



HAL
open science

On bandwidth selection problems in nonparametric trend estimation under martingale difference errors

Karim Benhenni, Didier A. Girard, Sana Louhichi

► **To cite this version:**

Karim Benhenni, Didier A. Girard, Sana Louhichi. On bandwidth selection problems in nonparametric trend estimation under martingale difference errors. 2020. hal-02514827v1

HAL Id: hal-02514827

<https://hal.science/hal-02514827v1>

Preprint submitted on 23 Mar 2020 (v1), last revised 10 Apr 2021 (v2)

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

On bandwidth selection problems in nonparametric trend estimation under martingale difference errors

Karim Benhenni*, Didier A. Girard†, Sana Louhichi‡

March 23, 2020

Abstract

In this paper, we are interested in the problem of smoothing parameter selection in nonparametric curve estimation under dependent errors. We focus on kernel estimation and the case when the errors form a general stationary sequence of martingale difference random variables where neither linearity assumption nor "all moments are finite" are required. We compare the behaviors of the smoothing bandwidths obtained by minimizing three criteria: the average squared error, the mean average squared error and a Mallows-type criterion adapted to the dependent case. We prove that these three minimizers are first-order equivalent in probability. We give also a normal asymptotic behavior of the gap between the minimizer of the average square error and that of the Mallows-type criterion. Finally, we apply our theoretical results to a specific case of martingale difference sequence, namely the Auto-Regressive Conditional Heteroscedastic (ARCH(1)) process. A Monte-carlo simulation study, for this regression model with ARCH(1) process, is conducted.

Keywords. Nonparametric trend estimation, Kernel nonparametric models, Smoothing parameter selection, Martingale difference sequences, Average squared error, Mean average squared error, Mallows criterion, Cross validation, Generalized cross validation, ARCH(1).

2010 Mathematics Subject Classification. 62G08. 62G20. 60G10

*Laboratoire Jean Kuntzmann. Université Grenoble Alpes, 700 Avenue Centrale, 38401 Saint-Martin-d'Hères, France. E.mail: karim.benhenni@univ-grenoble-alpes.fr

†CNRS, Laboratoire Jean Kuntzmann. E.mail: didier.girard@univ-grenoble-alpes.fr

‡Corresponding author. Laboratoire Jean Kuntzmann, Université Grenoble Alpes, 700 Avenue Centrale, 38401 Saint-Martin-d'Hères, France. E.mail: sana.louhichi@univ-grenoble-alpes.fr.

Contents

1	Introduction	2
2	Model and notations	4
3	Main results and applications	7
3.1	Application to ARCH(1) processes	9
3.2	A Monte-carlo simulation study for a “trend plus ARCH(1) process”	10
4	Proofs	13
4.1	Main lemmas for the proof of Proposition 1	13
4.1.1	Proof of Lemma 3	14
4.1.2	Proof of Lemma 4	18
4.2	Proof of Proposition 1	20
4.3	Proof of Theorem 1	22
4.3.1	Step 1: convergence in distribution of $a_n \delta'_2(h_n)$	23
4.3.2	Step 2: convergence in probability of $a_n(\delta'_2(h_n) - \delta'_2(\hat{h}_n))$	25
4.3.3	Step 3: convergence in probability of $\frac{CL''(h^*)}{\mathbb{E}(T''_n(h_n))}$	25
4.3.4	End of the proof of Theorem 1	27
5	Acknowledgements	27
6	References	27
A	Auxiliary technical tools	30
A.1	Proof of Lemma 1	30
A.2	Other useful lemmas	33
B	Central limit theorem for some triangular arrays of a quadratic form of a stationary MDS	35
B.1	Control of the variance	36
B.2	Central limit theorem	45
C	Tools for martingale difference sequences	53

1 Introduction

This paper is about nonparametric regression model (known also as a machine learning function) which is used as a tool to describe and to analyse the trend between a response variable and one or more explanatory random variables. This subject was studied by several authors

since 1964 (Nadaraya, E. A. (1964), Watson, G. S. (1964)) and is still relevant, due to the fact that nonparametric regression has a lot of applications in different fields, such as economics, medicine, biology, physics, environment, social sciences, \dots , see for instance Hastie, T., Tibshirani, R. and Friedman, J. (2009).

Several estimate of the nonparametric regression function are proposed in the literature such as kernel smoothing, local polynomial regression, spline-based regression models, and regression trees (see for instance Hastie, T., Tibshirani, R. and Friedman, J. (2009)). In this paper, we are interested in kernel nonparametric estimations. These estimate depend on some smoothing parameter h which has to be chosen according to some criteria. For independent observations, two first criteria, to select h , are known as the Cross Validation (CV) criterion and its rotation-invariant version called Generalized Cross-Validation (GCV) criterion. The GCV criterion has different variants, see for instance Akaike, H. (1970), Craven, P. and Wahba, G. (1979), Shibata, R. (1981), Rice, J. (1984), Mallows, C. L. (1973). We refer the reader to Härdle, W., Hall, P. and Marron, J. S. (1988) who studied this problem in the case of independent, equally spaced, observations. They gave, in particular, the behaviors of the minimizers over h of the average squared errors, the mean average squared errors, the cross-validation score CV or the generalized cross-validation GCV. They also studied the deviation between these selected smoothing parameters.

In general, independence of the observations is, however, not a realistic modeling of observed data since, in practice, they are often correlated. Autoregressive models, autoregressive conditional heteroscedasticity models, Markov chains are examples of dependent models (see for instance Doukhan, P. and Louhichi, S. (1999)). We focus, in this paper, on the case of kernel nonparametric models with particular dependent errors, more precisely, the case when the errors form a stationary sequence of martingale difference random variables (MDS). They are, essentially, two reasons that motivated us to restrict our study of dependence to the case of stationary MDS.

- The first reason is that, studying MDS is a promising step for studying the general case of stationary dependent errors. In fact, MDS plays an important role in establishing the results for arbitrary stationary sequences, see for instance Peligrad, M., Utev, S. and Wu, W. B. (2007) (for moment inequalities purpose).
- The second reason is that MDS is not an abstract notion. Indeed, there are a lot of well known stationary MDS models which are used in practice, such as ARCH(1) or more general GARCH(1,1) stochastic volatility models.

We compare, in the case of nonparametric regression model with MDS errors, the behaviors of the smoothing bandwidths obtained by minimizing three criteria: the average squared error, the mean average squared error and a Mallows-type criterion adapted to our dependent case. We prove that these three minimizers are first-order equivalent. We give also a normal asymptotic

behavior of the gap between the minimizer of the average squared error and that of the Mallows-type criterion. The obtained results generalize those under independent errors, as in [Härdle, W., Hall, P. and Marron, J. S. \(1988\)](#), to MDS ones. Finally, we apply our results to a specific case of MDS namely the ARCH(1) processes.

The adaptation to the dependent case from the independent one is not trivial and needs to establish more theoretical and technical results such as maximal inequalities or limit theorems for quadratic forms of dependent data. To establish our theoretical results, we make use of some ingredients adapted to our case of dependent observations taken from [Burkholder, D. L. \(1988\)](#), [Doukhan, P. and Louhichi, S. \(1999\)](#), [McLeish, D. L. \(1974\)](#) and [Rio, E. \(1993\)](#). Our proofs are based, in particular, on maximal moment inequalities for quadratic forms for MDS that we establish using Burkholder-type moment inequalities together with some chaining arguments. Recall that chaining is a nice approach to approximate the supremum, over a non countable set, of stochastic processes (used in the theory of empirical processes see for instance [Andrews, D. W. K. and Pollard, D. \(1994\)](#), [Louhichi, S. \(2000\)](#), or [Pollard, D. \(1990\)](#)). A central limit theorem for triangular arrays of quadratic forms for MDS is also needed for the proof of our results. We prove this central limit theorem by checking the technical conditions of [McLeish, D. L. \(1974\)](#).

Our paper is organized as follows. In [Section 2](#), we introduce the regression model and the different criteria for the selection of the smoothing parameter h . In [Section 3](#), we state our main results. We apply our theoretical results, in [Subsection 3.1](#), to ARCH(1) processes. A Monte-carlo simulation study is conducted in [Subsection 3.2](#). The proofs of our results are given in [Section 4](#). [Appendix A](#) and [B](#) are dedicated to the proofs of the main tools needed to establish our main results. [Appendix C](#) gives and proves some ingredients for MDS used throughout the proofs of the main results.

2 Model and notations

Let $(\epsilon_i)_{i \geq 0}$ be a stationary sequence of centered random variables with finite second moment. Let $\sigma^2 = \text{Var}(\epsilon_1)$ and R be the correlation matrix of the vector $(\epsilon_1, \dots, \epsilon_n)$. Consider the following regression model, defined for $i = 1, \dots, n$, by

$$Y_i = r(x_i) + \epsilon_i, \quad x_i = \frac{i}{n}, \tag{1}$$

where r is an unknown regression function of class \mathcal{C}^2 and the x_i 's are equally spaced fixed design. We are interested in this paper by the Priestley-Chao estimator of r defined, for $x \in \mathbb{R}$, by

$$\hat{r}(x) = \sum_{i=1}^n l_i(x) Y_i, \quad \text{with} \quad l_i(x) = \frac{1}{nh} K\left(\frac{x - x_i}{h}\right),$$

where K is a compactly supported even kernel with class $\mathcal{C}^1([-1, 1])$ and h is a positive bandwidth less than $1/2$. The above curve estimator entails the following smoothing, in the matrix form,

$$\hat{r} = LY,$$

with

$$\hat{r} = (\hat{r}(x_1), \dots, \hat{r}(x_n))^t, \quad Y = (Y_1, \dots, Y_n)^t$$

and $L = (l_j(x_i))_{1 \leq i, j \leq n}$ is known as the smoothing matrix or the hat matrix. Since the estimator \hat{r} depends on some smoothing parameter h , we will need some procedure for choosing h . For this, we recall some known criteria of selecting this parameter h .

In order to eliminate the boundary effects of the compactly supported kernel K , we introduce, as was done in the literature (see for instance [Gasser, T. and Muller, H.G. \(1979\)](#)), a known function supported on a subinterval of the unit interval. For this, suppose without loss of generality that $h < \epsilon$ where ϵ is a fixed positive real number less than $1/2$. Let $u := u_\epsilon$ be a positive function, of class \mathcal{C}^1 and $[\epsilon, 1 - \epsilon]$ -compactly supported.

Define the average squared error

$$T_n(h) = \frac{1}{n} \sum_{i=1}^n u(x_i) (\hat{r}(x_i) - r(x_i))^2 = \frac{1}{n} \|U^{1/2}(\hat{r} - r)\|^2,$$

where U is the diagonal matrix $U = \text{diag}(u(x_1), \dots, u(x_n))$ and for any vector v , $\|v\|^2 = v^t v$. The following lemma (its proof is given in [Appendix A.1](#)) evaluates its mean, $\mathbb{E}(T_n(h))$, for finite variance of stationary errors $(\epsilon_i)_{i \in \mathbb{N}}$.

Lemma 1. *Suppose that $\sum_{k=1}^{\infty} k |\text{Cov}(\epsilon_0, \epsilon_k)| < \infty$. Define,*

$$\begin{aligned} D_n(h) &= \frac{h^4}{4} \int_0^1 u(x) r''^2(x) dx \left(\int_{-1}^1 t^2 K(t) dt \right)^2 \\ &+ \frac{1}{nh} \left(\int_0^1 u(x) dx \right) \int_{-1}^1 K^2(y) dy \left(\sigma^2 + 2 \sum_{k=1}^{\infty} \text{Cov}(\epsilon_0, \epsilon_k) \right). \end{aligned}$$

Then for any $n \geq 1$ and $h \in]0, \epsilon[$,

$$\mathbb{E}(T_n(h)) = D_n(h) + O\left(\frac{1}{n}\right) + o(h^4) + O\left(\frac{1}{n^2 h^4}\right) + \frac{\gamma(h)}{nh},$$

where O is uniformly on n and h , $\gamma(h)$ depends on h (but not on n) and tends to 0 when h tends to 0.

Let $h_n^* = \text{argmin}_{h>0} D_n(h)$. Clearly, if $\int_0^1 u(x) r''^2(x) dx \neq 0$ then

$$h_n^* = n^{-1/5} \left(\frac{\left(\int_0^1 u(x) dx \right) \int_{-1}^1 K^2(y) dy \left(\sigma^2 + 2 \sum_{k=1}^{\infty} \text{Cov}(\epsilon_0, \epsilon_k) \right)}{\int_0^1 u(x) r''^2(x) dx \left(\int_{-1}^1 t^2 K(t) dt \right)^2} \right)^{1/5} =: cn^{-1/5}.$$

Let, as in [Hall, P., Lahiri, S. N. and Polzehl, J. \(1995\)](#) and [Rice, J. \(1984\)](#), H_n be a neighborhood of h_n^* , i.e, $H_n = [an^{-1/5}, bn^{-1/5}]$ for some fixed $a < c < b$. Define also,

$$h_n = \operatorname{argmin}_{h \in H_n} \mathbb{E}(T_n(h)) \quad \text{and} \quad \hat{h}_n = \operatorname{argmin}_{h \in H_n} T_n(h).$$

Of course these three “optimal” parameters h_n , h_n^* and \hat{h}_n depend on the unknown function r , since the criteria that they respectively minimise, depend themselves on the regression function r . Many authors agree that, among these ones, \hat{h}_n should be the target (see [Girard, D. \(1998\)](#), page 316). For this reason, an important literature considered minimizers of “good” estimators of $T_n(h)$ and studied their asymptotic behavior.

For i.i.d. errors $(\epsilon_i)_{1 \leq i \leq n}$ with all finite moments, this question is solved. A nearly unbiased estimate of $\mathbb{E}(T_n(h))$ is constructed allowing to define a criterion that selects an observable choice for h . The first criterion is known as the Cross Validation criterion which was extended to Generalized Cross-Validation (GCV) criteria. All those different forms of the GCV criteria are second order equivalent (in the sense defined below as a footnote) to the following:

$$C_p := C_p(h) = \frac{1}{\sum_{i=1}^n u(x_i)} \sum_{i=1}^n u(x_i) (Y_i - \hat{r}(x_i))^2 + 2 \frac{\nu}{n} \hat{\sigma}^2, \quad (2)$$

where,

$$\hat{\sigma}^2 = \frac{1}{\sum_{i=1}^n u(x_i)} \sum_{i=1}^n u(x_i) (Y_i - \hat{r}(x_i))^2, \quad \nu := n \frac{\operatorname{tr}(UL)}{\operatorname{tr}(U)} = \frac{1}{h} K(0).$$

The above notation C_p (where ν is the “weighted” degree of freedom) is related to the C_p -statistics introduced by [Mallows, C. L. \(1973\)](#) for variable selection in linear regression models. Let \hat{h} be a minimizer over $h \in H_n$ of the function C_p . [Härdle, W., Hall, P. and Marron, J. S. \(1988\)](#) proved, in the context of i.i.d errors $(\epsilon_i)_{1 \leq i \leq n}$ with all finite moments, that $\hat{h}, h_n^*, \hat{h}_n, h_n$ are all equivalent in probability and that $\hat{h} - \hat{h}_n, h_n - \hat{h}_n$ are also close in distribution as n tends to infinity.¹

The above criteria can hardly be considered a priori as adapted to the case of dependent errors since they take into account only the variance σ^2 of the errors and not their overall dependence structure. Several authors extended Mallow’s criterion to some cases of stationary dependent errors. [Wang, Y. \(1988\)](#) and [Han, Ch. and Gu, Ch. \(2008\)](#), among others, generalized Mallow’s criteria in (2) (but for other purposes than ours) to stationary dependent errors with known covariance matrix $\sigma^2 R$ of the vector $(\epsilon_1, \dots, \epsilon_n)^t$, by

$$\operatorname{CL}(h) = n^{-1} \|U^{1/2}(I - L)Y\|^2 + 2\sigma^2 n^{-1} \operatorname{tr}(URL), \quad (3)$$

which is linked to the average squared error $T_n(h)$ due to the following relation,

$$\operatorname{CL}(h) = T_n(h) + \delta_2(h) + n^{-1} \|U^{1/2}(Y - r)\|^2,$$

¹The second order equivalence means that the asymptotic law of $\hat{h} - \hat{h}_n$ is unchanged if C_p is replaced by GCV.

where

$$\delta_2(h) = 2n^{-1}(Y - r)'U(r - \hat{r}) + 2\sigma^2 n^{-1} \text{tr}(URL). \quad (4)$$

Let us consider, according to our purpose, \hat{h}_M to be the minimizer of the dependent version of the Mallows criteria (3),

$$\hat{h}_M = \operatorname{argmin}_{h \in H_n} \text{CL}(h).$$

Recall that we are interested in the problem of selecting the parameter h when the errors form a sequence of stationary and dependent random variables. As we mentioned in the introduction, we consider through all this paper, the above regression model with stationary MDS errors (defined in Assumptions (B) of Section 3 below). Since MDS is a sequence of non-correlated and centered random variables, R , which represents the correlation of the errors, is nothing else but the identity matrix. We restrict ourselves to the case where the variance of the errors σ^2 is known as is done in several works (see for instance Wang, Y. (1988)).

3 Main results and applications

The following two assumptions are required to establish our main results.

Assumptions (A). Suppose that both the functions $h \mapsto T_n(h)$ and $h \mapsto \text{CL}(h)$ have continuous first derivatives, that $T'_n(\hat{h}_n) = 0$ and $\text{CL}'(\hat{h}_M) = 0$ almost surely. Suppose also that the function $h \mapsto \mathbb{E}(T_n(h))$ is twice differentiable with continuous second derivative and that $\frac{\partial^2}{\partial h^2} \mathbb{E}(T_n(h)) = \mathbb{E}(T''_n(h))$.

Assumptions (B). Assume that the errors $(\epsilon_i)_{i \geq 0}$ form a stationary MDS with respect to some natural filtration $(\mathcal{F}_i)_{i \geq 1}$, i.e, for any $i > 0$, ϵ_i is \mathcal{F}_i -measurable and $\mathbb{E}(\epsilon_i | \mathcal{F}_{i-1}) = 0$. Suppose also that $\mathbb{E}(\epsilon_1^{2p}) < \infty$ for some $p > 8$.

Our first result states that for MDS errors, the bandwidths h_n, h_n^*, \hat{h}_n and \hat{h}_M are first-order equivalent in probability (and the CL enjoys the "asymptotic optimality" property).

Proposition 1. *Suppose that Assumptions (A) and (B) are satisfied. Then*

$$\frac{h_n^*}{h_n}, \frac{\hat{h}_n}{h_n}, \frac{\hat{h}_M}{h_n}$$

all converge in probability to 1 as n tends to infinity.

Notice that Hall, P., Lahiri, S. N. and Polzehl, J. (1995) gave two theorems for two bandwidth selection methods (precisely a block-bootstrap method and the classical leave- k -out technique) under a rather general dependence assumption on the error sequence, namely the Rosenblatt

mixing condition (see their Section 2.2). Each of these two theorems is a first-order optimality like Proposition 1 above, and it could be applied, in particular, to certain stationary MD sequences. However we point out that these two theorems also require that all moments of the marginal law of the errors are finite. Thus the results of Hall, P., Lahiri, S. N. and Polzehl, J. (1995) cannot be applied to any ARCH process except the trivial one ($\alpha = 0$ in the notation of Section 3.1).

Our second result gives, under an additional dependence condition, the rate at which $\hat{h}_n - \hat{h}_M$ converges in distribution to a centered normal law, and furthermore states that the martingale difference dependence doesn't impact $\hat{h}_n - \hat{h}_M$ up to second-order.

Theorem 1. *Suppose that Assumptions (A) and (B) are satisfied. Moreover, suppose that there exists a positive decreasing function Φ defined on \mathbb{R}^+ satisfying*

$$\sum_{s=1}^{\infty} s^4 \Phi(s) < \infty,$$

and for any positive integer $q \leq 6$, $1 \leq i_1 \leq \dots \leq i_k < i_{k+1} \leq \dots \leq i_q \leq n$ such that $i_{k+1} - i_k \geq \max_{1 \leq l \leq q-1} (i_{l+1} - i_l)$,

$$|\text{Cov}(\epsilon_{i_1} \dots \epsilon_{i_k}, \epsilon_{i_{k+1}} \dots \epsilon_{i_q})| \leq \Phi(i_{k+1} - i_k). \quad (5)$$

Then

$$n^{3/10}(\hat{h}_n - \hat{h}_M)$$

converges in distribution to a centered normal law with variance Σ^2 given by

$$\begin{aligned} \Sigma^2 &= \frac{4\sigma^{6/5}}{5^2 A^{8/5} B^{2/5}} \left(\int t^2 K(t) dt \right)^2 \int_0^1 u^2(x) r''^2(x) dx \\ &+ \frac{8\sigma^{6/5}}{5^2 A^{3/5} B^{7/5}} \int_0^1 u^2(x) dx \int (K - G)^2(u) du, \end{aligned}$$

where $\sigma^2 = \mathbb{E}(\epsilon_1^2)$, G is the function defined for any $x \in \mathbb{R}$ by $G(x) = -xK'(x)$ and

$$A = \int_0^1 u(x) r''^2(x) dx \left(\int t^2 K(t) dt \right)^2, \quad B = \int_0^1 u(x) dx \int K^2(t) dt.$$

Remarque 1. *The control of the covariance quantity $|\text{Cov}(\epsilon_{i_1} \dots \epsilon_{i_k}, \epsilon_{i_{k+1}} \dots \epsilon_{i_q})|$ appearing in (5) is well known in the literature. It was used, for instance in Doukhan, P. and Louhichi, S. (1999), in order to obtain Marcinkiewicz-Zygmund type moments inequalities of an even order of the partial sum $\sum_{i=1}^n \epsilon_i$. If the sequence $(\epsilon_n)_n$ is strongly mixing with mixing coefficients $(\alpha_s)_{s \in \mathbb{N}}$, then it is proved (Rio, E. (1993), see also Lemma 9 in Doukhan, P. and Louhichi, S. (1999)) that, for $1 \leq i_1 \leq \dots \leq i_k < i_{k+1} \leq \dots \leq i_q \leq n$ such that $s := i_{k+1} - i_k \geq \max_{1 \leq l \leq q-1} (i_{l+1} - i_l)$,*

$$|\text{Cov}(\epsilon_{i_1} \dots \epsilon_{i_k}, \epsilon_{i_{k+1}} \dots \epsilon_{i_q})| \leq 4 \int_0^{\alpha_s} Q^q(u) du,$$

where Q is the quantile function of $|\epsilon_1|$, i.e. the inverse of the tail function $t \mapsto \mathbb{P}(|\epsilon_1| > t)$.

3.1 Application to ARCH(1) processes

We consider the regression model defined in (1) with an ARCH(1) error process $(\epsilon_n)_{n \geq 1}$ defined, for $n \geq 1$, by the following stochastic difference equation,

$$\epsilon_n = \eta_n \sqrt{\sigma^2(1 - \alpha) + \alpha \epsilon_{n-1}^2}, \quad 0 \leq \alpha < 1, \quad \sigma^2 > 0 \quad (6)$$

where $(\eta_n)_{n \geq 1}$ is an i.i.d. centered sequence distributed as a standard normal law and such that η_n is independent of $(\epsilon_1, \dots, \epsilon_{n-1})$.

Proposition 2. *Let $(\epsilon_n)_{n \geq 1}$ be a strictly stationary ARCH(1) process as defined in (6) with α such that $\alpha^8 \prod_{i=1}^8 (2i - 1) < 1$ (this is equivalent to $\alpha < 0.162796$). Suppose that Assumptions (A) are satisfied. Then the conclusions of Proposition 1 and Theorem 1 hold.*

Proof of Proposition 2. We first recall the following well known properties in the literature (see for instance Engle, R. F. (1982), Lindner, A.M. (2009) and the references therein).

Lemma 2. *Consider the process $(\epsilon_n)_n$ as defined in (6). Then*

1. $(\epsilon_n)_n$ is a geometric ergodic homogeneous Markov chain with a unique stationary distribution π . The stationary distribution π is continuous and symmetric.
2. $(\epsilon_n)_n$ is strongly mixing with mixing coefficients $(\alpha_l)_{l > 0}$

$$\alpha_l := \sup_{A \in \sigma(\epsilon_s, s \leq 0), B \in \sigma(\epsilon_s, s \geq l)} |\text{Cov}(\mathbf{1}_A, \mathbf{1}_B)| = O(\rho^l),$$

for some $\rho \in]0, 1[$.

3. $\mathbb{E}(\epsilon_1^{2r}) < \infty$, for $r \in \mathbb{N}^*$, if and only if $\alpha^r \prod_{i=1}^r (2i - 1) < 1$.
4. $\mathbb{P}(|\epsilon_1| > x) \sim cx^{-\kappa}$ as x tends to infinity², for some $c > 0$ and κ is given as the unique positive solution to $\alpha^{\kappa/2} \mathbb{E}(|\eta_1|^\kappa) = 1$.

Letting $\mathcal{F}_i = \sigma(\eta_1, \dots, \eta_i)$, then ϵ_i is \mathcal{F}_i -measurable and

$$\mathbb{E}(\epsilon_i | \mathcal{F}_{i-1}) = \sqrt{\sigma^2(1 - \alpha) + \alpha \epsilon_{i-1}^2} \mathbb{E}(\eta_i | \mathcal{F}_{i-1}) = 0.$$

The sequence $(\epsilon_n)_n$ is then a martingale-difference. Since it is strongly mixing with $\alpha_s \leq C\rho^s$, we get, from Remark 1, (5) with

$$\Phi(s) = 4 \int_0^{\alpha_s} Q^q(u) du \leq 4 \int_0^{C\rho^s} Q^q(u) du.$$

²In all this paper the notation $a(x) \sim b(x)$ means that $\lim_{x \rightarrow \infty} \frac{a(x)}{b(x)} = 1$.

We deduce from $\mathbb{P}(|\epsilon_1| > x) \sim cx^{-\kappa}$ as x tends to infinity that $Q(u) = O(u^{-1/\kappa})$, $q < \kappa$ (since $\mathbb{E}(|\epsilon_1|^q) < \infty$) and

$$\int_0^{C\rho^s} Q^q(u)du = O(\rho^{s(1-q/\kappa)}).$$

Consequently, for some positive constant C ,

$$\sum_{s=1}^{\infty} s^4 \Phi(s) \leq 4 \sum_{s=1}^{\infty} s^4 \int_0^{C\rho^s} Q^q(u)du \leq C \sum_{s=1}^{\infty} s^4 \rho^{s(1-q/\kappa)} < \infty, \text{ since } \rho \in]0, 1[.$$

3.2 A Monte-carlo simulation study for a “trend plus ARCH(1) process”

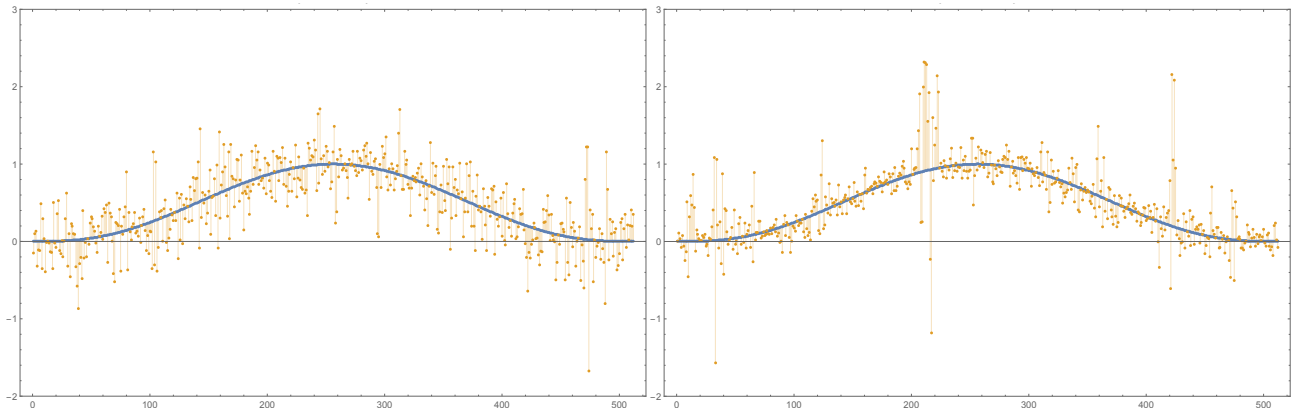


Figure 1 : $n = 2^9$. Each of these 2 panels displays one data set Y and the “smooth” deterministic trend $r(x)$. The 2 panels only differ by $\alpha = 0.577$ (left) and $\alpha = 0.9$ (right)

For our experiments, we consider an example of regression function, the “deterministic trend” here, which is very often used (e.g. [Rice, J. \(1984\)](#), [Härdle, W., Hall, P. and Marron, J. S. \(1988\)](#)) and a noise level for which the noise-to-signal ratio is “moderate” (precisely 0.32^2). So, the chosen trend function is

$$r(x) = (4x(1-x))^3,$$

and we use an equispaced design for simulating data sets with an ARCH(1) noise as defined in the above Subsection 3.1. That is, each data set is the sum of this trend $r(x)$ evaluated at $x_i = i/n, i = 1, \dots, n$, plus an ARCH(1) sequence with a “persistence” parameter α . We consider 6 settings for the ARCH(1) noise, precisely

$$\alpha \in \{0.01, 0.162, 0.577, 0.75, 0.9, 0.98\}$$

with a common value $\sigma = 0.32$. The first value of α nearly corresponds to i.i.d. normal observation noises (this setting will be referred to as the “quasi-iid-normal” case) and the last one generates noise sequences for which a strong serial correlation is always present when the sequence is squared. Recall that the moment of order 16 no longer exists as soon as α is slightly above 0.162, but the moment of order 4 still exists for $\alpha < \sqrt{1/3} \approx 0.57735$.

The kernel function we used here is the classical biweight $K(x) = (15/8)(1 - 4x^2)^2 1_{[-.5,.5]}(x)$. As is well known, its precise specification, among possible positive “bell shaped” kernels, has a weak impact on the behavior of bandwidth selection techniques.

The data sets size n was chosen in $\{2^9, 2^{12}, 2^{15}\} = \{512, 4096, 32768\}$. We generated 1000 replicated data sets for each of these 3×6 settings. For each data set, the minimizer of $T_n(h)$ and the one of $CL(h)$ were numerically computed by a simple grid-search over the domain $[0.019, 1.30]$, the grid-step being chosen fine enough so that the “granularity” in the 2000 computed h 's has a very weak impact on the conclusions.

As in [Rice, J. \(1984\)](#) and [Härdle, W., Hall, P. and Marron, J. S. \(1988\)](#), a periodic version of the Priestley-Chao kernel estimate was considered (see e.g. [Härdle, W., Hall, P. and Marron, J. S. \(1988\)](#) for details). This is appropriate because r is “smoothly” periodic with period 1. We used $u(t) \equiv 1$ as a weight function, which is possible in this periodic setting. Notice that the computational cost of the simulation study (CPU time) is then greatly reduced by using fast Fourier transforms.

The “a.o.” property. First, let us analyze the asymptotic optimality (a.o.) result. As is well known, a result like [Proposition 1](#) generally stems from a uniform relative accuracy result which states that $CL(h) - n^{-1} \|U^{1/2}(Y - r)\|^2$ uniformly approximates $T_n(h)$ (or its expectation $MASE(h)$) with a small (in probability and in *sup* norm over the domain of candidate h 's) error, “small” being defined relatively to $MASE(h)$.

We resume in [Figure 2](#) that a uniform relative accuracy is well observed and, above all, this accuracy in the case $\alpha = 0.162$ is of the same order as the accuracy observed in the quasi-iid-normal case ($\alpha = 0.01$). Furthermore, an interesting observation is that this accuracy is not deteriorated when $\alpha = 0.577$. However there is clearly a deterioration for larger α , especially for $\alpha = 0.98$.

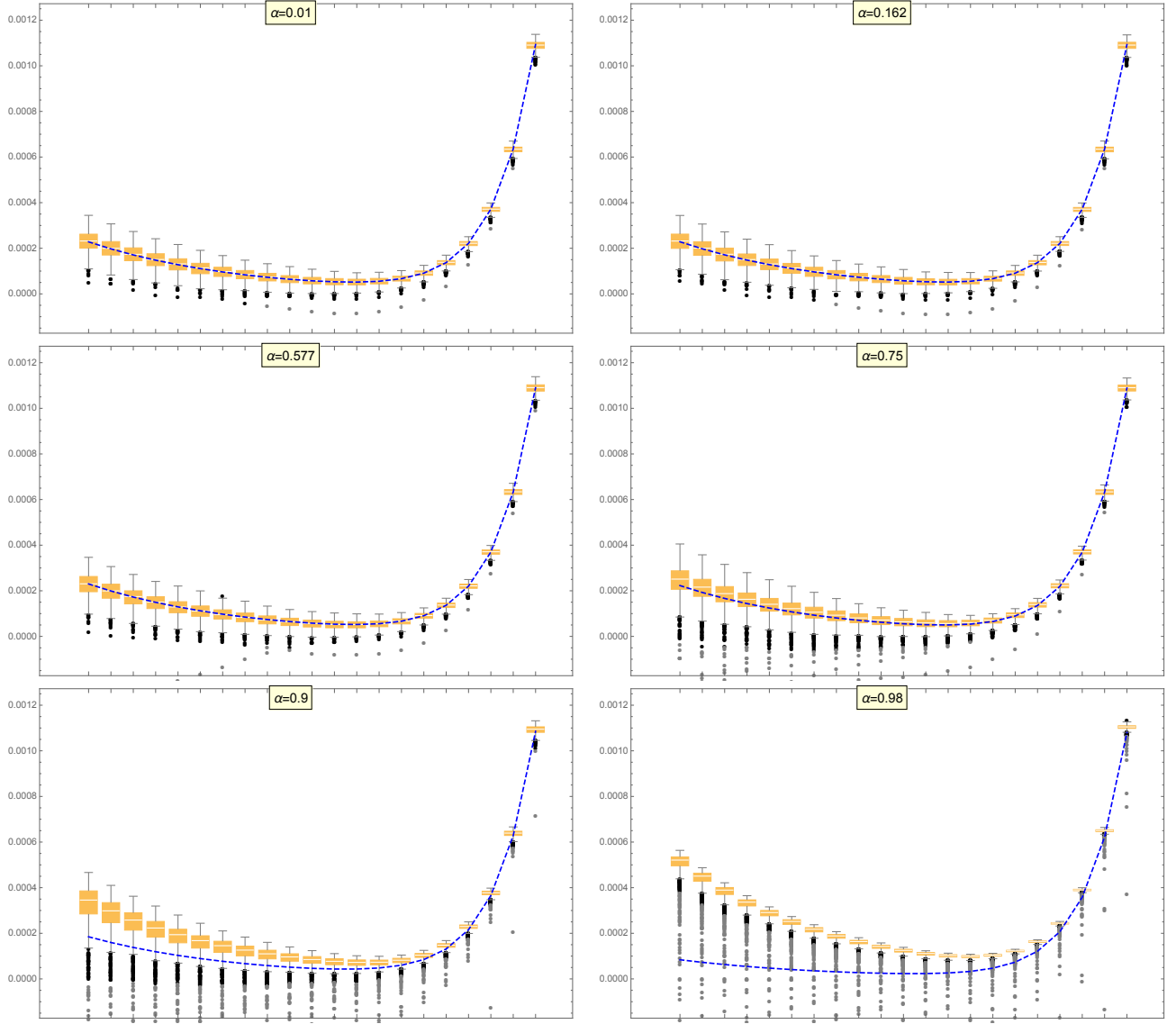


Figure 2 : $n = 2^{15}$. These 6 panels only differ by α varying in $\{0.01, 0.162, 0.577, 0.75, 0.9, 0.98\}$. In each panel, the dashed blue curve is “empirical MASE”, precisely the average (over the 1000 replicates) of the $T_n(h)$ curves. Each of the 21 boxplots (located at 21 discrete values for h) are built from the 1000 replicates of $\text{CL}(h) - n^{-1}\|U^{1/2}(Y - r)\|^2$.

It may be thus conjectured that, at least for ARCH(1) processes, the restriction $p > 8$ of our Assumptions (B) could be weakened to $p > 2$. However, the poor behavior of CL (even with quite large n) in cases $\alpha = 0.75, 0.9$ or 0.98 , leads us to conjecture that $p > 2$ should be considered as a necessary condition for the a.o. of CL under general MDS observation errors.

Asymptotic normal distribution. Now, let us look at the usefulness of the asymptotic normal approximation stated in Theorem 1. By inspecting Figure 3, we clearly see, in the left-bottom panel, that this asymptotic approximation fits very well for $n = 2^{15}$ and $\alpha = 0.577$.

We have also made such a comparison for $\alpha = 0.01$ and $\alpha = 0.162$ (not shown in Figure 3), and, as expected by Theorem 1 and Section 3.1, the fit is also very good. For settings with the much smaller $n = 512$, the fit is still surprisingly good for $\alpha = 0.577$, but this is no longer true for $\alpha = 0.75$ or $\alpha = 0.9$ (and the fit is even worse for $\alpha = 0.98$, not shown in Figure 3). Notice that, as expected, the range of the abscissae (h -differences) increases by moving from $n = 2^{15}$ (bottom) to $n = 2^9$ (top). It is good news that Theorem 1 is thus useful also with $\alpha = 0.577$, since this gives support to the conjecture that Theorem 1 could be extended to an ARCH process under the only existence of the fourth moment of the marginal law.

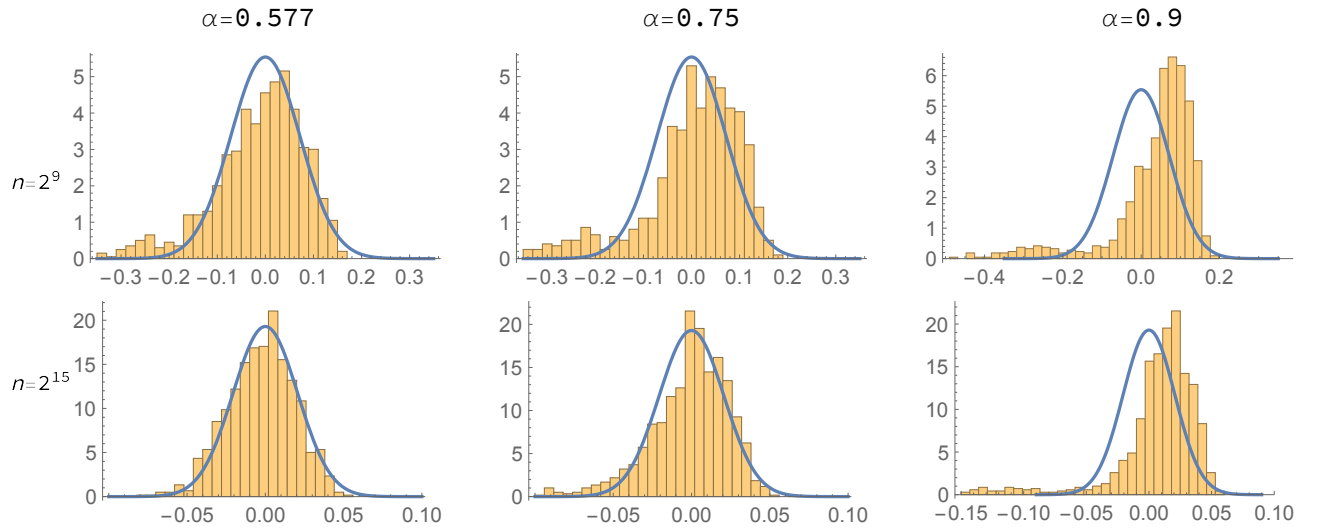


Figure 3 : $n = 2^9$ (top) and $n = 2^{15}$ (bottom). These 6 panels only differ by n and by α varying in $\{0.577, 0.75, 0.9\}$. In each panel, the displayed histogram is that of the 1000 replicates of $\hat{h}_M - \hat{h}_n$; the histograms are normalized so that their integrals are equal to 1. The superposed blue curve is the normal distribution of $\hat{h}_M - \hat{h}_n$ predicted by the asymptotic theory.

4 Proofs

4.1 Main lemmas for the proof of Proposition 1

The following two lemmas are very useful for the proof of Proposition 1. We denote by $\|\cdot\|_p$ the p -norm, i.e, for a random variable X , $\|X\|_p = (\mathbb{E}(|X|^p))^{1/p}$ and we recall that $\delta_2(h)$ is defined as in (4) and $H_n = [an^{-1/5}, bn^{-1/5}]$ for some fixed $a < c < b$.

Lemma 3. *It holds, for $p > 8$,*

$$\lim_{n \rightarrow \infty} \left\| \sup_{h \in H_n} nh |\delta_2(h)| \right\|_p = 0, \quad (7)$$

$$\lim_{n \rightarrow \infty} \left\| \sup_{h \in H_n} nh^3 |\delta_2''(h)| \right\|_p = 0, \quad (8)$$

$$\lim_{M \rightarrow \infty} \limsup_{n \rightarrow \infty} \mathbb{P} \left(\sup_{h \in A_\epsilon} \sqrt{n} |\delta_2''(h)| \geq M \right) = 0, \quad (9)$$

where for fixed $\epsilon > 0$, A_ϵ is a subset of H_n defined by,

$$A_\epsilon = \left\{ h \in H_n, \left| \frac{h}{h_n} - 1 \right| \leq \epsilon \right\}.$$

Lemma 4. *It holds, for $p > 8$,*

$$\lim_{n \rightarrow \infty} \left\| \sup_{h \in H_n} nh |T_n(h) - \mathbb{E}(T_n(h))| \right\|_p = 0, \quad (10)$$

$$\lim_{n \rightarrow \infty} \left\| \sup_{h \in H_n} nh^3 |T_n''(h) - \mathbb{E}(T_n''(h))| \right\|_p = 0. \quad (11)$$

4.1.1 Proof of Lemma 3

We have the following decomposition

$$\begin{aligned} \delta_2(h) &= (2n^{-1}(Y - r)'U(r - \hat{r}) + 2\sigma^2 n^{-1} \text{tr}(URL)) \\ &= 2 \sum_{i=1}^n A_i(h) u(x_i) \epsilon_i + 2 \sum_{i=1}^n \sum_{j=1}^{i-1} (u(x_i) + u(x_j)) B_{i,j}(h) \epsilon_i \epsilon_j \\ &\quad + 2 \sum_{i=1}^n u(x_i) B_{i,i}(h) (\epsilon_i^2 - \mathbb{E}(\epsilon_i^2)), \end{aligned}$$

with,

$$\begin{aligned} A_i(h) &= \frac{1}{n} (r(x_i) - \mathbb{E}(\hat{r}(x_i))) = -\frac{1}{n} B(x_i, h) \\ B_{i,j}(h) &= -n^{-1} \frac{1}{nh} K\left(\frac{x_i - x_j}{h}\right), \quad B_{i,i}(h) = -n^{-1} \frac{1}{nh} K(0). \end{aligned}$$

Proof of (7). We have, for any $h, h' \in H_n$, (using the same calculations yielding to (34)),

$$\begin{aligned} u(x_i) \sup_{h \in H_n} nh |A_i(h)| &= O(n^{-3/5}), \\ u(x_i) |nh A_i(h) - nh' A_i(h')| \\ &\leq hu(x_i) |B(x_i, h) - B(x_i, h')| + u(x_i) B(x_i, h') |h - h'| \leq cst n^{-2/5} |h - h'| \end{aligned}$$

and,

$$\begin{aligned} nh|B_{i,j}(h)| &\leq cst n^{-1} \mathbb{1}_{|i-j|\leq nh}, \\ |nhB_{i,j}(h) - nh'B_{i,j}(h')| &\leq cst n^{-4/5} |h - h'| \mathbb{1}_{|i-j|\leq n \max(h,h')} \\ nhB_{i,i}(h) &= \frac{K(0)}{n}. \end{aligned}$$

Using Lemmas 14, 16 and the fact that

$$\frac{1}{n} \sum_{i=1}^n u(x_i) (\epsilon_i^2 - \mathbb{E}(\epsilon_i^2)) \rightarrow 0, \text{ in probability as } n \rightarrow \infty,$$

we have

$$\lim_{n \rightarrow \infty} \sup_{h \in H_n} nh |\delta_2(h)| = 0, \text{ in probability as } n \rightarrow \infty.$$

Proof of (8). For any $h \in]0, \epsilon[$ and any $n \geq 1$,

$$\begin{aligned} \delta_2''(h) &= 2 \sum_{i=1}^n c_i(h) u(x_i) \epsilon_i + 2 \sum_{i=1}^n \sum_{j=1}^{i-1} (u(x_i) + u(x_j)) c_{i,j}(h) \epsilon_i \epsilon_j \\ &\quad + 2 \sum_{i=1}^n c_{i,i}(h) u(x_i) (\epsilon_i^2 - \mathbb{E}(\epsilon_i^2)), \end{aligned}$$

where, letting $K_1 = K - G$ and $G_1(u) = -uK_1'(u)$,

$$\begin{aligned} c_i(h) &= -\frac{1}{n} \frac{\partial^2}{\partial h^2} \mathbb{E}(\hat{r}(x_i)) = -\frac{1}{n} \frac{\partial^2}{\partial h^2} B(x_i, h), \\ c_{i,j}(h) &= -\frac{2}{n^2 h^3} K_1\left(\frac{x_i - x_j}{h}\right) + \frac{1}{n^2 h^3} G_1\left(\frac{x_i - x_j}{h}\right), \\ c_{i,i}(h) &= \frac{-2}{n^2 h^3} K(0). \end{aligned}$$

Now,

$$nh^3 \sum_{i=1}^n c_{i,i}(h) u(x_i) (\epsilon_i^2 - \mathbb{E}(\epsilon_i^2)) = cst \frac{1}{n} \sum_{i=1}^n u(x_i) (\epsilon_i^2 - \mathbb{E}(\epsilon_i^2))$$

which converges in probability to 0 by an analogous to Lemma 6.

We also have, for any $h, h' \in H_n$, (see Lemma 10) the following bounds,

$$\begin{aligned} u(x_i) |nh^3 c_i(h)| &\leq cst n^{-3/5}, \\ u(x_i) |nh^3 c_i(h) - nh'^3 c_{i,n}(h')| &\leq cst n^{-2/5} |h - h'|, \\ u(x_i) |nh^3 c_{i,j}(h)| &\leq cst \frac{1}{n} \mathbb{1}_{|i-j|\leq nh}, \\ u(x_i) |nh^3 c_{i,j}(h) - nh'^3 c_{i,j}(h')| &\leq cst n^{-4/5} |h - h'| \mathbb{1}_{|i-j|\leq n \max(h,h')}. \end{aligned}$$

All the requirements of Lemmas 14 and 16 are satisfied. We deduce that,

$$\lim_{n \rightarrow \infty} \sup_{h \in H_n} nh^3 |\delta_2''(h)| = 0, \text{ in probability.}$$

Proof of (9). We have,

$$\begin{aligned} \sqrt{n} \delta_2''(h) &= 2 \sum_{i=1}^n e_i(h) u(x_i) \epsilon_i + 2 \sum_{i=1}^n \sum_{j=1}^{i-1} (u(x_i) + u(x_j)) e_{i,j}(h) \epsilon_i \epsilon_j \\ &+ 2 \sum_{i=1}^n e_{i,i}(h) u(x_i) (\epsilon_i^2 - \mathbb{E}(\epsilon_i^2)), \end{aligned}$$

where,

$$\begin{aligned} e_i(h) &= -\frac{1}{\sqrt{n}} \frac{\partial^2}{\partial h^2} \mathbb{E}(\hat{r}(x_i)) = -\frac{1}{\sqrt{n}} \frac{\partial^2}{\partial h^2} B(x_i, h), \\ e_{i,j}(h) &= -\frac{2}{n\sqrt{n}h^3} K_1\left(\frac{x_i - x_j}{h}\right) + \frac{1}{n\sqrt{n}h^3} G_1\left(\frac{x_i - x_j}{h}\right), \\ e_{i,i}(h) &= \frac{-2}{n\sqrt{n}h^3} K(0). \end{aligned}$$

We have, since $\sum_{i=1}^{\infty} |\text{Cov}(\epsilon_1^2, \epsilon_i^2)| < \infty$,

$$\begin{aligned} \sup_{h \in H_n} \left| \sum_{i=1}^n e_{i,i}(h) u(x_i) (\epsilon_i^2 - \mathbb{E}(\epsilon_i^2)) \right| &\leq \frac{cst}{n^{9/10}} \left| \sum_{i=1}^n u(x_i) (\epsilon_i^2 - \mathbb{E}(\epsilon_i^2)) \right| \\ \left\| \sup_{h \in H_n} \left| \sum_{i=1}^n e_{i,i}(h) u(x_i) (\epsilon_i^2 - \mathbb{E}(\epsilon_i^2)) \right| \right\|_2^2 &\leq \frac{cst}{n^{18/10}} \left\| \sum_{i=1}^n u(x_i) (\epsilon_i^2 - \mathbb{E}(\epsilon_i^2)) \right\|_2^2 \\ &\leq cst \frac{n}{n^{18/10}}. \end{aligned}$$

Hence,

$$\lim_{n \rightarrow \infty} \left\| \sup_{h \in H_n} \left| \sum_{i=1}^n e_{i,i}(h) u(x_i) (\epsilon_i^2 - \mathbb{E}(\epsilon_i^2)) \right| \right\|_2 = 0. \quad (12)$$

Let, for $h \in H_n, h' \in H_n, \tilde{e}_i(h) = e_i(h) - e_i(h_n)$, then

$$\begin{aligned} u(x_i) |\tilde{e}_i(h)| &\leq cst \frac{1}{\sqrt{n}} |h - h_n| \leq cst n^{-7/10}, \\ u(x_i) |\tilde{e}_i(h) - \tilde{e}_i(h')| &\leq cst \frac{1}{\sqrt{n}} |h - h'|. \end{aligned}$$

Hence,

$$\sup_{h \in H_n} \left| \sum_{i=1}^n e_i(h) u(x_i) \epsilon_i \right| \leq \sup_{h \in H_n} \left| \sum_{i=1}^n (e_i(h) - e_i(h_n)) u(x_i) \epsilon_i \right| + \left| \sum_{i=1}^n e_i(h_n) u(x_i) \epsilon_i \right|.$$

It follows that

$$\begin{aligned} & \left\| \sup_{h \in H_n} \left\| \sum_{i=1}^n e_i(h) u(x_i) \epsilon_i \right\| \right\|_p \\ & \leq \left\| \sup_{h \in H_n} \left\| \sum_{i=1}^n (e_i(h) - e_i(h_n)) u(x_i) \epsilon_i \right\| \right\|_p + \left\| \sum_{i=1}^n e_i(h_n) u(x_i) \epsilon_i \right\|_p. \end{aligned}$$

Applying Lemmas 14 and Corollary 1 of the Appendix C, we deduce that,

$$\limsup_{n \rightarrow \infty} \left\| \sup_{h \in H_n} \left\| \sum_{i=1}^n e_i(h) u(x_i) \epsilon_i \right\| \right\|_p < \infty. \quad (13)$$

Let, for $h \in A_\epsilon = \{h \in H_n, |\frac{h}{h_n} - 1| \leq \epsilon\}$, $H_{i,j}(h) = \frac{1}{h^3}(G_1 - 2K_1)(\frac{x_i - x_j}{h})$

$$\tilde{e}_{i,j}(h) = e_{i,j}(h) - e_{i,j}(h_n) = \frac{1}{n^{3/2}}(H_{i,j}(h) - H_{i,j}(h_n)).$$

Since, $|\frac{\partial H_{i,j}}{\partial h}(h)| \leq cst h^{-4}$, then for any $h, h' \in H_n$

$$|H_{i,j}(h) - H_{i,j}(h')| \leq n^{4/5} |h - h'| \mathbb{1}_{|i-j| \leq n \max(h, h')}$$

and, for any $h \in A_\epsilon$,

$$\begin{aligned} |\tilde{e}_{i,j}(h)| & \leq cst \frac{\epsilon}{n^{9/10}} |\mathbb{1}_{|i-j| \leq n \max(h, h')}|, \\ |\tilde{e}_{i,j}(h) - \tilde{e}_{i,j}(h')| & \leq cst \frac{1}{n^{7/10}} |h - h'| |\mathbb{1}_{|i-j| \leq n \max(h, h')}|. \end{aligned}$$

We have,

$$\begin{aligned} & \sup_{h \in A_\epsilon} \left| \sum_{i=1}^n \sum_{j=1}^{i-1} (u(x_i) + u(x_j)) e_{i,j}(h) \epsilon_i \epsilon_j \right| \\ & \leq \sup_{h \in A_\epsilon} \left| \sum_{i=1}^n \sum_{j=1}^{i-1} (u(x_i) + u(x_j)) (e_{i,j}(h) - e_{i,j}(h_n)) \epsilon_i \epsilon_j \right| \\ & \quad + \left| \sum_{i=1}^n \sum_{j=1}^{i-1} (u(x_i) + u(x_j)) e_{i,j}(h_n) \epsilon_i \epsilon_j \right|. \end{aligned}$$

Arguing as in Lemma 17, we have

$$\limsup_{n \rightarrow \infty} \left\| \sup_{h \in A_\epsilon} \left\| \sum_{i=1}^n \sum_{j=1}^{i-1} (u(x_i) + u(x_j)) (e_{i,j}(h) - e_{i,j}(h_n)) \epsilon_i \epsilon_j \right\| \right\|_p < \infty,$$

and by Proposition 8,

$$\limsup_{n \rightarrow \infty} \left\| \sum_{i=1}^n \sum_{j=1}^{i-1} (u(x_i) + u(x_j)) e_{i,j}(h_n) \epsilon_i \epsilon_j \right\|_p < \infty.$$

Consequently,

$$\lim_{n \rightarrow \infty} \left\| \sup_{h \in A_\epsilon} \left| \sum_{i=1}^n \sum_{j=1}^{i-1} (u(x_i) + u(x_j)) e_{i,j}(h) \epsilon_i \epsilon_j \right| \right\|_p < \infty. \quad (14)$$

The limit (9) is proved by collecting (12), (13) and (14).

4.1.2 Proof of Lemma 4

We can write the following decomposition,

$$\begin{aligned} T_n(h) - \mathbb{E}(T_n(h)) &= \frac{1}{n} \sum_{i=1}^n u(x_i) [(\hat{r}(x_i) - \mathbb{E}(\hat{r}(x_i)))^2 - \mathbb{E}[(\hat{r}(x_i) - \mathbb{E}(\hat{r}(x_i)))^2]] \\ &+ \frac{2}{n} \sum_{i=1}^n u(x_i) (\hat{r}(x_i) - \mathbb{E}(\hat{r}(x_i))) (\mathbb{E}(\hat{r}(x_i)) - r(x_i)) \\ &= \sum_{j=1}^n C_{j,n}(h) \epsilon_j + \sum_{j=1}^n \sum_{l=1}^{j-1} B_{j,l}(h) \epsilon_j \epsilon_l + \sum_{j=1}^n D_{j,n}(h) (\epsilon_j^2 - \mathbb{E}(\epsilon_j^2)), \end{aligned} \quad (15)$$

where,

$$\begin{aligned} C_{j,n}(h) &= \frac{2}{n^2 h} \sum_{i=1}^n u(x_i) K\left(\frac{x_i - x_j}{h}\right) B(x_i, h), \quad B(x_i, h) = \mathbb{E}(\hat{r}(x_i)) - r(x_i), \\ B_{j,l}(h) &= \frac{2}{n^3 h^2} \sum_{i=1}^n u(x_i) K\left(\frac{x_i - x_j}{h}\right) K\left(\frac{x_i - x_l}{h}\right), \\ D_{j,n}(h) &= \frac{1}{n^3 h^2} \sum_{i=1}^n u(x_i) K^2\left(\frac{x_i - x_j}{h}\right). \end{aligned}$$

Proof of (10). Let

$$c_{j,n}(h) = nh C_{j,n}(h) = \frac{2}{n} \sum_{i=1}^n u(x_i) K\left(\frac{x_i - x_j}{h}\right) B(x_i, h)$$

with $B(x_i, h) = \mathbb{E}(\hat{r}(x_i)) - r(x_i) = \frac{1}{nh} \sum_{l=1}^n K\left(\frac{x_i - x_l}{h}\right) r(x_l) - r(x_i)$. We get, for $h, h' \in H_n$,

$$c_{j,n}(h) - c_{j,n}(h') = \frac{2}{n} \sum_{i=1}^n u(x_i) \left(K\left(\frac{x_i - x_j}{h}\right) B(x_i, h) - K\left(\frac{x_i - x_j}{h'}\right) B(x_i, h') \right).$$

Now, since K is a Lipschitz function,

$$\begin{aligned} & \left| K\left(\frac{x_i - x_j}{h}\right)B(x_i, h) - K\left(\frac{x_i - x_j}{h'}\right)B(x_i, h') \right| \\ & \leq cst \left(|B(x_i, h) - B(x_i, h')| + |x_i - x_j| \sup_{h \in H_n} |B(x_i, h)| \frac{|h - h'|}{hh'} \right) \mathbb{1}_{|x_i - x_j| \leq \max(h, h')}. \end{aligned}$$

We have, for any $h, h' \in H_n$,

$$u(x_i)|B(x_i, h) - B(x_i, h')| \leq cst n^{-1/5}|h - h'|,$$

and by the proof of Lemma 8 in Appendix A.1, $|B(x_i, h)| \leq cst h^2$. Hence, for $h, h' \in H_n$,

$$|c_{j,n}(h) - c_{j,n}(h')| \leq cst n^{-2/5}|h - h'|.$$

Since K is compactly supported, we have,

$$\sup_{h \in H_n} |c_{j,n}(h)| \leq cst \max_{1 \leq i \leq n} \sup_{h \in H_n} (h|B(x_i, h)|) = O(n^{-\frac{3}{5}}).$$

Consequently, we obtain using Lemma 14,

$$\lim_{n \rightarrow \infty} \left\| \sup_{h \in H_n} nh \left\| \sum_{j=1}^n C_{j,n}(h) \epsilon_j \right\|_2 \right\| = 0. \quad (16)$$

Now, let

$$d_{j,n}(h) = nhD_{j,n}(h) = \frac{1}{n^2 h} \sum_{i=1}^n u(x_i) K^2\left(\frac{x_i - x_j}{h}\right).$$

We have, $|d_{j,n}(h)| \leq \frac{cst}{n}$ and

$$|d_{j,n}(h) - d_{j,n}(h')| \leq n^{-4/5}|h - h'|,$$

Then Lemma 15 gives,

$$\lim_{n \rightarrow \infty} \left\| \sup_{h \in H_n} nh \left\| \sum_{j=1}^n D_{j,n}(h) (\epsilon_j^2 - \mathbb{E}(\epsilon_j^2)) \right\|_p \right\| = 0. \quad (17)$$

Now, let

$$b_{j,l}(h) = nhB_{j,l}(h) = \frac{1}{n^2 h} \sum_{i=1}^n u(x_i) K\left(\frac{x_i - x_j}{h}\right) K\left(\frac{x_i - x_l}{h}\right),$$

we have,

$$b_{j,l}^2(h) = \frac{1}{n^4 h^2} \left(\sum_{i=1}^n u(x_i) K\left(\frac{x_i - x_j}{h}\right) K\left(\frac{x_i - x_l}{h}\right) \right)^2 \leq \frac{1}{n^2} \mathbb{1}_{|j-l| \leq 2nh}.$$

Our purpose now is to control, for $h, h' \in H_n$, the increment $|b_{j,l}(h) - b_{j,l}(h')|$. We have,

$$|b_{j,l}(h) - b_{j,l}(h')| \leq \frac{cst \max(h, h')}{nhh'} |h - h'| \leq cst n^{-4/5} |h - h'| \mathbb{1}_{|j-l| \leq 2nh}.$$

Then by Lemma 16, we obtain

$$\lim_{n \rightarrow \infty} \left\| \sup_{h \in H_n} nh \left\| \sum_{j=1}^n \sum_{l=1}^{j-1} B_{j,l}(h) \epsilon_j \epsilon_l \right\| \right\|_p = 0. \quad (18)$$

Collecting (15), (16), (17) and (18), we finally deduce (10).

Proof of (11). Taking the second derivative over h in (15), we have

$$\begin{aligned} & T_n''(h) - \mathbb{E}(T_n''(h)) \\ &= \sum_{j=1}^n C_{j,n}''(h) \epsilon_j + \sum_{j=1}^n \sum_{l=1}^{j-1} B_{j,l}''(h) \epsilon_j \epsilon_l + \sum_{j=1}^n D_{j,n}''(h) (\epsilon_j^2 - \mathbb{E}(\epsilon_j^2)), \end{aligned}$$

where, letting $B(x_i, h) = \mathbb{E}(\hat{r}(x_i)) - r(x_i)$,

$$\begin{aligned} & C_{j,n}''(h) \\ &= \frac{1}{n^2 h} \sum_{i=1}^n u(x_i) \left(\frac{B(x_i, h)}{h^2} G_1\left(\frac{x_i - x_j}{h}\right) + \frac{B'(x_i, h)}{h} G_2\left(\frac{x_i - x_j}{h}\right) + B''(x_i, h) G_3\left(\frac{x_i - x_j}{h}\right) \right), \\ & B_{j,l}''(h) = \frac{1}{n^3 h^4} \sum_{i=1}^n u(x_i) F_1\left(\frac{x_i - x_j}{h}\right) F_2\left(\frac{x_i - x_l}{h}\right) \\ & D_{j,n}''(h) = \frac{1}{n^3 h^4} \sum_{i=1}^n u(x_i) F\left(\frac{x_i - x_j}{h}\right), \end{aligned}$$

where F_1, F_2, F, G_1, G_2 and G_3 are bounded functions of class C^1 , $[-1, 1]$ -compactly supported. The proof of (11) is analogous to (9) and (10).

4.2 Proof of Proposition 1

We have from Lemma 1,

$$\lim_{n \rightarrow \infty} \sup_{h \in H_n} \left| \frac{\mathbb{E}(T_n(h))}{D_n(h)} - 1 \right| = 0. \quad (19)$$

From this, we claim that $\frac{h_n}{h_n^*}$ converges to 1, as n tends to infinity. In fact, it holds by the definition of h_n ,

$$\mathbb{E}(T_n(h_n)) \leq \mathbb{E}(T_n(h_n^*)).$$

Hence,

$$D_n(h_n) \frac{\mathbb{E}(T_n(h_n))}{D_n(h_n)} \leq \frac{\mathbb{E}(T_n(h_n^*))}{D_n(h_n^*)} D_n(h_n^*),$$

so by (19) and the definition of h_n^* , we deduce that, for a fixed $\epsilon > 0$ there exists n_0 such that for any $n \geq n_0$,

$$(1 - \epsilon)D_n(h_n) \leq (1 + \epsilon)D_n(h_n^*) \leq (1 + \epsilon)D_n(h_n),$$

so that $\lim_{n \rightarrow \infty} \frac{D_n(h_n^*)}{D_n(h_n)} = 1$, which ensures that $\lim_{n \rightarrow \infty} \frac{h_n^*}{h_n} = 1$, in fact (supposing without loss of generality that $D_n(h_n^*) - D_n(h_n) \neq 0$),

$$\begin{aligned} \frac{h_n^* - h_n}{h_n} &= \frac{h_n^* - h_n}{D_n(h_n^*) - D_n(h_n)} \frac{D_n(h_n^*) - D_n(h_n)}{h_n} \\ &= \frac{h_n^* - h_n}{D_n(h_n^*) - D_n(h_n)} \frac{D_n(h_n)}{h_n} \left(\frac{D_n(h_n^*)}{D_n(h_n)} - 1 \right) \\ &= \frac{D_n(h_n)}{h_n D_n'(h_n^*)} \left(\frac{D_n(h_n^*)}{D_n(h_n)} - 1 \right), \end{aligned}$$

where h^* is between h_n and h_n^* which are all in H_n , consequently $\limsup_{n \rightarrow \infty} \left| \frac{D_n(h_n)}{h_n D_n'(h_n^*)} \right| < \infty$ and then the behavior of $\frac{h_n^* - h_n}{h_n}$ is deduced from the fact that $\frac{D_n(h_n^*)}{D_n(h_n)} - 1$ tends to 0 as n tends to infinity.

In order to complete the proof of Proposition 1, we only need to prove that both $\frac{D_n(\hat{h}_n)}{D_n(h_n)}$ and $\frac{D_n(\hat{h}_M)}{D_n(h_n)}$ converge in probability to 1 as n tends to infinity (recall that both \hat{h}_M and \hat{h}_n belong to H_n). We refer the reader to Rice, J. (1984) for similar arguments. For this, we have to prove an analogous to the limit (19),

$$\lim_{n \rightarrow \infty} \left\| \sup_{h \in H_n} \left| \frac{T_n(h)}{D_n(h)} - 1 \right| \right\|_p = 0, \quad \text{for some } p > 8 \quad (20)$$

which gives, from the same previous arguments, that, for any $\epsilon > 0$,

$$\lim_{n \rightarrow \infty} \mathbb{P} \left((1 - \epsilon) \leq (1 + \epsilon) \frac{D_n(h_n^*)}{D_n(\hat{h}_n)} \leq (1 + \epsilon) \right) = 1.$$

Since $\inf_{h \in H_n} nhD_n(h) > 0$ and by (19), the limit (20) is proved as soon as,

$$\lim_{n \rightarrow \infty} \left\| \sup_{h \in H_n} nh |T_n(h) - \mathbb{E}(T_n(h))| \right\|_p = 0,$$

which immediately follows from Lemma 4 (more precisely (10) of Subsection 4.1). Our purpose now is to prove that $\frac{\hat{h}_M}{h_n}$ converges in probability to 1 as n tends to infinity. Recall that $\text{CL}(h) = T_n(h) + \delta_2(h) + n^{-1} \|U^{1/2}(Y - r)\|^2$, where $\delta_2(h) = 2n^{-1}(Y - r)'U(r - \hat{r}) + 2\sigma^2 n^{-1} \text{tr}(URL)$. We have, using Lemma 3 of Subsection 4.1,

$$\lim_{n \rightarrow \infty} \left\| \sup_{h \in H_n} nh |\delta_2(h)| \right\|_p = 0,$$

or equivalently, since $\inf_{h \in H_n} nhD_n(h) > 0$,

$$\lim_{n \rightarrow \infty} \left\| \sup_{h \in H_n} \left| \frac{\delta_2(h)}{D_n(h)} \right| \right\|_p = 0.$$

This last limit, together with (20), give

$$\lim_{n \rightarrow \infty} \left\| \sup_{h \in H_n} \left| \frac{T_n(h) + \delta_2(h)}{D_n(h)} - 1 \right| \right\|_p = 0. \quad (21)$$

Now, since $n^{-1} \|U^{1/2}(Y - r)\|^2$ doesn't depend on h ,

$$\hat{h}_M := \operatorname{argmin}_{h \in H_n} \operatorname{CL}(h) = \operatorname{argmin}_{h \in H_n} (T_n(h) + \delta_2(h)),$$

so that using (21) and the same previous arguments, we prove that

$$\frac{D_n(\hat{h}_M)}{D_n(\hat{h}_n)} \rightarrow 1, \text{ in probability as } n \rightarrow \infty.$$

The proof of Proposition 1 is completed.

4.3 Proof of Theorem 1

The following technical lemma is crucial for the proof of Theorem 1. It gives conditions under which $v_n(\hat{h}_M - \hat{h}_n)$ converges to a normal law with some rate v_n .

Lemma 5. *Suppose that Assumptions (A) hold. If, as n tends to infinity and for some positive rate a_n , (recall that $\delta_2(h)$ is defined in Equation (4)),*

1. $a_n \delta_2'(h_n)$ converges to a centered normal law with variance V ,
2. $a_n (\delta_2'(h_n) - \delta_2'(\hat{h}_n))$ converges in probability to 0,
3. $\frac{\operatorname{CL}''(h^*)}{\mathbb{E}(T_n''(h_n))}$ tends in probability to 1 for any h^* between \hat{h}_n and \hat{h}_M ,

then

$$a_n \mathbb{E}(T_n''(h_n))(\hat{h}_n - \hat{h}_M)$$

converges in distribution to a centered normal law with variance V .

Proof of Lemma 5. We have,

$$\operatorname{CL}(h) = T_n(h) + \delta_2(h) + n^{-1} \|U^{1/2}(Y - r)\|^2.$$

Hence $\operatorname{CL}'(h) = T_n'(h) + \delta_2'(h)$ and (recall that $\operatorname{CL}'(\hat{h}_M) = 0, T_n'(\hat{h}_n) = 0$),

$$\delta_2'(\hat{h}_n) = \operatorname{CL}'(\hat{h}_n) - T_n'(\hat{h}_n) = \operatorname{CL}'(\hat{h}_n) - \operatorname{CL}'(\hat{h}_M).$$

We also have, that there exists some sequence h^* between \hat{h}_n and \hat{h}_M such that

$$\operatorname{CL}'(\hat{h}_n) - \operatorname{CL}'(\hat{h}_M) = (\hat{h}_n - \hat{h}_M) \operatorname{CL}''(h^*) = (\hat{h}_n - \hat{h}_M) \mathbb{E}(T_n''(h_n)) \frac{\operatorname{CL}''(h^*)}{\mathbb{E}(T_n''(h_n))}.$$

Consequently,

$$\begin{aligned} a_n \mathbb{E}(T_n''(h_n))(\hat{h}_n - \hat{h}_M) &= a_n \frac{\mathbb{E}(T_n''(h_n))}{\text{CL}''(h^*)} (\text{CL}'(\hat{h}_n) - \text{CL}'(\hat{h}_M)) \\ &= \frac{\mathbb{E}(T_n''(h_n))}{\text{CL}''(h^*)} a_n \delta_2'(\hat{h}_n) = \frac{\mathbb{E}(T_n''(h_n))}{\text{CL}''(h^*)} a_n \delta_2'(h_n) + \frac{\mathbb{E}(T_n''(h_n))}{\text{CL}''(h^*)} a_n (\delta_2'(\hat{h}_n) - \delta_2'(h_n)). \end{aligned}$$

The last equality together with the assumptions of Lemma 5 complete the proof of this lemma. According to Lemma 5, we have to consider three steps. We study each of them in the following three subsections. The fourth subsection concludes the proof of Theorem 1

4.3.1 Step 1: convergence in distribution of $a_n \delta_2'(h_n)$.

The following proposition studies the asymptotic distribution of $a_n \delta_2'(h_n)$ for $a_n = \sqrt{\frac{n}{h_n^2}}$.

Proposition 3. *Suppose that the assumptions of Theorem 1 are satisfied. Then the following two assertions are equivalent.*

- $\sqrt{\frac{n}{h_n^2}} \delta_2'(h_n)$ converges in distribution as n tends to infinity to a centered normal law with variance $4V$
- $\sqrt{\frac{n}{h_n^2}} \sum_{i=1}^n \left(\tilde{a}_{i,n}(h_n) u(x_i) \epsilon_i + \sum_{j=1}^{i-1} (u(x_i) + u(x_j)) b_{i,j}(h_n) \epsilon_i \epsilon_j \right)$ converges to a centered normal law with variance V , where

$$\tilde{a}_{i,n}(h_n) = -C_K \frac{h_n}{n} r''(x_i).$$

Proof of Proposition 3. Recall that, for $G(u) = -uK'(u)$, for any $h \in]0, \epsilon[$ and any $n \geq 1$,

$$\begin{aligned} \delta_2'(h) &= 2 \sum_{i=1}^n a_i(h) u(x_i) \epsilon_i + 2 \sum_{i=1}^n \sum_{j=1}^{i-1} (u(x_i) + u(x_j)) b_{i,j}(h) \epsilon_i \epsilon_j \\ &+ 2 \sum_{i=1}^n b_{i,i}(h) u(x_i) (\epsilon_i^2 - \mathbb{E}(\epsilon_i^2)), \end{aligned} \tag{22}$$

where,

$$\begin{aligned} a_i(h) &= -\frac{1}{n} \frac{\partial}{\partial h} \mathbb{E}(\hat{r}(x_i)) = \frac{1}{n^2 h^2} \sum_{j=1}^n (K - G)\left(\frac{x_i - x_j}{h}\right) r(x_j), \\ b_{i,j}(h) &= \frac{1}{n^2 h^2} K\left(\frac{x_i - x_j}{h}\right) - \frac{1}{n^2 h^2} G\left(\frac{x_i - x_j}{h}\right), \\ b_{i,i}(h) &= \frac{1}{n^2 h^2} K(0). \end{aligned}$$

We also need, for the proof of Proposition 3, the following two lemmas.

Lemma 6. Recall that $h_n = cn^{-1/5}$ and suppose that $\sum_{j=1}^{\infty} |\text{Cov}(\epsilon_1^2, \epsilon_j^2)| < \infty$, then

$$\lim_{n \rightarrow \infty} \frac{n}{h_n^2} \text{Var} \left(\sum_{i=1}^n u(x_i) b_{i,i}(h_n) (\epsilon_i^2 - \mathbb{E}(\epsilon_i^2)) \right) = 0.$$

Proof of Lemma 6. We have,

$$\begin{aligned} \frac{n}{h_n^2} \text{Var} \left(\sum_{i=1}^n u(x_i) b_{i,i}(h_n) (\epsilon_i^2 - \mathbb{E}(\epsilon_i^2)) \right) &= \frac{n}{h_n^2 n^4 h_n^4} K^2(0) \sum_{i=1}^n \sum_{j=1}^n u(x_i) u(x_j) \text{Cov}(\epsilon_i^2, \epsilon_j^2) \\ &\leq \frac{1}{n^2 h_n^6} K^2(0) \|u\|_{\infty}^2 \sup_i \sum_{j=1}^{\infty} |\text{Cov}(\epsilon_i^2, \epsilon_j^2)|. \end{aligned}$$

The proof of this lemma is achieved since $\lim_{n \rightarrow \infty} n^2 h_n^6 = \lim_{n \rightarrow \infty} n^{4/5} = \infty$.

Lemma 7. Recall that $h_n = cn^{-1/5}$. We have, noting $C_K = \int x^2 K(x) dx$,

$$\lim_{n \rightarrow \infty} \frac{n}{h_n^2} \text{Var} \left(\sum_{i=1}^n \left(a_i(h_n) + C_K \frac{h_n}{n} r''(x_i) \right) u(x_i) \epsilon_i \right) = 0.$$

Proof of Lemma 7. Clearly, we have using Lemma 10 of Appendix A.2,

$$a_i(h_n) + C_K \frac{h_n}{n} r''(x_i) = O\left(\frac{h_n^2}{n} + \frac{1}{n^2 h_n^3}\right),$$

and

$$\begin{aligned} &\text{Var} \left(\sum_{i=1}^n \left(a_i(h_n) + C_K \frac{h_n}{n} r''(x_i) \right) u(x_i) \epsilon_i \right) \\ &= \sum_{i=1}^n \sum_{j=1}^n \left(a_i(h_n) + C_K \frac{h_n}{n} r''(x_i) \right) \left(a_j(h_n) + C_K \frac{h_n}{n} r''(x_j) \right) u(x_i) u(x_j) \text{Cov}(\epsilon_i, \epsilon_j) \\ &\leq cst \sup_i \left(\left(a_i(h_n) + C_K \frac{h_n}{n} r''(x_i) \right) u(x_i) \right)^2 n \sigma^2 \\ &= O\left(n \left(\frac{h_n^2}{n} + \frac{1}{n^2 h_n^3} \right)^2 \right) = O(n^{-9/5}) = o\left(\frac{h_n^2}{n}\right). \end{aligned}$$

The proof of Lemma 7 is complete.

End of the proof of Proposition 3. We have, using (22),

$$\begin{aligned} \sqrt{\frac{n}{h_n^2}} \delta_2'(h_n) &= 2 \sqrt{\frac{n}{h_n^2}} \sum_{i=1}^n \left(\tilde{a}_{i,n}(h_n) u(x_i) \epsilon_i + \sum_{j=1}^{i-1} (u(x_i) + u(x_j)) b_{i,j}(h_n) \epsilon_i \epsilon_j \right) \\ &\quad + 2 \sqrt{\frac{n}{h_n^2}} \sum_{i=1}^n (a_i(h_n) - \tilde{a}_{i,n}(h_n)) u(x_i) \epsilon_i + 2 \sqrt{\frac{n}{h_n^2}} \sum_{i=1}^n b_{i,i}(h_n) u(x_i) (\epsilon_i^2 - \mathbb{E}(\epsilon_i^2)). \end{aligned}$$

The proof of Proposition 3 is complete if $\sqrt{\frac{n}{h_n^2}} \sum_{i=1}^n (a_i(h_n) - \tilde{a}_{i,n}(h_n)) u(x_i) \epsilon_i$ and $\sqrt{\frac{n}{h_n^2}} \sum_{i=1}^n b_{i,i}(h_n) u(x_i) (\epsilon_i^2 - \mathbb{E}(\epsilon_i^2))$ converge in probability to 0 as n tends to infinity, which are satisfied due to Lemmas 6 and 7.

4.3.2 Step 2: convergence in probability of $a_n(\delta'_2(h_n) - \delta'_2(\hat{h}_n))$

The following proposition checks step 2 of Lemma 5.

Proposition 4. *Under the assumptions of Theorem 1, $\sqrt{\frac{n}{h_n^2}}(\delta'_2(h_n) - \delta'_2(\hat{h}_n))$ converges in probability to 0 as n tends to infinity.*

Proof of Proposition 4. We have,

$$\sqrt{\frac{n}{h_n^2}}(\delta'_2(h_n) - \delta'_2(\hat{h}_n)) = \sqrt{\frac{n}{h_n^2}}(h_n - \hat{h}_n)\delta''_2(h^*) = \sqrt{n}\left(1 - \frac{\hat{h}_n}{h_n}\right)\delta''_2(h^*),$$

where h^* is an element of H_n between h_n and \hat{h}_n and since \hat{h}_n/h_n converges in probability to 1 as n tends to infinity (by Proposition 1), we deduce that

$$\lim_{n \rightarrow \infty} \mathbb{P}(h^* \notin A_\epsilon) = 0, \quad \forall \epsilon > 0, \quad (23)$$

where, for fixed $\epsilon > 0$, $A_\epsilon = \{h \in H_n, | \frac{h}{h_n} - 1 | \leq \epsilon\}$. Now, we have for any $M > 0$,

$$\begin{aligned} & \mathbb{P}\left(a_n|\delta'_2(h_n) - \delta'_2(\hat{h}_n)| \geq \epsilon^2\right) \\ & \leq \mathbb{P}\left(\sqrt{n}\left|1 - \frac{\hat{h}_n}{h_n}\right||\delta''_2(h^*)| \geq \epsilon^2, h^* \in A_\epsilon\right) + \mathbb{P}\left(\sqrt{n}\left|1 - \frac{\hat{h}_n}{h_n}\right||\delta''_2(h^*)| \geq \epsilon^2, h^* \notin A_\epsilon\right) \\ & \leq \mathbb{P}\left(\sqrt{n}\left|1 - \frac{\hat{h}_n}{h_n}\right||\delta''_2(h^*)|\mathbb{1}_{h^* \in A_\epsilon} \geq \epsilon^2\right) + \mathbb{P}(h^* \notin A_\epsilon) \\ & \leq \mathbb{P}\left(\sqrt{n}\left|1 - \frac{\hat{h}_n}{h_n}\right| \sup_{h \in A_\epsilon} |\delta''_2(h)| \geq \epsilon^2\right) + \mathbb{P}(h^* \notin A_\epsilon) \\ & \leq \mathbb{P}\left(\sqrt{n}\left|1 - \frac{\hat{h}_n}{h_n}\right| \sup_{h \in A_\epsilon} |\delta''_2(h)| \geq \epsilon^2, \sqrt{n} \sup_{h \in A_\epsilon} |\delta''_2(h)| \geq M\right) \\ & \quad + \mathbb{P}\left(\sqrt{n}\left|1 - \frac{\hat{h}_n}{h_n}\right| \sup_{h \in A_\epsilon} |\delta''_2(h)| \geq \epsilon^2, \sqrt{n} \sup_{h \in A_\epsilon} |\delta''_2(h)| < M\right) + \mathbb{P}(h^* \notin A_\epsilon) \\ & \leq \mathbb{P}\left(\sqrt{n} \sup_{h \in A_\epsilon} |\delta''_2(h)| \geq M\right) + \mathbb{P}\left(M\left|1 - \frac{\hat{h}_n}{h_n}\right| \geq \epsilon^2\right) + \mathbb{P}(h^* \notin A_\epsilon), \end{aligned}$$

which tends to 0 by letting first n tends to infinity and then M tends to infinity, due to Proposition 1, (9) and (23).

4.3.3 Step 3: convergence in probability of $\frac{\text{CL}''(h^*)}{\mathbb{E}(T_n''(h_n))}$

Proposition 5. *Under the assumptions of Theorem 1, $\frac{\text{CL}''(h^*)}{\mathbb{E}(T_n''(h_n))}$ tends in probability to 1, as $n \rightarrow \infty$, for any h^* between \hat{h}_n and \hat{h}_M .*

Proof of Proposition 5. We have, for any $h > 0$, $\text{CL}''(h) = T_n''(h) + \delta_2''(h)$, and

$$\frac{\text{CL}''(h)}{\mathbb{E}(T_n''(h))} = \frac{T_n''(h)}{\mathbb{E}(T_n''(h))} + \frac{\delta_2''(h)}{\mathbb{E}(T_n''(h))}.$$

Our first purpose is to prove that, $\sup_{h \in H_n} \frac{|\delta_2''(h)|}{|\mathbb{E}(T_n''(h))|}$ converges to 0, in probability, as n tends to infinity. Since $\sup_{h \in H_n} \left| \frac{D_n''(h)}{|\mathbb{E}(T_n''(h))|} - 1 \right|$ converges to 0 as n tends to infinity, with

$$D_n''(h) = 3h^2 \int_0^1 u(x)r''(x)dx \int_{-1}^1 t^2 K(t)dt + \frac{2}{nh^3} \left(\int_0^1 u(x)dx \right) \int_{-1}^1 K^2(y)dy\sigma^2,$$

(see Lemma 11 of Appendix A.2) it remains then to prove that, (since $\inf_{h \in H_n} nh^3 |D_n''(h)| > 0$),

$$\sup_{h \in H_n} (nh^3 |\delta_2''(h)|) \longrightarrow 0, \text{ in probability as } n \rightarrow \infty,$$

which is proved due to Lemma 3 of Subsection 4.1. It remains to prove that

$$\sup_{h \in H_n} \left| \frac{T_n''(h)}{\mathbb{E}(T_n''(h))} - 1 \right| \longrightarrow 0, \text{ in probability as } n \rightarrow \infty,$$

or equivalently,

$$\sup_{h \in H_n} nh^3 |T_n''(h) - \mathbb{E}(T_n''(h))| \longrightarrow 0, \text{ in probability as } n \rightarrow \infty,$$

which is proved due to Lemma 4 of Subsection 4.1. Consequently,

$$\sup_{h \in H_n} \left| \frac{\text{CL}''(h)}{\mathbb{E}(T_n''(h))} - 1 \right| \longrightarrow 0, \text{ in probability as } n \rightarrow \infty,$$

Finally,

$$\left| \frac{\text{CL}''(h^*)}{D_n''(h_n)} - 1 \right| \leq cst \sup_{h \in H_n} \left| \frac{\text{CL}''(h)}{D_n''(h)} - 1 \right| + \left| \frac{D_n''(h^*)}{D_n''(h_n)} - 1 \right|.$$

Since by definition of h^* and by Proposition 1, we deduce that

$$\left| \frac{D_n''(h^*)}{D_n''(h_n)} - 1 \right| \longrightarrow 0 \text{ in probability as } n \rightarrow \infty,$$

and then

$$\left| \frac{\text{CL}''(h^*)}{D_n''(h_n)} - 1 \right| \longrightarrow 0 \text{ in probability as } n \rightarrow \infty.$$

4.3.4 End of the proof of Theorem 1

We have to check the three items of Lemma 5. We have, from Proposition 7,

$$n^{7/10} \sum_{i=1}^n \left(\tilde{a}_{i,n}(h_n) u(x_i) \epsilon_i + \sum_{j=1}^{i-1} (u(x_i) + u(x_j)) b_{i,j}(h_n) \epsilon_i \epsilon_j \right) \Longrightarrow \mathcal{N}(0, V).$$

It follows from Proposition 3 that

$$n^{7/10} \delta_2'(h_n) \Longrightarrow \mathcal{N}(0, 4V),$$

where $V = c^2 C_K^2 \sigma^2 \int_0^1 u^2(x) r''^2(x) dx + \frac{4}{c^3} \sigma^4 \int_0^1 u^2(x) dx \int_0^1 (K - G)^2(u) du$. The two other items of Lemma 5 are satisfied using Propositions 4 and 5. The proof of Theorem 1 is complete using Lemma 5 together with the fact that (see Lemma 11), $\mathbb{E}(T_n''(h_n))$ is equivalent to

$$n^{-2/5} \left(3c^2 \int_0^1 u(x) r''^2(x) dx \left(\int_{-1}^1 t^2 K(t) dt \right)^2 + \frac{2}{c^3} \left(\int_0^1 u(x) dx \right) \int_{-1}^1 K^2(y) dy \sigma^2 \right).$$

5 Acknowledgements

This work was developed in the framework of Grenoble Alpes Data Institute (ANR-15-IDEX-02).

6 References

- Akaike, H. (1970). Statistical predictor information. *Annals of the Institute of Statistical Mathematics*, **22**, 203-217.
- Andrew, D. W. K. and Pollard, D. (1994). An introduction to functional central limit theorems for dependent stochastic processes. *Int. Stat. Rev.* **62**, 119-132.
- Burkholder, D. L. (1988). Sharp inequalities for martingales and stochastic integrals. *Astérisque* **157-158**, 75-94.
- Craven, P. and Wahba, G. (1979). Smoothing noisy data with spline functions. *Numerische Mathematik*, **31**, 377-403.

- Engle, R. F. (1982). Autoregressive Conditional Heteroscedasticity with Estimates of the Variance of United Kingdom Inflation. *Econometrica* **50-4**, 987-1007.
- Doukhan, P. and Louhichi, S. (1999). A new weak dependence condition and applications to moment inequalities. *Stochastic Process. Appl.*, **84**, 2, 313-342.
- T. Gasser and H.G. Muller (1979). Kernel estimation of regression functions, in Smoothing Techniques in Curve Estimation. *Lecture Notes in Mathematics* **757**, 23-68.
- Girard, D. (1998). Asymptotic comparison of (partial) cross-validation, GCV and randomized GCV in non-parametric regression. *Ann. Stat.*, **26**, 315-334.
- Han, Ch. and Gu, Ch. (2008). Optimal Smoothing with Correlated Data. *Sankhya: The Indian Journal of Statistics, Series A*, **70**, 1, 38-72.
- Härdle, W., Hall, P. and Marron, J.S. (1988). How far are automatically chosen regression smoothing parameters from their optimum? *Journal of the American Statistical Association* **83**, 86-95.
- Hall, P., Lahiri, S. N. and Polzehl, J. (1995). On bandwidth choice in nonparametric regression with both short and long-range dependent errors. *Ann. Statist.* **23-6**, 1921-1936.
- Hastie, T., Tibshirani, R. and Friedman, J. (2009). The Elements of Statistical Learning: Data Mining, Inference, and Prediction. Springer.
- Lindner, A.M. (2009). Stationarity, Mixing, Distributional Properties and Moments of GARCH(p, q)-Processes. In: Mikosch T., Kreib JP., Davis R., Andersen T. (eds) Handbook of Financial Time Series. Springer, Berlin, Heidelberg.
- Louhichi, S. (2000). Weak convergence for empirical processes of associated sequences. *Ann. Inst. Henri Poincaré Probabilités et Statistiques*, **36**, 5, 547-567.
- Mallows, C. L. (1973). Some comments on Cp. *Technometrics* **15**, 661-675.
- McLeish, D. L. (1974). Dependent Central Limit Theorems and Invariance Principles. *Ann. Probab.* **2**, 4, 620-628.
- Peligrad, M., Utev, S. and Wu, W. B. (2007). A maximal Lp-inequality for stationary sequences and its applications. *Proc. Am. Math. Soc.* **135**, 541-550.
- Pollard, D. (1990). Empirical Processes: Theory and Applications. NSF-CBMS Regional Conference Series in Probability and Statistics, Institute of Mathematical Statistics, Hayward, CA.

- Rice, J. (1984). Bandwidth choice for nonparametric regression. *Annals of Statistics* **12**, 1215-1230.
- Rio, E. (1993). Covariance inequalities for strongly mixing processes. *Ann. Inst. H. Poincaré Probab. Statist.*, **29** 587-597.
- Rio, E. (2017). Asymptotic theory of weakly dependent random processes. *Probability Theory and Stochastic Modelling*, 80. Springer.
- Shibata, R. (1981). An optimal selection of regression variables. *Biometrika*, **68**, 45-54.
- Nadaraya, E. A. (1964). On Estimating Regression. *Theory of Probability and Its Applications*. **9** (1), 141-2.
- Watson, G. S. (1964). Smooth regression analysis. *Sankhya: The Indian Journal of Statistics, Series A*. **26**, 359-372.
- Y. Wang (1998). Smoothing Spline Models with Correlated Random Errors. *Journal of the American Statistical Association*, **93**, 441, 41-348.

A Auxiliary technical tools

A.1 Proof of Lemma 1

In all the following proofs we denote by *cst* a generic constant independent of n and of h . We need the following two lemmas for the proof of Lemma 1.

Lemma 8. *One has, under the requirements of Lemma 1,*

$$\begin{aligned} & \frac{1}{n} \sum_{i=1}^n u(x_i) (\mathbb{E}(\hat{r}(x_i)) - r(x_i))^2 \\ &= \frac{h^4}{4} \int_0^1 u(x) r''^2(x) dx \left(\int_{-1}^1 t^2 K(t) dt \right)^2 + O\left(\frac{1}{n}\right) + o(h^4) + O\left(\frac{1}{n^2 h^4}\right). \end{aligned}$$

Lemma 9. *It the assumptions of Lemma 1 are satisfied, then*

$$\frac{1}{n} \sum_{i=1}^n u(x_i) \text{Var}(\hat{r}(x_i)) = \frac{\sigma^2}{nh} \left(\int_0^1 u(x) dx \right) \int_{-1}^1 K^2(y) dy + O\left(\frac{1}{n^2 h^3}\right) + \frac{\gamma(h)}{nh}.$$

where $\gamma(h)$ depends on h (not on n) and tends to 0 as h tends to 0.

The proof of Lemma 1 is complete using Lemmas 8, 9 and the trivial fact

$$\mathbb{E}(T_n(h)) = \frac{1}{n} \sum_{i=1}^n u(x_i) (\mathbb{E}(\hat{r}(x_i)) - r(x_i))^2 + \frac{1}{n} \sum_{i=1}^n u(x_i) \text{Var}(\hat{r}(x_i)).$$

Proof of Lemma 8. We have,

$$\begin{aligned} \mathbb{E}(\hat{r}(x_i)) - r(x_i) &= \frac{1}{nh} \sum_{j=1}^n K\left(\frac{x_i - x_j}{h}\right) r(x_j) - r(x_i) \\ &= \frac{1}{h} \int_0^1 K\left(\frac{x_i - s}{h}\right) r(s) ds - r(x_i) + \Delta_1(h) \\ &= \int_{(x_i-1)/h}^{x_i/h} K(y) r(x_i - hy) dy - r(x_i) + \Delta_1(h), \end{aligned} \tag{24}$$

where $\Delta_1(h) = \frac{1}{nh} \sum_{j=1}^n K\left(\frac{x_i - x_j}{h}\right) r(x_j) - \frac{1}{h} \int_0^1 K\left(\frac{x_i - s}{h}\right) r(s) ds$. In order to control $\Delta_1(h)$, we apply the following bound, true for any \mathcal{C}^1 function f on $[0, 1]$,

$$\left| \int_0^1 f(x) dx - \frac{1}{n} \sum_{i=1}^n f(x_i) \right| \leq \frac{1}{n} \|f'\|_\infty, \tag{25}$$

to the function $f(s) = K\left(\frac{x_i - s}{h}\right) r(s)$. We obtain

$$|\Delta_1(h)| \leq \frac{\|K'\|_\infty \|r\|_\infty}{nh^2} + \frac{\|K\|_\infty \|r'\|_\infty}{nh} \leq \frac{cst}{nh^2}. \tag{26}$$

Now, applying the Taylor's expansion, we have

$$|\Delta_2(h, y)| := \left| r(x_i - hy) - r(x_i) + hyr'(x_i) - \frac{1}{2}h^2y^2r''(x_i) \right| = o(h^2). \quad (27)$$

Hence,

$$\begin{aligned} & \int_{(x_i-1)/h}^{x_i/h} K(y)r(x_i - hy)dy - r(x_i) \\ &= \int_{(x_i-1)/h}^{x_i/h} K(y)(r(x_i) - hyr'(x_i) + \frac{1}{2}h^2y^2r''(x_i) + \Delta_2(h, y))dy - r(x_i) \\ &= r(x_i) \left(\int_{(x_i-1)/h}^{x_i/h} K(y)dy - 1 \right) - hr'(x_i) \int_{(x_i-1)/h}^{x_i/h} yK(y)dy \\ & \quad + \frac{1}{2}h^2r''(x_i) \int_{(x_i-1)/h}^{x_i/h} y^2K(y)dy + \int_{(x_i-1)/h}^{x_i/h} K(y)\Delta_2(h, y)dy. \end{aligned}$$

We obtain, collecting all the above equalities,

$$\begin{aligned} & \mathbb{E}(\hat{r}(x_i)) - r(x_i) = \\ & r(x_i) \left(\int_{(x_i-1)/h}^{x_i/h} K(y)dy - 1 \right) - hr'(x_i) \int_{(x_i-1)/h}^{x_i/h} yK(y)dy \\ & \quad + \frac{1}{2}h^2r''(x_i) \int_{(x_i-1)/h}^{x_i/h} y^2K(y)dy + \Delta_1(h) + \int_{(x_i-1)/h}^{x_i/h} \Delta_2(h, y)K(y)dy. \end{aligned} \quad (28)$$

We have, using (27),

$$\left| \int_{(x_i-1)/h}^{x_i/h} \Delta_2(h, y)K(y)dy \right| = o(h^2).$$

If $[-1, 1] \subset [(x_i - 1)/h, x_i/h]$ then $\int_{(x_i-1)/h}^{x_i/h} K(y)dy = 1$, $\int_{(x_i-1)/h}^{x_i/h} yK(y)dy = 0$, $\int_{(x_i-1)/h}^{x_i/h} y^2K(y)dy = \int_{-1}^1 y^2K(y)dy$ and (28) together with (26), (27) give,

$$\mathbb{E}(\hat{r}(x_i)) - r(x_i) = \frac{1}{2}h^2r''(x_i) \int_{-1}^1 u^2K(u)du + O\left(\frac{1}{nh^2}\right) + o(h^2),$$

and

$$\begin{aligned} & u(x_i) (\mathbb{E}(\hat{r}(x_i)) - r(x_i))^2 \\ &= \frac{1}{4}h^4u(x_i)r''^2(x_i) \left(\int_{-1}^1 u^2K(u)du \right)^2 + O\left(\frac{1}{n^2h^4}\right) + O\left(\frac{1}{n}\right) + o(h^4). \end{aligned}$$

Now if $x_i < h < \epsilon$ or if $1 - \epsilon < 1 - h < x_i$ then by definition of u , $u(x_i) = 0$. So that the bound (29) is true for any $i \in \{1, \dots, n\}$ and gives, due to (25),

$$\begin{aligned} & \frac{1}{n} \sum_{i=1}^n u(x_i) (\mathbb{E}(\hat{r}(x_i)) - r(x_i))^2 \\ &= \frac{1}{4} h^4 \int_0^1 u(x) r''^2(x) dx \left(\int_{-1}^1 u^2 K(u) du \right)^2 + O\left(\frac{1}{n}\right) + O\left(\frac{1}{n^2 h^4}\right) + o(h^4). \end{aligned}$$

This proves Lemma 8.

Proof of Lemma 9. We have, MDS is a sequence of non-correlated random variables,

$$\begin{aligned} u(x_i) \text{Var}(\hat{r}(x_i)) &= \frac{1}{n^2 h^2} \sum_{j_1=1}^n \sum_{j_2=1}^n u(x_i) K\left(\frac{x_i - x_{j_1}}{h}\right) K\left(\frac{x_i - x_{j_2}}{h}\right) \text{Cov}(\epsilon_{j_1}, \epsilon_{j_2}) \\ &= \sigma^2 \frac{1}{n^2 h^2} \sum_{j_1=1}^n u(x_i) K^2\left(\frac{x_i - x_{j_1}}{h}\right). \end{aligned}$$

We have, using (25) together with some properties of the kernel K ,

$$\left| \frac{1}{n} \sum_{j_1=1}^n K^2\left(\frac{x_i - x_{j_1}}{h}\right) - \int_0^1 K^2\left(\frac{x_i - s}{h}\right) ds \right| \leq \frac{cst}{nh}.$$

Hence,

$$\left| u(x_i) \text{Var}(\hat{r}(x_i)) - \sigma^2 \frac{1}{nh^2} \int_0^1 u(x_i) K^2\left(\frac{x_i - s}{h}\right) ds \right| \leq cst \frac{1}{n^2 h^3}.$$

This bound together with the following inequality (obtained using (25))

$$\left| \frac{1}{n} \sum_{i=1}^n u(x_i) K^2\left(\frac{x_i - s}{h}\right) - \int_0^1 u(x) K^2\left(\frac{x - s}{h}\right) dx \right| \leq \frac{cst}{nh},$$

give,

$$\left| \sum_{i=1}^n u(x_i) \text{Var}(\hat{r}(x_i)) - \frac{\sigma^2}{h^2} \int_0^1 \int_0^1 u(x) K^2\left(\frac{x - s}{h}\right) ds dx \right| \leq cst \frac{1}{nh^3}.$$

The last bound together with the following elementary calculation,

$$\begin{aligned} & \frac{1}{h} \int_0^1 \int_0^1 u(x) K^2\left(\frac{x - s}{h}\right) ds dx = \int_0^1 u(x) \left(\int_{(x-1)/h}^{x/h} K^2(y) dy \right) dx \\ &= \int_{-1}^1 K^2(y) \left(\int_{hy}^{hy+1} u(x) dx \right) dy \\ &= \int_{-1}^1 K^2(y) dy \left(\int_0^1 u(x) dx \right) + \int_{-1}^1 K^2(y) dy \left(\int_{hy}^{hy+1} u(x) dx - \int_0^1 u(x) dx \right) \\ &= \int_{-1}^1 K^2(y) dy \left(\int_0^1 u(x) dx \right) + \gamma(h), \end{aligned}$$

complete the proof of Lemma 9.

A.2 Other useful lemmas

Lemma 10 below is used for the proof of Lemma 7 and Lemma 3.

Lemma 10. *Let $h \in]0, \epsilon[$ be fixed. We have, for any $i \in \{1, \dots, n\}$,*

$$\begin{aligned} & -u(x_i) \frac{1}{n} \frac{\partial}{\partial h} (\mathbb{E}(\hat{r}(x_i))) \\ &= -u(x_i) \frac{h}{n} r''(x_i) \int u^2 K(u) du + O\left(\frac{h^2}{n}\right) + O\left(\frac{1}{n^2 h^3}\right). \end{aligned} \quad (29)$$

For any $h, h' \in H_n$, we have, letting $B(x_i, h) = \mathbb{E}(\hat{r}(x_i)) - r(x_i)$, for any $1 \leq i \leq n$,

$$u(x_i) |B(x_i, h) - B(x_i, h')| \leq cst n^{-1/5} |h - h'|, \quad (30)$$

$$u(x_i) \sup_{h \in H_n} |B''(x_i, h)| < \infty, \quad (31)$$

$$u(x_i) |B''(x_i, h) - B''(x_i, h')| \leq cst |h - h'|. \quad (32)$$

We also need for the proofs of Proposition 5 and Theorem 1, the following trivial lemma that we state without proof.

Lemma 11. *Let $h \mapsto D_n(h)$ be the function as defined in Lemma 1. Then this function is twice differentiable and*

$$D_n''(h) = 3h^2 \int_0^1 u(x) r''^2(x) dx \left(\int_{-1}^1 t^2 K(t) dt \right)^2 + \frac{2\sigma^2}{nh^3} \left(\int_0^1 u(x) dx \right) \int_{-1}^1 K^2(y) dy.$$

We also have, $\sup_{h \in H_n} \left| \frac{|D_n''(h)|}{|\mathbb{E}(T_n''(h))|} - 1 \right|$ converges to 0 as n tends to infinity.

Proof of Lemma 10. We take the derivative over h in (24) and we obtain,

$$\begin{aligned} & \frac{\partial \mathbb{E}(\hat{r}(x_i))}{\partial h} \\ &= -\frac{1}{h^2} \int_0^1 K\left(\frac{x_i - s}{h}\right) r(s) ds - \frac{1}{h} \int_0^1 \frac{x_i - s}{h^2} K'\left(\frac{x_i - s}{h}\right) r(s) ds + \Delta'_1(h) \\ &= -\frac{1}{h} \int_{(x_i-1)/h}^{x_i/h} K(y) r(x_i - yh) ds - \frac{1}{h} \int_{(x_i-1)/h}^{x_i/h} y K'(y) r(x_i - yh) dy + \Delta'_1(h). \end{aligned}$$

Hence if $[-1, 1] \subset [(x_i - 1)/h, x_i/h]$ then

$$\frac{\partial \mathbb{E}(\hat{r}(x_i))}{\partial h} = -\frac{1}{h} \int_{-1}^1 K(y) r(x_i - yh) ds - \frac{1}{h} \int_{-1}^1 y K'(y) r(x_i - yh) dy + \Delta'_1(h).$$

If $[-1, 1]$ is not a subset of $[(x_i - 1)/h, x_i/h]$, then $u(x_i) = 0$ so that in both cases, we have

$$\begin{aligned} & u(x_i) \frac{\partial \mathbb{E}(\hat{r}(x_i))}{\partial h} \\ &= -\frac{u(x_i)}{h} \int_{-1}^1 K(y)r(x_i - yh)ds - \frac{u(x_i)}{h} \int_{-1}^1 yK'(y)r(x_i - yh)dy + u(x_i)\Delta'_1(h). \end{aligned} \quad (33)$$

Now, using the same arguments as in (28),

$$\begin{aligned} & \frac{1}{h} \int_{-1}^1 K(y)r(x_i - yh)dy = \frac{r(x_i)}{h} + \frac{h}{2}r''(x_i) \int_{-1}^1 y^2K(y)dy + O(h^2), \\ & \frac{1}{h} \int_{-1}^1 yK'(y)r(x_i - yh)dy \\ &= \int_{-1}^1 yK'(y)dy \frac{r(x_i)}{h} - r'(x_i) \int_{-1}^1 y^2K'(y)dy + \frac{h}{2}r''(x_i) \int_{-1}^1 y^3K'(y)dy + O(h^2). \end{aligned} \quad (34)$$

Since $\int_{-1}^1 yK'(y)dy = -1$, $\int_{-1}^1 y^2K'(y)dy = 0$, $\int_{-1}^1 y^3K'(y)dy = -3 \int_{-1}^1 y^2K(y)dy$, we deduce that,

$$\begin{aligned} & \frac{1}{h} \int_{-1}^1 yK'(y)r(x_i - yh)dy \\ &= -\frac{r(x_i)}{h} - \frac{3h}{2}r''(x_i) \int_{-1}^1 y^2K(y)dy + O(h^2). \end{aligned} \quad (35)$$

We obtain, collecting (33), (34) and (36), for any $1 \leq i \leq n$,

$$u(x_i) \frac{\partial \mathbb{E}(\hat{r}(x_i))}{\partial h} = hr''(x_i)u(x_i) \int_{-1}^1 y^2K(y)dy + u(x_i)\Delta'_1(h) + O(h^2).$$

Let us now calculate $\Delta'_1(h)$. We have,

$$\begin{aligned} \Delta'_1(h) &= -\frac{1}{h}\Delta_1(h) \\ &+ \frac{1}{h^3} \left(\int_0^1 (x_i - s)K' \left(\frac{x_i - s}{h} \right) r(s)ds - \frac{1}{n} \sum_{j=1}^n (x_i - x_j)K' \left(\frac{x_i - x_j}{h} \right) r(x_j) \right). \end{aligned} \quad (36)$$

We have already proved that (see (26)),

$$\left| \frac{1}{h}\Delta_1(h) \right| \leq \frac{cst}{nh^3}. \quad (37)$$

Define $f(s) = (x_i - s)K' \left(\frac{x_i - s}{h} \right) r(s)$. Likewise, using (25), since f has a bounded first derivative,

$$\left| \int_0^1 (x_i - s)K' \left(\frac{x_i - s}{h} \right) r(s)ds - \frac{1}{n} \sum_{j=1}^n (x_i - x_j)K' \left(\frac{x_i - x_j}{h} \right) r(x_j) \right| \leq \frac{cst}{n},$$

and then by (36) and (37),

$$|\Delta'_1(h)| \leq \frac{cst}{nh^3}.$$

Consequently, for any $i = 1, \dots, n$,

$$u(x_i) \frac{\partial \mathbb{E}(\hat{r}(x_i))}{\partial h} = hu(x_i)r''(x_i) \int_{-1}^1 y^2 K(y) dy + O(h^2 + \frac{1}{nh^3}).$$

The proof of (29) is complete. Let us now prove (30), (31) and (32). Inequality (30) is true since

$$\sup_{h \in H_n} |u(x_i) \frac{\partial B(x_i, h)}{\partial h}| = \sup_{h \in H_n} |u(x_i) \frac{\partial \mathbb{E}(\hat{r}(x_i))}{\partial h}| \leq cst n^{-1/5}.$$

We have already proved that, for any $1 \leq i \leq n$, (see (28)),

$$\begin{aligned} & u(x_i)B(x_i, h) \\ &= \frac{h^2}{2}u(x_i)r''(x_i) \int y^2 K(y) dy + \Delta_1(h)u(x_i) + u(x_i) \int_{-1}^1 \Delta_2(h, y)K(y) dy, \end{aligned}$$

with,

$$|\Delta''_1(h)| \leq \frac{cst}{nh^4}, \quad |\Delta_1^{(3)}(h)| \leq \frac{cst}{nh^5}, \quad |\Delta''_2(h, y)| \leq cst, \quad |\Delta_2^{(3)}(h, y)| \leq cst.$$

Consequently, for any $h, h' \in H_n$, skipping the details,

$$\begin{aligned} u(x_i)|B''(x_i, h) - B''(x_i, h')| &\leq cst|h - h'| \\ u(x_i)|B''(x_i, h)| &\leq cst. \end{aligned}$$

The proof of Lemma 10 is complete.

B Central limit theorem for some triangular arrays of a quadratic form of a stationary MDS

Recall that $K - G$ is an even function, $[-1, 1]$ -supported, that the window h_n is a positive sequence satisfying

$$\lim_{n \rightarrow \infty} h_n = 0, \quad \lim_{n \rightarrow \infty} nh_n = \infty.$$

Define, for $i = 1, \dots, n$, $x_i = \frac{i}{n}$ and, for a positive constant C_K depending only on K ,

$$\begin{aligned} a_{i,n}(h_n) &= C_K \frac{h_n}{n} r''(x_i) u(x_i), \\ b_{i,j}(h_n) &= \frac{1}{n^2 h_n^2} (K - G) \left(\frac{x_i - x_j}{h_n} \right), \\ \tilde{b}_{i,j} &= b_{i,j}(h_n) (u(x_i) + u(x_j)). \end{aligned}$$

Let $(\epsilon_i)_{i \geq 0}$ be a centered sequence of stationary MD random variables with finite second moment σ^2 . The purpose of this section is to prove, letting

$$Y_{i,n}(h_n) = a_{i,n}(h_n)\epsilon_i + \sum_{j=1}^{i-1} \tilde{b}_{i,j}\epsilon_i\epsilon_j, \quad (38)$$

that

$$\frac{1}{s_n} \sum_{i=1}^n (Y_{i,n}(h_n) - \mathbb{E}(Y_{i,n}(h_n))),$$

converges in distribution to a normal law, with $s_n^2 = \text{Var}(\sum_{i=1}^n Y_{i,n}(h_n))$.

For this, we first control $\text{Var}(\sum_{i=1}^n Y_{i,n}(h_n))$. In all the proofs of this Appendix we denote by *cst* a generic constant independent of n and h_n and that may be different from line to line.

B.1 Control of the variance

Proposition 6. *Suppose that r is of class \mathcal{C}^2 on $[0, 1]$. Let $(\epsilon_i)_{i \geq 0}$ be a stationary sequence of centered MD random variables with finite fourth moment. Suppose that there exists a positive decreasing function Φ defined on \mathbb{R}^+ satisfying*

$$\sum_{s=1}^{\infty} s^2 \Phi(s) < \infty,$$

and for any $1 \leq i_1 \leq i_2 < i_3 \leq i_4 \leq i_5 \leq n$ such that $i_3 - i_2 \geq \max(i_2 - i_1, i_4 - i_3, i_5 - i_4)$

$$\begin{aligned} |\text{Cov}(\epsilon_{i_1}\epsilon_{i_2}, \epsilon_{i_3}\epsilon_{i_4})| &\leq \Phi(i_3 - i_2), \\ |\text{Cov}(\epsilon_{i_2}, \epsilon_{i_3}\epsilon_{i_4}\epsilon_{i_5})| &\leq \Phi(i_3 - i_2). \end{aligned}$$

Then

$$\begin{aligned} \text{Var} \left(\sum_{i=1}^n Y_{i,n}(h_n) \right) &= \frac{h_n^2 \sigma^2}{n} C_K^2 \int u^2(x) r''^2(x) dx + \frac{4\sigma^4}{n^2 h_n^3} \int_0^1 u^2(x) dx \int_0^1 (K - G)^2(u) du \\ &+ o\left(\frac{1}{n^2 h_n^3} + \frac{h_n^2}{n}\right). \end{aligned}$$

Proof of Proposition 6. We have,

$$\begin{aligned}
\text{Var}\left(\sum_{i=1}^n Y_{i,n}(h_n)\right) &= \sum_{i_1=1}^n \sum_{i_2=1}^n \text{Cov}(Y_{i_1,n}(h_n), Y_{i_2,n}(h_n)) \\
&= \sum_{i_1=1}^n \sum_{i_2=1}^n a_{i_1}(h_n) a_{i_2}(h_n) \text{Cov}(\epsilon_{i_1}, \epsilon_{i_2}) + \sum_{i_1=1}^n \sum_{i_2=1}^n a_{i_1}(h_n) \sum_{j_2=1}^{i_2-1} \tilde{b}_{i_2,j_2} \text{Cov}(\epsilon_{i_1}, \epsilon_{i_2} \epsilon_{j_2}) \\
&\quad + \sum_{i_1=1}^n \sum_{i_2=1}^n a_{i_2}(h_n) \sum_{j_1=1}^{i_1-1} \tilde{b}_{i_1,j_1} \text{Cov}(\epsilon_{i_2}, \epsilon_{i_1} \epsilon_{j_1}) \\
&\quad + \sum_{i_1=1}^n \sum_{i_2=1}^n \sum_{j_1=1}^{i_1-1} \sum_{j_2=1}^{i_2-1} \tilde{b}_{i_1,j_1} \tilde{b}_{i_2,j_2} \text{Cov}(\epsilon_{i_1} \epsilon_{j_1}, \epsilon_{i_2} \epsilon_{j_2}) =: I + II + III + IV.
\end{aligned}$$

Control of I. Clearly, since $\text{Cov}(\epsilon_{i_1}, \epsilon_{i_2}) = 0$ for $i_1 \neq i_2$,

$$I = \sum_{i_1=1}^n \sum_{i_2=1}^n a_{i_1}(h_n) a_{i_2}(h_n) \text{Cov}(\epsilon_{i_1}, \epsilon_{i_2}) = \sum_{i_1=1}^n a_{i_1}^2(h_n) \sigma^2. \quad (39)$$

We have,

$$\sum_{i_1=1}^n a_{i_1}^2(h_n) = C_K^2 \frac{h_n^2}{n^2} \sum_{i_1=1}^n r''^2(x_i) u^2(x_i),$$

and since

$$\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{i_1=1}^n r''^2(x_i) u^2(x_i) = \int_0^1 r''^2(x) u^2(x) dx$$

we deduce that

$$\sum_{i_1=1}^n a_{i_1}^2(h_n) \sigma^2 = \sigma^2 C_K^2 \frac{h_n^2}{n} \int_0^1 r''^2(x) u^2(x) dx + o\left(\frac{h_n^2}{n}\right). \quad (40)$$

Combining (39) and (40), it yields that

$$I = \sigma^2 C_K^2 \frac{h_n^2}{n} \int_0^1 u^2(x) r''^2(x) dx + o\left(\frac{h_n^2}{n}\right).$$

Control of II and III. We only concentrate on II (the control of III is similar). We need the following bound, (recall that $j_2 < i_2$ and that $\text{Cov}(\epsilon_{i_1}, \epsilon_{j_2} \epsilon_{i_2}) = \mathbb{E}(\epsilon_{i_1} \epsilon_{j_2} \epsilon_{i_2}) = \text{Cov}(\epsilon_{i_1} \epsilon_{j_2}, \epsilon_{i_2}) = \text{Cov}(\epsilon_{j_2}, \epsilon_{i_1} \epsilon_{i_2})$),

$$\begin{aligned}
|\text{Cov}(\epsilon_{i_1}, \epsilon_{i_2} \epsilon_{j_2})| &\leq \Phi(j_2 - i_1) \mathbb{1}_{i_1 < j_2 < i_2} \mathbb{1}_{i_2 - j_2 \leq j_2 - i_1} + \Phi(i_2 - j_2) \mathbb{1}_{i_1 < j_2 < i_2} \mathbb{1}_{i_2 - j_2 > j_2 - i_1} \\
&\quad + \Phi(i_1 - i_2) \mathbb{1}_{j_2 < i_2 < i_1} \mathbb{1}_{i_1 - i_2 > i_2 - j_2} + \Phi(i_2 - j_2) \mathbb{1}_{j_2 < i_2 < i_1} \mathbb{1}_{i_1 - i_2 \leq i_2 - j_2} \\
&\quad + \Phi(i_1 - j_2) \mathbb{1}_{j_2 < i_1 < i_2} \mathbb{1}_{i_1 - j_2 > i_2 - i_1} + \Phi(i_2 - i_1) \mathbb{1}_{j_2 < i_1 < i_2} \mathbb{1}_{i_1 - j_2 \leq i_2 - i_1}.
\end{aligned}$$

Thanks to the previous bound, we decompose the sum in II into six sums according to whether $i_1 < j_2 < i_2$ or $i_2 - j_2 \leq j_2 - i_1, \dots, j_2 < i_1 < i_2$ or $i_1 - j_2 \leq i_2 - i_1$. We only need to control one sum among the six, since the calculations are similar. We have (recall that $|\tilde{b}_{i,j}| \leq 2\|u\|_\infty\|K - G\|_\infty/n^2h_n^2$),

$$\begin{aligned}
& \sum_{i_1=1}^n \sum_{i_2=1}^n a_{i_1}(h_n) \sum_{j_2=1}^{i_2-1} \tilde{b}_{i_2, j_2} |\text{Cov}(\epsilon_{i_1}, \epsilon_{i_2} \epsilon_{j_2})| \mathbb{1}_{i_1 < j_2 < i_2} \mathbb{1}_{i_2 - j_2 \leq j_2 - i_1} \\
& \leq cst \frac{1}{n^2 h_n^2} \sum_{i_1=1}^n |a_{i_1}(h_n)| \sum_{i_2=1}^n \sum_{j_2=1}^{i_2-1} \Phi(j_2 - i_1) \mathbb{1}_{i_1 < j_2 < i_2} \mathbb{1}_{i_2 - j_2 \leq j_2 - i_1} \\
& \leq cst \frac{1}{n^2 h_n^2} \sum_{i_1=1}^n |a_{i_1}(h_n)| \sum_{j_2=i_1}^n \sum_{i_2=j_2+1}^{2j_2-i_1} \Phi(j_2 - i_1) \mathbb{1}_{i_1 < j_2 < i_2} \mathbb{1}_{i_2 - j_2 \leq j_2 - i_1} \\
& \leq cst \frac{1}{n^2 h_n^2} \sum_{i_1=1}^n |a_{i_1}(h_n)| \sum_{s=1}^{n-i_1} \sum_{i_2=j_2+1}^{j_2+s} \Phi(s) \mathbb{1}_{i_2 - j_2 \leq s} \\
& \leq cst \frac{1}{n^2 h_n^2} \sum_{i_1=1}^n |a_{i_1}(h_n)| \sum_{s=1}^{\infty} s \Phi(s),
\end{aligned}$$

which is of order $O(\frac{1}{n^2 h_n})$ since $\sum_{i_1=1}^n |a_{i_1}(h_n)| = O(h_n)$ and $\sum_{s=1}^{\infty} s \Phi(s) < \infty$. With the same way, we prove that

$$II + III = O\left(\frac{1}{n^2 h_n}\right) = o\left(\frac{1}{n^2 h_n^3}\right).$$

Control of IV. Recall that

$$IV = \sum_{i_1=1}^n \sum_{i_2=1}^n \sum_{j_1=1}^{i_1-1} \sum_{j_2=1}^{i_2-1} \tilde{b}_{i_1, j_1} \tilde{b}_{i_2, j_2} \text{Cov}(\epsilon_{i_1} \epsilon_{j_1}, \epsilon_{i_2} \epsilon_{j_2}).$$

We decompose the sums in IV according to the following cases:

- (a) sums over $j_1 < i_1$ and $(i_1, j_1) = (i_2, j_2)$,
- (b) sums over $j_1 < i_1 \leq j_2 < i_2$ (or $j_2 < i_2 \leq j_1 < i_1$)
- (c) sums over $j_1 \leq j_2 \leq i_1 \leq i_2$, $(i_1, j_1) \neq (i_2, j_2)$, $j_1 < i_1$ and $j_2 < i_2$
- (c') sums over $j_2 \leq j_1 \leq i_2 \leq i_1$, $(i_1, j_1) \neq (i_2, j_2)$, $j_1 < i_1$ and $j_2 < i_2$
- (d) sums over $j_1 \leq j_2 < i_2 \leq i_1$ and $(i_1, j_1) \neq (i_2, j_2)$
- (d') $j_2 \leq j_1 < i_1 \leq i_2$ and $(i_1, j_1) \neq (i_2, j_2)$

We denote by $S_a, S_b, S_c, S_{c'}, S_d, S_{d'}$ the sums corresponding to these items, so that

$$IV = S_a + S_b + S_c + S_{c'} + S_d + S_{d'}. \quad (41)$$

Then our purpose is to control each of those sums $S_a, S_b, S_c, S_{c'}, S_d$ and $S_{d'}$.

Control of S_a . We get, using the fact that $\text{Var}(\epsilon_{i_1} \epsilon_{j_1}) = \sigma^4 + \text{Cov}(\epsilon_{i_1}^2, \epsilon_{j_1}^2) - \text{Cov}^2(\epsilon_{i_1}, \epsilon_{j_1}) = \sigma^4 + \text{Cov}(\epsilon_{i_1}^2, \epsilon_{j_1}^2)$, for $i_1 \neq j_1$,

$$\begin{aligned} S_a &= \sum_{i_1=1}^n \sum_{j_1=1}^{i_1-1} \tilde{b}_{i_1, j_1}^2 \text{Var}(\epsilon_{i_1} \epsilon_{j_1}) \\ &= \sum_{i_1=1}^n \sum_{j_1=1}^{i_1-1} \tilde{b}_{i_1, j_1}^2 \sigma^4 + \sum_{i_1=1}^n \sum_{j_1=1}^{i_1-1} \tilde{b}_{i_1, j_1}^2 \text{Cov}(\epsilon_{i_1}^2, \epsilon_{j_1}^2). \end{aligned}$$

This term is controlled as in the i.i.d. case, see [Härdle, W., Hall, P. and Marron, J. S. \(1988\)](#) and [Girard, D. \(1998\)](#), nevertheless, we give the calculations details for the sake of clarity,

$$\begin{aligned} &\sum_{i_1=1}^n \sum_{j_1=1}^{i_1-1} \tilde{b}_{i_1, j_1}^2 \\ &= \frac{1}{n^4 h_n^4} \sum_{i=1}^n \sum_{j=1}^{i-1} (u(x_i) + u(x_j))^2 \left(K\left(\frac{x_i - x_j}{h_n}\right) - G\left(\frac{x_i - x_j}{h_n}\right) \right)^2. \end{aligned} \quad (42)$$

Clearly,

$$\begin{aligned} &\frac{1}{n^2 h_n} \sum_{i=1}^n \sum_{j=1}^{i-1} (u(i/n) + u(j/n))^2 \left(K\left(\frac{x_i - x_j}{h_n}\right) - G\left(\frac{x_i - x_j}{h_n}\right) \right)^2 \\ &= \frac{1}{n^2 h_n} \sum_{s=1}^{n-1} \sum_{i=s+1}^n (u(i/n) + u(\frac{i-s}{n}))^2 \left(K\left(\frac{s}{nh_n}\right) - G\left(\frac{s}{nh_n}\right) \right)^2 \\ &= \frac{4}{n^2 h_n} \sum_{s=1}^{n-1} \sum_{i=1}^n u^2\left(\frac{i}{n}\right) \left(K\left(\frac{s}{nh_n}\right) - G\left(\frac{s}{nh_n}\right) \right)^2 \\ &\quad + O\left(\frac{1}{n^2 h_n} \sum_{s=1}^{n-1} s \left(K\left(\frac{s}{nh_n}\right) - G\left(\frac{s}{nh_n}\right) \right)^2 \right), \end{aligned} \quad (43)$$

the last bound is obtained by noticing that

$$\left| \sum_{i=s+1}^n (u(i/n) + u(\frac{i-s}{n}))^2 - 4 \sum_{i=1}^n u^2\left(\frac{i}{n}\right) \right| \leq cst s.$$

Now, we have using (25) (recall that $K - G$ is $[-1, 1]$ -supported and that $[nh_n] \leq n - 1$)

$$\begin{aligned} &\frac{1}{n^2 h_n} \sum_{s=1}^{n-1} s \left(K\left(\frac{s}{nh_n}\right) - G\left(\frac{s}{nh_n}\right) \right)^2 = \frac{nh_n}{n^2 h_n} \sum_{s=1}^{nh_n} \frac{s}{nh_n} \left(K\left(\frac{s}{nh_n}\right) - G\left(\frac{s}{nh_n}\right) \right)^2 \\ &= O(h_n). \end{aligned} \quad (44)$$

We have, again due to (25),

$$\frac{1}{n^2 h_n} \sum_{s=1}^{n-1} \sum_{i=1}^n u^2\left(\frac{i}{n}\right) (K - G)^2\left(\frac{s}{nh_n}\right) = \int_0^1 u^2(x) dx \int_0^1 (K - G)^2(y) dy + o(1). \quad (45)$$

Collecting (42), (43), (44) and (45), we deduce that

$$\sum_{i_1=1}^n \sum_{j_1=1}^{i_1-1} \tilde{b}_{i_1, j_1}^2 = \frac{4}{n^2 h_n^3} \int_0^1 u^2(x) dx \int_0^1 (K - G)^2(u) du + o\left(\frac{1}{n^2 h_n^3}\right). \quad (46)$$

Our purpose now is to prove that,

$$\sum_{i_1=1}^n \sum_{j_1=1}^{i_1-1} \tilde{b}_{i_1, j_1}^2 \text{Cov}(\epsilon_{i_1}^2, \epsilon_{j_1}^2) = o\left(\frac{1}{n^2 h_n^3}\right), \quad (47)$$

so that,

$$S_a = \frac{4\sigma^4}{n^2 h_n^3} \int_0^1 u^2(x) dx \int_0^1 (K - G)^2(u) du + o\left(\frac{1}{n^2 h_n^3}\right). \quad (48)$$

We know that $|\text{Cov}(\epsilon_{i_1}^2, \epsilon_{j_1}^2)| \leq cst \Phi(i_1 - j_1)$. We get, arguing as before,

$$\begin{aligned} & \sum_{i_1=1}^n \sum_{j_1=1}^{i_1-1} \tilde{b}_{i_1, j_1}^2 |\text{Cov}(\epsilon_{i_1}^2, \epsilon_{j_1}^2)| \\ & \leq cst \sum_{i_1=1}^n \sum_{j_1=1}^{i_1-1} b_{i_1, j_1}^2(h_n) \Phi(i_1 - j_1) \\ & \leq \frac{cst}{n^4 h_n^4} \sum_{s=1}^{[nh_n]} (n - s) \left(K\left(\frac{s}{nh_n}\right) - G\left(\frac{s}{nh_n}\right) \right)^2 \Phi(s) \\ & \leq \frac{cst}{n^3 h_n^4} \sum_{s=1}^{\infty} \Phi(s) = o\left(\frac{1}{n^2 h_n^3}\right). \end{aligned}$$

This proves (47) and then (48) holds.

Control of S_b . We now consider the sums over $j_1 < i_1 \leq j_2 < i_2$ (the case $j_2 < i_2 \leq j_1 < i_1$ is exactly the same). We discuss the following subcases. If $i_2 - j_2 \geq \max(i_1 - j_1, j_2 - i_1)$, then

$$\begin{aligned} & |\text{Cov}(\epsilon_{i_1} \epsilon_{j_1}, \epsilon_{i_2} \epsilon_{j_2})| \\ & = |\text{Cov}(\epsilon_{j_1} \epsilon_{i_1} \epsilon_{j_2}, \epsilon_{i_2}) - \text{Cov}(\epsilon_{i_1}, \epsilon_{j_1}) \text{Cov}(\epsilon_{i_2}, \epsilon_{j_2})| \\ & = |\text{Cov}(\epsilon_{j_1} \epsilon_{i_1} \epsilon_{j_2}, \epsilon_{i_2})| \leq \Phi(i_2 - j_2). \end{aligned}$$

If $j_1 < i_1 \leq j_2 < i_2$ and $i_1 - j_1 \geq \max(i_2 - j_2, j_2 - i_1)$, then

$$|\text{Cov}(\epsilon_{i_1} \epsilon_{j_1}, \epsilon_{i_2} \epsilon_{j_2})| \leq \Phi(i_1 - j_1).$$

If $j_1 < i_1 \leq j_2 < i_2$ and $j_2 - i_1 \geq \max(i_2 - j_2, i_1 - j_1)$, then

$$|\text{Cov}(\epsilon_{i_1} \epsilon_{j_1}, \epsilon_{i_2} \epsilon_{j_2})| \leq \Phi(j_2 - i_1).$$

The sum $\sum_{j_1 < i_1 \leq j_2 < i_2} \tilde{b}_{i_1, j_1} \tilde{b}_{i_2, j_2} |\text{Cov}(\epsilon_{i_1} \epsilon_{j_1}, \epsilon_{i_2} \epsilon_{j_2})|$ is then decomposed into three sums according to whether $i_2 - j_2 \geq \max(i_1 - j_1, j_2 - i_1)$, $i_1 - j_1 \geq \max(i_2 - j_2, j_2 - i_1)$ or $j_2 - i_1 \geq \max(i_2 - j_2, i_1 - j_1)$. We only discuss one sum among these three sums (since the calculations are similar). We have,

$$\begin{aligned} & \sum_{j_1 < i_1 \leq j_2 < i_2} \tilde{b}_{i_1, j_1} \tilde{b}_{i_2, j_2} |\text{Cov}(\epsilon_{i_1} \epsilon_{j_1}, \epsilon_{i_2} \epsilon_{j_2})| \mathbb{1}_{i_2 - j_2 \geq \max(i_1 - j_1, j_2 - i_1)} \\ & \leq \frac{cst}{n^4 h_n^4} \sum_{j_1 < i_1 \leq j_2 < i_2} \mathbb{1}_{i_2 - j_2 \geq \max(i_1 - j_1, j_2 - i_1)} \Phi(i_2 - j_2) \\ & \leq \frac{cst}{n^4 h_n^4} \sum_{j_1=1}^{n-1} \sum_{i_1=j_1+1}^n \sum_{j_2=i_1}^{n-1} \sum_{i_2=j_2+1}^n \mathbb{1}_{i_2 - j_2 \geq \max(i_1 - j_1, j_2 - i_1)} \Phi(i_2 - j_2) \\ & \leq \frac{cst}{n^4 h_n^4} \sum_{j_1=1}^{n-1} \sum_{l=1}^{n-j_1} \sum_{j_2=l+j_1}^{n-1} \sum_{s=1}^{n-j_2} \mathbb{1}_{s \geq \max(l, j_2 - j_1 - l)} \Phi(s) \\ & \leq \frac{cst}{n^3 h_n^4} \sum_{s=1}^{\infty} s^2 \Phi(s) = o\left(\frac{1}{n^2 h_n^3}\right). \end{aligned}$$

From this, we deduce that

$$S_b = o\left(\frac{1}{n^2 h_n^3}\right). \quad (49)$$

Control of S_c . We now discuss the sums over $(i_1, j_1) \neq (i_2, j_2)$, $j_1 < i_1$, $j_2 < i_2$ and $j_1 \leq j_2 \leq i_1 \leq i_2$. We write,

$$\begin{aligned} \text{Cov}(\epsilon_{i_1} \epsilon_{j_1}, \epsilon_{i_2} \epsilon_{j_2}) &= \mathbb{E}(\epsilon_{j_1} \epsilon_{j_2} \epsilon_{i_1} \epsilon_{i_2}) - \text{Cov}(\epsilon_{i_1}, \epsilon_{j_1}) \text{Cov}(\epsilon_{i_2}, \epsilon_{j_2}) \\ &= \mathbb{E}(\epsilon_{j_1} \epsilon_{j_2} \epsilon_{i_1} \epsilon_{i_2}). \end{aligned}$$

We discuss the following subcases.

- $j_1 \leq j_2 \leq i_1 \leq i_2$, $j_2 - j_1 \geq \max(i_1 - j_2, i_2 - i_1)$. In this case

$$|\mathbb{E}(\epsilon_{j_1} \epsilon_{j_2} \epsilon_{i_1} \epsilon_{i_2})| \leq \Phi(j_2 - j_1)$$

so that $|\text{Cov}(\epsilon_{i_1} \epsilon_{j_1}, \epsilon_{i_2} \epsilon_{j_2})| \leq \Phi(j_2 - j_1)$. We have,

$$\begin{aligned}
& \sum_{j_1 \leq j_2 < i_1 \leq i_2} \tilde{b}_{i_1, j_1} \tilde{b}_{i_2, j_2} \mathbb{1}_{j_2 - j_1 \geq \max(i_1 - j_2, i_2 - i_1)} |\text{Cov}(\epsilon_{i_1} \epsilon_{j_1}, \epsilon_{i_2} \epsilon_{j_2})| \\
& \leq \frac{cst}{n^4 h_n^4} \sum_{j_1 \leq j_2 < i_1 \leq i_2} \mathbb{1}_{j_2 - j_1 \geq \max(i_1 - j_2, i_2 - i_1)} \Phi(j_2 - j_1) \\
& \leq \frac{cst}{n^4 h_n^4} \sum_{j_1=1}^{n-1} \sum_{j_2=j_1}^{n-1} \sum_{i_1=j_2}^{n-1} \sum_{i_2=i_1}^n \mathbb{1}_{j_2 - j_1 \geq \max(i_1 - j_2, i_2 - i_1)} \Phi(j_2 - j_1) \\
& \leq \frac{cst}{n^4 h_n^4} \sum_{j_1=1}^{n-1} \sum_{s=1}^{n-j_1} \sum_{i_1=j_2}^{s+j_2} \sum_{i_2=i_1}^{s+i_1} \mathbb{1}_{s \geq \max(i_1 - j_2, i_2 - i_1)} \Phi(s) \\
& \leq \frac{cst}{n^3 h_n^4} \sum_{s=1}^{\infty} s^2 \Phi(s) = o\left(\frac{1}{n^2 h_n^3}\right). \tag{50}
\end{aligned}$$

- $j_1 \leq j_2 \leq i_1 \leq i_2$, $i_2 - i_1 \geq \max(j_2 - j_1, i_1 - j_2)$. Then

$$|\mathbb{E}(\epsilon_{j_1} \epsilon_{j_2} \epsilon_{i_1} \epsilon_{i_2})| \leq \Phi(i_2 - i_1) \text{Cov}(\epsilon_{j_1}, \epsilon_{i_1}) \text{Cov}(\epsilon_{j_2}, \epsilon_{i_2}) = 0,$$

and then

$$|\text{Cov}(\epsilon_{i_1} \epsilon_{j_1}, \epsilon_{i_2} \epsilon_{j_2})| \leq \Phi(i_2 - i_1).$$

We prove, noting that in this case $i_2 - i_1 \geq \frac{i_1 - j_1}{2}$ and $i_1 - 2(i_2 - i_1) \leq j_1 \leq i_1$,

$$\begin{aligned}
& \sum_{j_1 \leq j_2 \leq i_1 \leq i_2} \tilde{b}_{i_1, j_1} \tilde{b}_{i_2, j_2} \mathbb{1}_{i_2 - i_1 \geq \max(j_2 - j_1, i_1 - j_2)} |\text{Cov}(\epsilon_{i_1} \epsilon_{j_1}, \epsilon_{i_2} \epsilon_{j_2})| \\
& \leq cst \sum_{j_1 \leq j_2 < i_1 \leq i_2} b_{i_1, j_1}(h_n) b_{i_2, j_2}(h_n) \mathbb{1}_{i_2 - i_1 \geq \max(j_2 - j_1, i_1 - j_2)} \Phi(i_2 - i_1) \\
& \leq \frac{cst}{n^4 h_n^4} \sum_{i_1=1}^n \sum_{i_2=i_1}^n \sum_{j_1=i_1-2(i_2-i_1)}^{i_1} \sum_{j_2=j_1}^{j_1+(i_2-i_1)} \Phi(i_2 - i_1) \\
& \leq \frac{cst}{n^3 h_n^4} \sum_{s=1}^{\infty} s^2 \Phi(s) = o\left(\frac{1}{n^2 h_n^3}\right). \tag{51}
\end{aligned}$$

- $j_1 \leq j_2 \leq i_1 \leq i_2$, $i_1 - j_2 \geq \max(j_2 - j_1, i_2 - i_1)$, $(i_1, j_1) \neq (i_2, j_2)$, $j_1 < i_1, j_2 < i_2$. We write,

$$\begin{aligned}
& \text{Cov}(\epsilon_{i_1} \epsilon_{j_1}, \epsilon_{i_2} \epsilon_{j_2}) = \mathbb{E}(\epsilon_{j_1} \epsilon_{j_2} \epsilon_{i_1} \epsilon_{i_2}) - \text{Cov}(\epsilon_{i_1}, \epsilon_{j_1}) \text{Cov}(\epsilon_{i_2}, \epsilon_{j_2}) \\
& = \text{Cov}(\epsilon_{j_1} \epsilon_{j_2}, \epsilon_{i_1} \epsilon_{i_2}) + \text{Cov}(\epsilon_{j_1}, \epsilon_{j_2}) \text{Cov}(\epsilon_{i_1}, \epsilon_{i_2}).
\end{aligned}$$

We have,

$$|\text{Cov}(\epsilon_{j_1} \epsilon_{j_2}, \epsilon_{i_1} \epsilon_{i_2})| \leq \Phi(i_1 - j_2).$$

We get, arguing as before,

$$\sum_{j_1 \leq j_2 \leq i_1 \leq i_2} \tilde{b}_{i_1, j_1} \tilde{b}_{i_2, j_2} \mathbb{1}_{i_1 - j_2 \geq \max(j_2 - j_1, i_2 - i_1)} \Phi(i_1 - j_2) = o\left(\frac{1}{n^2 h_n^3}\right)$$

so that, (noting by $\sum_{j_1 \leq j_2 \leq i_1 \leq i_2}^*$ the sum over $j_1 \leq j_2 \leq i_1 \leq i_2$, $(i_1, j_1) \neq (i_2, j_2)$, $j_1 < i_1, j_2 < i_2$)

$$\begin{aligned} & \sum_{j_1 \leq j_2 \leq i_1 \leq i_2}^* \tilde{b}_{i_1, j_1} \tilde{b}_{i_2, j_2} \mathbb{1}_{i_1 - j_2 \geq \max(j_2 - j_1, i_2 - i_1)} \text{Cov}(\epsilon_{i_1} \epsilon_{j_1}, \epsilon_{i_2} \epsilon_{j_2}) \\ &= \sum_{j_1 \leq j_2 \leq i_1 \leq i_2}^* \tilde{b}_{i_1, j_1} \tilde{b}_{i_2, j_2} \mathbb{1}_{i_1 - j_2 \geq \max(j_2 - j_1, i_2 - i_1)} \text{Cov}(\epsilon_{j_1}, \epsilon_{j_2}) \text{Cov}(\epsilon_{i_1}, \epsilon_{i_2}) + o\left(\frac{1}{n^2 h_n^3}\right) \\ &= 0 + o\left(\frac{1}{n^2 h_n^3}\right), \quad \text{by definition of the sum } \sum_{j_1 \leq j_2 \leq i_1 \leq i_2}^*. \end{aligned} \tag{52}$$

We deduce, collecting all the bounds (50), (51) and (52)

$$S_c = o\left(\frac{1}{n^2 h_n^3}\right).$$

Control of S'_c . We now discuss the sums over $(i_1, j_1) \neq (i_2, j_2)$, $j_1 < i_1$, $j_2 < i_2$ and $j_2 \leq j_1 \leq i_2 \leq i_1$. As for the control of S_c , we prove that

$$S'_c = o\left(\frac{1}{n^2 h_n^3}\right). \tag{53}$$

Control of S_d . We now control the item (d), i.e, the sums over $j_1 \leq j_2 < i_2 \leq i_1$ $(i_1, j_1) \neq (i_2, j_2)$. Here again, we discuss the following subcases,

- If $i_1 - i_2 \geq \max(i_2 - j_2, j_2 - j_1)$ then $|\mathbb{E}(\epsilon_{j_1} \epsilon_{j_2} \epsilon_{i_2} \epsilon_{i_1})| \leq \Phi(i_1 - i_2)$, so that

$$|\text{Cov}(\epsilon_{i_1} \epsilon_{j_1}, \epsilon_{i_2} \epsilon_{j_2})| \leq \Phi(i_1 - i_2).$$

We have,

$$\begin{aligned} & \sum_{j_1 \leq j_2 < i_2 \leq i_1} \tilde{b}_{i_1, j_1} \tilde{b}_{i_2, j_2} \mathbb{1}_{i_1 - i_2 \geq \max(i_2 - j_2, j_2 - j_1)} |\text{Cov}(\epsilon_{i_1} \epsilon_{j_1}, \epsilon_{i_2} \epsilon_{j_2})| \\ & \leq \frac{cst}{n^4 h_n^4} \sum_{j_1=1}^n \sum_{j_2=j_1}^n \sum_{i_2=j_2}^n \sum_{i_1=i_2}^n \Phi(i_1 - i_2) \mathbb{1}_{i_1 - i_2 \geq \max(i_2 - j_2, j_2 - j_1)} \\ & \leq \frac{cst}{n^4 h_n^4} cst \sum_{j_1=1}^n \sum_{j_2=j_1}^n \sum_{i_2=j_2}^n \sum_{s=0}^{n-i_2} \Phi(s) \mathbb{1}_{s \geq \max(i_2 - j_2, j_2 - j_1)} \\ & \leq \frac{cst}{n^3 h_n^4} cst \sum_{s=1}^{\infty} s^2 \Phi(s) = o\left(\frac{1}{n^2 h_n^3}\right). \end{aligned}$$

- If $j_2 - j_1 \geq \max(i_1 - i_2, i_2 - j_2)$ then $|\mathbb{E}(\epsilon_{j_1} \epsilon_{j_2} \epsilon_{i_2} \epsilon_{i_1})| \leq \Phi(j_2 - j_1)$ and $\text{Cov}(\epsilon_{i_1}, \epsilon_{j_1})\text{Cov}(\epsilon_{i_2}, \epsilon_{j_2}) = 0$, so that

$$|\text{Cov}(\epsilon_{i_1} \epsilon_{j_1}, \epsilon_{i_2} \epsilon_{j_2})| \leq \Phi(j_2 - j_1).$$

As before,

$$\begin{aligned} & \sum_{j_1 \leq j_2 < i_2 \leq i_1} \tilde{b}_{i_1, j_1} \tilde{b}_{i_2, j_2} \mathbb{1}_{j_2 - j_1 \geq \max(i_1 - i_2, i_2 - j_2)} |\text{Cov}(\epsilon_{i_1} \epsilon_{j_1}, \epsilon_{i_2} \epsilon_{j_2})| \\ & \leq \frac{cst}{n^4 h_n^4} \sum_{j_1=1}^n \sum_{j_2=j_1}^n \sum_{i_2=j_2}^n \sum_{i_1=i_2}^n \Phi(j_2 - j_1) \mathbb{1}_{j_2 - j_1 \geq \max(i_1 - i_2, i_2 - j_2)} \\ & \leq \frac{cst}{n^3 h_n^4} \sum_{s=1}^{\infty} s^2 \Phi(s) = o\left(\frac{1}{n^2 h_n^3}\right). \end{aligned}$$

- If $i_2 - j_2 \geq \max(i_1 - i_2, j_2 - j_1)$ then $|\text{Cov}(\epsilon_{j_1} \epsilon_{j_2}, \epsilon_{i_2} \epsilon_{i_1})| \leq \Phi(i_2 - j_2)$ and $\text{Cov}(\epsilon_{i_1}, \epsilon_{j_1})\text{Cov}(\epsilon_{i_2}, \epsilon_{j_2}) = 0$. Write,

$$\begin{aligned} \text{Cov}(\epsilon_{i_1} \epsilon_{j_1}, \epsilon_{i_2} \epsilon_{j_2}) &= \mathbb{E}(\epsilon_{i_1} \epsilon_{j_1} \epsilon_{i_2} \epsilon_{j_2}) \\ &= \text{Cov}(\epsilon_{j_1} \epsilon_{j_2}, \epsilon_{i_2} \epsilon_{i_1}) + \text{Cov}(\epsilon_{j_1}, \epsilon_{j_2})\text{Cov}(\epsilon_{i_2}, \epsilon_{i_1}). \end{aligned}$$

We have, as before,

$$\sum_{j_1 \leq j_2 < i_2 \leq i_1} \tilde{b}_{i_1, j_1} \tilde{b}_{i_2, j_2} \mathbb{1}_{i_2 - j_2 \geq \max(i_1 - i_2, j_2 - j_1)} \Phi(i_2 - j_2) = o\left(\frac{1}{n^2 h_n^3}\right),$$

and thus

$$\begin{aligned} & \sum_{j_1 \leq j_2 < i_2 \leq i_1} \tilde{b}_{i_1, j_1} \tilde{b}_{i_2, j_2} \mathbb{1}_{i_2 - j_2 \geq \max(i_1 - i_2, j_2 - j_1)} \text{Cov}(\epsilon_{i_1} \epsilon_{j_1}, \epsilon_{i_2} \epsilon_{j_2}) \\ &= \sum_{j_1 \leq j_2 < i_2 \leq i_1} \tilde{b}_{i_1, j_1} \tilde{b}_{i_2, j_2} \mathbb{1}_{i_2 - j_2 \geq \max(i_1 - i_2, j_2 - j_1)} \mathbb{E}(\epsilon_{j_1} \epsilon_{j_2}) \mathbb{E}(\epsilon_{i_2} \epsilon_{i_1}) \mathbb{1}_{(i_1, j_1) \neq (i_2, j_2)} \\ &+ o\left(\frac{1}{n^2 h_n^3}\right) = o\left(\frac{1}{n^2 h_n^3}\right). \end{aligned}$$

Collecting all the previous inequalities, we deduce that

$$S_d = o\left(\frac{1}{n^2 h_n^3}\right).$$

Control of S'_d . We now control the item (d'), i.e, the sums over $j_2 \leq j_1 < i_1 \leq i_2$ (i_1, j_1) \neq (i_2, j_2). As for the control of S_d , we obtain

$$S'_d = o\left(\frac{1}{n^2 h_n^3}\right).$$

Since $IV = S_a + S_b + S_c + S'_c + S_d + S'_d$, we deduce that

$$IV = \frac{4\sigma^4}{n^2 h_n^3} \int_0^1 u^2(x) dx \int_0^1 (K - G)^2(u) du + o\left(\frac{1}{n^2 h_n^3}\right).$$

The proof of Proposition 6 is complete by collecting the terms I, II, III, IV .

B.2 Central limit theorem

The following proposition proves a central limit theorem for the partial sum of $(Y_{in}(h_n))_{1 \leq i \leq n}$, as defined in (38).

Proposition 7. *Let $(\epsilon_i)_{i \geq 0}$ be a stationary sequence of centered martingale difference random variables relative to the filtration $\mathcal{F}_i = \sigma(\epsilon_1, \dots, \epsilon_i)$ such that $\mathbb{E}(\epsilon_1^8) < \infty$. Moreover, suppose that there exists a positive decreasing function Φ defined on \mathbb{R}^+ satisfying*

$$\sum_{s=1}^{\infty} s^4 \Phi(s) < \infty,$$

and for any positive integer $q \leq 6$, $1 \leq i_1 \leq \dots \leq i_k < i_{k+1} \leq \dots \leq i_q \leq n$ such that $i_{k+1} - i_k \geq \max_{1 \leq l \leq k} (i_{l+1} - i_l)$

$$|\text{Cov}(\epsilon_{i_1} \dots \epsilon_{i_k}, \epsilon_{i_{k+1}} \dots \epsilon_{i_q})| \leq \Phi(i_{k+1} - i_k).$$

Let $Y_{i,n}(h_n)$ be defined as in (38) with $h_n = cn^{-1/5}$. Then

$$n^{7/10} \sum_{i=1}^n Y_{i,n}(h_n) \Longrightarrow \mathcal{N}(0, V),$$

where \Longrightarrow denotes the convergence in distribution when n tends to infinity and the variance V is defined by,

$$V = c^2 C_K^2 \sigma^2 \int_0^1 u^2(x) r''^2(x) dx + \frac{4}{c^3} \sigma^4 \int_0^1 u^2(x) dx \int_0^1 (K - G)^2(u) du,$$

with $\sigma^2 = \mathbb{E}(\epsilon_1^2)$.

Proof of Proposition 7.

Since $(\epsilon_i)_{i \geq 0}$ is a stationary sequence of centered martingale-difference relative to the filtration $\mathcal{F}_i = \sigma(\epsilon_1, \dots, \epsilon_i)$ then the sequence $(Y_{i,n}(h_n))_{1 \leq i \leq n}$ (defined in (38)) is also a martingale-difference relative to the filtration \mathcal{F}_i , in fact $Y_{i,n}(h_n)$ is \mathcal{F}_i -measurable and

$$\mathbb{E}(Y_{i,n}(h_n) | \mathcal{F}_{i-1}) = a_{i,n}(h_n) \mathbb{E}(\epsilon_i | \mathcal{F}_{i-1}) + \sum_{j=1}^{i-1} \tilde{b}_{i,j} \epsilon_j \mathbb{E}(\epsilon_i | \mathcal{F}_{i-1}) = 0.$$

For a martingale-difference sequence $(Y_{i,n})$, the central limit theorem (CLT, in short) for $\frac{1}{s_n} \sum_{i=1}^n Y_{i,n}$, with $s_n^2 = \text{Var}(\sum_{i=1}^n Y_{i,n})$, follows by using the following result due to [McLeish, D. L. \(1974\)](#).

Theorem 2. *Let $(Y_{i,n})_{i \geq 0}$ be a martingale-difference series relative to the filtration \mathcal{F}_i . If the following conditions hold,*

$$(T_1) \max_{1 \leq i \leq n} \frac{|Y_{i,n}|}{s_n} \text{ is uniformly bounded in } L_2\text{-norm,}$$

$$(T_2) \lim_{n \rightarrow \infty} \max_{1 \leq i \leq n} \frac{|Y_{i,n}|}{s_n} = 0 \text{ in probability,}$$

$$(T_3) \lim_{n \rightarrow \infty} \frac{1}{s_n^2} \sum_{i=1}^n Y_{i,n}^2 = 1 \text{ in probability,}$$

then $\frac{1}{s_n} \sum_{i=1}^n Y_{i,n}$ converges in distribution to the standard normal law.

Proposition 7 is then proved if Conditions (T_1) , (T_2) and (T_3) are satisfied by the sequence $Y_{i,n}(h_n) =: Y_{i,n}$ as defined in (38). These Conditions (T_1) , (T_2) and (T_3) follow immediately from the following Lemmas 12 and 13.

Lemma 12. *We have, for any $\epsilon > 0$,*

$$\lim_{n \rightarrow \infty} \sum_{i=1}^n \mathbb{E} \left(\frac{Y_{i,n}^2}{s_n^2} \mathbb{1}_{\frac{|Y_{i,n}|}{s_n} \geq \epsilon} \right) = 0.$$

Conditions (T_1) and (T_2) are then satisfied.

Proof of Lemma 12. We have, for any $\delta > 0$ and $\epsilon > 0$,

$$\sum_{i=1}^n \mathbb{E} \left(\frac{Y_{i,n}^2}{s_n^2} \mathbb{1}_{\frac{|Y_{i,n}|}{s_n} \geq \epsilon} \right) \leq \frac{1}{\epsilon^\delta s_n^{2+\delta}} \sum_{i=1}^n \mathbb{E} (|Y_{i,n}|^{2+\delta}) \leq \frac{n}{\epsilon^\delta s_n^{2+\delta}} \max_{1 \leq i \leq n} \mathbb{E} (|Y_{i,n}|^{2+\delta}).$$

Now we use the inequality $(a+b)^{2+\delta} \leq 2^{1+\delta}(a^{2+\delta} + b^{2+\delta})$ for any $a, b \geq 0$ together with Hölder's inequality with conjugate integers $p, q \geq 2$ for which $\mathbb{E} (|\epsilon_1|^{(2+\delta)q}) < \infty$. We get,

$$\begin{aligned} \mathbb{E}(|Y_{i,n}|^{2+\delta}) &\leq 2^{1+\delta} a_{i,n}^{2+\delta}(h_n) \mathbb{E}(|\epsilon_i|^{2+\delta}) + 2^{1+\delta} \mathbb{E} \left(|\epsilon_i|^{2+\delta} \left| \sum_{j=1}^{i-1} \tilde{b}_{i,j} \epsilon_j \right|^{2+\delta} \right) \\ &\leq 2^{1+\delta} a_{i,n}^{2+\delta}(h_n) \mathbb{E}(|\epsilon_i|^{2+\delta}) + 2^{1+\delta} [\mathbb{E} (|\epsilon_i|^{(2+\delta)q})]^{1/q} \left[\mathbb{E} \left(\left| \sum_{j=1}^{i-1} \tilde{b}_{i,j} \epsilon_j \right|^{(2+\delta)p} \right) \right]^{1/p}. \end{aligned}$$

Corollary 1 of the Appendix C (recall that $|\tilde{b}_{i,j}| \leq cst \frac{1}{n^2 h_n^2} \mathbb{1}_{|i-j| \leq nh_n}$) gives

$$\left[\mathbb{E} \left(\left| \sum_{j=1}^{i-1} \tilde{b}_{i,j} \epsilon_j \right|^{(2+\delta)p} \right) \right]^{1/p} \leq cst \left(\sum_{j=1}^{i-1} \tilde{b}_{i,j}^2 \right)^{(2+\delta)/2} \leq cst \left(\frac{nh_n}{n^4 h_n^4} \right)^{(2+\delta)/2}.$$

Hence,

$$\mathbb{E}(|Y_{i,n}|^{2+\delta}) \leq cst \left(\frac{h_n}{n}\right)^{2+\delta} + cst \left(\frac{nh_n}{n^4 h_n^4}\right)^{(2+\delta)/2}$$

and,

$$\sum_{i=1}^n \mathbb{E} \left(\frac{Y_{i,n}^2}{s_n^2} \mathbb{1}_{\frac{|Y_{i,n}|}{s_n} \geq \epsilon} \right) \leq cst \frac{n}{s_n^{2+\delta}} \left(\frac{h_n}{n}\right)^{2+\delta} + cst \frac{n}{s_n^{2+\delta}} \left(\frac{nh_n}{n^4 h_n^4}\right)^{(2+\delta)/2}$$

Finally, (recall that, by Proposition 6 $\lim_{n \rightarrow \infty} s_n^2 \frac{n}{h_n^2} = cst$ and that $nh_n^5 = c$)

$$\sum_{i=1}^n \mathbb{E} \left(\frac{Y_{i,n}^2}{s_n^2} \mathbb{1}_{\frac{|Y_{i,n}|}{s_n} \geq \epsilon} \right) = O\left(\frac{1}{n^{\delta/2}}\right).$$

Hence $\lim_{n \rightarrow \infty} \sum_{i=1}^n \mathbb{E} \left(\frac{Y_{i,n}^2}{s_n^2} \mathbb{1}_{\frac{|Y_{i,n}|}{s_n} \geq \epsilon} \right) = 0$. This ensures, using some elementary calculations, that (T_1) and (T_2) hold.

Lemma 13. *Let $(\epsilon_i)_{i \geq 0}$ be a stationary sequence of martingale-difference relative to the filtration \mathcal{F}_i . Suppose that all the requirements of Proposition 7 are satisfied. Recall that $h_n = cn^{-1/5}$. Then $\frac{1}{s_n^2} \sum_{i=1}^n Y_{i,n}^2$ converges in probability to 1 as n tends to infinity.*

Proof of Lemma 13. Define $V_n^2 = \sum_{i=1}^n Y_{i,n}^2$. Our purpose is to prove that $\frac{1}{s_n^2} V_n^2$ converges in probability to 1 as n tends to infinity. Thanks to Markov's inequality, it suffices to prove that

$$\lim_{n \rightarrow \infty} \frac{1}{s_n^4} \mathbb{E} \left[(V_n^2 - s_n^2)^2 \right] = 0.$$

Let $A_n = \sum_{i=1}^n a_{i,n}^2 (h_n) \sigma^2 + \sum_{i=1}^n \sum_{j=1}^{i-1} \tilde{b}_{i,j}^2 \sigma^4$. We have, noting that by Proposition 6, (40) and (46), $s_n^2 = A_n + o\left(\frac{1}{n^2 h_n^3} + \frac{h_n^2}{n}\right) = A_n + o(n^{-7/5})$, so that

$$(V_n^2 - s_n^2)^2 \leq 2(V_n^2 - A_n)^2 + 2(A_n - s_n^2)^2 \leq 2(V_n^2 - A_n)^2 + o(n^{-14/5}).$$

The last bound together with the fact that $s_n^2 \sim cst \left(\frac{1}{n^2 h_n^3} + \frac{h_n^2}{n}\right) \sim cst(n^{-7/5})$ prove that

$$\frac{1}{s_n^4} \mathbb{E} \left[(V_n^2 - s_n^2)^2 \right] \leq \frac{A_n^2}{s_n^4} \frac{\mathbb{E} \left[(V_n^2 - A_n)^2 \right]}{A_n^2} + o(1).$$

From this and the fact that $\frac{A_n^2}{s_n^4} = 1 + o(1)$, we deduce that, in order to prove Lemma 13, it suffices to prove that

$$\lim_{n \rightarrow \infty} \frac{\mathbb{E} \left[(V_n^2 - A_n)^2 \right]}{A_n^2} = 0. \tag{54}$$

Indeed, we have,

$$\begin{aligned}
V_n^2 &= \sum_{i=1}^n Y_{i,n}^2 = \sum_{i=1}^n \left(a_{i,n}(h_n) + \sum_{j=1}^{i-1} \tilde{b}_{i,j} \epsilon_j \right)^2 \epsilon_i^2 \\
&= \sum_{i=1}^n a_{i,n}^2(h_n) \epsilon_i^2 + \sum_{i=1}^n \left(\sum_{j=1}^{i-1} \tilde{b}_{i,j} \epsilon_j \right)^2 \epsilon_i^2 + 2 \sum_{i=1}^n a_{i,n}(h_n) \sum_{j=1}^{i-1} \tilde{b}_{i,j} \epsilon_j \epsilon_i^2 \\
&= \sum_{i=1}^n a_{i,n}^2(h_n) \epsilon_i^2 + \sum_{i=1}^n \sum_{j=1}^{i-1} \tilde{b}_{i,j}^2 \epsilon_j^2 \epsilon_i^2 + 2 \sum_{i=1}^n \sum_{1 \leq j_1 < j_2 \leq i-1} \tilde{b}_{i,j_1} \tilde{b}_{i,j_2} \epsilon_{j_1} \epsilon_{j_2} \epsilon_i^2 \\
&\quad + 2 \sum_{i=1}^n a_{i,n}(h_n) \sum_{j=1}^{i-1} \tilde{b}_{i,j} \epsilon_j \epsilon_i^2,
\end{aligned}$$

and

$$\begin{aligned}
V_n^2 - A_n &= V_n^2 - \sum_{i=1}^n a_{i,n}^2(h_n) \sigma^2 - \sum_{i=1}^n \sum_{j=1}^{i-1} \tilde{b}_{i,j}^2 \sigma^4 \\
&= \sum_{i=1}^n a_{i,n}^2(h_n) (\epsilon_i^2 - \sigma^2) + \sum_{i=1}^n \sum_{j=1}^{i-1} \tilde{b}_{i,j}^2 (\epsilon_j^2 \epsilon_i^2 - \sigma^4) \\
&\quad + 2 \sum_{i=1}^n \sum_{1 \leq j_1 < j_2 \leq i-1} \tilde{b}_{i,j_1} \tilde{b}_{i,j_2} \epsilon_{j_1} \epsilon_{j_2} \epsilon_i^2 + 2 \sum_{i=1}^n a_{i,n}(h_n) \sum_{j=1}^{i-1} \tilde{b}_{i,j} \epsilon_j \epsilon_i^2.
\end{aligned}$$

Hence,

$$\begin{aligned}
&\mathbb{E} \left[(V_n^2 - A_n)^2 \right] \\
&\leq 4\mathbb{E} \left[\left(\sum_{i=1}^n a_{i,n}^2(h_n) (\epsilon_i^2 - \sigma^2) \right)^2 \right] + 4\mathbb{E} \left[\left(\sum_{i=1}^n \sum_{j=1}^{i-1} \tilde{b}_{i,j}^2 (\epsilon_j^2 \epsilon_i^2 - \sigma^4) \right)^2 \right] \\
&\quad + 16\mathbb{E} \left[\left(\sum_{i=1}^n a_{i,n}(h_n) \sum_{j=1}^{i-1} \tilde{b}_{i,j} \epsilon_j \epsilon_i^2 \right)^2 \right] + 16\mathbb{E} \left[\left(\sum_{i=1}^n \sum_{1 \leq j_1 < j_2 \leq i-1} \tilde{b}_{i,j_1} \tilde{b}_{i,j_2} \epsilon_{j_1} \epsilon_{j_2} \epsilon_i^2 \right)^2 \right] \\
&= I + II + III + IV. \tag{55}
\end{aligned}$$

Control of I.

$$\begin{aligned}
& \mathbb{E} \left[\left(\sum_{i=1}^n a_{i,n}^2(h_n) (\epsilon_i^2 - \sigma^2) \right)^2 \right] = \sum_{i=1}^n a_{i,n}^4(h_n) \mathbb{E} \left[(\epsilon_i^2 - \sigma^2)^2 \right] \\
& + 2 \sum_{1 \leq i < j \leq n} a_{i,n}^2(h_n) a_{j,n}^2(h_n) \mathbb{E} \left[(\epsilon_i^2 - \sigma^2) (\epsilon_j^2 - \sigma^2) \right] \\
& \leq cst \frac{h_n^4}{n^4} n \text{Var}(\epsilon_1^2) + cst \frac{h_n^4}{n^4} n \sup_i \sum_{j=1}^{\infty} |\text{Cov}(\epsilon_i^2, \epsilon_j^2)| \\
& = O\left(\frac{h_n^4}{n^3}\right).
\end{aligned}$$

Consequently,

$$\frac{1}{A_n^2} \mathbb{E} \left[\left(\sum_{i=1}^n a_{i,n}^2(h_n) (\epsilon_i^2 - \sigma^2) \right)^2 \right] = O(n^{-1}). \quad (56)$$

Control of II.

$$\begin{aligned}
& \mathbb{E} \left[\left(\sum_{i=1}^n \sum_{j=1}^{i-1} \tilde{b}_{i,j}^2 (\epsilon_j^2 \epsilon_i^2 - \sigma^4) \right)^2 \right] \\
& = \sum_{i_1=1}^n \sum_{j_1=1}^{i_1-1} \sum_{i_2=2}^n \sum_{j_2=1}^{i_2-1} \tilde{b}_{i_1,j_1}^2 \tilde{b}_{i_2,j_2}^2 \mathbb{E} \left[(\epsilon_{j_1}^2 \epsilon_{i_1}^2 - \sigma^4) (\epsilon_{j_2}^2 \epsilon_{i_2}^2 - \sigma^4) \right] \\
& \leq \frac{1}{n^8 h_n^8} \sum_{i_1=1}^n \sum_{j_1=1}^{i_1-1} \sum_{i_2=1}^n \sum_{j_2=1}^{i_2-1} \mathbb{1}_{|i_1-j_1| \leq nh_n} \mathbb{1}_{|i_2-j_2| \leq nh_n} |\mathbb{E} \left[(\epsilon_{j_1}^2 \epsilon_{i_1}^2 - \sigma^4) (\epsilon_{j_2}^2 \epsilon_{i_2}^2 - \sigma^4) \right]|.
\end{aligned}$$

We have to prove that

$$\begin{aligned}
& \sum_{i_1=1}^n \sum_{j_1=1}^{i_1-1} \sum_{i_2=2}^n \sum_{j_2=1}^{i_2-1} \mathbb{1}_{|i_1-j_1| \leq nh_n} \mathbb{1}_{|i_2-j_2| \leq nh_n} |\mathbb{E} \left[(\epsilon_{j_1}^2 \epsilon_{i_1}^2 - \sigma^4) (\epsilon_{j_2}^2 \epsilon_{i_2}^2 - \sigma^4) \right]| \\
& = o(n^6 h_n^{12}). \quad (57)
\end{aligned}$$

We evaluate now the sums over $i_1 \neq i_2, j_1 \neq j_2$ (if $(i_1, j_1) = (i_2, j_2)$ or $i_1 = i_2, j_1 \neq j_2$ or $i_1 \neq i_2, j_1 = j_2$ then it is easy to check that the order $o(n^6 h_n^{12})$ is obtained). We suppose also without loss of generality, that $j_1 < i_1 < j_2 < i_2$. Write,

$$\begin{aligned}
& \mathbb{E} \left[(\epsilon_{j_1}^2 \epsilon_{i_1}^2 - \sigma^4) (\epsilon_{j_2}^2 \epsilon_{i_2}^2 - \sigma^4) \right] \\
& = \mathbb{E}(\epsilon_{j_1}^2 \epsilon_{i_1}^2 \epsilon_{j_2}^2 \epsilon_{i_2}^2) - \sigma^4 (\text{Cov}(\epsilon_{j_2}^2, \epsilon_{i_2}^2) + \sigma^4) - \sigma^4 (\text{Cov}(\epsilon_{j_1}^2, \epsilon_{i_1}^2) + \sigma^4) + \sigma^8 \\
& = \mathbb{E}(\epsilon_{j_1}^2 \epsilon_{i_1}^2 \epsilon_{j_2}^2 \epsilon_{i_2}^2) - \sigma^8 - \sigma^4 \text{Cov}(\epsilon_{j_2}^2, \epsilon_{i_2}^2) - \sigma^4 \text{Cov}(\epsilon_{j_1}^2, \epsilon_{i_1}^2). \quad (58)
\end{aligned}$$

We have,

$$\begin{aligned}
& \sum_{i_1=1}^n \sum_{j_1=1}^{i_1-1} \sum_{i_2=1}^n \sum_{j_2=1}^{i_2-1} \mathbb{1}_{|i_1-j_1| \leq nh_n} \mathbb{1}_{|i_2-j_2| \leq nh_n} |\text{Cov}(\epsilon_{j_2}^2, \epsilon_{i_2}^2)| \\
& \leq \sum_{i_1=1}^n \sum_{j_1=1}^{i_1-1} \sum_{i_2=1}^n \sum_{j_2=1}^{i_2-1} \mathbb{1}_{|i_1-j_1| \leq nh_n} \mathbb{1}_{|i_2-j_2| \leq nh_n} \Phi(i_2 - j_2) \leq n^3 h_n \sum_{r=1}^{\infty} \Phi(r) = o(n^6 h_n^{12}).
\end{aligned} \tag{59}$$

Now, since

$$\begin{aligned}
& \mathbb{E}(\epsilon_{j_1}^2 \epsilon_{i_1}^2 \epsilon_{j_2}^2 \epsilon_{i_2}^2) - \sigma^8 \\
& = \text{Cov}(\epsilon_{j_1}^2, \epsilon_{i_1}^2 \epsilon_{j_2}^2 \epsilon_{i_2}^2) + \sigma^2 \mathbb{E}(\epsilon_{i_1}^2 \epsilon_{j_2}^2 \epsilon_{i_2}^2) - \sigma^8 \\
& = \text{Cov}(\epsilon_{j_1}^2, \epsilon_{i_1}^2 \epsilon_{j_2}^2 \epsilon_{i_2}^2) + \sigma^2 (\text{Cov}(\epsilon_{i_1}^2, \epsilon_{j_2}^2 \epsilon_{i_2}^2) + \sigma^2 \mathbb{E}(\epsilon_{j_2}^2 \epsilon_{i_2}^2)) - \sigma^8 \\
& = \text{Cov}(\epsilon_{j_1}^2, \epsilon_{i_1}^2 \epsilon_{j_2}^2 \epsilon_{i_2}^2) + \sigma^2 \text{Cov}(\epsilon_{i_1}^2, \epsilon_{j_2}^2 \epsilon_{i_2}^2) + \sigma^4 \text{Cov}(\epsilon_{j_2}^2, \epsilon_{i_2}^2),
\end{aligned}$$

we obtain,

$$|\mathbb{E}(\epsilon_{j_1}^2 \epsilon_{i_1}^2 \epsilon_{j_2}^2 \epsilon_{i_2}^2) - \sigma^8| \leq \Phi(i_1 - j_1) + \sigma^2 \Phi(j_2 - i_1) + \sigma^4 \Phi(i_2 - j_2).$$

Consequently,

$$\begin{aligned}
& \sum_{i_1=2}^n \sum_{j_1=1}^{i_1-1} \sum_{j_2: i_1 < j_2}^n \sum_{i_2: j_2 < i_2} \mathbb{1}_{|i_1-j_1| \leq nh_n} \mathbb{1}_{|i_2-j_2| \leq nh_n} |\mathbb{E}(\epsilon_{j_1}^2 \epsilon_{i_1}^2 \epsilon_{j_2}^2 \epsilon_{i_2}^2) - \sigma^8| \\
& \leq \sum_{i_1=2}^n \sum_{j_1=1}^{i_1-1} \sum_{i_1 < j_2}^n \sum_{j_2 < i_2} \mathbb{1}_{|i_1-j_1| \leq nh_n} \mathbb{1}_{|i_2-j_2| \leq nh_n} (\Phi(i_1 - j_1) + \sigma^2 \Phi(j_2 - i_1) + \sigma^4 \Phi(i_2 - j_2)) \\
& \leq 3n^3 h \sum_{r=1}^{\infty} \Phi(r) = o(n^6 h_n^{12}).
\end{aligned} \tag{60}$$

Hence (57) is proved by collecting (58), (59) and (60). Consequently,

$$\frac{1}{A_n^2} \mathbb{E} \left[\left(\sum_{i=1}^n \sum_{j=1}^{i-1} \tilde{b}_{i,j}^2 (\epsilon_j^2 \epsilon_i^2 - \sigma^4) \right)^2 \right] = o(1). \tag{61}$$

Control of III. We have,

$$\begin{aligned}
& \mathbb{E} \left[\left(\sum_{i=1}^n a_{i,n}(h_n) \sum_{j=1}^{i-1} \tilde{b}_{i,j} \epsilon_j \epsilon_i^2 \right)^2 \right] \\
& = \sum_{i_1=1}^n \sum_{j_1=1}^{i_1-1} \sum_{i_2=1}^n \sum_{j_2=1}^{i_2-1} a_{i_1,n}(h_n) a_{i_2,n}(h_n) \tilde{b}_{i_1,j_1} \tilde{b}_{i_2,j_2} \mathbb{E}(\epsilon_{j_1} \epsilon_{i_1}^2 \epsilon_{j_2} \epsilon_{i_2}^2) \\
& \leq cst \frac{1}{n^6 h_n^2} \sum_{i_1=1}^n \sum_{j_1=1}^{i_1-1} \sum_{i_2=1}^n \sum_{j_2=1}^{i_2-1} |\mathbb{E}(\epsilon_{j_1} \epsilon_{i_1}^2 \epsilon_{j_2} \epsilon_{i_2}^2)| \mathbb{1}_{|i_1-j_1| \leq nh_n} \mathbb{1}_{|i_2-j_2| \leq nh_n}.
\end{aligned}$$

It suffices to prove that

$$\sum_{i_1=2}^n \sum_{j_1=1}^{i_1-1} \sum_{i_2=2}^n \sum_{j_2=1}^{i_2-1} |\mathbb{E}(\epsilon_{j_1} \epsilon_{i_1}^2 \epsilon_{j_2} \epsilon_{i_2}^2)| \mathbb{1}_{|i_1-j_1| \leq nh_n} \mathbb{1}_{|i_2-j_2| \leq nh_n} = o\left(\frac{n^2}{h_n^4}\right), \quad (62)$$

in order to obtain

$$\frac{1}{A_n^2} \mathbb{E} \left[\left(\sum_{i=1}^n a_{i,n}(h_n) \sum_{j=1}^{i-1} \tilde{b}_{i,j} \epsilon_j \epsilon_i^2 \right)^2 \right] = o(1). \quad (63)$$

So we concentrate in proving (62). The sums when $(i_1, j_1) = (i_2, j_2)$, $i_1 = i_2, j_1 \neq j_2, i_1 \neq i_2, j_1 = j_2$ are respectively of order $n^2 h$, $n^3 h^2$ and $n^3 h^2$. So they are all of order $o\left(\frac{n^2}{h_n^4}\right)$ and we have only to consider the case where the sum in (62) is taken over $i_1 \neq i_2$ and $j_1 \neq j_2$.

We only suppose that $j_1 < j_2 < i_2 < i_1$ (the other cases are similar). We write

$$\mathbb{E}(\epsilon_{j_1} \epsilon_{i_1}^2 \epsilon_{j_2} \epsilon_{i_2}^2) = \mathbb{E}(\epsilon_{j_1} \epsilon_{j_2} \epsilon_{i_2}^2 \epsilon_{i_1}^2).$$

We have to discuss the subcases $j_2 - j_1 \geq \max(i_1 - i_2, i_2 - j_2)$, $i_1 - i_2 \geq \max(j_2 - j_1, i_2 - j_2)$ or $i_2 - j_2 \geq \max(i_1 - i_2, j_2 - j_1)$.

- if $j_2 - j_1 \geq \max(i_1 - i_2, i_2 - j_2)$ then we write

$$|\mathbb{E}(\epsilon_{j_1} \epsilon_{j_2} \epsilon_{i_2}^2 \epsilon_{i_1}^2)| = |\text{Cov}(\epsilon_{j_1}, \epsilon_{j_2} \epsilon_{i_2}^2 \epsilon_{i_1}^2)| \leq \Phi(j_2 - j_1).$$

Hence,

$$\begin{aligned} & \sum_{j_1 < j_2 < i_2 < i_1} |\mathbb{E}(\epsilon_{j_1} \epsilon_{i_1}^2 \epsilon_{j_2} \epsilon_{i_2}^2)| \mathbb{1}_{|i_1-j_1| \leq nh_n} \mathbb{1}_{|i_2-j_2| \leq nh_n} \mathbb{1}_{j_2-j_1 \geq \max(i_1-i_2, i_2-j_2)} \\ & \leq \sum_{j_1 < j_2 < i_2 < i_1} \mathbb{1}_{|i_1-j_1| \leq nh_n} \mathbb{1}_{|i_2-j_2| \leq nh_n} \mathbb{1}_{j_2-j_1 \geq \max(i_1-i_2, i_2-j_2)} \Phi(j_2 - j_1) \\ & \leq \sum_{j_1=1}^n \sum_{j_2=j_1+1}^n (j_2 - j_1)^2 \Phi(j_2 - j_1) \leq n \sum_{r=1}^{\infty} r^2 \Phi(r) = o\left(\frac{n^2}{h_n^4}\right). \end{aligned}$$

- if $i_1 - i_2 \geq \max(j_2 - j_1, i_2 - j_2)$ and $j_2 - j_1 \geq i_2 - j_2$ (the case $i_2 - j_2 \geq j_2 - j_1$ is similar noting that $\mathbb{E}(\epsilon_{j_1} \epsilon_{j_2}) = 0$) then we write

$$\begin{aligned} & |\mathbb{E}(\epsilon_{j_1} \epsilon_{j_2} \epsilon_{i_2}^2 \epsilon_{i_1}^2)| \leq |\text{Cov}(\epsilon_{j_1} \epsilon_{j_2} \epsilon_{i_2}^2, \epsilon_{i_1}^2)| + \sigma^2 |\text{Cov}(\epsilon_{j_1}, \epsilon_{j_2} \epsilon_{i_2}^2)| \\ & \leq \Phi(i_1 - i_2) + \sigma^2 \Phi(j_2 - j_1) \leq cst \Phi(j_2 - j_1). \end{aligned}$$

Hence,

$$\begin{aligned} & \sum_{j_1 < j_2 < i_2 < i_1} |\mathbb{E}(\epsilon_{j_1} \epsilon_{i_1}^2 \epsilon_{j_2} \epsilon_{i_2}^2)| \mathbb{1}_{|i_1-j_1| \leq nh_n} \mathbb{1}_{|i_2-j_2| \leq nh_n} \mathbb{1}_{i_1-i_2 \geq \max(j_2-j_1, i_2-j_2)} \mathbb{1}_{j_2-j_1 \geq i_2-j_2} \\ & \leq cst \sum_{j_1 < j_2 < i_2 < i_1} \Phi(j_2 - j_1) \mathbb{1}_{|i_1-j_1| \leq nh_n} \mathbb{1}_{|i_2-j_2| \leq nh_n} \mathbb{1}_{i_1-i_2 \geq \max(j_2-j_1, i_2-j_2)} \mathbb{1}_{j_2-j_1 \geq i_2-j_2} \\ & \leq cst n^2 h \sup_{j_1} \sum_{j_2 > j_1} (j_2 - j_1) \Phi(j_2 - j_1) \leq cst n^2 h \sum_{r=1}^{\infty} r \Phi(r) = o\left(\frac{n^2}{h_n^4}\right). \end{aligned}$$

- if $i_2 - j_2 \geq \max(i_1 - i_2, j_2 - j_1)$ then we write (since $\mathbb{E}(\epsilon_{j_1} \epsilon_{j_2}) = 0$)

$$|\mathbb{E}(\epsilon_{j_1} \epsilon_{j_2} \epsilon_{i_2}^2 \epsilon_{i_1}^2)| = |\text{Cov}(\epsilon_{j_1} \epsilon_{j_2}, \epsilon_{i_2}^2 \epsilon_{i_1}^2)| \leq \Phi(i_2 - j_2),$$

and

$$\begin{aligned} & \sum_{j_1 < j_2 < i_2 < i_1} |\mathbb{E}(\epsilon_{j_1} \epsilon_{i_1}^2 \epsilon_{j_2} \epsilon_{i_2}^2)| \mathbb{1}_{|i_1 - j_1| \leq nh_n} \mathbb{1}_{|i_2 - j_2| \leq nh_n} \mathbb{1}_{i_2 - j_2 \geq \max(i_1 - i_2, j_2 - j_1)} \\ & \leq \sum_{j_1 < j_2 < i_2 < i_1} \Phi(i_2 - j_2) \mathbb{1}_{|i_1 - j_1| \leq nh_n} \mathbb{1}_{i_2 - j_2 \geq \max(i_1 - i_2, j_2 - j_1)} \\ & \leq n^2 h \sum_{r=1}^{\infty} r \Phi(r) = o\left(\frac{n^2}{h_n^4}\right). \end{aligned}$$

Collecting all the previous bounds, the proof of (62) is achieved.

Control of IV.

$$\begin{aligned} IV &= \mathbb{E} \left[\left(\sum_{i=1}^n \sum_{1 \leq j_1 < j_2 \leq i-1} \tilde{b}_{i,j_1} \tilde{b}_{i,j_2} \epsilon_{j_1} \epsilon_{j_2} \epsilon_i^2 \right)^2 \right] \\ &= \sum_{i=1}^n \sum_{1 \leq j_1 < j_2 \leq i-1} \sum_{l=2}^n \sum_{1 \leq l_1 < l_2 \leq l-1} \tilde{b}_{i,j_1} \tilde{b}_{i,j_2} \tilde{b}_{l,l_1} \tilde{b}_{l,l_2} \mathbb{E}(\epsilon_{j_1} \epsilon_{j_2} \epsilon_i^2 \epsilon_{l_1} \epsilon_{l_2} \epsilon_l^2). \end{aligned}$$

We have to prove (letting $\mathbb{1}_{i,j_1,j_2,l,l_1,l_2} := \mathbb{1}_{|i-j_1| \leq nh_n} \mathbb{1}_{|i-j_2| \leq nh_n} \mathbb{1}_{|l-l_1| \leq nh_n} \mathbb{1}_{|l-l_2| \leq nh_n}$)

$$\begin{aligned} & \sum_{i=1}^n \sum_{1 \leq j_1 < j_2 \leq i-1} \sum_{l=2}^n \sum_{1 \leq l_1 < l_2 \leq l-1} \mathbb{1}_{i,j_1,j_2,l,l_1,l_2} \mathbb{E}(\epsilon_{j_1} \epsilon_{j_2} \epsilon_i^2 \epsilon_{l_1} \epsilon_{l_2} \epsilon_l^2) \\ &= o(n^4 h_n^2), \end{aligned} \tag{64}$$

so that

$$\frac{1}{A_n^2} \mathbb{E} \left[\left(\sum_{i=1}^n \sum_{1 \leq j_1 < j_2 \leq i-1} \tilde{b}_{i,j_1} \tilde{b}_{i,j_2} \epsilon_{j_1} \epsilon_{j_2} \epsilon_i^2 \right)^2 \right] = o(1). \tag{65}$$

To prove (64), we write, if all the index i, l, j_1, j_2, l_1, l_2 are different, (the other cases are similar)

$$\begin{aligned} & \mathbb{E}(\epsilon_{j_1} \epsilon_{j_2} \epsilon_i^2 \epsilon_{l_1} \epsilon_{l_2} \epsilon_l^2) \\ &= \mathbb{E}(\epsilon_{j_1} \epsilon_{j_2} (\epsilon_i^2 - \sigma^2) \epsilon_{l_1} \epsilon_{l_2} (\epsilon_l^2 - \sigma^2)) + \sigma^2 \mathbb{E}(\epsilon_{j_1} \epsilon_{j_2} (\epsilon_i^2 - \sigma^2) \epsilon_{l_1} \epsilon_{l_2}) \\ &+ \sigma^2 \mathbb{E}(\epsilon_{j_1} \epsilon_{j_2} \epsilon_{l_1} \epsilon_{l_2} (\epsilon_l^2 - \sigma^2)) + \sigma^4 \mathbb{E}(\epsilon_{j_1} \epsilon_{j_2} \epsilon_{l_1} \epsilon_{l_2}), \end{aligned}$$

so that $\mathbb{E}(\epsilon_{j_1} \epsilon_{j_2} \epsilon_i^2 \epsilon_{l_1} \epsilon_{l_2} \epsilon_l^2)$ is written as a sum of expectations of products of centered random variables, i.e, a sum of terms of the form $\mathbb{E}(Y_{t_1} \times \cdots \times Y_{t_q})$ where $q \in \{4, 5, 6\}$, $\{t_1, \dots, t_q\} = \{i, l, j_1, j_2, l_1, l_2\}$, $t_1 < \cdots < t_q$ and $(Y_{t_i})_i$ is a centered sequence of random variables. In order

to evaluate $\mathbb{E}(Y_{t_1} \times \cdots \times Y_{t_q})$, we use the classical techniques (used to get moment inequalities for sums of even orders for example in [Rio, E. \(2017\)](#), see also [Doukhan, P. and Louhichi, S. \(1999\)](#)). Letting k be such that $t_{k+1} - t_k = \max_i(t_{i+1} - t_i)$,

$$\begin{aligned} & |\mathbb{E}(Y_{t_1} \times \cdots \times Y_{t_q})| \\ & \leq |\text{Cov}(Y_{t_1} \times \cdots \times Y_{t_k}, Y_{t_{k+1}} \cdots \times Y_{t_q})| + |\mathbb{E}(Y_{t_1} \times \cdots \times Y_{t_k})| |\mathbb{E}(Y_{t_{k+1}} \cdots \times Y_{t_q})| \\ & \leq \Phi(t_{k+1} - t_k) + |\mathbb{E}(Y_{t_1} \times \cdots \times Y_{t_k})| |\mathbb{E}(Y_{t_{k+1}} \cdots \times Y_{t_q})| \\ & \sum_{1 \leq t_1 < \cdots < t_q \leq m} \Phi(t_{k+1} - t_k) \leq m \sum_{r=1}^{\infty} r^{q-2} \Phi(r). \end{aligned}$$

Hence

$$\begin{aligned} & \sum_{1 \leq t_1 < \cdots < t_q \leq m} |\mathbb{E}(Y_{t_1} \times \cdots \times Y_{t_q})| \leq m \sum_{r=1}^{\infty} r^{q-2} \Phi(r) \\ & + \sum_{t_k=1}^m \sum_{(t_1, \dots, t_{k-1})} \sum_{t_1 < \cdots < t_{k-1} < t_k} |\mathbb{E}(Y_{t_1} \times \cdots \times Y_{t_k})| \sum_{t_{k+1} < \cdots < t_q} |\mathbb{E}(Y_{t_{k+1}} \cdots \times Y_{t_q})| \end{aligned}$$

and we can control $|\mathbb{E}(Y_{t_1} \times \cdots \times Y_{t_k})|$ and $|\mathbb{E}(Y_{t_{k+1}} \cdots \times Y_{t_q})|$ as it was done for $|\mathbb{E}(Y_{t_1} \times \cdots \times Y_{t_q})|$.

We prove then (54) by collecting (55) together with (56), (61), (65) and (63). With this, the proof of Lemma 13 is complete.

C Tools for martingale difference sequences

We recall the following Marcinkiewicz-Zygmund type inequality which is a simple consequence of the Minkowski and the Burkholder inequalities (see [Burkholder, D. L. \(1988\)](#)).

Theorem 3. *Let $(\eta_i)_{i \geq 0}$ be a stationary sequence of martingale difference of finite p th moment with $p \geq 2$. Then there exists a positive constant c_p such that for any positive integer n ,*

$$\left\| \sum_{i=1}^n \eta_i \right\|_p^2 \leq c_p \sum_{i=1}^n \|\eta_i\|_p^2.$$

An immediate consequence of Theorem 3 is the following corollary.

Corollary 1. *Let $(\eta_i)_{i \geq 0}$ be a stationary sequence of martingale difference of finite p th moment with $p \geq 2$. Then there exists a positive constant c_p such that for any positive integer n and for any sequence of real numbers $(d_{i,n})_{1 \leq i \leq n}$*

$$\left\| \sum_{i=1}^n d_{i,n} \eta_i \right\|_p^2 \leq c_p \sum_{i=1}^n d_{i,n}^2.$$

We also need the following propositions for which their proofs use Theorem 3 above.

Proposition 8. *Let $(\eta_i)_{i \geq 0}$ be a stationary sequence of martingale difference such that $\|\eta_i\|_{2p} < \infty$ for some $p \geq 2$. Then, there exists a positive constant c_p such that for any positive integer n , and for any sequence of real numbers $(b_{i,j,n})_{1 \leq i,j \leq n}$,*

$$\left\| \sum_{i=1}^n \sum_{j=1}^{i-1} b_{i,j,n} \eta_j \eta_i \right\|_p^2 \leq c_p \sum_{i=1}^n \sum_{j=1}^{i-1} b_{i,j,n}^2,$$

Proof of Proposition 8. Let $X_i = \sum_{j=1}^{i-1} b_{i,j,n} \eta_j \eta_i$. The sequence $(X_i)_i$ is a martingale difference relative to the filtration $\mathcal{F}_i = \sigma(\eta_j, 1 \leq j \leq i)$. We apply Theorem 3 to this sequence $(X_i)_i$, we obtain

$$\begin{aligned} & \left\| \sum_{i=1}^n \sum_{j=1}^{i-1} b_{i,j,n} \eta_j \eta_i \right\|_p^2 = \left\| \sum_{i=1}^n X_i \right\|_p^2 \\ & \leq c_p \sum_{i=1}^n \|X_i\|_p^2 = c_p \sum_{i=1}^n \left\| \eta_i \sum_{j=1}^{i-1} b_{i,j,n} \eta_j \right\|_p^2 \\ & \leq c_p \sum_{i=1}^n \|\eta_i\|_{2p}^2 \left\| \sum_{j=1}^{i-1} b_{i,j,n} \eta_j \right\|_{2p}^2 \quad \text{by Hölder inequality} \\ & \leq c_p \sum_{i=1}^n \sum_{j=1}^{i-1} b_{i,j,n}^2 \quad \text{by Corollary 1.} \end{aligned}$$

The following maximal limits, for weighed sums of martingale difference or weighted sums of quadratic forms of martingale differences, are also very needed in the proofs. Their proofs need some chaining arguments (as used for instance in [Andrews, D. W. K. and Pollard, D. \(1994\)](#), [Louhichi, S. \(2000\)](#) or [Pollard, D. \(1990\)](#)).

Lemma 14. *Let $(\eta_i)_{i \geq 0}$ be a sequence of stationary martingale difference with $\|\eta_i\|_p < \infty$ for some $p \geq 2$. Let $(c_{i,n}(h))_{i,n,h}$ be a sequence of weights satisfying, for any $h, h' \in H_n = [an^{-1/5}, bn^{-1/5}]$,*

$$|c_{i,n}(h) - c_{i,n}(h')| \leq cst |h - h'|$$

and

$$\max_{i \leq n} \sup_{h \in H_n} |c_{i,n}(h)| \leq cst n^{-\alpha}, \quad \alpha > \frac{5p-2}{10(p-1)}.$$

Then,

$$\lim_{n \rightarrow \infty} \left\| \sup_{h \in H_n} \left| \sum_{i=1}^n c_{i,n}(h) \eta_i \right| \right\|_p = 0.$$

Proof of Lemma 14. Define, for $h \in H_n = [an^{-1/5}, bn^{-1/5}]$, $Z_n(h) = \sum_{i=1}^n c_{i,n}(h)\eta_i$. Clearly, for any fixed $k \in \mathbb{N}$ (to be chosen later)

$$H_n = [an^{-1/5}, bn^{-1/5}] = \cup_{p=0}^{2^k-1} H_{n,k,p},$$

where

$$H_{n,k,p} = [(a + p2^{-k}(b-a))n^{-1/5}, (a + (p+1)2^{-k}(b-a))n^{-1/5}],$$

so that $(H_{n,k,p})_p$ is a covering set of H_n , of 2^k cardinality, such that for any $h \in H_n$ there exists $p \in \{0, \dots, 2^k - 1\}$ for which, letting $h_{k,p} = (a + p2^{-k}(b-a))n^{-1/5}$, one has,

$$|h - h_{k,p}| \leq n^{-1/5}2^{-k}(b-a) =: \rho_{n,k}.$$

Define

$$I_k = \{(a + p2^{-k}(b-a))n^{-1/5}, p = 0, \dots, 2^k - 1\}.$$

Let $h \in H_n$ be fixed. We define inductively, for $m < k$ (to be chosen later), the sequence $h_k \in I_k, \dots, h_l \in I_l, \dots, h_m \in I_m$ in such a way that $|h - h_k| \leq n^{-1/5}2^{-k}(b-a)$,

$$|h_k - h_{k-1}| \leq n^{-1/5}2^{-(k-1)}(b-a), \dots, |h_l - h_{l-1}| \leq n^{-1/5}2^{-(l-1)}(b-a).$$

We have,

$$|Z_n(h)| \leq |Z_n(h) - Z_n(h_k)| + \sum_{l=m+1}^k |Z_n(h_l) - Z_n(h_{l-1})| + |Z_n(h_m)|.$$

Hence,

$$\begin{aligned} & \sup_{h \in H_n} |Z_n(h)| \\ & \leq \sup_{h: |h-h_k| \leq \rho_{n,k}} |Z_n(h) - Z_n(h_k)| + \sum_{l=m+1}^k \max_{h_l \in I_l} |Z_n(h_l) - Z_n(h_{l-1})| + \max_{h \in I_m} |Z_n(h)| \\ & \leq \sum_{i=1}^n \sup_{h: |h-h_k| \leq \rho_{n,k}} |c_{i,n}(h) - c_{i,n}(h_k)| |\eta_i| \\ & \quad + \sum_{l=m+1}^k \max_{h_l \in I_l} |Z_n(h_l) - Z_n(h_{l-1})| + \max_{h \in I_m} |Z_n(h)|, \end{aligned} \tag{66}$$

so that, taking the p -norm in the last inequality and using the requirements of Lemma 14 together with the inequality

$$\| \max_{i \leq N} |Y_i| \|_p \leq N^{1/p} \max_{i \leq N} \|Y_i\|_p, \tag{67}$$

we obtain

$$\begin{aligned}
& \left\| \sup_{h \in H_n} |Z_n(h)| \right\|_p \\
& \leq cst \rho_{n,k} \sum_{i=1}^n \|\eta_i\|_p + \sum_{l=m+1}^k \left\| \max_{h_l \in I_l} |Z_n(h_l) - Z_n(h_{l-1})| \right\|_p + \left\| \max_{h \in I_m} |Z_n(h)| \right\|_p \\
& \leq cst n \rho_{n,k} + \sum_{l=m+1}^k |I_l|^{1/p} \max_{h_l \in I_l} \|Z_n(h_l) - Z_n(h_{l-1})\|_p + |I_m|^{1/p} \max_{h \in I_m} \|Z_n(h)\|_p
\end{aligned}$$

The last bound, together with Corollary 1, give (recall that $\rho_{n,k} = n^{-1/5} 2^{-k}(b-a)$),

$$\begin{aligned}
& \left\| \sup_{h \in H_n} |Z_n(h)| \right\|_p \leq cst n \rho_{n,k} + cst \sum_{l=m+1}^k |I_l|^{1/p} \max_{h_l \in I_l} \left(\sum_{i=1}^n |c_{i,n}(h_l) - c_{i,n}(h_{l-1})|^2 \right)^{1/2} \\
& + cst |I_m|^{1/p} \max_{h \in I_m} \left(\sum_{i=1}^n |c_{i,n}(h)|^2 \right)^{1/2} \\
& \leq cst n \rho_{n,k} + cst \sum_{l=m+1}^k 2^{l/p} \sqrt{n} \rho_{n,l-1} + 2^{m/p} \sqrt{n} \max_{i \leq n} \sup_{h \in H_n} |c_{i,n}(h)| \\
& \leq cst n^{4/5} 2^{-k} + cst n^{3/10} \sum_{l=m+1}^{\infty} 2^{l/p} 2^{-l} + 2^{m/p} \sqrt{n} \max_{i \leq n} \sup_{h \in H_n} |c_{i,n}(h)| \\
& \leq cst n^{4/5} 2^{-k} + cst n^{3/10} 2^{-m(1-1/p)} + 2^{m/p} \sqrt{nn}^{-\alpha}.
\end{aligned}$$

Our task now is to choose m and k ($m < k$) in such a way that the last hand side of this last inequality tends to 0 as n tends to infinity. This choice is possible, by taking for instance, $\gamma > \beta > 0$, such that,

$$\begin{aligned}
2^m &= n^\beta, \text{ with } \frac{3}{10} \frac{p}{p-1} < \beta < \left(\alpha - \frac{1}{2}\right)p, \\
2^k &= n^\gamma, \text{ with, } \gamma > \max(4/5, \left(\alpha - \frac{1}{2}\right)p).
\end{aligned}$$

Lemma 15. *Let $(\epsilon_j)_j$ be a sequence of random variables with finite fourth moment and such that,*

$$\sup_i \sum_{j=1}^{\infty} |\text{Cov}(\epsilon_i^2, \epsilon_j^2)| < \infty.$$

Let for $h \in H_n = [an^{-1/5}, bn^{-1/5}]$, $(d_{j,n}(h))_{1 \leq j \leq n}$ be a sequence of real numbers satisfying for any $1 \leq j \leq n$,

$$|d_{j,n}(h)| \leq \frac{cst}{n}, \text{ and } |d_{j,n}(h) - d_{j,n}(h')| \leq cst n^{-2/5} |h - h'|.$$

Then,

$$\lim_{n \rightarrow \infty} \left\| \sup_{h \in H_n} \left| \sum_{i=1}^n d_{i,n}(h) (\epsilon_i^2 - \mathbb{E}(\epsilon_i^2)) \right| \right\|_2 = 0.$$

Proof of Lemma 15. Let

$$Z_n(h) = \sum_{i=1}^n d_{i,n}(h) (\epsilon_i^2 - \mathbb{E}(\epsilon_i^2)).$$

Hence, for any $h, h' \in H_n$,

$$Z_n(h) - Z_n(h') = \sum_{i=1}^n (d_{i,n}(h) - d_{i,n}(h')) (\epsilon_i^2 - \mathbb{E}(\epsilon_i^2))$$

We have,

$$\begin{aligned} \|Z_n(h)\|_2^2 &= \text{Var}(Z_n(h)) = \sum_{i=1}^n \sum_{j=1}^n d_{i,n}(h) d_{j,n}(h) \text{Cov}(\epsilon_i^2, \epsilon_j^2) \\ &\leq n \max_i d_{i,n}^2 \sup_i \sum_{j=1}^{\infty} |\text{Cov}(\epsilon_i^2, \epsilon_j^2)|. \end{aligned}$$

Let h_k be related to h as in the proof of Lemma 14. Then,

$$\sup_{h \in H_n} |Z_n(h)| \leq \sup_{h: |h-h_k| \leq \rho_{n,k}} |Z_n(h) - Z_n(h_k)| + \max_{h \in I_k} |Z_n(h)|.$$

We take the 2-norm in the last inequality and we use some similar calculations as in the proof of Lemma 14 and Inequality (67). We get for k to be chosen later,

$$\begin{aligned} \left\| \sup_{h \in H_n} |Z_n(h)| \right\|_2 &\leq \left\| \sup_{h: |h-h_k| \leq \rho_{n,k}} |Z_n(h) - Z_n(h_k)| \right\|_2 + \left\| \max_{h \in I_k} |Z_n(h)| \right\|_2 \\ &\leq cst n \max_j \sup_{h: |h-h_k| \leq \rho_{n,k}} |d_{j,n}(h) - d_{j,n}(h_k)| + cst 2^{k/2} \sqrt{n} \max_i |d_{i,n}| \\ &\leq cst \left(n^{3/5} \rho_{n,k} + \frac{2^{k/2}}{\sqrt{n}} \right) \leq cst \left(n^{2/5} 2^{-k} + \frac{2^{k/2}}{\sqrt{n}} \right), \end{aligned}$$

which tends to 0 as n tends to infinity if we choose k such that $2^k = n^\beta$ for $2/5 < \beta < 1$.

Lemma 16. Let $(\eta_i)_{i \geq 0}$ be a stationary sequence of martingale difference random variables with finite moment of order $2p$, for some $p > 8$. Suppose that, for any $h, h' \in H_n$

$$|b_{i,j,n}(h)| \leq \frac{cst}{n} \mathbb{1}_{|i-j| \leq 2nh}, \quad |b_{i,j,n}(h) - b_{i,j,n}(h')| \leq cst n^{-4/5} |h - h'| \mathbb{1}_{|i-j| \leq 2n \max(h, h')}.$$

Then,

$$\lim_{n \rightarrow \infty} \left\| \sup_{h \in H_n} \left| \sum_{i=1}^n \sum_{j=1}^{i-1} b_{i,j,n}(h) \eta_j \eta_i \right| \right\|_p = 0.$$

Proof of Lemma 16. Define $Z_n(h) = \sum_{i=1}^n \sum_{j=1}^{i-1} b_{i,j,n}(h) \eta_j \eta_i$. We have, using the requirements of Lemma 16 together with Proposition 8,

$$\begin{aligned} \|Z_n(h)\|_p &\leq cst \left(\sum_{i=1}^n \sum_{j=1}^{i-1} b_{i,j,n}^2(h) \right)^{1/2} \leq cst \sqrt{h}, \\ |Z_n(h) - Z_n(h')| &\leq cst \sum_{i=1}^n \sum_{j=1}^{i-1} |b_{i,j,n}(h) - b_{i,j,n}(h')| |\eta_i \eta_j|. \end{aligned}$$

We have, for $p > 8$, using the notations of the proof of Lemma 14,

$$\begin{aligned} &\| \sup_{h \in H_n} |Z_n(h)| \|_p \\ &\leq \| \sup_{h: |h-h_k| \leq \rho_{n,k}} |Z_n(h) - Z_n(h_k)| \|_p + \| \max_{h \in I_k} |Z_n(h)| \|_p \\ &\leq \| \sup_{h: |h-h_k| \leq \rho_{n,k}} |Z_n(h) - Z_n(h_k)| \|_p + |I_k|^{1/p} \max_{h \in I_k} \| |Z_n(h)| \|_p \\ &\leq \sum_{i=1}^n \sum_{j=1}^{i-1} \sup_{h: |h-h_k| \leq \rho_{n,k}} |b_{i,j,n}(h) - b_{i,j,n}(h_k)| |\eta_j \eta_i|_p + 2^{k/p} \max_{h \in I_k} \| |Z_n(h)| \|_p \\ &\leq cst 2^{-k} n^{4/5} + cst 2^{k/p} n^{-1/10}. \end{aligned}$$

Now, we choose k such that $2^k = n^\beta$ with $\frac{4}{5} < \beta < \frac{p}{10}$, so that

$$\lim_{n \rightarrow \infty} \| \sup_{h \in H_n} |Z_n(h)| \|_p = 0.$$

Lemma 17. Let $(\eta_i)_{i \geq 0}$ be a stationary sequence of martingale difference random variables with finite moment of order $2p$, for some $p \geq 1$. Suppose that, for any $h, h' \in H_n$

$$|b_{i,j,n}(h)| \leq \frac{cst}{n^{9/10}} \mathbb{1}_{|i-j| \leq 2nh}, \quad |b_{i,j,n}(h) - b_{i,j,n}(h')| \leq cst n^{-7/10} |h - h'| \mathbb{1}_{|i-j| \leq 2n \max(h, h')}.$$

Then,

$$\limsup_{n \rightarrow \infty} \left\| \sup_{h \in H_n} \left| \sum_{i=1}^n \sum_{j=1}^{i-1} b_{i,j,n}(h) \eta_j \eta_i \right| \right\|_p < \infty.$$

Proof of Lemma 17. Define $Z_n(h) = \sum_{i=1}^n \sum_{j=1}^{i-1} b_{i,j,n}(h) \eta_j \eta_i$. Let m be fixed, we have, using

(66), together with the notations of the proof of Lemma 14,

$$\begin{aligned}
& \left\| \sup_{h \in H_n} |Z_n(h)| \right\|_p \leq \left\| \sup_{h: |h-h_k| \leq \rho_{n,k}} |Z_n(h) - Z_n(h_k)| \right\|_p \\
& + \left\| \sum_{l=m+1}^k \max_{h_l \in I_l} |Z_n(h_l) - Z_n(h_{l-1})| \right\|_p + \left\| \max_{h \in I_m} |Z_n(h)| \right\|_p \\
& \leq \sum_{i=1}^n \sum_{j=1}^{i-1} \sup_{h: |h-h_k| \leq \rho_{n,k}} |b_{i,j,n}(h) - b_{i,j,n}(h_k)| \|\eta_i \eta_j\|_p \\
& + \left\| \sum_{l=m+1}^k \max_{h_l \in I_l} |Z_n(h_l) - Z_n(h_{l-1})| \right\|_p + \left\| \max_{h \in I_m} |Z_n(h)| \right\|_p.
\end{aligned}$$

Now,

$$\begin{aligned}
& \sum_{i=1}^n \sum_{j=1}^{i-1} \sup_{h: |h-h_k| \leq \rho_{n,k}} |b_{i,j,n}(h) - b_{i,j,n}(h_k)| \|\eta_i \eta_j\|_p \\
& \leq cst \sum_{i=1}^n \sum_{j=1}^{i-1} n^{-7/10} 2^{-k} n^{-1/5} \mathbb{1}_{|i-j| \leq 2n^{4/5}} \leq cst n^{1+4/5-7/10-1/5} 2^{-k} \leq cst n^{9/10} 2^{-k},
\end{aligned}$$

and

$$\begin{aligned}
& \left\| \sum_{l=m+1}^k \max_{h_l \in I_l} |Z_n(h_l) - Z_n(h_{l-1})| \right\|_p \leq \sum_{l=m+1}^k \left\| \max_{h_l \in I_l} |Z_n(h_l) - Z_n(h_{l-1})| \right\|_p \\
& \leq \sum_{l=m+1}^k 2^{l/p} \max_{h_l \in I_l} \| |Z_n(h_l) - Z_n(h_{l-1})| \|_p \\
& \leq \sum_{l=m+1}^k 2^{l/p} \max_{h_l \in I_l} \left(\sum_{i=1}^n \sum_{j=1}^{i-1} |b_{i,j,n}(h_l) - b_{i,j,n}(h_{l-1})|^2 \right)^{1/2} \\
& \leq \sum_{l=m+1}^k 2^{l/p} n^{-7/10} 2^{-l} n^{-1/5} \sqrt{n^2 n^{-1/5}} \leq \sum_{l=m+1}^k 2^{-l(1-1/p)} \leq \sum_{l=m+1}^{\infty} 2^{-l(1-1/p)} < \infty.
\end{aligned}$$

Finally,

$$\begin{aligned}
& \left\| \max_{h \in I_m} |Z_n(h)| \right\|_p \leq 2^{m/p} \max_{h \in I_m} \| |Z_n(h)| \|_p \leq 2^{m/p} \max_{h \in I_m} \left(\sum_{i=1}^n \sum_{j=1}^{i-1} |b_{i,j,n}(h)|^2 \right)^{1/2} \\
& \leq 2^{m/p} n^{-9/10} \sqrt{n^2 n^{-1/5}} \leq 2^{m/p}.
\end{aligned}$$

Hence,

$$\left\| \sup_{h \in H_n} |Z_n(h)| \right\|_p \leq cst n^{9/10} 2^{-k} + \sum_{l=m+1}^{\infty} 2^{-l(1-1/p)} + 2^{m/p}.$$

The right hand side of the last inequality is uniformly bounded over n if we choose a constant $m < k$ and $2^k = n^\beta$ with $9/10 < \beta$.