Supplemental Material For
"Confidence surfaces for the mean of locally stationary functional time series"

Ruhr-Universität Bochum and Tsinghua University

Supplementary Material

Section S1 provides additional remarks regarding the noisy and multivariate locally stationary functional time series. Section S2 provides approaches to the selection of tuning parameters and simulation results, including the simulation results for both boundary and interior regions, and the simulation results checking the Gaussian approximation and the long-run variance function estimator. Section S3 discusses possible alternative assumptions for Theorem 2. Section S4 contains some details about simultaneous confidence bands for the regression function in model (2.1), where one of the arguments is fixed (Section S4.1) including additional numerical results for this case (see Sections S4.3 and S4.4). In Section S5 we provide examples of locally stationary functional processes, illustrating our approach of modeling non-stationary functional data. Section S6 contains the proof of all Theorems, while Section S7 provides propositions. Finally, Section S8 presents auxiliary results for the proofs.

S1 Additional remarks

In this section, we provide two remarks which briefly discuss how to build simultaneous confidence surface for noisy data and multivariate locally stationary functional time series using our method.

Remark S1. Indeed several authors consider (stationary) functional data models with noisy observation (see Cao et al., 2012; Chen and Song, 2015, among others) and we expect that the results presented in this section can be extended to this scenario. More precisely, consider the model

$$Y_{ij} = X_{i,n}(\tfrac{j}{N}) + \sigma(\tfrac{j}{N})z_{ij},\ 1 \leq i \leq n, 1 \leq j \leq N\ ,$$

where $X_{i,n}$ is the functional time series defined in (2.1), $\{z_{ij}\}_{i=1,\ldots,n,j=1,\ldots,N}$ is an array of centered independent identically distributed observations and $\sigma(\cdot)$ is a positive function on the interval $[0,1]$. This means that one can not observe the full trajectory of $\{X_{i,n}(t) \mid t \in [0,1]\}$, but only the function $X_{i,n}$ evaluated at the discrete time points $1/N, 2/N, \ldots, (N-1)/N, 1$ subject to some random error. If $N \to \infty$ as $n \to \infty$, and the regression function m in (2.1) is sufficiently smooth, we expect that we can construct simultaneous confidence bands and surfaces by applying the procedure described in this section to smoothed trajectories.

For example, we can consider the smooth estimate

$$\tilde{m}(u,\cdot) = \underset{g\in\mathcal{S}_p}{\operatorname{argmin}} \sum_{i=\lfloor nu-\sqrt{n}\rfloor}^{\lceil nu+\sqrt{n}\rceil} \sum_{j=1}^{N} \big(Y_{i,j} - g(\tfrac{j}{N})\big)^2\ ,$$

where $\mathcal{S}_p$ denotes the set of splines of order p, which depends on the smoothness of the function $t \to m(u,t)$. We can now construct confidence bands applying the methodology to the data $\tilde{X}_{i,n}(\cdot) = \tilde{m}(\frac{i}{\sqrt{n}},\cdot)$, $i = 1,\ldots,\sqrt{n}$ due to the asymptotic efficiency of the spline estimate (see Proposition 3.2-3.4 in Cao et al., 2012).

Alternatively, we can also obtain smooth estimates $t \to \check{X}_{i,n}(t)$ of the trajectory using local polynomials, and we expect that the proposed methodology applied to the data $\check{X}_{1,n}, \ldots, \check{X}_{n,n}$ will yield valid simultaneous confidence bands and surfaces, where the range for the variable t is restricted to the interval $[c_n, 1 - c_n]$ and c_n denotes the bandwidth of the local polynomial estimator used in smooth estimator of the trajectory.

Remark S2. The methodology presented so far can be extended to construct a simultaneous confidence surfaces for the vector of mean functions of a multivariate locally stationary functional time series. For simplicity we consider a 2-dimensional series of the form

$$\begin{pmatrix} X^1_{i,n}(t) \\ \\ X^2_{i,n}(t) \end{pmatrix} = \begin{pmatrix} m_1(\frac{i}{n}, t) \\ \\ m_2(\frac{i}{n}, t) \end{pmatrix} + \begin{pmatrix} \varepsilon^1_{i,n}(t) \\ \\ \varepsilon^2_{i,n}(t) \end{pmatrix} ,$$

and define for $a = 1, 2$

$$\begin{aligned} \hat{Z}_i^{a,\hat{\sigma}}(u) &= (\hat{Z}_{i,1}^{a,\hat{\sigma}}(u), \ldots, \hat{Z}_{i,p}^{a,\hat{\sigma}}(u))^\top \\ &= K\Big(\frac{\frac{i}{n} - u}{b_n}\Big)\Big(\frac{\hat{\varepsilon}^a_{i,n}(\frac{1}{p})}{\hat{\sigma}_a(\frac{i}{n}, \frac{1}{p})}, \frac{\hat{\varepsilon}^a_{i,n}(\frac{2}{p})}{\hat{\sigma}_a(\frac{i}{n}, \frac{2}{p})}, \ldots, \frac{\hat{\varepsilon}^a_{i,n}(\frac{p-1}{p})}{\hat{\sigma}_a(\frac{i}{n}, \frac{p-1}{p})}, \frac{\hat{\varepsilon}^a_{i,n}(1)}{\hat{\sigma}_a(\frac{i}{n}, 1)}\Big)^\top, \end{aligned}$$

where $\hat{\varepsilon}^a_{i,n}(t) = X^a_{i,n}(t) - \hat{m}_a(\frac{i}{n}, t)$ and $\hat{\sigma}^2_a(\frac{i}{n}, t)$ is the estimator of long-variance of $\varepsilon^a_{i,n}$ defined in (2.16). Next we consider the $2(n - 2\lceil nb_n \rceil + 1)p$-dimensional vector

$$\hat{\hat{Z}}_j^{\hat{\sigma}} = \big(\hat{Z}^{\hat{\sigma},\top}_{j,\lceil nb_n \rceil}, \hat{Z}^{\hat{\sigma},\top}_{j+1,\lceil nb_n \rceil+1} \ldots, \hat{Z}^{\hat{\sigma},\top}_{n-2\lceil nb_n \rceil+j, n-\lceil nb_n \rceil}\big)^\top ,$$

where $\hat{Z}^{\hat{\sigma}}_{i,l} = \hat{Z}^{\hat{\sigma}}_i(\frac{l}{n}) = (\hat{Z}^{1,\hat{\sigma}}_{i,l,1}, \hat{Z}^{2,\hat{\sigma}}_{i,l,1} \ldots \hat{Z}^{1,\hat{\sigma}}_{i,l,p}, \hat{Z}^{2,\hat{\sigma}}_{i,l,p})^\top$ contains information from both components. Define for $a = 1, 2$

$$\hat{L}^{\hat{\sigma}}_{3,a}(u,t) = \hat{m}_a(u,t) - \hat{r}_{3,a}(u,t), \quad \hat{U}^{\hat{\sigma}}_{3,a}(u,t) = \hat{m}_a(u,t) + \hat{r}_{3,a}(u,t) \ ,$$

where

$$\hat{r}_{3,a}(u,t) = \frac{\hat{\sigma}_a(u,t)\sqrt{2}T^{\hat{\sigma}}_{\lfloor(1-\alpha)B\rfloor}}{\sqrt{nb_n}\sqrt{2\lceil nb_n\rceil - m'_n}}$$

and $T^{\hat{\sigma}}_{\lfloor(1-\alpha)B\rfloor}$ is generated in the same way as in step (d) of Algorithm 2 with p replaced by $2p$, $\hat{m}_a$ is the kernel estimator of m_a defined in (2.3). Further, define for $a = 1, 2$ the set of functions

$$\mathcal{C}^{\hat{\sigma}}_{a,n} = \big\{ f \in \mathcal{C}^{3,0} : [0,1]^2 \to \mathbb{R} \mid \ \hat{L}_{3,a}(u,t) \leq f(u,t) \leq \hat{U}_{3,a}(u,t) \\ \forall u \in [b_n, 1-b_n] \ \forall t \in [0,1] \big\}.$$

Suppose that the mean functions and error processes of $X^1_{i,n}(t)$ and $X^2_{i,n}(t)$ satisfy the conditions of Theorem 4, then it can be proved that the set $\mathcal{C}^{\hat{\sigma}}_{1,n} \times \mathcal{C}^{\hat{\sigma}}_{2,n}$ defines an asymptotic $(1-\alpha)$ simultaneous confidence surface for the vector function $(m_1, m_2)^\top$. The details are omitted for the sake of brevity.

S2 Finite Sample Performance

In this section we study the finite sample performance of the simultaneous confidence surfaces proposed in the previous sections. We start giving some more details regarding the general implementation of the algorithms, and present the simulation study.

S2.1 Implementation

For the estimator of the regression function in (2.3) we use the kernel (of order 4) in $[b_n, 1-b_n]$

$$K(x) = (45/32 - 150x^2/32 + 105x^4/32)\mathbf{1}(|x| \le 1) ,$$

and for the boundary we use the kernel function $K_l(x) = (420x^2 - 480x + 120)x(1 - x)\mathbf{1}(0 \le x \le 1)$. We choose the bandwidth as the minimizer of

$$MGCV(b) = \max_{1 \le s \le p} \frac{\sum_{i=1}^{n}(\hat{m}_b(\frac{i}{n}, \frac{s}{p}) - X_{i,n}(\frac{s}{p}))^2}{(1 - \mathrm{tr}(Q_s(b))/n)^2} , \tag{S2.1}$$

$Q_s(b)$ is an $n \times n$ matrix such that

$$\big(\hat{m}_b(\tfrac{1}{n}, \tfrac{s}{p}), \hat{m}_b(\tfrac{2}{n}, \tfrac{s}{p}), \ldots, \hat{m}_b(1, \tfrac{s}{p})\big)^\top = Q_s(b)\big(X_{1,n}(\tfrac{s}{p}), \ldots, X_{n,n}(\tfrac{s}{p})\big)^\top.$$

Here $\hat{m}_b(u,t)$ is the NW estimator with bandwidth b defined in (2.3).

The criterion (S2.1) is motivated by the generalized cross validation criterion introduced by Craven and Wahba (1978) and will be called Maximal Generalized Cross Validation (MGCV) method throughout this paper.

For the estimator of the long-run variance in (2.16) we use $w = \lfloor n^{2/7} \rfloor$ and $\tau_n = n^{-1/7}$ as recommended in Dette and Wu (2019). The window size in the multiplier bootstrap is then selected by the minimal volatility method advocated by Politis et al. (1999). For the sake of brevity, we discuss this method only for Algorithm 2 in detail (the method for Algorithm 1 is similar). We consider a grid of window sizes $\tilde{m}_1 < \ldots <$

$\tilde{m}_M$ (for some integer M). We first calculate $\hat{S}^{\hat{\sigma}}_{j\tilde{m}_s} = (\hat{S}^{\hat{\sigma}}_{j\tilde{m}_s,r}, 1 \leq r \leq (n-2\lceil nb_n\rceil+1)p)$ defined in step (c) of Algorithm 2 for each $\tilde{m}_s$. Let $\hat{S}^{\hat{\sigma},\diamond}_{\tilde{m}_s}$ denote the $(n-2\lceil nb_n\rceil+1)p$ dimensional vector with r_{th} entry defined by

$$\hat{S}^{\hat{\sigma},\diamond}_{\tilde{m}_s,r} = \frac{1}{2\lceil nb_n\rceil - \tilde{m}_s} \sum_{j=1}^{2\lceil nb_n\rceil-\tilde{m}_s} (\hat{S}^{\hat{\sigma}}_{j\tilde{m}_s,r})^2 \ ,$$

and consider the standard error of $\{\hat{S}^{\hat{\sigma},\diamond}_{\tilde{m}_s,r}\}_{s=k-2}^{k+2}$, that is

$$\mathrm{se}\big(\{\hat{S}^{\hat{\sigma},\diamond}_{\tilde{m}_s,r}\}_{s=k-2}^{k+2}\big) = \Big(\frac{1}{4}\sum_{s=k-2}^{k+2}\Big(\hat{S}^{\hat{\sigma},\diamond}_{\tilde{m}_s,r} - \frac{1}{5}\sum_{s=k-2}^{k+2}\hat{S}^{\hat{\sigma},\diamond}_{\tilde{m}_s,r}\Big)^2\Big)^{1/2}.$$

Then we choose $m_n' = \tilde{m}_j$ where j is defined as the minimizer of the function

$$MV(k) = \frac{1}{(n-2\lceil nb_n\rceil+1)p} \sum_{r=1}^{(n-2\lceil nb_n\rceil+1)p} se\big(\{\hat{S}^{\hat{\sigma},\diamond}_{\tilde{m}_s,r}\}_{s=k-2}^{k+2}\big)$$

in the set $\{3, \ldots, M-2\}$. Throughout this section we consider $p = \lfloor\sqrt{n}\rfloor$.

S2.2 Simulated data

We consider two regression functions

$$m_1(u,t) = (u+2t)^2/2,$$

$$m_2(u,t) = (1+u^2)(6(t-0.5)^2(1+\mathbf{1}(t>0.3))+1)$$

(note that m_2 is discontinuous at the point $t = 0.3$). For the definition of the error processes let $\{\varepsilon_i\}_{i\in\mathbb{Z}}$ be a sequence of independent standard normally distributed random variables and $\{\eta_i\}_{i\in\mathbb{Z}}$ be a sequence of independent t-distributed random variables with

8 degrees of freedom. Define the functions

$$a(t) = 0.5\cos(\pi t/3), \qquad b(t) = 0.4t, \qquad c(t) = 0.3t^2,$$

$$d_1(t) = 1 + 0.5\sin(\pi t), \qquad d_{2,1}(t) = 2t - 1, \qquad d_{2,2}(t) = 6t^2 - 6t + 1,$$

and $\mathcal{F}_i^1 = (\ldots, \varepsilon_{i-1}, \varepsilon_i)$, $\mathcal{F}_i^2 = (\ldots, \eta_{i-1}, \eta_i)$. We consider the following two locally stationary functional time series errors G_1 and G_2 are defined by

$$G_1(u, t, \mathcal{F}_i^1) = G_0(u, t, \mathcal{F}_i^1)d_1(t)/3, \text{where } G_0(u, t, \mathcal{F}_i^1) = (a(u) - 0.1t)G_0(u, t, \mathcal{F}_i^1) + \epsilon_i,$$

$$G_2(u, t, \mathcal{F}_i^1, \mathcal{F}_i^2) = \tilde{G}_1(u, \mathcal{F}_i^1)d_{2,1}(t)/2 + \tilde{G}_2(u, \mathcal{F}_i^2)d_{2,2}(t)/2$$

where the locally stationary time series $\tilde{G}_1$ and $\tilde{G}_2$ are defined as

$$\tilde{G}_1(u, \mathcal{F}_i^1) = a(u)\tilde{G}_1(u, \mathcal{F}_{i-1}^1) + \varepsilon_i, \ \tilde{G}_2(u, \mathcal{F}_i^2) = b(u)\tilde{G}_2(u, \mathcal{F}_{i-1}^2) + \eta_i - c(u)\eta_{i-1}.$$

Note that $\tilde{G}_1$ is a locally stationary AR(1) process (or equivalently a locally stationary MA(∞) process), and that $\tilde{G}_2$ is a locally stationary ARMA(1, 1) model. With these processes we define the following functional time series model (for $1 \le i \le n$, $0 \le t \le 1$)

$$\text{(a)} \quad X_{i,n}(t) = m_1(\tfrac{i}{n}, t) + G_1(\tfrac{i}{n}, t, \mathcal{F}_i^1) \quad \text{(b)} X_{i,n}(t) = m_1(\tfrac{i}{n}, t) + G_2(\tfrac{i}{n}, t, \mathcal{F}_i^1, \mathcal{F}_i^2)$$

$$\text{(c)} \quad X_{i,n}(t) = m_2(\tfrac{i}{n}, t) + G_1(\tfrac{i}{n}, t, \mathcal{F}_i^1) \quad \text{(d)} X_{i,n}(t) = m_2(\tfrac{i}{n}, t) + G_2(\tfrac{i}{n}, t, \mathcal{F}_i^1, \mathcal{F}_i^2).$$

In Figure S1 we display typical 95% simultaneous confidence surfaces of the form (2.2) from one simulation run for model (a) with sample size $n = 800$ and $B = 1000$ bootstrap replications, which are calculated by Algorithm 1 (constant width) and Algorithm 2 (varying width). We observe that there exist differences between the surfaces with constant and variable width, but they are not substantial.

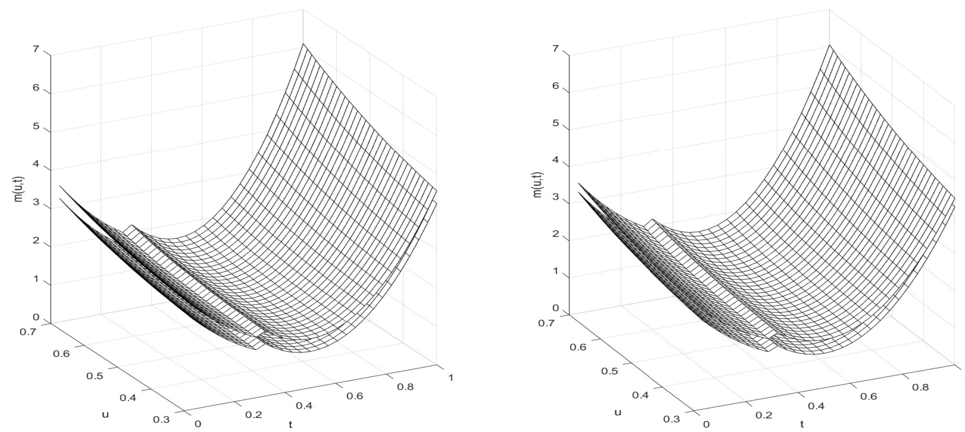

Figure S1: 95% *simultaneous confidence surfaces* (2.13) *and* (2.18) *for the regression function in model (c) from* $n = 800$ *observations. Left panel: constant width (Algorithm 1); Right panel: varying width (Algorithm 2)*

We next investigate the coverage probabilities of the different surfaces constructed in this paper for sample sizes $n = 500$ and $n = 800$. All results are based on 1000 simulation runs and $B = 1000$ bootstrap replications. The left part of Table S1 shows the coverage probabilities of the surfaces with constant width while the results in the right part correspond to the bands with varying width. We observe that the simulated coverage probabilities are close to their nominal levels in all cases under consideration, which illustrates the validity of our methods for finite sample sizes.

We conclude this section mentioning that confidence bands for the regression function m for a fixed u or a fixed t can be constructed in a similar manner and details and some additional numerical results for these bands are discussed in Section S4.

Table S1: *Simulated coverage probabilities of the simultaneous confidence bands* (2.13) *and* (2.18) *calculated by Algorithm 1 (constant width) and Algorithm 2 (varying width), respectively.*

	constant width				varying width			
	model (a)		model (b)		model (a)		model (b)	
level	90%	95%	90%	95%	90%	95%	90%	95%
n=500	88.0%	94.2%	90.1%	93.8%	91.2%	95.3%	87.9%	93.6%
n=800	89.9%	95.8%	88.3%	93.9%	90.9%	96.1%	90.7%	96.0%
	model (c)		model (d)		model (c)		model (d)	
level	90%	95%	90%	95%	90%	95%	90%	95%
n=500	87.9%	93.9%	91.3%	95.4%	87.5%	95.1%	87.7%	94.8%
n=800	88.6%	94.2%	89.9%	95.9%	90.8%	95.0%	90.1%	94.9%

S2.3 Simulation results in the boundary

We examine the proposed method for the simultaneous inference in the boundary region in Remark 2. We summarize our results in table S2, and find that our method in boundary works reasonably well.

Table S2: Simulated coverage probabilities of simultaneous confidence surface in the boundary using methods in Remark 2.

	constant width				varying width			
	model (a)		model (b)		model (a)		model (b)	
level	90%	95%	90%	95%	90%	95%	90%	95%
n=500	87.6%	93.9%	87.2%	93.7%	91.9%	96.0%	91.2%	96.2%
n=800	89.4%	94.4%	90.6%	96.1%	90.2%	95.4%	89.7%	95.0%
	model (c)		model (d)		model (c)		model (d)	
level	90%	95%	90%	95%	90%	95%	90%	95%
n=500	91.4%	95.1%	89.8%	94.6%	90.4%	95.5%	90.8%	95.3%
n=800	89.8%	95.7%	90.1%	94.9%	90.7%	95.4%	88.4%	94.1%

S2.4 Empirical investigation of Theorem 1

In this section we investigate the finite sample accuracy of the Gaussian approximation in Theorem 1. We consider model (d), the sample size $n = 500, 800$ and $b = 0.1, 0.2$ and compare the simulated quantiles of the maximum deviation of $\max_{\substack{b_n \leq u \leq 1-b_n \\ 0 \leq t \leq 1}} \sqrt{nb_n}|\hat{\Delta}(u,t)|$ and that of the maximum norm of the sum of corresponding high-dimensional Gaussian vectors with the auto-covariance structure described in Theorem 1. The results are presented in Figure S2, which shows that the approximation accuracy of Theorem 1 is quite high.

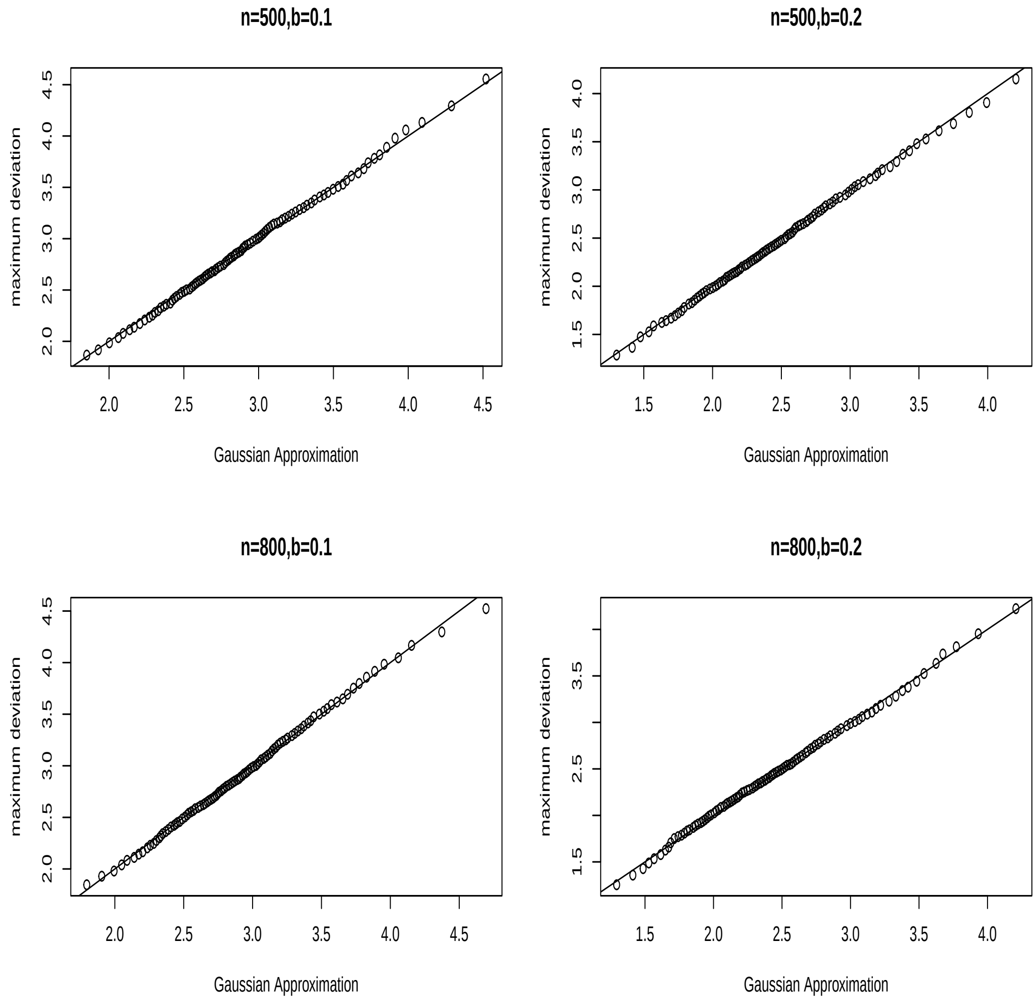

Figure S2: *Quantile-quantile plot of the* $\max_{b_n \le u \le 1-b_n, 0 \le t \le 1} \sqrt{nb_n}|\hat{\Delta}(u,t)|$ *versus* $\left|\frac{1}{\sqrt{nb_n}} \sum_{i=1}^{2\lceil nb_n \rceil -1} \tilde{Y}_i\right|_\infty$ *as described in Theorem 1.*

S2.5 Empirical performance of the long-run variance estimator

In this section we investigate the finite sample performance of the difference based long-run variance estimator (4.1). We examine the maximum error

$$\max_{1\leq i\leq n, 1\leq j\leq p} |\hat{\sigma}(i/n, j/p) - \sigma(i/n, j/p)| \tag{S2.2}$$

where $p = \lfloor n^{1/2} \rfloor$ as mentioned in Section S2.1. We consider model (a), (c), (b), (d) with sample size $n = 500$ and 800, respectively. The results are shown in Figure S3, where we display for each case the box plot of 2000 simulations of (S2.2). We observe that the estimator works reasonably well and in all simulation scenarios the estimation error decreases as the sample size increases.

S3 Discussion on the alternative assumptions of Theorem 2

In this section, we discuss alternative assumptions for Theorem 2. Some assumptions in the main paper can be relaxed yielding different approximation rates.

Remark S3.

(i) A careful inspection of the proofs in Section S6 shows that it is possible to prove similar results under alternative moment assumptions. For example, Theorem 1 holds under the assumption

$$\mathbb{E}\Big[\sup_{0\leq u,t\leq 1} (G(u, t, \mathcal{F}_0))^4\Big] < \infty \quad . \tag{S3.1}$$

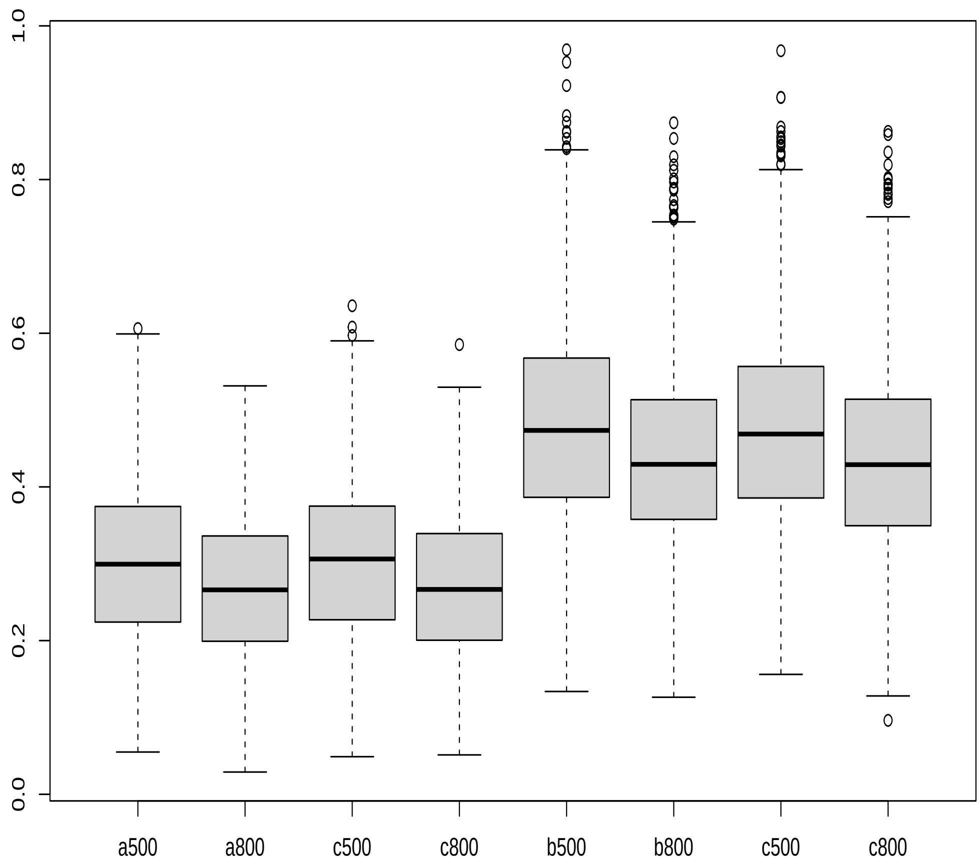

Figure S3: *Box plot of the simulated estimation error* (S2.2) *for model (a), (c), (b), (d) with sample size* 500 *and* 800*, respectively. The label* a500 *means model (a) for sample size* 500*. Other labels can be understood similarly.*

The details are omitted for the sake of brevity. Note that the sup in (S3.1) appears inside the expectation, while it appears outside the expectation in (3.2). Thus neither (3.2) implies (S3.1) nor vice versa.

(ii) Assumption 3.2(2) requires geometric decay of the dependence measure $\delta_q(G,i)$ and a careful inspection of the proofs in Section S6 shows that similar (but weaker) results can be obtained under less restrictive assumptions. To be precise, define $\Delta_{k,q} = \sum_{i=k}^{\infty} \delta_q(G,i)$, $\Xi_M = \sum_{i=M}^{\infty} i\delta_2(G,i)$ and consider the following assumptions.

(a) $\sum_{i=0}^{\infty} i\delta_3(G,i) < \infty$.

(b) There exist constants $M = M(n) > 0$, $\gamma = \gamma(n) \in (0,1)$ and $C_2 > 0$ such that

$$(2\lceil nb_n \rceil)^{3/8} M^{-1/2} {l'_n}^{-5/8} \geq C_2 l'_n$$

where $l'_n = \max(\log(2\lceil nb_n\rceil(n - 2\lceil nb_n\rceil + 1)p/\gamma), 1)$.

Then under the conditions of Theorem 1 with Assumption 3.2(2) replaced by (a) and (b), we have

$$\mathfrak{P}_n = O\Big(\eta'_n + \Theta\Big(\sqrt{nb_n}(b_n^4 + \frac{1}{n}), np\Big) + \Theta\Big(\big((np)^{1/q^*}((nb_n)^{-1} + 1/p)\big)^{\frac{q^*}{q^*+1}}, np\Big)\Big)$$

with

$$\eta_n' = (nb_n)^{-1/8}M^{1/2}{l'}_n^{7/8} + \gamma + \left((nb_n)^{1/8}M^{-1/2}{l'}_n^{-3/8}\right)^{q/(1+q)} \left(np\Delta_{M,q}^q\right)^{1/(1+q)}$$
$$+\Xi_M^{1/3}\left(1 \vee \log\left(np/\Xi_M\right)\right)^{2/3}.$$

The same arguments as given in the proof of Theorem 2 show that (under the other conditions in this theorem) the set $\mathcal{C}_n$ defined by (2.13) defines an (asymptotic) $(1-\alpha)$ simultaneous confidence surface if $\eta_n' = o(1)$. For example, if $\delta_q(G,i) = O(i^{-1-\alpha})$ for some $\alpha > 0$, $p = n^\beta$ for some $\beta > 0$ and $b_n = n^{-\gamma}$ for some $0 < \gamma < 1$, then $\eta_n' = o(1)$ if $(1+\beta) - (1-\gamma)q\alpha/4 < 0$, which gives a lower bound on q.

S4 Simultaneous confidence bands for fixed u or t

S4.1 Theoretical background and algorithms

In this section, we present the simultaneous confidence band for the regression function $(u,t) \to m(u,t)$ in model (2.1), where one of the arguments u and t is fixed. Let $\mathcal{C}^a$ be the class of functions with Lipschitz continuous a_{th} order derivatives with bounded Lipschitz constant. Consider

(1) simultaneous confidence bands *for fixed* t, which have the form

$$\mathcal{C}(t) = \left\{ f \in \mathcal{C}^3 \mid \hat{L}_1(u,t) \le f(u) \le \hat{U}_1(u,t) \;\; \forall u \right\}, \tag{S4.1}$$

where $\hat{L}_1$ and $\hat{U}_1$ are appropriate lower and upper bounds calculated from the data. As $t \in [0,1]$ is fixed these bounds can be derived generalizing results for confidence bands in nonparametric regression from the independent (see Konakov and Piterbarg, 1984; Xia, 1998; Proksch, 2014, among others) to the locally stationary case (see also Wu and Zhao, 2007, for results in a model with a stationary error process). An alternative approach based on multiplier bootstrap will be given below.

(2) simultaneous confidence bands *for fixed* u, which have the form

$$\mathcal{C}(u) = \left\{ f \in \mathcal{C}^0 \mid \ \hat{L}_2(u,t) \leq f(t) \leq \hat{U}_2(u,t) \ \ \forall t \in [0,1] \right\} , \tag{S4.2}$$

where $\hat{L}_2$ and $\hat{U}_2$ are appropriate lower and upper bounds calculated from the data. Note that these bounds can not be directly calculated using results of Dette et al. (2020) as these authors develop their methodology under the assumption of stationarity.

Recall the definition of the residuals $\hat{\varepsilon}_{i,n}(t)$ and the long-run variance estimator $\hat{\sigma}$ in the main article. For the construction of a simultaneous confidence bands for a fixed $t \in [0,1]$ of the form (S4.1) we define

$$\hat{Z}_i(u,t) = K\Big(\frac{\frac{i}{n}-u}{b_n}\Big)\hat{\varepsilon}_{i,n}(t), \hat{Z}_{i,l}(t) = \hat{Z}_i(\tfrac{l}{n},t),$$

$$\hat{Z}_i^{\hat{\sigma}}(u,t) = K\Big(\frac{\frac{i}{n}-u}{b_n}\Big)\frac{\hat{\varepsilon}_{i,n}(t)}{\hat{\sigma}(\frac{i}{n},t)}, \quad \hat{Z}_{i,l}^{\hat{\sigma}}(t) = \hat{Z}_i^{\hat{\sigma}}(\tfrac{l}{n},t).$$

Next we consider the $(n - 2\lceil nb_n \rceil + 1)$-dimensional vectors

$$\hat{\hat{Z}}_j(t) = \big(\hat{Z}_{j,\lceil nb_n \rceil}(t), \hat{Z}_{j+1,\lceil nb_n \rceil+1}(t), \ldots, \hat{Z}_{n-2\lceil nb_n \rceil+j,n-\lceil nb_n \rceil}(t)\big)^\top, \tag{S4.3}$$

$$\hat{\hat{Z}}_j^{\hat{\sigma}}(t) = \big(\hat{Z}^{\hat{\sigma}}_{j,\lceil nb_n \rceil}(t), \hat{Z}^{\hat{\sigma}}_{j+1,\lceil nb_n \rceil+1}(t), \ldots, \hat{Z}^{\hat{\sigma}}_{n-2\lceil nb_n \rceil+j,n-\lceil nb_n \rceil}(t)\big)^\top \tag{S4.4}$$

$(1 \leq j \leq 2\lceil nb_n \rceil - 1)$, then a simultaneous confidence band for fixed $t \in [0, 1]$ can be generated by the Algorithms S1 (constant width) and Algorithm S2 (varying width).

Algorithm S1:

Result: simultaneous confidence band of the form (S4.1) with fixed width

(a) Calculate the the $(n-2\lceil nb_n\rceil+1)$-dimensional vector $\hat{\hat{Z}}_j(t)$ in (S4.3);

(b) For window size m_n, let $m_n' = 2\lfloor m_n/2\rfloor$, define

$$\hat{S}_{jm_n'}(t) = \frac{1}{\sqrt{m_n'}}\sum_{r=j}^{j+\lfloor m_n/2\rfloor-1}\hat{\hat{Z}}_r(t) - \frac{1}{\sqrt{m_n'}}\sum_{r=j+\lfloor m_n/2\rfloor}^{j+m_n'-1}\hat{\hat{Z}}_r(t)$$

Let $\hat{\varepsilon}_{j:j+m_n',k}(t)$ be the k_{th} component of $\hat{S}_{jm_n'}(t)$.

(c) **for** r=1, ..., B **do**

- Generate independent standard normal distributed random variables $\{R_i^{(r)}\}_{i\in[1,n-m_n']}$. -

Calculate

$$T_k^{(r)}(t) = \sum_{j=1}^{2\lceil nb_n\rceil - m_n'}\hat{\varepsilon}_{j:j+m_n',k}(t)R_{k+j-1}^{(r)}\ ,\quad k=1,\dots,n-2\lceil nb_n\rceil+1,$$

$$T^{(r)}(t) = \max_{1\le k\le n-2\lceil nb_n\rceil+1}|T_k^{(r)}(t)|.$$

end

(d) Define $T_{\lfloor(1-\alpha)B\rfloor}(t)$ as the empirical $(1-\alpha)$-quantile of the sample $T^{(1)}(t),\dots,T^{(B)}(t)$ and

$$\hat{L}_3(u,t) = \hat{m}(u,t) - \hat{r}_3(t)\quad ,\qquad \hat{U}_3(u,t) = \hat{m}(u,t) + \hat{r}_3(t)$$

where

$$\hat{r}_3(t) = \frac{\sqrt{2}T_{\lfloor(1-\alpha)B\rfloor}(t)}{\sqrt{nb_n}\sqrt{2\lceil nb_n\rceil - m_n'}}$$

Output: $\mathcal{C}_n(t) = \big\{f\in\mathcal{C}^3 : [0,1]^2\to\mathbb{R}\ |\ \ \hat{L}_3(u,t)\le f(u)\le\hat{U}_3(u,t)\ \ \forall u\in[b_n,1-b_n]\big\}$

Algorithm S2:

Result: simultaneous confidence band of the form (S4.1) with varying width

(a) Calculate the estimate of the long-run variance $\hat{\sigma}^2$ in (2.16)

(b) Calculate the $(n - 2\lceil nb_n \rceil + 1)$-dimensional vectors $\hat{\tilde{Z}}_j^{\hat{\sigma}}(t)$ in (S4.4)

(c) For window size m_n, let $m_n' = 2\lfloor m_n/2 \rfloor$, define

$$\hat{S}^{\hat{\sigma}}_{jm_n'}(t) = \frac{1}{\sqrt{m_n'}} \sum_{r=j}^{j+\lfloor m_n/2 \rfloor - 1} \hat{\tilde{Z}}_r^{\hat{\sigma}}(t) - \frac{1}{\sqrt{m_n'}} \sum_{r=j+\lfloor m_n/2 \rfloor}^{j+m_n'-1} \hat{\tilde{Z}}_r^{\hat{\sigma}}(t)$$

Let $\hat{S}^{\hat{\sigma}}_{jm_n',k}(t)$ be the k_{th} component of $\hat{S}^{\hat{\sigma}}_{jm_n'}(t)$. Let $\hat{\varepsilon}^{\hat{\sigma}}_{j:j+m_n',k}(t)$ be the k_{th} component of $\hat{S}^{\hat{\sigma}}_{jm_n'}(t)$.

(d) **for** r=1, ..., B **do**

- Generate independent standard normal distributed random variables $\{R_i^{(r)}\}_{i \in [1, n-m_n']}$.

- Calculate

$$T_k^{\hat{\sigma},(r)}(t) = \sum_{j=1}^{2\lceil nb_n \rceil - m_n'} \hat{\varepsilon}^{\hat{\sigma}}_{j:j+m_n',k}(t) R^{(r)}_{k+j-1}\ , \quad k = 1, \ldots, n - 2\lceil nb_n \rceil + 1,$$

$$T^{\hat{\sigma},(r)}(t) = \max_{1 \le k \le n - 2\lceil nb_n \rceil + 1} |T_k^{\hat{\sigma},(r)}(t)|.$$

end

(e) Define $T^{\hat{\sigma}}_{\lfloor (1-\alpha)B \rfloor}(t)$ as the empirical $(1-\alpha)$-quantile of the sample $T^{\hat{\sigma},(1)}(t), \ldots, T^{\hat{\sigma},(B)}(t)$ and

$$\hat{L}_4^{\hat{\sigma}}(u,t) = \hat{m}(u,t) - \hat{r}_4(u,t), \qquad \hat{U}_4^{\hat{\sigma}}(u,t) = \hat{m}(u,t) + \hat{r}_4(u,t)$$

where

$$\hat{r}_4(u,t) = \frac{\hat{\sigma}(u,t)\sqrt{2} T^{\hat{\sigma}}_{\lfloor (1-\alpha)B \rfloor}(t)}{\sqrt{nb_n}\sqrt{2\lceil nb_n \rceil - m_n'}}$$

Output:

$$\mathcal{C}_n^{\hat{\sigma}}(t) = \left\{ f \in \mathcal{C}^3 : [0,1]^2 \to \mathbb{R} \mid \ \hat{L}_4^{\hat{\sigma}}(u,t) \le f(u) \le \hat{U}_4^{\hat{\sigma}}(u,t) \ \ \forall u \in [b_n, 1-b_n] \right\}.$$

The following result shows that the sets constructed by Algorithms S1 and S2 are asymptotic $(1-\alpha)$-confidence bands of the form (S4.1). The proof is similar to but easier than the proof of Theorems 2 and 3 is therefore omitted for the sake of brevity.

Theorem S1. *Assume that the conditions of Theorem 1 hold. Define*

$$\vartheta_n^\dagger = \frac{\log^2 n}{m_n} + \frac{m_n \log n}{nb_n} + \sqrt{\frac{m_n}{nb_n}} n^{4/q}.$$

(i) If $\vartheta_n^{\dagger,1/3}\{1 \vee \log(\frac{n}{\vartheta_n^\dagger})\}^{2/3} + \Theta\big(\big(\sqrt{m_n \log n}(\frac{1}{\sqrt{nb_n}} + b_n^3)(n)^{\frac{1}{q}}\big)^{q/(q+1)}, n\big) = o(1)$ *we have that for any* $\alpha \in (0,1)$ *and any* $t \in [0,1]$

$$\lim_{n\to\infty} \lim_{B\to\infty} \mathbb{P}(m \in \mathcal{C}_n(t) \mid \mathcal{F}_n) = 1 - \alpha$$

in probability.

(ii) If further the conditions of Theorem 3 and Proposition 1 hold, then

$$\lim_{n\to\infty} \lim_{B\to\infty} \mathbb{P}(m \in \mathcal{C}_n^{\hat{\sigma}}(t) \mid \mathcal{F}_n) = 1 - \alpha$$

in probability.

The next theorem presents a Gaussian approximation in the case where u is fixed. It is the basis for the construction of a confidence band for fixed u and its proof follows by similar (but easier) arguments as given in the proof of Theorem 1.

Theorem S2. *Let Assumptions 3.1 - 2.1 be satisfied and assume that the bandwidth in* (2.3) *satisfies that* $n^{1+a}b_n^9 = o(1)$, $n^{a-1}b_n^{-1} = o(1)$ *for some* $0 < a < 4/5$. *For any fixed* $u \in (0,1)$ *there exists a sequence of centered* p*-dimensional Gaussian vectors* $(Y_i(u))_{i\in\mathbb{N}}$

with the same covariance structure as the vector $Z_i(u)$ *in* (2.5)*, such that*

$$\mathfrak{P}_n(u) := \sup_{x\in\mathbb{R}} \Big| \mathbb{P}\Big(\max_{0\le t\le 1} \sqrt{nb_n} |\hat{\Delta}(u,t)| \le x \Big) - \mathbb{P}\Big(\Big| \frac{1}{\sqrt{nb_n}} \sum_{i=1}^{n} Y_i(u) \Big|_\infty \le x \Big) \Big|$$
$$= O\Big((nb_n)^{-(1-11\iota)/8} + \Theta\Big(\sqrt{nb_n} \Big(b_n^4 + \frac{1}{n} \Big), p \Big) + \Theta\Big(p^{\frac{1-q^*}{1+q^*}}, p \Big) \Big)$$

for any sequence $p \to \infty$ *with* $p = O(\exp(n^\iota))$ *for some* $0 \le \iota < 1/11$*. In particular,* $\mathfrak{P}_n(u) = o(1)$ *if* $p = n^c$ *for some* $c > 0$ *and the constant* q^* *in Assumption 3.3 is sufficiently large.*

Algorithm S3:

Result: simultaneous confidence band for fixed $u \in [b_n, 1-b_n]$ as defined in (S4.2)

(a) Calculate the p-dimensional vectors $\hat{Z}_i(u)$ in (2.1)

(b) For window size m_n, let $m_n' = 2\lfloor m_n/2 \rfloor$, define

$$\hat{S}_{jm_n'}(u) = \frac{1}{\sqrt{m_n'}} \sum_{r=j}^{j+\lfloor m_n/2\rfloor-1} \hat{Z}_r(u) - \frac{1}{\sqrt{m_n'}} \sum_{r=j+\lfloor m_n/2\rfloor}^{j+m_n'-1} \hat{Z}_r(u)$$

(c) **for** r=1, ..., B **do**

- Generate independent standard normal distributed random variables $\{R_i^{(r)}\}_{i=\lceil nu-nb_n\rceil}^{\lfloor nu+nb_n\rfloor}$ -

Calculate the bootstrap statistic

$$T^{(r)}(u) = \Big| \sum_{j=\lceil nu-nb_n\rceil}^{\lfloor nu+nb_n\rfloor-m_n'+1} \hat{S}_{jm_n'}(u) R_j^{(r)} \Big|_\infty$$

end

(d) Define $T_{\lfloor(1-\alpha)B\rfloor}(u)$ as the empirical $(1-\alpha)$-quantile of the sample $T^{(1)}(u), \dots, T^{(B)}(u)$ and

$$\hat{L}_5(u,t) = \hat{m}(u,t) - \hat{r}_5(u) \quad , \qquad \hat{U}_5(u,t) = \hat{m}(u,t) + \hat{r}_5(u) \ ,$$

where

$$\hat{r}_5(u) = \frac{\sqrt{2} T_{\lfloor(1-\alpha)B\rfloor}(u)}{\sqrt{nb_n}\sqrt{(\lfloor nu+nb_n\rfloor - \lceil nu-nb_n\rceil - m_n' + 2)}}$$

Output:

$$\mathcal{C}_n(u) = \big\{ f \in \mathcal{C}^0 : [0,1]^2 \to \mathbb{R} \mid \ \hat{L}_5(u,t) \le f(t) \le \hat{U}_5(u,t) \ \ \forall t \in [0,1] \big\}. \tag{S4.5}$$

Algorithm S4:

Result: simultaneous confidence band of the form (S4.2) with varying width.

(a) For given $u \in [b_n, 1-b_n]$, calculate the the estimate of the long-run variance $\hat{\sigma}^2(u,\cdot)$ in (2.16)

(b) Calculate the vector $\hat{Z}_i^{\hat{\sigma}_u}(u)$ in (S4.7);

(c) For window size m_n, let $m_n' = 2\lfloor m_n/2 \rfloor$ and define the p-dimensional random vectors

$$\hat{S}_{jm_n'}^{\hat{\sigma}_u}(u) = \frac{1}{\sqrt{m_n'}} \sum_{r=j}^{j+\lfloor m_n/2\rfloor-1} \hat{Z}_r^{\hat{\sigma}_u}(u) - \frac{1}{\sqrt{m_n'}} \sum_{r=j+\lfloor m_n/2\rfloor}^{j+m_n'-1} \hat{Z}_r^{\hat{\sigma}_u}(u)$$

(d) **for** r=1, ..., B **do**

- Generate independent standard normal distributed random variables $\{R_i^{(r)}\}_{i=\lceil nu-nb_n\rceil}^{\lfloor nu+nb_n\rfloor}$ -

Calculate the bootstrap statistic

$$T^{\hat{\sigma}_u,(r)}(u) = \Big| \sum_{j=\lceil nu-nb_n\rceil}^{\lfloor nu+nb_n\rfloor-m_n'+1} \hat{S}_{jm_n'}^{\hat{\sigma}_u}(u) R_j^{(r)} \Big|_\infty$$

end

(e) Define $T_{\lfloor(1-\alpha)B\rfloor}^{\hat{\sigma}_u}(u)$ as the empirical $(1-\alpha)$-quantile of the sample $T^{\hat{\sigma}_u,(1)}(u), \ldots, T^{\hat{\sigma}_u,(B)}(u)$ and

$$\hat{L}_6^{\hat{\sigma}_u}(u,t) = \hat{m}(u,t) - \hat{r}_6^{\hat{\sigma}_u}(u,t) \quad , \quad \hat{U}_6^{\hat{\sigma}_u}(u,t) = \hat{m}(u,t) + \hat{r}_6^{\hat{\sigma}_u}(u,t),$$

where

$$\hat{r}_6^{\hat{\sigma}_u}(u,t) = \frac{\hat{\sigma}(u,t)\sqrt{2}T_{\lfloor(1-\alpha)B\rfloor}^{\hat{\sigma}_u}(u)}{\sqrt{nb_n}\sqrt{(\lfloor nu+nb_n\rfloor - \lceil nu-nb_n\rceil - m_n' + 2)}}$$

Output:

$$\mathcal{C}_n^{\hat{\sigma}_u}(u) = \big\{ f \in \mathcal{C}^0 : [0,1]^2 \to \mathbb{R} \mid \ \hat{L}_6^{\hat{\sigma}_u}(u,t) \le f(t) \le \hat{U}_6^{\hat{\sigma}_u}(u,t) \ \ \forall t \in [0,1] \big\}. \tag{S4.6}$$

Next we present details of the algorithms for a simultaneous confidence band for a fixed u (of the form (S4.2)) with fixed and varying width. For this purpose we define the p-dimensional vector

$$\hat{Z}_i^{\hat{\sigma}_u}(u) = (\hat{Z}_{i,1}^{\hat{\sigma}_u}(u), \ldots, \hat{Z}_{i,p}^{\hat{\sigma}_u}(u))^\top \qquad (S4.7)$$
$$= K\Big(\frac{\frac{i}{n}-u}{b_n}\Big)\Big(\frac{\hat{\varepsilon}_{i,n}(\frac{1}{p})}{\hat{\sigma}(u,\frac{1}{p})}, \frac{\hat{\varepsilon}_{i,n}(\frac{2}{p})}{\hat{\sigma}(u,\frac{2}{p})}, \ldots, \frac{\hat{\varepsilon}_{i,n}(\frac{p-1}{p})}{\hat{\sigma}(u,\frac{p-1}{p})}, \frac{\hat{\varepsilon}_{i,n}(1)}{\hat{\sigma}(u,1)}\Big)^\top,$$

where $\hat{\varepsilon}_{i,n}$ and $\hat{\sigma}$ are defined in the main article, respectively. Algorithms S3 and S4 provides asymptotically correct the confidence bands of type (S4.2). The next Theorem S3 yields the validity of Algorithms S3 and S4, which is a consequence of Theorem S2.

Theorem S3. *Assume that the conditions of Theorem 1 hold.Define*

$$\vartheta_n' = \frac{\log^2 n}{m_n} + \frac{m_n \log n}{nb_n} + \sqrt{\frac{m_n}{nb_n}} p^{4/q}$$

and assume that $p \to \infty$ *such that* $p = O(\exp(n^\iota))$ *for some* $0 \le \iota < 1/11$.

(1) If $\alpha \in (0,1)$ *and*

$$\vartheta_n'^{1/3}\Big\{1 \vee \log\Big(\frac{p}{\vartheta_n'}\Big)\Big\}^{2/3} + \Theta\Big(\Big(\sqrt{m_n \log p}\Big(\frac{1}{\sqrt{nb_n}} + b_n^3\Big)p^{\frac{1}{q}}\Big)^{q/(q+1)}, p\Big) = o(1),$$

then we have for the confidence band in (S4.5)

$$\lim_{n\to\infty}\lim_{B\to\infty} \mathbb{P}(m \in \mathcal{C}_n(u) \mid \mathcal{F}_n) = 1-\alpha$$

in probability.

(ii) If further the conditions of Theorem 3 and Proposition 1 hold, then for the confidence band in (2.13)

$$\lim_{n\to\infty}\lim_{B\to\infty} \mathbb{P}(m \in \mathcal{C}_n^{\hat{\sigma}}(u) \mid \mathcal{F}_n) = 1-\alpha$$

in probability.

The proof of Theorem S3 follows by similar (but easier) arguments as given in the proof of Theorem 2 and Theorem 3.

Remark S4. One can prove similar results under alternative moment assumptions. In fact, Theorem S2 remains valid if condition (3.2) is replaced by

$$\mathbb{E}\Big[\sup_{0\le t\le 1}(G(u,t,\mathcal{F}_0))^4\Big] < \infty \quad .$$

Moreover, one can prove Theorem S2 under weaker assumptions than Assumption 3.2 (ii), which requires geometrically decaying dependence measure. More precisely, If the assumptions of Theorem S2 hold, where Assumption 3.2 (ii) is replaced by assumption (a) in (ii) of Remark S3 and the following conditions

(b1) There exist constants $M = M(n) > 0$, $\gamma = \gamma(n) \in (0,1)$ and $C_1 > 0$ such that

$$(2\lceil nb_n\rceil)^{3/8} M^{-1/2} l_n^{-5/8} \ge C_1 l_n$$

where $l_n = \max(\log(2\lceil nb_n\rceil p/\gamma), 1)$.

Recall the quantity Ξ_M and $\Delta_{M,q}$ defined in Remark S3. Then we have

$$\mathfrak{P}_n(u) = O\Big(\eta_n + \Theta\Big(\sqrt{nb_n}(b_n^4 + \frac{1}{n}), p\Big) + \Theta\Big(p^{\frac{1-q^*}{1+q^*}}, p\Big)\Big)$$

with

$$\begin{aligned}\eta_n = {}& (nb_n)^{-1/8} M^{1/2} l_n^{7/8} + \gamma + \left((nb_n)^{1/8} M^{-1/2} l_n^{-3/8}\right)^{q/(1+q)} \left(p\Delta_{M,q}^q\right)^{1/(1+q)} \\ & + \Xi_M^{1/3} \left(1 \vee \log\left(p/\Xi_M\right)\right)^{2/3}.\end{aligned}$$

By similar arguments as given in Remark S3, the sets $\mathcal{C}_n(u)$ and $\mathcal{C}_n^{\hat{\sigma}_u}(u)$ defined by (S4.5) and (S4.6), respectively, define an (asymptotic) $(1-\alpha)$ simultaneous confidence surface if $\eta_n = o(1)$. For example, if $\delta_q(G,i) = O(i^{-1-\alpha})$ for some $\alpha > 0$, $p = n^{\beta}$ for some $\beta > 0$ and $b_n = n^{-\gamma}$ for some $0 < \gamma < 1$, then $\eta_n = o(1)$ if $\beta - (1-\gamma)q\alpha/4 < 0$, which gives a lower bound on q.

S4.2 Finite sample properties

In this section we provide numerical results for the confidence bands for the regression function m with fixed u or t derived in Algorithms S1 - S4. As in the main part of the paper we consider simulated and real data.

For the simultaneous confidence band for a fixed $t \in [0,1]$ in (S4.1) and a fixed $u \in (0,1)$ in (S4.2), the tuning parameters are chosen in a similar way as described in Section S2.1. In particular for a fixed $u \in (0,1)$ use the bandwidth b_n as the minmizer of the loss function

$$MGCV(b) = \max_{1 \le s \le p} \frac{\sum_{i=\lceil nu-nb_n \rceil}^{\lfloor nu+nb_n \rfloor} (\hat{m}_b(\frac{i}{n}, \frac{s}{p}) - X_{i,n}(\frac{s}{p}))^2}{(1 - \mathrm{tr}(Q_s(b,u))/(\lfloor nu+nb_n \rfloor - \lceil nu-nb_n \rceil + 1))^2}. \quad \text{(S4.8)}$$

and $Q_s(b,u)$ is the submatrix of $Q_s(b)$ defined in (S2.1) consisting of $\lceil nu-nb_n \rceil : \lfloor nu+nb_n \rfloor_{th}$ rows and lines. The criterion (S4.8) is also motivated by the generalized cross validation criterion introduced by Craven and Wahba (1978).

S4.3 Simulated data

For simulated data, the regression functions and locally stationary functional time series are stated in Section S2.2. We begin displaying typical 95% simultaneous confidence bands obtained from one simulation run for model (a) with sample size $n = 800$. Figure S4 shows the simultaneous band of the type (S4.1) with constant width (Algorithm S1) and variable width (Algorithm S2), while in Figure S5 we display the simultaneous confidence bands of the form (S4.2) (for fixed u) with constant width (Algorithm S3) and variable width (Algorithm S4). We observe that in all cases there exist differences between the bands with constant and variable width, but they are not substantial.

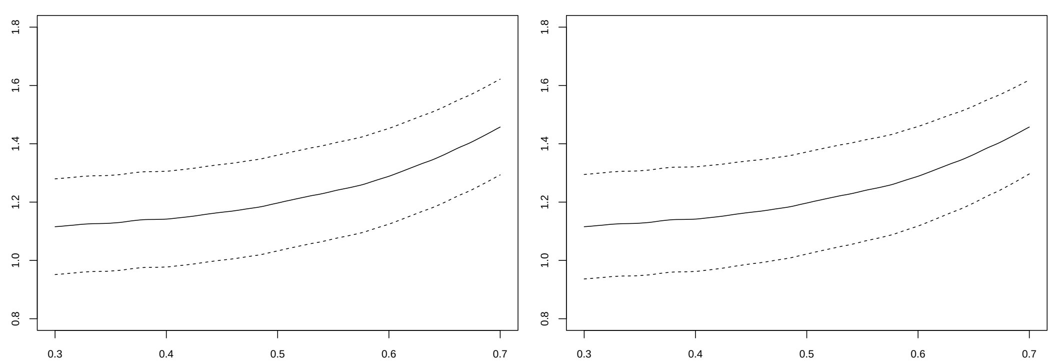

Figure S4: 95% *simultaneous confidence bands of the form* (S4.1) *(fixed* $t = 0.5$*) for the regression function in model (c) from* $n = 800$ *observations. Left panel: constant width (Algorithm S1); Right panel: varying width (Algorithm S2).*

We next investigate the coverage probabilities of confidence bands constructed for fixed $t = 0.5$ and $u = 0.5$ for sample sizes $n = 500$ and $n = 800$. All results presented in the following discussion are based on 1000 simulation runs and $B = 1000$ bootstrap replications. In all tables the left part shows the coverage probabilities of the bands with constant width while the results in the right part correspond to the bands with varying width.

In Table S3 we give some results for the confidence bands of the form (S4.1) (for fixed $t = 0.5$) with constant and variable width (c.f. Algorithm S1 and Algorithm S2), while we present in Table S4 the simulated coverage probabilities of the simultaneous confidence bands of the form (S4.2), where $u = 0.5$ is fixed (c.f. Algorithm S3 and Algorithm S4). We observe that the simulated coverage probabilities are close to their nominal levels in all cases under consideration, which illustrates the validity of our methods for finite sample sizes.

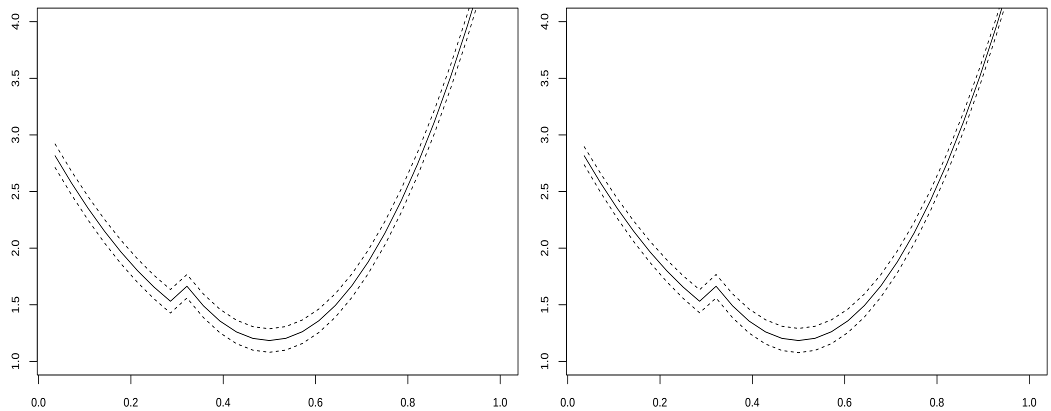

Figure S5: 95% *simultaneous confidence band of the form* (S4.2) *(fixed* $u = 0.5$*) for the regression function in model (c) from* $n = 800$ *observations. Left panel: constant width (Algorithm S3); Right panel: varying width (Algorithm S4).*

Table S3: *Simulated coverage probabilities of the simultaneous confidence band of the form* (S4.1) *for fixed* $t = 0.5$ *calculated by Algorithm S1 (constant width) and S2 (varying width).*

	Constant Width				Varying Width			
	Model (a)		Model (b)		Model (a)		Model (b)	
Level	90%	95%	90%	95%	90%	95%	90%	95%
n=500	90.3%	95.0%	91.7%	96.0%	91.2%	95.6%	91.3%	96.2%
n=800	88.5%	95.4%	88.7%	94.4%	88.8%	94.5%	88.4%	94.0%
	Model (c)		Model (d)		Model (c)		Model (d)	
Level	90%	95%	90%	95%	90%	95%	90%	95%
n=500	91.7%	96.3%	91.4%	95.6%	91.5%	95.4%	90.4%	94.1%
n=800	89.1%	94.8%	89.8%	94.5%	87.5%	93.4%	88.7%	94.4%

Table S4: *Simulated coverage probabilities of the simultaneous confidence band of the form* (S4.2) *for fixed* $u = 0.5$ *calculated by Algorithms S3 (constant width) and S4 (varying width).*

	Constant Width				Varying Width			
	Model (a)		Model (b)		Model (a)		Model (b)	
Level	90%	95%	90%	95%	90%	95%	90%	95%
n=500	87.0%	93.4%	88.4%	93.5%	86.9%	92.2%	88.7%	93.7%
n=800	88.7%	93.7%	88.4%	94.7%	89.4%	94.4%	88.9%	94.1%
	Model (c)		Model (d)		Model (c)		Model (d)	
Level	90%	95%	90%	95%	90%	95%	90%	95%
n=500	86.6%	92.3%	90.2%	94.0%	90.2%	94.5%	89.5%	94.2%
n=800	89.6%	94.7%	87.8%	93.3%	88.9%	93.4%	89.8%	94.1%

S4.4 Real data

In this section we further study the well documented volatility smile for implied volatility of the European call option of SP500 data set considered in Section 4 of the main article. In Figure S6 of we display 95% simultaneous confidence bands of the form (S4.1) for fixed $t = 0.5$ (which corresponds to Moneyness=1.1) where the parameters are chosen as $b_n = 0.12$ and $m_n = 18$. We observe that the implied volatility changes with time (or precisely the time to maturity) when moneyness (or equivalently, the strike price and underlying asset price) is specified. We also calculate confidence bands of the form (S4.2) for fixed $u = 0.5$, by Algorithm S3 (constant width) and Algorithm S4 (varying width). The parameter selection procedure yields $b_n = 0.1$ and $m_n = 32$, and the resulting simultaneous confidence bands of the form (S4.2) are presented in Figure S7. We observe that both 95% simultaneous confidence bands indicate that the implied volatility is a quadratic function of moneyness, which supports the well documented phenomenon of 'volatility smile'. We observe that the differences between the bands with constant and variable width are rather small.

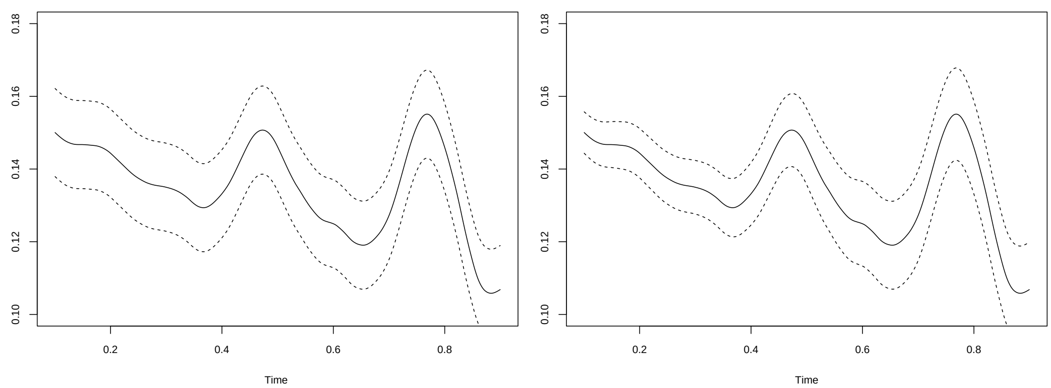

Figure S6: *95% simultaneous confidence bands of the form* (S4.1) *(fixed* $t = 0.5$*) for the data example in Section 4. Left panel: constant width (Algorithm S1); Right panel: variable width (Algorithm S2).*

S5 Examples of locally stationary error processes

In this section we present several examples for the error processes, which satisfy the assumptions of the main article.

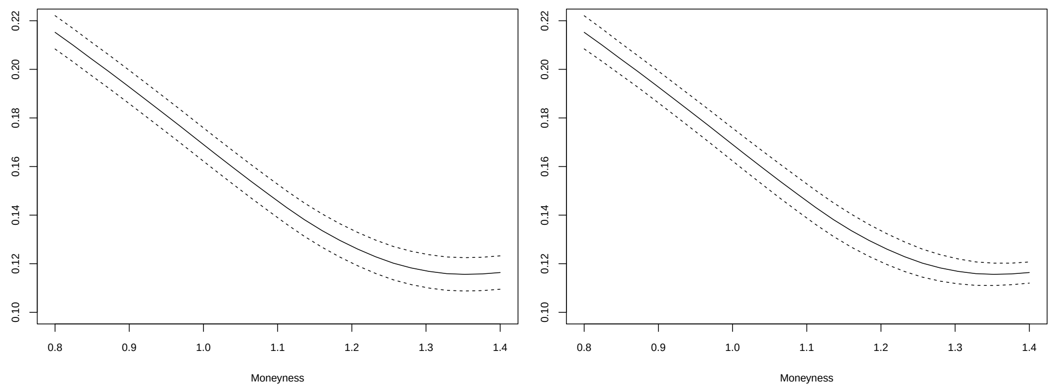

Figure S7: *95% simultaneous confidence bands of the form* (S4.2) *(fixed* $u = 0.5$*) for the IV surface. Left panel: constant width (Algorithm S3); Right panel: variable width (Algorithm S4).*

Example S1. Let $(B_j)_{j\geq 0}$ denote a basis of $L^2\big([0,1]^2\big)$ and let $(\eta_{i,j})_{i\geq 0, j\geq 0}$ denote an array of independent identically distributed centered random variables with variance σ^2. We define the error process

$$\epsilon_i(u,v) = \sum_{j=0}^{\infty} \eta_{i,j} B_j(u,v),$$

assume that

$$\sup_{u\in[0,1]} \int_0^1 \mathbb{E}(\epsilon_i^2(u,v))dv = \sigma^2 \sup_{u\in[0,1]} \sum_{s=0}^{\infty} \int B_s^2(u,v)dv < \infty.$$

Next, consider the locally stationary MA(∞) functional linear model

$$\varepsilon_{i,n}(t) = \sum_{j=0}^{\infty} \int_0^1 a_j(t,v)\epsilon_{i-j}(\tfrac{i}{n},v)dv \ , \tag{S5.1}$$

where $(a_j)_{j\geq 0}$ is a sequence of square integrable functions $a_j : [0,1]^2 \to \mathbb{R}$ satisfying

$$\sum_{j=0}^{\infty} \sup_{u,v\in[0,1]} |a_j(u,v)| < \infty.$$

Define $\mathcal{F}_i = (\ldots, \eta_{i-1}, \eta_i)$, then we obtain from (S5.1) the representation of the form $\varepsilon_{i,n}(t) = G(\frac{i}{n}, t, \mathcal{F}_i)$, where

$$G(u,t,\mathcal{F}_i) = \sum_{j=0}^{\infty} \int_0^1 a_j(t,v) \sum_{s=0}^{\infty} \eta_{i-j,s} B_s(u,v)dv.$$

Further, assume that $\|\eta_{1,1}\|_q < \infty$ for some $q > 2$, then by Burkholder's and Cauchy's inequality the physical dependence measure defined in (3.2) satisfies

$$\begin{aligned}
\delta_q(G,i) &= \sup_{u,t\in[0,1]} \Big\| \sum_{s=0}^{\infty} \int_0^1 a_i(t,v) B_s(u,v)dv(\eta_{0,s} - \eta'_{0,s}) \Big\|_q \\
&= O\Big(\sup_{u,t\in[0,1]} \Big(\sum_{s=0}^{\infty} \Big(\int_0^1 a_i(t,v)B_s(u,v)dv \Big)^2 \Big)^{1/2} \Big) \\
&= O\Big(\sup_{t\in[0,1]} \Big[\int_0^1 a_i^2(t,v)dv \Big]^{1/2} \Big) \ .
\end{aligned}$$

Therefore Assumption 3.2(2) will be satisfied if

$$\sup_{t\in[0,1]} \Big[\int_0^1 a_i^2(t,v)dv \Big]^{1/2} = O(\chi^i) \ .$$

Similarly, it follows for $q \geq 2$ that

$$\begin{aligned}
\|G(u,t,\mathcal{F}_0)\|_q^2 &\leq Mq \sum_{j=0}^{\infty} \sum_{s=0}^{\infty} (\int_0^1 a_j(t,v)B_s(u,v)dv)^2 \|\eta_{1,1}\|_q^2 \\
&\leq Mq \sum_{j=0}^{\infty} \int_0^1 a_j^2(t,v)dv \sum_{s=0}^{\infty} \int_0^1 B_s^2(u,v)dv \|\eta_{1,1}\|_q^2 \qquad \text{(S5.2)}
\end{aligned}$$

for some sufficiently large constant M. Consequently, the filter G has finite moment of order q, if

$$\sum_{j=0}^{\infty} \int_0^1 a_j^2(t,v)dv < \infty \ . \qquad \text{(S5.3)}$$

Furthermore, if there exists positive constants M_0 and α such that $\|\eta_{1,1}\|_q \le M_0 q^{1/2-\alpha}$, Assumption 3.2(1) is also satisfied, because for any fixed t_0, the sequence

$$\frac{t_0^q \|G(u,t,\mathcal{F}_0)\|_q^q}{q!} = O\Big(\frac{C^q t_0^q q^{q-\alpha q}}{q!}\Big) = O\Big(\frac{1}{\sqrt{2\pi q}}(\frac{Ct_0 e}{q^\alpha})^q\Big)$$

is summable, where

$$C = \sup_{t\in[0,1],u\in[0,1]} M_0 \sqrt{M\sum_{j=0}^{\infty}\int_0^1 a_j^2(t,v)dv \sum_{s=0}^{\infty}\int_0^1 B_s^2(u,v)dv}.$$

Moreover, if $b_s(u,v) := \frac{\partial}{\partial u}B_s(u,v)$ exists for $u\in(0,1), v\in[0,1]$, then it follows observing (S5.2) that Assumption 3.2(3) holds under (S5.3) and

$$\sup_{u\in[0,1]}\sum_{s=0}^{\infty}\int b_s^2(u,v)dv < \infty.$$

Finally, if $\|\eta_{1,1}\|_{q^*} < \infty$ and

$$\sup_{t\in[0,1]}\Big[\int_0^1 \Big(\frac{\partial}{\partial t}a_i(t,v)\Big)^2 dv\Big]^{1/2} = O(\chi^i) \ ,$$

it can be shown by similar arguments as given above that Assumption 3.3 is satisfied.

Example S2. For a given orthonormal basis $(\phi_k(t))_{k\ge1}$ of $L^2([0,1])$ consider the functional time series $(G(u,t,\mathcal{F}_i))_{i\in\mathbb{Z}}$ defined by

$$G(u,t,\mathcal{F}_i) = \sum_{k=1}^{\infty} H_k(u,\mathcal{F}_i)\phi_k(t) \ , \tag{S5.4}$$

where for each $k\in\mathbb{N}$ and $u\in[0,1]$ the random coefficients $(H_k(u,\mathcal{F}_i))_{i\in\mathbb{Z}}$ are stationary time series. A parsimonious choice of (S5.4) is to consider $\mathcal{F}_i = \cup_{k=1}^{\infty}\mathcal{F}_{i,k}$ where $\{\mathcal{F}_{i,k}\}_{k=1}^{\infty}$

are independent filtrations. In this case we obtain

$$G(u,t,\mathcal{F}_i)=\sum_{k=1}^{\infty}H_k(u,\mathcal{F}_{i,k})\phi_k(t), \tag{S5.5}$$

and the random coefficients $H_k(u,\mathcal{F}_{i,k})$ are stochastically independent. A sufficient condition for Assumption 3.2(2) in model (S5.5) is

$$\sup_{t\in[0,1]}\sum_{k=0}^{\infty}|\phi_k(t)|\delta_q(H_k,i)=O(\chi^i)\ ,$$

where $\delta_q(H_k,i):=\sup_{u\in[0,1]}\|H_k(u,\mathcal{F}_{i,k})-H_k(u,\mathcal{F}_{i,k}^*)\|_q$. The qth moment of the process G in (S5.5) exists for $q\geq 2$, if

$$\Delta_q:=\sup_{t\in[0,1],u\in[0,1]}\sum_{k=0}^{\infty}\phi_k^2(t)\|H_k(u,\mathcal{F}_{0,k})\|_q^2<\infty.$$

If further $\Delta_q=O(q^{1/2-\alpha})$ for some $\alpha>0$, then similar arguments as given in Example S1 show that Assumption 3.2(1) is satisfied as well. Finally, if the inequality

$$\sum_{k=0}^{\infty}\phi_k^2(t)\Big\|\frac{\partial}{\partial u}H_k(u,\mathcal{F}_{0,k})\Big\|_q^2<\infty$$

holds uniformly with respect to $t,u\in[0,1]$, Assumption 3.2(3) is also satisfied.

On the other hand, in model (S5.4) we have $H_k(u,\mathcal{F}_i)=\int_0^1 G(u,t,\mathcal{F}_i)\phi_k(t)dt$, and consequently the magnitude of $\|H_k\|_q$ and $\delta_q(H_k,i)$ can be determined by Assumption 3.2. For example, if the basis of $L^2([0,1])$ is given by $\phi_k(t)=\cos(k\pi t)$ $(k=0,1,\ldots)$ and the inequality

$$\|G(u,0,\mathcal{F}_1)\|_q+\Big\|\frac{\partial}{\partial t}G(u,0,\mathcal{F}_1)\Big\|_q+\sup_{t\in[0,1]}\Big\|\frac{\partial^2}{\partial t^2}G(u,t,\mathcal{F}_1)\Big\|_q<\infty,$$

holds for $u \in [0,1]$, it follows by similar arguments as given in Zhou and Dette (2020) that

$$\sup_{u\in[0,1]} \big\|H_k(u,\mathcal{F}_k)\big\|_q = O(k^{-2}), \quad \delta_q(H_k,i) = O\big(\min(k^{-2},\delta_G(i,q))\big). \qquad \text{(S5.6)}$$

Similarly, assume that the basis of $L^2([0,1])$ is given by the Legendre polynomials and that

$$\sup_{u\in[0,1]} \max_{s=1,2,3} \Big\| \int_{-1}^{1} \frac{|\frac{\partial^s}{\partial t^s} G(u,t,\mathcal{F}_0)|}{\sqrt{1-x^2}} dx \Big\|_q < \infty.$$

If additionally for every $\varepsilon > 0$, there exists a constant $\delta > 0$ such that

$$\sum_{s=1,2} \sum_k \Big\| \frac{\partial^s}{\partial t^s} G\left(u, x_k, \mathcal{F}_i\right) - \frac{\partial^s}{\partial t^s} G\left(u, x_{k-1}, \mathcal{F}_i\right)\Big) \Big\|_q < \varepsilon$$

for any finite sequence of pairwise disjoint sub-intervals (x_{k-1}, x_k) of the interval $(0,1)$ such that $\sum_k (x_k - x_{k-1}) < \delta$, it follows from Theorem 2.1 of Wang and Xiang (2012) that (S5.6) holds as well.

Finally, if

$$\sup_{t\in[0,1]} \sum_{k=0}^{\infty} |\phi_k'(t)| \delta_{q^*}(H_k,i) = O(\chi^i),$$

it can be shown by similar arguments as given above that Assumption 3.3 is also satisfied.

S6 Proofs of Theorems

In the proofs, for two real sequence a_n and b_n we write $a_n \lesssim b_n$, if there exists a universal positive constant M such that $a_n \le M b_n$. Let $\mathbf{1}(\cdot)$ be the usual indicator function. For simplicity let $\tilde{K}(u) = \frac{1}{nb_n}\sum_{i=1}^n K\big(\frac{i/n-u}{b_n}\big)$.

S6.1 Proof of Theorem 1

For $p \in \mathbb{N}$ define by $t_v = \frac{v}{p}$, $(v = 0, \ldots, p)$ an equidistant partition of the interval $[0, 1]$ and let M be a sufficiently large generic constant which may vary from line to line. Define

$$W_n(u,t) = \sqrt{nb_n}\big(\hat{m}(u,t) - \mathbb{E}(\hat{m}(u,t))\big) = \frac{1}{\sqrt{nb_n}\tilde{K}(u)}\sum_{i=1}^n G(\tfrac{i}{n}, t, \mathcal{F}_i)K\Big(\frac{\frac{i}{n}-u}{b_n}\Big), \tag{S6.1}$$

we have by triangle inequality

$$\Big|\sup_{b_n \le u \le 1-b_n, 0\le t\le 1} |W_n(u,t)| - \max_{\substack{\lceil nb_n\rceil \le l_1 \le n-\lceil nb_n\rceil \\ 1\le s\le p}} |W_n(\tfrac{l_1}{n}, \tfrac{s}{p})|\Big| \le \tilde{W}_n \ ,$$

where

$$\tilde{W}_n \ = \max_{\substack{\lceil nb_n\rceil \le l_1\le n-\lceil nb_n\rceil, 1\le s\le p, \\ |u-\frac{l_1}{n}|\le 1/n, |t-\frac{s}{p}|\le 1/p, u,t\in[0,1]}} |W_n(u,t) - W_n(\tfrac{l_1}{n}, \tfrac{s}{p})|.$$

By Assumption 3.3, Burkholder's inequality and similar arguments as given in the proof of Proposition 1.1 of Dette and Wu (2022) we obtain

$$\begin{aligned}&\sup_{u,t\in[0,1]}\Big\|\frac{\partial}{\partial u}W_n(u,t)\Big\|_{q^*} \le \frac{M}{b_n}, \ \sup_{u,t\in[0,1]}\Big\|\frac{\partial}{\partial t}W_n(u,t)\Big\|_{q^*} \le M, \\ &\sup_{u,t\in[0,1]}\Big\|\frac{\partial^2}{\partial u\partial t}W_n(u,t)\Big\|_{q^*} \le \frac{M}{b_n}.\end{aligned} \tag{S6.2}$$

Note that we have for $\tau_s > 0$, $s = 1, 2$ and $x, y \in [0, 1)$,

$$\left\| \sup_{\substack{0 \le t_1 \le \tau_1 \\ 0 \le t_2 \le \tau_2}} |W_n(t_1 + x, t_2 + y) - W_n(x, y)| \right\|_{q*} \le \int_0^{\tau_1} \left\| \frac{\partial}{\partial u} W_n(x + u, y) \right\|_{q^*} du$$
$$+ \int_0^{\tau_2} \left\| \frac{\partial}{\partial t} W_n(x, y + v) \right\|_{q^*} dv + \int_0^{\tau_1} \int_0^{\tau_2} \left\| \frac{\partial^2}{\partial x \partial t} W_n(x + u, y + v) \right\|_{q^*} dudv.$$

Therefore, (S6.2) and similar arguments as in the proof of Proposition B.2 of Dette et al. (2019) show

$$\|\tilde{W}_n\|_{q^*} = O((np)^{1/q^*}((nb_n)^{-1} + 1/p)). \tag{S6.3}$$

Observing (S7.7) and (S7.8). Lemma S1 and (S6.3) it therefore follows that

$$\begin{aligned}
\mathfrak{P}_n &\lesssim (nb_n)^{-(1-11\iota)/8} + \Theta\Big(\sqrt{nb_n}\Big(b_n^4 + \frac{1}{n}\Big), np\Big) + \Theta(\delta, np) + \mathbb{P}(\tilde{W}_n > \delta) \\
&\lesssim (nb_n)^{-(1-11\iota)/8} + \Theta\Big(\sqrt{nb_n}\Big(b_n^4 + \frac{1}{n}\Big), np\Big) + \Theta(\delta, np) \\
&\quad + \big((np)^{1/q^*}((nb_n)^{-1} + 1/p)/\delta\big)^{q^*}.
\end{aligned}$$

Solving $\delta = \big((np)^{1/q^*}((nb_n)^{-1} + 1/p)/\delta\big)^{q^*}$ we get $\delta = \big((np)^{1/q^*}((nb_n)^{-1} + 1/p)\big)^{\frac{q^*}{q^*+1}}$ and the assertion of the theorem follows. □

S6.2 Proof of Theorem 2

Proof. In the following discussion we use the following notation. For any vector y_n indexed by n, let $y_{n,r}$ be its r_{th} component. For example, $\hat{S}_{rm_n,j}$ is the j_{th} entry of the vector $\hat{S}_{rm_n}$.

Let T_k denote the statistic generated by (2.15) in one bootstrap iteration of Algorithm 1 and define for integers a, b the quantities

$$T^\diamond_{ap+b} = \sum_{j=1}^{2\lceil nb_n\rceil - m'_n} \hat{S}_{jm'_n,(a-1)p+b} R_{k+j-1}, a = 1, ...n - 2\lceil nb_n\rceil + 1, 1 \leq b \leq p$$

$$T^\diamond := ((T^\diamond_1)^\top, \ldots, (T^\diamond_{(n-2\lceil nb_n\rceil+1)p})^\top)^\top = \big({T_1}^\top, \ldots, {T_{n-2\lceil nb_n\rceil+1}}^\top\big)^\top$$

$$T = |T^\diamond|_\infty = \max_{1\leq k\leq n-2\lceil nb_n\rceil+1} |T_k|_\infty$$

It suffices to show that the following inequality holds

$$\sup_{x\in\mathbb{R}} \Big|\mathbb{P}(|T^\diamond/\sqrt{2\lceil nb_n\rceil - m'_n}|_\infty \leq x|\mathcal{F}_n) - \mathbb{P}\Big(\frac{1}{\sqrt{2nb_n}}\Big|\sum_{i=1}^{2\lceil nb_n\rceil-1} \tilde{Y}_i\Big|_\infty \leq x\Big)\Big|$$

$$= O_p\Big(\vartheta_n^{1/3}\{1 \vee \log(\frac{np}{\vartheta_n})\}^{2/3} + \Theta\Big(\Big(\sqrt{m_n \log np}\Big(\frac{1}{\sqrt{nb_n}} + b_n^3\Big)(np)^{\frac{1}{q}}\Big)^{q/(q+1)}, np\Big)\Big) \quad \text{(S6.4)}$$

If this estimate has been established, Theorem 2 follows from Theorem 1, which shows that the probabilities $\mathbb{P}\big(\max_{b_n\leq u\leq 1-b_n, 0\leq t\leq 1} \sqrt{nb_n}|\hat{\Delta}(u,t)| \leq x\big)$ can be approximated by the probabilities

$$\mathbb{P}\big(\frac{1}{\sqrt{nb_n}}|\sum_{i=1}^{2\lceil nb_n\rceil-1} \tilde{Y}_i|_\infty \leq x\big)$$

uniformly with respect to $x \in \mathbb{R}$.

For a proof of (S6.4) we assume without loss of generality that m_n is even so that $m'_n = m_n$. For convenience, let $\sum_{i=a}^b Z_i = 0$ if the indices a and b satisfy $a > b$. Given the data, it follows for the conditional covariance

$$
\begin{aligned}
((2\lceil nb_n\rceil-1)-m_n+1)\sigma^{T^\diamond}_{(k_1-1)p+j_1,(k_2-1)p+j_2} &:= \mathbb{E}(T^\diamond_{(k_1-1)p+j_1}T^\diamond_{(k_2-1)p+j_2}|\mathcal{F}_n) \quad \text{(S6.5)}\\
&= \mathbb{E}\Big(\sum_{r=1}^{2\lceil nb_n\rceil-m_n}\hat{S}_{rm_n,(k_1-1)p+j_1}R_{k_1+r-1}\sum_{r=1}^{2\lceil nb_n\rceil-m_n}\hat{S}_{rm_n,(k_2-1)p+j_2}R_{k_2+r-1}\Big|\mathcal{F}_n\Big)\\
&= \sum_{r=1}^{2\lceil nb_n\rceil-m_n-(k_2-k_1)}\hat{S}_{(r+k_2-k_1)m_n,(k_1-1)p+j_1}\hat{S}_{rm_n,(k_2-1)p+j_2}.
\end{aligned}
$$

where $1\leq k_1\leq k_2\leq(n-2\lceil nb_n\rceil+1)$, $1\leq j_1,j_2\leq p$. Here, without generality, we assume $k_1\leq k_2$. Define $\tilde{T}^\diamond$, and $\tilde{S}_{jm_n}$ in the same way as $T^\diamond$, and $\hat{S}_{jm_n}$ in (2.15) and (2.14), respectively, where the residuals $\hat{\tilde{Z}}_i$ defined in (2.12) and used in step (a) of Algorithm 1 have been replaced by quantities $\tilde{Z}_i$ defined in (2.7).Then we obtain by similar arguments

$$
\begin{aligned}
((2\lceil nb_n\rceil-1)-m_n+1)\sigma^{\tilde{T}^\diamond}_{(k_1-1)p+j_1,(k_2-1)p+j_2} &:= \mathbb{E}(\tilde{T}^\diamond_{(k_1-1)p+j_1}\tilde{T}^\diamond_{(k_2-1)p+j_2}|\mathcal{F}_n)\\
&= \sum_{r=1}^{\lceil 2nb_n\rceil-m_n-(k_2-k_1)}\tilde{S}_{(r+k_2-k_1)m_n,(k_1-1)p+j_1}\tilde{S}_{rm_n,(k_2-1)p+j_2}. \quad \text{(S6.6)}
\end{aligned}
$$

Recall the definition of the random variable $\tilde{Y}_j$ in Proposition S1 and denote by $\tilde{Z}_{j,i}$, $\tilde{Y}_{j,i}$ the ith component of the vectors $\tilde{Z}_j$ and $\tilde{Y}_j$, respectively ($1\leq i\leq(n-2\lceil nb_n\rceil+1)p$, $1\leq j\leq 2\lceil nb_n\rceil-1$). Then we obtain

$$
\begin{aligned}
\sigma^{\tilde{Y}}_{(k_1-1)p+j_1,(k_2-1)p+j_2} &:= \mathbb{E}\Big(\frac{1}{2\lceil nb_n\rceil-1}\sum_{i_1=1}^{2\lceil nb_n\rceil-1}\tilde{Y}_{i_1,(k_1-1)p+j_1}\sum_{i_2=1}^{2\lceil nb_n\rceil-1}\tilde{Y}_{i_2,(k_2-1)p+j_2}\Big)\\
&= \frac{\mathbb{E}(\sum_{i_1=1}^{2\lceil nb_n\rceil-1}\tilde{Z}_{i_1,(k_1-1)p+j_1}\sum_{i_2=1}^{2\lceil nb_n\rceil-1}\tilde{Z}_{i_2,(k_2-1)p+j_2})}{2\lceil nb_n\rceil-1}\\
&= \frac{\mathbb{E}(\sum_{i_1=1}^{2\lceil nb_n\rceil-1}Z_{i_1+(k_1-1),\lceil nb_n\rceil+(k_1-1),j_1}\sum_{i_2=1}^{2\lceil nb_n\rceil-1}Z_{i_2+(k_2-1),\lceil nb_n\rceil+(k_2-1),j_2})}{2\lceil nb_n\rceil-1}, \quad \text{(S6.7)}
\end{aligned}
$$

where $Z_{i_1+(k_1-1),\lceil nb_n\rceil,j_1}$ is the j_1th entry of the $p-$dimensional random vector $Z_{i_1+(k_1-1),\lceil nb_n\rceil}$ and $Z_{i_2+(k_2-1),\lceil nb_n\rceil,j_2}$ is defined similarly. We will show at the end of this section that

$$\Big\| \max_{k_1,k_2,j_1,j_2} |\sigma^{\tilde{Y}}_{(k_1-1)p+j_1,(k_2-1)p+j_2} - \sigma^{\tilde{T}^\diamond}_{(k_1-1)p+j_1,(k_2-1)p+j_2}\Big\|_{q/2} = O(\vartheta_n). \qquad \text{(S6.8)}$$

If (S6.8) holds, it follows from Lemma S3 that there exists a constant $\eta_0 > 0$ such that

$$\mathbb{P}\Big(\min_{\substack{1\le k\le (n-2\lceil nb_n\rceil+1),\\ 1\le j\le p}} \sigma^{\tilde{T}^\diamond}_{(k-1)p+j,(k-1)p+j} \ge \eta_0\Big) \ge 1 - O(\vartheta_n^{q/2}).$$

Then, by Theorem 2 of Chernozhukov et al. (2015), we have

$$\sup_{x\in\mathbb{R}} \Big|\mathbb{P}\Big(\frac{|\tilde{T}^\diamond|_\infty}{\sqrt{2\lceil nb_n\rceil - m_n}} \le x \;\Big|\mathcal{F}_n\Big) - \mathbb{P}\Big(\frac{1}{\sqrt{2nb_n}}\Big|\sum_{i=1}^{2\lceil nb_n\rceil -1} \tilde{Y}_i\Big|_\infty \le x\Big)\Big|$$

$$= O_p(\vartheta_n^{1/3}\{1\vee \log(\frac{np}{\vartheta_n})\}^{2/3}). \qquad \text{(S6.9)}$$

Since conditional on $\mathcal{F}_n$, $\big(\tilde{T}^\diamond - T^\diamond\big)$ is an $(n-2\lceil nb_n\rceil+1)p$ dimensional Gaussian random vector we obtain by the (conditional) Jensen inequality and conditional inequality for the concentration of the maximum of a Gaussian process (see Chapter 5 in Appendix A of Chatterjee, 2014, where a similar result has been derived in Lemma A.1) that

$$\mathbb{E}(|\tilde{T}^\diamond - T^\diamond|^q_\infty|\mathcal{F}_n) \le M|\sqrt{\log np} \max_{r=1}^{(n-2\lceil nb_n\rceil+1)p} \Big(\sum_{j=1}^{2\lceil nb_n\rceil - m_n'} (\hat{S}_{jm_n',r} - S_{jm_n',r})^2\Big)^{1/2}|^q \qquad \text{(S6.10)}$$

for some large constant M almost surely. Observing that

$$\max_{1\le i\le n} |Z_i|^l \le \sum_{1\le i\le n} |Z_i|^l \qquad \text{for any } l>0, n\in\mathbb{N} \qquad \text{(S6.11)}$$

and using a similar argument as given in the proof of Proposition 1.1 in Dette and Wu

(2022) and the fact that K_l and K_r are both three order kernels, we have

$$\frac{1}{\sqrt{2\lceil nb_n\rceil - m_n}}\Big\|\max_{r=1}^{(n-2\lceil nb_n\rceil+1)p}\Big(\sum_{j=1}^{\lceil 2nb_n\rceil - m_n'}(\hat{S}_{jm_n',r}-S_{jm_n',r})^2\Big)^{1/2}\Big\|_q = O\Big(\sqrt{m_n}\Big(\frac{1}{\sqrt{nb_n}}+b_n^3\Big)(np)^{\frac{1}{q}}\Big),$$

and combining this result with the (conditional version) of Lemma S1 and (S6.10) yields

$$\begin{aligned}
&\sup_{x\in\mathbb{R}}\Big|\mathbb{P}\Big(\frac{|T^\diamond|_\infty}{\sqrt{2\lceil nb_n\rceil - m_n}} > x \,\Big|\mathcal{F}_n\Big) - \mathbb{P}\Big(\frac{1}{\sqrt{2nb_n}}\Big|\sum_{i=1}^{2\lceil nb_n\rceil -1}\tilde{Y}_i\Big|_\infty > x\Big)\Big| \\
&\leq \sup_{x\in\mathbb{R}}\Big|\mathbb{P}\Big(\frac{|\tilde{T}^\diamond|_\infty}{\sqrt{2\lceil nb_n\rceil - m_n}} > x \,\Big|\mathcal{F}_n\Big) - \mathbb{P}\Big(\frac{1}{\sqrt{2nb_n}}\Big|\sum_{i=1}^{2\lceil nb_n\rceil -1}\tilde{Y}_i\Big|_\infty > x\Big)\Big| \\
&\qquad + \mathbb{P}\Big(\frac{|\tilde{T}^\diamond - T^\diamond|_\infty}{\sqrt{2\lceil nb_n\rceil - m_n}} > \delta\Big|\mathcal{F}_n\Big) + O\big(\Theta(\delta, np)\big) \\
&\leq \sup_{x\in\mathbb{R}}|\mathbb{P}\Big(\frac{|\tilde{T}^\diamond|_\infty}{\sqrt{2\lceil nb_n\rceil - m_n}} > x \,\Big|\mathcal{F}_n\Big) - \mathbb{P}\Big(\frac{1}{\sqrt{2nb_n}}\Big|\sum_{i=1}^{2\lceil nb_n\rceil -1}\tilde{Y}_i\Big|_\infty > x\Big)\Big| \\
&\qquad + O_p\big(\delta^{-q}\big(\sqrt{m_n\log np}\big(\frac{1}{\sqrt{nb_n}}+b_n^3\big)(np)^{\frac{1}{q}}\big)^q\big) + O\big(\Theta(\delta, np)\big), \quad \text{(S6.12)}
\end{aligned}$$

where we have used the Markov's inequality. Taking $\delta = \big(\sqrt{m_n\log np}(\frac{1}{\sqrt{nb_n}}+b_n^3)(np)^{\frac{1}{q}}\big)^{q/(q+1)}$ in (S6.12), and combining this estimate with (S6.9) yields (S6.4) completes the proof.

Proof of (S6.8)**.** To simplify the notation, write

$$G_{j,i,k} = G(\tfrac{i+k-1}{n}, j/p, \mathcal{F}_{i+k-1}), \quad G_{j,i,k,u} = G(\tfrac{i+k-1+u}{n}, j/p, \mathcal{F}_u)$$

Without loss of generality, we consider the case $k_1 \leq k_2$. We calculate $\sigma^{\tilde{Y}}_{(k_1-1)p+j_1,(k_2-1)p+j_2}$ observing the representation

$$Z_{i_1+(k_1-1),\lceil nb_n\rceil+(k_1-1),j_1} = G_{j_1,i_1,k_1}K\Big(\frac{i_1-\lceil nb_n\rceil}{nb_n}\Big).$$

By Lemma S2 it follows that

$$\mathbb{E}\big[Z_{i_1+(k_1-1),\lceil nb_n\rceil+(k_1-1),j_1} Z_{i_2+(k_2-1),\lceil nb_n\rceil+(k_2-1),j_2}\big] = O(\chi^{|i_1-i_2+k_1-k_2|}). \quad \text{(S6.13)}$$

uniformly for $1 \le i_1, i_2 \le 2\lceil nb_n\rceil - 1$, $1 \le j_1, j_2 \le p$, $1 \le k_1, k_2 \le n - 2\lceil nb_n\rceil + 1$. We first show that (S6.8) holds whenever $k_2 - k_1 > 2\lceil nb_n\rceil - m_n$. On the one hand, observing and (S6.5) and (S6.6) that if $2\lceil nb_n\rceil - m_n - (k_2 - k_1) < 0$ then

$$\sigma^{\tilde{T}^\diamond}_{(k_1-1)p+j_1,(k_2-1)p+j_2} = 0 \quad a.s. \quad \text{(S6.14)}$$

Moreover, by (S6.7) and (S6.13), straightforward calculations show that

$$\sigma^{\tilde{Y}}_{(k_1-1)p+j_1,(k_2-1)p+j_2} = \frac{1}{2\lceil nb_n\rceil - 1} O\Big(\sum_{i_1=1}^{2\lceil nb_n\rceil-1}\sum_{i_2=1}^{2\lceil nb_n\rceil-1} \chi^{|i_1-i_2+k_1-k_2|}\Big) = O\Big(\frac{m_n}{nb_n}\Big) \quad \text{(S6.15)}$$

Combining (S6.14), (S6.15) and by applying similar argument to $k_1 \ge k_2$, we obtain

$$\Big\| \max_{\substack{k_1,k_2,j_1,j_2 \\ |k_2-k_1|>2\lceil nb_n\rceil-m_n}} |\sigma^{\tilde{Y}}_{(k_1-1)p+j_1,(k_2-1)p+j_2} - \sigma^{\tilde{T}^\diamond}_{(k_1-1)p+j_1,(k_2-1)p+j_2}\Big\|_{q/2} = O\Big(\frac{m_n}{nb_n}\Big) \quad \text{(S6.16)}$$

Now consider the case that $k_2 - k_1 \le 2\lceil nb_n\rceil - m_n$. Without losing generality we consider $k_1 \le k_2$. Again by (S6.7)

$$\mathbb{E}\Big(\sum_{i_1=1}^{k_2-k_1} Z_{i_1+(k_1-1),\lceil nb_n\rceil+(k_1-1),j_1} \sum_{i_2=1}^{2\lceil nb_n\rceil-1} Z_{i_2+(k_2-1),\lceil nb_n\rceil+(k_2-1),j_2}\Big)$$
$$= O\Big(\sum_{i_1=1}^{k_2-k_1}\sum_{i_2=1}^{2\lceil nb_n\rceil-1} \chi^{|i_2-i_1+k_2-k_1|}\Big) = O\Big(\sum_{i_1=1}^{k_2-k_1}\sum_{i_2=1}^{2\lceil nb_n\rceil-1} \chi^{i_2-i_1+k_2-k_1}\Big) = O(1),$$
$$\mathbb{E}\Big(\sum_{i_1=1}^{2\lceil nb_n\rceil-1} Z_{i_1+(k_1-1),\lceil nb_n\rceil+(k_1-1),j_1} \sum_{i_2=2\lceil nb_n\rceil-(k_2-k_1)}^{2\lceil nb_n\rceil-1} Z_{i_2+(k_2-1),\lceil nb_n\rceil+(k_2-1),j_2}\Big)$$
$$= O\Big(\sum_{i_1=1}^{2\lceil nb_n\rceil-1}\sum_{i_2=2\lceil nb_n\rceil-(k_2-k_1)}^{2\lceil nb_n\rceil-1} \chi^{|i_2-i_1+k_2-k_1|}\Big) = O\Big(\sum_{i_1=1}^{2\lceil nb_n\rceil-1}\sum_{i_2=2\lceil nb_n\rceil-(k_2-k_1)}^{2\lceil nb_n\rceil-1} \chi^{i_2-i_1+k_2-k_1}\Big) = O(1).$$

Let $a = \lfloor M \log n \rfloor$ for a sufficiently large constant M. Using (S6.7), it follows (considering the lags up to a) that

$$
\begin{aligned}
&\sigma^{\tilde{Y}}_{(k_1-1)p+j_1,(k_2-1)p+j_2} \\
&= \tfrac{1}{2\lceil nb_n\rceil-1}\mathbb{E}\Big(\sum_{i_1=k_2-k_1+1}^{2\lceil nb_n\rceil-1} Z_{i_1+(k_1-1),\lceil nb_n\rceil+(k_1-1),j_1} \sum_{i_2=1}^{2\lceil nb_n\rceil-(k_2-k_1)-1} Z_{i_2+(k_2-1),\lceil nb_n\rceil+(k_2-1),j_2}\Big) \\
&+ O((nb_n)^{-1}) \\
&= \tfrac{1}{2\lceil nb_n\rceil-1}\mathbb{E}\Big(\sum_{i_1,i_2=1}^{2\lceil nb_n\rceil-(k_2-k_1)-1} G_{j_1,i_1,k_2}K(\tfrac{i_1+k_2-k_1-\lceil nb_n\rceil}{nb_n})G_{j_2,i_2,k_2}K(\tfrac{i_2-\lceil nb_n\rceil}{nb_n})\Big) + O((nb_n)^{-1}) \\
&= A + B + O(nb_n\chi^a + (nb_n)^{-1}), \qquad \text{(S6.17)}
\end{aligned}
$$

where the terms A and B are defined by

$$
\begin{aligned}
A &:= \tfrac{1}{(2\lceil nb_n\rceil-1)} \sum_{i=1}^{2\lceil nb_n\rceil-(k_2-k_1)-1} A_i, \qquad \text{(S6.18)} \\
A_i &= \mathbb{E}(G_{j_1,i,k_2,0}G_{j_2,i,k_2,0})K(\tfrac{i+k_2-k_1-\lceil nb_n\rceil}{nb_n})K(\tfrac{i-\lceil nb_n\rceil}{nb_n}) \\
B &= \tfrac{1}{(2\lceil nb_n\rceil-1)} \sum_{u=1}^{a}(B_{1,u}+B_{2,u}), \\
B_{1,u} &= \sum_{i=1}^{2\lceil nb_n\rceil-(k_2-k_1)-1-u} B_{1,u,i}, \qquad \text{(S6.19)} \\
B_{2,u} &=: \sum_{i=1}^{2\lceil nb_n\rceil-(k_2-k_1)-1-u} B_{2,u,i}. \qquad \text{(S6.20)}
\end{aligned}
$$

and

$$
B_{1,u,i} = \mathbb{E}(G_{j_1,i,k_2,u}G_{j_1,i,k_2,0})K(\tfrac{i+u+k_2-k_1-\lceil nb_n\rceil}{nb_n})K(\tfrac{i-\lceil nb_n\rceil}{nb_n}))
$$

$$
B_{2,u,i} = \mathbb{E}(G_{j_1,i,k_2,0}G_{j_2,i,k_2,u})K(\tfrac{i+k_2-k_1-\lceil nb_n\rceil}{nb_n})K(\tfrac{i+u-\lceil nb_n\rceil}{nb_n})
$$

Therefore, by (S6.17), we have that

$$\sigma^{\tilde{Y}}_{(k_1-1)p+j_1,(k_2-1)p+j_2} = \frac{1}{2\lceil nb_n\rceil - 1}\Big(\sum_{i=1}^{2\lceil nb_n\rceil-1-(k_2-k_1)} A_i + \sum_{u=1}^{a}\sum_{i=1}^{2\lceil nb_n\rceil-1-(k_2-k_1)-u}(B_{1,u,i}+B_{2,u,i})\Big) + O(nb_n\chi^a + (nb_n)^{-1}). \tag{S6.21}$$

Now for the term in (S6.6) we have

$$m_n\tilde{S}_{(r+k_2-k_1)m_n,(k_1-1)p+j_1}\tilde{S}_{rm_n,(k_2-1)p+j_2} = \Big(\sum_{i=r+k_2-k_1}^{r+k_2-k_1+m_n/2-1} - \sum_{i=r+k_2-k_1+m_n/2}^{r+k_2-k_1+m_n}\Big)Z_{i+k_1-1,\lceil nb_n\rceil+k_1-1,j_1}$$
$$\times\Big(\sum_{i=r}^{r+m_n/2-1} - \sum_{i=r+m_n/2}^{r+m_n}\Big)Z_{i+k_2-1,\lceil nb_n\rceil+k_2-1,j_2}$$
$$=\Big(\sum_{i=r}^{r+m_n/2-1} - \sum_{i=r+m_n/2}^{r+m_n}\Big)G_{j_1,i,k_2}K(\tfrac{i+k_2-k_1-\lceil nb_n\rceil}{nb_n}) \times \Big(\sum_{i=r}^{r+m_n/2-1} - \sum_{i=r+m_n/2}^{r+m_n}\Big)G_{j_2,i,k_2}K(\tfrac{i-\lceil nb_n\rceil}{nb_n}).$$

By Lemma S2, it follows that uniformly for $|k_2-k_1| \le 2\lceil nb_n\rceil - m_n$ and $1 \le r \le \lceil 2nb_n\rceil - m_n - (k_2-k_1)$,

$$\begin{aligned}
&m_n\mathbb{E}\tilde{S}_{(r+k_2-k_1)m_n,(k_1-1)p+j_1}\tilde{S}_{rm_n,(k_2-1)p+j_2}\\
&=\sum_{i=r}^{r+m_n}\mathbb{E}(G_{j_1,i,k_2}G_{j_2,i,k_2})K(\tfrac{i+k_2-k_1-\lceil nb_n\rceil}{nb_n})K(\tfrac{i-\lceil nb_n\rceil}{nb_n})\\
&+\sum_{u=1}^{a}\Big(\sum_{i=r}^{r+m_n-u}\big(\mathbb{E}(G_{j_1,i,(k_2+u)}G_{j_2,i,k_2})K(\tfrac{i+u+k_2-k_1-\lceil nb_n\rceil}{nb_n})K(\tfrac{i-\lceil nb_n\rceil}{nb_n})\\
&+\mathbb{E}(G_{j_2,i,(k_2+u)}G_{j_1,i,k_2})K(\tfrac{i+k_2-k_1-\lceil nb_n\rceil}{nb_n})K(\tfrac{i+u-\lceil nb_n\rceil}{nb_n})\big)\Big) + O(m_n\chi^a + a^2),
\end{aligned} \tag{S6.22}$$

where the the term $m_n\chi^a$ corresponds to the error of omitting terms in the sum with a large index a, and the term a^2 summarizes the error due to ignoring different signs in the product $\tilde{S}_{(r+k_2-k_1)m_n,(k_1-1)p+j_1}\tilde{S}_{rm_n,(k_2-1)p+j_2}$ (for each index u, we omit $2u$). Furthermore, by Assumption 2.1 and 3.2(3) it follows that uniformaly for $|u| \leq a$

$$\frac{1}{m_n}\sum_{i=r}^{r+m_n}\mathbb{E}(G_{j_1,i,k_2}G_{j_2,i,k_2})K(\tfrac{i+k_2-k_1-\lceil nb_n\rceil}{nb_n})K(\tfrac{i-\lceil nb_n\rceil}{nb_n}) = A_r + O(\frac{m_n}{nb_n}), \tag{S6.23}$$

$$\frac{1}{m_n}\sum_{i=r}^{r+m_n-u}\mathbb{E}(G_{j_1,i,(k_2+u)}G_{j_2,i,k_2})K(\tfrac{i+u+k_2-k_1-\lceil nb_n\rceil}{nb_n})K(\tfrac{i-\lceil nb_n\rceil}{nb_n}) = B_{1,u,r} + O(\frac{m_n}{nb_n}+\frac{a}{m_n}), \tag{S6.24}$$

$$\frac{1}{m_n}\sum_{i=r}^{r+m_n-u}\mathbb{E}(G_{j_2,i,(k_2+u)}G_{j_1,i,k_2})K(\tfrac{i+k_2-k_1-\lceil nb_n\rceil}{nb_n})K(\tfrac{i+u-\lceil nb_n\rceil}{nb_n}) = B_{2,u,r} + O(\frac{m_n}{nb_n}+\frac{a}{m_n})), \tag{S6.25}$$

where terms $A_r, B_{1,u,r}$ and $B_{2,u,r}$ are defined in equations (S6.18), (S6.19) and (S6.20), respectively. Notice that (S6.6) and expressions (S6.22), (S6.23), (S6.24) and (S6.25) yield that

$$\mathbb{E}\sigma^{\tilde{T}^\diamond}_{(k_1-1)p+j_1,(k_2-1)p+j_2} = \frac{1}{2\lceil nb_n\rceil - m_n}\Big\{\sum_{r=1}^{2\lceil nb_n\rceil-m_n-(k_2-k_1)}(A_r + O(\frac{m_n}{nb_n}))$$
$$+\sum_{u=1}^{a}\sum_{r=1}^{2\lceil nb_n\rceil-m_n-(k_2-k_1)}(B_{1,u,r}+B_{2,u,r}+O(\frac{m_n}{nb_n}+\frac{a}{m_n}))\Big\} + O\big(\chi^a + \frac{a^2}{m_n}\big). \tag{S6.26}$$

Lemma S2 implies

$$\max_{\substack{1\leq r\leq 2\lceil nb_n\rceil-(k_2-k_1)-1,\\ 1\leq k_1\leq k_2\leq (n-2\lceil nb_n\rceil+1), s=1,2}} B_{s,u,r} = O(\chi^u),$$

which yields in combination with equations (S6.21), (S6.26) with $a = M\log n$ for a sufficiently large constant M, and a similar argument applied to the case that $k_1 \geq k_2$,

$$\max_{\substack{1 \leq k_1, k_2 \leq (n-2\lceil nb_n \rceil+1) \\ |k_2-k_1| \leq 2\lceil nb_n \rceil - m_n, 1 \leq j_1, j_2 \leq p}} \Big| \mathbb{E}\sigma^{\tilde{T}^\diamond}_{(k_1-1)p+j_1,(k_2-1)p+j_2} - \sigma^{\tilde{Y}}_{(k_1-1)p+j_1,(k_2-1)p+j_2} \Big| = O\Big(\frac{\log^2 n}{m_n} + \frac{m_n \log n}{nb_n} \Big). \tag{S6.27}$$

Furthermore, using (S6.11), the Cauchy-Schwartz inequality, a similar argument as given in the proof of Lemma 1 of Zhou (2013) and Assumption 3.2(2) yield that

$$\Big\| \max_{\substack{1 \leq k_1 \leq k_2 \leq (n-2\lceil nb_n \rceil+1), \\ 1 \leq j_1, j_2 \leq p}} |\mathbb{E}\sigma^{\tilde{T}^\diamond}_{(k_1-1)p+j_1,(k_2-1)p+j_2} - \sigma^{\tilde{T}^\diamond}_{(k_1-1)p+j_1,(k_2-1)p+j_2}| \Big\|_{q/2} = O\Big(\sqrt{\frac{m_n}{nb_n}} (np)^{4/q} \Big). \tag{S6.28}$$

Combining (S6.27) and (S6.28), we obtain

$$\Big\| \max_{\substack{k_1, k_2, j_1, j_2 \\ |k_2-k_1| \leq 2\lceil nb_n \rceil - m_n}} |\sigma^{\tilde{Y}}_{(k_1-1)p+j_1,(k_2-1)p+j_2} - \sigma^{\tilde{T}^\diamond}_{(k_1-1)p+j_1,(k_2-1)p+j_2} \Big\|_{q/2} = O\Big(\frac{\log^2 n}{m_n} + \frac{m_n \log n}{nb_n} + \sqrt{\frac{m_n}{nb_n}} (np)^{4/q} \Big). \tag{S6.29}$$

Therefore the estimate (S6.8) follows combining (S6.16) and (S6.29). ◇

S6.3 Proof of Theorem 3

Similarly to (S7.1) and (S7.2) in the proof of Proposition S1 we obtain

$$\sup_{\substack{u \in [b_n, 1-b_n] \\ t \in [0,1]}} \frac{1}{\sigma(u,t)} \Big| \mathbb{E}(\hat{m}(u,t)) - m(u,t) \Big| \leq M \Big(\frac{1}{n} + b_n^4 \Big) \tag{S6.30}$$

for some constant M, where we have used the fact that, by Assumption 2.1, $\int K(v)v^2 dv = 0$. Moreover, by a similar but simpler argument as given in the proof of equation (B.7) in Lemma B.3 of Dette et al. (2019) we have for the quantity

$$\frac{(\hat{m}(u,t)-\mathbb{E}(\hat{m}(u,t)))}{\sigma(u,t)} = \frac{1}{nb_n\tilde{K}(u)}\sum_{i=1}^{n}\frac{G(\frac{i}{n},t,\mathcal{F}_i)}{\sigma(u,t)}K\Big(\frac{\frac{i}{n}-u}{b_n}\Big) := \Psi^{\sigma}(u,t)$$

the estimate

$$\Big\|\sup_{u\in[b_n,1-b_n],t\in[0,1]}\sqrt{nb_n}|\Phi^{\sigma}(u,t)-\Psi^{\sigma}(u,t)|\Big\|_q = O(b_n^{1-2/q}), \tag{S6.31}$$

where

$$\Phi^{\sigma}(u,t) = \frac{1}{nb_n\tilde{K}(u)}\sum_{i=1}^{n}\frac{G(\frac{i}{n},t,\mathcal{F}_i)}{\sigma(\frac{i}{n},t)}K(\frac{\frac{i}{n}-u}{b_n}).$$

Following the proof of Theorem 1 we find that

$$\sup_{x\in\mathbb{R}}\Big|\mathbb{P}\Big(\max_{b_n\le u\le 1-b_n,0\le t\le 1}\sqrt{nb_n}\big|\Phi^{\sigma}(u,t)\big|\le x\Big)-\mathbb{P}\Big(\Big|\frac{1}{\sqrt{nb_n}}\sum_{i=1}^{2\lceil nb_n\rceil-1}\tilde{Y}_i^{\sigma}\Big|_\infty\le x\Big)\Big|$$
$$= O\Big((nb_n)^{-(1-11\iota)/8}+\Theta\Big(\big((np)^{1/q^*}((nb_n)^{-1}+1/p)\big)^{\frac{q^*}{q^*+1}},np\Big)\Big).$$

Combining this result with Lemma S1 (with $X=\max_{\substack{b_n\le u\le 1-b_n\\0\le t\le 1}}\sqrt{nb_n}\big|\Phi^{\sigma}(u,t)\big|$, $Y=\frac{1}{\sqrt{nb_n}}\sum_{i=1}^{2\lceil nb_n\rceil-1}\tilde{Y}_i^{\sigma}$, $X'=\max_{b_n\le u\le 1-b_n,0\le t\le 1}\sqrt{nb_n}\big|\Psi^{\sigma}(u,t)\big|$) and (S6.31) gives

$$\sup_{x\in\mathbb{R}}\Big|\mathbb{P}\Big(\max_{b_n\le u\le 1-b_n,0\le t\le 1}\sqrt{nb_n}\big|\Psi^{\sigma}(u,t)\big|\le x\Big)-\mathbb{P}\Big(\Big|\frac{1}{\sqrt{nb_n}}\sum_{i=1}^{2\lceil nb_n\rceil-1}\tilde{Y}_i^{\sigma}\Big|_\infty\le x\Big)\Big|$$
$$= O\Big((nb_n)^{-(1-11\iota)/8}+\Theta(\big((np)^{1/q^*}((nb_n)^{-1}+1/p)\big)^{\frac{q^*}{q^*+1}},np)$$
$$+\mathbb{P}\Big(\sup_{u\in[b_n,1-b_n],t\in[0,1]}\sqrt{nb_n}|\Phi^{\sigma}(u,t)-\Psi^{\sigma}(u,t)|>\delta\Big)+\Theta(\delta,np)\Big)$$
$$= O\Big((nb_n)^{-(1-11\iota)/8}+\Theta(\big((np)^{1/q^*}((nb_n)^{-1}+1/p)\big)^{\frac{q^*}{q^*+1}},np)+\Theta(\delta,np)+\frac{b_n^{q-2}}{\delta^q}\Big) \tag{S6.32}$$

Taking $\delta=b_n^{\frac{q-2}{q+1}}$ we obtain for the last two terms in (S6.32)

$$\Theta(\delta,np)+\frac{b_n^{q-2}}{\delta^q} = O\Big(\Theta(b_n^{\frac{q-2}{q+1}},np)\Big).$$

On the other hand, (S6.30), (S6.32) and Lemma S1 (with $X = \max_{b_n \le u \le 1-b_n, 0 \le t \le 1} \sqrt{nb_n}\big|\Psi^\sigma(u,t)\big|$, $Y = \frac{1}{\sqrt{nb_n}} \sum_{i=1}^{2\lceil nb_n \rceil - 1} \tilde{Y}_i^\sigma$, $X' = \max_{b_n \le u \le 1-b_n, 0 \le t \le 1} \sqrt{nb_n}\big|\hat{\Delta}^\sigma(u,t)\big|$ and $\delta = M\sqrt{nb_n}(\frac{1}{n} + b_n^4)$ with a sufficiently large constant M) yield

$$\sup_{x \in \mathbb{R}} \Big|\mathbb{P}\Big(\max_{b_n \le u \le 1-b_n, 0 \le t \le 1} \sqrt{nb_n}\big|\hat{\Delta}^\sigma(u,t)\big| \le x\Big) - \mathbb{P}\Big(\Big|\frac{1}{\sqrt{nb_n}} \sum_{i=1}^{2\lceil nb_n \rceil - 1} \tilde{Y}_i^\sigma\Big|_\infty \le x\Big)\Big|$$
$$= O\Big((nb_n)^{-(1-11\iota)/8} + \Theta\big(\big((np)^{1/q^*}((nb_n)^{-1} + 1/p)\big)^{\frac{q^*}{q^*+1}}, np\big)$$
$$+ \Theta\Big(\sqrt{nb_n}(b_n^4 + \frac{1}{n}), np\Big) + \Theta(b_n^{\frac{q-2}{q+1}}, np)\Big).$$

□

S6.4 Proof of Theorem 4

Proof. Recall that $g_n = \frac{w^{5/2}}{n}\tau_n^{-1/q'} + w^{1/2}n^{-1/2}\tau_n^{-1/2-2/q'} + w^{-1}$ and let η_n be a sequence of positive numbers such that $\eta_n \to \infty$ and $(g_n + \tau_n)\eta_n \to 0$ (note that $g_n + \tau_n$ is the convergence rate of the estimator $\hat{\sigma}^2$ in Proposition 1). Define the $\mathcal{F}_n$ measurable event

$$A_n = \Big\{\sup_{u \in [0,1], t \in [0,1]} |\hat{\sigma}^2(u,t) - \sigma^2(u,t)| > (g_n + \tau_n)\eta_n\Big\},$$

then Proposition 1 and Markov's inequality yield

$$\mathbb{P}(A_n) = O\big(\eta_n^{-q'}\big). \tag{S6.33}$$

Then by Theorem 3, Proposition 1 and Lemma S1 we have

$$\mathfrak{P}^{\hat{\sigma}} = \sup_{x \in \mathbb{R}} \Big|\mathbb{P}\Big(\max_{b_n \le u \le 1-b_n, 0 \le t \le 1} \sqrt{nb_n}|\hat{\Delta}^{\hat{\sigma}(u,t)}| \le x\Big) - \mathbb{P}\Big(\Big|\frac{1}{\sqrt{nb_n}} \sum_{i=1}^{2\lceil nb_n \rceil - 1} \tilde{Y}_i^\sigma\Big|_\infty \le x\Big)\Big| = o_p(1). \tag{S6.34}$$

Let $T_k^{\hat{\sigma}}$ denote the statistic $T_k^{\hat{\sigma},(r)}$ in step (d) of Algorithm 2 generated by one bootstrap iteration and define for integers a, b the quantities

$$T_{ap+b}^{\hat{\sigma},\diamond} = \sum_{j=1}^{2\lceil nb_n\rceil - m_n'} \hat{S}^{\hat{\sigma}}_{jm_n',(a-1)p+b} R_{k+j-1}, a = 1, ...n - 2\lceil nb_n\rceil + 1, 1 \le b \le p$$

$$T^{\hat{\sigma},\diamond} := ((T_1^{\hat{\sigma},\diamond})^\top, \dots, (T^{\hat{\sigma},\diamond}_{(n-2\lceil nb_n\rceil+1)p})^\top)^\top = \big(T_1^{\hat{\sigma}\top}, \dots, T_{n-2\lceil nb_n\rceil+1}^{\hat{\sigma}}{}^\top\big)^\top$$

and therefore

$$T^{\hat{\sigma}} = |T^{\hat{\sigma},\diamond}|_\infty = \max_{1\le k\le n-2\lceil nb_n\rceil+1} |T_k^{\hat{\sigma}}|_\infty$$

We recall the notation (4.2), introduce the $(n-2\lceil nb_n\rceil+1)p$-dimensional random vectors $\hat{S}^{\sigma,*}_{jm_n} = \sum_{r=j}^{j+m_n-1} \tilde{Z}^\sigma_r$, and

$$\hat{S}^\sigma_{jm_n'} = \frac{1}{\sqrt{m_n'}} \hat{S}^{\sigma,*}_{j,\lfloor m_n/2\rfloor} - \frac{1}{\sqrt{m_n'}} \hat{S}^{\sigma,*}_{j+\lfloor m_n/2\rfloor+1,\lfloor m_n/2\rfloor} ,$$

and consider

$$T_k^\sigma = \sum_{j=1}^{2\lceil nb_n\rceil - m_n'} \hat{S}^\sigma_{jm_n',[(k-1)p+1:kp]} R_{k+j-1} , \quad k = 1, \dots, n - 2\lceil nb_n\rceil + 1,$$

$$T^{\sigma,\diamond} := ((T_1^{\sigma,\diamond})^\top, \dots, (T^{\sigma,\diamond}_{(n-2\lceil nb_n\rceil+1)p})^\top)^\top = \big(T_1^{\sigma\top}, \dots, T_{n-2\lceil nb_n\rceil+1}^{\sigma}{}^\top\big)^\top ,$$

where $T^{\sigma,\diamond}$ is obtained from $T^{\hat{\sigma},\diamond}$ by replacing $\hat{\sigma}$ by σ. Similar arguments as given in the proof of Theorem 2 show, that it is sufficient to show the estimate

$$\begin{aligned} &\sup_{x\in\mathbb{R}} \Big| \mathbb{P}(|T^{\hat{\sigma},\diamond}/\sqrt{2\lceil nb_n\rceil - m_n'}|_\infty \le x|\mathcal{F}_n) - \mathbb{P}\Big(\frac{1}{\sqrt{2nb_n}}\Big| \sum_{i=1}^{2\lceil nb_n\rceil-1} \tilde{Y}_i^\sigma\Big|_\infty \le x\Big)\Big| \\ &= O_p\Big(\vartheta_n^{1/3}\{1 \vee \log(\frac{np}{\vartheta_n})\}^{2/3} + \Theta\big(\sqrt{m_n \log np}(\frac{1}{\sqrt{nb_n}} + b_n^3)(np)^{\frac{1}{q}}\big)^{q/(q+1)}, np\Big) \\ &+ \Theta\big(\big(\sqrt{m_n \log np}((g_n + \tau_n)\eta_n)(np)^{\frac{1}{q}}\big)^{q/(q+1)}, np\big) + \eta_n^{-q'}\Big) \qquad \text{(S6.35)} \end{aligned}$$

where ϑ_n is defined in Theorem 2. The assertion of Theorem 4 then follows from (S6.34).

Now we prove (S6.35). By the first step in the proof of Theorem 2 it follows that

$$\begin{aligned}
&\sup_{x\in\mathbb{R}}\Big|\mathbb{P}(|T^{\sigma,\diamond}/\sqrt{2\lceil nb_n\rceil - m_n'}|_\infty \le x|\mathcal{F}_n) - \mathbb{P}\Big(\frac{1}{\sqrt{2nb_n}}\Big|\sum_{i=1}^{2\lceil nb_n\rceil-1}\tilde{Y}_i^{\sigma}\Big|_\infty \le x\Big)\Big| \\
&= O_p\Big(\vartheta_n^{1/3}\{1\vee\log(\frac{np}{\vartheta_n})\}^{2/3} \\
&+\Theta\Big(\big(\sqrt{m_n\log np}(\frac{1}{\sqrt{nb_n}}+b_n^3)(np)^{\frac{1}{q}}\big)^{q/(q+1)}, np\Big)\Big). \qquad \text{(S6.36)}
\end{aligned}$$

By similar arguments as given in the proof of Theorem 2 we have

$$\mathbb{E}\big(|T^{\sigma,\diamond}-T^{\hat{\sigma},\diamond}|_\infty^q\mathbf{1}(A_n)\big|\mathcal{F}_n\big) \le M\Big|\sqrt{\log np}\max_{r=1}^{(n-2\lceil nb_n\rceil+1)p}\Big(\sum_{j=1}^{\lceil 2nb_n\rceil-m_n'}(\hat{S}^{\sigma}_{jm_n',r}-\hat{S}^{\hat{\sigma}}_{jm_n',r})^2\mathbf{1}(A_n)\Big)^{1/2}\Big|^q \qquad \text{(S6.37)}$$

for some large constant M almost surely, and the triangle inequality, a similar argument as given in the proof of Proposition 1.1 in Dette and Wu (2022) and (S6.11) yield

$$\frac{1}{\sqrt{2\lceil nb_n\rceil - m_n}}\Big\|\max_{r=1}^{(n-2\lceil nb_n\rceil+1)p}\Big(\sum_{j=1}^{\lceil 2nb_n\rceil-m_n'}(\hat{S}^{\sigma}_{jm_n',r}-\hat{S}^{\hat{\sigma}}_{jm_n',r})^2\mathbf{1}(A)\Big)^{1/2}\Big\|_q = O\big(\sqrt{m_n}(g_n+\tau_n)\eta_n(np)^{\frac{1}{q}}\big).$$

This together with the (conditional version) of Lemma S1 and (S6.37) shows that

$$\sup_{x\in\mathbb{R}}\Big|\mathbb{P}\Big(\frac{|T^{\hat\sigma,\diamond}|_\infty}{\sqrt{2\lceil nb_n\rceil-m_n}}>x\Big|\mathcal{F}_n\Big)-\mathbb{P}\Big(\frac{1}{\sqrt{2nb_n}}\Big|\sum_{i=1}^{2\lceil nb_n\rceil-1}\tilde Y_i^\sigma\Big|_\infty>x\Big)\Big|$$
$$\leq\sup_{x\in\mathbb{R}}\Big|\mathbb{P}\Big(\frac{|T^{\sigma,\diamond}|_\infty}{\sqrt{2\lceil nb_n\rceil-m_n}}>x\Big|\mathcal{F}_n\Big)-\mathbb{P}\Big(\frac{1}{\sqrt{2nb_n}}\Big|\sum_{i=1}^{2\lceil nb_n\rceil-1}\tilde Y_i^\sigma\Big|_\infty>x)\Big|$$
$$+\mathbb{P}\Big(\frac{|T^{\diamond,\sigma}-T^{\diamond,\hat\sigma}|_\infty}{\sqrt{2\lceil nb_n\rceil-m_n}}>\delta\Big|\mathcal{F}_n\Big)+O\big(\Theta(\delta,np)\big)$$
$$\leq\sup_{x\in\mathbb{R}}\Big|\mathbb{P}\Big(\frac{|T^{\sigma,\diamond}|_\infty}{\sqrt{2\lceil nb_n\rceil-m_n}}>x\Big|\mathcal{F}_n\Big)-\mathbb{P}\Big(\frac{1}{\sqrt{2nb_n}}\Big|\sum_{i=1}^{2\lceil nb_n\rceil-1}\tilde Y_i^\sigma\Big|_\infty>x\Big)\Big|$$
$$+O_p\big(\delta^{-q}\big(\sqrt{m_n\log np}((g_n+\tau_n)\eta_n)(np)^{\frac{1}{q}}\big)^q\big)+O\big(\Theta(\delta,np)+\eta_n^{-q'}\big),$$

where we used Markov's inequality and (S6.33). Taking

$$\delta=\big(\sqrt{m_n\log np}((g_n+\tau_n)\eta_n)(np)^{\frac{1}{q}}\big)^{q/(q+1)}$$

and observing (S6.36) yields (S6.35) and proves the assertion. ◇

S7 Proposition S1 and Proof of Proposition 1

S7.1 Proposition S1

The proof of Theorems 1 is based on the following auxiliary result providing a Gaussian approximation for the maximum deviation of the quantity $\sqrt{nb_n}|\hat\Delta(u,t_v)|$ over the grid of $\{1/n,...,n/n\}\times\{t_1,...,t_p\}$ where $t_v=\frac{v}{p}$ $(v=1,\dots,p)$.

Proposition S1. *Assume that* $n^{1+a}b_n^9=o(1)$, $n^{a-1}b_n^{-1}=o(1)$ *for some* $0<a<4/5$, *and let Assumptions 3.1, 3.2 and 2.1 be satisfied.*

(i) For a fixed $u \in (0,1)$*, let* $Y_1(u), \ldots, Y_n(u)$ *denote a sequence of centered p-dimensional Gaussian vectors such that* $Y_i(u)$ *has the same auto-covariance structure of the vector* $Z_i(u)$ *defined in (2.5). If* $p = O(\exp(n^\iota))$ *for some* $0 \le \iota < 1/11$*, then*

$$\mathfrak{P}_{p,n}(u) := \sup_{x \in \mathbb{R}} \Big| \mathbb{P}\Big(\max_{1 \le v \le p} \sqrt{nb_n} |\hat{\Delta}(u, t_v)| \le x \Big) - \mathbb{P}\Big(\Big| \frac{1}{\sqrt{nb_n}} \sum_{i=1}^n Y_i(u) \Big|_\infty \le x \Big) \Big|$$
$$= O\Big((nb_n)^{-(1-11\iota)/8} + \Theta\Big(\sqrt{nb_n} \Big(b_n^4 + \frac{1}{n} \Big), p \Big) \Big)$$

(ii) Let $\tilde{Y}_1, \ldots, \tilde{Y}_{2\lceil nb_n \rceil - 1}$ *denote independent* $(n - 2\lceil nb_n \rceil + 1)p$*-dimensional centered Gaussian vectors with the same auto-covariance structure as the vector* $\tilde{Z}_i$ *in (2.7). If* $np = O(\exp(n^\iota))$ *for some* $0 \le \iota < 1/11$*, then*

$$\mathfrak{P}_{p,n} := \sup_{x \in \mathbb{R}} \Big| \mathbb{P}\Big(\max_{\lceil nb_n \rceil \le l \le n - \lceil nb_n \rceil, 1 \le v \le p} \sqrt{nb_n} |\hat{\Delta}(\tfrac{l}{n}, t_v)| \le x \Big) - \mathbb{P}\Big(\Big| \frac{1}{\sqrt{nb_n}} \sum_{i=1}^{2\lceil nb_n \rceil - 1} \tilde{Y}_i \Big|_\infty \le x \Big) \Big|$$
$$= O\Big((nb_n)^{-(1-11\iota)/8} + \Theta\Big(\sqrt{nb_n} \Big(b_n^4 + \frac{1}{n} \Big), np \Big) \Big)$$

Proof. Using Assumptions 3.1, 2.1 and a Taylor expansion we obtain

$$\sup_{\substack{u \in [b_n, 1-b_n] \\ t \in [0,1]}} \Big| \mathbb{E}(\hat{m}(u,t)) - m(u,t) - b_n^2 \int K(v) v^2 dv \frac{\partial^2}{\partial u^2} m(u,t)/2 \Big| \le M \Big(\frac{1}{n} + b_n^4 \Big) \quad \text{(S7.1)}$$

for some constant M. Notice that by assumption $\int K(v)v^2 dv = 0$. Notice that for $u \in [b_n, 1 - b_n]$,

$$\hat{m}(u,t) - \mathbb{E}(\hat{m}(u,t)) = \frac{1}{nb_n \tilde{K}(u)} \sum_{i=1}^n G(\tfrac{i}{n}, t, \mathcal{F}_i) K\Big(\frac{\frac{i}{n} - u}{b_n} \Big) \quad \text{(S7.2)}$$
$$= \frac{1}{nb_n \tilde{K}(u)} \sum_{i=\lceil n(u-b_n) \rceil}^{\lfloor n(u+b_n) \rfloor} G(\tfrac{i}{n}, t, \mathcal{F}_i) K\Big(\frac{\frac{i}{n} - u}{b_n} \Big).$$

Therefore, observing the definition of $Z_i(u)$ in (2.5) we have (notice that $Z_i(u)$ is a vector of zero if $|\frac{i}{n}-u|\geq b_n$)

$$\max_{1\leq v\leq p}\sqrt{nb_n}|\hat{m}(u,t_v)-\mathbb{E}(\hat{m}(u,t_v))|\tilde{K}(u)=\Big|\frac{1}{\sqrt{nb_n}}\sum_{i=\lceil n(u-b_n)\rceil}^{\lfloor n(u+b_n)\rfloor}Z_i(u)\Big|_\infty.$$

We will now apply Corollary 2.2 of Zhang and Cheng (2018) and check its assumptions first. By Assumption 3.2(2) and the fact that the kernel is bounded it follows that

$$\max_{1\leq l\leq p}\sup_i\|Z_{i,l}(u)-Z_{i,l}^{(i-j)}(u)\|_2=O(\chi^j),$$

where for any (measurable function) $g=g(\mathcal{F}_i)$, we define for $j\leq i$ the function $g^{(j)}$ by $g^{(j)}=g(\mathcal{F}_i^{(j)})$, where $\mathcal{F}_i^{(j)}=(\ldots,\eta_{j-1},\eta_j',\eta_{j+1},\ldots,\eta_i)$ and $\{\eta_i'\}_{i\in\mathbb{Z}}$ is an independent copy of $\{\eta_i\}_{i\in\mathbb{Z}}$ (recall that $\mathcal{F}_i=(\eta_{-\infty},...,\eta_i)$). Lemma S3 in Section S8 shows that condition (9) in the paper of Zhang and Cheng (2018) is satisfied. Moreover Assumption 3.2(1) implies condition (13) in this reference. Observing that for random vector $v=(v_1,...,v_p)^\top$ and all $x\in\mathbb{R}$

$$\{|v|_\infty\leq x\}=\Big\{\max(v_1,...,v_p,-v_1,...,-v_p)\leq x\Big\},$$

we can use Corollary 2.2 of Zhang and Cheng (2018)

$$\sup_{x\in\mathbb{R}}\Big|\mathbb{P}\Big(\Big|\frac{1}{\sqrt{nb_n}}\sum_{i=1}^nY_i(u)\Big|_\infty\leq x\Big)-\mathbb{P}\Big(\frac{1}{\sqrt{nb_n}}\Big|\sum_{i=1}^nZ_i(u)\Big|_\infty\leq x\Big)\Big|=O((nb_n)^{-(1-11\iota')/8}).\tag{S7.3}$$

Therefore by (S7.1), (S7.3) and Lemma S1, and the fact that $\tilde{K}(u)=1+O(\frac{1}{nb_n})$ for

$b_n \le u \le 1-b_n$

$$\sup_{x\in\mathbb{R}}\Big|\mathbb{P}\Big(\max_{1\le v\le p}\sqrt{nb_n}|\hat{\Delta}(u,t_v)|\le x\Big)-\mathbb{P}\Big(\Big|\frac{1}{\sqrt{nb_n}}\sum_{i=1}^{n}Y_i(u)\tilde{K}(u)\Big|_\infty\le x\Big)\Big|$$
$$=O\Big((nb_n)^{-(1-11\iota)/8}+\Theta\Big(\sqrt{nb_n}\Big(b_n^4+\frac{1}{nb_n}\Big),p\Big)\Big). \tag{S7.4}$$

Using Theorem 2 of Chernozhukov et al. (2015), it follows that

$$\sup_{x\in\mathbb{R}}\Big|\mathbb{P}\Big(\Big|\frac{1}{\sqrt{nb_n}}\sum_{i=1}^{n}Y_i(u)\Big|_\infty\le x\Big)-\mathbb{P}\Big(\Big|\frac{1}{\sqrt{nb_n}}\sum_{i=1}^{n}Y_i(u)\tilde{K}(u)\Big|_\infty\le x\Big)\Big|$$
$$=O\Big((nb_n)^{-1/3}\log^{2/3}(npb_n)\Big). \tag{S7.5}$$

Since $p=O(\exp(n^\iota))$, Then part (i) of the assertion follows from (S7.4) and (S7.5). For part (ii), notice that $\tilde{K}(i/n)=\tilde{K}(j/n)$ for $i,j\in\mathbb{Z}$ such that $b_n\le i/n, j/n\le 1-b_n$. Let $\tilde{K}=\tilde{K}(\lfloor n/2\rfloor/n)$. Further note that by the definition of the vector $\tilde{Z}_i$ in (2.7) we have that (Recall the notation $W_n(u,t)$ in (S6.1))

$$\max_{1\le v\le p}\max_{\lceil nb_n\rceil\le l\le n-\lceil nb_n\rceil}\tilde{K}|W_n(\tfrac{l}{n},t_v)|=\max_{\lceil nb_n\rceil\le l\le n-\lceil nb_n\rceil}\Big|\frac{1}{\sqrt{nb_n}}\sum_{i=1}^{n}Z_i(\frac{l}{n})\Big|_\infty=\Big|\frac{1}{\sqrt{nb_n}}\sum_{i=1}^{2\lceil nb_n\rceil-1}\tilde{Z}_i\Big|_\infty. \tag{S7.6}$$

Let $\tilde{Z}_{i,s}$ denote the sth entry of the vector $\tilde{Z}_i$ defined in (2.7) $(1\le s\le (n-2\lceil nb_n\rceil+1)p)$. By Assumption 3.2(2) it follows that

$$\max_{1\le s\le (n-2\lceil nb_n\rceil+1)p}\sup_i\|\tilde{Z}_{i,s}-\tilde{Z}_{i,s}^{(i-j)}\|_2=O(\chi^j).$$

By Lemma S3 in Section S8 we obtain the inequality

$$c_1\le\min_{1\le j\le (n-2\lceil nb_n\rceil+1)p}\tilde{\sigma}_{j,j}\le\max_{1\le j\le (n-2\lceil nb_n\rceil+1)p}\tilde{\sigma}_{j,j}\le c_2$$

for the quantities

$$\tilde{\sigma}_{j,j} := \frac{1}{2\lceil nb_n \rceil - 1} \sum_{i,l=1}^{2\lceil nb_n \rceil - 1} \mathrm{Cov}(\tilde{Z}_{i,j}, \tilde{Z}_{l,j}).$$

Therefore condition (9) in the paper of Zhang and Cheng (2018) holds, and condition (13) in this reference follows from Assumption 3.2(1). As a consequence, Corollary 2.2 in Zhang and Cheng (2018) (the validity of Corollary 2.2 of Zhang and Cheng (2018) for $\tilde{Z}_i$ can be verified via the argument of Proposition 2.1, A.1 and Theorem 2.1 of that paper and via (S7.6); details are omitted for the sake of brevity) yields

$$\sup_{x\in\mathbb{R}} \Big| \mathbb{P}\Big(\max_{\substack{\lceil nb_n \rceil \le l_1 \le n - \lceil nb_n \rceil \\ 1 \le l_2 \le p}} \tilde{K} |W_n(\tfrac{l_1}{n}, \tfrac{l_2}{p})| \le x \Big) - \mathbb{P}\Big(\Big| \frac{1}{\sqrt{nb_n}} \sum_{i=1}^{2\lceil nb_n \rceil - 1} \tilde{Y}_i \Big|_\infty \le x \Big) \Big| = O((nb_n)^{-(1-11\iota)/8}). \tag{S7.7}$$

Using Theorem 2 of Chernozhukov et al. (2015) and the the fact that $\tilde{K} = 1 + O(\frac{1}{nb_n})$, it follows that

$$\sup_{x\in\mathbb{R}} \Big| \mathbb{P}\Big(\Big| \frac{1}{\sqrt{nb_n}} \sum_{i=1}^{n} \tilde{Y}_i \Big|_\infty \le x \Big) - \mathbb{P}\Big(\Big| \frac{1}{\sqrt{nb_n}} \sum_{i=1}^{n} \tilde{Y}_i \tilde{K} \Big|_\infty \le x \Big) \Big| = O\Big((nb_n)^{-1/3} \log^{2/3}(n^2 p b_n) \Big). \tag{S7.8}$$

Consequently part (ii) follows by the same arguments given in the proof of part (i) via an application of Lemma S1. ♢

S7.2 Proof of Proposition 1

Proof. Define $\tilde{S}^G_{k,r}(t) = \frac{1}{\sqrt{r}} \sum_{i=k}^{k+r-1} G(i/n, t, \mathcal{F}_i)$, and define for $u \in [w/n, 1 - w/n]$

$$\tilde{\Delta}_j(t) = \frac{\tilde{S}^G_{j-w+1,w}(t) - \tilde{S}^G_{j+1,w}(t)}{\sqrt{w}} \ , \quad \tilde{\sigma}^2(u,t) = \sum_{j=1}^{n} \frac{w \tilde{\Delta}_j^2(t)}{2} \bar{\omega}(u,j)$$

as the analogs of $\Delta_j(t)$ defined in the main article and the quantities in (2.16), respectively. We also use the convention $\tilde{\sigma}^2(u,t) = \tilde{\sigma}^2(w/n,t)$ and $\tilde{\sigma}^2(u,t) = \tilde{\sigma}^2(1-w/n,t)$ if $u \in [0, w/n)$ and $u \in (1-w/n, 1]$, respectively. Assumption 3.1 and the mean value theorem yield

$$\max_{w \le j \le n-w} \sup_{0 \le t \le 1} |\tilde{\Delta}_j(t) - \Delta_j(t)| = \max_{w \le j \le n-w} \sup_{0 \le t \le 1} \Big| \sum_{r=j-w+1}^{j} m(r/n,t) - \sum_{r=j+1}^{j+w} m(r/n,t) \Big| = O(w/n). \tag{S7.9}$$

On the other hand, Assumption 3.2 and Assumption 3.3 and similar arguments as given in the proof of Lemma 3 of Zhou and Wu (2010) give

$$\max_j \|\tilde{\Delta}_j(t)\|_{q'} = O(\sqrt{w}), \ \ \max_j \Big\| \frac{\partial}{\partial t} \tilde{\Delta}_j(t) \Big\|_{q'} = O(\sqrt{w}). \tag{S7.10}$$

Here we use the convention that $\frac{\partial}{\partial t}\tilde{\Delta}_j|_{t=0} = \frac{\partial}{\partial t}\tilde{\Delta}_j|_{t=0+}$, $\frac{\partial}{\partial t}\tilde{\Delta}_j|_{t=1} = \frac{\partial}{\partial t}\tilde{\Delta}_j|_{t=1-}$. Moreover, Proposition B.1. of Dette et al. (2019) yields

$$\max_j \Big\| \sup_t |\tilde{\Delta}_j(t)| \Big\|_{q'} = O(\sqrt{w}). \tag{S7.11}$$

Now we introduce the notation $C_j(t) = \tilde{\Delta}_j(t) - \Delta_j(t)$ (note that this quantity is not random) and obtain by (S7.9) the representation

$$\begin{aligned} \tilde{\sigma}^2(u,t) - \hat{\sigma}^2(u,t) &= \sum_{j=1}^{n} \frac{w(2\tilde{\Delta}_j(t) - C_j(t))C_j(t)}{2} \bar{w}(u,j) \\ &= \sum_{j=1}^{n} w\tilde{\Delta}_j(t)C_j(t)\bar{\omega}(u,j) + O(w^3/n^2) \end{aligned} \tag{S7.12}$$

uniformly with respect to u, t. Furthermore, by (S7.9) we have

$$\sup_{t\in[0,1]}\Big|\sum_{j=1}^{n} w\tilde{\Delta}_j(t)C_j(t)\bar{\omega}(u,j)\Big| \le W^{\diamond}(u) := M(w/n)\sum_{j=1}^{n} w \sup_{t\in[0,1]}|\Delta_j(t)|\bar{\omega}(u,j)\ ,$$

where M is a sufficiently large constant. Notice that $W^{\diamond}(u)$ is differentiable with respect to the variable u. Therefore it follows from the triangle inequality, (S7.11) and Proposition B.1 of Dette et al. (2019), that

$$\Big\|\sup_{u\in[\gamma_n,1-\gamma_n]}|W^{\diamond}(u)|\Big\|_{q'} = O\Big(\frac{w^{5/2}}{n}\tau_n^{-1/q'}\Big). \tag{S7.13}$$

Combining (S7.12)– (S7.13), we obtain

$$\Big\|\sup_{\substack{u\in[\gamma_n,1-\gamma_n]\\ t\in[0,1]}}|\tilde{\sigma}^2(u,t)-\hat{\sigma}^2(u,t)|\Big\|_{q'} = O\Big(\frac{w^{5/2}}{n}\tau_n^{-1/q'} + w^3/n^2\Big). \tag{S7.14}$$

By Burkholder inequality (see for example Wu, 2005) in $\mathcal{L}^{q'/2}$ norm, (S7.10) and similar arguments as given in the proof of Lemma 3 in Zhou and Wu (2010) we have

$$\sup_{\substack{u\in[\gamma_n,1-\gamma_n]\\ t\in[0,1]}}\Big\|\tilde{\sigma}^2(u,t)-\mathbb{E}(\tilde{\sigma}^2(u,t))\Big\|_{q'/2} = O\big(w^{1/2}n^{-1/2}\tau_n^{-1/2}\big),$$

$$\sup_{\substack{u\in[\gamma_n,1-\gamma_n]\\ t\in[0,1]}}\Big\|\frac{\partial}{\partial t}(\tilde{\sigma}^2(u,t)-\mathbb{E}(\tilde{\sigma}^2(u,t)))\Big\|_{q'/2} = O\big(w^{1/2}n^{-1/2}\tau_n^{-1/2}\big),$$

$$\sup_{\substack{u\in[\gamma_n,1-\gamma_n]\\ t\in[0,1]}}\Big\|\big(\frac{\partial}{\partial u}+\frac{\partial^2}{\partial u\partial t}\big(\tilde{\sigma}^2(u,t)-\mathbb{E}(\tilde{\sigma}^2(u,t))\big)\Big\|_{q'/2} = O\big(w^{1/2}n^{-1/2}\tau_n^{-1/2-1}\big).$$

It can be shown by similar but simpler argument as given in the proof of Proposition B.2 of Dette et al. (2019) that these estimates imply

$$\|\sup_{\substack{u\in[\gamma_n,1-\gamma_n]\\ t\in[0,1]}}|\tilde{\sigma}^2(u,t)-\mathbb{E}(\tilde{\sigma}^2(u,t))|\|_{q'/2} = O\big(w^{1/2}n^{-1/2}\tau_n^{-1/2-2/q'}\big). \tag{S7.15}$$

Moreover, it follows from the proof of Theorem 4.4 of Dette and Wu (2019) that

$$\sup_{\substack{u\in[\gamma_n,1-\gamma_n]\\ t\in[0,1]}}\Big|\mathbb{E}\tilde{\sigma}^2(u,t)-\sigma^2(u,t)\Big| = O\Big(\sqrt{w/n}+w^{-1}+\tau_n^2\Big),$$

$$\sup_{\substack{u\in[0,\gamma_n)\cup(1-\gamma_n,1]\\ t\in[0,1]}}\Big|\mathbb{E}\tilde{\sigma}^2(u,t)-\sigma^2(u,t)\Big| = O\Big(\sqrt{w/n}+w^{-1}+\tau_n\Big) \tag{S7.16}$$

and the assertion is a consequence of (S7.14), (S7.15) and (S7.16). ◊

S8 Some auxiliary results

This section contains several technical lemmas, which will be used in the proofs of the main results in Section S6.

Lemma S1. *For any random vectors X, X', Y, and $\delta\in\mathbb{R}$, we have that*

$$\sup_{x\in\mathbb{R}}|\mathbb{P}(|X'|>x)-\mathbb{P}(|Y|>x)| \le \sup_{x\in\mathbb{R}}|\mathbb{P}(|X|>x)-\mathbb{P}(|Y|>x)| + \mathbb{P}(|X-X'|>\delta)+2\sup_{x\in\mathbb{R}}\mathbb{P}(|Y-x|\le\delta). \tag{S8.1}$$

Furthermore, if $Y=(Y_1,...,Y_p)^\top$ is a p-dimensional Gaussian vector and there exist positive constants $c_1\le c_2$ such that for all $1\le j\le p$, $c_1\le\mathbb{E}(Y_j^2)\le c_2$, then

$$\sup_{x\in\mathbb{R}}|\mathbb{P}(|X'|>x)-\mathbb{P}(|Y|_\infty>x)| \le \sup_{x\in\mathbb{R}}|\mathbb{P}(|X|>x)-\mathbb{P}(|Y|_\infty>x)|+\mathbb{P}(|X-X'|>\delta) + C\Theta(\delta,p), \tag{S8.2}$$

where C is a constant only dependent on c_1 and c_2.

Proof of Lemma S1. By triangle inequality, we shall see that

$$\mathbb{P}(|X'| > x) - \mathbb{P}(|Y| > x) \leq \mathbb{P}(|X' - X| > \delta) + \mathbb{P}(|X| > x - \delta) - \mathbb{P}(|Y| > x), \quad \text{(S8.3)}$$

$$\mathbb{P}(|X'| > x) - \mathbb{P}(|Y| > x) \geq -\mathbb{P}(|X' - X| > \delta) + \mathbb{P}(|X| > x + \delta) - \mathbb{P}(|Y| > x). \quad \text{(S8.4)}$$

Notice that right-hand side of (S8.3) is

$$\mathbb{P}(|X' - X| > \delta) + \mathbb{P}(|X| > x - \delta) - \mathbb{P}(|Y| > x - \delta) + \mathbb{P}(|Y| > x - \delta) - \mathbb{P}(|Y| > x).$$

The absolute value of the above expression is then uniformly bounded by

$$\mathbb{P}(|X' - X| > \delta) + \sup_{x\in\mathbb{R}} |\mathbb{P}(|X| > x) - \mathbb{P}(|Y| > x)| + 2 \sup_{x\in\mathbb{R}} \mathbb{P}(|Y - x| \leq \delta). \quad \text{(S8.5)}$$

Similarly, the absolute value of right-hand side of (S8.4) is also uniformly bounded by (S8.5), which proves (S8.1). Finally, (S8.2) follows from (S8.1) and an application of Corollary 1 in Chernozhukov et al. (2015). Note that in this result the constant C is determined by $\max_{1\leq j\leq p} \mathbb{E}(Y_j^2) \leq c_2$ and $\min_{1\leq j\leq p} \mathbb{E}(Y_j^2) \geq c_1$. ♢

The following result is a consequence of of Lemma 5 of Zhou and Wu (2010).

Lemma S2. *Under the assumption 3.2(2), we have that*

$$\sup_{u_1,u_2,t_1,t_2\in[0,1]} |\mathbb{E}(G(u_1, t, \mathcal{F}_i)G(u_2, t_2, \mathcal{F}_j))| = O(\chi^{|i-j|}).$$

Lemma S3. *Define*

$$\sigma_{j,j}(u) = \frac{1}{nb_n} \sum_{i,l=1}^{n} \mathrm{Cov}(Z_{i,j}(u), Z_{l,j}(u))$$

where $Z_{i,j}$ are the components of the vector $Z_i(u)$ defined in (2.5)*. If $b_n = o(1)$, $\frac{\log n}{nb_n} = o(1)$ and Assumption 3.2 and Assumption 2.1 are satisfied, then there exist positive constants c_1 and c_2 such that for sufficiently large n*

$$0 < c_1 \leq \min_{1\leq j\leq p} \sigma_{j,j}(u) \leq \max_{1\leq j\leq p} \sigma_{j,j}(u) \leq c_2 < \infty.$$

for all $u \in [b_n, 1-b_n]$. Moreover, we have for

$$\tilde{\sigma}_{j,j} := \frac{1}{2\lceil nb_n\rceil - 1} \sum_{i,l=1}^{2\lceil nb_n\rceil-1} \mathrm{Cov}(\tilde{Z}_{i,j}, \tilde{Z}_{l,j}), \tag{S8.6}$$

the estimates

$$c_1 \leq \min_{1\leq j\leq (n-2\lceil nb_n\rceil+1)p} \tilde{\sigma}_{j,j} \leq \max_{1\leq j\leq (n-2\lceil nb_n\rceil+1)p} \tilde{\sigma}_{j,j} \leq c_2.$$

Proof of Lemma S3. By definition,

$$\sigma_{j,j}(u) = \frac{1}{nb_n}\sum_{i,l=1}^{n} \mathbb{E}\Big(G(\tfrac{i}{n}, t_j, \mathcal{F}_i)K\Big(\frac{\frac{i}{n}-u}{b_n}\Big)G(\tfrac{l}{n}, t_j, \mathcal{F}_l)K\Big(\frac{\frac{l}{n}-u}{b_n}\Big)\Big).$$

Observing Assumption 3.2 and Lemma S2, we have

$$\mathbb{E}(G(\tfrac{i}{n}, t_j, \mathcal{F}_i)G(\tfrac{l}{n}, t_j, \mathcal{F}_l) - G(u, t_j, \mathcal{F}_i)G(u, t_j, \mathcal{F}_l)) = O\big(\min(\chi^{|l-i|}, b_n)\big)$$

uniformly with respect to $u \in [b_n, 1-b_n]$, $|\frac{i}{n} - u| \leq b_n$ and $|\frac{l}{n} - u| \leq b_n$. Consequently, observing Assumption 2.1 it follows that

$$\sigma_{j,j}(u) = \frac{1}{nb_n}\sum_{i,l=1}^{n} \mathbb{E}\Big(G(u, t_j, \mathcal{F}_i)K\Big(\frac{\frac{i}{n}-u}{b_n}\Big)G(u, t_j, \mathcal{F}_l)K\Big(\frac{\frac{l}{n}-u}{b_n}\Big)\Big) + O(-b_n \log b_n) \tag{S8.7}$$

On the other hand, if r_n is a sequence such that $r_n = o(1)$ and $nb_nr_n \to \infty$, $A(u, r_n) := \{l : |\frac{\frac{l}{n}-u}{b_n}| \leq 1 - r_n, u \in [b_n, 1-b_n]\}$ we obtain by (S8.7) and Lemma S2 that

$$
\begin{aligned}
\sigma_{j,j}(u) = &\frac{1}{nb_n}\sum_{l=1}^{n}\sum_{i=1}^{n}\mathbf{1}(|i-l|\le nb_nr_n)\mathbb{E}\Big(G(u,t_j,\mathcal{F}_i)K\Big(\frac{\frac{i}{n}-u}{b_n}\Big)G(u,t_j,\mathcal{F}_l)K\Big(\frac{\frac{l}{n}-u}{b_n}\Big)\Big)\\
&+O(-b_n\log b_n+\chi^{nb_nr_n})\\
&=\frac{1}{nb_n}\sum_{l=1}^{n}K^2\Big(\frac{\frac{l}{n}-u}{b_n}\Big)\sum_{\substack{1\le i\le n,\\ |i-l|\le nb_nr_n}}\mathbb{E}\Big(G(u,t_j,\mathcal{F}_i)G(u,t_j,\mathcal{F}_l)\mathbf{1}\Big(\Big|\frac{\frac{i}{n}-u}{b_n}\Big|\le 1\Big)\Big)\\
&+O(-b_n\log b_n+\chi^{nb_nr_n}+r_n)\\
&=\frac{1}{nb_n}\sum_{\substack{1\le l\le n,\\ l\in A(u,r_n)}}K^2\Big(\frac{\frac{l}{n}-u}{b_n}\Big)\sum_{\substack{1\le i\le n,\\ |i-l|\le nb_nr_n}}\mathbb{E}\Big(G(u,t_j,\mathcal{F}_i)G(u,t_j,\mathcal{F}_l)\mathbf{1}\Big(\Big|\frac{\frac{i}{n}-u}{b_n}\Big|\le 1\Big)\Big)\\
&+O(-b_n\log b_n+\chi^{nb_nr_n}+r_n) \qquad\qquad (S8.8)
\end{aligned}
$$

uniformly for $j\in\{1,\ldots,p\}$. We obtain by the definition of the long-run variance $\sigma^2(u,t)$ in Assumption 3.2(4) and Lemma S2 that

$$
\Big|\sum_{i=1}^{n}\mathbb{E}\Big(G(u,t_j,\mathcal{F}_i)G(u,t_j,\mathcal{F}_l)\mathbf{1}\Big(\Big|\frac{\frac{i}{n}-u}{b_n}\Big|\le 1,|i-l|\le nb_nr_n\Big)\Big)-\sigma^2(u,t_j)\Big| = O(\chi^{nb_nr_n}) \qquad (S8.9)
$$

uniformly with respect to $l\in A(u,r_n)=\{l:|\frac{\frac{l}{n}-u}{b_n}|\le 1-r_n, u\in[b_n,1-b_n]\}$ and $j\in\{1,\ldots,p\}$. Combining (S8.8) and (S8.9) and using Lemma S2 yields

$$
\begin{aligned}
\sigma_{j,j}(u) &= \frac{1}{nb_n}\sum_{l=1}^{n}K^2\Big(\frac{\frac{l}{n}-u}{b_n}\Big)\sigma^2(u,t_j)+O(-b_n\log b_n+\chi^{nb_nr_n}+r_n)\\
&=\sigma^2(u,t_j)\int_{-1}^{1}K^2(t)dt+O\Big(-b_n\log b_n+\chi^{nb_nr_n}+r_n+\frac{1}{nb_n}\Big).
\end{aligned}
$$

Let $r_n=\frac{a\log n}{nb_n}$ for some sufficiently large positive constant a, then the assertion of the lemma follows in view of Assumption 3.2(4)).

For the second assertion, consider the case that $j = k_1 p + k_2$ for some $0 \le k_1 \le n - 2\lceil nb_n \rceil$ and $1 \le k_2 \le p$. Therefore by definition (2.7) in the main article,

$$\tilde{Z}_{i,k_1p+k_2} = G(\tfrac{i+k_1}{n}, \tfrac{k_2}{p}, \mathcal{F}_{i+k_1})K(\tfrac{i-\lceil nb_n \rceil}{nb_n}),$$

which gives for the quantity in (S8.6)

$$\tilde{\sigma}_{k_1p+k_2,k_1p+k_2} = \frac{1}{2\lceil nb_n \rceil - 1} \sum_{i,l=1}^{2\lceil nb_n \rceil - 1} \mathbb{E}\Big(G(\tfrac{i+k_1}{n}, \tfrac{k_2}{p}, \mathcal{F}_{i+k_1})K(\tfrac{i-\lceil nb_n \rceil}{nb_n})G(\tfrac{l+k_1}{n}, \tfrac{k_2}{p}, \mathcal{F}_{l+k_1})K(\tfrac{l-\lceil nb_n \rceil}{nb_n})\Big)$$

Consequently, putting $i + k_1 = s_1$ and $l + k_1 = s_2$ and using a change of variable, we obtain that

$$\tilde{\sigma}_{k_1p+k_2,k_1p+k_2} = \sigma_{k_2,k_2}\Big(\tfrac{k_1+\lceil nb_n \rceil}{n}\Big),$$

which finishes the proof. $\diamondsuit$

Bibliography

Cao, G., L. Yang, and D. Todem (2012). Simultaneous inference for the mean function based on dense functional data. *Journal of Nonparametric Statistics 24*(2), 359–377.

Chatterjee, S. (2014). *Superconcentration and related topics*, Volume 15. Springer.

Chen, M. and Q. Song (2015). Simultaneous inference of the mean of functional time series. *Electronic Journal of Statistics 9*(2), 1779–1798.

Chernozhukov, V., D. Chetverikov, and K. Kato (2015). Comparison and anti-concentration bounds for maxima of gaussian random vectors. *Probability Theory and Related Fields 162*(1), 47–70.

Craven, P. and G. Wahba (1978). Smoothing noisy data with spline functions. *Numerische mathematik 31*(4), 377–403.

Dette, H., K. Kokot, and A. Aue (2020). Functional data analysis in the Banach space of continuous functions. *Annals of Statistics 48*(2), 1168–1192.

Dette, H. and W. Wu (2019). Detecting relevant changes in the mean of nonstationary processes—a mass excess approach. *The Annals of Statistics 47*(6), 3578–3608.

Dette, H. and W. Wu (2022). Prediction in locally stationary time series. *Journal of Business & Economic Statistics 40*(1), 370–381.

Dette, H., W. Wu, and Z. Zhou (2019). Change point analysis of correlation in non-stationary time series. *Statistica Sinica 29*(2), 611–643.

Konakov, V. D. and V. I. Piterbarg (1984). On the convergence rate of maximal deviation distribution for kernel regression estimate. *Journal of Multivariate Analysis 15*, 279–294.

Politis, D. N., J. P. Romano, and M. Wolf (1999). *Subsampling*. Springer Science & Business Media.

Proksch, K. (2014). On confidence bands for multivariate nonparametric regression. *Annals of the Institute of Statistical Mathematics 68*, 209–236.

Wang, H. and S. Xiang (2012). On the convergence rates of legendre approximation. *Mathematics of Computation 81*(278), 861–877.

Wu, W. B. (2005). Nonlinear system theory: Another look at dependence. *Proceedings of the National Academy of Sciences 102*(40), 14150–14154.

Wu, W. B. and Z. Zhao (2007). Inference of trends in time series. *Journal of the Royal Statistical Society: Series B (Statistical Methodology) 69*(3), 391–410.

Xia, Y. (1998). Bias-corrected confidence bands in nonparametric regression. *Journal of the Royal Statistical Society, Ser. B 60*, 797–811.

Zhang, X. and G. Cheng (2018). Gaussian approximation for high dimensional vector under physical dependence. *Bernoulli 24*(4A), 2640–2675.

Zhou, Z. (2013). Heteroscedasticity and autocorrelation robust structural change detection. *Journal of the American Statistical Association 108*(502), 726–740.

Zhou, Z. and H. Dette (2020). Statistical inference for high dimensional panel functional time series. *arXiv preprint arXiv:2003.05968*.

Zhou, Z. and W. B. Wu (2010). Simultaneous inference of linear models with time varying coefficients. *Journal of the Royal Statistical Society: Series B (Statistical Methodology) 72*(4), 513–531.