Adaptive estimation for Weakly Dependent Functional Times Series
Supplementary Material

Hassan Maissoro[1,2], Valentin Patilea[1], and Myriam Vimond*[1]

[1]Univ Rennes, Ensai, CNRS, CREST (Center for Research in Economics and Statistics) - UMR 9194, F-35000 Rennes, France
[2]Datastorm, F-91120 Palaiseau - France

February 1, 2024

Abstract

In this supplement we provide the proofs of the lemmas and additional technical statements given in the Appendix of the main document. We also provide further empirical results and details on the construction of our simulation setups and the real data case.

In section S.1 the proofs of the technical lemmas stated in the Appendix section A are given. Additional results for the local regularity estimation in the case of differentiable sample paths are stated and proved in section S.2. The proof of the lemmas used in the Appendix section C are given in section S.3 below. Details of the simulation setups, additional simulation results and insight on the choice of the tuning parameters involved in the local regularity estimation are given in section S.4.

Contents

*Corresponding author. E-mail: myriam.vimond@ensai.fr

S.1 Proofs of lemmas for the local regularity and examples

In this section we provide the proofs of lemmas in the Appendix of the manuscript and a formal justification for the Examples 2 to 6. For the sake of readability, we reproduce each of the statements before providing the proof.

S.1.1 Local regularity properties

Lemma 1. *Assume that X belongs to $\mathcal{X}(\delta + H_\delta, \boldsymbol{L}_\delta, J)$ for some $\delta \in \mathbb{N}^*$, J an open sub-interval of I, $0 < H_\delta < 1$, and a bounded vector-valued function $\boldsymbol{L}_\delta \in \mathbb{R}_+^{\delta+1}$. Then, for any $d \in \{0, \dots, \delta - 1\}$, X belongs to $\mathcal{X}(d + H_d, \boldsymbol{L}_d, J)$ with $H_d \equiv 1$ and some bounded vector-valued function $\boldsymbol{L}_d \in \mathbb{R}_+^{d+1}$.*

Proof of Lemma 1. By the definition of $\mathcal{X}(\delta + H_\delta, \boldsymbol{L}_\delta, J)$, the Assumption (H6)-(a) is satisfied for any $d < \delta$. Let us next fix $t \in J$, and $d \in \{0, \dots, \delta - 1\}$. By definition, $\Delta_{\delta,0} > 0$ exists such that $[t - \Delta_{\delta,0}/2, t + \Delta_{\delta,0}/2] \subset J$ and (4) holds true for δ. By the Mean-Value Theorem, $\forall u, v \in [t - \Delta_{\delta,0}/2, t + \Delta_{\delta,0}/2]$ such that $u \le t \le v$, there exists $w \in (u, v)$, which may depend on d, such that

$$\mathbb{E}\left[\left|\nabla^d X(u) - \nabla^d X(v)\right|^2\right] = (u-v)^2 \mathbb{E}\left[\left(\nabla^{d+1} X(w)\right)^2\right] = (u-v)^2 \left\{L_{d,t}^2 + 2E_1(d) + E_2(d)\right\},$$

where $L_{d,t}^2 := \mathbb{E}\left[\left(\nabla^{d+1} X(t)\right)^2\right] \in [\,\underline{a}_{d+1}, \overline{a}_{d+1}]$ and

$$E_1(d) := \mathbb{E}\left[\nabla^{d+1} X(t) \left(\nabla^{d+1} X(w) - \nabla^{d+1} X(t)\right)\right], \quad E_2(d) := \mathbb{E}\left[\left(\nabla^{d+1} X(w) - \nabla^{d+1} X(t)\right)^2\right].$$

By Cauchy-Schwartz inequality we get

$$\begin{aligned}
\left|\mathbb{E}\left[\left|\nabla^d X(u) - \nabla^d X(v)\right|^2\right] - L_{d,t}^2 (u-v)^2\right| &= \left|2E_1(d) + E_2(d)\right| (u-v)^2 \\
&\le \left(2L_{d,t}\sqrt{E_2(d)} + E_2(d)\right)(u-v)^2 \\
&\le \left(2\sqrt{\overline{a}_{d+1}}\sqrt{E_2(d)} + E_2(d)\right)(u-v)^2. \qquad \text{(S.1)}
\end{aligned}$$

It thus remains to bound $E_2(d)$. Without loss of generality, the length of J is assumed smaller than 1.

The case of $d \le \delta - 2$. By the Mean-Value Theorem, condition (3), and since $|w-t| \le |u-v| \le \Delta_{\delta,0}$,

$$E_2(d) = \mathbb{E}\left[\left|\nabla^{d+1} X(w) - \nabla^{d+1} X(t)\right|^2\right] \le \overline{a}_{d+2}(w-t)^2 \le \overline{a}_{d+2}|u-v|^2.$$

Then (S.1) implies (4) with $\Delta_{d,0} = \Delta_{\delta,0}$, $H_{d,t} = 1$, $S_d^2 = 2\sqrt{\overline{a}_{d+1}}\sqrt{\overline{a}_{d+2}} + \overline{a}_{d+2}$ and $\beta_d = 1/2$.

The case $d = \delta - 1$. Since $|w-t| \le |u-v| \le \Delta_{\delta,0} < 1$, condition (4) considered with δ implies

$$E_2(\delta-1) = \mathbb{E}\left[\left|\nabla^\delta X(w) - \nabla^\delta X(t)\right|^2\right] \le L_{\delta,t}^2 |w-t|^{2H_{\delta,t}} + S_\delta^2 |w-t|^{2H_{\delta,t} + 2\beta_\delta} \le \{L_{\delta,t}^2 + S_\delta^2\}|u-v|^{2H_{\delta,t}}.$$

We then deduce from (S.1) that, $\forall |u-v| \le \Delta_{\delta-1,0} = \Delta_{\delta,0}$,

$$\left|\mathbb{E}\left[\left|\nabla^d X(u) - \nabla^d X(v)\right|^2\right] - L_{d,t}^2(u-v)^2\right| \le S_{\delta-1}^2 |u-v|^{2H_{d,t} + 2\beta_d} \quad \text{with} \quad H_{d,t} = 1,\ \beta_d = \underline{H},$$

$S_{\delta-1}^2 = 2\sqrt{\overline{a}_{d+1}}\sqrt{\overline{L}^2 + S_\delta^2} + \overline{L}^2 + S_\delta^2$, where $\sup_{t \in J} L_{\delta,t} \le \overline{L}$, $0 < \underline{H} \le \inf_{t \in J} H_{\delta,t}$. Then (4) follows. □

Example 2. [Local regularity of FAR(1)] *Let $\{X_n\}$ be the stationary FAR(1) time series defined by an integral operator with kernel ψ and with an MfBm functional white noise with Hurst exponent function H_ξ. Under the conditions stated in Example 2, $\{X_n\}$ belongs to $\mathcal{X}(H_\xi, 1; I)$.*

Proof of the statement in Example 2. Let t in the interior of I, and $u, v \in I$ such that $u \leq t \leq v$. Without loss of generality, assume the length of I is equal to 1. By Jensen's inequality,

$$|\{X_n(u) - X_n(v)\} - \{\xi_n(u) - \xi_n(v)\}|^2 \leq \int_I |\psi(s,u) - \psi(s,v)|^2 X_{n-1}^2(s)ds \leq C|u-v|^{2H_\psi} \int_I X_{n-1}^2(s)ds.$$

Using the stationarity of $\{X_n\}$, we have

$$\begin{aligned}&\Big|\nu_2\left(X_n(u) - X_n(v)\right) - \nu_2\left(\xi_n(u) - \xi_n(v)\right)\Big| \\ &\qquad\qquad \leq \nu_2\left(\{X_n(u) - X_n(v)\} - \{\xi_n(u) - \xi_n(v)\}\right) \leq C^{1/2}\nu_2(\|X\|_\infty)|u-v|^{H_\psi}.\end{aligned}$$

By the properties of the MfBm, assuming $\sup_{u\in I} H_{\xi,u} < 1$,

$$\nu_2^2\left(\xi_n(u) - \xi_n(v)\right) = |u-v|^{2H_{\xi,t}}\left\{1 + O(|u-v|^{2\beta_\xi})\right\},$$

for some $\beta_\xi > 0$ (Wei et al., 2023). Next, since $|x^2 - y^2| \leq |x-y|^2 + 2|y||x-y|$, we get

$$\left|\mathbb{E}\left[|X_n(u) - X_n(v)|^2\right] - |u-v|^{2H_{\xi,t}}\right| \leq C_0(C_0+2)|u-v|^{2H_{\xi,t}+2\beta_0},$$

with $C_0 = C^{1/2}\nu_2(\|X\|_\infty)$ and $\beta_0 = \min\{\beta_\xi, H_\psi - H_{\xi,t}\} > 0$. Hence, $\{X_n\}$ belongs to $\mathcal{X}(H_\xi, 1; I)$. □

S.1.2 Lemmas on $\mathbb{L}^p - m-$approximability

We first provide the proofs for the results stated in the Appendix of the manuscript. For the sake of readability, we recall the notation and the statements. The multiplication operator $\otimes$ is defined as

$$(f \otimes g)(s,t) = f(s)g(t) \qquad \forall s,t \in I \quad \text{and} \quad \ell \in \mathbb{Z}.$$

Meanwhile, the tensor product $\circ$ is defined as

$$\left(X_n \circ Y_n\right)(g) = \langle Y_n, g\rangle_{\mathcal{H}} X_n, \qquad \forall X_n, Y_n, g \in \mathcal{C}.$$

Finally, $\mathcal{L} = \mathcal{L}(\mathcal{C},\mathcal{C})$ is the space of bounded linear operators on $\mathcal{C}(I)$ equipped with the sup-norm.

Lemma 2. *Let $\{X_n\}$ and $\{Y_n\}$ be two $\mathbb{L}^p_{\mathcal{C}} - m-$approximable sequences in $\mathcal{C}$, for some $p \geq 4$. Define :*

1. $Z_n^{(1)} = A(X_n)$, *where* $A \in \mathcal{L}$;
2. $Z_n^{(2)} = X_n + Y_n$;
3. $Z_n^{(3)} = X_n Y_n$;
4. $Z_n^{(4)} = \langle X_n, Y_n\rangle_{\mathcal{H}} \in \mathbb{R}$;
5. $Z_n^{(5)} = X_n \circ Y_n \in \mathcal{L}$;
6. $Z_n^{(6)} = X_n \otimes X_{n+\ell}$, *where here* $\{X_n\}$ *is* $\mathbb{L}^p_{\mathcal{C}} - m-$*approximable for some* $p \geq 8$.

Then $\{Z_n^{(1)}\}$, $\{Z_n^{(2)}\}$ are $\mathbb{L}^p_{\mathcal{C}} - m-$approximable sequences in $\mathcal{C}$, and $\{Z_n^{(6)}\}$ is $\mathbb{L}^{p/2}_{\mathcal{C}} - m-$approximable sequences in $\mathcal{C}$ and its $\mathbb{L}^{p/2}_{\mathcal{C}} - m-$approximation is $X_n^{(m)} \otimes X_{n+\ell}^{(m+\ell)}$. If X_n and Y_n are independent, then $\{Z_n^{(3)}\}$, $\{Z_n^{(4)}\}$ and $\{Z_n^{(5)}\}$ are $\mathbb{L}^p - m-$approximable in their respective spaces.

Proof of Lemma 2. We use the simplified notation Z_n for all the points of the Lemma. Moreover, without loss of generality, we assume the length of I is equal to 1.

1) Let $Z_n = A(X_n)$, and let $Z_n^{(m)} = A(X_n^{(m)})$ be its coupled version. The definitions of $\mathcal{L}$ and $|||A|||_\infty$ entail that

$$\nu_p\left(\|Z_n - Z_n^{(m)}\|_\infty\right) \leq |||A|||_\infty \nu_p\left(\|X_n - X_n^{(m)}\|_\infty\right).$$

Since $\{X_n\}$ is $\mathbb{L}^p_{\mathcal{C}} - m-$approximable, the sequence $\{\nu_p(\|Z_n - Z_n^{(m)}\|_\infty), m \in \mathbb{Z}\}$ thus converges in the sense of condition 4 in Definition 3. As a consequence, $\{Z_n\}$ is $\mathbb{L}^p_{\mathcal{C}} - m-$approximable.

2) If $Z_n = X_n + Y_n$, we have

$$\nu_p\left(\|Z_n - Z_n^{(m)}\|_\infty\right) \leq \nu_p\left(\|X_n - X_n^{(m)}\|_\infty\right) + \nu_p\left(\|Y_n - Y_n^{(m)}\|_\infty\right),$$

and the statement is a direct consequence of the fact that $\{X_n\}$ and $\{Y_n\}$ are $\mathbb{L}^p_{\mathcal{C}} - m-$approximable.

3) When $Z_n(t) = X_n(t)Y_n(t),\ \forall t \in I$, we note that

$$Z_n - Z_n^{(m)} = X_n\left(Y_n - Y_n^{(m)}\right) + Y_n^{(m)}\left(Y_n - Y_n^{(m)}\right).$$

By the independence between X_n and Y_n, we have

$$\nu_p\left(\|Z_n - Z_n^{(m)}\|_\infty\right) \leq \nu_p\left(\|X_n\|_\infty\right)\nu_p\left(\|Y_n - Y_n^{(m)}\|_\infty\right) + \nu_p\left(\|Y_n^{(m)}\|_\infty\right)\nu_p\left(\|X_n - X_n^{(m)}\|_\infty\right).$$

Using the stationarity, $\nu_p(\|Y_n^{(m)}\|_\infty) = \nu_p\left(\|Y_n\|_\infty\right)$ and $\nu_p\left(\|X_n\|_\infty\right)$ are constants. Hence, $\{Z_n\}$ is $\mathbb{L}^p_{\mathcal{C}} - m-$approximable.

4) If $Z_n = \langle X_n, Y_n\rangle_{\mathcal{H}}$, we have

$$\left|Z_n - Z_n^{(m)}\right| = \left|\langle X_n, Y_n\rangle_{\mathcal{H}} - \langle X_n^{(m)}, Y_n^{(m)}\rangle_{\mathcal{H}}\right| \leq \left\|X_nY_n - X_n^{(m)}Y_n^{(m)}\right\|_\infty,$$

and thus

$$\nu_p\left(Z_n - Z_n^{(m)}\right) \leq \nu_p\left(\|X_nY_n - X_n^{(m)}Y_n^{(m)}\|_\infty\right).$$

By the property at point 3), $\{Z_n\}$ is $\mathbb{L}^p - m-$approximable.

5) Here Z_n is the Hilbert-Schmidt operator defined by the tensor product $(X_n \circ Y_n)(\cdot) = \langle X_n, \cdot\rangle Y_n$. Thus the notion of $\mathbb{L}^p - m-$approximability is considered with $\mathcal{C}$ replaced by the space $\mathcal{L}$ equipped with $|||\cdot|||_\infty$, which is a Banach space. Since $|||\cdot|||_\infty \leq |||\cdot|||_2$, and

$$\left[Z_n(e_j) - Z_n^{(m)}(e_j)\right](t) = \langle X_n, e_j\rangle_{\mathcal{H}}Y_n(t) - \langle X_n^{(m)}, e_j\rangle_{\mathcal{H}}Y_n^{(m)}(t) = \left\langle X_nY_n(t) - X_n^{(m)}Y_n^{(m)}(t), e_j\right\rangle,$$

using Parseval's identity, we get

$$\begin{aligned}\left|\left|\left|Z_n - Z_n^{(m)}\right|\right|\right|_\infty^2 &\leq \sum_{j=1}^\infty \left\|Z_n(e_j) - Z_n^{(m)}(e_j)\right\|_{\mathcal{H}}^2 \leq \int_I \left(\sum_{j=1}^\infty \left\langle X_nY_n(t) - X_n^{(m)}Y_n^{(m)}(t), e_j\right\rangle^2\right) dt \\ &= \int_I \|X_nY_n(t) - X_n^{(m)}Y_n^{(m)}(t)\|_{\mathcal{H}}^2 dt = \iint_{I\times I}\left(X_n(s)Y_n(t) - X_n^{(m)}(s)Y_n^{(m)}(t)\right)^2 dsdt.\end{aligned}$$

Next, since $(ab - cd)^2 \leq 2a^2(b-d)^2 + 2d^2(a-c)^2$, we get

$$\left(X_n(s)Y_n(t) - X_n^{(m)}(s)Y_n^{(m)}(t)\right)^2 \leq 2\left[X_n(s)\left(Y_n(t) - Y_n^{(m)}(t)\right)\right]^2 + 2\left[Y_n^{(m)}(t)\left(X_n(s) - X_n^{(m)}(s)\right)\right]^2.$$

By Cauchy-Schwarz inequality and the subadditivity of $x \mapsto \sqrt{x}$,

$$\begin{aligned}\left|\left|\left|Z_n - Z_n^{(m)}\right|\right|\right|_\infty &\leq \sqrt{2\left(\|X_n\|_\infty^2\|Y - Y_n^{(m)}\|_\infty^2 + \|Y_n^{(m)}\|_\infty^2\|X_n - Y_n^{(m)}\|_\infty^2\right)}, \\ &\leq \sqrt{2}\left(\|X_n\|_\infty\|Y - Y_n^{(m)}\|_\infty + \|Y_n^{(m)}\|_\infty\|X_n - Y_n^{(m)}\|_\infty\right).\end{aligned}$$

Since $\nu_p(\cdot)$ is a norm, and the processes X_n ans Y_n are independent, we get,

$$\nu_p\left(\left|\left|\left|Z_n - Z_n^{(m)}\right|\right|\right|_\infty\right) \leq \nu_p\left(\|X_n\|_\infty\right)\nu_p\left(\|Y_n - Y_n^{(m)}\|_\infty\right) + \nu_p\left(\|Y_n\|_\infty\right)\nu_p\left(\|X_n - X_n^{(m)}\|_\infty\right).$$

We then conclude that $\{Z_n\}$ is $\mathbb{L}^p_{\mathcal{L}} - m-$approximable.

6) Here, $Z_n = X_n \otimes X_{n+\ell}$ and we have

$$Z_n = f(\xi_n, \xi_{n-1}, \dots) \circ f(\xi_{n+\ell}, \xi_{(n+\ell)-1}, \dots) = g\left(\xi_{n+\ell}, \xi_{(n+\ell)-1}, \dots, \xi_n, \xi_{n-1}, \dots\right).$$

Thus, we can define

$$Z_n^{(m)} = g\left(\xi_{n+\ell}, \xi_{(n+\ell)-1}, \dots, \xi_n, \xi_{n-1}, \dots, \xi_{n-m}^{(m)}, \xi_{n-(m+1)}^{(m)}, \dots\right) = X_n^{(m)} \otimes X_{n+\ell}^{(m+\ell)},$$

which entails that

$$\begin{aligned} Z_n - Z_n^{(m)} &= X_n \otimes X_{n+\ell} - X_n^{(m)} \otimes X_{n+\ell}^{(m+\ell)} \\ &= \left(X_n - X_n^{(m)}\right) \otimes X_{n+\ell} + X_n^{(m)} \otimes \left(X_{n+\ell} - X_{n+\ell}^{(m+\ell)}\right), \end{aligned}$$

and

$$\|Z_n - Z_n^{(m)}\|_\infty \le \|X_n - X_n^{(m)}\|_\infty \|X_{n+\ell}\|_\infty + \|X_n^{(m)}\|_\infty \|X_{n+\ell} - X_{n+\ell}^{(m+\ell)}\|_\infty.$$

By Cauchy-Schwartz inequality, we then have

$$\begin{aligned} \nu_{p/2}\left(\|Z_n - Z_n^{(m)}\|_\infty\right) \le \nu_p\left(\|X_n - X_n^{(m)}\|_\infty\right) \nu_p\left(\|X_{n+\ell}\|_\infty\right) \\ + \nu_p\left(\|X_n^{(m)}\|_\infty\right) \nu_p\left(\|X_{n+\ell} - X_{n+\ell}^{(m+\ell)}\|_\infty\right), \end{aligned}$$

and thus $\{Z_n\}$ is $\mathbb{L}_{\mathcal{C}}^{p/2} - m-$approximable. This conclude the proof of the Lemma. □

Lemma 3. *Let $\{X_n\}_{n\in\mathbb{Z}}$ be a $\mathbb{L}_{\mathcal{C}}^p - m$-approximable sequence. Let $s, t \in I$, $t \neq s$, and let c be a constant. Define*

$$F_n = X_n(t) \in \mathbb{R} \qquad \text{and} \qquad G_n = (X_n(s) - X_n(t))^2 + c.$$

Then $\{F_n\}$ is $\mathbb{L}^p - m-$approximable in $\mathbb{L}^p$ and $\{G_n\}$ is $\mathbb{L}^{p/2} - m-$approximable in $\mathbb{L}^{p/2}$.

Proof of Lemma 3. With $F_n^{(m)} = X_n^{(m)}(t)$, we have

$$\nu_p\left(F_n - F_n^{(m)}\right) \le \nu_p\left(\|X_n - X_n^{(m)}\|_\infty\right).$$

The $\mathbb{L}_{\mathcal{C}}^p - m-$approximability of $\{X_n\}$ therefore involves the $\mathbb{L}^p - m-$approximability of $\{F_n\}$. For G_n, let $G_n^{(m)} = \{X_n^{(m)}(s) - X_n^{(m)}(t)\}^2$. By Lemma A.2, $\{X_n(s) - X_n(t), n \in \mathbb{Z}\}$ is $\mathbb{L}^p - m-$approximable. Using the Cauchy Schwarz inequality and the stationarity, it is straightforward to deduce that G_n is $\mathbb{L}^{p/2} - m-$approximable. □

S.1.3 Examples of $\mathbb{L}^p - m-$approximable FTS

Example 3 ($\mathbb{L}^p - m-$approximability of FAR(1))**.** *Let Ψ be a bounded linear operator such that $\|\Psi\|_\infty < 1$ and $\{\xi_n\} \subset \mathbb{L}_{\mathcal{C}}^2$ be i.i.d. with mean zero. A zero mean sequence $\{X_n\}$ of elements of $\mathcal{C}$ follows a FAR(1) model if*

$$X_n(t) = \Psi(X_{n-1})(t) + \xi_n(t), \qquad t \in I, \quad n \in \mathbb{Z},$$

see Bosq (2000, Theorem 3.1). Then $\{X_n\}$ is $\mathbb{L}_{\mathcal{C}}^p - m-$approximable.

Proof of the statement in Example 3. According to the Theorem 3.1. of Bosq (2000), the FAR model has a unique stationary solution $\{X_n\} \subset \mathbb{L}_{\mathcal{H}}^2$ and admits a moving average (linear) representation

$$X_n = \sum_{j=0}^{\infty} \Psi^j(\xi_{n-j}),$$

where Ψ^j is the j^{th} iterate of Ψ. For each n, let $\{\xi_j^{(n)}, j \in \mathbb{Z}\}$ be an independent copy of $\{\xi_j, j \in \mathbb{Z}\}$. Then, the approximation of X_n is given by

$$X_n^{(m)} = \sum_{j=0}^{m-1} \Psi^j(\xi_{n-j}) + \sum_{j=m}^{\infty} \Psi^j(\xi_{n-j}^{(n)}),$$

and using the linearity of Ψ, we get

$$X_n - X_n^{(m)} = \sum_{j=m}^{\infty} \Psi^j \left(\xi_{n-j} - \xi_{n-j}^{(n)}\right).$$

Applying the sup-norm and ν_p on both sides, using the triangle inequality, noting that $\{\xi_j\}$ and $\{\xi_j^{(n)}\}$ are i.i.d., and $|||\Psi|||_\infty < 1$, we get

$$\nu_p\left(\|X_n - X_n^{(m)}\|_\infty\right) \leq 2\nu_p(\xi) \sum_{j=m}^{\infty} |||\Psi|||_\infty^j \leq |||\Psi|||_\infty^m \frac{2\nu_p(\xi)}{1 - |||\Psi|||_\infty} \to 0, \qquad \text{as } m \to \infty,$$

where ξ has the same distribution as ξ_j, and

$$\sum_{m=1}^{\infty} \nu_p\left(\|X_n - X_n^{(m)}\|_\infty\right) < \infty.$$

It follows that $\{X_n\}$ is $\mathbb{L}_{\mathcal{C}}^p - m$-approximable. □

Example 4 ($\mathbb{L}^p - m$-approximability of general linear process)**.** *Suppose $\{X_n\} \in \mathbb{L}_{\mathcal{C}}^2$ is a linear process like in the Example 4, with the errors distributed as ξ and satisfying $\nu_p(\|\xi\|_\infty) < \infty$, $p \geq 2$. Moreover, the operators satisfy the condition*

$$\sum_{m=1}^{\infty} \sum_{j=m}^{\infty} |||\Psi_j|||_\infty < \infty.$$

Then $\{X_n\}$ is $\mathbb{L}_{\mathcal{C}}^p - m$-approximable.

Proof of statement in Example 4. Let $\{\xi_j^{(n)}\}$ be an independent copy of $\{\xi_j\}$ for each n. Then, the $\mathbb{L}_{\mathcal{C}}^p - m-$approximation of X_n is given by

$$X_n^{(m)} = \sum_{j=0}^{m-1} \Psi_j(\xi_{n-j}) + \sum_{j=m}^{\infty} \Psi_j(\xi_{n-j}^{(n)}).$$

Thus, following closely the same steps as the FAR(1), we have

$$\nu_p\left(\|X_n - X_n^{(m)}\|_\infty\right) \leq \nu_p\left(\left\|\sum_{j=m}^{\infty} \Psi_j\left(\xi_{n-j} - \xi_{n-j}^{(n)}\right)\right\|_\infty\right) \leq 2\nu_p(\|\xi\|_\infty) \sum_{j=m}^{\infty} |||\Psi_j|||_\infty,$$

and thus

$$\sum_{m=1}^{\infty} \nu_p\left(\|X_n - X_n^{(m)}\|_\infty\right) \leq 2\nu_p(\xi) \sum_{m=1}^{\infty} \sum_{j=m}^{\infty} |||\Psi_j|||_\infty < \infty.$$

This shows that $\{X_n\}$ is $\mathbb{L}_{\mathcal{C}}^p - m-$approximable. □

Example 5 ($\mathbb{L}^p - m$-approximability of product process)**.** *Suppose that $\{Y_n\} \subset \mathbb{L}_{\mathcal{C}}^p$ and $\{U_n\} \subset \mathbb{L}^p$ are two independent $\mathbb{L}^p - m-$approximable sequences in the respective spaces. Their representations are $Y_n = g_Y(\eta_1, \eta_2, \dots)$ and $U_n = g_U(\gamma_1, \gamma_2, \dots)$, respectively, where $\{\eta_k\}_k$ and $\{\gamma_k\}_k$ are two i.i.d. random sequences. Then, the sequence $X_n(\cdot) = U_n Y_n(\cdot)$ is $\mathbb{L}_{\mathcal{C}}^p - m-$approximable sequence.*

Proof of statement in Example 5. Let $X_n^{(m)} = U_n^{(m)} Y_n^{(m)}$ be the $\mathbb{L}^p_{\mathcal{C}} - m-$approximation of X_n. Then,

$$X_n - X_n^{(m)} = U_n Y_n - U_n^{(m)} Y_n^{(m)} = U_n \left(Y_n - Y_n^{(m)}\right) + Y_n^{(m)} \left(U_n - U_n^{(m)}\right),$$

taking the ν_p norm both sides, we have

$$\nu_p \left(\|X_n - X_n^{(m)}\|_\infty\right) \leq \nu_p \left(\|U_n \left(Y_n - Y_n^{(m)}\right)\|_\infty\right) + \nu_p \left(\|Y_n^{(m)} \left(U_n - U_n^{(m)}\right)\|_\infty\right).$$

Using stationarity and the independence between Y_n and U_n, we get

$$\nu_p \left(\|X_n - X_n^{(m)}\|_\infty\right) \leq \nu_p \left(U_n\right) \nu_p \left(\|Y_n - Y_n^{(m)}\|_\infty\right) + \nu_p \left(\|Y_n\|_\infty\right) \nu_p \left(U_n - U_n^{(m)}\right).$$

Thus,

$$\sum_{m=1}^{\infty} \nu_p \left(X_n - X_n^{(m)}\right) \leq \nu_p \left(U_n\right) \sum_{m=1}^{\infty} \nu_p \left(Y_n - Y_n^{(m)}\right) + \nu_p \left(Y_n\right) \sum_{m=1}^{\infty} \nu_p \left(U_n - U_n^{(m)}\right) < \infty,$$

since Y_n and U_n are $\mathbb{L}^p - m-$approximable the respective spaces. This concludes the proof. □

Example 6 ($\mathbb{L}^p - m-$approximability of ARCH model)**.** *Let $\delta \in \mathcal{C}$ be a positive function and $\{\xi_n\}$ a sequence of independent copies of $\xi \in \mathbb{L}^p_{\mathcal{C}}$. Let $\beta(s,t)$ be a continuous non-negative kernel function in $\mathbb{L}^2(I \times I)$. Then,*

$$Y_n(t) = \xi_n(t)\sigma_n(t) \quad \textit{where} \quad \sigma_n^2(t) = \delta(t) + \int_I \beta(s,t) Y_{n-1}^2(s) ds, \tag{S.2}$$

is the so-called functional ARCH(1) *series. If for some $p > 0$*

$$\mathbb{E}\left\{H(\xi^2)\right\}^{p/2} < 1 \qquad \textit{with} \qquad H(\xi^2) = \sup_{t \in I} \int_0^1 \beta(s,t)\xi^2(s) ds,$$

then (S.2) *has a unique, strictly stationary solution $\{Y_n\}$, which is is $\mathbb{L}^p_{\mathcal{C}} - m-$approximable.*

Proof of statement in Example 6. The existence and uniqueness of the solution of (S.2) was proved by Hörmann et al. (2013). Theorem 2.2 of Hörmann et al. (2013) shows that $\{\sigma_n\}$ admits the MA representation

$$\sigma_n^2 = g(\xi_{n-1}, \xi_{n-2}, \ldots),$$

with some positive, measurable $g \in \mathcal{C}$. For each n, let $\{\xi_j^{(n)}\}$ be an independent copy of $\{\xi_j\}$. Then, the coupled version of Y_n is given by

$$Y_n^{(m)} = \xi_n \sigma_n^{(m)},$$

where $\{\sigma_n^{(m)}\}^2 = g(\xi_n, \ldots, \xi_{n-m+1}, \xi_{n-m+1}^{(n)}, \xi_{n-m-1}^{(n)}, \ldots)$ is the coupled version of σ_n^2. Note that

$$\|Y_n - Y_n^{(m)}\|_\infty \leq \|\xi_n\|_\infty \left\|\sigma_n^2 - \{\sigma_n^{(m)}\}^2\right\|_\infty^{1/2},$$

because $|\sigma_n + \sigma_n^{(m)}| \geq |\sigma_n - \sigma_n^{(m)}|$. Since ξ_n is independent of $\sigma_n^2 - \{\sigma_n^{(m)}\}^2$, we get

$$\mathbb{E}\left\{\|Y_n - Y_n^{(m)}\|_\infty^p\right\} \leq \mathbb{E}\left\{\|\xi\|_\infty^p\right\} \mathbb{E}\left\{\left\|\sigma_n^2 - \{\sigma_n^{(m)}\}^2\right\|_\infty^{p/2}\right\}.$$

Theorem 2.3 of Hörmann et al. (2013) provides the following upper bound

$$\mathbb{E}\left\{\left\|\sigma_n^2 - \{\sigma_n^{(m)}\}^2\right\|_\infty^{p/2}\right\} \leq c r^m, \qquad \forall n, m,$$

where $0 < r = r(p/2) < 1$ and $c = c(p/2) < \infty$. Consequently, we have,

$$\nu_p\left(\|Y_n - Y_n^{(m)}\|_\infty\right) \leq c^{1/p} r^{m/p} \nu_p\left(\|\xi\|_\infty\right), \qquad \forall n, m.$$

Then the series of general term $\nu_p\left(\|Y_m - Y_m^{(m)}\|_\infty\right)$ is convergent, and this shows that $\{Y_n\}$ is $\mathbb{L}^p - m-$approximable. □

S.1.4 Technical lemma: proxies error

Lemma 4 (Proxies accuracy)**.** *Let $t \in J$.*

1. *For any $\varphi \in (0,1)$ and $0 < \Delta \leq \Delta_{0,0}$ such that $4\Delta^{2\beta_0} S_0^2 < L_t^2 \log(2)\varphi$, we have*

$$|\widetilde{H}_t - H_t| < \varphi/2.$$

2. *Let $H \in (0,1]$ such that $|H - H_t| < \varphi < 1$. For any $\psi \in (0,1)$ and $0 < \Delta \leq \Delta_{0,0}$ such that $S_0^2 \Delta^{2\beta_0 - 2\varphi} < \psi/3$, we have*

$$\left| \frac{\theta(t_1,t_3) - L_t \Delta^{2H_t}}{\Delta^{2H}} \right| < \psi/3.$$

Proof of Lemma 4. 1. By condition (4), we may rewrite $\theta(u,v)$ for $u, v \in [t - \Delta/2, t + \Delta/2]$ as

$$\theta(u,v) = L_t^2 |u-v|^{2H_t} \{1 + \rho(u,v)\},$$

where $|\rho(u,v)| \leq (S_0/L_t)^2 \Delta^{2\beta_0}$. Since $4\Delta^{2\beta_0} S_0^2 < L_t^2 \log(2)$, we have $|\rho(u,v)| < 1/2$. Using the fact that $x \mapsto \log(1+x)$ is Lipschitz continuous on $x \in (-1/2, \infty)$, we get

$$\begin{aligned} \left| \widetilde{H}_t - H_t \right| &= \frac{|\log(1 + \rho(t_1,t_3)) - \log(1+\rho(t_1,t_2))|}{2\log(2)} \\ &\leq \frac{1}{\log(2)} (|\rho(t_1,t_3)| + |\rho(t_1,t_2)|) \leq \frac{2}{\log(2)} (S_0/L_t)^2 \Delta^{2\beta_0} < \varphi/2. \end{aligned}$$

2. By condition (4), if $\Delta^{2\beta_0 - 2\varphi} S_0^2 < \psi/3$, we get

$$\left| \frac{\theta(t_1,t_3) - L_t^2 \Delta^{2H_t}}{\Delta^{2H}} \right| < S_0^2 \Delta^{2\beta_0 + 2(H_t - H)} < S_0^2 \Delta^{2\beta_0 - 2\varphi} < \psi/3.$$

□

S.1.5 Technical lemma: Nagaev inequality

The local regularity estimators in Section 3 are functions of

$$\widehat{\theta}(u,v) = \frac{1}{N} \sum_{n=1}^{N} \left(\widetilde{X}_n(v) - \widetilde{X}_n(u) \right)^2, \qquad u, v \in J.$$

To study the properties of $\widehat{\theta}(u,v)$, we use the Nagaev-type inequality for sums of dependent random variables, see Liu et al. (2013). When dealing with real valued random variable, the dependence measure used in $\mathbb{L}^p - m-$approximation is slightly more restrictive that the *functional dependence measure* defined in Wu (2005, Definition 1).

Wu (2005, Theorem 1) establishes that the measure of dependence of a stationary causal random variable X_n on $\{\xi_j, j \geq m\}$ can be bounded by the dependence measures of X_n on individual ξ_j's. Therefore, the author considers only the element-wise dependence in the sequel of his work. In the same way, Liu et al. (2013) adopt the functional dependence measure on individual ξ_j and state Nagaev inequality in this framework.

Nagaev inequality of Liu et al. (2013). Let $\{U_n, n \in \mathbb{Z}\}$ be a stationary, centered, real-valued causal process of the form

$$U_n = g(\xi_n, \xi_{n-1}, \cdots),$$

where $\{\xi_n\}_{n \in \mathbb{Z}}$ are i.i.d. real random variables and $g : \mathbb{R}^\infty \to \mathbb{R}$ is a measurable function. Let $\{\xi'_n, n \in \mathbb{Z}\}$ be an independent copy of $\{\xi_n, n \in \mathbb{Z}\}$. A coupled version of U_n is denoted by

$$U'_m = g(\xi_m, \xi_{m-1}, \cdots, \xi_1, \xi'_0, \xi_{-1}, \cdots),$$

and the corresponding distance is measured with

$$\lambda_{m,p} = \nu_p(U_m - U'_m), \qquad \upsilon = \sum_{m=1}^{\infty} \left(m^{p/2-1} \lambda_{m,p}^p \right)^{1/(p+1)}.$$

We assume that a short-range dependence condition is satisfied, *i.e.,* $\sum_{m \geq 0} \lambda_{m,p} < \infty$. Let $S_n = U_1 + \cdots + U_n$ be the partial sum of the process. Liu et al. (2013, Theorem 2) provides Nagaev-type inequality for $S_N^* = \max\{|S_n|,\ n = 1, \ldots, N\}$,

$$\forall \varepsilon > 0, \quad \mathbb{P}\left(S_N^* \geq \varepsilon\right) \leq c_p \frac{N}{\varepsilon^p} \left(\upsilon^{p+1} + \|U_1\|_p^p \right) + c'_p \exp\left(-\frac{c_p \varepsilon^2}{N \upsilon^{2+2/p}} \right) + 2 \exp\left(-\frac{c_p \varepsilon^2}{N \|U_1\|_2^2} \right), \tag{S.3}$$

where $c_p = 29p/\log(p)$ and c'_p are two positives constants. The expression of the constant c'_p depends on the Gaussian-like tail function defined as:

$$G_q(y) = \sum_{j=1}^{\infty} \exp\left(-j^q y^2\right), \qquad y > 0,\ q > 0.$$

For instance, if $\varepsilon = \sqrt{N} \upsilon^{1+1/p} y$ and $y \geq 1$, then we get $c'_p = 4G_{1-2/p}(1)e$. Now, if $y < 1$, we can take a fix and very small y_0 such that $y \geq y_0$ and obtain $c'_p = 4G_{1-2/p}\left(\sqrt{c_p} y_0\right) \exp\{c_p y_0^2\}$. Finally, we can consider,

$$c'_p = \max\left\{4G_{1-2/p}(1)e; 4G_{1-2/p}\left(\sqrt{c_p} y_0\right) \exp\{c_p y_0^2\}\right\}.$$

Nagaev inequality under our weak dependency assumption. The inequality (S.3) involves only element-wise dependence coefficients whereas the $\mathbb{L}^p - m$-approximation measures the dependence of U_n on the whole sequence $\{\xi_j, j \geq m\}$. Let us now consider $\{U_n\}$ a $\mathbb{L}^p - m-$approximable stationary process, and let the associated coupled version of U_m be defined as

$$U_m^{(m)} = g(\xi_m, \xi_{m-1}, \cdots, \xi_1, \xi_0^{(m)}, \xi_{-1}^{(m)}, \xi_{-2}^{(m)}, \cdots),$$

where, for each $m \geq 0$, $\{\xi_i^{(m)}, i \in \mathbb{Z}\}$ is an independent copy of $\{\xi_i, i \in \mathbb{Z}\}$. Let

$$\nu_{m,p} = \nu_p\left(U_m - U_m^{(m)}\right).$$

Lemma 5 states a version of Liu et al. (2013, Theorem 2) under $\mathbb{L}^p - m-$approximability assumption.

Lemma 5 (Nagaev inequality)**.** *Let $\{U_n\}$ be a real centered valued $\mathbb{L}^p - m-$approximable stationary process such that*

$$\upsilon := \sum_{m=1}^{\infty} \left(m^{p/2-1} \nu_{m,p}^p \right)^{1/(p+1)} < \infty.$$

The Nagaev-type inequality remains true, that is

$$\mathbb{P}\left(S_N^* \geq \varepsilon\right) \leq c_p \frac{N}{\varepsilon^p} \left(\upsilon^{p+1} + \|U_1\|_p^p \right) + c'_p \exp\left(-\frac{c_p \varepsilon^2}{N \upsilon^{2+2/p}} \right) + 2 \exp\left(-\frac{c_p \varepsilon^2}{N \|U_1\|_2^2} \right),$$

where $c_p = 29p/\log(p)$ and c'_p are two positives constants.

Proof of Lemma 5. The proof of this lemma follows closely the lines of the proofs of Liu et al. (2013, Theorem 1 and Theorem 2), therefore some similar parts will be omitted. The key step is their Equation (2.12) in the proof of Theorem 1 where $\|U_{1,j} - U_{1,j-1}\|_p$ is bounded by $\lambda_{j,p}$, with $U_{k,j} = \mathbb{E}\left[U_k | \xi_k, \cdots, \xi_{k-j}\right]$. So it remains to show that $\|U_{1,j} - U_{1,j-1}\|_p$ is also bounded by $\nu_{j,p}$. Note that

$$\begin{aligned} U_{1,j} - U_{1,j-1} &= \mathbb{E}\left[U_1 | \xi_1, \cdots, \xi_{1-j}\right] - \mathbb{E}\left[U_1 | \xi_1, \cdots, \xi_{1-(j-1)}\right] \\ &= \mathbb{E}\left[g(\xi_1, \xi_0, \cdots, \xi_{1-(j-1)}, \xi_{1-j}, \xi_{1-(j+1)}, \cdots) | \xi_1, \cdots, \xi_{1-j}\right] \\ &\quad - \mathbb{E}\left[g(\xi_1, \xi_0, \cdots, \xi_{1-(j-1)}, \xi_{1-j}, \xi_{1-(j+1)}, \cdots) | \xi_1, \cdots, \xi_{1-(j-1)}\right]. \end{aligned}$$

Since $\{\xi_i^{(j)}\}_{i\in\mathbb{Z}}$ is an independent copy of $\{\xi_i\}_{i\in\mathbb{Z}}$, the second conditional expectation of the last display stay unchanged if we replace $\{\xi_{1-j}, \xi_{1-(j+1)}, \cdots\}$ by $\left\{\xi_{1-j}^{(j)}, \xi_{1-(j+1)}^{(j)}, \cdots\right\}$, that is

$$U_{1,j} - U_{1,j-1} = \mathbb{E}\left[U_1 | \xi_1, \cdots, \xi_{1-j}\right] - \mathbb{E}\left[U_1^{(j)} | \xi_1, \cdots, \xi_{1-(j-1)}\right]$$

Since $U_1^{(j)}$ is independent of ξ_{1-j} (it depends no longer on ξ_{1-j} but on $\xi_{1-j}^{(j)}$), the variable ξ_{1-j} can be added in the conditioning part without changing the expression,

$$U_{1,j} - U_{1,j-1} = \mathbb{E}\left[U_1 - U_1^{(j)} | \xi_1, \cdots, \xi_{1-j}\right].$$

Now, using Jensen's inequality, we get

$$\mathbb{E}\left[|U_{1,j} - U_{1,j-1}|^p\right] \leq \mathbb{E}\left[\left|U_1 - U_1^{(j)}\right|^p\right].$$

Using the stationarity of $\{U_n\}$ we obtain

$$\|U_{1,j} - U_{1,j-1}\|_p \leq \nu_{j,p},$$

which conclude the proof. □

S.1.6 Technical lemma: concentration of $\widehat{\theta}(u,v)$

We now study the concentration of $\widehat{\theta}(u,v)$ and $\widehat{\theta}(u,v)/\theta(u,v)$.

Lemma 6. *Assume the conditions of Theorem 1 hold true. Let $u, v \in J$, $u \leq t \leq v$, be fixed points such that $\Delta/2 \leq |u-v| \leq \Delta$ and let*

$$\eta_0 = \eta_0(\lambda) = 8\left(2\sqrt{\overline{a}_0} + \sqrt{R_2(\lambda)}\right)\sqrt{R_2(\lambda)}.$$

For any $\kappa > 0$, define the probabilities

$$p_0^+(u,v;\kappa) = \mathbb{P}\left[\widehat{\theta}(u,v) > (1+\kappa)\theta(u,v)\right], \qquad p_0^-(u,v;\kappa) = \mathbb{P}\left[\widehat{\theta}(u,v) < (1-\kappa)\theta(u,v)\right].$$

Then, for any η such that $\eta_0 < \eta < 1$, we have

$$\mathbb{P}\left(\left|\widehat{\theta}(u,v) - \theta(u,v)\right| > \eta\right) \leq \frac{\mathfrak{a}}{N\eta^2} + \mathfrak{b}\exp\left(-\mathfrak{c}N\eta^2\right),$$

where $\mathfrak{b}$ is a universal constant, and $\mathfrak{a}$ and $\mathfrak{c}$ are two positive constants depending on the dependence measure and the bound of the fourth-order moment of $\widetilde{X}(u)$. Moreover, for any κ such that $\eta_0 < \kappa\theta(u,v) < 1$, we have:

$$\max\left[p_0^+(u,v;\kappa), p_0^-(u,v;\kappa)\right] \leq \frac{2^{2H_t+2}\mathfrak{a}}{N\kappa^2 L_t^4 \Delta^{4H_t}} + \mathfrak{b}\exp\left(-\frac{\mathfrak{c}}{2^{H_t+2}}N\kappa^2 L_t^4 \Delta^{4H_t}\right).$$

Proof of Lemma 6. We write $\widehat{\theta}(u,v) - \theta(u,v)$ as the sum of a bias term and a centered stochastic term :

$$\widehat{\theta}(u,v) - \theta(u,v) = \frac{1}{N}\sum_{n=1}^{N} Z_n(u,v) + \left\{\mathbb{E}\left(\widehat{\theta}(u,v)\right) - \theta(u,v)\right\}, \qquad u, v \in J,$$

where $Z_n = Z_n(u,v) = \left(\widetilde{X}_n(u) - \widetilde{X}_n(v)\right)^2 - \mathbb{E}\left(\widetilde{X}_n(u) - \widetilde{X}_n(v)\right)^2$.

Bounds for the bias term. Since $\{X_n\}$ is stationary and the sequence

$$\{\zeta_n = (M_n, T_{n,1}, \ldots, T_{n,M_n}, \varepsilon_{n,1}, \ldots, \varepsilon_{n,M_n}), n \geq 1\},$$

are i.i.d. (see (H2), (H3), (H4) and (H5)), the process $\{\widetilde{X}_n\}$ is stationary and we thus have

$$\mathbb{E}\left[\widehat{\theta}(u,v)\right] - \theta(u,v) = 2\mathbb{E}\left[\{G_n(u) - G_n(v)\}\{X_n(u) - X_n(v)\}\right] + \mathbb{E}\left[\{G_n(u) - G_n(v)\}^2\right],$$

where $G_n(u) = \widetilde{X}_n(u) - X_n(u)$. By Assumption (H10) and the inequality $(x+y)^2 \leq 2(x^2+y^2)$, we get $\mathbb{E}\left[\{G_n(u) - G_n(v)\}^2\right] \leq 4R_2(\lambda) \leq 4B\lambda^{-\tau}$. Cauchy-Schwarz inequality then implies

$$\left|\mathbb{E}\left[\widehat{\theta}(u,v)\right] - \theta(u,v)\right| \leq \eta_0/2. \tag{S.4}$$

Concentration bounds for the stochastic term. Recall that, for any $N \geq 1$, the finite sequence $\{\zeta_n, 1 \leq n \leq N\}$ is i.i.d. and this implies that the finite sequence $\{\widetilde{X}_n, 1 \leq n \leq N\}$ is also stationary. We now complete these finite sequences to infinite ones, $\{\zeta_n, n \in \mathbb{Z}\}$ and $\{\widetilde{X}_n, n \in \mathbb{Z}\}$, by generating independent M_n from the same distribution as $M_1, \ldots, M_N$, and independent copies $(T_{n,1}, \varepsilon_{n,1}), \ldots, (T_{n,M_n}, \varepsilon_{n,M_n})$ of (T, ε), for any $n \notin \{1, \ldots, N\}$. By the definition (9) and using the MA representation of $\{X_n\}$, see (6) in Definition 3, we can rewrite $\widetilde{X}_n$ as,

$$\begin{aligned}
\widetilde{X}_n(u) &= \sum_{i=1}^{M_n} W_{n,i}(u) X_n(T_{n,i}) + \sum_{k=1}^{M_n} W_{n,i}(u)\sigma(T_{n,i})\varepsilon_{n,i} \\
&= \sum_{i=1}^{M_n} W_{n,i}(u) f(\xi_n, \xi_{n-1}, \ldots)(T_{n,i}) + \sum_{i=1}^{M_n} W_{n,i}(u)\sigma(T_{n,i})\varepsilon_{n,i} = g\left((\zeta_n, \xi_n), (\zeta_{n-1}, \xi_{n-1}), \ldots\right),
\end{aligned}$$

where $\{(\zeta_n, \xi_n)\}$ is an i.i.d. sequence in the measurable space $\widetilde{\mathcal{S}} = \mathbb{N}^* \times \{\cup_{m\geq 1}[0,1]^m \times \mathbb{R}^m\} \times \mathcal{S}$ and $g : \widetilde{\mathcal{S}}^\infty \to \mathcal{H}$ is some measurable function. Then a coupled version of $\widetilde{X}_m(u)$ is

$$\widetilde{X}_n^{(m)}(u) = \sum_{i=1}^{M_n} W_{n,i}(u) X_n^{(m)}(T_{n,i}) + \sum_{i=1}^{M_n} W_{n,i}(u)\sigma(T_{n,i})\varepsilon_{n,i}, \qquad m \geq 1.$$

From this and (H9), a constant C exists such that $|\widetilde{X}_n^{(m)}(u) - \widetilde{X}_n(u)| \leq C\|X_n - X_n^{(m)}\|_\infty$. According to Definition 3, the sequence is $\mathbb{L}^4 - m-$approximable. Lemma 3 then entails that the sequence $\{Z_n\} = \{Z_n(u,v)\}$ is $\mathbb{L}^2 - m-$approximable. Let $\nu_{m,2} = \nu_2(Z_m - Z_m^{(m)})$ be its dependence coefficient, where $Z_m^{(m)}$ is the associated coupled version of Z_m. Using Cauchy-Schwartz inequality, we get

$$\nu_{m,2} \leq 8\nu_4\left(\|\widetilde{X}_m\|_\infty\right)\nu_4\left(\left\|X_m - X_m^{(m)}\right\|_\infty\right).$$

Since $\{X_n\}$ satisfies (H7), $\sigma(\cdot)$ is bounded and (H8) guarantees $\nu_4(\varepsilon) < \infty$, we necessarily have $\nu_4(\|\widetilde{X}_n\|_\infty) \leq C$, for some constant C independent of n. Finally, since $\nu_4(\|X_m - X_m^{(m)}\|_\infty) = \mathcal{O}(1/m^\alpha)$ with $\alpha > 3/2$, the dependence coefficient of $\{Z_n\}$ satisfies the condition $\nu := \sum_{m=1}^\infty \nu_{m,2}^{2/3} < \infty$, which will allow us to apply Lemma 5. More precisely, in view of (S.4), we deduce that $\forall \eta \in (\eta_0, 1)$,

$$\mathbb{P}\left(\widehat{\theta}(u,v) - \theta(u,v) > \eta\right) \leq \mathbb{P}\left(\frac{1}{N}\sum_{n=1}^N Z_n > \eta/2\right).$$

Applying then Nagaev-type inequality from Lemma 5, we get

$$\begin{aligned}
\mathbb{P}\left(\widehat{\theta}(u,v) - \theta(u,v) > \eta\right) &\leq \frac{4c_2\left(\nu^3 + \nu_2^2(Z_1)\right)}{N\eta^2} + c_2' \exp\left(-\frac{c_2}{4\nu^3}N\eta^2\right) + 2\exp\left(-\frac{c_2}{4\nu_2^2(Z_1)}N\eta^2\right) \\
&\leq \frac{\mathfrak{a}}{N\eta^2} + \mathfrak{b}\exp\left(-\mathfrak{c}N\eta^2\right),
\end{aligned}$$

where $\mathfrak{a} = 4c_2\left(\nu^3 + \nu_2^2(Z_1)\right)$, $\mathfrak{b} = c_2' + 2$ and $\mathfrak{c} = \min\left(c_2/(4\nu^3), c_2/(4\nu_2^2(Z_1))\right)$. Moreover, by (4) and (12) we have

$$\theta(u,v) \geq |u-v|^{2H_t} L_t^2/2 > 0.$$

This implies $\theta(u,v) > \eta_0$, provided λ is sufficiently large, and $\kappa > 0$ exists such that $\eta_0 < \kappa\theta(u,v) < 1$. We can then consider $\eta = \kappa\theta(u,v)$ in Lemma 5 and deduce

$$
\begin{aligned}
p_0^+(u,v;\kappa) = \mathbb{P}\left(\widehat{\theta}(u,v) > (1+\kappa)\theta(u,v)\right) &\leq \frac{\mathfrak{a}}{N\kappa^2\theta^2(u,v)} + \mathfrak{b}\exp\left(-\mathfrak{c}N\kappa^2\theta^2(u,v)\right), \\
&\leq \frac{2^{4H_t+2}\mathfrak{a}}{N\kappa^2L_t^4\Delta^{4H_t}} + \mathfrak{b}\exp\left(-\frac{\mathfrak{c}}{2^{4H_t+2}}N\kappa^2L_t^4\Delta^{4H_t}\right).
\end{aligned}
$$

Similar arguments apply for bounding $p_0^-(u,v;\kappa)$. The proof of Lemma 6 is thus complete. □

S.2 Local regularity estimation for smooth trajectories

Let us recall that, following the lines of Golovkine et al. (2022), in Section 3.2 we defined

$$
\widehat{\delta} = \min\left\{d \in \mathbb{N} : \widehat{H}_{d,t} < 1 - \varphi(\widehat{\lambda})\right\},
$$

where $\widehat{H}_{d,t}$ is an estimator of the local regularity exponent parameter of $\{\nabla^d X_n\}$ at t, estimator to be defined below. A natural estimator of the local regularity parameter α_t is then

$$
\widehat{\alpha}_t = \widehat{\delta} + \widehat{H}_{\widehat{\delta},t}.
$$

The sequential procedure based on $\widehat{\delta}$ was summarized in Algorithm 1. It thus remains to study the estimators for the H_d, $d = 1, 2\ldots$, introduced in Section 3.2. Like in the non-differentiable case, we first define proxies for these quantities that we next estimate nonparametrically.

Proxy values of $H_{d,t}$ and $L^2_{d,t}$. Let $d \geq 1$, $\Delta \leq \Delta_{d,0}$ and $t_1, t_2, t_3 \in J$ such that $t_3 - t_1 = \Delta$ and $t_2 = t = (t_1 + t_3)/2$. In view of (H6), we consider the following proxy values of $H_{d,t}$ and $L^2_{d,t}$:

$$
\begin{aligned}
\widetilde{H}_{d,t} &= \widetilde{H}_d(\Delta) = \frac{\log(\theta_d(t_1,t_3)) - \log(\theta_d(t_1,t_2))}{2\log(2)}, \\
\widetilde{L}^2_{d,t} &= \widetilde{L}^2_{d,t}(\Delta) = \frac{\theta_d\left(t_1,t_3\right)}{\Delta^{2H_{d,t}}}, \qquad \text{where} \qquad \theta_d(u,v) = \mathbb{E}\left[\left\{\nabla^d X(u) - \nabla^d X(v)\right\}^2\right].
\end{aligned}
$$

Like in the non-differentiable case, an estimator of $\theta_d(u,v)$, $u,v \in J$, is easily obtained from the estimates of the $d-$th derivative of the samples paths.

Presmoothing the derivatives. Let $d \geq 1$. Given the data points $(Y_{n,i}, T_{n,i})$, $1 \leq i \leq M_n$, we consider a linear smoother under the form

$$
\widetilde{\nabla^d X}_n(u) = \sum_{i=1}^{M_n} W_{n,i}^{(d)}(u) Y_{n,i}, \qquad u \in J,\ n = 1,\ldots,N, \tag{S.5}
$$

where the weights $\{W_{n,i}^{(d)}\}_{i=1\ldots M_n}$ are built from the data points. The local smoother we have in mind is the local polynomials. We consider the following assumptions for the presmoothing of the derivatives.

(D1) A constant $\mathfrak{c}_W > 0$ exists such that

$$
\sup_{n=1\ldots N}\sup_{u\in J}\sum_{i=1}^{M_n}\left|W_{n,i}^{(d)}(u)\right| \leq \mathfrak{c}_W, \qquad \forall d \in \{0,\ldots,\delta\}.
$$

(D2) Constants $B > 0$ and $\tau > 0$ exit such that

$$
R_{2,d}(\lambda) = \sup_{u\in J}\mathbb{E}\left(|\widetilde{\nabla^d X}(u) - \nabla^d X(u)|^2\right) \leq B\lambda^{-\tau}, \qquad \forall d \in \{0,\ldots,\delta\}.
$$

For instance, up to a slight modification, local polynomial smoothers satisfy the conditions (1) and (2).

Local regularity estimators of the d-th derivatives. For $d \geq 1$, given a presmoothing estimator $\widetilde{\nabla^d X}_n(u)$ of $\nabla^d X_n(u)$, for $u \in J$, we define the estimators of $H_{d,t}$ and $L^2_{d,t}$ as

$$\widehat{H}_{d,t} = \frac{\log \widehat{\theta}_d(t_1,t_3) - \log \widehat{\theta}_d(t_1,t_2)}{2\log(2)},$$

$$\widehat{L}^2_{d,t} = \frac{\widehat{\theta}_d\,(t_1,t_3)}{\Delta^{2\widehat{H}_{d,t}}} \qquad \text{where} \qquad \widehat{\theta}_d(u,v) = \frac{1}{N}\sum_{n=1}^{N}\left(\widetilde{\nabla^d X}_n(u) - \widetilde{\nabla^d X}_n(v)\right)^2.$$

In view of the proof of Lemma 1, let us define

$$\beta_d = 1/2 \quad \text{if } \ d \leq \delta - 2, \qquad \text{and} \qquad \beta_{\delta-1} = \underline{H},$$

$$L^2_{d,t} := \mathbb{E}\left[\left(\nabla^{d+1}X(t)\right)^2\right] \in [\,\underline{a}_{d+1}, \overline{a}_{d+1}], \quad 1 \leq d \leq \delta - 1,$$

and

$$S_d^2 = 2\sqrt{\overline{a}_{d+1}}\sqrt{\overline{a}_{d+2}} + \overline{a}_{d+2} \quad \text{if } \ d \leq \delta - 2, \qquad \text{and} \qquad S^2_{\delta-1} = 2\sqrt{\overline{a}_{d+1}}\sqrt{\overline{L}^2 + S^2_\delta} + \overline{L}^2 + S^2_\delta.$$

We now state the counterparts of Theorems 1 and 2 for the case of differentiable sample paths. The proofs are provided in the next section.

Proposition 1. *Assume that (H1) – (H8), (D1) – (D2) hold true. Let $d \in \{0,\dots,\delta\}$ and $\widetilde{H}_{d,t}, \widetilde{L}^2_{d,t}$ are defined with $\Delta \leq \Delta_{\delta,0}$. Constants C_d exist such that, for any $\varphi \in (0,1)$ satisfying the conditions*

$$\Delta^{2\beta_d} S_d^2 < \frac{L^2_{d,t}\log(2)}{4}\varphi, \tag{S.6}$$

$$\lambda^{-\tau/2} < C_d L^2_{d,t}\varphi\Delta^{2H_{d,t}}, \tag{S.7}$$

we have

$$\mathbb{P}(|\widehat{\alpha}_t - \alpha_t| > \varphi) \leq (2+\delta)\left[\frac{\mathfrak{f}}{N\varphi^2\Delta^4} + \mathfrak{b}\exp\left(-\mathfrak{g}N\varphi^2\Delta^4\right)\right],$$

for some universal constant $\mathfrak{b}$, provided λ is sufficiently large. The constants C_d depend on the $\overline{a}_d$'s from (3) and B from (D2), while the positive constants $\mathfrak{f}$ and $\mathfrak{g}$ depend one the dependence measure.

Proposition 2. *Assume the conditions of Proposition 1 hold true. Moreover, constants $\widetilde{C}_d > 0$, $d \in \{0,\dots,\delta\}$, exist such that for any $\varphi, \psi \in (0,1)$ satisfying*

$$3\Delta^{-2\varphi}\Delta^{2\beta_d}S_d^2 < \psi, \tag{S.8}$$

$$6L^2_{d,t}\Delta^{-2\varphi}\varphi|\log\Delta| < \psi, \tag{S.9}$$

$$\lambda^{-\tau/2} < \widetilde{C}_d\Delta^{2\varphi}\psi\Delta^{2H_{d,t}}, \tag{S.10}$$

we have

$$\mathbb{P}\left(\left|\widehat{L}^2_{d,t} - L^2_{d,t}\right| > \psi\right) \leq \frac{\mathfrak{c}_d}{N\psi^2\Delta^{4H_d+4\varphi}} + \frac{\mathfrak{f}_d}{N\varphi^2\Delta^{4H_d}} + 4\mathfrak{b}\exp\left(-\mathfrak{g}_d N\varphi^2\Delta^{4H_d}\right) + \mathfrak{b}\exp\left(-\mathfrak{l}_d N\psi^2\Delta^{4H_d+4\varphi}\right),$$

for some universal constant $\mathfrak{b}$, provided λ is sufficiently large. The constants $\widetilde{C}_d$ depend on the $\overline{a}_d$'s and B, while the constants $\mathfrak{c}_d$, $\mathfrak{f}_d$, $\mathfrak{g}_d$, $\mathfrak{l}_d$ are determined by the dependence structure of X.

S.2.1 Proofs of the concentration bounds for regularity estimators

The proofs below are using conditions (D1) and (D2). In order to hold with local polynomials, condition (D1) requires to modify the smoother, for instance to set it equal to zero, when the smallest eigenvalue of the design matrix used to define it is too close to zero. See (see Tsybakov, 2009, equation (1.66) and Assumption (LP), page 37). Under our assumptions, the probability of the event of the smallest eigenvalue close to zero is exponentially small. See Golovkine et al. (2022). For simplicity, we omit exponentially small probability events and assume (D1) holds true.

Let us recall, for $d \geq 1$, $\Delta \leq \Delta_{d,0}$ and $t_1, t_2, t_3 \in J$ such that $t_3 - t_1 = \Delta$ and $t_2 = t = (t_1 + t_3)/2$, the proxy values of $H_{d,t}$ and $L_{d,t}^2$ are

$$\widetilde{H}_{d,t} = \widetilde{H}_d(\Delta) = \frac{\log(\theta_d(t_1,t_3)) - \log(\theta_d(t_1,t_2))}{2\log(2)},$$

$$\widetilde{L}_{d,t}^2 = \widetilde{L}_{d,t}^2(\Delta) = \frac{\theta_d\left(t_1,t_3\right)}{\Delta^{2H_{d,t}}}, \qquad \text{where} \qquad \theta_d(u,v) = \mathbb{E}\left[\left(\nabla^d X(u) - \nabla^d X(v)\right)^2\right].$$

Moreover, given a presmoothing estimator $\widetilde{\nabla^d X}_n(u)$ of $\nabla^d X_n(u)$, for $u \in J$, the estimators of $H_{d,t}$ and $L_{d,t}^2$ are defined as

$$\widehat{H}_{d,t} = \widehat{H}_{d,t}(\Delta) = \frac{\log\widehat{\theta}_d(t_1,t_3) - \log\widehat{\theta}_d(t_1,t_2)}{2\log(2)},$$

$$\widehat{L}_{d,t}^2 = \widehat{L}_{d,t}^2(\Delta) = \frac{\widehat{\theta}_d\left(t_1,t_3\right)}{\Delta^{2\widehat{H}_{d,t}}}, \qquad \text{where} \qquad \widehat{\theta}_d(u,v) = \frac{1}{N}\sum_{n=1}^{N}\left(\widetilde{\nabla^d X}_n(u) - \widetilde{\nabla^d X}_n(v)\right)^2.$$

Lemma S.1. *Assume that the assumptions of Corrolary 1 are satisfied.*

Let $u, v \in J$, $u \leq t \leq v$, be fixed points such that $\Delta/2 \leq |u-v| \leq \Delta \leq \Delta_{\delta,0}$ and, for any $d = 1, \ldots, \delta$, let

$$\eta_d = \eta_d(\lambda) = 8\left(2\sqrt{\overline{a}_d} + \sqrt{R_{2,d}(\lambda)}\right)\sqrt{R_{2,d}(\lambda)},$$

and, for any $\kappa > 0$,

$$p_d^+(u,v;\kappa) = \mathbb{P}\left[\widehat{\theta}_d(u,v) > (1+\kappa)\theta_d(u,v)\right], \qquad p_d^-(u,v;\kappa) = \mathbb{P}\left[\widehat{\theta}_d(u,v) < (1-\kappa)\theta_d(u,v)\right].$$

For any η such that $\eta_d < \eta < 1$,

$$\mathbb{P}\left(\left|\widehat{\theta}_d(u,v) - \theta_d(u,v)\right| > \eta\right) \leq \frac{\mathfrak{a}_d}{N\eta^2} + \mathfrak{b}\exp\left(-\mathfrak{c}_d N\eta^2\right), \quad d = 1, \ldots, \delta,$$

where $\mathfrak{b}$ is some universal constant, and $\mathfrak{a}_d$ and $\mathfrak{c}_d$ are two positive constants determined by the dependence measure. Moreover, for any κ such that $\eta_d < \kappa\theta_d(u,v) < 1$, we have:

$$\max\left[p_d^+(u,v;\kappa), p_d^-(u,v;\kappa)\right] \leq \frac{2^{4H_d+2}\mathfrak{a}_d}{N\kappa^2 L_{d,t}^4\Delta^{4H_{d,t}}} + \mathfrak{b}\exp\left(-\frac{\mathfrak{c}_d}{2^{4H_{d,t}+2}}N\kappa^2 L_{d,t}^4\Delta^{4H_{d,t}}\right), \quad d = 1, \ldots, \delta.$$

Proof of Lemma S.1. Following the lines of the proof of Lemma 6, we can rewrite $\widehat{\theta}_d(u,v) - \theta_d(u,v)$ as the sum of a zero mean stochastic term and a bias term,

$$\widehat{\theta}_d(u,v) - \theta_d(u,v) = \frac{1}{N}\sum_{n=1}^{N} Z_{n,d}(u,v) + \left\{\mathbb{E}\left[\widehat{\theta}_d(u,v)\right] - \theta_d(u,v)\right\},$$

where, for any $n = 1, \ldots, N$,

$$Z_{n,d} = Z_{n,d}(u,v) = \left(\widetilde{\nabla^d X}_n(u) - \widetilde{\nabla^d X}_n(v)\right)^2 - \mathbb{E}\left(\widetilde{\nabla^d X}_n(u) - \widetilde{\nabla^d X}_n(v)\right)^2.$$

Bounds for the bias term. Since $\{X_n\}$ is stationary and the sequence

$$\{\zeta_n = (M_n, T_{n,1}, \ldots, T_{n,M_n}, \varepsilon_{n,1}, \ldots, \varepsilon_{n,M_n}), n \geq 1\}$$

is i.i.d. (see assumptions (H2),(H3),(H5) and (H4)), the processes $\left\{\widetilde{\nabla^d X}_n\right\}$, $1 \le d \le \delta$, are also stationary, and thus

$$\mathbb{E}\left[\widehat{\theta}_d(u,v)\right] - \theta_d(u,v) = 2\mathbb{E}\left[\{G_{n,d}(u) - G_{n,d}(v)\}\left\{\nabla^d X_n(u) - \nabla^d X_n(v)\right\}\right] + \mathbb{E}\left[\{G_{n,d}(u) - G_{n,d}(v)\}^2\right],$$

where $G_{n,d}(u) = \widetilde{\nabla^d X}_n(u) - \nabla^d X_n(u)$. Since $(x+y)^2 \le 2(x^2+y^2)$, by Assumption (D2), we get

$$\mathbb{E}\left[\{G_{n,d}(u) - G_{n,d}(v)\}^2\right] \le 4R_{2,d}(\lambda).$$

Cauchy-Schwarz inequality then implies

$$\left|\mathbb{E}\left[\widehat{\theta}_d(u,v)\right] - \theta_d(u,v)\right| \le \eta_d/2. \tag{S.11}$$

Concentration bounds for the stochastic term. Recall that, for any $N \ge 1$, the finite sequence $\{\zeta_n, 1 \le n \le N\}$ is i.i.d. and this implies that the finite sequence $\{\widetilde{\nabla^d X}_n, 1 \le n \le N\}$ is also stationary. We now complete these finite sequences to infinite ones, $\{\zeta_n, n \in \mathbb{Z}\}$ and $\{\widetilde{\nabla^d X}_n, n \in \mathbb{Z}\}$, by generating independent M_n from the same distribution as $M_1, \ldots, M_N$, and independent copies $(T_{n,1}, \varepsilon_{n,1}), \ldots, (T_{n,M_n}, \varepsilon_{n,M_n})$ of (T, ε), for any $n \notin \{1, \ldots, N\}$. By the definition (S.5) and using the MA representation of $\{X_n\}$, see (6) in Definition 3, we can rewrite $\widetilde{\nabla^d X}_n$ as,

$$\begin{aligned}
\widetilde{\nabla^d X}_n(u) &= \sum_{i=1}^{M_n} W_{n,i}^{(d)}(u) X_n(T_{n,i}) + \sum_{i=1}^{M_n} W_{n,i}^{(d)}(u)\sigma(T_{n,i})\varepsilon_{n,i}, \\
&= \sum_{i=1}^{M_n} W_{n,i}^{(d)}(u) f(\xi_n, \xi_{n-1}, \ldots)(T_{n,i}) + \sum_{i=1}^{M_n} W_{n,i}^{(d)}(u)\sigma(T_{n,i})\varepsilon_{n,i}, \\
&= g_d\left((\zeta_n, \xi_n), (\zeta_{n-1}, \xi_{n-1}), \ldots\right)
\end{aligned}$$

where $\{(\zeta_n, \xi_n)\}$ are i.i.d. in the measurable space $\widetilde{\mathcal{S}} = \mathbb{N}^* \times \{\cup_{m\ge1}[0,1]^m \times \mathbb{R}^m\} \times \mathcal{S}$ and $g_d : \widetilde{\mathcal{S}}^\infty \to \mathcal{H}$ is a measurable function. Then a coupled version of $\widetilde{\nabla^d X}_n(u)$ is

$$\widetilde{\nabla^d X}_n^{(m)}(u) = \sum_{i=1}^{M_n} W_{n,i}^{(d)} X_n^{(m)}(T_{n,i}) + \sum_{i=1}^{M_n} W_{n,i}^{(d)} \sigma(T_{n,i})\varepsilon_{n,i}, \qquad m \ge 1.$$

A consequence is that

$$\left|\widetilde{\nabla^d X}_n^{(m)}(u) - \widetilde{\nabla^d X}_n(u)\right| \le \mathfrak{c}_W \|X_n - X_n^{(m)}\|_\infty,$$

according to Assumption (D1). By Definition 3, the sequence is then $\mathbb{L}^4 - m-$approximable. Lemma 3 then entails that the sequence $\{Z_{n,d}\}$ is $\mathbb{L}^2 - m-$approximable. Let

$$\nu_{m,2} = \nu_2\left(Z_{m,d} - Z_{m,d}^{(m)}\right),$$

be its dependence coefficient, where $Z_{m,d}^{(m)}$ is the associated coupled version of $Z_{m,d}$. Using Cauchy-Schwartz inequality and the stationary, we get

$$\nu_{m,2} \le 8\nu_4\left(\left\|\widetilde{\nabla^d X}_m\right\|_\infty\right)\nu_4\left(\left\|X_m - X_m^{(m)}\right\|_\infty\right).$$

Since $\{\widetilde{\nabla^d X}_n\}$ satisfies (D1), $\{X_n\}$ satisfies (H7), $\sigma(\cdot)$ is bounded and (H8) guarantees $\nu_4(\varepsilon) < \infty$, we necessarily have $\nu_4(\|\widetilde{\nabla^d X}_m\|_\infty) \le C$, for some constant C independent of n. Finally, since by our conditions $\nu_4(\|X_m - X_m^{(m)}\|_\infty) = \mathcal{O}(1/m^\alpha)$ with $\alpha > 3/2$, the dependence coefficient of $\{Z_{m,d}\}$ satisfies the following condition

$$\nu := \sum_{m=1}^{\infty} \nu_{m,2}^{2/3} < \infty,$$

and thus allows us to apply Lemma 5 above. More precisely, using (S.11), we first get

$$\mathbb{P}\left(\widehat{\theta}_d(u,v)-\theta_d(u,v)>\eta\right)\leq \mathbb{P}\left(\frac{1}{N}\sum_{n=1}^{N}Z_{n,d}>\eta/2\right).$$

Applying next Lemma 5, for any $\eta\in(\eta_d,1)$, we get

$$\begin{aligned}\mathbb{P}\left(\widehat{\theta}_d(u,v)-\theta_d(u,v)>\eta\right)&\leq \frac{4c_2\left(\nu^3+\nu_2^2(Z_{1,d})\right)}{N\eta^2}+c_2'\exp\left(-\frac{c_2}{4\nu^3}N\eta^2\right)+2\exp\left(-\frac{c_2}{4\nu_2^2(Z_{1,d})}N\eta^2\right)\\&\leq \frac{\mathfrak{a}_d}{N\eta^2}+\mathfrak{b}\exp\left(-\mathfrak{e}_d N\eta^2\right),\end{aligned}$$

where $\mathfrak{a}_d=4c_2\left(\nu^3+\nu_2^2(Z_{1,d})\right)$, $\mathfrak{b}=c_2'+2$ and $\mathfrak{e}_d=\min\left(c_2/(4\nu^3),c_2/(4\nu_2^2(Z_{1,d}))\right)$. Using (4) (with $L_{d,t}^2=\mathbb{E}[\{\nabla^d X(t)\}^2]$ and $\Delta_{d,0}=\Delta_{\delta,0}$ when $d<\delta$; see also the proof of Lemma 1), and the condition (S.6), we get

$$\theta_d(u,v)\geq |u-v|^{2H_{d,t}}L_{d,t}^2/2>0.$$

We thus get that $\eta_d<\theta_d(u,v)$, for a sufficiently large λ, values $\kappa\in(0,1)$ exist such that $\eta_d<\kappa\theta_d(u,v)$. We can then consider $\eta=\kappa\theta_d(u,v)$ in the Nagaev-type inequality and deduce

$$\begin{aligned}p_d^+(u,v;\kappa)\mathbb{P}\left(\widehat{\theta}_d(u,v)>(1+\kappa)\theta_d(u,v)\right)&\leq \frac{\mathfrak{a}_d}{N\kappa^2\theta_d^2(u,v)}+\mathfrak{b}\exp\left(-\mathfrak{e}_d N\kappa^2\theta_d^2(u,v)\right)\\&\leq \frac{2^{4H_{d,t}+2}\mathfrak{a}_d}{N\kappa^2L_{d,t}^4\Delta^{4H_{d,t}}}+\mathfrak{b}\exp\left(-\frac{\mathfrak{e}_d}{2^{4H_{d,t}+2}}N\kappa^2L_{d,t}^4\Delta^{4H_{d,t}}\right).\end{aligned}$$

Similar arguments apply for bounding $p_d^-(u,v;\kappa)$. The proof of Lemma S.1 is now complete. □

In view of the proof of Lemma 1, let us define

$$\beta_d=1/2\quad \text{if }\ d\leq\delta-2,\qquad\text{and}\qquad \beta_{\delta-1}=\underline{H},\tag{S.12}$$

$$L_{d,t}^2:=\mathbb{E}\left[\left(\nabla^{d+1}X(t)\right)^2\right]\in[\,\underline{a}_{d+1},\overline{a}_{d+1}],\quad 1\leq d\leq\delta-1,\tag{S.13}$$

and

$$S_d^2=2\sqrt{\overline{a}_{d+1}}\sqrt{\overline{a}_{d+2}}+\overline{a}_{d+2}\quad\text{if }\ d\leq\delta-2,\qquad\text{and}\qquad S_{\delta-1}^2=2\sqrt{\overline{a}_{d+1}}\sqrt{\overline{L}^2+S_\delta^2}+\overline{L}^2+S_\delta^2.\tag{S.14}$$

Lemma S.2. *1. For any $\varphi\in(0,1)$ and $0<\Delta\leq\Delta_{\delta,0}$ such that*

$$\Delta^{2\beta_d}S_d^2<\frac{L_{d,t}^2\log(2)}{4}\varphi,\qquad \forall 1\leq d\leq\delta,$$

with β_d, $L_{d,t}^2$ and S_d^2 in (S.12), (S.13) and (S.14), respectively, then

$$\left|\widetilde{H}_{d,t}-H_{d,t}\right|<\varphi/2,\qquad \forall 1\leq d\leq\delta.$$

2. Let $1\leq d\leq\delta$, and let $H\in(0,1]$ such that $|H-H_{d,t}|<\varphi<1$. For any $\psi\in(0,1)$ and $0<\Delta\leq\Delta_{\delta,0}$ such that $S_d^2\Delta^{2\beta_d-2\varphi}<\psi/3$, we have

$$\left|\frac{\theta_d(t_1,t_3)-L_{d,t}\Delta^{2H_{d,t}}}{\Delta^{2H}}\right|<\psi/3.$$

Proof of Lemma S.2. 1. Let $1\leq d\leq\delta$ be a fixed integer. From the proof of Lemma 1 and condition (4), we can rewrite $\theta_d(u,v)$, with $u,v\in[t-\Delta/2,t+\Delta/2]\subset[t-\Delta_{\delta,0}/2,t+\Delta_{\delta,0}/2]$ as

$$\theta_d(u,v)=L_{d,t}^2|u-v|^{2H_d}(1+\rho_d(u,v)),$$

where $|\rho_d(u,v)| \leq (S_{d,t}/L_{d,t})^2 \Delta^{2\beta_d}$. Since our conditions imply $\Delta^{2\beta_d} S_{d,t}^2 < L_{d,t}^2 \log(2)/4$, we deduce $|\rho_d(u,v)| < 1/2$. Using the fact that $x \mapsto \log(1+x)$ is Lipschitz for $x \in (-1/2, +\infty)$, we get:

$$\begin{aligned} |\widetilde{H}_{d,t} - H_{d,t}| &= \frac{|\log(1+\rho_d(t_1,t_3)) - \log(1+\rho_d(t_1,t_2))|}{\log(2)} \\ &\leq \frac{|\rho_d(t_1,t_3)| + |\rho_d(t_1,t_2)|}{\log(2)} \\ &\leq \frac{2}{\log(2)} \left(\frac{S_d}{L_{d,t}}\right)^2 \Delta^{2\beta_d}. \end{aligned}$$

We then deduce from the condition on φ that $|\widetilde{H}_{d,t} - H_{d,t}| < \varphi/2$.

2. By condition (4), if $\Delta^{2\beta_d - 2\varphi} S_d^2 < \psi/3$, we get

$$\left| \frac{\theta_d(t_1,t_3) - L_{d,t}^2 \Delta^{2H_{d,t}}}{\Delta^{2H}} \right| < S_d^2 \Delta^{2\beta_d + 2(H_{d,t} - H)} < S_d^2 \Delta^{2\beta_d - 2\varphi} < \psi/3.$$

□

Lemma S.3. *Assume that the conditions of Proposition 1 hold true. For any $d \in \{1,\dots,\delta\}$, there exists a universal positive constant $\mathfrak{b}$, and positive constants $\mathfrak{f}_d$ and $\mathfrak{g}_d$ depending on dependence measure such that the following inequality holds:*

$$\mathbb{P}\left(|\widehat{H}_{d,t} - H_{d,t}| > \varphi\right) \leq \frac{\mathfrak{f}_d}{N\varphi^2 \Delta^{4H_{d,t}}} + 4\mathfrak{b} \exp\left(-\mathfrak{g}_d N \varphi^2 \Delta^{4H_{d,t}}\right).$$

Proof of Lemma S.3. According to condition (S.6) and Lemma S.2, we have that $|\widetilde{H}_{d,t} - H_{d,t}| \leq \varphi/2$. It then follows that,

$$\begin{aligned} \mathbb{P}(|\widehat{H}_{d,t} - H_{d,t}| > \varphi) &\leq \mathbb{P}\left(\left|\widehat{H}_{d,t} - \widetilde{H}_{d,t}\right| > \varphi/2\right) \\ &\leq \mathbb{P}\left(\left|\log \frac{\widehat{\theta}_d(t_1,t_3)}{\theta_d(t_1,t_3)} \frac{\theta_d(t_1,t_2)}{\widehat{\theta}_d(t_1,t_2)}\right| > \varphi \log(2)\right) \\ &\leq \mathbb{P}\left(\frac{\widehat{\theta}_d(t_1,t_3)}{\theta_d(t_1,t_3)} \frac{\theta_d(t_1,t_2)}{\widehat{\theta}_d(t_1,t_2)} > 2^{-\varphi}\right) + \mathbb{P}\left(\frac{\widehat{\theta}_d(t_1,t_3)}{\theta_d(t_1,t_3)} \frac{\theta_d(t_1,t_2)}{\widehat{\theta}_d(t_1,t_2)} < 2^{-\varphi}\right). \end{aligned}$$

By simple algebra and the definition of the functions p_d^+ and p_d^- introduced in Lemma S.1, we get:

$$\begin{aligned} \mathbb{P}(|\widehat{H}_{d,t} - H_{d,t}| > \varphi) \leq p_d^+(t_1,t_3; 2^{\varphi/2} - 1) + p_d^-(t_1,t_3; 1 - 2^{-\varphi/2}) \\ + p_d^+(t_1,t_2; 2^{\varphi/2} - 1) + p_d^-(t_1,t_2; 1 - 2^{-\varphi/2}), \end{aligned} \tag{S.15}$$

provided that $\eta_d(\lambda) < |2^{\pm\varphi/2} - 1| \theta_d(u,v) < 1$ which is guaranteed by condition (S.7) with $C_d = 5B^{-1/2}(2\sqrt{\overline{a}_d} + \sqrt{B})^{-1} \log(2)/2^{11/2}$. To see this, first note that for any $\varphi \in (0,1)$, $|2^{\pm\varphi/2} - 1| \leq \varphi \log(2)/2^{1/2}$. Thus, by (4) and (S.6), we have

$$|2^{\pm\varphi/2} - 1| \theta_d(u,v) \leq \left(5\log(2)/2^{5/2}\right) \varphi L_{d,t}^2 \Delta^{2H_{d,t}} < 1 \quad \text{as} \quad \Delta \to 0.$$

Second, (D2) entails that $\eta_d(\lambda) < 8\left(2\sqrt{\overline{a}_d} + \sqrt{B}\right) B^{1/2} \lambda^{-\tau/2}$. Gathering the two bounds, we obtain

$$\lambda^{-\tau/2} < \left(5B^{-1/2}\left(2\sqrt{\overline{a}_d} + \sqrt{B}\right)^{-1} \log(2)/2^{11/2}\right) \varphi L_{d,t}^2 \Delta^{2H_{d,t}},$$

which is exactly the condition (S.7). Now, with $t_k = t_2$ or $t_k = t_3$, we have

$$p_d^+(t_1,t_k; 2^{\varphi/2} - 1) \leq \frac{2^{4H_{d,t}+2} \mathfrak{a}_d}{N(2^{\varphi/2} - 1)^2 L_{d,t}^4 \Delta^{4H_{d,t}}} + \mathfrak{b} \exp\left(-\frac{\mathfrak{e}_d}{2^{4H_{d,t}+2}} N (2^{\varphi/2} - 1)^2 L_{d,t}^4 \Delta^{4H_{d,t}}\right).$$

Since $\log x \leq x-1$ for any $x>0$, we get $\log(2^{\varphi/2}) \leq 2^{\varphi/2}-1$. We obtain,

$$p_d^+(t_1,t_k;2^{\varphi/2}-1) \leq \frac{2^{4H_{d,t}+4}\mathfrak{a}_d/\log(2)^2}{N\varphi^2 L_{d,t}^4 \Delta^{4H_{d,t}}} + \mathfrak{b}\exp\left(-\frac{\mathfrak{e}_d \log(2)^2}{2^{4H_{d,t}+4}} N\varphi^2 L_{d,t}^4 \Delta^{4H_{d,t}}\right).$$

Setting $\mathfrak{f}_d = 2^{4H_{d,t}+6}\mathfrak{a}_d/(\log(2)^2 L_{d,t}^4)$ and $\mathfrak{g}_d = \mathfrak{e}_d L_{d,t}^4 \log(2)^2/2^{4H_{d,t}+4}$, we finally get:

$$p_d^+(t_1,t_k;2^{\varphi/2}-1) \leq \frac{\mathfrak{f}_d/4}{N\varphi^2\Delta^{4H_{d,t}}} + \mathfrak{b}\exp\left(-\mathfrak{g}_d N\varphi^2\Delta^{4H_{d,t}}\right).$$

The same reasoning can be applied to bound the other three terms on the right-hand side of (S.15). See also the arguments used in the proof of Theorem 1. □

Proof of Proposition 1. Note that:

$$\begin{aligned}
\mathbb{P}\left(|\widehat{\alpha}_t-\alpha_t|>\varphi\right) &\leq \mathbb{P}\left(|\widehat{\alpha}_t-\alpha_t|>\varphi, \widehat{\delta}=\delta\right) + \mathbb{P}\left(\widehat{\delta}\neq\delta\right) \\
&\leq \mathbb{P}\left(\left|\widehat{H}_{\delta,t}-H_{\delta,t}\right|>\varphi\right) + \mathbb{P}\left(\widehat{\delta}<\delta\right) + \mathbb{P}(\widehat{\delta}>\delta) \\
&\leq \mathbb{P}\left(\left|\widehat{H}_{\delta,t}-H_{\delta,t}\right|>\varphi\right) + \sum_{d=0}^{\delta-1}\mathbb{P}\left(\widehat{H}_{d,t}<1-\varphi\right) + \mathbb{P}\left(\widehat{H}_{\delta,t}>1-\varphi\right) \\
&\leq \sum_{d=0}^{\delta}\mathbb{P}\left(\left|\widehat{H}_{d,t}-H_{d,t}\right|>\varphi\right) + \mathbb{P}\left(\left|\widehat{H}_{\delta,t}-H_{\delta,t}\right|>1-H_{\delta,t}-\varphi\right).
\end{aligned}$$

For the last inequality we use the fact that, for $d<\delta$ we have $H_{d,t}=1$, while $H_{\delta,t}<1$. Since $1-H_{\delta,t}>2\varphi$ for sufficiently large λ, repeatedly applying Lemma S.3, we have

$$\begin{aligned}
\mathbb{P}(|\widehat{\alpha}_t-\alpha_t|>\varphi(\lambda)) &\leq \frac{\mathfrak{f}_\delta}{N\varphi^2\Delta^{4H_{\delta,t}}} + 4\mathfrak{b}\exp\left(-4\mathfrak{g}_\delta N\varphi^2\Delta^{4H_{\delta,t}}\right) \\
&\quad + \sum_{d=0}^{\delta}\frac{\mathfrak{f}_d}{N\varphi^2\Delta^{4H_{d,t}}} + 4\mathfrak{b}\exp\left(-\mathfrak{g}_d N\varphi^2\Delta^{4H_{d,t}}\right)
\end{aligned}$$

Setting $\mathfrak{f}=\max\{\mathfrak{f}_0,\ldots,\mathfrak{f}_\delta\}$ and $\mathfrak{g}=\min\{\mathfrak{g}_0,\ldots,\mathfrak{g}_\delta\}$, after changing $4\mathfrak{b}$ to $\mathfrak{b}$, we get:

$$\mathbb{P}(|\widehat{\alpha}_t-\alpha_t|>\varphi) \leq (2+\delta)\left[\frac{\mathfrak{f}}{N\varphi^2\Delta^4} + \mathfrak{b}\exp\left(-\mathfrak{g}N\varphi^2\Delta^4\right)\right].$$

□

Proof of Proposition 2. First, we may rewrite $\widehat{L}_{d,t}^2 - L_{d,t}^2$ as the sum of three terms such that :

$$\left|\widehat{L}_{d,t}^2 - L_{d,t}^2\right| \leq \frac{\left|\widehat{\theta}_d(t_1,t_3)-\theta_d(t_1,t_3)\right|}{\Delta^{2\widehat{H}_d}} + \frac{\left|\theta_d(t_1,t_3)-L_{d,t}^2\Delta^{2H_d}\right|}{\Delta^{2\widehat{H}_d}} + L_{d,t}^2\left|1-\Delta^{2H_d-2\widehat{H}_d}\right|.$$

It then follows that,

$$\mathbb{P}\left(\left|\widehat{L}_{d,t}^2-L_{d,t}^2\right|>\psi\right) \leq \mathbb{P}\left(\left|\widehat{H}_{d,t}-H_{d,t}\right|\leq\varphi, \left|\widehat{L}_{d,t}^2-L_{d,t}^2\right|>\psi\right) + \mathbb{P}\left(\left|\widehat{H}_{d,t}-H_{d,t}\right|>\varphi\right).$$

On the event $\left|\widehat{H}_{d,t}-H_{d,t}\right|\leq\varphi$, using (S.8) and Lemma S.2, we get

$$\frac{|\theta_d(t_1,t_3)-L_{d,t}^2\Delta^{2H_d}|}{\Delta^{2\widehat{H}_d}} \leq \psi/3.$$

Furthermore, the function $x \in [-\varphi, \varphi] \to \Delta^{2x}$ is Lipschitz. Consequently we get,

$$L_{d,t}^2 |1 - \Delta^{2H_d - 2\widehat{H}_d}| \leq \psi/3,$$

provided $\left|\widehat{H}_{d,t} - H_{d,t}\right| \leq \varphi$ and condition (S.9) holds true. We then deduce that,

$$\begin{aligned}
\mathbb{P}\left(|\widehat{L}_{d,t}^2 - L_{d,t}^2| > \psi\right) &\leq \mathbb{P}\left(\left|\widehat{H}_{d,t} - H_{d,t}\right| \leq \varphi, \frac{|\widehat{\theta}_d(t_1,t_3) - \theta_d(t_1,t_3)|}{\Delta^{2\widehat{H}_{d,t}}} > \psi/3\right) \\
&\quad + \mathbb{P}\left(\left|\widehat{H}_{d,t} - H_{d,t}\right| > \varphi\right), \\
&\leq \mathbb{P}\left(\left|\widehat{H}_{d,t} - H_{d,t}\right| \leq \varphi, \left|\widehat{\theta}_d(t_1,t_3) - \theta_d(t_1,t_3)\right| > \Delta^{2H_{d,t}+2\varphi}\psi/3\right) \\
&\quad + \mathbb{P}\left(\left|\widehat{H}_{d,t} - H_{d,t}\right| > \varphi\right), \\
&\leq \mathbb{P}\left(\left|\widehat{\theta}_d(t_1,t_3) - \theta_d(t_1,t_3)\right| > \Delta^{2H_{d,t}+2\varphi}\psi/3\right) + \mathbb{P}\left(\left|\widehat{H}_{d,t} - H_{d,t}\right| > \varphi\right).
\end{aligned}$$

The second probability of the right-hand side of the last inequality can be bounded using Lemma S.3 and the first probability using Lemma S.1, provided that $\eta_d(\lambda) < \Delta^{2H_{d,t}+2\varphi}\psi/3 < 1$ which is guaranteed by condition (S.10) with $\widetilde{C}_d = B^{-1/2}(2\sqrt{\overline{a}_d} + \sqrt{B})^{-1}/(3 \times 2^3)$. In fact, note that the Assumption (D2) implies that $\eta_d(\lambda) \leq 8\left(2\sqrt{\overline{a}_d} + \sqrt{B}\right) B^{1/2}\lambda^{-\tau/2}$, hence

$$\lambda^{-\tau/2} < B^{-1/2}(2\sqrt{\overline{a}_d} + \sqrt{B})^{-1}/(3 \times 2^3)\Delta^{2\varphi}\psi\Delta^{2H_{d,t}} < 1 \quad \text{as} \quad \Delta \to 0.$$

Proposition 2 then follows:

$$\begin{aligned}
\mathbb{P}\left(\left|\widehat{L}_{d,t}^2 - L_{d,t}^2\right| > \psi\right) \leq \frac{\mathfrak{c}_d}{N\psi^2\Delta^{4H_d+4\varphi}} + \mathfrak{b}\exp\left(-\mathfrak{l}_d N\psi^2\Delta^{4H_d+4\varphi}\right) \\
+ \frac{\mathfrak{f}_d}{N\varphi^2\Delta^{4H_d}} + 4\mathfrak{b}\exp\left(-\mathfrak{g}_d N\varphi^2\Delta^{4H_d}\right),
\end{aligned}$$

where $\mathfrak{c}_d = 9\mathfrak{a}_d$ and $\mathfrak{l}_d = \mathfrak{e}_d/9$. □

S.3 Proofs for adaptive estimation

In this section we provide proofs for the results in Appendix C.

S.3.1 Technical lemmas

Lemma 7. *Assume that $X \in \mathcal{X}(H, \boldsymbol{L}; J)$ and let $\widehat{X}_n(t,h)$ be defined as in (17). Assume (H1) to (H6), (H11) and (H12) hold true. Then :*

1. *$\{B_n(t;h)\}$ and $\{V_n(t;h)\}$ are conditionally independent given $\{M_n\}$ and $\{T_{n,i}\}$;*
2. *$\{V_n(t;h)\}$ are conditionally independent given $\{M_n\}$ and $\{T_{n,i}\}$;*
3. *$\mathbb{E}_{M,T}\left[V_n^2(t;h)\right] \leq \{1 + o(1)\}\sigma^2(t)\max_{1\leq i\leq M_n} W_{n,i}(t;h)$, with $o(1)$ uniform with respect to $h \in \mathcal{H}_N$;*
4. *$\mathbb{E}_{M,T}\left[B_n^2(t;h)\right] \leq L_t^2 h^{2H_t} b_n(t;h,2H_t)\{1 + o(1)\}$, with $o(1)$ uniform with respect to $h \in \mathcal{H}_N$;*

Proof of Lemma 7. By definition, we have

$$B_n(t;h) = \sum_{i=1}^{M_n} W_{n,i}(t;h)\{X_n(T_{n,i}) - X_n(t)\} \quad \text{and} \quad V_n(t;h) = \sum_{i=1}^{M_n} \varepsilon_{n,i} W_{n,i}(t;h).$$

1) Direct consequence of the definitions and assumption (H5).

2) Direct consequence of the definitions and assumption (H4).

3) By elementary calculations and (H4), and since $\left|\sigma^2(T_{n,i}) - \sigma^2(t)\right| \leq L_\sigma |T_{n,i} - t|$ for some $L_\sigma > 0$, uniformly with respect to h, we have

$$\mathbb{E}_{M,T}\left[V_n^2(t;h)\right] = \sum_{i=1}^{M_n} W_{n,i}^2(t;h)\sigma^2(T_{n,i}) \leq \{1+o(1)\}\sigma^2(t) \max_{1\leq i\leq M_n} W_{n,i}(t;h).$$

4) By Jensen's inequality,

$$B_n^2(t;h) \leq \sum_{i=1}^{M_n} W_{n,i}(t;h)\left\{X_n(T_{n,i}) - X_n(t)\right\}^2.$$

Then condition (4) in (H6) implies

$$\begin{aligned}\mathbb{E}_{M,T}\left[B_n^2(t;h)\right] &\leq L_t^2 \sum_{i=1}^{M_n} W_{n,i}(t;h)|T_{n,i} - t|^{2H_t} \times \left\{1 + h^{2\beta_0} S_0^2/L_t^2\right\} \\ &= L_t^2 h^{2H_t} b_n(t;h,2H_t) \times \left\{1 + h^{2\beta_0} S_0^2/L_t^2\right\}.\end{aligned}$$

Moreover, for any $t \in I$ and $h \in \mathcal{H}_N$, by (H12) we get

$$0 < h^{2\beta_0} S_0^2/L_t^2 \leq \left(\max \mathcal{H}_N\right)^{2\beta_0} S_0^2/\underline{L} \to 0.$$

The statement then follows. □

Lemma 8. *Assume that the assumptions (H1) to (H5), and (H12) to (H14) hold true.*

1. *For any $t \in (0,1)$ and $h \in \mathcal{H}_N$ such that $\int_{t-h}^{t+h} g(u)du \leq 1/2$, we have*

$$1 - \exp\left(-M_n p(t;h)\right) \leq \mathbb{E}[\pi_n(t;h) \mid M_n] \leq 1 - \exp\left(-2M_n p(t;h)\right), \qquad \forall 1 \leq n \leq N.$$

2. *There exists two constants $\underline{C}_\mu$ and $\overline{C}_\mu$ such that for all $h \in \mathcal{H}_N$,*

$$\underline{C}_\mu\{1+o(1)\} \leq \frac{\mathbb{E}[P_N(t;h)]}{N\min(1,\lambda h)} \leq \overline{C}_\mu\{1+o(1)\},$$

and $P_N(t;h) = \mathbb{E}[P_N(t;h)]\{1+o_{\mathbb{P}}(1)\}$, with $o(1)$ and $o_{\mathbb{P}}(1)$ uniform with respect to $h \in \mathcal{H}_N$.

3. *Moreover if (H16) holds, constants $\underline{C}_\gamma$ and $\overline{C}_\gamma$ exist such that $\forall h \in \mathcal{H}_N$,*

$$\underline{C}_\gamma\{1+o(1)\} \leq \frac{\mathbb{E}[P_{N,\ell}(s,t;h)]}{(N-\ell)\min(1,(\lambda h)^2)} \leq \overline{C}_\gamma\{1+o(1)\},$$

and $P_{N,\ell}(s,t;h) = \mathbb{E}[P_{N,\ell}(s,t;h)]\{1+o_{\mathbb{P}}(1)\}$, with $o(1)$ and $o_{\mathbb{P}}(1)$ uniform with respect to $h \in \mathcal{H}_N$.

Proof of Lemma 8. 1) For simplicity, we omit the subscript n, and write M and $\pi(t;h)$. Since $\{T_{n,i}\}$ are independent by (H3), we have

$$\mathbb{E}[\pi(t;h) \mid M] = 1 - (1 - p(t;h))^M, \qquad \text{with } \; p(t;h) = \int_{t-h}^{t+h} g(u)du.$$

We remark that, using the Hölder continuity of g (H14), we obtain $p(t;h) = 2hg(t)\{1+o(1)\}$ where $o(1)$ converges to 0 uniformly with respect to $h \in \mathcal{H}_N$. Thus, for a sufficiently small $\max \mathcal{H}_N$, we can assume that $p(t;h) < 1/2$, $\forall h \in \mathcal{H}_N$. Using the following elementary inequality,

$$-\frac{u}{1-u} \leq \log(1-u) \leq -u, \quad \forall u \in (0,1),$$

we deduce that for any $u \in (0, 1/2)$, and for any $M \geq 0$

$$1 - \exp(-Mu) \leq 1 - (1-u)^M < 1 - \exp(-2Mu).$$

Replacing u by $p(t;h)$, the first statement follows.

2) Note that $P_N(t;h) = \sum_{n=1}^N \pi_n(t;h)$ is a sum of N independent Bernoulli random variables, and $\mathbb{E}[P_N(t;h)] = N\mathbb{E}[\pi_1(t;h)]$. Assumption (H13) implies that for all $u \in (0,1)$, $\exp(-\overline{c}\lambda u) \leq \mathbb{E}(e^{-uM}) \leq \exp(-\underline{c}u\lambda)$. Then, by 1) above we get,

$$1 - \exp(-\lambda h(1+o(1))2g(t)\underline{c}) \leq \frac{\mathbb{E}[P_N(t;h)]}{N} \leq 1 - \exp(-\lambda h(1+o(1))2g(t)\overline{c}).$$

If $\lambda h \geq 1$, the following inequality hold by (H14),

$$\{1 - \exp(-2\underline{c}_g\underline{c})\}(1+o(1)) \leq \frac{\mathbb{E}[P_N(t;h)]}{N} \leq 1.$$

If $\lambda h < 1$, we remark that if $C > 0$, $1 - e^{-Cu} \geq (1-e^C)u$ for all $u \in (0,1)$ and that $1 - e^{-x} \leq x$ for all $x \in \mathbb{R}$. We deduce by (H14),

$$\left\{1 - \exp\left(-2\underline{c}_g\underline{c}\right)\right\} \lambda h(1+o(1)) \leq \frac{\mathbb{E}[P_N(t;h)]}{N} \leq \lambda h\, 2\overline{c}_g\overline{c}\{1+o(1)\}.$$

Gathering facts, the double inequality in 2) follows by setting $\underline{C}_\mu = 1 - \exp\left(-2\underline{c}_g\underline{c}\right)$ and $\overline{C}_\mu = \max(1, 2\overline{c}_g\overline{c})$. Next, using the left bound of this double inequality and Chernoff's (see, for instance, Vershynin, 2018, Section 2.3) exponential bound, for any $0 < \eta < 1$,

$$\begin{aligned}\mathbb{P}\left(\left|\frac{P_N(t;h)}{\mathbb{E}[P_N(t;h)]} - 1\right| > \eta\right) &\leq 2\exp\left(-\eta^2\mathbb{E}[P_N(t;h)]/3\right)\\ &\leq 2\exp\left(-\underline{C}_\mu\eta^2 N\min(1, \lambda\min\mathcal{H}_N)/3\right).\end{aligned}$$

Since $\mathcal{H}_N$ is a grid of at most $(N\lambda)^c$ points for some $c > 0$, we deduce that

$$\begin{aligned}\mathbb{P}\left(\sup_{h\in\mathcal{H}_N}\left|\frac{P_N(t;h)}{\mathbb{E}[P_N(t;h)]} - 1\right| > \eta\right) &\leq 2(N\lambda)^c\exp\left[-\underline{C}_\mu\eta^2 N\min(1,\lambda\min\mathcal{H}_N)/3\right]\\ &\leq 2\exp\left[-N\min(1,\lambda\min\mathcal{H}_N)\left(\underline{C}_\mu\eta^2/3 - \frac{c\log(N\lambda)}{N\min(1,\lambda\min\mathcal{H}_N)}\right)\right].\end{aligned}$$

Since by Assumption (H12) $N\min(1,\lambda\min\mathcal{H}_N)/\log(N\lambda) \to \infty$, we deduce that $P_N(t;h)/\mathbb{E}[P_N(t;h)]$ converges in probability to 1 uniformly over $h \in \mathcal{H}_N$.

3) Let $\lfloor x \rfloor$ denote the largest integer smaller than x. Note that, for a fixed ℓ, we can decompose

$$P_{N,\ell}(s,t;h) = \sum_{n=1}^{N-\ell}\pi_n(s;h)\pi_{n+\ell}(t;h) = \sum_{l=1}^{\ell+1}P_{N,\ell,l}(s,t;h),$$

where

$$P_{N,\ell,l}(s,t;h) = \sum_{n'=0}^{\lfloor (N-\ell-l)/(\ell+1)\rfloor}\pi_{n'(\ell+1)+l}(s;h)\pi_{n'(\ell+1)+\ell+l}(t;h).$$

Each $P_{N,\ell,l}(s,t;h)$ is a sum of independent Bernoulli random variables because, by definition and the condition we imposed, the $\pi_n(s;h)$'s, $1 \leq n \leq N$, are independent. Moreover,

$$\mathbb{E}[P_{N,\ell,l}(s,t;h)] = \{\lfloor (N-\ell-l)/(\ell+1)\rfloor + 1\}\mathbb{E}[\pi_1(s;h)]\mathbb{E}[\pi_{1+\ell}(t;h)].$$

By arguments as used for 2), constants $\underline{C}$ and $\overline{C}$ exist such that

$$\underline{C}^2 \leq \frac{\mathbb{E}[P_{N,\ell,l}(s,t;h)]}{\{\lfloor (N-\ell-l)/(\ell+1)\rfloor + 1\}\min(1,(\lambda h)^2)} \leq \overline{C}^2, \qquad 1 \leq l \leq \ell.$$

By little algebra, summing over the integers l between 1 and $\ell+1$, we get the first part of the statement. For the second part, using Chernoff's exponential bound (see, for instance, Vershynin, 2018, Section 2.3) and the fact that $\mathcal{H}_N$ is a grid of up to $(N\lambda)^c$ points, for each $1 \leq l \leq \ell+1$, we deduce that

$$\begin{aligned}\mathbb{P}\left(\sup_{h\in\mathcal{H}_N}\left|\frac{P_{N,\ell,l}(s,t;h)}{\mathbb{E}[P_{N,\ell}(s,t;h)]}-1\right|>\eta\right)\\ &\leq 2\exp\Big[-\lfloor (N-\ell-l)/(\ell+1)\rfloor \min(1,(\lambda\min\mathcal{H}_N)^2)\\ &\qquad\times\left(\frac{C\eta^2}{3}-\frac{c\log(N\lambda)}{\lfloor (N-\ell-l)/(\ell+1)\rfloor\min(1,(\lambda\min\mathcal{H}_N)^2)}\right)\Big]\end{aligned}$$

Using Assumption (H16) and summing over $1 \leq l \leq \ell+1$, the statement follows by elementary algebra. □

Lemma 9. *Let assumptions (H1) to (H7) and (H14) hold true. Then, for any $t \in I$, $\widehat{\sigma}^2(t) = \sigma^2(t)\{1+o_{\mathbb{P}}(1)\}$.*

Proof of Lemma 9. For any $t\in I$,

$$\widehat{\sigma}^2(t)-\sigma^2(t)=\frac{1}{N}\sum_{n=1}^{N}\{Z_n-\mathbb{E}Z_n\}+\{\mathbb{E}Z_n-\sigma^2(t)\},$$

where $Z_n=\{Y_{n,i(t)}-Y_{n,i(t)+1}\}^2/2$. Thus,

$$\mathbb{P}\left(|\widehat{\sigma}^2(t)-\sigma^2(t)|>\eta\right)\leq Q_1+Q_2,$$

where

$$Q_1=\mathbb{P}\left(\sum_{n=1}^{N}|Z_n-\mathbb{E}Z_n|>N\eta/2\right)\quad\text{and}\quad Q_2=\mathbb{P}\left(\left|\mathbb{E}Z_n-\sigma^2\right|>\eta/2\right).$$

Study of Q_2. Using the assumptions (H3), (H4) and (H5), we have

$$\begin{aligned}0\leq \mathbb{E}_{M,T}Z_n-\sigma^2(t)=|\mathbb{E}_{M,T}Z_n-\sigma^2(t)|&=\mathbb{E}_{M,T}\{X_n(T_{n,i(t)})-X_n(T_{n,i(t)+1})\}^2/2\\ &=\frac{1}{2}\mathbb{E}_{M,T}\left[\{X_n(T_{n,i(t)})-X_n(T_{n,i(t)+1})\}^2[\mathbb{1}_{\mathcal{A}}+\mathbb{1}_{\overline{\mathcal{A}}}]\right]\\ &\leq\frac{1}{2}\mathbb{E}_{M,T}\left[\{X_n(T_{n,i(t)})-X_n(T_{n,i(t)+1})\}^2\mathbb{1}_{\mathcal{A}}\right]+2\sup_{u\in I}\mathbb{E}\left[X^2(u)\right]\mathbb{P}(\overline{\mathcal{A}}),\end{aligned}$$

where $\mathcal{A}=\{\left|T_{n,i(t)}-T_{n,i(t)+1}\right|\leq\Delta_{0,0}\}\cap\{T_{n,i(t)},T_{n,i(t)+1}\in J\}$ with J and $\Delta_{0,0}$ the set and the constant from condition (4) in (H6). Then constants C_1 and C_2 exist such that

$$0\leq\mathbb{E}Z_n-\sigma^2(t)\leq C_1L_t^2\mathbb{E}\left[\left|T_{n,i(t)}-T_{n,i(t)+1}\right|^{2H_t}\right]+C_2\mathbb{P}(\overline{\mathcal{A}}).$$

Since the $T_{n,i}$'s are independently drawn and admit a density g bounded and bounded away from zero (H14), it is easy to check that the bound of $\mathbb{E}Z_n-\sigma^2(t)$ tends to zero, provided $\lambda\to\infty$. See also (Golovkine et al., 2022, Section F) for the moments of the spacings between the ordered $T_{n,i}$.

Study of Q_1. By (1),

$$Y_{n,i(t)}=X_n(T_{n,i(t)})+\sigma^2((T_{n,i(t)})\varepsilon_{n,i(t)},\qquad 1\leq n\leq N.$$

The infinite sequence $\{X_n,n\in\mathbb{Z}\}$ is stationary. Recall that in our setup, $T_{n,i}$, $1\leq i\leq M_n$, $1\leq n\leq N$ is a triangular array, with $M_1,\ldots,M_N$ independent copies of M which has a distribution which changes with N, while the $T_{n,i}$'s are independent copies of T with a fixed distribution. Thus, for any $N\geq 1$, the finite sequence $\{\zeta_n=(T_{n,i(t)},\varepsilon_{n,i(t)}),1\leq n\leq N\}$ is i.i.d., see assumptions (H3), (H4) and (H5). This implies that the finite sequence $\{Y_{n,i(t)},1\leq n\leq N\}$ is also stationary. We next complete these finite sequences to infinite ones, $\{\zeta_n, n \in \mathbb{Z}\}$ and $\{Y_{n,i(t)}, n \in \mathbb{Z}\}$, by generating independent M_n from the same distribution as $M_1, \dots, M_N$, and independent copies $(T_{n,1}, \varepsilon_{n,1}), \dots, (T_{n,M_n}, \varepsilon_{n,M_n})$ of (T, ε), for any $n \notin \{1, \dots, N\}$. Using the MA representation of $\{X_n\}$, see (6) in Definition 3, we then rewrite $Y_{n,i(t)}$

$$Y_{n,i(t)} = f(\xi_n, \xi_{n-1}, \dots)(T_{n,i(t)}) + \sigma^2((T_{n,i(t)})\varepsilon_{n,i(t)} = g\left((\zeta_n, \xi_n)(\zeta_{n-1}, \xi_{n-1}), \dots\right), \quad n \geq 1,$$

where $\{(\zeta_n, \xi_n), n \in \mathbb{Z}\}$ is an i.i.d. sequence taking values in a measurable space $\widetilde{\mathcal{S}} = \{[0,1] \times \mathbb{R}\} \times \mathcal{S}$ and $g : \widetilde{\mathcal{S}}^\infty \to \mathcal{C}$ a measurable function. A coupled version of $Y_{n,i(t)}$ is then

$$Y_{n,i(t)}^{(m)} = X_n^{(m)}(T_{n,i(t)}) + \sigma^2((T_{n,i(t)})\varepsilon_{n,i(t)}, \qquad m \geq 1,$$

and we have

$$\left| Y_{n,i(t)}^{(m)} - Y_{n,i(t)} \right| \leq \|X_n^{(m)} - X_n\|_\infty,$$

and deduce that $\{Y_{n,i(t)}, n \geq 1\}$ is $\mathbb{L}^4 - m-$approximable. The same facts hold true for $\{Y_{n,i(t)+1}, n \geq 1\}$. This entails that $\{Z_n\}_n$ is $\mathbb{L}^2 - m$-approximable, since

$$|Z_n - Z_n^{(m)}| \leq 2\left(\|X_n\|_\infty + \|X_n^{(m)}\|_\infty\right) \|X_n^{(m)} - X_n\|_\infty.$$

By Cauchy-Schwarz inequality and Jensen's inequality, we get

$$\nu_2\left(Z_n - Z_n^{(m)}\right) \leq 4\nu_4(\|X_n\|_\infty)\nu_4\left(\|X_n^{(m)} - X_n\|_\infty\right).$$

Condition (H7) then implies

$$\upsilon := \sum_{m=1}^{\infty} \nu_2\left(Z_m - Z_m^{(m)}\right)^{2/3} < \infty.$$

Applying Nagaev's inequality, see Lemma 5, constants c_2 and c_2' exist such that

$$\mathbb{P}\left(\sum_{n=1}^{N} |Z_n - \mathbb{E}Z_n| > N\eta/2\right) \leq c_2' \exp\left(-\frac{c_2}{4\upsilon^3} N\eta^2\right) + \frac{4c_2\left(\upsilon^3 + \nu_2(Z_n - \mathbb{E}Z_n)^2\right)}{N\eta^2} + 2\exp\left(-\frac{c_2 N\eta^2}{4\nu_2(Z_n - \mathbb{E}Z_n)^2}\right).$$

This shows that Q_1 tends to zero. The proof is now complete. □

Lemma 10. *Assume the assumptions (H1) to (H5), and (H11) to (H14) hold true. For each $N \geq 1$, we have*

$$0 \leq \max_{n,i} W_{n,i}(t;h) \leq S_{n,W}(h) \min\left(1, (\lambda h)^{-1}\right), \quad 1 \leq n \leq N,$$

where $S_{n,W}(h) \geq 1$ is a random variable with the mean and the variance bounded by constants which do not depend on h and n. Moreover, the variables $\{S_{n,W}(h), 1 \leq n \leq N\}$ are independent.

Proof of Lemma 10. By construction, the weights of the NW estimator with a non-negative kernel satisfy

$$0 \leq \min_{1 \leq n \leq N} \min_{1 \leq i \leq M_n} W_{n,i}(t;h) \leq \max_{1 \leq n \leq N} \max_{1 \leq i \leq M_n} W_{n,i}(t;h) \leq 1, \qquad \forall t \in I, h \in \mathcal{H}_N.$$

It thus remains to study more carefully the case where $\lambda h \geq C$ for some constant $C > 0$. Using the fact that the kernel is bounded and bounded away from zero on $[-1,1]$, for each $1 \leq n \leq N$, we have

$$\max_{1 \leq i \leq M_n} W_{n,i}(t;h) \leq \frac{\|K\|_\infty}{\tau} \frac{\mathbb{1}\{S(M_n,t,h) > 0\}}{S(M_n,t,h)},$$

where

$$\tau = \inf_{|t|\leq 1} K(t) \quad \text{ and } \quad S(M_n, t, h) = \sum_{k=1}^{M_n} \mathbb{1}\{|T_{n,k} - t| \leq h\} \leq \overline{c}\lambda.$$

Note that $S(M_n, t, h)$ is an integer-valued variable, non-decreasing as function of h. Conditionally given M_i, the variable $S(M_n, t, h)$ is a Binomial variable with parameters M_n and

$$\mathbb{P}(|T_{n,i} - t| \leq h) = \int_{(t-h)\vee 1}^{(t+h)\wedge 1} g(u)du \geq h \times \inf_{u\in I} g(u) \geq h \times \underline{c}_g.$$

Let us now recall a result of Chao and Strawderman (1972) : if S is a non-degenerate Binomial random variable $B(n, p)$, then

$$\mathbb{E}[(1+S)^{-1}] = \frac{1-q^{n+1}}{(n+1)p}, \qquad \text{where } q = 1-p. \tag{S.16}$$

In our context, we have $S = S(M_n, t, h)$, $n = M_n$ and $p = \mathbb{P}(|T_{n,i} - t| \leq h)$. From (S.16) and Cauchy-Schwarz inequality, we deduce

$$\frac{1}{1+np} \leq \mathbb{E}[(1+S)^{-1}] \leq \frac{1}{(n+1)p}. \tag{S.17}$$

On the other hand, we can write

$$\mathbb{E}\left[\frac{1}{1+S}\right] = \mathbb{P}(S=0) + \mathbb{E}\left[\frac{\mathbb{1}\{S\geq 1\}}{S(1+1/S)}\right] \geq \mathbb{P}(S=0) + \frac{1}{2}\mathbb{E}\left[\frac{\mathbb{1}\{S>0\}}{S}\right],$$

and

$$\mathbb{E}\left[\frac{1}{1+S}\right] = \mathbb{P}(S=0) + \mathbb{E}\left[\frac{\mathbb{1}\{S\geq 1\}}{S(1+1/S)}\right] \leq \mathbb{P}(S=0) + \mathbb{E}\left[\frac{\mathbb{1}\{S>0\}}{S}\right].$$

Using (S.17) and the fact that, in our context, np and λh have the same rate, we deduce that constants $\mathfrak{c}_1, \mathfrak{c}_2 > 0$ exist such that

$$\frac{\mathfrak{c}_1}{np} \leq \mathbb{E}\left[\frac{\mathbb{1}\{S>0\}}{S}\right] \leq \frac{\mathfrak{c}_2}{np},$$

provided $\lambda h \geq C$ and the constant C is sufficiently large. Indeed, the upper bound is obvious. For the lower bound, let us consider $C > \{\underline{c}_g \underline{c}(e-1)\}^{-1}$. We then have

$$\begin{aligned}
\mathbb{E}\left[\frac{\mathbb{1}\{S>0\}}{S}\right] &\geq \mathbb{E}\left[\frac{1}{1+S}\right] - \mathbb{P}(S=0) \\
&= \frac{1}{1+np} - (1-p)^n \geq \frac{1}{1+np} - \exp(-np) \qquad \text{using } \log(1-x) \leq -x, \ \forall x < 1 \\
&\geq \frac{1}{1+np} - \frac{e^{-1}}{np} \qquad \text{using } \exp(-x) \leq \frac{e^{-1}}{x} \ \forall x > 0 \\
&\geq \frac{\mathfrak{c}_1}{np},
\end{aligned}$$

for some constant $\mathfrak{c}_2$. Replacing S by $S(M_n, t, h)$, and using the independence between M and T, we get

$$\frac{c_{1,W}}{\lambda h} \leq \mathbb{E}\left[\frac{\mathbb{1}\{S(M_n,t,h)>0\}}{S(M_n,t,h)}\right] \leq \frac{c_{2,W}}{\lambda h},$$

where $c_{1,W}, c_{2,W}$ are positive constant, depending only on C, $\underline{c}_g$, $\overline{c}_g$, $\underline{c}$, $\overline{c}$. Similarly, it can be shown that a constant c'_W exists, depending only on C, $\underline{c}_g$ and $\underline{c}$, such that

$$\mathbb{E}\left[\frac{\mathbb{1}\{S(M_n,t,h)>0\}}{S^2(M_n,t,h)}\right] \leq \frac{c'_W}{(\lambda h)^2}.$$

The result then follows by defining

$$S_{n,W}(h) = \max\left\{1, \ (\lambda h)(\|K\|_\infty/\tau)\mathbb{1}\{S(M_n,t,h)>0\}S^{-1}(M_n,t,h)\right\},$$

because under our assumptions, the $\{S_{n,W}(h), 1 \leq n \leq N\}$ are clearly independent and we showed that their mean and variance are uniformly bounded with respect to $h \in \mathcal{H}_N$. $\square$

Lemma 11. *Assume that the assumptions (H1) to (H5), (H7) for $p \geq 8$ and (H12) hold. For each $h \in \mathcal{H}_N$, $\{\pi_n(h), n \geq 1\}$ be a sequence of i.i.d. Bernoulli random variables which is independent of $\{X_n, n \in \mathbb{Z}\}$. Then, for any $t \in I$,*

$$\frac{1}{N}\sum_{n=1}^{N} \pi_n(h) X_n^2(t) = \mathbb{E}\left[\pi_1(h) X_1^2(t)\right]\{1 + o_{\mathbb{P}}(1)\} \quad \textit{uniformly with respect to} \;\; h \in \mathcal{H}_N.$$

Proof of Lemma 11. First we decompose the following sum in two terms

$$\frac{1}{N}\sum_{n=1}^{N} \pi_n(h) X_n^2(t) = Z_{1,N}(h;t) + Z_{2,N}(h;t),$$

where

$$Z_{1,N}(h;t) = \frac{1}{N}\sum_{n=1}^{N}\{\pi_n(h) - \mathbb{E}\left[\pi_n(h)\right]\} X_n^2(t) \quad \text{and} \quad Z_{2,N}(h;t) = \frac{1}{N}\sum_{n=1}^{N} \mathbb{E}\left[\pi_n(h)\right] X_n^2(t).$$

Our assumption clearly entails $Z_{2,N}(h;t) = \mathbb{E}\left[\pi_n(h)\right]\mathbb{E}\left[X_n^2(t)\right]\{1 + o_{\mathbb{P}}(1)\}$, uniformly with respect to $h \in \mathcal{H}_N$. To study the uniform convergence of $Z_{1,N}(h;t)$, we first condition on the realization of the sequence $\{X_n, n \geq 1\}$, derive an exponential concentration bound for the weighted sequence of $\pi_n(h)$. Finally, we integrate this bound on a suitable set of realizations of $\{X_n, n \geq 1\}$ with high probability, and provide a bound for the complement of this suitable set.

For the exponential concentration bound for the weighted sequence of $\pi_n(h)$, we use the following general Hoeffding's inequality, a suitable result for our study (see Vershynin, 2018, Theorem 2.6.3). Let $\{U_n, n \geq 1\}$ independent copies of a standardized, sub-gaussian variable U. Let $\boldsymbol{a} = (a_1, \ldots, a_N) \in \mathbb{R}^N$ be a non-random vector. We then have

$$\mathbb{P}\left(\left|\sum_{n=1}^{N} a_n U_n\right| \geq v\right) \leq 2\exp\left[-\frac{Cv^2}{K^2\|\boldsymbol{a}\|_2^2}\right], \qquad \forall v > 0, \tag{S.18}$$

where $C > 0$ is an absolute constant, and $K = \inf\{u > 0 : \mathbb{E}(\exp(U^2/u^2)) \leq 2\}$. When $U_n = \pi_n(h) - \mathbb{E}\left[\pi_n(h)\right]$, we have

$$\begin{aligned}\mathbb{E}(\exp(U^2/u^2)) = \exp(\mathbb{E}\left[\pi_n(h)\right]^2/u^2) \times \{1 - \mathbb{E}\left[\pi_n(h)\right]\} + \exp(\{1 - \mathbb{E}\left[\pi_n(h)\right]\}^2/u^2) \times \mathbb{E}\left[\pi_n(h)\right] \\ \geq 1 + \{\exp(\{1 - \mathbb{E}\left[\pi_n(h)\right]\}^2/u^2) - 1\} \times \mathbb{E}\left[\pi_n(h)\right],\end{aligned}$$

and thus deduce

$$K \leq \frac{1 - \mathbb{E}\left[\pi_n(h)\right]}{\sqrt{\log\left(1 + 1/\mathbb{E}\left[\pi_n(h)\right]\right)}} \leq \frac{1}{\sqrt{\log(2)}}.$$

We apply (S.18) with $U_n = \pi_n(h) - \mathbb{E}\left[\pi_n(h)\right]$ for each $h \in \mathcal{H}_N$. Since $\{\pi_n(h)\}$ and $\{X_n\}$ are independent sequences, and using Boole's (union bound) and (H12) inequality, we deduce

$$\begin{aligned}\mathbb{P}\left(\sup_{h \in \mathcal{H}_N}\left|\sum_{n=1}^{N}\{\pi_n(h) - \mathbb{E}\left[\pi_n(h)\right]\} X_n^2(t)\right| \geq Nv \;\middle|\; (X_1^2(t), \ldots, X_N^2(t)) = \boldsymbol{a}\right) \\ \leq 2\exp\left[c\log(N\lambda) - \frac{CN^2v^2}{K^2\|\boldsymbol{a}\|_2^2}\right].\end{aligned}$$

We next define the event $\mathcal{A} = \{(X_1^2(t), \ldots, X_N^2(t)), \|(X_1^2(t), \ldots, X_N^2(t))\|_2^2 \leq N\mathfrak{c}\}$, for some real number $N\mathfrak{c}$ such that $\mathfrak{c} > \mathbb{E}[X_1^4(t)]$. Then, by (H7) and the Nagaev inequality stated in Lemma 5 we have

$$\begin{aligned}\mathbb{P}\left(\overline{\mathcal{A}}\right) = \mathbb{P}\left(\sum_{n=1}^{N}\{X_n^4(t) - \mathbb{E}[X_1^4(t)]\} > N(\mathfrak{c} - \mathbb{E}[X_1^4(t)])\right) \leq \frac{c_2\left(\vartheta^3 + \|X_1^4(t) - \mathbb{E}[X_1^4(t)]\|_2^2\right)}{N(\mathfrak{c} - \mathbb{E}[X_1^4(t)])^2} \\ + c_2'\exp\left(-\frac{c_2 N(\mathfrak{c} - \mathbb{E}[X_1^4(t)])^2}{\vartheta^3}\right) + 2\exp\left(-\frac{c_2 N(\mathfrak{c} - \mathbb{E}[X_1^4(t)])^2}{\|X_1^4(t) - \mathbb{E}[X_1^4(t)]\|_2^2}\right),\end{aligned}$$

where c_2 and c_2' are two positives constants and ϑ corresponds the dependency coefficient. Thus gathering facts, we obtain

$$\begin{aligned}
&\mathbb{P}\left(\sup_{h\in\mathcal{H}_N}\left|\sum_{n=1}^{N}\{\pi_n(h)-\mathbb{E}\left[\pi_n(h)\right]\}X_n^2(t)\right|\geq Nv\right)\\
&\leq \mathbb{P}\left(\sup_{h\in\mathcal{H}_N}\left|\sum_{n=1}^{N}\{\pi_n(h)-\mathbb{E}\left[\pi_n(h)\right]\}X_n^2(t)\right|\geq Nv \;\middle|\; \|X_1^2(t),\ldots,X_N^2(t)\|_2^2\leq N\mathfrak{c}\right)\mathbb{P}\left(\mathcal{A}\right)+\mathbb{P}\left(\overline{\mathcal{A}}\right)\\
&\qquad\leq 2\exp\left[c\log(N\lambda)-\frac{C\log(2)Nv^2}{\mathfrak{c}}\right]+\frac{c_2\left(\vartheta^3+\|X_1^4(t)-\mathbb{E}[X_1^4(t)]\|_2^2\right)}{N(\mathfrak{c}-\mathbb{E}[X_1^4(t)])^2}\\
&\qquad\qquad+c_2'\exp\left(-\frac{c_2N(\mathfrak{c}-\mathbb{E}[X_1^4(t)])^2}{\vartheta^3}\right)+2\exp\left(-\frac{c_2N(\mathfrak{c}-\mathbb{E}[X_1^4(t)])^2}{\|X_1^4(t)-\mathbb{E}[X_1^4(t)]\|_2^2}\right).
\end{aligned}$$

The proof of Lemma 11 is now complete. □

S.3.2 Mean estimator: risk bounds

Lemma 12. *Under the assumptions (H1) to (H7), (H11) and (H12), we have*

$$\mathbb{E}_{M,T}\left[\{\widehat{\mu}_N(t;h)-\mu(t)\}^2\right]\leq 2R_\mu(t;h)\{1+o(1)\},$$

with $o(1)$ uniform with respect to $h\in\mathcal{H}_N$ and

$$R_\mu(t;h)=L_t^2h^{2H_t}\mathbb{B}(t;h,2H_t)+\sigma^2(t)\mathbb{V}_\mu(t;h)+\mathbb{D}_\mu(t;h)/P_N(t;h).$$

Proof of Lemma 12. Let $\widetilde{\mu}_N(t;h)$ be the infeasible mean estimator

$$\widetilde{\mu}_N(t;h)=\frac{1}{P_N(t;h)}\sum_{n=1}^{N}\pi_n(t;h)X_n(t).$$

Then

$$\mathbb{E}_{M,T}\left[\{\widehat{\mu}_N(t;h)-\mu(t)\}^2\right]\leq 2G_1(t;h)+2G_2(t;h)$$

where

$$G_1(t;h)=\mathbb{E}_{M,T}\left[\{\widehat{\mu}_N(t;h)-\widetilde{\mu}_N(t;h)\}^2\right]\quad\text{and}\quad G_2(t;h)=\mathbb{E}_{M,T}\left[\{\widetilde{\mu}_N(t;h)-\mu(t)\}^2\right].$$

Bound for G_1. We rewrite

$$\widehat{\mu}_N(t;h)-\widetilde{\mu}_N(t;h)=\frac{1}{P_N(t;h)}\sum_{n=1}^{N}\pi_n(t;h)B_n(t;h)+\frac{1}{P_N(t;h)}\sum_{n=1}^{N}\pi_n(t;h)V_n(t;h).$$

Using (C.1), Cauchy-Schwarz inequality, the fact that $\pi_n=\pi_n^2$ and Lemma 7, we have

$$\begin{aligned}
G_1(t;h)&\leq\frac{1}{P_N(t;h)}\sum_{n=1}^{N}\pi_n(t;h)\mathbb{E}_{M,T}B_n^2(t;h)+\frac{1}{P_N^2(t;h)}\sum_{n=1}^{N}\pi_n(t;h)\mathbb{E}_{M,T}V_n^2(t;h)\\
&\leq\{1+o(1)\}\left\{L_t^2h^{2H_t}\mathbb{B}(t;h,2H_t)+\sigma^2(t)\mathbb{V}_\mu(t;h)\right\},
\end{aligned}$$

with $o(1)$ uniform with respect to $h\in\mathcal{H}_N$.

Bound for G_2. Let us first note that

$$\begin{aligned}
P_N(t;h)G_2(t;h)&=\mathbb{E}\left[\{X_0(t)-\mu(t)\}^2\right]\\
&\quad+2\sum_{\ell=1}^{N-1}\mathbb{E}\left[\{X_0(t)-\mu(t)\}\{X_\ell(t)-\mu(t)\}\right]\left\{\sum_{i=1}^{N-\ell}\frac{\pi_i(t;h)\pi_{i+\ell}(t;h)}{P_N(t;h)}\right\}=:\mathbb{D}_\mu(t;h).
\end{aligned}$$

Now we show that $\mathbb{D}_\mu(t;h)$ is finite under the $\mathbb{L}^4_{\mathcal{C}} - m-$approximation. After completing the sequence $\{X_\ell(t), \ell \in \mathbb{Z}\}$, we first have that $\mathbb{D}_\mu(t;h)$ is bounded. Indeed, taking absolute values and using the fact that the autocovariance function is absolutely summable (see Hörmann and Kokoszka, 2010, Lemma 4.1), we get

$$\mathbb{D}_\mu(t;h) \leq \mathbb{E}\left[\{X_0(t)-\mu(t)\}^2\right] + 2\sum_{\ell\geq 1} \left|\mathbb{E}\left[\{X_0(t)-\mu(t)\}\{X_\ell(t)-\mu(t)\}\right]\right| < \infty.$$

Second, the process $\{X_\ell, \ell \in \mathbb{Z}\}$ is $\mathbb{L}^4_{\mathcal{C}} - m-$approximable and the $\mathbb{L}^4_{\mathcal{C}} - \ell-$approximation of X_ℓ is $X_\ell^{(\ell)}$. It is easy to show that for each $\ell \geq 1$, $X_\ell^{(\ell)}$ is independent of X_0, see condition 3) in Definition 3. Therefore, we get

$$\mathbb{E}\left[\{X_0(t)-\mu(t)\}\{X_\ell(t)-\mu(t)\}\right] = \mathbb{E}\left[(X_0(t)-\mu(t))\left(X_\ell(t) - X_\ell^{(\ell)}(t)\right)\right].$$

Then, by Cauchy-Schwartz inequality and by (H7), we get

$$\mathbb{D}_\mu(t;h) \leq \nu_2\left(X_0(t)-\mu(t)\right)\left\{\nu_2\left(X_0(t)-\mu(t)\right) + 2\sum_{\ell\geq 1}\nu_2\left(X_\ell(t)-X_\ell^{(\ell)}(t)\right)\right\} \leq \infty.$$

This conclude the proof. □

Lemma 13. *Assume the assumptions (H1) to (H7), (H12), (H14), (H15) hold true. Let*

$$\widehat{R}_\mu(t;h) = \widehat{L}_t^2 h^{2\widehat{H}_t}\mathbb{B}(t;h,2\widehat{H}_t) + \widehat{\sigma}^2(t)\mathbb{V}_\mu(t;h) + \mathbb{D}_\mu(t;h)/P_N(t;h).$$

Then

$$\sup_{h\in\mathcal{H}_N}\left|\frac{\widehat{R}_\mu(t;h)}{R_\mu(t;h)} - 1\right| = o_{\mathbb{P}}(1).$$

Proof of Lemma 13. Lemma 9 states that $\widehat{\sigma}^2(t) = \sigma^2(t)\{1+o_{\mathbb{P}}(1)\}$, for all $t \in I$. Moreover, $\widehat{\sigma}^2(t)$ does not depend on h. Since by (H15) $\widehat{L}_t$ concentrates to $L_t > 0$, it thus suffices to show that $h^{2\widehat{H}_t} = h^{2H_t}\{1+o_{\mathbb{P}}(1)\}$ uniformly over the range $\mathcal{H}_N$. For any $\epsilon > 0$, we can write

$$\mathbb{P}\left(\sup_{h\in\mathcal{H}_N}\left|h^{2(\widehat{H}_t - H_t)} - 1\right| > \epsilon\right) = Q_3 + Q_4,$$

where

$$Q_3 = \mathbb{P}\left(\sup_{h\in\mathcal{H}_N}\left|h^{2(\widehat{H}_t - H_t)} - 1\right| > \epsilon, |\widehat{H}_t - H_t| \leq \varphi\right),$$
$$Q_4 = \mathbb{P}\left(\sup_{h\in\mathcal{H}_N}\left|h^{2(\widehat{H}_t - H_t)} - 1\right| > \epsilon, |\widehat{H}_t - H_t| > \varphi\right).$$

Without loss of generality we consider $0 < h < 1$, and we define the function $x \to g(x) = h^{2x}$ which is defined and continuously differentiable on $\mathbb{R}$. Then if $|\widehat{H}_t - H_t| \leq \varphi$, we get,

$$\left|h^{2(\widehat{H}_t - H_t)} - 1\right| \leq 2|\log h| h^{-2\varphi}|\widehat{H}_t - H_t|.$$

and then,

$$\left\{\sup_{h\in\mathcal{H}_N}\left|h^{2(\widehat{H}_t - H_t)} - 1\right| > \epsilon, |\widehat{H}_t - H_t| \leq \varphi\right\} \subseteq \left\{(\min\mathcal{H}_N)^{-2\varphi}\left|\log\left((\min\mathcal{H}_N)^{-2\varphi}\right)\right| > \epsilon\right\}.$$

Choosing $\varphi(\lambda) = C_\varphi(\log\lambda)^{-2}$ for some constant $C_\varphi > 0$, thanks to Assumption (H12), which implies $\log(\min\mathcal{H}_N)/\log^2(\lambda) \to 0$, we have

$$(\min\mathcal{H}_N)^{-2\varphi} = \exp\left\{-2C_\varphi\frac{\log(\min\mathcal{H}_N)}{\log^2(\lambda)}\right\} \to 1.$$

Since the continuous function $x \mapsto x\log(x)$, $x > 0$, vanishes at $x = 1$, we deduce that $Q_3 = 0$ for sufficiently large values of λ. On the other hand, (H15) guarantees $Q_4 \to 0$. □

S.3.3 Covariance estimator: risk bound

Lemma 14. *Under the assumptions (H1) to (H6), (H7) for $p \geq 8$, (H11) to (H14), and (H16) we have*

$$\mathbb{E}_{M,T}\left[\{\widehat{\gamma}_{N,\ell}(s,t;h) - \gamma_\ell(s,t)\}^2\right] \leq 2R_\gamma(s,t;h)\{1+o_{\mathbb{P}}(1)\},$$

with $o_{\mathbb{P}}(1)$ uniform with respect to $h \in \mathcal{H}_N$ and

$$\begin{aligned}R_\gamma(s,t;h) = {} & 3\nu_2^2\left(X_{1+\ell}(t)\right) L_s^2 h^{2H_s}\mathbb{B}(s|t;h,2H_s,0) + 3\nu_2^2\left(X_1(s)\right) L_t^2 h^{2H_t}\mathbb{B}(t|s;h,2H_t,\ell)\\ & + 3\left\{\sigma^2(s)\nu_2^2(X_{1+\ell}(t))\mathbb{V}_{\gamma,1}(s,t;h) + \sigma^2(t)\nu_2^2(X_1(s))\mathbb{V}_{\gamma,2}(s,t;h)\right\} + 3\sigma^2(s)\sigma^2(t)\mathbb{V}_\gamma(s,t;h)\\ & + \mathbb{D}(s,t;h)/P_{N,\ell}(s,t;h).\end{aligned}$$

Proof of Lemma 14. Recall that

$$\mathbb{B}(t|s;h,\alpha,\ell) = \sum_{n=1}^{N-\ell}\frac{\pi_n(s;h)\pi_{n+\ell}(t;h)}{P_{N,\ell}(s,t;h)} b_{n+\ell}(t;h,\alpha) \quad \text{with} \quad b_n(t;h,\alpha) = \sum_{i=1}^{M_n}\left|\frac{T_{n,i}-t}{h}\right|^\alpha W_{n,i}(t;h).$$

Recall that $g \otimes f(s,t) := g(s)f(t)$. Let $\widetilde{\gamma}_{N,\ell}(s,t;h)$ be the weighted mean of the unobserved curves $X_n \otimes X_{n+\ell}$, $n = 1\dots N$, *i.e.*,

$$\widetilde{\gamma}_{N,\ell}(s,t;h) = \frac{1}{P_{N,\ell}(s,t;h)}\sum_{n=1}^{N-\ell}\pi_n(s;h)\pi_{n+\ell}(t;h)X_n(s)X_{n+\ell}(t).$$

The quadratic risk of $\widehat{\gamma}_{N,\ell}(s,t;h)$ is then bounded by two terms :

$$\mathbb{E}_{M,T}\left[\{\widehat{\gamma}_{N,\ell}(s,t;h) - \gamma_\ell(s,t)\}^2\right] \leq 2G_1(s,t;h) + 2G_2(s,t;h),$$

where

$$G_1(s,t;h) = \mathbb{E}_{M,T}\left[\{\widehat{\gamma}_{N,\ell}(s,t;h) - \widetilde{\gamma}_{N,\ell}(s,t;h)\}^2\right], \qquad G_2(s,t;h) = \mathbb{E}_{M,T}\left[\{\widetilde{\gamma}_{N,\ell}(s,t;h) - \gamma_\ell(s,t)\}^2\right].$$

We next derive bounds for G_1 and G_2, respectively.

Bound for G_1. We decompose $\widehat{\gamma}_\ell - \widetilde{\gamma}_\ell$ as $\widehat{\gamma}_{N,\ell}(s,t;h) - \widetilde{\gamma}_{N,\ell}(s,t;h) = \mathfrak{a} + \mathfrak{b} + \mathfrak{c}$, where

$$\begin{aligned}\mathfrak{a} &= \frac{1}{P_{N,\ell}(s,t;h)}\sum_{n=1}^{N-\ell}\pi_n(s;h)\pi_{n+\ell}(t;h)X_n \otimes (B_{n+\ell}+V_{n+\ell})(s;(t;h)),\\ \mathfrak{b} &= \frac{1}{P_{N,\ell}(s,t;h)}\sum_{n=1}^{N-\ell}\pi_n(s;h)\pi_{n+\ell}(t;h)(B_n+V_n)\otimes X_{n+\ell}((s;h);t),\\ \mathfrak{c} &= \frac{1}{P_{N,\ell}(s,t;h)}\sum_{n=1}^{N-\ell}\pi_n(s;h)\pi_{n+\ell}(t;h)(B_n+V_n)\otimes(B_{n+\ell}+V_{n+\ell})((s;h);(t;h)).\end{aligned}$$

Since $(\mathfrak{a}+\mathfrak{b}+\mathfrak{c})^2 \leq 3(\mathfrak{a}^2+\mathfrak{b}^2+\mathfrak{c}^2)$, it is sufficient to control the expectations of $\mathfrak{a}^2$, $\mathfrak{b}^2$ and $\mathfrak{c}^2$. Using Lemma 7-(1) and (2), we have

$$\begin{aligned}\mathbb{E}_{M,T}\left(\mathfrak{a}^2\right) &= \mathbb{E}_{M,T}\left(\frac{1}{P_{N,\ell}(s,t;h)}\sum_{i=1}^{N-\ell}\pi_i(s;h)\pi_{i+\ell}(t;h)X_i(s)B_{i+\ell}(t;h)\right)^2\\ &\quad + \frac{1}{P_{N,\ell}^2(s,t;h)}\sum_{i=1}^{N-\ell}\pi_i^2(s;h)\pi_{i+\ell}^2(t;h)\mathbb{E}X_i^2(s)\mathbb{E}_{M,T}V_{i+\ell}^2(t;h),\\ \mathbb{E}_{M,T}\left(\mathfrak{b}^2\right) &= \mathbb{E}_{M,T}\left(\frac{1}{P_{N,\ell}(s,t;h)}\sum_{i=1}^{N-\ell}\pi_i(s;h)\pi_{i+\ell}(t;h)B_i(s;h)X_{i+\ell}(t)\right)^2\\ &\quad + \frac{1}{P_{N,\ell}^2(s,t;h)}\sum_{i=1}^{N-\ell}\pi_i^2(s;h)\pi_{i+\ell}^2(t;h)\mathbb{E}_{M,T}V_i^2(s;h)\mathbb{E}X_{i+\ell}^2(t).\end{aligned}$$

By Cauchy-Schwartz inequality for sums we get,

$$\left(\sum_{i=1}^{N-\ell} \pi_i(s;h)\pi_{i+\ell}(t;h)X_i(s)B_{i+\ell}(t;h)\right)^2 \\ \leq \left(\sum_{i=1}^{N-\ell} \pi_i(s;h)\pi_{i+\ell}(t;h)X_i^2(s)\right)\left(\sum_{i=1}^{N-\ell} \pi_i(s;h)\pi_{i+\ell}(t;h)B_{i+\ell}^2(t;h)\right).$$

Let $\mathfrak{p}_i(h) = \pi_i(s;h)\pi_{i+\ell}(t;h)$. We apply again the idea used in the proof of Lemma 8-(3), that we decompose $\sum_{i=1}^{N-\ell} \mathfrak{p}_i(h)X_i^2(s)$ in $\ell+1$ sub-sums such that in each sub-sum the $\mathfrak{p}_i(h)$ are independent. Applying next Lemma 11 to each sub-sum separately and gathering the facts, we deduce

$$\sum_{i=1}^{N-\ell} \mathfrak{p}_i(h)X_i^2(s) = (N-\ell)\mathbb{E}\left[\pi_1(s;h)\pi_{1+\ell}(t;h)X_1^2(s)\right]\{1+o_{\mathbb{P}}(1)\} = \mathbb{E}\left[P_{N,\ell}(s,t;h)\right]\mathbb{E}\left[X_1(s)^2\right]\{1+o_{\mathbb{P}}(1)\},$$

and the $o_{\mathbb{P}}(1)$ is uniform with respect to $h \in \mathcal{H}_N$. From this and Lemma 8-(3), we next get

$$\left(\frac{1}{P_{N,\ell}(s,t;h)}\sum_{i=1}^{N-\ell} \pi_i(s;h)\pi_{i+\ell}(t;h)X_i(s)B_{i+\ell}(t;h)\right)^2 \\ \leq \frac{\mathbb{E}\left[P_{N,\ell}(s,t;h)\right]\mathbb{E}\left[X_1^2(s)\right]\{1+o_{\mathbb{P}}(1)\}}{P_{N,\ell}^2(s,t;h)}\sum_{i=1}^{N-\ell} \pi_i(s;h)\pi_{i+\ell}(t;h)B_{i+\ell}^2(t;h) \\ = \{1+o_{\mathbb{P}}(1)\}\frac{\mathbb{E}[X_1^2(s)]}{P_{N,\ell}(s,t;h)}\sum_{i=1}^{N-\ell} \pi_i(s;h)\pi_{i+\ell}(t;h)B_{i+\ell}^2(t;h),$$

uniformly with respect to $h \in \mathcal{H}_N$. We thus have the bounds

$$\mathbb{E}_{M,T}\left(\mathfrak{a}^2\right) \leq \frac{\{1+o_{\mathbb{P}}(1)\}\nu_2^2(X_1(s))}{P_{N,\ell}(s,t;h)}\sum_{i=1}^{N-\ell} \pi_i(s;h)\pi_{i+\ell}(t;h)\mathbb{E}_{M,T}B_{i+\ell}^2(t;h) \\ + \frac{1}{P_{N,\ell}^2(s,t;h)}\sum_{i=1}^{N-\ell} \pi_i(s;h)\pi_{i+\ell}(t;h)\mathbb{E}X_i(s)^2\mathbb{E}_{M,T}V_{i+\ell}^2(t;h),$$

$$\mathbb{E}_{M,T}\left(\mathfrak{b}^2\right) \leq \frac{\{1+o_{\mathbb{P}}(1)\}\nu_2^2(X_{1+\ell}(t))}{P_{N,\ell}(s,t;h)}\sum_{i=1}^{N-\ell} \pi_i(s;h)\pi_{i+\ell}(t;h)\mathbb{E}_{M,T}B_i^2(s;h) \\ + \frac{1}{P_{N,\ell}(s,t;h)^2}\sum_{i=1}^{N-\ell} \pi_i(s;h)\pi_{i+\ell}(t;h)\mathbb{E}_{M,T}V_i^2(s;h)\mathbb{E}X_{i+\ell}^2(t).$$

Finally, by Lemma 7-(3) and (4),

$$\mathbb{E}_{M,T}\left(\mathfrak{a}^2\right) \leq \frac{\{1+o_{\mathbb{P}}(1)\}L_t^2h^{2H_t}\nu_2^2(X_1(s))}{P_{N,\ell}(s,t;h)}\sum_{i=1}^{N-\ell} \pi_i(s;h)\pi_{i+\ell}(t;h)b_{i+\ell}(t,h,2H_t) \\ + \frac{\{1+o(1)\}\sigma^2(t)\nu_2^2(X_1(s))}{P_{N,\ell}(s,t;h)^2}\sum_{i=1}^{N-\ell} \pi_i(s;h)\pi_{i+\ell}(t;h)\max_{1\leq k\leq M_{i+\ell}}|W_{i+\ell,k}(t;h)| \\ = \{1+o_{\mathbb{P}}(1)\}\left[\nu_2\left(X_1(s)\right)^2 L_t^2h^{2H_t}\mathbb{B}(t|s;h,2H_t,\ell) + \sigma^2(t)\nu_2^2(X_1(s))\mathbb{V}_{\gamma,2}(s,t;h)\right],$$

$$\mathbb{E}_{M,T}\left(\mathfrak{b}^2\right) \leq \frac{\{1+o_{\mathbb{P}}(1)\}L_s^2h^{2H_s}\nu_2^2(X_{1+\ell}(t))}{P_{N,\ell}(s,t;h)}\sum_{i=1}^{N-\ell} \pi_i(s;h)\pi_{i+\ell}(t;h)b_i(s,h,2H_s) \\ + \frac{\{1+o(1)\}\sigma^2(s)\nu_2^2(X_{1+\ell}(t))}{P_{N,\ell}^2(s,t;h)}\sum_{i=1}^{N-\ell} \pi_i(s;h)\pi_{i+\ell}(t;h)c_i(s;h)\max_{1\leq k\leq M_i}|W_{i,k}(s;h)| \\ = \{1+o_{\mathbb{P}}(1)\}\left[\nu_2^2\left(X_{1+\ell}(t)\right)L_s^2h^{2H_s}\mathbb{B}(s|t;h,2H_s,0) + \sigma^2(s)\nu_2^2(X_{1+\ell}(t))\mathbb{V}_{\gamma,1}(s,t;h)\right],$$

and the $o_{\mathbb{P}}(1)$ factors are uniform with respect to $h \in \mathcal{H}_N$.

To derive a bound for $\mathbb{E}_{M,T}\left[\mathfrak{c}^2\right]$, we decompose this expectation as follows :

$$\mathbb{E}_{M,T}\left(\mathfrak{c}^2\right) = \mathbb{E}_{M,T}\left(\frac{1}{P_{N,\ell}(s,t;h)}\sum_{i=1}^{N-\ell}\pi_i(s;h)\pi_{i+\ell}(t;h)B_i(s;h)B_{i+\ell}(t;h)\right)^2 \tag{S.19}$$

$$+\frac{1}{P^2_{N,\ell}(s,t;h)}\sum_{i=1}^{N-\ell}\pi_i^2(s;h)\pi_{i+\ell}^2(t;h)\mathbb{E}_{M,T}V_i^2(s;h)\mathbb{E}_{M,T}V_{i+\ell}^2(t;h) \tag{S.20}$$

$$+\frac{1}{P^2_{N,\ell}(s,t;h)}\sum_{i=1}^{N-\ell}\pi_i^2(s;h)\pi_{i+\ell}^2(t;h)\mathbb{E}_{M,T}B_i^2(s;h)\mathbb{E}_{M,T}V_{i+\ell}^2(t;h) \tag{S.21}$$

$$+\frac{1}{P^2_{N,\ell}(s,t;h)}\sum_{i=1}^{N-\ell}\pi_i^2(s;h)\pi_{i+\ell}^2(t;h)\mathbb{E}_{M,T}V_i^2(s;h)\mathbb{E}_{M,T}B_{i+\ell}^2(t;h) \tag{S.22}$$

$$+\frac{2}{P^2_{N,\ell}(s,t;h)}\sum_{i=1}^{N-2\ell}\mathbb{E}_{M,T}(\pi_i\pi_{i+\ell}B_iV_{i+\ell})\otimes(\pi_{i+\ell}\pi_{i+2\ell}B_{i+2\ell}V_{i+\ell})((s;h);(t;h)). \tag{S.23}$$

Lemma 7 ensures that the sum of the terms (S.21), (S.22) and (S.23) are negligible compared to $\mathbb{E}_{M,T}(\mathfrak{a}^2+\mathfrak{b}^2)$, uniformly with respect to $h \in \mathcal{H}_N$.

Applying again Lemma 7, the term (S.20) is bounded by,

$$\text{(S.20)} \leq \{1+o(1)\}\sigma^2(s)\sigma^2(t)\mathbb{V}_\gamma(s,t;h).$$

Finally, for the term (S.19), we first note that by Jensen's inequality and the fact that $x^2y^2 \leq (x^4+y^4)/2$,

$$\begin{aligned}\text{(S.19)} &\leq \frac{1}{P_{N,\ell}(s,t;h)}\sum_{i=1}^{N-\ell}\pi_i(s;h)\pi_{i+\ell}(t;h)\mathbb{E}_{M,T}\left[B_i^2(s;h)B_{i+\ell}^2(t;h)\right]\\ &\leq \frac{1}{2P_{N,\ell}(s,t;h)}\sum_{i=1}^{N-\ell}\pi_i(s;h)\pi_{i+\ell}(t;h)\left\{\mathbb{E}_{M,T}[B_i^4(s;h)]+\mathbb{E}_{M,T}[B_{i+\ell}^4(t;h)]\right\}.\end{aligned}$$

To bound the 4th order moment of the bias term, we use condition (5), that is a constant $\mathfrak{C}>0$ exists such that

$$\mathbb{E}(X(u)-X(v))^4 \leq \mathfrak{C}\left[\mathbb{E}(X(u)-X(v))^2\right]^2, \qquad \forall u,v \in I.$$

More precisely, by Jensen's inequality we have

$$B_n^4(t;h) \leq \sum_{i=1}^{M_n} W_{n,i}(t;h)\left\{X_n(T_{n,i})-X_n(t)\right\}^4.$$

Taking expectation both sides, and applying conditions (5) and (4) we can write

$$\begin{aligned}\mathbb{E}_{M,T}\left[B_n^4(t;h)\right] &\leq \mathfrak{C}\sum_{i=1}^{M_n}W_{n,i}(t;h)\mathbb{E}^2_{M,T}\left[\left\{X_n(T_{n,i})-X_n(t)\right\}^2\right]\\ &\leq \mathfrak{C}L_t^4\sum_{i=1}^{M_n}W_{n,i}(t;h)|T_{n,i}-t|^{4H_t}\times\left\{1+h^{2\beta_0}S_0^2/L_t^2\right\}^2\\ &= \mathfrak{C}L_t^4h^{4H_t}\mathfrak{b}_n(t;h,4H_t)\times\left\{1+h^{2\beta_0}S_0^2/L_t^2\right\}^2.\end{aligned}$$

We deduce

$$\text{(S.19)} \leq \mathfrak{C}\max(L_t^4,L_s^4)h^{4\min(H_t,H_s)}\left\{\mathbb{B}(s|t;h,4H_s,0)+\mathbb{B}(t|s;h,4H_t,\ell)\right\}\times\{1+o(1)\},$$

and the $o(1)$ factor is uniform with respect to $h \in \mathcal{H}_N$. This ensures that the term (S.19) is also negligible in comparison of $\mathbb{E}_{M,T}(\mathfrak{a}^2+\mathfrak{b}^2)$ uniformly with respect to $h \in \mathcal{H}_N$.

Putting the three bounds together, we get

$$\begin{aligned}G_1(s,t;h)/3 \leq\ & \nu_2^2\left(X_{1+\ell}(t)\right) L_s^2 h^{2H_s}\mathbb{B}(s|t;h,2H_s,0) + \nu_2^2\left(X_1(s)\right) L_t^2 h^{2H_t}\mathbb{B}(t|s;h,2H_t,\ell)\\ &+ \nu_2^2(X_{1+\ell}(t))\sigma^2(s)\mathbb{V}_{\gamma,1}(s,t;h) + \nu_2^2(X_1(s))\sigma^2(t)\mathbb{V}_{\gamma,2}(s,t;h)\\ &+ \sigma^2(s)\sigma^2(t)\mathbb{V}_\gamma(s,t;h) + \text{uniformly negligible terms.}\end{aligned}$$

Bound for G_2. For each $k\in\{1,\ldots,N-\ell-1\}$, we define the positive real number $p_k(s,t;h)\in[0,1]$,

$$p_k(s,t;h) = \sum_{i=1}^{N-k-\ell}\frac{\pi_i(s;h)\pi_{i+k}(s;h)\pi_{i+\ell}(t;h)\pi_{i+\ell+k}(t;h)}{P_{N,\ell}(s,t;h)} \leq 1.$$

Then we rewrite G_2 as,

$$\begin{aligned}P_{N,\ell}(s,t;h)G_2(s,t;h) =\ & \mathbb{E}(X_0\otimes X_\ell-\gamma_\ell)^2(s,t)\\ &+ 2\sum_{k=1}^{N-\ell-1} p_k(s,t;h)\mathbb{E}(X_0\otimes X_\ell-\gamma_\ell)(X_k\otimes X_{k+\ell}-\gamma_\ell)(s,t) =: \mathbb{D}(s,t;h).\end{aligned}$$

Now we show that $\mathbb{D}(s,t;h)$ is finite under the $\mathbb{L}^4_{\mathcal{C}}-m-$approximation. First, if we complete the sequence $\{X_k(s)X_{k+\ell}(t),\ k\in\mathbb{Z}\}$, $\mathbb{D}(s,t;h)$ can be bounded by the convergent series of absolute values of the terms in the long-run variance of the time series $\{X_k(s)X_{k+\ell}(t),\ k\in\mathbb{Z}\}$. More precisely,

$$\mathbb{D}(s,t,h) \leq \nu_2^2((X_0\otimes X_\ell-\gamma_\ell)(s,t)) + 2\sum_{k\geq 1}\left|\mathbb{E}(X_0\otimes X_\ell-\gamma_\ell)(X_k\otimes X_{k+\ell}-\gamma_\ell)(s,t)\right| \tag{S.24}$$

Moreover, according to Lemma 2, the process $\{X_k\otimes X_{k+\ell}, k\in\mathbb{Z}\}$ is $\mathbb{L}^2_{\mathcal{C}}-m-$approximable, and the $\mathbb{L}^2_{\mathcal{C}}-(k-\ell)-$approximation of $X_k\otimes X_{k+\ell}$ is $X_k^{(k-\ell)}\otimes X_{k+\ell}^{(k)}$. It is easy to show that the variables $X_0\otimes X_\ell$ and $X_k^{(k-\ell)}\otimes X_{k+\ell}^{(k)}$ are independent. Therefore, we get

$$\begin{aligned}\left|\mathbb{E}(X_0\otimes X_\ell-\gamma_\ell)(X_k\otimes X_{k+\ell}-\gamma_\ell)(s,t)\right| &= \left|\mathbb{E}(X_0\otimes X_\ell-\gamma_\ell)\left(X_k\otimes X_{k+\ell}-X_k^{(k-\ell)}\otimes X_{k+\ell}^{(k)}\right)(s,t)\right|\\ &\leq \nu_2((X_0\otimes X_\ell-\gamma_\ell)(s,t))\nu_2\left(\left(X_k\otimes X_{k+\ell}-X_k^{(k-\ell)}\otimes X_{k+\ell}^{(k)}\right)(s,t)\right).\end{aligned}$$

We next use this in (S.24) to get

$$\begin{aligned}\mathbb{D}(s,t;h) \leq\ & \nu_2^2((X_0\otimes X_\ell-\gamma_\ell)(s,t))\\ &+ 2\nu_2((X_0\otimes X_\ell-\gamma_\ell)(s,t))\sum_{k=1}^{N-\ell-1}\nu_2\left(\left(X_k\otimes X_{k+\ell}-X_k^{(k-\ell)}\otimes X_{k+\ell}^{(k)}\right)(s,t)\right) \leq \infty.\end{aligned}$$

This concludes the proof. □

S.4 Details on numerical study

In this section we give more details on the simulation setting of the Section 5 and report the results of the FTS Model 1 and FTS Model 3 simulation setups.

S.4.1 Details on some quantities used in the simulations

Simple estimators of $\Sigma(t)$ and $\mathbb{S}_\mu(t)$ using the $R=400$ replications of the generated data from the FTS Model 2 (or FTS Model 3) are

$$\widehat{\Sigma}(t) = \frac{1}{R}\sum_{r=1}^{R}\widehat{\Sigma}_{r,N}(t) \quad\text{and}\quad \widehat{\mathbb{S}}_\mu(t) = \frac{1}{R}\sum_{r=1}^{R}\widehat{\mathbb{S}}_{\mu,r,N}(t), \tag{S.25}$$

where $\widehat{\Sigma}_{r,N}(t)$ and $\widehat{\mathbb{S}}_{\mu,r,N}(t)$ are the $r-$th replication of

$$\widehat{\Sigma}_N(t) = \frac{\widehat{\sigma}^2(t)}{P_N(t;h_N)} \sum_{n=1}^{N} \pi_n(t;h_N) \left\{ \sum_{i=1}^{M_n} W_{n,i}^2(t;h_N) \right\},$$

$$\text{and } \widehat{\mathbb{S}}_{\mu,N}(t) = \frac{1}{\sqrt{P_N(t;h_N)}} \sum_{n=1}^{N} \pi_n(t;h_N)\{X_n(t;h_N) - \mu(t)\},$$

respectively.

The approximation $\widetilde{\gamma}_1(s,t)$ of $\gamma_1(s,t) = \mathbb{E}[X_n(s)X_{n+1}(t)]$ in Figure 1 was obtained as the lag-1 empirical autocovariance function calculated from a very large sample generated from FTS Model 2 with $\mu \equiv 0$. More precisely, we generate $R_\gamma = 30$ replications of the functional time series $\{X_n, n = 1, \dots, N\}$ with a large $N = 5000$ and a burn-in period of 500 curves to remove the initialization effect. In the simulation framework it is then possible to accurately approximate $\gamma_1(s,t)$ by

$$\widetilde{\gamma}_1(s,t) = \frac{1}{R_\gamma} \sum_{r=1}^{R_\gamma} \left\{ \frac{1}{N-1} \sum_{n=1}^{N-1} X_{r,n}(s)X_{r,n+1}(t) \right\} \qquad \forall (s,t) \in G \times G,$$

where G is a fine grid of design points, and $X_{r,n}$ denotes the r-th replication of the curve X_n.

S.4.2 Simulation setting : FTS Model 3

The FTS Model 3 setup is based on the Individual Household Electricity Consumption dataset from the UC Irvine Machine Learning Repository (Hebrail and Berard, 2012). It contains various measurements of electricity consumption in a household near Paris, with a sampling rate of one minute from December 2006 to November 2010. The data of interest here are the daily voltages curves, considering only the days without missing values in the measurements, see Figure S.1. The extracted dataset contains 1358 voltage curves with a uniform common design of 1440 points, normalized so that $I = (0, 1]$.

We use this real dataset to build a data generation setup and simulate functional time series with patterns similar to the voltage curves. We use the FAR(1) equation

$$X_n(u) = \mu(u) + \int_0^1 \psi(u,s)(X_{n-1}(s) - \mu(s))ds + L_t\xi_n(u), \tag{S.26}$$

with the mean function and the kernel of the autoregressive operator estimated from the real curves. The Fourier expansion was used to estimate the mean and kernel functions. More precisely, we consider

$$\mu(t) = \beta_0 + \sum_{k=1}^{10} \beta_k \zeta_k(t), \qquad t \in (0,1], \tag{S.27}$$

$$\psi(s,t) = \sum_{k=0}^{4} \sum_{l=0}^{4} \theta_{k,l} \zeta_k(s) \zeta_l(t), \qquad s,t \in (0,1], \tag{S.28}$$

where $\{\zeta_k, k \geq 0\} = \{1, \sqrt{2}\cos(2\pi t), \sqrt{2}\sin(2\pi t), \sqrt{2}\cos(4\pi t), \sqrt{2}\sin(4\pi t), ...\}$ is the Fourier orthonormal basis on the unit interval. The β and θ coefficients are obtained by LASSO regression using the `R` package `glmnet` (Friedman et al., 2010). For the mean function, the coefficients β are estimated using the 1440 values of the empirical mean of the 1358 curves and t on the regular grid. Similarly, for the integral operator kernel function $\psi(\cdot,\cdot)$, the empirical covariance and lag-1 autocovariance functions are used to estimate the θ coefficients using a representation that we explain below; see (S.29). Figure S.1 shows the estimates of the mean function and Figure S.2 shows the level plot of the kernel function $\psi(\cdot,\cdot)$.

We now explain how we derive the representation used to build an estimate of the integral operator kernel function $\psi(\cdot,\cdot)$. Let $t_1, \dots, t_{1440}$ be the common design points for this data set. For all $\forall\, s,t \in (0,1]$ and $\ell \geq 0$, let

$$C_\ell(s,t) = \mathbb{E}\{(X_n(s) - \mu(s))(X_{n+\ell}(t) - \mu(t))\},$$

denote the lag-ℓ autocovariance function of $\{X_n\}$. By Fubini's Theorem and (S.26), we get

$$C_1(s,t) = \int_0^1 \psi(t,u)C_0(s,u)du.$$

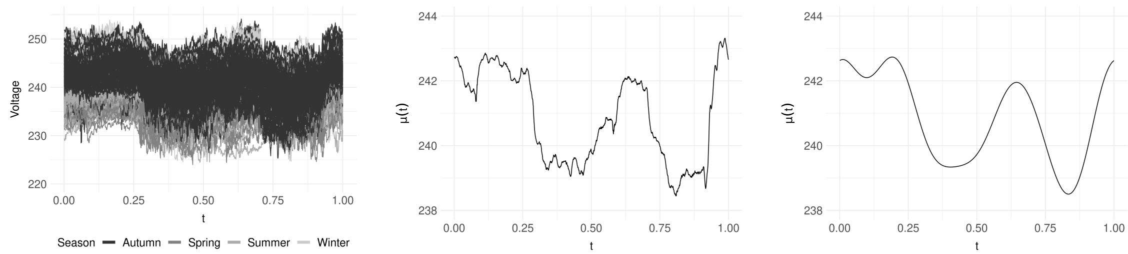

Figure S.1: Curves and mean functions of the daily voltage curves with no missing. **Right:** The raw daily voltage curves. **Middle:** Empirical mean function of the daily voltage curves. **Right:** Smooth mean function of the daily voltage curves obtained from the model (S.27).

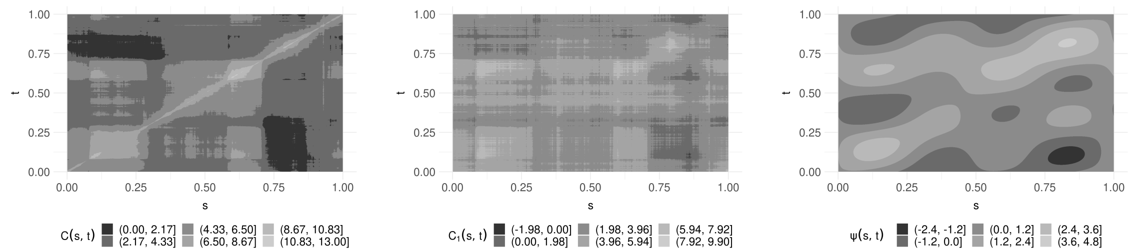

Figure S.2: Level plots of covariance and autocovariance functions of the daily voltage curves with no missing values and the estimated FAR kernel function. **Left:** Empirical covariance. **Middle:** Empirical lag-1 autocovariance. **Right:** Estimated FAR kernel from the models (26) and (S.28).

Then, with the function $\psi(\cdot,\cdot)$ from (S.28), we get

$$C_1(s,t) = \sum_{k=0}^{4}\sum_{l=0}^{4} \theta_{k,l}\zeta_k(t)Z_l(s), \qquad \text{where} \qquad Z_l(s) = \int_0^1 C_0(s,u)\zeta_l(u)du.$$

The values $Z_l(\cdot)$ can be simply approximated by $\widehat{Z}_l(\cdot)$ using the Riemann sums approximation and the empirical covariance function $\widehat{C}_0(s,u)$ calculated at $t_1,\ldots,t_{1440}$ (see Figure S.2). Let

$$\mathbf{Z} = \begin{bmatrix} \widehat{Z}_0(t_1) & \cdots & \widehat{Z}_4(t_1) \\ \widehat{Z}_0(t_2) & \cdots & \widehat{Z}_4(t_2) \\ \vdots & \ddots & \vdots \\ \widehat{Z}_0(t_{1440}) & \cdots & \widehat{Z}_4(t_{1440}) \end{bmatrix}, \qquad \boldsymbol{\zeta} = \begin{bmatrix} \zeta_0(t_1) & \cdots & \zeta_4(t_1) \\ \zeta_0(t_2) & \cdots & \zeta_4(t_2) \\ \vdots & \ddots & \vdots \\ \zeta_0(t_{1440}) & \cdots & \zeta_4(t_{1440}) \end{bmatrix},$$

and $\boldsymbol{\Theta} = (\theta_{k,l})_{0\leq k,l\leq 4}$ the $5\times 5-$matrix of coefficient to be determined using the real data set, and let

$$\mathbb{C}_1(\boldsymbol{\Theta}) = \boldsymbol{\zeta}\boldsymbol{\Theta}\mathbf{Z}^\top, \tag{S.29}$$

be the lag-1 autocovariance function we consider, computed at the common design pairs of points. The elements of $\boldsymbol{\Theta}$ are chosen such that $\mathbb{C}_1(\boldsymbol{\Theta})$ is the closest, in terms of Frobenius norm, to the empirical lag-1 autocovariance computed at the same common design pairs.

S.4.3 Additional simulation results on local regularity estimation

Our estimation approach for estimating H_t and L_t^2 depends on two tuning parameters : the presmoothing bandwidth used in (9) and the window length Δ used in (8). The following paragraphs explain how we tune these parameters and give the simulation results of the local regularity estimates of FTS Model 1 and FTS Model 3.

Choice of the presmoothing bandwidth The presmoothing step consists of smoothing each curve of the time series individually using a bandwidth parameter. To reduce the computation time, we use the median of the bandwidths selected by cross-validation on the last 30 curves of the series as the smoothing parameter of all curves. Recall that given the sample points $\{(Y_{n,1}, T_{n,1}), \dots, (Y_{n,M_n}, T_{n,M_n})\}$ of a curve X_n, the cross-validation optimal bandwidth for the presmoothing estimator (9) is defined as

$$h^* \in \arg\min_h \sum_{i=1}^{M_n} \left\{ Y_{n,i} - \widetilde{X}_n^{(-i)}(T_{n,i}) \right\}^2,$$

where $\widetilde{X}_n^{(-i)}(T_{n,i}) = \widetilde{X}_n^{(-i)}(T_{n,i}; h)$ denotes the estimator (9) that is computed without the observations corresponding to the $i-$th design point, and the bandwidth h.

Choice of Δ The choice of Δ is a crucial point for the local regularity estimation, and extensive empirical experiences have been devoted to the investigation of how to fix it. The study of the choice of Δ was carried out independently of the three FTS Models introduced in the main manuscript. It was based on a zero-mean FAR(1) where the innovation process is the MfBm with a Hurst logistic function. The autoregressive operator of the process is an integral operator with a smooth kernel function chosen such that the conditions of the Example 2 hold. The main idea of the investigation is to use 200 replications of data generated from the FAR(1) with N curves, each with λ mean points, to compute the local exponent $\widehat{H}_t$ from (11) and compare it over a chosen risk with the estimate of $\widetilde{H}_t$ from (8) for a given Δ. So, given a grid of Δ candidates, the best one is the one that minimises the following relative risk. Let $t_1, t_2, t_3 \in J \subset I$ such that $t_3 - t_1 = \Delta$ and $t_2 = t = (t_1 + t_3)/2$, then

$$\Delta^* \in \arg\min_\Delta \frac{1}{200} \sum_{r=1}^{200} \frac{(\widehat{\theta}_r(t_1,t_2) - \widetilde{\theta}_r(t_1,t_2))^2 + (\widehat{\theta}_r(t_1,t_3) - \widetilde{\theta}_r(t_1,t_3))^2}{\widetilde{\theta}_r(t_1,t_2)^2 + \widetilde{\theta}_r(t_1,t_3)^2}, \tag{S.30}$$

where r denotes the r-th replication of the data set of N curves and λ mean points per curve, $\widehat{\theta}$ is as defined in (10), and since we are in a simulation framework, it is possible to get the true X_n and estimate

$$\widetilde{\theta}(u,v) = N^{-1} \sum_{n=1}^{N} (X_n(v) - X_n(u))^2, \qquad u, v \in \{t_1, t_2, t_3\}.$$

The investigation is carried out by testing Δ values for $I = (0,1]$. Namely, for each Δ in an equidistant grid of 30 values between 0.01 and 0.2, and for each $t \in \{0.2, 0.4, 0.7, 0.8\}$, and using 200 replications of data generated from the setups $(N, \lambda) \in \{100, 200, 300, 400\} \times \{\lambda_1 = 30,\ \lambda_{i+1} = \lambda_i + 15,\ 2 \leq i \leq 30\}$, we estimate Δ^* according to (S.30). The result is that any $\Delta^* \in [0.1, 0.2]$ gives a reasonably small risk and all Δ values within this interval give relatively the same risk as defined in (S.30). Moreover, if $\Delta^* \leq 0.1$ the risk increases slowly. Therefore, we propose to chose $\gamma = 1/3$ and set $\Delta^* = \min\{\exp(-\log(\widehat{\lambda})^{1/3})\}, 0.2\}$.

Additional simulation results for the regularity parameters Here we present the simulation results in the setup of FTS Model 1 and FTS Model 3. Figure S.3 and Figure S.4 show the boxplots of $\widehat{H}_t$ and $\widehat{L}_t^2$ defined in (11) for the four pairs (N, λ) at four points $t \in I = (0,1]$ for Model 1 and FTS Model 3 respectively. The results are similar to those of FTS Model 2. Indeed, the bias of the regularity parameters estimates decreases as λ increases, and the boxplot are more concentrated as N increases. Overall, the local regularity estimators show good finite sample performance.

S.4.4 Additional results on mean function estimation

This section presents the results of the mean function estimation using data generated according to the FTS Model 3 setup. Recall that the FTS Model 1 setup is similar to FTS Model 2, the results of which are already presented in the main paper, and that it contains twelve setups (H_t, N, λ), so the results associated with FTS Model 1 are not reported. Figure S.5 presents the average of the risk function $\widehat{R}_\mu(t;h)$ over 400 independent functional time series generated according to FTS Model 3, with four setups (N, λ). As for FTS Model 2, the plots provide evidence that $h \to R_\mu(t;h)$ is a convex function which converges to zero as N and λ become larger.

Table S.1 presents the bias and standard deviation of the estimates of $\widehat{\mu}_N^*(t) = \widehat{\mu}_N(t; h_\mu^*)$ obtained for functional time series generated according to FTS Model 3. As expected the bias and the variance

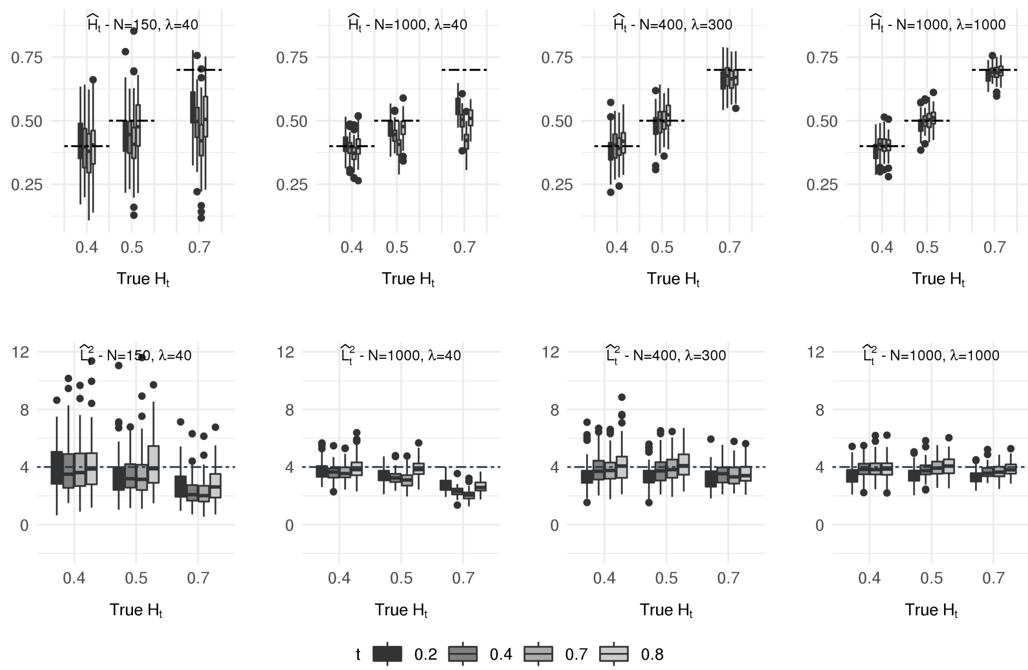

Figure S.3: Boxplots of $R = 100$ pointwise estimates of $\widehat{H}_t$ and $\widehat{L}_t^2$, for $t \in \{0.2, 0.4, 0.7, 0.8\}$ and four pairs (N, λ), in FTS Model 1. The dashed lines indicate the true values of H_t and L_t^2.

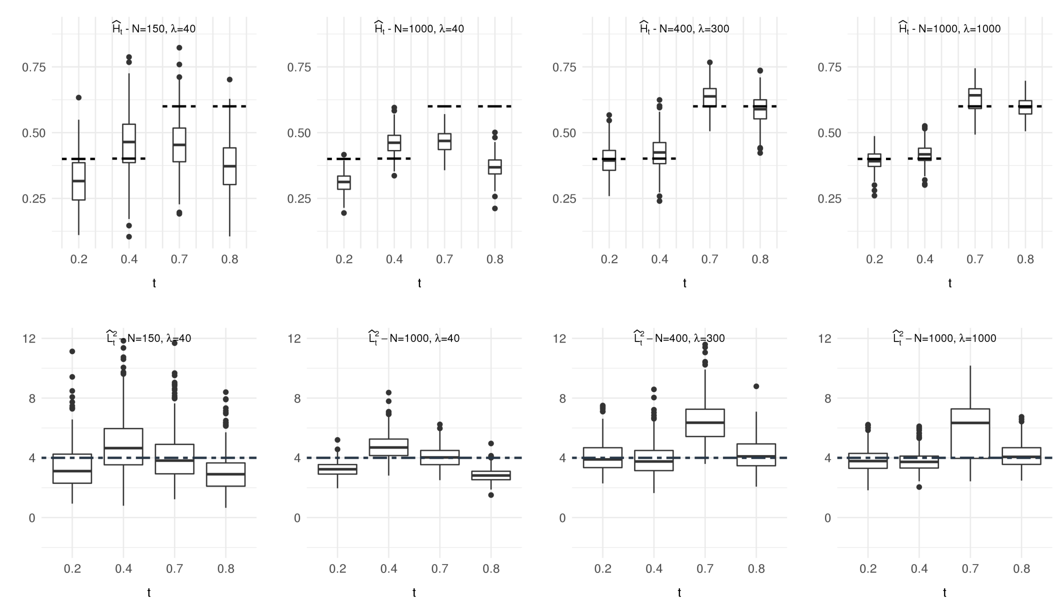

Figure S.4: Boxplots of $R = 400$ pointwise estimates of $\widehat{H}_t$ and $\widehat{L}_t^2$, for $t \in \{0.2, 0.4, 0.7, 0.8\}$ and four pairs (N, λ), in FTS Model 3. The dashed lines indicate the true values of H_t and L_t^2.

decreases as $N \to \infty$ and as $\lambda \to \infty$. However, larger t also means larger $\text{Var}(X_t)$ (see the Figure S.6). We next study the asymptotic distribution of $\widehat{\mu}_N^*(t)$. The $Q - Q$ plots Figure S.7 show that, as stated by Theorem 4, the standard normal distribution is an accurate approximation of the distribution of $\sqrt{P_N(t; h_N)/\{\mathbb{S}_\mu(t) + \Sigma(t)\}}\,\{\widehat{\mu}_N(t; h_N) - \mu(t)\}$.

We end this section on empirical evidence for the mean function estimation by a comparison with the procedure of Rubín and Panaretos (2020), procedure refer to as `RP20`, in the context of the FTS

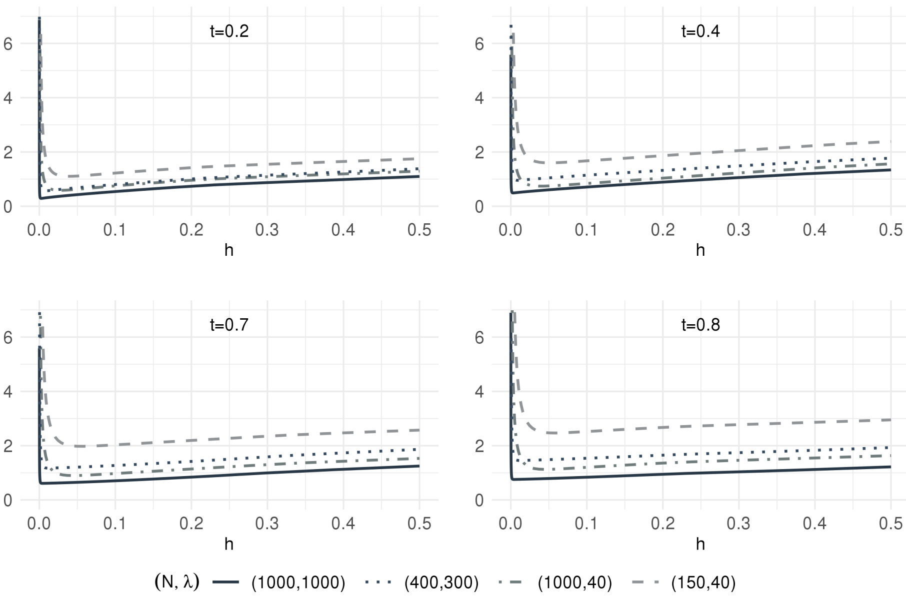

Figure S.5: Empirical average of the risk function $\widehat{R}_{\mu}(t;h)$ at $t \in \{0.2, 0.4, 0.7, 0.8\}$ over 400 independent functional time series generated according to FTS Model 3, with four setups (N, λ).

		$t = 0.2$		$t = 0.4$		$t = 0.7$		$t = 0.8$	
N	λ	Bias	Sd	Bias	Sd	Bias	Sd	Bias	Sd
150	40	-0.0714	0.2815	0.0515	0.3681	-0.0799	0.4209	0.1585	0.4598
1000	40	-0.0401	0.1085	0.0275	0.1354	-0.0580	0.1544	0.1113	0.1706
400	300	-0.0158	0.1758	-0.0216	0.2295	-0.0327	0.2595	-0.0288	0.2852
1000	1000	-0.0016	0.0937	-0.0039	0.1206	-0.0008	0.1340	0.0018	0.1477

Table S.1: Bias and standard deviation (Sd) of the mean function estimates obtained from 400 independent functional time series generated according to FTS Model 3.

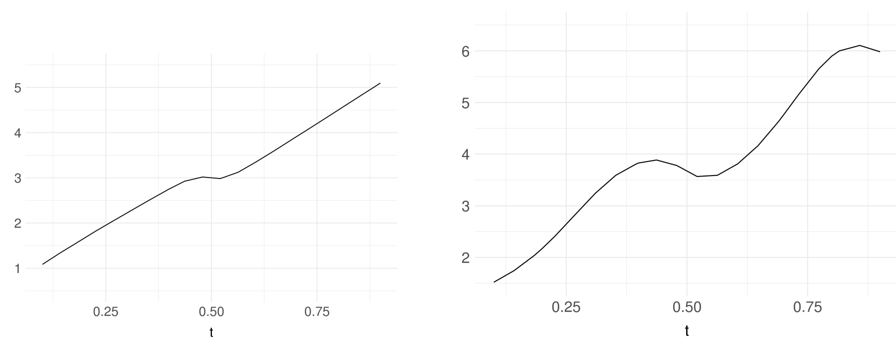

Figure S.6: Estimates of the variance function of $\mathrm{Var}(X_n)$. **Left:** Variance function in FTS Model 2 process. **Right:** Variance function in FTS Model 3 process.

Model 3. We present in Figure S.8 the boxplots of the selected bandwidths according to `RP20`'s global approach and to our local approach. The selected bandwidths have comparable magnitudes in almost all setups (N, λ). As expected given the increasing shape of the function H, our local bandwidths are smaller for t in the first half of I and increase when t is closer to 1. Table S.2 presents the ratio of the Monte-Carlo estimates of the Mean Square Error (MSE) of our mean function estimator and the `RP20` locally linear estimator. Although the ratio is close to 1, our estimator shows slightly better performance (ratio smaller than 1) in most of the setups (N, λ).

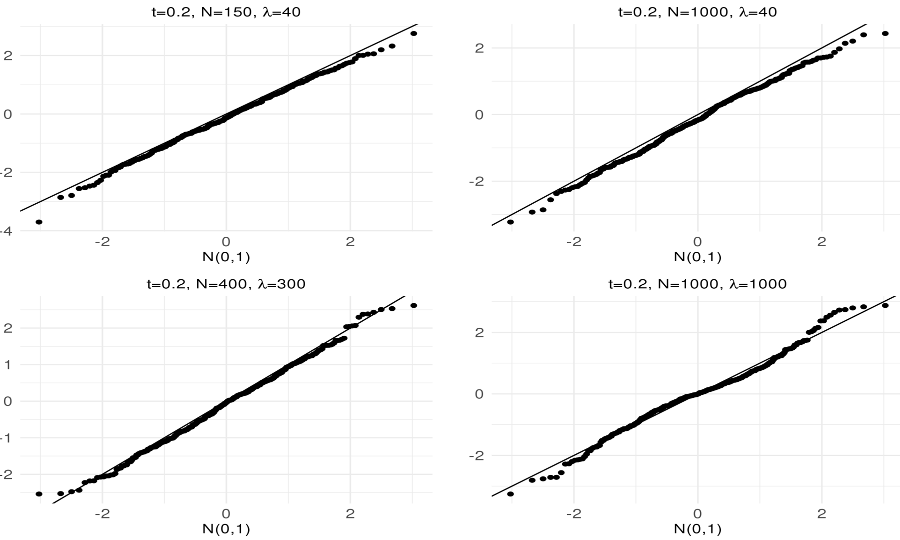

Figure S.7: Normal $Q-Q$ plots of $\sqrt{P_N(t;h_N)}\,(\widehat{\mu}_N(t;h_N)-\mu(t))\,/\sqrt{\widehat{\mathbb{S}}_\mu(t)+\widehat{\Sigma}(t)}$ at $t=0.2$, with $h_N=\{h_\mu^*\}^{1.1}$ and $\widehat{\mathbb{S}}_\mu(t)+\widehat{\Sigma}(t)$ computed according to (S.25). Results obtained with 400 independent time series generated according to FTS Model 3.

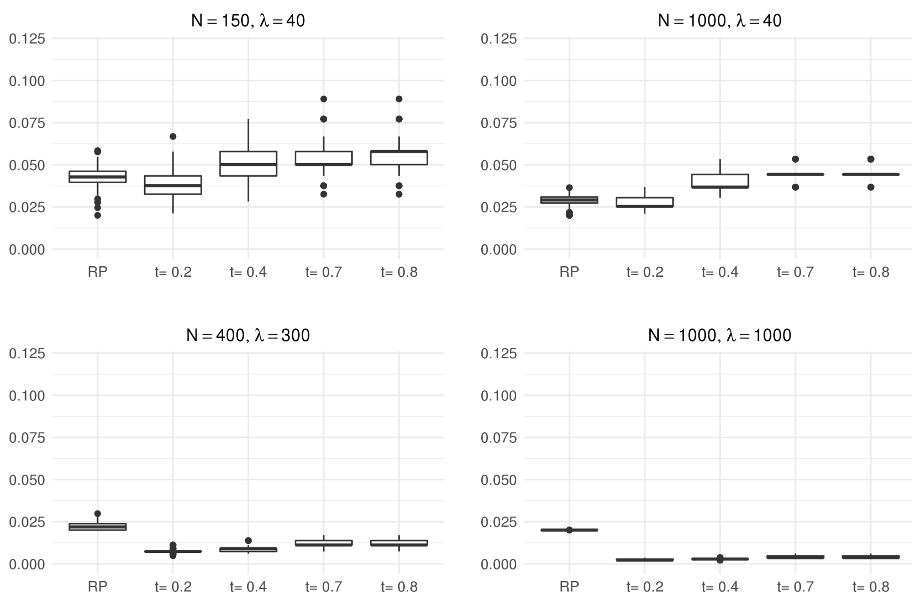

Figure S.8: Bandwidths selected by `RP20` (left boxplot) and by our local approach for the mean estimation at $t\in\{0.2,0.4,0.7,0.8\}$; results from 400 independent series generated according to the FTS Model 3.

N	λ	$t=0.2$	$t=0.4$	$t=0.7$	$t=0.8$
150	40	0.9601	0.9786	0.9971	1.0309
1000	40	1.0036	0.9838	1.0103	1.1388
400	300	0.9762	0.9993	0.9884	0.9913
1000	1000	0.9689	1.0017	0.9840	0.9886

Table S.2: MSE ratio for our mean estimator and `RP20`; results from 400 independent series generated in FTS Model 3.

S.4.5 Additional results on autocovariance function estimation

Similar to the mean function, our adaptive 'smooth first, then estimate' estimator of the autocovariance function is built with the bandwidth h_γ^* defined as in (24), obtained by minimizing the estimated bound $3\widehat{R}_\gamma(s,t;h)$ of the pointwise quadratic risk. Again, instead of the dependence coefficient $\mathbb{D}(s,t;h)$, we simply consider

$$\widehat{\mathbb{D}}(s,t;h) = \frac{1}{N-\ell}\sum_{n=1}^{N-\ell}\left\{\widetilde{X}_n(s)\widetilde{X}_{n+\ell}(t) - \widehat{\mathfrak{g}}_\ell(s,t)\right\}^2 + \sum_{k=1}^{N-\ell-1}\frac{2}{N-\ell-k}\left|\sum_{n=1}^{N-\ell-k-1}\left\{\widetilde{X}_n(s)\widetilde{X}_{n+\ell}(t) - \widehat{\mathfrak{g}}_\ell(s,t)\right\}\left\{\widetilde{X}_{n+k}(s)\widetilde{X}_{n+\ell+k}(t) - \widehat{\mathfrak{g}}_\ell(s,t)\right\}\right|,$$

where $\widehat{\mathfrak{g}}_\ell(s,t)$ is an estimator of $\gamma_\ell(s,t)$,

$$\widehat{\mathfrak{g}}_\ell(s,t) = \frac{1}{N-\ell}\sum_{n=1}^{N-\ell}\widetilde{X}_n(s)\widetilde{X}_{n+\ell}(t), \quad \ell \geq 1,$$

with $\{\widetilde{X}_n\}$ the presmoothed curves as defined in (9). Figure S.9 presents the average of the risk function $\widehat{R}_\gamma(s,t;h)$ over 400 independent functional time series generated according to FTS Model 2 with $\mu \equiv 0$, with four setups (N,λ). The plots provide evidence that $h \to R_\gamma(s,t;h)$ is a convex function which converges to zero as N and λ become larger.

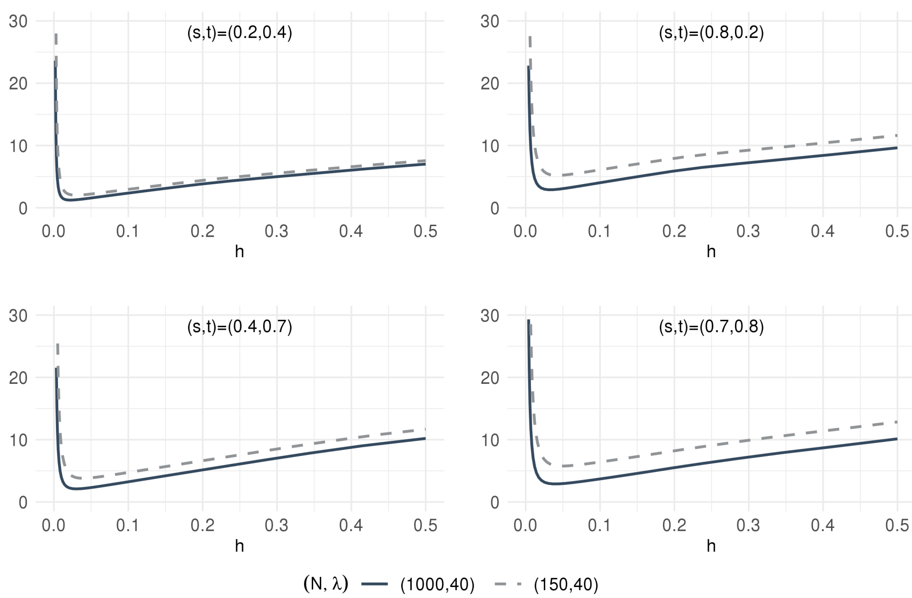

Figure S.9: Empirical average of the risk function $\widehat{R}_\gamma(h,\widehat{H}_s,\widehat{L}_s^2,\widehat{H}_t,\widehat{L}_t^2)$ of the lag-1 cross-product function $\gamma_1(s,t)$ at $(s,t) \in \{(0.2,0.4),(0.8,0.2)(0.4,0.7),(0.7,0.8)\}$ over 400 independent functional time series generated according to FTS Model 2 with $\mu \equiv 0$, with four setups (N,λ).

References

Bosq, D. (2000). *Linear processes in function spaces*, volume 149 of *Lecture Notes in Statistics*. Springer-Verlag, New York. Theory and applications.

Chao, M. T. and Strawderman, W. E. (1972). Negative moments of positive random variables. *Journal of the American Statistical Association*, 67(338):429–431.

Friedman, J., Hastie, T., and Tibshirani, R. (2010). Regularization paths for generalized linear models via coordinate descent. *Journal of Statistical Software*, 33(1):1–22.

Golovkine, S., Klutchnikoff, N., and Patilea, V. (2022). Learning the smoothness of noisy curves with application to online curve estimation. *Electron. J. Stat.*, 16(1):1485–1560.

Hebrail, G. and Berard, A. (2012). Individual household electric power consumption. UCI Machine Learning Repository. DOI: https://doi.org/10.24432/C58K54.

Hörmann, S., Horváth, L., and Reeder, R. (2013). A functional version of the ARCH model. *Econometric Theory*, 29(2):267–288.

Hörmann, S. and Kokoszka, P. (2010). Weakly dependent functional data. *Ann. Statist.*, 38(3):1845–1884.

Liu, W., Xiao, H., and Wu, W. B. (2013). Probability and moment inequalities under dependence. *Statist. Sinica*, 23(3):1257–1272.

Rubín, T. and Panaretos, V. M. (2020). Sparsely observed functional time series: estimation and prediction. *Electron. J. Stat.*, 14(1):1137–1210.

Tsybakov, A. B. (2009). *Introduction to nonparametric estimation.* Springer Series in Statistics. Springer, New York. Revised and extended from the 2004 French original, Translated by Vladimir Zaiats.

Vershynin, R. (2018). *High-dimensional probability*, volume 47 of *Cambridge Series in Statistical and Probabilistic Mathematics.* Cambridge University Press, Cambridge. An introduction with applications in data science, With a foreword by Sara van de Geer.

Wei, S. W. G., Patilea, V., and Klutchnikoff, N. (2023). Adaptive functional principal components analysis. *arXiv preprint arXiv:2306.16091.*

Wu, W. B. (2005). Nonlinear system theory: another look at dependence. *Proc. Natl. Acad. Sci. USA*, 102(40):14150–14154.