Electronic Journal of Statistics
Vol. 12 (2018) 2323–2354
ISSN: 1935-7524
https://doi.org/10.1214/18-EJS1459

Non-parametric estimation of time varying AR(1)–processes with local stationarity and periodicity

Jean-Marc Bardet[1,*] **and Paul Doukhan**[2,**]

[1] *SAMM EA4543, University Panthéon-Sorbonne, 90, rue de Tolbiac, 75634, Paris, France.*
e-mail: [*]bardet@univ-paris1.fr

[2] *AGM-UMR8088, University Cergy-Pontoise, France, and CIMFAV, Valparaiso, Chile.*
e-mail: [**]doukhan@u-cergy.fr

Abstract: Extending the ideas of [7], this paper aims at providing a kernel based non-parametric estimation of a new class of time varying AR(1) processes (X_t), with local stationarity and periodic features (with a known period T), inducing the definition $X_t = a_t(t/nT)X_{t-1} + \xi_t$ for $t \in \mathbb{N}$ and with $a_{t+T} \equiv a_t$. Central limit theorems are established for kernel estimators $\widehat{a}_s(u)$ reaching classical minimax rates and only requiring low order moment conditions of the white noise $(\xi_t)_t$ up to the second order.

MSC 2010 subject classifications: Primary 62G05, 62M10; secondary60F05.
Keywords and phrases: Local stationarity, nonparametric estimation, central limit theorem.

Received February 2018.

This paper is dedicated to the memory of Jean Bretagnolle

1. Introduction

Since the seminal paper [5], the local-stationarity property provides new models and approaches for introducing non-stationarity in times series. The recently published handbook [7] gives a complete survey about new results obtained since 20 years on this topics.

An interesting new kind of models is obtained from a natural extension of usual ARMA processes, so called tvARMA(p,q)–processes defined in [8], as:

$$\sum_{j=0}^{p} \alpha_j\left(\frac{t}{n}\right) X_{t-j}^{(n)} = \sum_{k=0}^{q} \beta_k\left(\frac{t}{n}\right) \xi_{t-k}, \qquad 1 \le t \le n, \tag{1.1}$$

where α_j and β_k are bounded functions. This is a special case of locally stationary linear process defined by $X_t^{(n)} = \sum_{j=0}^{\infty} \gamma_j\left(\frac{t}{n}\right) \xi_{t-j}$. Such models have been studied in many papers, especially concerning the parametric, semi-parametric or non-parametric estimations of functions α_j, β_k or γ_j, or other functions depending on these functions; see, for instance references [6], [8], [7], or [12], [3],

[11], [17] or [2].

For simplicity, we restrict in this first work to time-varying AR(1)–processes $(X_t^{(n)})$ including a periodic component:

$$X_t^{(n)} = a_t\left(\frac{t}{nT}\right) X_{t-1}^{(n)} + \xi_t, \quad \text{with } a_{t+T} \equiv a_t, \text{ for any} \begin{cases} 1 \le t \le nT \\ n \in \mathbb{N}^* \end{cases}, \quad (1.2)$$

where $T \in \mathbb{N}^*$ is a fixed and known integer number, and (ξ_t) a white noise. Note that given the functions $a_1, \ldots, a_T$, one may even build a periodic sequence $(a_t)_{t\in\mathbb{Z}}$ through the relation $a_{t+T} = a_t$.

The choice of such extension of the tvAR(1) processes is relative to modelling considerations: for instance, in the climatic framework, [4] considered models of air temperatures where the function of interest writes as the product of a periodic sequence by a locally varying function. This choice provide an interesting extension of more classical periodic models of air temperature such as those proposed in [14].

Other periodic representation for locally stationary processes can also be found in for instance in the paper [19], but the seasonal component is treated as an additive deterministic trend and is not included in the dynamic of the process, which is the case for model (1.2).

We then study non-parametric estimators $\widehat{a}_s(u)$, for $s = 1, 2, \ldots, T$, $u \in (0,1)$ from an observed trajectory $(X_1^{(n)}, \ldots, X_{nT}^{(n)})$. We consider kernel-based estimators which are naturally induced from covariance relationships satisfied by the process (see Section 2). Central limit theorems are established for these estimators under some regularity conditions on the functions $a_s(\cdot)$ for $s = 1, 2, \ldots, T$. The results are only obtained by assuming second-order moments on the white noise (ξ_t). This is a main improvement with respect to usual limit theorems on locally-stationary processes which are obtained with the assumption that any moment exists for (ξ_t). This is due to the new ideas developed in our proof which combines a central limit theorem for martingale increment arrays as well as an embedding in an Orlicz space (see details in Section 4).

The obtained convergence rate is optimal with respect to the minimax rate up to a logarithmic term. Simulations based on Monte-Carlo experiments illustrate the accuracy of the estimators. An application to real-life data, *i.e.* monthly average temperature readings in London from 1659 to 1998, shows the interest of using our new model (1.2).

This paper is also a first step concerning new results for new class of nonstationary processes. Indeed, we can extend the definition (1.2) to processes $(X_t^{(n)})$ such as:

$$X_t^{(n)} = F_t\left(\frac{t}{n}, \xi_t, Z_t; X_{t-1}^{(n)}, X_{t-2}^{(n)}, \ldots\right), \qquad 1 \le t \le n, \qquad (1.3)$$

where (Z_t) is a sequence of i.i.d. random vectors modelling for instance exogenous inputs. This more tough case is deferred to forthcoming papers.

Other time-varying models with an infinite memory may also be processed as GARCH-type models (see for instance [9]). Remark also that [10] introduced INGARCH-models. Those models are GLM models; non-stationary versions of which also may be considered. They will be considered in further works.

The structure of the paper is as follows. In Section 2, we define and study asymptotic properties of non-parametric estimators for the process (1.2). Section 3 provides the results of some Monte-Carlo experiments and real-life data application, while the proofs are reported in Section 4.

2. Asymptotic normality of a non-parametric estimator for periodic tvAR(1) processes

2.1. Definition and first properties of the process

Denote classically $\mathbb{N} = \{0, 1, \ldots\}$ and $\mathbb{N}^* = \{1, 2, \ldots\}$. Here we consider $T \in \mathbb{N}^*$ a fixed and known period. We will write $s \equiv t[T]$ if $t - s$ is a multiple of T.

The paper is dedicated to the simplest case $X = (X_t^{(n)})_{1\le t\le nT,\, n\in\mathbb{N}}$, of a $T-$periodic locally stationary AR(1)$-$process, defined in (1.2) where $X_0^{(n)} = X_0$ with $\mathbb{E}(X_0^2) < \infty$. Here $(\xi_t)_{t\in\mathbb{N}}$ is a sequence of i.i.d. r.v.s satisfying $\mathbb{E}(\xi_t) = 0$ and $\mathrm{Var}\,(\xi_t) = \sigma^2$ for any $t \in \mathbb{N}^*$, with $(\xi_t)_t$ independent of $X_0^{(n)}$.

The functions $(a_s(\cdot))_{1\le s\le T}$, $[0, 1] \to \mathbb{R}$ are supposed to satisfy some regularity assumption. Hence, we provide the forthcoming definition usually made in a non-parametric framework:

Definition 2.1. *For $\rho > 0$, we denote $\lceil\rho\rceil \in \mathbb{N}$ the largest integer such that $\lceil\rho\rceil < \rho$. A function $f : x \in \mathbb{R} \mapsto f(x) \in \mathbb{R}$ is said to belong to the class $\mathcal{C}^\rho(\mathcal{V}_u)$ where $\mathcal{V}_u$ is a neighbourhood of $u \in \mathbb{R}$, if $f \in \mathcal{C}^{\lceil\rho\rceil}(\mathcal{V}_u)$ and if $f^{(\lceil\rho\rceil)}$ is a $(\rho - \lceil\rho\rceil)$-Hölderian function, i.e. there exists $C_f \ge 0$ such as*

$$\left|f^{(\lceil\rho\rceil)}(u_1) - f^{(\lceil\rho\rceil)}(u_2)\right| \le C_f\, |u_1 - u_2|^{\rho-\lceil\rho\rceil}, \quad \textit{for any } u_1, u_2 \in \mathcal{V}_u.$$

In case ρ is an integer we simply assume that $f^{(\rho)}$ exists and is a continuous and bounded function on the neighbourhood of u. As a consequence we specify the assumptions on functions (a_t) using a fixed positive real number $\rho > 0$:

Assumption (A(ρ)): The functions $\{a_t(\cdot);\, t \in \mathbb{N}\}$ are such as:

1. (Periodicity) There exists $T \in \mathbb{N}^*$ such that $a_t(v) = a_{t+T}(v)$ for any $(t, v) \in \mathbb{Z} \times [0, 1]$.
2. (Contractivity) There exists $\alpha = \sup_{\{t\in\mathbb{Z},\, v\in[0,1]\}} |a_t(v)| < 1$.
3. (Regularity) For any $t \in \mathbb{Z}$, assume that $a_t \in \mathcal{C}^\rho$.

Remark 2.1. *Quote that $T = 1$ corresponds to a non-periodic case and $(X_t^{(n)})$ is then a usual tvAR(1) process defined in (1.1).*

First it is clear that the conditions on functions (a_s) ensure the existence of a causal linear process $(X_t^{(n)})_{1\leq t\leq nT}$ for any $n\in\mathbb{N}$ satisfying (1.2). More precisely, we obtain the following moment relationships:

Proposition 2.1. *Let $X=(X_t^{(n)})_{1\leq t\leq nT,\, n\in\mathbb{N}^*}$ satisfy (1.2) under Assumption (A(ρ)) with $\rho\geq 1$. Then for some convenient constant $c>0$,*

1. *For any $n\in\mathbb{N}^*$ and $1\leq t\leq nT$, $\big|\mathbb{E}\big(X_t^{(n)}\big)\big|\leq \alpha^t\,\big|\mathbb{E}(X_0)\big|$.*
2. *Let $s\in\{1,\dots,T\}$. There exists functions $\gamma_s^{(2)}\in\mathcal{C}^\rho([0,1])$ such as if $t\in\{[c\log n],\dots,nT\}$ and $t\equiv s\ [T]$:*

$$\mathbb{E}\big((X_t^{(n)})^2\big)=\gamma_s^{(2)}\big(\frac{t}{nT}\big)+\mathcal{O}\big(\frac{1}{n}\big),$$
$$\text{with}\quad\left\{\begin{array}{rcl}\gamma_s^{(2)}(v)&=&\sigma^2\,\dfrac{1+\sum_{i=0}^{T-2}\beta_{s,i}(v)}{1-\beta_{s,T-1}(v)},\\ \beta_{s,i}(v)&=&\prod_{j=0}^{i}a_{s-j}^2(v)\leq\alpha^{2i}<1.\end{array}\right.\tag{2.1}$$

3. *Assume $\mathbb{E}(\xi_0^4)=\mu_4<\infty$ and $\mathbb{E}(\xi_0^3)=0$ (this holds e.g. if ξ_0 admits a symmetric distribution).*
   For $s\in\{1,\dots,T\}$, there exist functions $\gamma_s^{(4)}\in\mathcal{C}^\rho([0,1])$ such as, for $t\in\{[c\log n],\dots,nT\}$ with $t\equiv s\ [T]$,

$$\mathbb{E}\big((X_t^{(n)})^4\big)=\gamma_s^{(4)}\big(\frac{t}{nT}\big)+\mathcal{O}\big(\frac{1}{n}\big),$$
$$\text{with}\quad\left\{\begin{array}{rcl}\gamma_s^{(4)}(v)&=&\big(\mu_4+6\sigma^2\gamma_s^{(2)}(v)-6\sigma^4\big)\dfrac{1+\sum_{i=0}^{T-2}\delta_{s,i}(v)}{1-\delta_{s,T-1}(v)},\\ \delta_{s,i}(v)&=&\prod_{j=0}^{i}a_{s-j}^4(v)\leq\alpha^{4i}<1.\end{array}\right.\tag{2.2}$$

   Moreover, for any $(t,t')\in\{[c\log n],\dots,nT\}^2$ with $t>t'$,

$$Cov\big((X_t^{(n)})^2,(X_{t'}^{(n)})^2\big)=\Big(\gamma_{s'}^{(4)}\big(\frac{t'}{nT}\big)+\mathcal{O}\big(\frac{1}{n}\big)\Big)\prod_{i=1}^{t-t'}a_{t'+i}^2\big(\frac{t'+i}{n}\big).\tag{2.3}$$

We will now assume $X_0=0$.

In addition of the previous proposition, another relation can be easily established. Indeed, for $t\in\{1,2,\dots,nT\}$, with $s=t\ [T]$, by multiplying (1.2) by $X_{t-1}^{(n)}$ and taking the expectation:

$$a_t\Big(\frac{t}{nT}\Big)=a_s\Big(\frac{t}{nT}\Big)=\frac{\mathbb{E}\big(X_tX_{t-1}\big)}{\mathbb{E}\big(X_{t-1}^2\big)}.\tag{2.4}$$

The relation (2.4) is at the origin of the definition of the following non-parametric estimators of the functions $a_s(\cdot)$.

2.2. Asymptotic normality of the estimator

Assume that the sample $(X_1, \ldots, X_{nT})$ is observed for some $n \geq 1$; this condition entails a reasonable loss of at most T data and allows us for a more comprehensive study.

For each $s \in \{1, \ldots, T\}$, we define $I_{n,s} = \big\{s, s+T, \ldots, s+(n-1)T\big\}$, a set with $\#I_{n,s} = n$. Now (2.4) writes:

$$a_s\Big(\frac{t}{nT}\Big) = \frac{\mathbb{E}\big(X_t X_{t-1}\big)}{\mathbb{E}\big(X_{t-1}^2\big)}, \qquad \forall t \in I_{n,s}.$$

A convolution kernel $K : \mathbb{R} \to \mathbb{R}$ will be required in the sequel and it satisfies one of both the following assumptions:

Assumption (K)**:** Let $K : \mathbb{R} \to \mathbb{R}$ be a Borel bounded function such that:

- $\displaystyle\int_{\mathbb{R}} K(t)dt = 1$ and $K(-x) = K(x)$ for any $x \in \mathbb{R}$;
- there exists $\beta > 0$ such as $\lim_{|t| \to +\infty} e^{\beta\,|t|} K(t) = 0$.

Assumption $(\widetilde{K})$**:** Let $K : \mathbb{R} \to \mathbb{R}$ be a Borel bounded function such that:

- $\displaystyle\int_{\mathbb{R}} K(t)dt = 1$ and $K(-x) = K(x)$ for any $x \in \mathbb{R}$;
- there exists some $B > 0$ such as $K(t) = 0$, if $|t| > B$.

Typical examples of kernel functions are $K(t) = (2\pi)^{-1/2} e^{-t^2/2}$ and $K(t) = \frac{1}{2}\, \mathbb{I}_{[-1,1]}(t)$ satisfying respectively Assumptions (K) and $(\widetilde{K})$. Note also the $K \geq 0$ would exclude dealing with a regularity $\rho > 2$.

For $r \geq 1$, we also specify another condition satisfied by such a function:

Assumption $\ker(r)$: Let $K : \mathbb{R} \to \mathbb{R}$ be a Borel bounded function such that:

- $\int_{\mathbb{R}} (|x|^r + 1)\, |K(x)|\, dx < \infty$ and $\int_{\mathbb{R}} x^p K(x)\, dx = 0$, if $p \in \{1, 2, \ldots, \lceil r \rceil - 1\}$;
- $\|K\|_\infty = \sup_{x \in \mathbb{R}} |K(x)| < \infty$ and $\mathrm{Lip}\,(K) = \sup_{x \neq y} \frac{|K(x)-K(y)|}{|x-y|} < \infty$.

Assume that a sequence of positive bandwidths $(b_n)_{n \in \mathbb{N}}$ is chosen in such a way that

$$\lim_{n \to \infty} b_n = 0, \qquad \lim_{n \to \infty} n b_n = \infty.$$

Now, keeping in mind the expression (2.4) and following the same ideas as with Nararaya-Watson estimator (see [18] and [22]), for $s \in \{1, \ldots, T\}$ and $u \in (0,1)$, we set

$$\widehat{a}_s^{(n)}(u) = \frac{\widehat{N}_s^{(n)}(u)}{\widehat{D}_s^{(n)}(u)}, \text{ with } \left\{ \begin{array}{rcl} \widehat{N}_s^{(n)}(u) & = & \frac{1}{nb_n} \sum_{j \in I_{n,s}} K\left(\frac{\frac{j}{nT} - u}{b_n}\right) X_j X_{j-1}, \\ \widehat{D}_s^{(n)}(u) & = & \frac{1}{nb_n} \sum_{j \in I_{n,s}} K\left(\frac{\frac{j}{nT} - u}{b_n}\right) X_{j-1}^2. \end{array} \right. \tag{2.5}$$

Since extremities are omitted we avoid the corresponding edge effects due to the fact that at the extremities, summations are not considered over a symmetric interval of times containing nu. The case $u=0$ does not make any contribution while the case $u=1$ corresponds with simple periodic behaviours and such results should be found in [14].

Using essentially a martingale central limit theorem (the steps of the proofs are precisely detailed in Section 4), we obtain:

Theorem 2.1. *Let $0<\rho\le 2$ and Assumption $(A(\rho))$, let K satisfy Assumption (K) or $(\widetilde{K})$ as well as Assumption ker$(\rho\vee 1)$. Then, for a sequence $(b_n)_{n\in\mathbb{N}}$ of positive real numbers such as $\lim_{n\to\infty} b_n\, n^{\frac{1}{1+2(\rho\wedge 1)}}=0$,*

$$\sqrt{nb_n}\big(\widehat{a}_s(u)-a_s(u)\big)\ \underset{n\to+\infty}{\overset{\mathcal{L}}{\longrightarrow}}\ \mathcal{N}\Big(0\,,\frac{\sigma^2}{\gamma_s^{(2)}(u)}\int_{\mathbb{R}}K^2(x)\,dx\Big), \tag{2.6}$$

for any $u\in(0,1)$, $s\in\{1,\dots,T\}$, with $\gamma_s^{(2)}(u)=\sigma^2\,\dfrac{1+\sum_{i=0}^{T-2}\beta_{s,i}(u)}{1-\beta_{s,T-1}(u)}$.

Note that for $\rho\le 1$ the classical optimal semi-parametric minimax rate is reached.

This is not the case if $\rho\in(1,2]$. In that case, another moment condition is needed in order to improve the convergence rate of $\widehat{a}_s(u)$.

Theorem 2.2. *Let $1\le\rho\le 2$ and Assumption $(A(\rho))$, let K satisfy Assumption (K) or $(\widetilde{K})$ as well as Assumption ker(ρ). Moreover, suppose that $\mathbb{E}|\xi_0|^\beta<\infty$ with $\beta=4-\dfrac{2\rho}{5\rho-4}\in\Big[2,\dfrac{10}{3}\Big]$ (Note that $\beta=2$ if $\rho=2$) and that ξ_0 admits a symmetric distribution. Then (2.6) holds for a sequence $(b_n)_{n\in\mathbb{N}}$ of positive real numbers such as $b_n\,n^{\frac{1}{2\rho+1}}\underset{n\to+\infty}{\longrightarrow}0$.*

Moreover in case $\rho=2$ and if $b_n=c\,n^{-\frac{1}{5}}$ then the central limit still holds but the limit distribution is now non-centred:

$$\mathcal{N}\Big(\mu(u)\,,\frac{\sigma^2}{\gamma_s^{(2)}(u)}\int_{\mathbb{R}}K^2(x)\,dx\Big)$$

with $\mu(u)=\dfrac{c^{\frac{5}{2}}}{\gamma_s^{(2)}(u)}\Big(\dfrac{1}{2}a_s''(u)\gamma_s^{(2)}(u)+a_s'(u)(\gamma_s^{(2)})'(u)\Big)\displaystyle\int_{\mathbb{R}}z^2K(z)\,dz.$

Remark 2.2. *Optimal window widths write as $b_n\sim cn^{-\frac{1}{2\rho+1}}$ thus the above result holds with a suboptimal window width. Moreover the symmetry assumption is discussed in Remark 4.2. Now for the case $\rho=2$ in case the derivatives of a_s are regular around the point u, then the optimal window width actually may be used and the central limit theorem again holds with a non-centred Gaussian limit.*

Quote that the proposed normalisation yields the standard minimax rates $n^{-\frac{\rho}{2\rho+1}}$, in the case of compactly supported symmetric kernel (a $(\log n)-$loss *is observed for the Gaussian kernel); the obtained rates are in probability and further work is needed to prove that this is the minimax $\mathbb{L}^2-$rate.*

Moreover for large T the convergence rate is degraded with a factor $T^{\frac{\rho}{2\rho+1}}$ since the sample size is $N = nT$ and thus $n = N/T$.

Remark 2.3. *Of course, if $T = 1$, Theorems 2.1 and 2.2 hold, which provide another minimax estimation of the function $u \mapsto a_1(u)$ ($u \in [0,1]$) requiring sharper moment and regularity conditions than the ones proposed in Theorem 4.1 of [8].*

Remark 2.4. *If T is unknown we better consider an N-sample and set $n = [N/T]$, the proof of previous central limit theorem 2.1 provides an approach for estimating this period T. First fix $T_{\max} \geq 2$ (typically $T_{\max} = 12$ for monthly data). Then, for each $1 \leq \tau \leq T_{\max}$, we define an estimator $\widehat{a}_s^{(\tau)}(u)$ for any $1 \leq s \leq \tau$ and $u \in (0,1)$. It is clear that when τ is not a multiple of T, then the sums in (2.5) that are done on the set $I_{n,s}$, which depends on τ, is now a sum involving other a_k with $k \neq s$. As a consequence, $\widehat{a}_s(u)$ is not a convergent estimator of $a_s(u)$.*

Then, using a classical cross-validation, for each $1 \leq \tau \leq T_{\max}$, we compute

$$\widehat{CV}(\tau) = \sum_{j=2}^{N} \Big(X_j^{(n)} - \widehat{a}_j^{(\tau)}\big(\frac{j}{N}\big) X_{j-1}^{(n)} \Big)^2.$$

Finally, define $\widehat{T}$ as the smallest value such as

$$\widehat{T} = \operatorname*{Argmin}_{1 \leq \tau \leq T_{\max}} \widehat{CV}(\tau).$$

Remark 2.5. *The central limit theorem 2.1 naturally provides a test statistics $\widehat{A}_s$ for solving the test problem: $H_0 : a_s(u) = c_a$ versus $H_0 : a_s(u) \neq c_a$, where $c_a \in (0,1)$. Indeed, from (2.6) and Slutsky Lemma we deduce:*

$$\sqrt{nb_n \int_{\mathbb{R}} K^2(x)\,dx}\,\sqrt{\frac{1 + \sum_{i=0}^{T-2} \prod_{j=0}^{i} \widehat{a}_{s-j}^2(u)}{1 - \prod_{j=0}^{T-1} \widehat{a}_{s-j}^2(u)}} \Big(\widehat{a}_s(u) - a_s(u)\Big) \xrightarrow[n \to +\infty]{\mathcal{L}} \mathcal{N}(0\,,1).$$

Then if we consider

$$\widehat{A}_s = \sqrt{nb_n \int_{\mathbb{R}} K^2(x)\,dx}\,\sqrt{\frac{1 + \sum_{i=0}^{T-2} \prod_{j=0}^{i} \widehat{a}_{s-j}^2(u)}{1 - \prod_{j=0}^{T-1} \widehat{a}_{s-j}^2(u)}} \Big(\widehat{a}_s(u) - c_a\Big),$$

this provides a natural statistics test with usual standard Gaussian quantile as asymptotic threshold.

3. Monte-Carlo experiments and an application to climatic data

3.1. Monte-Carlo experiments

In this section, numerous Monte-Carlo experiments have been made for studying the accuracy of the new non-parametric estimator $\widehat{a}_s(\cdot)$.

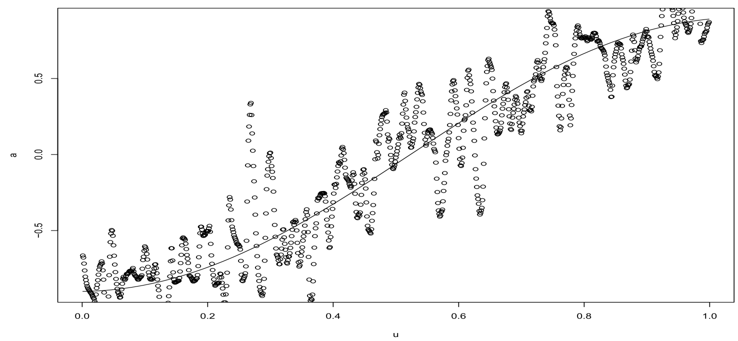

FIG 1. *Graph of the function* $a_1^{(2)}$ *and an example of its estimation (for* $n = 1000$*).*

Firstly, we considered 3 typical functions $[0,1] \to [-1,1]$, $a_s^{(\rho)}(u) \in \mathcal{C}^\rho([0,1])$ and such as $\sup_{u\in[0,1],s\in\mathbb{N}} |a_s^{(\rho)}(u)| \le \alpha < 1$:

- For $\rho = 2$, we choose $a_s^{(2)}(u) = 0.9 \cos\big(2\pi \dfrac{ns}{T}\big) \cos(3u)$. Figure 1 exhibits the graph of the function $a_1^{(2)}$ and an example of its estimation (for $n = 1000$);
- For $\rho = 1.5$, we choose $a_s^{(1.5)}(u) = 0.9 \cos\big(2\pi \dfrac{ns}{T}\big) \dfrac{\int_0^u W_t(\omega)\, dt}{\sup_{x\in[0,1]} |W_x(\omega)|}$ where $(W_t)_{t\in[0,1]}$ is an observed trajectory of a Wiener Brownian motion;
- For $\rho = 0.8$, we choose $a_s^{(0.8)}(u) = 0.9 \cos\big(2\pi \dfrac{ns}{T}\big) \dfrac{B_{0.8}(\omega,u)}{\sup_{x\in[0,1]} |B_{0.8}(\omega,x)|}$ where $B_H(\omega,t))_{t\in[0,1]}$ is an observed trajectory of a fractional Brownian motion with Hurst exponent $H = 0.8$ (Figure 2 exhibits the graph of this chosen function $a_1^{(0.8)}$). It is well known that a trajectory of a fractional Brownian motion with Hurst exponent $H \in (0,1)$ is almost surely α-Höderian for any $\alpha < H$;
- For $\rho = 0.5$, we choose $a_s^{(0.5)}(u) = 0.9 \cos\big(2\pi \dfrac{ns}{T}\big) \dfrac{W_u(\omega)}{\sup_{x\in[0,1]} |W_x(\omega)|}$ where $(W_t(\omega))_{t\in[0,1]}$ is an observed trajectory of a Wiener Brownian motion.

We also consider two "typical" kernels:

- A bounded supported kernel, the well-known Epanechnikov kernel defined by $K_E(x) = \frac{3}{4}\,(1-x^2)\, \mathbb{1}_{\{|x|\le 1\}}$, which is known to minimize the asymptotic MISE in the kernel density estimation frame;
- The unbounded supported Gaussian kernel with $K_G(x) = \frac{1}{\sqrt{2\pi}} \exp\big(-\frac{x^2}{2}\big)$.

We considered the cases $n = 100$, 200, 500 and 1000, and we fixed $T = 2$. Finally 1000 independent replications of $(X^{(n)})$ are generated with two different cases of innovations (ξ_t):

- Firstly, the case where the probability distribution of ξ_0 is a Gaussian

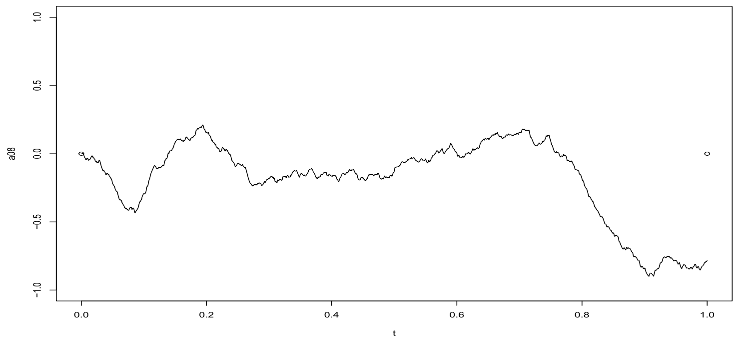

FIG 2. *Graph of the chosen function* $a_1^{(0.8)}$.

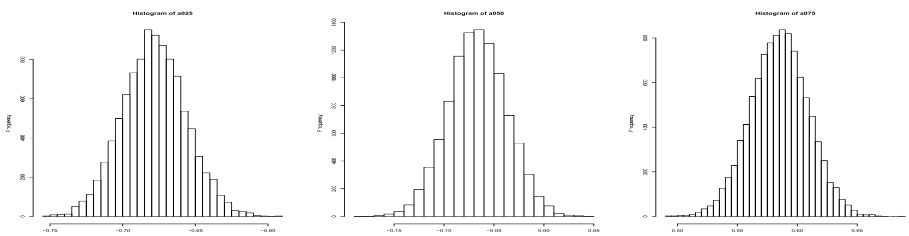

FIG 3. *Histograms of* $\widehat{a}_s^{(2)}(u)$ *for* $u = 0.25,\ 0.5$ *and* 0.75 *from* 10000 *independent replications*

$\mathcal{N}(0,4)$ distribution, then $\mathbb{E}|\xi_0|^4 < \infty$ and therefore Theorem 2.1 holds for $\rho = 0.5$ and Theorem 2.2 holds for $\rho = 1.5$ and $\rho = 2$.

- Secondly, the case where the probability distribution of ξ_0 is a Student $t(3)$ (with 3 degrees of freedom) distribution implying $\mathbb{E}|\xi_0|^\beta < \infty$ for any $\beta < 3$ but $\mathbb{E}|\xi_0|^3 = \infty$. Then if $\rho = 0.5$, Theorem 2.1 holds but if $\rho = 1.5$ and $\rho = 2$, Theorem 2.2 does not hold.

Finally, for each n, each functions $a_s^{(\rho)}$ and kernel K, and each probability distributions of ξ_0, we present the results computed from 1000 replications and the following methodology:

1. For each replication j, we defined $b_n = n^{-\lambda}$ with $\lambda = 0.10,\ 0.11, \ldots, 0.80$, $(u_i)_{1\le i\le 99} = 0.01,\ 0.02, \ldots, 0.99$, $s = 1, 2, \ldots, T$, and the estimators $\widehat{a}_s(u_i)$ are computed.
2. For each replication j and each $\lambda = 0.10, 0.11, \ldots, 0.80$, an estimator of the $MISE$ is computed:

$$\widehat{MISE}_s(\lambda) = \frac{1}{99}\sum_{i=1}^{99}\big(\widehat{a}_s(u_i) - a_s(u_i)\big)^2.$$

3. For each replication j, we minimised an estimator of the global square

TABLE 1
Results of the Monte Carlo experiments providing the accuracy of $\widehat{a}_s$ for the three chosen functions the three chosen functions with ξ_0 following a $\mathcal{N}(0,4)$ distribution, 1000 *independent replications are generated.*

	$a^{(\rho)}$	$a_s^{(2)}$		$a_s^{(1.5)}$		$a_s^{(0.8)}$		$a_s^{(0.5)}$	
	Kernel	K_E	K_G	K_E	K_G	K_E	K_G	K_E	K_G
$n=100$	$\overline{\lambda}$	0.243	0.407	0.283	0.450	0.172	0.322	0.235	0.392
	$\overline{MISE}^{1/2}$	0.248	0.239	0.286	0.282	0.230	0.234	0.354	0.353
$n=200$	$\overline{\lambda}$	0.227	0.363	0.278	0.429	0.256	0.392	0.250	0.386
	$\overline{MISE}^{1/2}$	0.185	0.175	0.219	0.219	0.232	0.232	0.308	0.303
$n=500$	$\overline{\lambda}$	0.234	0.320	0.276	0.399	0.321	0.431	0.287	0.406
	$\overline{MISE}^{1/2}$	0.129	0.119	0.154	0.156	0.213	0.210	0.256	0.254
$n=1000$	$\overline{\lambda}$	0.240	0.321	0.270	0.384	0.373	0.476	0.328	0.438
	$\overline{MISE}^{1/2}$	0.098	0.093	0.124	0.122	0.207	0.202	0.226	0.221

TABLE 2
Results of the Monte Carlo experiments providing the accuracy of $\widehat{a}_s$ for the three chosen functions with ξ_0 following a $t(3)$ distribution, 1000 *independent replications are generated.*

	$a^{(\rho)}$	$a_s^{(2)}$		$a_s^{(1.5)}$		$a_s^{(0.8)}$		$a_s^{(0.5)}$	
	Kernel	K_E	K_G	K_E	K_G	K_E	K_G	K_E	K_G
$n=100$	$\overline{\lambda}$	0.226	0.394	0.267	0.430	0.161	0.295	0.220	0.360
	$\overline{MISE}^{1/2}$	0.341	0.320	0.350	0.340	0.311	0.309	0.418	0.405
$n=200$	$\overline{\lambda}$	0.207	0.343	0.259	0.402	0.231	0.355	0.225	0.362
	$\overline{MISE}^{1/2}$	0.261	0.258	0.281	0.287	0.296	0.293	0.353	0.346
$n=500$	$\overline{\lambda}$	0.194	0.304	0.252	0.373	0.286	0.383	0.239	0.360
	$\overline{MISE}^{1/2}$	0.214	0.201	0.213	0.217	0.269	0.261	0.302	0.296
$n=1000$	$\overline{\lambda}$	0.193	0.321	0.246	0.342	0.346	0.450	0.258	0.368
	$\overline{MISE}^{1/2}$	0.166	0.093	0.172	0.181	0.258	0.250	0.262	0.275

root of MISE:

$$\widehat{\lambda}_j = \operatorname*{Argmin}_{0.1\leq\lambda\leq 0.8} \sum_{s=1}^{T} \sqrt{\widehat{MISE}_s(\lambda)}$$

4. Then we computed $\overline{\lambda} = \frac{1}{1000}\sum_{j=1}^{1000}\widehat{\lambda}_j$ over all the replications.
5. Finally, we computed the estimator of the minimal global square root of MISE,

$$\overline{MISE}^{1/2} = \frac{1}{1000}\sum_{j=1}^{1000}\sum_{s=1}^{T}\sqrt{\widehat{MISE}_s(\widehat{\lambda}_j)}.$$

As a consequence, $\overline{\lambda}$ and $\overline{MISE}^{1/2}$ are two interesting estimators relative to Theorems 2.1 and 2.2. The first one specifies the link between the choice of an optimal bandwidth b_n qnd the regularity ρ of the functions $a_s(\cdot)$. The second one measures the optimal convergence rate of the estimators $\widehat{a}_s(\cdot)$ to $a_s(\cdot)$. All the results are printed in Tables 1 and 2.

Moreover, for exhibiting the asymptotic normality of the estimators provided in the central limit theorem (2.6), we draw in Figure the histograms of $\widehat{a}_s^{(2)}(u)$ for $u = 0.25$, 0.5 and 0.75 from 10000 independent replications for $n = 5000$. We

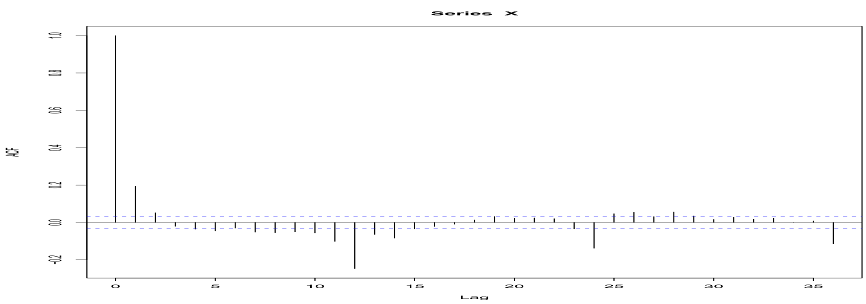

FIG 4. *Monthly average temperature readings in London from 1659 to 1998: correlogram of residual data*

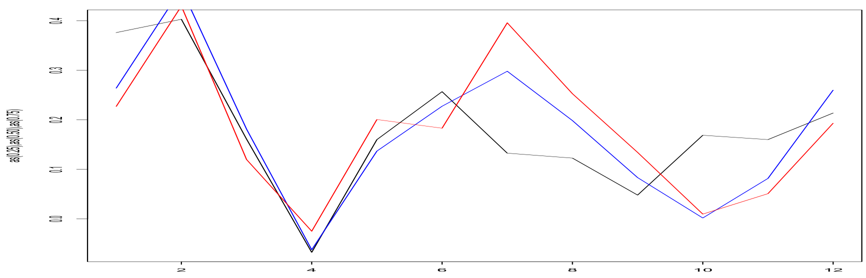

FIG 5. *Monthly average temperature readings in London from 1659 to 1998:* $\widehat{a}_s(u)$ *estimator for* $u = 0.25$ *(black),* $u = 0.50$ *(red) and* $u = 0.75$ *(blue) in terms of* $s = 1, \dots, 12$

also used a Jarque-Bera test to confirm the Gaussian asymptotic distribution since the p-values of this test are successively: $p - value = 0.105$, 0.927 and 0.345. Hence, the asymptotic normality of the estimator seems to be attested by Monte-Carlo experiments.

Conclusions of the simulations: Firstly, and as it should be deduced from Theorem 2.1 and 2.2, we observed the larger the regularity ρ, the smaller $\overline{\lambda}$ and therefore the larger the optimal bandwidth $\overline{b_n} = n^{-\overline{\lambda}}$, and the faster the convergence rate of $\widehat{a}_s$. Secondly, even if the choice of the optimal bandwidth is significantly different following the choice of the kernel (clearly smaller with the Epanechnikov kernel), the optimal convergence rate is almost the same for both the kernel. Finally, according also with Theorem 2.2, the convergence rate is clearly slower with a heavy tail distribution ($t(3)$) than with a Gaussian distribution, and this phenomenon increases when ρ increases.

3.2. Numerical application on climatic data

We also applied our model and its estimator to an example of real data, specifically the monthly average temperature readings in London from 1659 to 1998, or 340 years. Obviously in such a case one can expect that $T = 12$.

First, we removed an additive seasonal and trend component (estimated by LOESS) from these data and considered the residual data. On these, a global correlogram (see Figure 4) confirms a modelling by a process of type AR(1) and also the presence of a periodic phenomenon of period 12. As a consequence we may assume that these residual data can be modelled by the model (1.2). We then applied the $\widehat{a}_s(u)$ estimator for $u = 0.25$, 0.50 and 0.75 and $s = 1, \dots, 12$. Figure 5 summarizes these results and shows:

- The crucial interest of taking a pseudo-periodic model as we defined it in (1.2);
- The relatively small but not negligible change in the coefficient $a_t(t)$ as a function of t.

4. Proofs

We first provide the proof of Proposition 2.1.

Proof of Proposition 2.1.

1. We have $\mathbb{E}X_1^{(n)} = a_1\Big(\frac{1}{nT}\Big)\mathbb{E}(X_0)$ and $\mathbb{E}X_t^{(n)} = a_t\Big(\frac{t}{nT}\Big)\mathbb{E}X_{t-1}^{(n)})$ from the relation (1.2). From Assumption (A(ρ)) and since $\Big|a_1\Big(\frac{1}{nT}\Big)\Big| \leq \alpha < 1$, we deduce the first item of Proposition 2.1.
2. Below, for ease of reading, we will omit the exponent n. Set $v_t = \mathbb{E}\big(X_t^2\big)$, and $v = \sup_s v_s \in [0, +\infty]$; also write $\alpha_t = a_t^2\big(\frac{t}{nT}\big)$. We have:

$$v_t = \alpha_t\, v_{t-1} + \sigma^2 \leq \alpha^2 v_{t-1} + \sigma^2 \leq \alpha^2 \sup_s v_s + \sigma^2, \quad t > 0 \tag{4.1}$$

thus

$$\sup_s v_s \leq \frac{\sigma^2 + v_0}{1 - \alpha} < \infty. \tag{4.2}$$

Moreover, with $\delta_t = v_t - v_{t-T}$ for any $t > T$, we have

$$\begin{aligned} \delta_t &= \alpha_t \delta_{t-1} + (\alpha_t - \alpha_{t-T}) v_{t-T-1}, \\ \big|\delta_t\big| &\leq \alpha\big|\delta_{t-1}\big| + C|\alpha_t - \alpha_{t-T}|, \quad \text{with} \quad C > 0, \end{aligned} \tag{4.3}$$

from (4.2) and since for some constant $C > 0$,

$$\begin{aligned} \big|\alpha_t - \alpha_{t-T}\big| &= \Big|a_t^2\big(\frac{t}{nT}\big) - a_{t-T}^2\big(\frac{t-T}{nT}\big)\Big| \\ &\leq 2\alpha\ \Big|a_t\big(\frac{t}{nT}\big) - a_{t-T}\big(\frac{t-T}{nT}\big)\Big| \\ &\leq \frac{C}{n^{\rho\wedge 1}} \end{aligned} \tag{4.4}$$

from Assumption (A(ρ)). As a consequence of (4.3), we also obtain:

$$\left|\delta_t\right| \le \frac{C}{1-\alpha}\cdot\frac{1}{n^{\rho\wedge 1}} + \delta_{T+1}\alpha^{t-T+1}.$$

Thus for other constants $C, c > 0$ we derive

$$\left|\delta_t\right| \le C'\frac{1}{n^{\rho\wedge 1}}, \qquad \forall t \ge c\log n. \tag{4.5}$$

From now on, assume that $\rho \ge 1$.
Now use again the definition (1.2) of the model, and by iterating (4.1), we derive:

$$\begin{aligned} v_t &= \sigma^2 + \alpha_t\big(\sigma^2 + \alpha_{t-1}v_{t-2}\big) \\ &= \cdots \\ &= \sigma^2\Big(1 + \sum_{i=0}^{T-2}\alpha_t\cdots\alpha_{t-i}\Big) + \alpha_t\cdots\alpha_{t-T+1}\, v_{t-T} \\ &= \sigma^2\Big(1 + \sum_{i=0}^{T-2}\alpha_t\cdots\alpha_{t-i}\Big) + \alpha_t\cdots\alpha_{t-T+1}\, v_t + \mathcal{O}\big(\frac{1}{n}\big) \end{aligned}$$

from (4.5).
Hence,

$$v_t = \sigma^2\,\frac{1+\sum_{i=0}^{T-2}\alpha_t\cdots\alpha_{t-i}}{1-\alpha_t\cdots\alpha_{t-T+1}} + \mathcal{O}\big(\frac{1}{n}\big). \tag{4.6}$$

Now quoting that $\alpha_{t-j} = a^2_{t-j}\big(\frac{t-j}{nT}\big)$ we set $\widetilde{\alpha}_{t-j} = a^2_{t-j}\big(\frac{t}{nT}\big)$ for $1 \le j < T$, then since $\rho \ge 1$ and from (4.6) we derive

$$v_t = \sigma^2\,\frac{1+\sum_{i=0}^{T-2}\widetilde{\alpha}_t\cdots\widetilde{\alpha}_{t-i}}{1-\widetilde{\alpha}_t\cdots\widetilde{\alpha}_{t-T+1}} + \mathcal{O}\big(\frac{1}{n}\big) = \gamma_s^{(2)}\big(\frac{t}{nT}\big) + \mathcal{O}\big(\frac{1}{n}\big). \tag{4.7}$$

The conclusion follows.

3. The proof mimics the case of $\mathbb{E}(X_t^2)$. Denote $q_t = a_t^4\big(\frac{t}{nT}\big)$, and $\mu_k = \mathbb{E}(\xi_0^k)$, for $k = 1,2,3,4$. Then $\mu_1 = 0$ and

$$\begin{aligned} w_t = \mathbb{E}(X_t^4) &= \mathbb{E}(A_tX_{t-1}+\xi_t)^4 \\ &= q_t\, w_{t-1} + 4\,\mu_3\, A_t\,\mathbb{E}X_{t-1} + 6\,\sigma^2\,A_t^2\,v_{t-1} + \mu_4. \end{aligned}$$

Since $\mu_3 = 0$, we have:

$$w_t = q_tw_{t-1} + 6\sigma^2 v_t + \mu_4 - 6\sigma^4 \le \alpha^4\, w_{t-1} + r(t), \tag{4.8}$$

with $r(t) = 6\,\sigma^2\, v_t + \mu_4 - 6\sigma^4$ and this implies as previously $\sup_t w_t < \infty$. We also obtain for constants again denoted $C', C'' > 0$:

$$\left|w_t - w_{t-T}\right| \le \frac{C'}{n}, \qquad \forall t \ge c\log n. \tag{4.9}$$

Finally by iterating (4.8), we obtain:

$$\begin{aligned} w_t &= q_t \cdots q_{t-T+1}\, w_{t-T} + \Big(r(t) + \sum_{i=0}^{T-2} \big(q_t \cdots q_{t-i} \big) r(t-i-1) \Big) \\ &= q_t \cdots q_{t-T+1}\, w_t + \mathcal{O}\big(\frac{1}{n}\big) + \Big(r(t) + \sum_{i=0}^{T-2} \big(q_t \cdots q_{t-i} \big) r(t-i-1) \Big) \end{aligned}$$

from (4.9). Hence, always following the previous case

$$\begin{aligned} w_t &= \frac{r(t) + \sum_{i=0}^{T-2} \big(q_t \cdots q_{t-i} \big) r(t-i-1)}{1 - q_t \cdots q_{t-T+1}} + \mathcal{O}\big(\frac{1}{n}\big) \\ &= \frac{r(t) + \sum_{i=0}^{T-2} \big(q_t \cdots q_{t-i} \big) r(t-i-1)}{1 - q_t \cdots q_{t-T+1}} + \mathcal{O}\big(\frac{1}{n}\big), \end{aligned}$$

for $t \geq C'' \log n$, and this implies (2.2) from using again the regularity of the functions $(a_i)_{1 \leq i \leq T}$.
Finally, for any $t > t'$ such that $t, t' \in \{[c \log n], \ldots, nT\}$, since (X_t) is a causal process and by iteration,

$$\begin{aligned} \mathrm{Cov}\,(X_t^2, X_{t'}^2) &= \alpha_t \,\mathrm{Cov}\,(X_{t-1}^2, X_{t'}^2) + 0 + \mathrm{Cov}\,(\xi_t^2, X_{t'}^2) \\ &= \alpha_t \,\mathrm{Cov}\,(X_{t-1}^2, X_{t'}^2) \\ &= \big(\gamma_{s'}^{(4)}\big(\frac{t'}{nT}\big) + \mathcal{O}\big(\frac{1}{n}\big)\big) \prod_{i=1}^{t-t'} \alpha_{t'+i}, \end{aligned}$$

where $s' \equiv t'\ [T]$ and $\Big| \prod_{i=1}^{t-t'} \alpha_{t'+i} \Big| \leq \alpha^{2|t-t'|}$.

This completes the proof. ∎

Now we establish a technical lemma, which we were not able to find in the past literature (even if variants of this result may be found) and that will be extremely useful in the sequel. For a bounded continuous function c defined on $[0,1]$, and a kernel function H (see details below), an approximation of integral by appropriate Riemann sums yields (as for [20]'s estimator, see [21] for further developments):

$$\lim_{n \to \infty} \frac{1}{n b_n} \sum_{j \in I_{n,s}} H\left(\frac{\frac{j}{nT} - u}{b_n} \right) c\Big(\frac{j}{nT}\Big) = c(u),$$

where $u \in (0,1)$, $I_{n,s} = \big\{ s, s+T, \ldots, s+(n-1)T \big\}$ with $s \in \{1, \ldots, T\}$ and $T \in \mathbb{N}^*$. More precisely we would like to provide expansions of

$$\Delta_n = \frac{1}{n b_n} \sum_{j \in I_{n,s}} H\left(\frac{\frac{j}{nT} - u}{b_n} \right) c\Big(\frac{j}{nT}\Big) - c(u). \tag{4.10}$$

Lemma 4.1. *Let $u \in (0,1)$, $\rho > 0$, $c \in \mathcal{C}^\rho([0,1])$ a bounded function. Let H satisfy ker$(\rho \vee 1)$. Consider also a sequence of positive real numbers $(b_n)_n$ satisfying $\lim_{n\to\infty} b_n = 0$. Then, there exists $C > 0$ depending only on $\|H\|_\infty$, $\|c\|_\infty$ and Lip(H) such that for n large enough*

$$\big|\Delta_n\big| \le C\left(\frac{A_n}{nb_n} + b_n^\rho\right), \text{ with } \begin{cases} A_n = 1 & \text{under Assumption } (\widetilde{K}), \\ A_n = \log(n) & \text{under Assumption } (K). \end{cases} \tag{4.11}$$

Finally, if $\rho \in \mathbb{N}^*$ we have:

$$\Delta_n = b_n^\rho \cdot \frac{c^{(\rho)}(u)}{\rho!} \int_{\mathbb{R}} z^\rho H(z)\, dz \,\big(1 + o(1)\big) + \mathcal{O}\Big(\frac{A_n}{nb_n}\Big). \tag{4.12}$$

Proof of Lemma 4.1. In the sequel we will denote $h_n(v) = \frac{1}{b_n} H\big(b_n^{-1}(v-u)\big)$ for $v \in \mathbb{R}$. Then h_n is a Lipschitz function with Lip $h_n = \frac{1}{b_n^2}$ Lip H.

- *First assume that the function $c \equiv 1$ is a constant.* Set $v_i = i(nT)^{-1}$ for $i \in \mathbb{Z}$. For $1 \le s \le T$, we consider the sets

$$\begin{aligned} K_{n,s} &= \big\{j \in \mathbb{N},\, |v_{s+jT} - u| \le A_n b_n, 1 \le j \le n\big\} \\ &= \mathbb{N} \cap [1,n] \bigcap \Big[(u - A_n b_n)n - \frac{s}{T}\,,\, (u + A_n b_n)n - \frac{s}{T}\Big], \end{aligned}$$

and $L_{n,s} = I_{n,s} \setminus K_{n,s}$. Then, for n large enough,

$$\begin{aligned} \Delta_n &= \frac{1}{n} \sum_{i \in I_{n,s}} h_n(v_i) - \int_{\mathbb{R}} h_n(v)\, dv \\ &= \frac{1}{n} \sum_{i \in K_{n,s}} h_n(v_i) - \int_{\mathbb{R}} h_n(v)\, dv + \frac{1}{n} \sum_{i \in L_{n,s}} h_n(v_i) \\ &= \sum_{j=[(u-A_nb_n)n - \frac{s}{T}]+1}^{[(u+A_nb_n)n - \frac{s}{T}]} \int_{v_{s+jT}}^{v_{s+(j+1)T}} \big(h_n(v_{s+jT}) - h_n(v)\big) dv \\ &+ \frac{1}{n} \sum_{j \in L_{n,s}} h_n(v_{s+jT}) - \int_{(u+A_nb_n)+\frac{1}{n}}^{\infty} h_n(v)\, dv \\ &- \int_{-\infty}^{(u-A_nb_n)+1/n} h_n(v)\, dv. \end{aligned}$$

But $|h_n(v_{s+jT})| \le \frac{C}{b_n} \exp\Big(-\beta\Big|\frac{j/n - u + s/nT}{b_n}\Big|\Big)$ from Assumption (K) and using the usual comparison between sums and integrals for monotonic functions, we obtain:

$$\begin{aligned} \Big|\frac{1}{n} \sum_{j \in L_{n,s}} h_n(v_{s+jT})\Big| &\le \frac{C}{b_n} \int_{|v-u| \ge A_n b_n}^{\infty} \exp\Big(-\beta\, \frac{|v-u|}{b_n}\Big)\, dv \\ &\le 2\,C \exp(-\beta\, A_n). \end{aligned}$$

Thus

$$\begin{aligned}
\big|\Delta_n\big| &\leq \operatorname{Lip}(h_n) \sum_{j=[(u-A_nb_n)n-\frac{s}{T}]+1}^{[(u+A_nb_n)n-\frac{s}{T}]} \int_{v_{s+jT}}^{v_{s+(j+1)T}} (v - v_{s+jT})dv \\
&\quad +2\,C\,\exp(-\beta\,A_n) + \int_{u+A_nb_n}^{\infty} \big|h_n(v)\big|\,dv + \int_{-\infty}^{u-A_nb_n} \big|h_n(v)\big|\,dv \\
&\leq \frac{\operatorname{Lip}(H)}{b_n^2}\,\frac{2A_nb_nn}{2n^2} + 4\,C\,\exp(-\beta\,A_n) \\
&\leq \operatorname{Lip}(H)\,\frac{A_n}{nb_n} + 4\,C\,\exp(-\beta\,A_n).
\end{aligned}$$

because since $u \in (0,1)$, the above indices remain in the index set $[-n,n]$ for n large enough.
Then, if $A_n \geq \beta^{-1}\log n$ then $\exp(-\beta\,A_n) \leq 1/n$ and we deduce (4.11).

- *We now turn to the case of a non-constant function c.* First, if $\rho > 0$, for $(u,v) \in (0,1)^2$ the Taylor-Lagrange formula implies:

$$c(v) - c(u) = (v-u)c'(u) + \cdots + \frac{(v-u)^\ell}{\ell!}\,c^{(\ell)}\big(u + \lambda(v-u)\big),$$

with $\ell = \lceil \rho \rceil$ and $\lambda \in (0,1)$. Since $c \in \mathcal{C}^\rho([0,T])$,

$$\big|c^{(\ell)}\big(u+\lambda(v-u)\big) - c^{(\ell)}(u)\big| \leq C_\rho\,\big|\lambda(v-u)\big|^{\rho-\ell} \leq C_\rho\,\big|v-u\big|^{\rho-\ell}.$$

Therefore,

$$c(v) - c(u) = (v-u)c'(u) + \cdots + \frac{(v-u)^\ell}{\ell!}\,c^{(\ell)}(u) + R(u,v), \tag{4.13}$$

with $|R(u,v)| \leq C_\rho\,|u-v|^\rho$. Then for any $u \in (0,1)$, using Assumption $\ker(\rho \vee 1)$ and especially the relation $\int z^p H(z)dz = 0$ for $p = 1,\ldots,\ell$,

$$\begin{aligned}
&\Big|\int_{\mathbb{R}} h_n(v)\,c(v)\,dv - c(u)\int_{\mathbb{R}} h_n(v)\,dv\Big| \\
&= \Big|\int_{-\infty}^{\infty} H(z)\big(c(u+b_nz) - c(u)\big)\,dz\Big| \\
&= \Big|\int_{-\infty}^{\infty} H(z)R(u,u+b_nz)dz\Big| && (4.14)\\
&\leq C_\rho\,b_n^\rho\,C\int_{-\infty}^{\infty} e^{-\beta|z|}|z|^\rho dz \\
&\leq C'\,b_n^\rho && (4.15)
\end{aligned}$$

with $C' > 0$. Here we denote $k_n(v) = h_n(v)c(v)$ for $v \in [0,1]$.
Now, if $\rho \in (0,1)$, we have

$$|k_n(v_1) - k_n(v_2)| \leq \|c\|_\infty \operatorname{Lip}(h_n)\,|v_1 - v_2| + \frac{\|H\|_\infty}{b_n}\,C_\rho\,|v_1 - v_2|^\rho,$$

and therefore using the previous results:

$$\begin{aligned}\big|\Delta_n\big| &\leq \sum_{j=[(u-A_nb_n)n-\frac{s}{T}]+1}^{[(u+A_nb_n)n-\frac{s}{T}]} \int_{v_{s+jT}}^{v_{s+(j+1)T}} C\Big(\mathrm{Lip}\,(h_n)\,|v-v_{s+jT}| \\ &\qquad +\frac{1}{b_n}|v-v_{s+jT}|^{\rho}\Big)\,dv + C\,\|c\|_\infty\,\exp(-\beta A_n) \\ &\qquad +\Big|\int_{\mathbb{R}} h_n(v)\,c(v)\,dv - c(u)\int_{\mathbb{R}} h_n(v)\,dv\Big| \\ &\leq C\Big(\frac{A_n}{nb_n}+\frac{A_n}{n^{\rho}}+\exp(-\beta A_n)+b_n^{\rho}\Big).\end{aligned}$$

from (4.15) and this implies (4.11) since $nb_n \to \infty$ and therefore $n^{-\rho}$ is negligible with respect from b_n^{ρ}.
Now, if $\rho \geq 1$ and since H and c are bounded continuous Lipschitz functions, we obtain the inequality

$$\mathrm{Lip}\,(k_n) \leq \|c\|_\infty \mathrm{Lip}\,(h_n) + \frac{1}{b_n}\,\|H\|_\infty \mathrm{Lip}\,(c) < \infty.$$

Then, using the same computations than previously (replace h_n by $h_n \times c$),

$$\begin{aligned}|\Delta_n| &\leq \sum_{j=[(u-A_nb_n)n-\frac{s}{T}]+1}^{[(u+A_nb_n)n-\frac{s}{T}]} \int_{v_{s+jT}}^{v_{s+(j+1)T}} \big|k_n(v_{s+jT})-k_n(v)\big|\,dv \\ &+\ C\,\|c\|_\infty\,\exp(-\beta A_n) + \Big|\int_{\mathbb{R}} h_n(v)\,c(v)\,dv - c(u)\int_{\mathbb{R}} h_n(v)\,dv\Big| \\ &\leq\ C\frac{A_n}{nb_n}\Big(\|c\|_\infty \mathrm{Lip}\,(H) + b_n\,\mathrm{Lip}\,(c)\|H\|_\infty\Big) + C\|c\|_\infty e^{-\beta A_n} + C' b_n^{\rho},\end{aligned}$$

from (4.15) and this completes the first item since b_n is supposed to converge to 0. The proof is now easily completed.

- *Finally, in the case* $\rho \in \mathbb{N}^*$, we can use the previous case an a Taylor-Lagrange expansion of the function c, implying $R(u,v) = \dfrac{c^{(\rho)}(\theta)}{\rho!}\,\big|u-v\big|^{\rho}$ with $\theta = \lambda u + (1-\lambda)v$ and $\lambda \in [0,1]$.
Then, using (4.14) and with $\mu_u(z) \in [0,1]$, and $\zeta_n = \displaystyle\int_{\mathbb{R}} h_n(v)\,c(v)\,dv - c(u)\int_{\mathbb{R}} h_n(v)\,dv$

$$\begin{aligned}\zeta_n &= \frac{b_n^{\rho}}{\rho!}\int_{-\infty}^{\infty} H(z)z^{\rho}c^{(\rho)}\big(u+\mu_u(z)b_nz\big)\,dz \\ &= \frac{b_n^{\rho}}{\rho!}\,c^{(\rho)}(u)\int_{-\infty}^{\infty} H(z)z^{\rho}\,dz\,\big(1+o(1)\big)\end{aligned}$$

from Lebesgue theorem on dominated convergence. ∎

In the sequel we will denote the σ-algebra

$$\mathcal{F}_t^{(s)} = \sigma\big((\xi_i)_{i\le s+(t-1)T}\big). \tag{4.16}$$

Lemma 4.2. *Let H satisfy Assumption ker(1) and $(X_t^{(n)})$ be a solution of (1.2) under Assumption ($A(\rho)$) with $\rho > 0$. Then for any $u \in (0,1)$, and $s \in \{1,\dots,T\}$,*

$$\frac{1}{nb_n}\sum_{j=1}^{n} H\Big(\frac{\frac{s+(j-1)T}{nT}-u}{b_n}\Big)\big(X^{(n)}_{s+(j-1)T-1}\big)^2 \xrightarrow[n\to+\infty]{\mathbb{P}} \sigma^2\,\frac{1+\sum_{i=0}^{T-2}\beta_{s,i}(u)}{1-\beta_{s,T-1}(u)}.$$

Proof of Lemma 4.2. We use here a limit theorem for $\mathbb{L}^1$-mixingales established in [1]. Indeed, for $u\in(0,1)$, $s\in\{1,\dots,T\}$, let

$$Z_{n,t} = \frac{1}{b_n}\,H\Big(\frac{\frac{s+(t-1)T}{nT}-u}{b_n}\Big)\Big(\big(X^{(n)}_{s+(t-1)T-1}\big)^2 - \mathbb{E}\big(X^{(n)}_{s+(t-1)T-1}\big)^2\Big). \tag{4.17}$$

Then, set

$$c_0(t)=1, \quad \text{and} \quad c_k(t)=\prod_{i=1}^{k} a_{t+1-i}\Big(\frac{t+1-i}{nT}\Big), \qquad \text{for } k\ge 1,$$

we have:

$$X_t^{(n)} = \sum_{k=0}^{\infty} c_k(t)\,\xi_{t-k}. \tag{4.18}$$

Therefore, with $(\mathcal{F}^{(s)}_{n,t})$ defined in (4.16),

$$\begin{aligned}
\mathbb{E}\big[Z_{n,t}|\mathcal{F}^{(s)}_{n,t-m}\big] &= \frac{1}{b_n}\,H\Big(\frac{\frac{s+(j-1)T}{nT}-u}{b_n}\Big)\\
&\quad\times\Big\{\mathbb{E}\Big[\Big(\sum_{k=0}^{\infty} c_k(s+(t-1)T-1)\,\xi_{s+(t-1)T-1-k}\Big)^2\,|\;\mathcal{F}^{(s)}_{n,t-m}\Big]\\
&\qquad -\sigma^2\sum_{k=0}^{\infty} c_k^2(s+(t-1)T-1)\Big\}\\
&= \frac{1}{b_n}\,H\Big(\frac{\frac{s+(j-1)T}{nT}-u}{b_n}\Big)\Big\{\Big(\sum_{k=mT-1}^{\infty} c_k(s+(t-1)T-1)\,\xi_{s+(t-1)T-1-k}\Big)^2\\
&\qquad -\sigma^2\sum_{k=mT-1}^{\infty} c_k^2(s+(t-1)T-1)\Big\}.
\end{aligned}$$

But for any $t\in\mathbb{N}$, we have $|c_k(t)|\le\alpha^k$ from Assumption (A(ρ)). Then,

$$\Big\|\mathbb{E}\big[Z_{n,t}\mid\mathcal{F}^{(s)}_{n,t-m}\big]\Big\|_1 \le \frac{1}{b_n}\left|H\Big(\frac{\frac{s+(j-1)T}{nT}-u}{b_n}\Big)\right|$$
$$\times\Big\{\mathbb{E}\Big[\Big(\sum_{k=mT-1}^{\infty}c_k(s+(t-1)T-1)\,\xi_{s+(t-1)T-1-k}\Big)^2\Big]+\sigma^2\sum_{k=mT-1}^{\infty}\alpha^{2k}\Big\}$$
$$\le\frac{2\,\sigma^2}{b_n}\left|H\Big(\frac{\frac{s+(j-1)T}{nT}-u}{b_n}\Big)\right|\times\frac{\alpha^{2mT-2}}{1-\alpha^2}.$$

Thus, using the notations of Definition 2 in [1], it is easy to derive that $(Z_{n,t})$ is a triangular array such that $\phi_m=\alpha^{2mT-2}\|H\|_1\to 0$ (as $m\to\infty$) since $0\le\alpha<1$ and:

$$\frac{1}{n}\sum_{t=1}^{n}|c_{nt}|\underset{n\to+\infty}{\longrightarrow}\frac{2\,\sigma^2}{(1-\alpha^2)}\|H\|_1<\infty,\quad\text{with}$$
$$c_{nt}=\frac{2\,\sigma^2}{(1-\alpha^2)b_n}H\Big(\frac{\frac{s+(j-1)T}{nT}-u}{b_n}\Big).$$

As a consequence,

$$\frac{1}{n}\sum_{t=1}^{n}Z_{n,t}\underset{n\to+\infty}{\overset{\mathbb{P}}{\longrightarrow}}0,$$

implies

$$\frac{1}{nb_n}\sum_{j=1}^{n}H\Big(\frac{\frac{s+(j-1)T}{nT}-u}{b_n}\Big)\Big(\big(X^{(n)}_{s+(j-1)T-1}\big)^2-\mathbb{E}\big(\big(X^{(n)}_{s+(j-1)T-1}\big)^2\big)\Big)\underset{n\to+\infty}{\overset{\mathbb{P}}{\longrightarrow}}0.$$

Now, we collect the above relations. Lemma 4.1 and Proposition 2.1 with the $\rho-$regularity of the function $c(v)$, together conclude the proof. ■

Lemma 4.3. *Under the conditions of Theorem 2.1, with $(Y_{n,i})_{1\le i\le n,\,n\in\mathbb{N}}$ defined in (4.28), for any $\varepsilon>0$,*

$$\sum_{j=1}^{n}\mathbb{E}\big(Y^2_{n,j}\,1\!\!1_{\{|Y_{n,j}|\ge\varepsilon\}}\,|\,\mathcal{F}^{(s)}_{j-1}\big)\underset{n\to+\infty}{\overset{\mathbb{P}}{\longrightarrow}}0.\qquad(4.19)$$

Proof of Lemma 4.3. Since $\mathbb{E}\xi_0^2=\sigma^2<\infty$ this is easy to exhibit an increasing sequence $(c_k)_k$ with

$$c_0=1,\;\;c_1=2\text{ and }c_{k+1}\ge c_k^2,\;\;\text{where }\mathbb{E}\big(\xi_0^2\,1\!\!1_{\{|\xi_0|\ge c_k\}}\big)\le\frac{1}{k^3},\quad\text{for all }k\in\mathbb{N}^*.$$

Define $g(\cdot)$ as the piecewise affine function such that $g(c_k)=k$ for $k\in\mathbb{N}$ and $g(0)=0$. Then the function ψ defined by $\psi(x)=x^2g(x)$ for $x\ge0$ satisfies $\psi(0)=0$ and it is a continuous and non-decreasing function (for almost all

$x > 0$, $\psi'(x) = x^2 g'(x) + 2xg(x) > 0$) and convex function (indeed, for almost all $x > 0$, $\psi''(x) = 4xg'(x) + 2g(x) > 0$). Hence, we have:

$$\sum_{k=1}^{\infty} \mathbb{E}\Big(\xi_0^2 g(|\xi_0|) \mathbb{1}_{\{c_k \le |\xi_0| < c_{k+1}\}}\Big) \le \sum_{k=0}^{\infty} \mathbb{E}\Big(\xi_0^2 g(|\xi_0|) \mathbb{1}_{\{k \le g(|\xi_0|) < k+1\}}\Big)$$
$$\le \sum_{k=1}^{\infty} (k+1) \mathbb{E}\big(\xi_0^2 \mathbb{1}_{\{c_k \le |\xi_0|\}}\big) \le \sum_{k=1}^{\infty} \frac{k+1}{k^3} < \infty.$$

Therefore,

$$\mathbb{E}\psi(|\xi_0|) \le \mathbb{E}\big(\xi_0^2 g(|\xi_0|) \mathbb{1}_{\{0 \le |\xi_0| < 2\}}\big) + \sum_{k=0}^{\infty} \mathbb{E}\big(\xi_0^2 g(|\xi_0|) \mathbb{1}_{\{c_k \le |\xi_0|\big) < c_{k+1}\}} < \infty. \tag{4.20}$$

The construction of $(c_k)_k$ and the relation $c_{k+1} \ge c_k^2$ together imply:

$$\psi(xy) \le \psi(x)\psi(y). \tag{4.21}$$

Indeed, this relationship is equivalent to

$$g(xy) \le g(x)g(y), \quad \text{for any } 0 \le x \le y. \tag{4.22}$$

But if $0 \le x \le 1$ and $y \ge x$, then $xy \le y$: therefore $g(xy) \le g(y) \le g(x)g(y)$ since g is an increasing function and $g(x) \ge 1$ for any $x \ge 0$. Moreover, if $1 < x \le y$, there exists $0 \le k$ and $\lambda \in [0,1]$ such as $y = \lambda c_k + (1-\lambda)c_{k+1}$. But $h : [0,\infty) \to \mathbb{R}^+$ defined by $x \mapsto h(x) = g(x^2)$ is a convex function since $h'' \ge 0$ a.e. As a consequence,

$$g(y^2) = h(\lambda c_k + (1-\lambda)c_{k+1}) \le \lambda g(c_k^2) + (1-\lambda) g(c_{k+1}^2)$$
$$\le \lambda g(c_{k+1}) + (1-\lambda) g(c_{k+2}) \le \lambda(k+1) + (1-\lambda)(k+2) = k + 2 - \lambda,$$

from the construction of (c_k). Since $g(y) = \lambda g(c_k) + (1-\lambda) g(c_{k+1}) = k + 1 - \lambda$ because g is a piecewise function, we finally obtain $g(y^2) \le g(y)+1$. We conclude with $g(xy) \le g(y^2)$ for any $1 \le x \le y$ and $g(x) \ge 2$ (since $c_1 = 2$).

Hence the function ψ is a Orlicz function and $\|\xi_0\|_\psi < \infty$ with

$$\|V\|_\psi = \inf\Big\{z > 0;\ \mathbb{E}\psi\Big(\frac{|V|}{z}\Big) \le 1\Big\}, \qquad \text{for any random variable } V. \tag{4.23}$$

Now Theorem 1.1 in [16] implies:

$$\|V\|_\psi \le \inf_{z>0} \frac{1}{z}\big(1 + \mathbb{E}\big[\psi(z|V|)\big]\big) \le 2\,\|V\|_\psi. \tag{4.24}$$

Therefore $\|V\|_\psi \le 1 + \mathbb{E}\psi(|V|)$, and $\frac{1}{z}\mathbb{E}\psi(z|V|) \le 2\|V\|_\psi$ for any $z > 0$ since from convexity

$$\mathbb{E}\psi(|V|) \le \frac{z-1}{z} \cdot \mathbb{E}\psi(0) + \frac{1}{z} \cdot \mathbb{E}\psi(z|V|) \le 2\|V\|_\psi$$

and $\psi(0)=0$. Then, from the definition of $(X_t^{(n)})$ and the triangular inequality

$$\|X_t^{(n)}\|_\psi \leq \alpha\|X_{t-1}^{(n)}\|_\psi + \|\xi_t\|_\psi \leq \sum_{j=0}^{t-1}\alpha^j\|\xi_{t-j}\|_\psi \quad \text{for any } t\in\mathbb{N}^*,$$

with $0\leq\alpha<1$. Since $\|\xi_s\|_\psi=\|\xi_0\|_\psi$ for any $s\in\mathbb{N}$, we finally obtain

$$\sup_{t\in\mathbb{N}}\left\{\|X_t^{(n)}\|_\psi\right\} \leq \frac{1}{1-\alpha}\|\xi_0\|_\psi < \infty.$$

Thus (4.21) implies with the independence of ξ_t and $X_{t-1}^{(n)}$ that:

$$\mathbb{E}\psi(|\xi_t X_{t-1}^{(n)}|) \leq \mathbb{E}\psi(|\xi_t|)\cdot\mathbb{E}\psi(|X_{t-1}^{(n)}|).$$

Now relation (4.24) with $z=1$ entails

$$\sup_{t\in\mathbb{N}^*}\left\{\|\xi_t X_{t-1}^{(n)}\|_\psi\right\} < \infty.$$

Thus with $t=s+(j-1)T$ we have from (4.24),

$$\|Y_{n,j}\|_\psi \leq \frac{1}{\sqrt{nb_n}}\left|K\left(\frac{\frac{t}{nT}-u}{b_n}\right)\right|\|\xi_t\|_\psi\|X_{t-1}^{(n)}\|_\psi < \infty.$$

Again using (4.21) and with $K_t=K\left(\frac{\frac{t}{nT}-u}{b_n}\right)$,

$$\begin{aligned}
\mathbb{E}\big(Y_j^2\mathbb{1}_{\{|Y_j|\geq\varepsilon\}}\big) &= \frac{1}{nb_n}\mathbb{E}\Big((K_t\xi_tX_{t-1}^{(n)})^2\mathbb{1}_{\left\{g(|K_t\xi_tX_{t-1}^{(n)}|)\geq g(\varepsilon\sqrt{nb_n})\right\}}\Big)\\
&\leq \frac{1}{nb_n}\mathbb{E}\Big((K_t\xi_tX_{t-1}^{(n)})^2\\
&\qquad\cdot\frac{g(|K_t\xi_tX_{t-1}^{(n)}|)}{g(\varepsilon\sqrt{nb_n})}\mathbb{1}_{\left\{g(|K_t\xi_tX_{t-1}^{(n)}|)\geq g(\varepsilon\sqrt{nb_n})\right\}}\Big)\\
&\leq \frac{1}{\psi(\varepsilon\sqrt{nb_n})}\mathbb{E}\big(\psi(K_t\xi_tX_{t-1}^{(n)})\big)\\
&\leq \frac{2\psi(|K_t|)}{\psi(\varepsilon\sqrt{nb_n})}\sup_{t\in\mathbb{N}^*}\|\xi_tX_{t-1}^{(n)}\|_\psi.
\end{aligned}$$

As a consequence, for any $\varepsilon>0$,

$$\begin{aligned}
&\mathbb{E}\Big(\sum_{j=1}^n\mathbb{E}\big(Y_{n,j}^2\mathbb{1}_{\{|Y_{n,j}|>\varepsilon\}}|\mathcal{F}_{j-1}^{(s)}\big)\Big)\\
&\qquad\leq\frac{\sup_{t\in\mathbb{N}^*}\|\xi_tX_{t-1}^{(n)}\|_\psi}{\varepsilon^2\,g(\varepsilon\sqrt{nb_n})}\times\frac{1}{nb_n}\sum_{j=1}^n\psi\left(\left|K\left(\frac{\frac{s+(j-1)T}{nT}-u}{b_n}\right)\right|\right)
\end{aligned}$$

$$\leq 2 \times \frac{\sup_{t\in\mathbb{N}^*} \|\xi_t X_{t-1}^{(n)}\|_\psi}{\varepsilon^2\, g(\varepsilon\sqrt{nb_n})} \int_{\mathbb{R}} \psi(|K(x)|)\, dx$$

if n is large enough, from Lemma 4.1. As a consequence, for any $\varepsilon > 0$, since $g(\varepsilon\sqrt{nb_n}) \underset{n\to+\infty}{\longrightarrow} \infty$, then $\mathbb{E}\Big(\sum_{j=1}^n \mathbb{E}\big(Y_{n,j}^2 \mathbb{1}_{\{|Y_{n,j}|>\varepsilon\}} | \mathcal{F}_{j-1}^{(s)}\big)\Big) \underset{n\to+\infty}{\longrightarrow} 0$. Since $Y_{n,j}^2 \mathbb{1}_{\{|Y_{n,j}|>\varepsilon\}}$ is a non-negative triangular array, the proof of Lemma 4.3 is complete. ∎

Proof of Theorem 2.1. Using (1.2), write

$$\widehat{N}_s^{(n)}(u) = \frac{1}{nb_n} \sum_{j\in I_{n,s}} K\Big(\frac{\frac{j}{nT}-u}{b_n}\Big) X_{j-1}^{(n)} \Big(a_s\Big(\frac{j}{nT}\Big) X_{j-1}^{(n)} + \xi_j\Big)$$

we decompose it as: $\widehat{N}_s^{(n)}(u) = \widetilde{N}_s^{(n)}(u) + M_s^{(n)}(u)$, with

$$M_s^{(n)}(u) = \frac{1}{nb_n} \sum_{j\in I_{n,s}} K\Big(\frac{\frac{j}{nT}-u}{b_n}\Big) \xi_j X_{j-1}^{(n)},$$

$$\widetilde{N}_s^{(n)}(u) = \frac{1}{nb_n} \sum_{j\in I_{n,s}} K\Big(\frac{\frac{j}{nT}-u}{b_n}\Big) a_s\big(\frac{j}{nT}\big) (X_{j-1}^{(n)})^2$$

Therefore we obtain:

$$\sqrt{nb_n}\big(\widehat{a}_s(u) - a_s(u)\big) = \sqrt{nb_n}\, \frac{M_s^{(n)}(u)}{\widehat{D}_s^{(n)}(u)} + \frac{J_n}{\widehat{D}_s^{(n)}(u)}, \qquad (4.25)$$

with

$$\widehat{D}_s^{(n)}(u) = \frac{1}{nb_n} \sum_{j\in I_{n,s}} K\Big(\frac{\frac{j}{nT}-u}{b_n}\Big) X_{j-1}^2, \qquad (4.26)$$

$$J_n = \frac{1}{\sqrt{nb_n}} \sum_{j\in I_{n,s}} K\Big(\frac{\frac{j}{nT}-u}{b_n}\Big) (X_{j-1}^{(n)})^2 \Big(a_s\big(\frac{j}{nT}\big) - a_s(u)\Big). \qquad (4.27)$$

We are going to derive the consistency of the estimator $\widehat{a}_s(u)$ of $a_s(u)$, in two parts.

1/ We first prove that $\sqrt{nb_n} M_s^{(n)}(u) \Big/ \widehat{D}_s^{(n)}(u) \xrightarrow[n\to+\infty]{\mathcal{L}} \mathcal{N}\big(0, C\big)$ for some convenient constant $C > 0$.
Let $s \in \{1, \ldots, T\}$ and $u \in (0, 1)$. For $n \in \mathbb{N}^*$ and $j \in \{1, \ldots, n\}$, we denote

$$Y_{n,j} = \frac{1}{\sqrt{nb_n}} K\Big(\frac{\frac{s+(j-1)T}{nT}-u}{b_n}\Big) \xi_{s+(j-1)T} X_{s+(j-1)T-1}^{(n)}. \qquad (4.28)$$

This is clear that $(Y_{n,j})_{\leq j\leq n,\ n\in\mathbb{N}^*}$ is a triangular array of martingale increments with respect to the σ-algebra $\mathcal{F}_t^{(s)}=\sigma\big((\xi_i)_{i\leq s+(t-1)T}\big)$. Indeed $(X_t^{(n)})_{t\geq 0}$ is a process, causal with respect to $(\xi_t)_{t\geq 0}$. This implies that ξ_t is independent of $(X_i^{(n)})_{i\leq t-1}$ and that $\mathbb{E}(\xi_0)=0$. We are going to use a central limit theorem for triangular arrays of martingale increments, see for example [13] and more recently [15].

Denote

$$\sigma_{n,j}^2=\mathbb{E}\big(Y_{n,j}^2\,|\,\mathcal{F}_{j-1}^{(s)}\big)=\frac{1}{nb_n}K^2\Big(\frac{\frac{s+(j-1)T}{nT}-u}{b_n}\Big)\big(X_{s+(j-1)T-1}^{(n)}\big)^2,$$

since $\mathbb{E}(\xi_0^2)=0$. Using Lemma 4.2, we obtain:

$$\sum_{j=1}^n\sigma_{n,j}^2\ \xrightarrow[n\to+\infty]{\mathbb{P}}\ \sigma^2\cdot\frac{1+\sum_{i=0}^{T-2}\beta_{s,i}(u)}{1-\beta_{s,T-1}(u)}\cdot\int_{\mathbb{R}}K^2(x)dx,\tag{4.29}$$

$\widehat{D}_s^{(n)}(u)$ is defined from (4.26) and satisfies

$$\widehat{D}_s^{(n)}(u)\ \xrightarrow[n\to+\infty]{\mathbb{P}}\ \sigma^2\,\frac{1+\sum_{i=0}^{T-2}\beta_{s,i}(u)}{1-\beta_{s,T-1}(u)}\equiv\gamma_s^{(2)}(u).\tag{4.30}$$

Moreover, from Lemma 4.3, then for any $\varepsilon>0$,

$$\sum_{j=1}^n\mathbb{E}\big(Y_{n,j}^2\mathbb{1}_{\{|Y_{n,j}|\geq\varepsilon\}}\,|\,\mathcal{F}_{j-1}^{(s)}\big)\ \xrightarrow[n\to+\infty]{\mathbb{P}}\ 0.$$

As a consequence, the conditions of the central limit theorem for triangular arrays of martingale increments, in [15]), are satisfied and this implies that $\dfrac{\sum_{j=1}^n Y_{n,j}}{\sqrt{\sum_{j=1}^n\sigma_{n,j}^2}}\ \xrightarrow[n\to+\infty]{\mathcal{L}}\ \mathcal{N}\big(0,1\big)$.

Therefore from Slutsky lemma entails:

$$\begin{aligned}&\sqrt{nb_n}\frac{M_s^{(n)}(u)}{\widehat{D}_s^{(n)}(u)}\\&\quad=\frac{\sum_{j=1}^n Y_{n,j}}{\sqrt{\sum_{j=1}^n\sigma_{n,j}^2}}\times\frac{\sqrt{\sum_{j=1}^n\sigma_{n,j}^2}}{\frac{1}{nb_n}\sum_{j=1}^n K\Big(\frac{\frac{s+(j-1)T}{nT}-u}{b_n}\Big)\big(X_{s+(j-1)T-1}^{(n)}\big)^2}\\&\qquad\xrightarrow[n\to+\infty]{\mathcal{L}}\ \mathcal{N}\left(0,\sigma^2\,\frac{1-\beta_{s,T-1}(u)}{1+\sum_{i=0}^{T-2}\beta_{s,i}(u)}\int_{\mathbb{R}}K^2(x)\,dx\right).\end{aligned}\tag{4.31}$$

2/ The second term $J_n/\widehat{D}_s^{(n)}(u)$ in the expansion of $\sqrt{nb_n}\big(\widehat{a}_s(u)-a_s(u)\big)$ depends on the non-martingale term J_n, see (4.27), and the consistent term $\widehat{D}_s^{(n)}(u)$, see (4.26) and (4.30). The asymptotic behavior of this second term can be first obtained following two steps.

a. A first step consists in establishing an expansion of $\mathbb{E}J_n$. Using Proposition 2.1 and with $\gamma_s^{(2)} \in \mathcal{C}^\rho([0,1])$ defined in (2.1), we have

$$\mathbb{E}J_n = \sqrt{nb_n}\,\frac{1}{nb_n}\sum_{j\in I_{n,s}} K\Big(\frac{\frac{j}{nT}-u}{b_n}\Big) \times \Big(\gamma_s^{(2)}(\frac{j}{nT}) + \mathcal{O}(\frac{1}{n})\Big)\Big(a_s(\frac{j}{nT}) - a_s(u)\Big).$$

Using twice Lemma 4.1, with firstly $c(x) = \gamma_s^{(2)}(x)(a_s(x)-a_s(u))$, and secondly $c(x) = (a_s(x)-a_s(u))$, we deduce:

$$\big|\mathbb{E}J_n\big| \leq C\,\sqrt{nb_n}\,\Big(\frac{A_n}{nb_n} + b_n^\rho\Big)\Big(1+\mathcal{O}(\frac{1}{n})\Big). \tag{4.32}$$

As a consequence, if $b_n = o\big(n^{-1/(1+2\rho)}\big)$, then $\mathbb{E}J_n \underset{n\to+\infty}{\longrightarrow} 0$.

In the case $\rho \in \{1,2\}$, we also obtain from (4.12) and with $d_s(v) = (a_s(v)-a_s(u))\gamma_s^{(2)}(v) \in \mathcal{C}^\rho([0,1])$,

$$\begin{aligned}
\mathbb{E}J_n &= \sqrt{nb_n}\,\Big(\mathcal{O}(\frac{A_n}{nb_n}) + b_n^\rho\,\frac{d_s^{(\rho)}(u)}{\rho!}\int_{\mathbb{R}} z^\rho K(z)\,dz\,\big(1+o(1)\big)\Big) \\
&= \frac{d_s^{(\rho)}(u)}{\rho!}\int_{\mathbb{R}} z^\rho K(z)\,dz\,\sqrt{nb_n^{2\rho+1}} + o(\sqrt{nb_n^{2\rho+1}}) + \mathcal{O}(\frac{A_n}{\sqrt{nb_n}}) \\
&= \begin{cases} o(\sqrt{nb_n^3}) + \mathcal{O}(\frac{A_n}{\sqrt{nb_n}}), & \text{if } \rho = 1, \\ B_s(u)\,\sqrt{nb_n^5} + o(\sqrt{nb_n^5}) + \mathcal{O}(\frac{A_n}{\sqrt{nb_n}}), & \text{if } \rho = 2. \end{cases}
\end{aligned} \tag{4.33}$$

with $B_s(u) = \dfrac{d_s''(u)}{2}\displaystyle\int_{\mathbb{R}} z^2 K(z)\,dz$.

b. Now we are going to prove a first consistency result for $J_n/\widehat{D}_s^{(n)}(u)$ using the Markov Inequality. Indeed,

$$\begin{aligned}
\mathbb{E}|J_n| &\leq \frac{1}{\sqrt{nb_n}}\sum_{j\in I_{n,s}} \Big|K\Big(\frac{\frac{j}{nT}-u}{b_n}\Big)\Big|\Big|a_s(\frac{j}{nT}) - a_s(u)\Big|\mathrm{Var}\,(X_{j-1}^{(n)}) \\
&\leq \frac{C}{\sqrt{nb_n}}\sum_{j\in I_{n,s}} \Big|K\Big(\frac{\frac{j}{nT}-u}{b_n}\Big)\Big|\Big|a_s(\frac{j}{nT}) - a_s(u)\Big|.
\end{aligned}$$

Now using Lemma 4.1 with $c(v) = |a_s(v)-a_s(u)|$ which also belongs in $\mathcal{C}^\rho([0,1])$ (this is clear if $\rho<1$ and, for $\rho=1$ the Lipschitz property of $z\mapsto|z|$ allows to conclude), and $c(u)=0$, we derive:

$$\mathbb{E}|J_n| \leq \sqrt{nb_n}\Big(\frac{A_n}{nb_n} + b_n^{\rho\wedge 1}\Big). \tag{4.34}$$

Therefore, if $b_n = o\big(n^{-\frac{1}{1+2(\rho\wedge 1)}}\big)$, then $\mathbb{E}J_n \underset{n\to+\infty}{\longrightarrow} 0$ and $\mathbb{E}|J_n| \underset{n\to+\infty}{\longrightarrow} 0$, implying from Markov Inequality, $J_n \overset{\mathbb{P}}{\underset{n\to+\infty}{\longrightarrow}} 0$. Finally, since (4.30) establishes the consistency of $\widehat{D}_s^{(n)}(u)$, from Slutsky lemma, we deduce

$$\frac{J_n}{\widehat{D}_s^{(n)}(u)} \overset{\mathbb{P}}{\underset{n\to+\infty}{\longrightarrow}} 0. \tag{4.35}$$

As a consequence, the proof of the Theorem results by using the decomposition (4.25), the consistency results (4.31) and (4.35). ∎

Proof of Theorem 2.2. We restrict to the case $\rho \in (1, 2]$.

a. Case $\mathbb{E}\big(\xi_0^4\big) < \infty$.

Denote again $K_t = K\Big(\frac{\frac{t}{nT} - u}{b_n}\Big)$, for $t \in \mathbb{Z}$. First remark that the symmetry assumption on ξ_0's distribution implies $\mathbb{E}\big(\xi_0\big) = \mathbb{E}\big(\xi_0^3\big) = 0$.

$$\begin{aligned}
\mathrm{Var}\,(J_n) &= \frac{1}{nb_n} \sum_{t\in I_{n,s}} \sum_{t'\in I_{n,s}} K_t K_{t'} \,\mathrm{Cov}\,(X_t^2, X_{t'}^2) \times \\
&\qquad\qquad \times \Big(a_s\big(\frac{t}{nT}\big) - a_s(u)\Big)\Big(a_s\big(\frac{t}{nT}\big) - a_s(u)\Big) \\
&= \frac{1}{nb_n} \sum_{(t,t')\in L_{n,s,\alpha}} K_t K_{t'} \,\mathrm{Cov}\,(X_t^2, X_{t'}^2) \times \\
&\qquad\qquad \times \Big(a_s\big(\frac{t}{nT}\big) - a_s(u)\Big)\Big(a_s\big(\frac{t'}{nT}\big) - a_s(u)\Big) \\
&+ \frac{1}{nb_n} \sum_{(t,t')\in I_{n,s}^2\setminus L_{n,s,\alpha}} K_t K_{t'} \mathrm{Cov}\,(X_t^2, X_{t'}^2) \times \\
&\qquad\qquad \times \Big(a_s\big(\frac{t}{nT}\big) - a_s(u)\Big)\Big(a_s\big(\frac{t'}{nT}\big) - a_s(u)\Big)
\end{aligned}$$

with $L_{n,s,\alpha} = \big\{(t,t') \in I_{n,s}^2,\ \ |t-t'| \le \frac{\log n}{\log \alpha}\big\}$.
Firstly, consider the first left side term of the last inequality. If $t \in I_{n,s}$ then Proposition 2.1 entails $\mathrm{Var}\,(X_t^2) = \gamma_s^{(2)}(t/(nT)) + \mathcal{O}(1/n)$ for an adequate function $\gamma_s^{(2)} \in \mathcal{C}^\rho([0,1])$.
Hence we also have $\mathrm{Var}\,(X_t^2) = \gamma_s^{(2)}(t/(nT)) + \mathcal{O}(\log(n)/n)$.
Here the fact that $(z \mapsto z^2)$ is a function in $\mathcal{C}^\rho$, implies that the function defined from $b(v) = \big(a_s(v) - a_s(u)\big)^2$ is in $\mathcal{C}^\rho([0,1])$ too, and again $b(u) = 0$ and $\int xH^2(x)dx = 0$.
Therefore, we use Lemma 4.1 to derive:

$$\sum_{t,t'\in L_{n,s,\alpha}} K_t K_{t'} \mathrm{Cov}\,(X_t^2, X_{t'}^2)\Big(a_s\big(\frac{t}{nT}\big) - a_s(u)\Big)\Big(a_s\big(\frac{t'}{nT}\big) - a_s(u)\Big)$$

$$= \sum_{t,t' \in L_{n,s,\alpha}} K_t K_{t'} \prod_{i=1}^{|t-t'|} a_{s+i}^2\big(\frac{t}{nT}\big)\Big(\gamma_s^{(4)}\big(\frac{t+i}{nT}\big) + \mathcal{O}\big(\frac{1}{n}\big)\Big)$$
$$\times \Big(a_s\big(\frac{t}{nT}\big) - a_s(u)\Big)\Big(a_s\big(\frac{t'}{nT}\big) - a_s(u)\Big)$$
$$= \sum_{j=0}^{\frac{\log n}{\log \alpha}} \sum_{t \in I_{n,s}} K_t\Big(K_t + \mathcal{O}\big(\frac{\log n}{nb_n^2}\big)\Big) \prod_{i=1}^{j} a_{s+i}^2\big(\frac{t}{nT}\big)\Big(\gamma_s^{(4)}\big(\frac{t}{nT}\big) + \mathcal{O}\big(\frac{\log n}{n}\big)\Big)$$
$$\times \Big(a_s\big(\frac{t}{nT}\big) - a_s(u)\Big)\Big(a_s\big(\frac{t'}{nT}\big) - a_s(u)\Big)$$
$$\leq \sum_{j=0}^{\frac{\log n}{\log \alpha}} \alpha^{2j}\Big(\sum_{t \in I_{n,s}} K_t^2 \prod_{i=1}^{j} \gamma_s^{(4)}\big(\frac{t}{nT}\big) \times \Big(a_s\big(\frac{t}{nT}\big) - a_s(u)\Big)^2 + \mathcal{O}\big(\frac{\log n}{nb_n^2}\big)\Big)\Big)$$
$$\leq 2\sum_{j=0}^{\infty} \alpha^{2j} \sum_{t \in I_{n,s}} K_t^2 \prod_{i=1}^{j} \gamma_s^{(4)}\big(\frac{t}{nT}\big) \times \Big(a_s\big(\frac{t}{nT}\big) - a_s(u)\Big)^2$$
$$\leq 2nb_n \sum_{j=0}^{\infty} \alpha^{2j} \frac{1}{nb_n} \sum_{t \in I_{n,s}} K^2\Big(\frac{\frac{t}{nT} - u}{b_n}\Big) g_j\big(\frac{t}{nT}\big),$$

with $g_j(x) = \big(a_s(x) - a_s(u)\big)^2 \prod_{i=1}^{j} \big(\gamma_s^{(4)}(x)$, since for n large enough the above expression satisfies $\big|\mathcal{O}\big(\frac{\log n}{nb_n^2}\big)\big| \leq 1$. Using Lemma 4.1, with functions $H = K^2$ and $c = g_j$ with $g_j \in \mathcal{C}^\rho([0,1])$ (quote that $\max_{i \leq j} \big(\|g_i\| \vee \mathrm{Lip}\,(g_i)\big) = \mathcal{O}(j)$), we finally obtain:

$$\Big|\frac{1}{nb_n} \sum_{t,t' \in L_{n,s,\alpha}} K_t K_{t'} \mathrm{Cov}\,(X_t^2, X_{t'}^2)\big(a_s\big(\frac{t}{nT}\big) - a_s(u)\big)\big(a_s\big(\frac{t'}{nT}\big) - a_s(u)\big)\Big|$$
$$\leq C\Big(\frac{A_n}{nb_n} + b_n^\rho\Big). \quad (4.36)$$

Secondly, from Proposition 2.1, for $t, t' \in I_{n,s}^2 \setminus L_{n,s,\alpha}$, we have

$$|\mathrm{Cov}\,(X_t^2, X_{t'}^2)| \leq C\,\alpha^{2\,|t-t'|} \leq \frac{C}{n^2}.$$

Thus,

$$\Big|\frac{1}{nb_n} \sum_{t,t' \in I_{n,s}^2 \setminus L_{n,s,\alpha}} K_t K_{t'} \mathrm{Cov}\,(X_t^2, X_{t'}^2)\Big(a_s\big(\frac{t}{nT}\big) - a_s(u)\Big)\Big(a_s\big(\frac{t'}{nT}\big) - a_s(u)\Big)\Big|$$
$$\leq \frac{nb_n}{n^2}\Big(\frac{1}{nb_n} \sum_{t \in I_{n,s}} K_t\big(a_s\big(\frac{t}{nT}\big) - a_s(u)\big)\Big)^2 \leq C\,\frac{b_n}{n}\Big(\frac{A_n}{nb_n} + b_n^\rho\Big), \quad (4.37)$$

from Lemma 4.1. Then, (4.36) and (4.37) provide

$$\operatorname{Var}\big(J_n\big) \leq C\left(\frac{A_n}{nb_n} + b_n^\rho\right) \tag{4.38}$$

implying $\operatorname{Var}\big(J_n\big) \underset{n\to+\infty}{\longrightarrow} 0$ for any (b_n) such as

$$\max(b_n, A_n(n\,b_n)^{-1}) \underset{n\to+\infty}{\longrightarrow} 0.$$

b. Case $\mathbb{E}\big(|\xi_0|^\beta\big) < \infty$, for some $\beta \in [2,4]$.

From its expression given in (4.27), J_n is a quadratic form of (X_t) and therefore, as X_t is a linear process with innovations (ξ_t), J_n is also a quadratic form of (ξ_t). As a consequence, the fourth order moment can be injected such as there exists a sequence $z_n \downarrow 0$ (as $n \uparrow \infty$) satisfying:

$$\operatorname{Var}\left(J_n\right) \leq z_n\left(\mathbb{E}(\xi_0^4) \vee 1\right) = z_n\left(\mu_4 \vee 1\right), \text{ and } z_n = \mathcal{O}\Big(\frac{A_n}{nb_n} + b_n^\rho\Big). \tag{4.39}$$

Now, assume only that $\mathbb{E}(\xi_0^2) < \infty$. The innovations (ξ_t) can be truncated at level M, and write

$$\xi_{t,M} = \xi_t\, \mathbb{1}_{|\xi_t| \leq M} \qquad \text{for any } t \in \mathbb{N}.$$

Note that the symmetry assumption entails $\mathbb{E}(\xi_{j,M}) = 0$. Define also Define also

$$\begin{aligned} X_{t,M}^{(n)} &= a_t\Big(\frac{t}{nT}\Big)X_{t-1,M}^{(n)} + \xi_{t,M}, \qquad 1 \leq t \leq nT,\ n \in \mathbb{N} \\ \text{and} \quad J_{n,M} &= \frac{1}{\sqrt{nb_n}} \sum_{j \in I_{n,s}} K\Big(\frac{\frac{j}{nT} - u}{b_n}\Big)(X_{j-1}^{(n)})^2\Big(a_s\Big(\frac{j}{nT}\Big) - a_s(u)\Big). \end{aligned}$$

A consequence of (4.39) is:

$$\operatorname{Var}\left(J_{n,M}\right) \leq z_n\, \mathbb{E}(\xi_{0,M}^4) \leq z_n\, M^2 h(M), \tag{4.40}$$

with $h(M) = \mathbb{E}\big(|\xi_0|^2 \mathbb{1}_{\{|\xi_0| > M\}}\big)$ which satisfies $\lim_{M\to\infty} h(M) = 0$. Moreover,

$$\begin{aligned} \big|J_n - J_{n,M}\big| = \frac{1}{\sqrt{nb_n}} \sum_{j \in I_{n,s}} K\Big(\frac{\frac{j}{nT} - u}{b_n}\Big) \\ \times \big|(X_{j-1}^{(n)})^2 - (X_{j-1,M}^{(n)})^2\big|\big|a_s\Big(\frac{j}{nT}\Big) - a_s(u)\big|. \end{aligned} \tag{4.41}$$

But

$$X_{j-1,M}^{(n)} - X_{j-1}^{(n)} \quad = \quad a_{j-1}\Big(\frac{j-1}{nT}\Big)\big(X_{j-2,M}^{(n)} - X_{j-2}^{(n)}\big) + \big(\xi_{j-1,M} - \xi_{j-1}\big),$$

$$\big|X^{(n)}_{j-1,M}-X^{(n)}_{j-1}\big| \quad\le\quad \alpha\,\big|X^{(n)}_{j-2,M}-X^{(n)}_{j-2}\big|+|\xi_{j-1}|\mathbb{I}_{\{|\xi_{j-1}|>M\}}. \qquad (4.42)$$

We first remark from Proposition 2.1 that $\mathbb{E}(X^{(n)}_{j-1})^2+\mathbb{E}(X^{(n)}_{j-1,M})^2\le c$ for some constant $c>0$. Hence, Cauchy-Schwartz Inequality shows that, for each j:

$$\mathbb{E}\big(\big|(X^{(n)}_{j-1})^2-(X^{(n)}_{j-1,M})^2\big|\big)\le\sqrt{c\delta_{j-1,M}}, \qquad (4.43)$$

with $\delta_{j-1,M}=\mathbb{E}\big(|X^{(n)}_{j-1}-X^{(n)}_{j-1,M}|^2\big)$.
We are going to bound $\delta_{j-1,M}$. A first simple bound is clearly $\delta_{j-1,M}\le 2\,c$ and we use it together with (4.42), and Cauchy-Schwartz inequality in order to derive

$$\begin{aligned}
\delta_{j-1,M} &\le \alpha^2\delta_{j-2,M}+2\alpha\,\mathbb{E}\big(|X^{(n)}_{j-2,M}-X^{(n)}_{j-2}||\xi_{j-1}|\mathbb{I}_{\{|\xi_{j-1}|>M\}}|\big)\\
&\qquad +\mathbb{E}\big(|\xi_{j-1}|^2\mathbb{I}_{\{|\xi_{j-1}|>M\}}\big)\\
&\le \alpha^2\delta_{j-2,M}+2\alpha\sqrt{2c}\sqrt{\mathbb{E}|\xi_{j-1}|^2\mathbb{I}_{\{|\xi_{j-1}|>M\}}}\\
&\qquad +\mathbb{E}|\xi_{j-1}|^2\mathbb{I}_{\{|\xi_{j-1}|>M\}}\\
&\le \alpha^2\delta_{j-2,M}+H(M)\quad(\text{with}\quad H(M)=2\alpha\sqrt{2c}\sqrt{h(M)}+h(M))\\
&\le \alpha^4\delta_{j-3,M}+(1+\alpha^2)H(M)\\
&\le \cdots\\
&\le \alpha^{2(j-1)}\delta_{0,M}+(1+\cdots+\alpha^{2(j-2)})H(M)\\
&\le \frac{2}{1-\alpha^2}H(M)
\end{aligned}$$

since $\delta_{0,M}\le h(M)\le H(M)$.
Now, from (4.43), we obtain for M large enough:

$$\mathbb{E}\big|(X^{(n)}_{j-1})^2-(X^{(n)}_{j-1,M})^2\big|\le\sqrt{\frac{2c}{1-\alpha^2}}\sqrt{H(M)}\le C\,h^{1/4}(M) \qquad (4.44)$$

with $C>0$ and always with $h(M)=\mathbb{E}\big(|\xi_0|^2\mathbb{I}_{\{|\xi_0|>M\}}\big)$. Now a careful use of (4.34) and (4.41) entails:

$$\mathbb{E}|J_n-J_{n,M}|\le C\sqrt{nb_n}\Big(\frac{A_n}{nb_n}+b_n\Big)\,h^{1/4}(M) \qquad (4.45)$$

since $x\to|a(x)-a(u)|$ is a $\mathcal{C}^1$ function (in the above defined sense). Finally, using Cauchy-Schwartz inequality in (4.40), we obtain for M large enough,

$$\begin{aligned}
\mathbb{E}|J_n| &\le \mathbb{E}|J_n-J_{n,M}|+\sqrt{\mathrm{Var}\,(J_{n,M})}\\
&\le C\left(\sqrt{nb_n}\Big(\frac{A_n}{nb_n}+b_n\Big)\,h^{1/4}(M)+\Big(\frac{A_n}{nb_n}+b_n^{\rho}\Big)^{1/2}M\,h^{1/2}(M)\right)\\
&\le C\left(\sqrt{nb_n^3}\,h^{1/4}(M)+b_n^{\rho/2}\,M\,h^{1/2}(M)\right) \qquad (4.46)
\end{aligned}$$

assuming $A_n/nb_n = o(b_n^{\rho/2})$ *i.e.* $(n/A_n)^{-2/(2+\rho)} = o(b_n)$ (and note that $-2/(2+\rho) \le 1/(1+2\rho)$).
Now, if $\mathbb{E}\big(|\xi_0|^\beta\big) < \infty$ with $\beta \in (2,4]$, then using Hölder and Markov Inequalities, there exists $C_\beta > 0$ such as

$$h(M) = \mathbb{E}\big(|\xi_0|^2 \mathbb{I}_{\{|\xi_0|>M\}}\big) \le C_\beta\, M^{2-\beta}.$$

Since here $b_n = o\big(n^{-1/(1+2\rho)}\big)$, does not yields the minimax rates, we deduce that

$$\begin{cases} \sqrt{nb_n^3}\, h^{1/4}(M) \underset{n\to+\infty}{\longrightarrow} 0 & \text{when} \quad M^{1+2\rho} \ge n^{(4\rho-4)/(\beta-2)} \\ b_n^{\rho/2}\, M\, h^{1/2}(M) \underset{n\to+\infty}{\longrightarrow} 0 & \text{when} \quad M^{1+2\rho} \le n^{\rho/(4-\beta)} \end{cases}.$$

Thus, from inequality (4.46), we deduce that the optimal choice is obtained when

$$\frac{4\rho-4}{\beta-2} = \frac{\rho}{4-\beta}, \quad \text{which entails } \beta = 4 - 2\cdot\frac{\rho}{5\rho-4}.$$

d. Case $\rho = 2$.
The expression of the non-central limit for the case of optimal window widths and the expansion of the bias (4.33) now the asymptotic expression for (4.35) yields the proposed non-centred Gaussian limit, see Remark 4.1. The same truncation step as above is also needed.

The proof is now complete. ∎

Remark 4.1. *Using the previous bound* (4.32) *of* $\mathbb{E}J_n$ *and Bienaymé-Tchebychev inequality, we deduce that if* $b_n = o\big(n^{-1/(1+2\rho)}\big)$ *then* $J_n \underset{n\to+\infty}{\overset{\mathbb{P}}{\longrightarrow}} 0$.

Moreover, if $\rho = 2$ *and* $b_n = c\,n^{-1/5}$, *using the expansion* (4.33) *of* $\mathbb{E}J_n$ *and again Bienaymé-Chebychev inequality, then* $J_n \underset{n\to+\infty}{\overset{\mathbb{P}}{\longrightarrow}} B_s(u)\,c^{5/2}$.

Therefore with the consistency result (4.30), *for any* $u \in (0,1)$ *and* $s \in \{1,\dots,T\}$,

$$\frac{J_n}{\widehat{D}_s^{(n)}(u)} \underset{n\to+\infty}{\overset{\mathbb{P}}{\longrightarrow}} B_s(u)\,\frac{c^{\frac{5}{2}}}{\sigma^2}\,\frac{1+\sum_{i=0}^{T-2}\beta_{s,i}(u)}{1-\beta_{s,T-1}(u)}.$$

Remark 4.2. *For the general case with maybe* ξ_0 *non symmetric and* $\mathbb{E}\xi_0 = 0$, *the item 3. of Proposition 2.1 needs some improvements. Denote* $w_t^{(k)} = \mathbb{E}(X_t^k)$ *for* $k = 1,3$, *then* $w_t^{(4)} = w_t$ *and* $w_t^{(2)} = v_t$, *then* (4.8) *turns to be written*

$$w_t = q_t w_{t-1} + 4\mathbb{E}A_t\mathbb{E}X_{t-1}\mu_3 + 6\sigma^2 v_t + \mu_4 \le \alpha^4 w_{t-1} + r(t), \tag{4.47}$$

as previously $\sup_t w_t < \infty$.

We need to derive suitable equivalents of $w_t^{(k)}$ *if* $k = 1$. *Firstly*

$$w_t^{(1)} = \mathbb{E}A_t w_{t-1}^{(1)} = \cdots = EA_t \cdots \mathbb{E}A_1 \cdot \mathbb{E}X_0,$$

and in fact this term is negligible and the proof of Proposition 2.1 and Lemma 3. remains unchanged.

In this case the proof of the above point **2/ c.** *needs a simple improvement and*

$$\xi_{j,M} = \xi_j \wedge M \vee (-M) - \mathbb{E}\left(\xi_j \wedge M \vee (-M)\right).$$

In this truncated setting, inequality (4.42) writes:

$$|X^{(n)}_{j-1,M} - X^{(n)}_{j-1}| \le \alpha |X^{(n)}_{j-2,M} - X^{(n)}_{j-2}| + |\xi_{j-1}| \mathbb{1}_{\{|\xi_{j-1}|>M\}} + \mathbb{E}(|\xi_{j-1}| \mathbb{1}_{\{|\xi_{j-1}|>M\}})$$

so that the end of the proof is unchanged by only setting $C = 2c\mathbb{E}\xi_0^2/(1-\alpha)$.

Remark 4.3. *Secondly, in case we even omit the condition $\mathbb{E}\xi_0 = 0$ one needs to also express an asymptotic expansion for $w^{(3)}_t = \mathbb{E}A_t^3 w^{(3)}_{t-1} + 3\mathbb{E}A_t w^{(1)}_{t-1}\sigma^2 + \mu_3 \sim \mathbb{E}A_t^3 w^{(3)}_{t-1} + \mu_3$; an analogue expansion to Proposition 2.1 and Lemma 3. may thus be derived. Namely $w^{(3)}_t = \gamma^{(3)}_s(\frac{t}{nT}) + \mathcal{O}(\frac{1}{n}),$, with*

$$\begin{aligned}
\gamma^{(3)}_s(v) &= \mu_3 \cdot \frac{1 + \sum_{i=0}^{T-2} \zeta_{s,i}(v)}{1 - \zeta_{s,T-1}(v)}, \\
\zeta_{t,i}(v) &= \prod_{j=0}^{i-1} a^3_{t-j}(v) \le \alpha^{3i} < 1, \qquad \text{for} \quad 1 \le i \le T, \quad v \in (0,1).
\end{aligned}$$

Then the expression of the equivalent of w_t is also adequately transformed up to the above relations.

Acknowledgments

This work has been developed within the "MME-DII centre of excellence" (ANR-11-LABEX-0023-01) and with the help of PAI- CONICYT MEC Nr. 80170072.

The authors thank the referees for their fruitful comments and suggestions, which notably improved the quality of the paper. The second author wishes to thank Rainer Dahlhaus for many interesting discussions. As well, numerous discussions with Karine Bertin were extremely useful.

References

[1] Andrews, D. Laws of large numbers for dependent non-identically distributed random variables. *Econometric Theory 4*, 3 (1988), 458–467. MR0985156

[2] Azrak, R. and Mélard, G. Asymptotic properties of quasi-maximum likelihood estimators for ARMA models with time-dependent coefficients. *Statistical Inference for Stochastic Processes 9* (2006), 279–330. MR2252243

[3] Bibi, A. and Francq, C. Consistent and asymptotically normal estimators for cyclically time-dependent linear models. *Annals of the Institute of Statistical Mathematics 55* (2003), 41–68. MR1965962

[4] Dacunha-Castelle, D., Huong Hoang, H. T. and Parey, S. Modeling of air temperatures: preprocessing and trends, reduced stationary process, extremes, simulation. *Journal de la Société Française de Statistique 156*, 2 (2015), 138–168. MR3338246

[5] Dahlhaus, R. On the Kullback-Leibler information divergence of locally stationary processes. *Stochastic Processes and Applications 62* (1996), 139–168. MR1388767

[6] Dahlhaus, R. Fitting time series models to nonstationary processes. *Annals of Statistics 25* (1997), 1–37. MR1429916

[7] Dahlhaus, R. *Locally Stationary Processes*, vol. 30. Time Series Analysis: Methods and Applications, Elsevier, 2012.

[8] Dahlhaus, R. and Polonik, W. Empirical spectral processes for locally stationary time series. *Bernoulli 15* (2009), 1–39. MR2546797

[9] Doukhan, P., Grublyté, I. and Surgailis, D. A nonlinear model for long memory conditional heteroscedasticity. *Lithuanian Journal of Mathematic 56* (2016), 164–188. MR3504254

[10] Fokianos, K. and Tjøstheim, D. Log-linear Poisson autoregression. *Journal of Multivariate Analysis 102* (2011), 563–578. MR2755016

[11] Francq, C. and Gautier, E. Large sample properties of parameter least squares estimates for time-varying ARMA models. *Journal of Time Series Analysis 25* (2004), 765–783. MR2089194

[12] Grenier, E. ARMA models with time-dependent coefficients: estimators and applications. *Traitement du Signal 3* (1986), 219–233.

[13] Hall, P. and Heyde, C. C. *Martingale Limit Theory and Its Application*. Academic Press, 1980. MR0624435

[14] Lund, R., Hurd, H., Bloomfield, P. and Smith, R. Climatological time series with periodic correlation. *Journal of Climat 8* (1995), 2787–2809.

[15] Major, P. Central limit theorems for martingales. `http://www.math-inst.hu/~major/probability/martingale.pdf` (2016), 1–22.

[16] Maligranda, L. Orlicz spaces and interpolation. In *Seminars in Math. 5*, B. Campinas Sao Paulo, Ed., vol. 5 of *IMS Lecture Notes-Monograph Series*. 1989, pp. 127–140. MR2264389

[17] Moulines, E., Priouret, P. and Roueff, F. On recursive estimation for locally stationary time varying autoregressive processes. *Annals of Statistics 33* (2005), 2610–2654. MR2253097

[18] Nadaraya, E. On estimating regression. *Theory of Probability and Applications 9* (1964), 1411-1412.

[19] Paraschakisa, K. and Dahlhaus, R. Frequency and phase estimation in time series with quasi periodic components. *Journal of Time Series Analysis 33* (2012), 13–31. MR2877605

[20] Priestley, M. E., and Chao, M. T. Nonparametric function fitting. *Journal of The Royal Statistical Society, Series B 34* (1972), 385–392.

MR0331616

[21] Rosenblatt, M. *Stochastic curve estimation*, vol. 3. NSF-CBMS Regional Conference Series in Probability and Statistics, 1991.

[22] Watson, G. Smooth regression analysis. *Sankhyä Ser. A 26* (1964), 359–372. MR0185765