(u)|=\mathcal{O}_p(\texttt{rate}_n)$. Taking the inverse of (24) gives

(26)
$$S_{Tn,u}^{-1}=\begin{pmatrix}1/f_U(u) & 0\\ 0 & 1/(f_U(u)\nu_2(\kappa))\end{pmatrix}+\mathcal{O}_p^{\text{Unif}}\left(h_\mu^2+\frac{1}{\sqrt{Tn\,h_\mu}}+\frac{1}{\sqrt{T}}\right).$$

Plugging (26) and (25) into (23) leads to

$$\sup_{u\in I_0}|\hat{\mu}(u;h_\mu)-\mu(u)|=\mathcal{O}_p\left(h_\mu^2+\frac{1}{\sqrt{Tn\,h_\mu}}+\frac{1}{\sqrt{T}}\right).$$

□

Proof of Theorem 4.1, part (b): Let us initially consider the infeasible estimator $\hat{\gamma}_C$ that is based on the infeasible "clean" dependent variables $C_{ijt}=(Y_{it}-\mu(U_{it}))(Y_{jt}-\mu(U_{jt}))$ instead of the estimator $\hat{\gamma}$ in (14) that is based on the "dirty" dependent variables $\hat{C}_{ijt}=(Y_{it}-\hat{\mu}(U_{it}))(Y_{jt}-\hat{\mu}(U_{jt}))$, which are contaminated through having to estimate the unknown mean function μ. Equivalently to the estimator $\hat{\mu}$ above, we can write the estimator $\hat{\gamma}_C$ as

$$\hat{\gamma}_C(u,v;h_\gamma)=e_1^\top\tilde{S}_{1,TN,(u,v)}^{-1}\tilde{S}_{2,TN,(u,v)}, \tag{27}$$

with

$$\tilde{S}_{1,TN,(u,v)}^{-1}=\begin{pmatrix}\Theta_{0,TN}(u,v;h_\gamma) & \Theta_{1,TN}(u,v;h_\gamma)\\ \Theta_{1,TN}(u,v;h_\gamma) & \Theta_{2,TN}(u,v;h_\gamma)\end{pmatrix}^{-1}$$
$$=\begin{pmatrix}1/f_{UU}(u,v) & 0\\ 0 & 1/f_{UU}(u,v)(\nu_2(\kappa))^2\end{pmatrix}+\mathcal{O}_p^{\text{Unif}}\left(h_\gamma^2+\frac{1}{\sqrt{TN\,h_\gamma^2}}+\frac{1}{\sqrt{T}}\right) \tag{28}$$

and $\tilde{S}_{2,TN,(u,v)}=$

(29)
$$=\begin{pmatrix}\Theta_{3,TN}(u,v;h_\gamma)\\ \Theta_{4,TN}(u,v;h_\gamma)\end{pmatrix}=\begin{pmatrix}\gamma(u,v)f_{UU}(u,v)\\ 0\end{pmatrix}+\mathcal{O}_p^{\text{Unif}}\left(h_\gamma^2+\frac{1}{\sqrt{TN\,h_\gamma^2}}+\frac{1}{\sqrt{T}}\right),$$

where we use the notation and the results from Lemma B.2, and where we write $\Theta_{q,TN}(u,v;h_\gamma)-\eta_q(u,v)=\mathcal{O}_p^{\text{Unif}}(\texttt{rate}_n)$ in order to denote that $\sup_{(u,v)\in I_0^2}|\Theta_{q,TN}(u,v;h_\gamma)-\eta_q(u,v)|=\mathcal{O}_p(\texttt{rate}_n)$.

Plugging (28) and (29) into (27) leads to

$$\sup_{(u,v)\in I_0^2} |\hat{\gamma}_C(u,v;h_\mu) - \gamma(u,v)| = \mathcal{O}_p\left(h_\gamma^2 + \frac{1}{\sqrt{TN\,h_\gamma^2}} + \frac{1}{\sqrt{T}}\right). \tag{30}$$

It remains to consider the additional estimation error, which comes from using the "dirty" dependent variables $\hat{C}_{ijt}$ instead of "clean" dependent variables C_{ijt}. Observe that we can expand $\hat{C}_{ijt}$ as following:

$$\begin{aligned}\hat{C}_{ijt} = C_{ijt} &+ (Y_{it} - \mu(U_{it}))(\mu(U_{jt}) - \hat{\mu}(U_{jt}))\\ &+ (Y_{jt} - \mu(U_{jt}))(\mu(U_{it}) - \hat{\mu}(U_{it}))\\ &+ (\mu(U_{it}) - \hat{\mu}(U_{it}))(\mu(U_{jt}) - \hat{\mu}(U_{jt})).\end{aligned}$$

Using our finite moment assumptions on Y_{it} (Assumption A1) and our result in Theorem 4.1, part (a), we have that

$$\begin{aligned}\hat{C}_{ijt} &= C_{ijt} + \mathcal{O}_p(1)\mathcal{O}_p\left(h_\mu^2 + \frac{1}{\sqrt{Tn\,h_\mu}} + \frac{1}{\sqrt{T}}\right)\\ &\quad + \mathcal{O}_p(1)\mathcal{O}_p\left(h_\mu^2 + \frac{1}{\sqrt{Tn\,h_\mu}} + \frac{1}{\sqrt{T}}\right)\\ &\quad + \left(\mathcal{O}_p\left(h_\mu^2 + \frac{1}{\sqrt{Tn\,h_\mu}} + \frac{1}{\sqrt{T}}\right)\right)^2 = C_{ijt} + \mathcal{O}_p\left(h_\mu^2 + \frac{1}{\sqrt{Tn\,h_\mu}} + \frac{1}{\sqrt{T}}\right)\end{aligned}$$

for all $i \neq j \in \{1,\dots,n\}$ and $t \in \{1,\dots,T\}$.

Case of non-parametric rates: As long as n does not diverge too fast with T, i.e., as long as θ in $n \sim T^\theta$, $0 < \theta < 1$, is sufficiently small, the non-parametric rate components, i.e., h_μ^2 and $(Tn\,h_\mu)^{-1/2}$ are of a larger order than the parametric $T^{-1/2}$ component. For the sake of a simple argument, let us focus on the case of optimal bandwidth choices, i.e., $h_\mu \sim (Tn)^{-1/5}$ and $h_\gamma \sim (TN)^{-1/6}$. Then the non-parametric rate components are dominating if $0 < \theta < 1/4$ and we have that $\hat{C}_{ijt} = C_{ijt} + \mathcal{O}_p\left(T^{(-2/5)(1+\theta)}\right)$ and $\sup_{(u,v)\in I_0^2} |\hat{\gamma}_C(u,v;h_\mu) - \gamma(u,v)| = \mathcal{O}_p\left(T^{(-1/3)(1+2\theta)}\right)$. But $(-2/5)(1+\theta) < (-1/3)(1+2\theta) \Leftrightarrow \theta < 1/4$ which implies that the difference between C_{ijt} and $\hat{C}_{ijt}$ is of an order of magnitude smaller than the approximation error between $\hat{\gamma}_C$ and γ. It follows then form standard arguments that $\hat{\gamma}(u,v;h_\mu)$ converges at the same rate as $\hat{\gamma}_C(u,v;h_\mu)$, i.e.,

$$\sup_{(u,v)\in I_0^2} |\hat{\gamma}(u,v;h_\mu) - \gamma(u,v)| = \mathcal{O}_p\left(h_\gamma^2 + \frac{1}{\sqrt{TN\,h_\gamma^2}} + \frac{1}{\sqrt{T}}\right).$$

Similar but more lengthy arguments apply for the case of non-optimal bandwidth choices and for the case of the parametric $T^{-1/2}$ rate. $\square$

LEMMA B.1. *Define*

$$\Psi_{q,Tn}(u;h_\mu) = \frac{1}{Tnh_\mu}\sum_{it}\kappa\left(\frac{U_{it}-u}{h_\mu}\right)\psi_q\left(U_{it}-u,Y_{it}\right), \tag{31}$$

where

$$\psi_q\left(U_{it}-u,Y_{it}\right) = \begin{cases} (U_{it}-u)^q & \text{for } q\in\{0,1,2\} \\ Y_{it} & \text{for } q=3 \\ (U_{it}-u)\,Y_{it} & \text{for } q=4. \end{cases}$$

Then, under Assumptions A1-A7,

$$\tau_{q,Tn} = \sup_{u\in I_0}\left|\Psi_{q,Tn}(u;h_\mu)-m_q(u)\right| = \mathcal{O}_p\left(h_\mu^2+\frac{1}{\sqrt{Tn\,h_\mu}}+\frac{1}{\sqrt{T}}\right),$$

where $m_0(u)=f_U(u)$, $m_1(u)=0$, $m_2(u)=f_U(u)\nu_2(\kappa)$, $m_3(u)=\mu(u)f_U(u)=\mathbb{E}(Y_{it}|U_{it}=u)f_U(u)$, *and* $m_4(u)=0$.

LEMMA B.2. *Define*

$$\Theta_{q,TN}(u,v;h_\gamma) = \frac{1}{TNh_\gamma}\sum_{i\neq j,t}\kappa\left(\frac{U_{it}-u}{h_\gamma}\right)\kappa\left(\frac{U_{jt}-v}{h_\gamma}\right)\vartheta_q\left(U_{it}-u,U_{jt}-u,C_{ijt}\right), \tag{32}$$

where

$$\vartheta_q\left(U_{it}-u,U_{jt}-v,C_{ijt}\right) = \begin{cases} (U_{it}-u)^q\,(U_{jt}-v)^q & \text{for } q\in\{0,1,2\} \\ C_{ijt} & \text{for } q=3 \\ (U_{it}-u)\,(U_{jt}-v)\,C_{ijt} & \text{for } q=4. \end{cases}$$

Then, under Assumptions A1-A7,

$$\varrho_{q,Tn} = \sup_{(u,v)\in I_0^2}\left|\Theta_{q,TN}(u,v;h_\gamma)-\eta_q(u,v)\right| = \mathcal{O}_p\left(h_\gamma^2+\frac{1}{\sqrt{TN\,h_\gamma^2}}+\frac{1}{\sqrt{T}}\right),$$

where $\eta_0(u,v) = f_{UU}(u,v)$, $\eta_1(u,v) = 0$, $\eta_2(u,v) = f_{UU}(u,v)(\nu_2(\kappa))^2$, $\eta_3(u,v)=\gamma(u,v)f_{UU}(u,v)=\mathbb{E}(C_{ijt}|(U_{it},U_{jt})=(u,v))f_{UU}(u,v)$, *and* $m_4(u,v)=0$.

Proof of Lemma B.1: Remember that $\mathbb{E}(|X_n|) = \mathcal{O}(\texttt{rate}_n)$ implies that $X_n = \mathcal{O}_p(\texttt{rate}_n)$, therefore, we focus in the following on "$\mathbb{E}(|X_n|)$". Adding a zero and applying the triangle inequality yields that $\mathbb{E}(\tau_{q,Tn}) =$

$$\begin{aligned}\mathbb{E}(\sup_{u\in I_0}|\Psi_{q,Tn}(u;h_\mu)-m_q(u)|) \leq & \sup_{u\in I_0}|\,\mathbb{E}(\Psi_{q,Tn}(u;h_\mu))-m_p(u)|+\\ &+\mathbb{E}(\sup_{u\in I_0}|\Psi_{q,Tn}(u;h_\mu)-\mathbb{E}(\Psi_{q,Tn}(u;h_\mu))|).\end{aligned} \tag{33}$$

Let us first focus on the second summand in (33). The next steps will make use of the Fourier transformation of the kernel function κ see, e.g., (see, e.g., Tsybakov, 2008, Ch. 1.3):

$$\kappa^{\text{ft}}(x) := \mathcal{F}[\kappa](x) = \int_{\mathbb{R}} \kappa(z)\exp(-\mathrm{i}zx)dz = \int_{-1}^{1}\kappa(z)\exp(-\mathrm{i}zx)dz$$

with $\mathrm{i} = \sqrt{-1}$. Remember that, by Assumption A7, $\kappa(.)$ has a compact support $[-1,1]$. The inverse transform gives then

$$\kappa\left(s\right) = \frac{1}{2\pi}\int_{\mathbb{R}}\kappa^{\text{ft}}(x)\exp\left(\mathrm{i}xs\right)dx = \frac{1}{2\pi}\int_{\mathbb{R}}\kappa^{\text{ft}}(x)\exp\left(\mathrm{i}xs\right)dx\ \mathbb{1}_{(|s|<1)}.$$

Furthermore, we can use that (see Tsybakov, 2008, Ch. 1.3, Eq. (1.34)) $\mathcal{F}[\kappa(./h_\mu)/h_\mu](x) = \mathcal{F}[\kappa](xh_\mu) = \kappa^{\text{ft}}(xh_\mu)$ which yields

$$\begin{aligned}\kappa\left(s/h_\mu\right)/h_\mu &= \frac{1}{2\pi}\int_{\mathbb{R}}\mathcal{F}[\kappa(./h_\mu)/h_\mu](x)\exp\left(\mathrm{i}xs\right)dx\ \mathbb{1}_{(|s|<h_\mu)}\\ &= \frac{1}{2\pi}\int_{\mathbb{R}}\kappa^{\text{ft}}(xh_\mu)\exp\left(\mathrm{i}xs\right)dx\ \mathbb{1}_{(|s|<h_\mu)}.\end{aligned} \tag{34}$$

Plugging (34) into (31) yields $\Psi_{q,Tn}(u;h_\mu) =$

$$\begin{aligned}&= \frac{1}{Tn}\sum_{it}\kappa\left(\frac{U_{it}-u}{h_\mu}\right)\frac{1}{h_\mu}\psi_q\left(U_{it}-u,Y_{it}\right)\\ &= \frac{1}{Tn}\sum_{it}\frac{1}{2\pi}\int_{\mathbb{R}}\kappa^{\text{ft}}(xh_\mu)\exp\big(\mathrm{i}x(U_{it}-u)\big)dx\ \mathbb{1}_{(|U_{it}-u|<h_\mu)}\ \psi_q\left(U_{it}-u,Y_{it}\right)\\ &= \frac{1}{2\pi}\int_{\mathbb{R}}\left[\frac{1}{Tn}\sum_{it}\exp\big(\mathrm{i}xU_{it}\big)\,\psi_q\left(U_{it}-u,Y_{it}\right)\ \mathbb{1}_{(|U_{it}-u|<h_\mu)}\right]\exp\big(\mathrm{i}xu\big)\kappa^{\text{ft}}(xh_\mu)dx.\end{aligned}$$

Using that $|\exp(ixu)| \leq 1$ leads to

$$\mathbb{E}(\sup_{u\in I_0}|\Psi_{q,Tn}(u;h_\mu)-\mathbb{E}(\Psi_{q,Tn}(u;h_\mu))|) \leq \frac{1}{2\pi}\,\mathbb{E}\left(\sup_{u\in I_0}\left|\int_{\mathbb{R}}\tilde{\omega}_{q,Tn}(u,x)\cdot\kappa^{\text{ft}}(xh_\mu)dx\right|\right),$$

where

$$\tilde{\omega}_{q,Tn}(u,x) = \frac{1}{Tn}\sum_{it}\Big[\exp\big(\mathrm{i}xU_{it}\big)\psi_q\left(U_{it}-u,Y_{it}\right)\,\mathbb{1}_{(|U_{it}-u|<h_\mu)} - \mathbb{E}\left(\exp\big(\mathrm{i}xU_{it}\big)\psi_q\left(U_{it}-u,Y_{it}\right)\mathbb{1}_{(|U_{it}-u|<h_\mu)}\right)\Big].$$

Using further that κ^{ft} is symmetric, since κ is symmetric by Assumption A7, and that $\exp\big(\mathrm{i}xU_{it}\big) = \cos\big(xU_{it}\big) + \mathrm{i}\sin\big(xU_{it}\big)$ leads to

$$\frac{1}{2\pi}\,\mathbb{E}\left(\sup_{u\in I_0}\left|\int_{\mathbb{R}}\tilde{\omega}_{q,Tn}(u,x)\cdot\kappa^{\mathrm{ft}}(xh_\mu)dx\right|\right) = \frac{1}{2\pi}\,\mathbb{E}\left(\sup_{u\in I_0}\left|\int_{\mathbb{R}}\omega_{q,Tn}(u,x)\cdot\kappa^{\mathrm{ft}}(xh_\mu)dx\right|\right),$$

where

$$\omega_{q,Tn}(u,x) = \frac{1}{Tn}\sum_{it}\Big[\cos\big(xU_{it}\big)\psi_q\left(U_{it}-u,Y_{it}\right)\,\mathbb{1}_{(|U_{it}-u|<h_\mu)} - \mathbb{E}\left(\cos\big(xU_{it}\big)\psi_q\left(U_{it}-u,Y_{it}\right)\mathbb{1}_{(|U_{it}-u|<h_\mu)}\right)\Big], \tag{35}$$

such that

$$\begin{aligned}
&\mathbb{E}(\sup_{u\in I_0}|\Psi_{q,Tn}(u;h_\mu)-\mathbb{E}(\Psi_{q,Tn}(u;h_\mu))|)\\
&\leq \frac{1}{2\pi}\int_{\mathbb{R}}\mathbb{E}\left(\sup_{u\in I_0}\big|\omega_{q,Tn}(u,x)\big|\right)\cdot\Big|\kappa^{\mathrm{ft}}(xh_\mu)\Big|\,dx\\
&\leq \frac{1}{2\pi}\int_{\mathbb{R}}\sqrt{\mathbb{E}\left(\left(\sup_{u\in I_0}\big|\omega_{q,Tn}(u,x)\big|\right)^2\right)}\cdot\Big|\kappa^{\mathrm{ft}}(xh_\mu)\Big|\,dx\\
&= \frac{1}{2\pi}\int_{\mathbb{R}}\sqrt{\mathbb{E}\left(\sup_{u\in I_0}\ \left(\omega_{q,Tn}(u,x)\right)^2\right)}\cdot\Big|\kappa^{\mathrm{ft}}(xh_\mu)\Big|\,dx.
\end{aligned} \tag{36}$$

In order to simplify the notation we will denote

$$W_{it}^q(x,u) = \cos\big(xU_{it}\big)\psi_q\left(U_{it}-u,Y_{it}\right),$$

such that $\mathbb{E}\left(\sup_{u\in I_0}\ \left(\omega_{q,Tn}(u,x)\right)^2\right) =$

$$\begin{aligned}
&\mathbb{E}\Bigg(\sup_{u\in I_0}\Bigg(\frac{1}{(Tn)^2}\sum_{it}\left[W_{it}^q(x,u)\mathbb{1}_{(|U_{it}-u|<h_\mu)}-\mathbb{E}(W_{it}^q(x,u)\mathbb{1}_{(|U_{it}-u|<h_\mu)})\right]^2 +\\
&\quad\frac{1}{(Tn)^2}\sum_{(i,t)\neq(j,s)}\Big[(W_{it}^q(x,u)\mathbb{1}_{(|U_{it}-u|<h_\mu)}-\mathbb{E}(W_{it}^q(x,u)\mathbb{1}_{(|U_{it}-u|<h_\mu)}))\cdot\\
&\quad\cdot(W_{js}^q(x,u)\mathbb{1}_{(|U_{js}-u|<h_\mu)}-\mathbb{E}(W_{js}^q(x,u)\mathbb{1}_{(|U_{js}-u|<h_\mu)}))\Big]\Bigg)\Bigg).
\end{aligned}$$

As u takes only values within the compact interval $[a,b]$, there exist constants C_1 and C_2 such that, uniformly for all $u \in [a,b]$, $\mathbb{P}(|U_{it}-u| < h_\mu) \leq C_1 h_\mu < \infty$, for all i,t, and $\mathbb{P}(|U_{it}-u| < h_\mu$ AND $|U_{js}-u| < h_\mu) \leq C_2 h_\mu^2 < \infty$, for all $(i,t) \neq (j,s)$. Together with the triangle inequality, this yields that $\mathbb{E}\left(\sup_{u \in I_0} \left(\omega_{q,Tn}(u,x)\right)^2\right) \leq$

$$\frac{C_1 h_\mu}{(Tn)^2} \sum_{it} \mathbb{E}\left(\sup_{u \in I_0} \left[W_{it}^q(x,u) - \mathbb{E}(W_{it}^q(x,u))\right]^2\right) +$$
$$\frac{C_2 h_\mu^2}{(Tn)^2} \sum_{(i,t)\neq(j,s)} \mathbb{E}\left(\sup_{u \in I_0} \left[(W_{it}^q(x,u) - \mathbb{E}(W_{it}^q(x,u)))(W_{js}^q(x,u) - \mathbb{E}(W_{js}^q(x,u)))\right]\right).$$

From our moment assumptions (Assumption A1) and the fact that $I_0 = [a,b]$ is compact we can conclude that there must exist a constant C_3 such that, point-wise for every $x \in \mathbb{R}$,

$$\mathbb{E}\left(\left(\sup_{u \in I_0} \left|W_{it}^q(x,u) - \mathbb{E}(W_{it}^q(x,u))\right|\right)^2\right) \leq C_3 < \infty \tag{37}$$

for all i,t.

Within function dependencies: By the same reasoning there must exist a constant C_4 such that, point-wise for every $x \in \mathbb{R}$,

$$\begin{aligned} &\mathbb{E}\left(\sup_{u \in I_0} \left|W_{it}^q(x,u) - \mathbb{E}(W_{it}^q(x,u))\right| \cdot \sup_{u \in I_0} \left|W_{jt}^q(x,u) - \mathbb{E}(W_{jt}^q(x,u))\right|\right) \\ &\leq C_4 < \infty \end{aligned} \tag{38}$$

for all $i \neq j$ and all t.

Between function dependencies: Our weak dependency assumption (Assumption A2) and the fact that I_0 is compact yields that point-wise for every $x \in \mathbb{R}$

$$\begin{aligned} &\mathbb{E}\left(\sup_{u \in I_0} \left|W_{it}^q(x,u) - \mathbb{E}(W_{it}^q(x,u))\right| \cdot \sup_{u \in I_0} \left|W_{js}^q(x,u) - \mathbb{E}(W_{js}^q(x,u))\right|\right) \\ &\leq c r^{|t-s|} \end{aligned} \tag{39}$$

for all i,j and $|t-s| \geq 1$, where $0 < c < \infty$ and $0 < r < 1$.

Eq.s (37), (38), and (39) yield that $\mathbb{E}\left(\sup_{u \in I_0} \left(\omega_{q,Tn}(u,x)\right)^2\right) \leq$

$$\begin{aligned} &\leq \frac{C_1 h_\mu}{(Tn)^2} \sum_{it} C_3 + \frac{C_2 h_\mu^2}{(Tn)^2} \sum_{i \neq j, t} C_4 + \frac{C_2 h_\mu^2}{(Tn)^2} \sum_{i,j,t \neq s} c r^{|t-s|} \\ &= \mathcal{O}\left(\frac{h_\mu}{Tn} + \frac{h_\mu^2 (n-1)}{Tn} + \frac{h_\mu^2}{T}\right) = \mathcal{O}\left(\frac{h_\mu}{Tn} + \frac{h_\mu^2}{T}\right), \end{aligned}$$

such that

$$\sqrt{\mathbb{E}\left(\sup_{u\in I_0}\left(\omega_{q,Tn}(u,x)\right)^2\right)} = \mathcal{O}\left(\sqrt{\frac{h_\mu}{Tn}} + \frac{h_\mu}{\sqrt{T}}\right). \tag{40}$$

Plugging (40) into (36) and integration by substitution leads to

$$\begin{aligned}\mathbb{E}(\sup_{u\in I_0}|\Psi_{q,Tn}(u;h_\mu) - \mathbb{E}(\Psi_{q,Tn}(u;h_\mu))|) = \\ \leq \frac{1}{2\pi}\int_{\mathbb{R}}\sqrt{\mathbb{E}\left(\sup_{u\in I_0}\left(\omega_{q,Tn}(u,x)\right)^2\right)}\cdot\left|\kappa^{\text{ft}}(xh_\mu)\right|dx = \mathcal{O}\left(\frac{1}{\sqrt{Tn\,h_\mu}} + \frac{1}{\sqrt{T}}\right).\end{aligned} \tag{41}$$

Let us now focus on the first summand in (33). From standard arguments in nonparametric statistics we know that

$$\mathbb{E}(\Psi_{q,Tn}(u;h_\mu)) - m_q(u) = \mathcal{O}(h_\mu^2)$$

for each $u \in I_0$ and for all $q \in \{0,\dots,4\}$. Under Assumption A3, the "$\mathcal{O}(h_\mu^2)$" terms become uniformly valid for all $u \in I_0$, since for $q \in \{0,1,2,4\}$, we have then $\sup_{\in I_0}|f_U''(u)| \leq c < \infty$ and additionally for $q \in \{3,4\}$, we have then $\sup_{(u,y)\in I_0\times\mathbb{R}}|(\partial^2/\partial u^2)g_{YU}(y,u)| \leq c < \infty$, where $c > 0$ are generic constants (see Assumption A3). We can conclude with respect to the first summand in (33) that

$$\sup_{u\in I_0}|\mathbb{E}(\Psi_{q,Tn}(u;h_\mu)) - m_p(u)| = \mathcal{O}(h_\mu^2) \quad \text{for all} \quad q \in \{0,\dots,4\}. \tag{42}$$

Finally, plugging our results (41) and (42) into (33) leads to

$$\tau_{q,Tn} = \sup_{u\in I_0}|\Psi_{q,Tn}(u;h_\mu) - m_q(u)| = \mathcal{O}_p\left(h_\mu^2 + \frac{1}{\sqrt{Tn\,h_\mu}} + \frac{1}{\sqrt{T}}\right) \tag{43}$$

for all $q \in \{0,\dots,4\}$. □

The proof of Lemma B.2: Analogously to that of B.1.

The proof of Theorem 4.1, parts (c) and (d) Under the additional regularity Assumption A9 it follows from the eigenvalue and eigenfunction expansions in Hall and Hosseini-Nasab (2006), that with probability 1,

$$\sup_{k\geq 1}|\hat{\lambda}_k - \lambda_k| \leq \hat{\Delta}_{TN} \quad \text{and}$$

$$\sup_{u\in I_0}|\hat{\phi}_k(u) - \phi_k(u)| \leq \sqrt{8}\frac{\hat{\Delta}_{TN}}{\delta_k} \quad \text{for all} \quad 1 \leq k \leq \bar{K}_{TN} - 1,$$

where $\hat{\Delta}_{TN} = \sqrt{\int_{(u,v)\in I_0^2}(\hat{\gamma}(u,v)-\gamma(u,v))^2 d(u,v)}$, $\delta_k = \min_{1\leq i\leq k}\{\lambda_i - \lambda_{i+1}\}$, and $\bar{K}_{TN} = \inf\{k\geq 1 : \lambda_k - \lambda_{k+1} \leq 2\hat{\Delta}_{TN}\}$. See also Theorem 2 in the in Appendix A.1. of Hall and Hosseini-Nasab (2006) and its discussion.

From part (b) of our Theorem 4.1, it follows that $\hat{\Delta}_{TN} = \mathcal{O}_p\left(h_\gamma^2 + \frac{1}{\sqrt{TN\,h_\gamma^2}} + \frac{1}{\sqrt{T}}\right)$ which leads to

$$\sup_{k\geq 1}|\hat{\lambda}_k - \lambda_k| = \mathcal{O}_p\left(h_\gamma^2 + \frac{1}{\sqrt{TN\,h_\gamma^2}} + \frac{1}{\sqrt{T}}\right) \quad \text{and} \tag{44}$$

$$\sup_{u\in I_0}|\hat{\phi}_k(u) - \phi_k(u)| = \mathcal{O}_p\left(\delta_k^{-1}\left(h_\gamma^2 + \frac{1}{\sqrt{TN\,h_\gamma^2}} + \frac{1}{\sqrt{T}}\right)\right), \tag{45}$$

where (45) holds for all $1 \leq k \leq \bar{K}_{TN} - 1$. This concludes our proof of Theorem 4.1, parts (c) and (d).

The proof of Theorem 4.2.
Estimating the pc-scores:

$$\hat{\xi}_{kt}^{\mathrm{sm}} = \sum_{i=1}^{n}\hat{\phi}_k^{\mathrm{sm}}(U_{it}^{\mathrm{sm}})(Y_{it}^{\mathrm{sm}} - \hat{\mu}(U_{it}^{\mathrm{sm}};h_\mu))(U_{it}^{\mathrm{sm}} - U_{i-1,t}^{\mathrm{sm}}), \text{ with } U_{0,t}^{\mathrm{sm}} = A_t.$$

Using that $Y_{it}^{\mathrm{sm}} = X_t^{\mathrm{sm}}(U_{it}^{\mathrm{sm}}) + \varepsilon_{it}$, a weak law of large numbers, and results (a) and (d) of Theorem 4.1 lead to

$$\begin{aligned}\hat{\xi}_{kt}^{\mathrm{sm}} &= \int_{A_t}^{B_t}\hat{\phi}_k^{\mathrm{sm}}(u)(X_t^{\mathrm{sm}}(u) - \hat{\mu}(u;h_\mu))\,du + \mathcal{O}_p\left(n^{-1/2}\right)\\ &= \xi_{kt}^{\mathrm{sm}} + \mathcal{O}_p\left(n^{-1/2} + r_{Tn}^{\mu} + \frac{r_{TN}^{\gamma}}{\delta_k}\right),\end{aligned} \tag{46}$$

where $r_{Tn}^{\mu} = h_\mu^2 + 1/\sqrt{Tn\,h_\mu} + 1/\sqrt{T}$ and $r_{TN}^{\gamma} = h_\gamma^2 + 1/\sqrt{TN\,h_\gamma^2} + 1/\sqrt{T}$.

Estimating the functional prediction model: Let us denote

$$\widetilde{X}_{t,K}^{\mathrm{la}}(u) = \mu(u) + \sum_{k=1}^{K}\frac{\xi_{tk}^{\mathrm{sm}}}{\lambda_k^{\mathrm{sm}}}\int_{I_{\mathrm{sm}}}\phi_k^{\mathrm{sm}}(v)\gamma^{\mathrm{la}}(u,v)dv \quad \text{and}$$

$$\widetilde{X}_{t,K_+}^{c,\mathrm{la}}(u) = \sum_{k=K+1}^{\infty}\frac{\xi_{tk}^{\mathrm{sm}}}{\lambda_k^{\mathrm{sm}}}\int_{I_{\mathrm{sm}}}\phi_k^{\mathrm{sm}}(v)\gamma^{\mathrm{la}}(u,v)dv$$

such that $\widetilde{X}_t^{\mathrm{la}}(u) = \widetilde{X}_{t,K}^{\mathrm{la}}(u) + \widetilde{X}_{t,K_+}^{c,\mathrm{la}}(u)$. This allows us to decompose the prediction error into an estimation error part and an regularization error part:

(47)
$$\sup_{u\in I_0}\left|\widetilde{X}_t^{\mathrm{la}}(u) - \hat{\widetilde{X}}_{t,K}^{\mathrm{la}}(u)\right| \leq \sup_{u\in I_0}\left|\widetilde{X}_{t,K}^{\mathrm{la}}(u) - \hat{\widetilde{X}}_{t,K}^{\mathrm{la}}(u)\right| + \sup_{u\in I_0}\left|\widetilde{X}_{t,K_+}^{c,\mathrm{la}}(u)\right|.$$

Let us focus on the estimation error, i.e., the first term on the right hand side of (47). Using the results (a) and (d) of Theorem 4.1 and from (46)

$$\sup_{u\in I_0}\left|\widetilde{X}_{t,K}^{\mathrm{la}}(u) - \hat{\widetilde{X}}_{t,K}^{\mathrm{la}}(u)\right| = \mathcal{O}_p\left(K\left(n^{-1/2} + r_{Tn}^{\mu}\right) + r_{TN}^{\gamma}\sum_{k=1}^{K}\delta_k^{-1}\right). \tag{48}$$

In order to approximate the regularization error, i.e., the second term on the right hand side of (47), we use that, by the boundedness of γ and of the eigenfunctions, there exists a constant C, $0 < C < \infty$, such that

$$\sup_{u\in I_{\mathrm{la}}}\left|\int_{I_{\mathrm{sm}}}\phi_k^{\mathrm{sm}}(v)\gamma^{\mathrm{la}}(u,v)dv\right| \leq \lambda^{\mathrm{sm}}\,C.$$

The later, the triangle inequality, and Jensen's inequality yield that

$$\begin{aligned}\mathbb{E}\left(\sup_{u\in I_0}\left|\widetilde{X}_{t,K_+}^{c,\mathrm{la}}(u)\right|\right) &= \mathbb{E}\left(\sup_{u\in I_0}\left|\sum_{k=K+1}^{\infty}\frac{\xi_{tk}^{\mathrm{sm}}}{\lambda_k^{\mathrm{sm}}}\int_{I_{\mathrm{sm}}}\phi_k^{\mathrm{sm}}(v)\gamma^{\mathrm{la}}(u,v)dv\right|\right)\\ &\leq C\sum_{k=K+1}^{\infty}\sqrt{\mathbb{E}\,|\xi_{tk}^{\mathrm{sm}}|^2} = \mathcal{O}\left(\sum_{k=K+1}^{\infty}\sqrt{\lambda_k^{\mathrm{sm}}}\right).\end{aligned}$$

Therefore,

$$\sup_{u\in I_0}\left|\widetilde{X}_{t,K_+}^{c,\mathrm{la}}(u)\right| = \mathcal{O}_p\left(\sum_{k=K+1}^{\infty}\sqrt{\lambda_k^{\mathrm{sm}}}\right). \tag{49}$$

Under our moment assumption (Assumption A1), we know that $\sum_{k=1}^{\infty}\sqrt{\lambda_k^{\mathrm{sm}}} < \infty$, which implies that

$$\begin{aligned}&\sup_{u\in I_0}\left|\widetilde{X}_{t,K_+}^{c,\mathrm{la}}(u)\right| = \mathcal{O}_p(1) \quad \text{for fixed } K \text{ and that}\\ &\sup_{u\in I_0}\left|\widetilde{X}_{t,K_+}^{c,\mathrm{la}}(u)\right| = o_p(1) \quad \text{for } K\to\infty.\end{aligned}$$

From (48) and (49), we can conclude that

$$\sup_{u\in I_0}\left|\widetilde{X}_t^{\mathrm{la}}(u)-\hat{\widetilde{X}}_{t,K}^{\mathrm{la}}(u)\right| =$$
$$\mathcal{O}_p\left(K\,\left(n^{-1/2}+r_{Tn}^{\mu}\right)+r_{TN}^{\gamma}\sum_{k=1}^{K}\delta_k^{-1}\right)+\mathcal{O}_p\left(\sum_{k=K+1}^{\infty}\sqrt{\lambda_k^{\mathrm{sm}}}\right). \tag{50}$$

Case-by-case considerations under optimal bandwidth choices: Let us in the following focus on the case of optimal bandwidth choices, i.e., $h_\mu \sim (Tn)^{-1/5}$ and $h_\gamma \sim (TN)^{-1/6}$. Then we have that for $0<\theta<1/4$: $r_{Tn}^{\mu} \sim (Tn)^{-2/5}$ and $r_{TN}^{\gamma}\sim (TN)^{-1/3}$, which implies that $r_{Tn}^{\mu}=o\left(r_{TN}^{\gamma}\right)$, but also that $r_{TN}^{\gamma}=o\left(n^{-1/2}\right)$. These considerations yield to

$$\sup_{u\in I_0}\left|\widetilde{X}_t^{\mathrm{la}}(u)-\hat{\widetilde{X}}_{t,K}^{\mathrm{la}}(u)\right| =$$
$$=\mathcal{O}_p\left(K\,n^{-1/2}+(TN)^{-1/3}\sum_{k=1}^{K}\delta_k^{-1}\right)+\mathcal{O}_p\left(\sum_{k=K+1}^{\infty}\sqrt{\lambda_k^{\mathrm{sm}}}\right). \tag{51}$$

Under optimal bandwidth choices and for $1/4\leq\theta<1$ we have that $r_{Tn}^{\mu}\sim r_{TN}^{\gamma}\sim T^{-1/2}$ and that $T^{-1/2}=o\left(n^{-1/2}\right)$. These considerations yield to

$$\sup_{u\in I_0}\left|\widetilde{X}_t^{\mathrm{la}}(u)-\hat{\widetilde{X}}_{t,K}^{\mathrm{la}}(u)\right| =$$
$$\mathcal{O}_p\left(K\,n^{-1/2}+T^{-1/2}\sum_{k=1}^{K}\delta_k^{-1}\right)+\mathcal{O}_p\left(\sum_{k=K+1}^{\infty}\sqrt{\lambda_k^{\mathrm{sm}}}\right). \tag{52}$$

Under optimal bandwidth choices and for $1\leq\theta<\infty$ we have that $r_{Tn}^{\mu}\sim r_{TN}^{\gamma}\sim T^{-1/2}$ and that $n^{-1/2}=\mathcal{O}\left(T^{-1/2}\right)$. But note also that $KT^{-1/2}=\mathcal{O}(T^{-1/2}\sum_{k=1}^{K}\delta_k^{-1})$. These considerations yield to

$$\sup_{u\in I_0}\left|\widetilde{X}_t^{\mathrm{la}}(u)-\hat{\widetilde{X}}_{t,K}^{\mathrm{la}}(u)\right| =$$
$$\mathcal{O}_p\left(T^{-1/2}\sum_{k=1}^{K}\delta_k^{-1}\right)+\mathcal{O}_p\left(\sum_{k=K+1}^{\infty}\sqrt{\lambda_k^{\mathrm{sm}}}\right). \tag{53}$$

Deriving explicit rates:
From Assumption A8 we have that $\lambda_k^{\mathrm{sm}} = \mathcal{O}(k^{-a})$, with $a > 1$, and $\lambda_k^{\mathrm{sm}} - \lambda_{k+1}^{\mathrm{sm}} \geq \text{const.} \times k^{-a-1}$ or equivalently, $(\lambda_k^{\mathrm{sm}} - \lambda_{k+1}^{\mathrm{sm}})^{-1} \leq (\text{const.} \times k^{a+1})$. From the latter it follows that

$$(\lambda_k^{\mathrm{sm}} - \lambda_{k+1}^{\mathrm{sm}})^{-1} \leq (\text{const.} \times k^{a+1}) \leq (\text{const.} \times K^{a+1}), \text{ for } 1 \leq k \leq K.$$

Observe furthermore, that

$$\delta_k^{\mathrm{sm}} = \min_{1 \leq i \leq k} \{\lambda_i^{\mathrm{sm}} - \lambda_{i+1}^{\mathrm{sm}}\} \quad \Leftrightarrow \quad (\delta_k^{\mathrm{sm}})^{-1} = \max_{1 \leq i \leq k} \{(\lambda_i^{\mathrm{sm}} - \lambda_{i+1}^{\mathrm{sm}})^{-1}\}.$$

The latter two results yield that $(\delta_k^{\mathrm{sm}})^{-1} \leq \text{const.} \times k^{a+1}$, but also that $(\delta_k^{\mathrm{sm}})^{-1} \leq \text{const.} \times K^{a+1}$, for $1 \leq k \leq K$. The latter result implies that

$$\sum_{k=1}^{K} (\delta_k^{\mathrm{sm}})^{-1} \leq K(\text{const.} \times K^{a+1}) = \text{const.} \times K^{a+2}. \tag{54}$$

The assumption $\lambda_k^{\mathrm{sm}} = \mathcal{O}(k^{-a})$ leads to $\sum_{k=K+1}^{\infty} \sqrt{\lambda_k^{\mathrm{sm}}} = \mathcal{O}(\sum_{k=K+1}^{\infty} k^{-a/2})$. Using this and the approximation in Eq. (54) we get from Eq.s (51)-(53) our results (a)-(c) of Theorem 4.2. □

Dominik Liebl and Alois Kneip
Statistische Abteilung
University of Bonn
Adenauerallee 24-26
53113 Bonn, Germany
E-mail: dliebl@uni-bonn.de