Electronic Journal of Statistics
Vol. 8 (2014) 1152–1187
ISSN: 1935-7524
DOI: 10.1214/14-EJS925

Global smoothness estimation of a Gaussian process from general sequence designs

Delphine Blanke

Avignon University, LMA EA2151
33 rue Louis Pasteur, F-84000 Avignon, France
e-mail: delphine.blanke@univ-avignon.fr

and

Céline Vial

Université de Lyon, CNRS UMR 5208, Polytech Lyon-Université de Lyon 1
Institut Camille Jordan, 43 blvd du 11 novembre 1918
69622 Villeurbanne Cedex, France
e-mail: celine.vial@univ-lyon1.fr

Abstract: We consider a real Gaussian process X with global unknown smoothness (r_0, β_0): more precisely $X^{(r_0)}$, $r_0 \in \mathbb{N}_0$, is supposed to be locally stationary with Hölder exponent β_0, $\beta_0 \in]0,1[$. For X observed at a finite set of points, we derive estimators of r_0 and β_0 based on the quadratic variations for the divided differences of X. Under mild conditions, we obtain an exponential bound for estimating r_0, as well as sharp rates of convergence (up to logarithmic factors) for the estimation of β_0. An extensive simulation study illustrates the finite-sample properties of both estimators for different types of processes and we also include two real data applications.

MSC 2010 subject classifications: Primary 62M05, 62M09, 62F12, 60G15; secondary 60G17, 65D05.
Keywords and phrases: Inference for Gaussian processes, locally stationary process, divided differences, quadratic variations, unevenly spaced data.

Received January 2014.

Contents

1 Introduction . . . 1153
2 The framework . . . 1155
  2.1 The process and its design . . . 1155
  2.2 The methodology . . . 1157
3 Asymptotic results . . . 1158
4 Approximation and integration . . . 1160
  4.1 Results for plug-in estimators . . . 1160
  4.2 Simulation results . . . 1162

5 Numerical results 1162
5.1 Results for estimators of r_0 1163
5.2 Estimation of H and β_0 1164
5.2.1 Quality of estimation 1164
5.2.2 Asymptotic properties 1165
5.2.3 Impact of misspecification of regularity 1166
5.2.4 Processes with varying trend or non constant function d_0 1166
5.3 Real data 1169
5.3.1 Roller data 1169
5.3.2 Biscuit data 1169
5.4 Irregularly spaced points 1170
6 Annexes 1172
6.1 Results for the regular sequence design 1172
6.2 Results on the mean behavior of $\overline{(D_r^{(u)} X)^2}$, $r \geq 1$ 1172
6.3 A general exponential bound for $|\overline{(D_r^{(u)} X)^2} - \mathbb{E}(\overline{(D_r^{(u)} X)^2})|$ . . 1175
6.4 Proofs of the main results of the Section 3 1179
6.5 Proof of the results on approximation and integration 1183
Acknowledgements 1185
References 1185

1. Introduction

In many application areas, such as geostatistics, chemistry and environmental science, Gaussian processes are a fundamental modeling tool. One reason for this is that the dependence structure is characterized by the covariance function. Often, the covariance is assumed to belong to some parametric family, where the unknown parameters determine the sample path regularity (Gneiting et al. [18]). It is therefore important to be able to estimate sample path regularity from observations of the process at discrete time points. In this paper, we introduce a new estimator of the global regularity of a Gaussian process X whose smoothness is described by a pair (r_0, β_0), where $r_0 \in \mathbb{N}_0$ and $\beta_0 \in]0, 1[$. More precisely, we suppose that $X^{(r_0)}$ is locally stationary with regularity β_0 where, if $r_0 \geq 1$, X is r_0-times mean-square differentiable with derivative denoted by $X^{(r_0)}$. In contrast to previous estimators of (r_0, β_0) proposed in the literature, we do not require X to be observed at equally spaced time points and X is not assumed to be stationary nor even to have stationary increments. In fact, most previous works on this topic have assumed r_0 to be known.

We consider designs where the sampled points can be associated with quantiles of some distribution, see Section 2 for details. Actually, many applications make use of non equidistant sampling and Stein [40] (chap. 6.9) gives a hint of how adding three points very near to the origin, among the already twenty equally spaced observations, improve drastically the estimation of the regularity parameter. Taking into account a non uniform design is innovating regarding other existing estimators and makes sense as to the remark above.

Estimation of r_0 is applied to interpolation and integration of the observed sampled path. A wide range of methods have been proposed in this field. For processes satisfying the so-called Sacks and Ylvisaker (SY) conditions, recent works include: Müller-Gronbach [27, orthogonal projection, optimal designs], Müller-Gronbach and Ritter [28, linear interpolation, optimal designs], Müller-Gronbach and Ritter [29, linear interpolation, adaptive designs]. Under Hölder type conditions, one may cite e.g. the works of Seleznjev [36, linear interpolation], Seleznjev [37, Hermite interpolation splines, optimal designs], Seleznjev and Buslaev [38, best approximation order], Blanke and Vial [6, 7, interpolation with piecewise Lagrange polynomials]. Note that a detailed survey may be found in the book by Ritter [33]. Also, in a time series context, Cambanis [11] analyzes three important problems (estimation of regression coefficients, estimation of random integrals and detection of signal in noise) for which he is looking for optimal designs. The latter two problems involve approximations of integrals, where knowledge of process regularity is particularly important: we provide a detailed discussion on this topic in Section 4. For other applications of the estimation of regularity, we refer also to Adler [2] where bounds of suprema distributions depend on the sample roughness, Istas [21] where it is involved in the choice of the best wavelet base in image analysis, or more generally to prediction area, see Cuzick [15], Lindgren [26], Bucklew [10].

Concerning the parameter β_0, it is closely linked to fractal dimension of sample paths. This relationship is developed in particular in the works by Adler [1] and Taylor and Taylor [41] and it gave rise to an important literature around estimation of β_0. Note that this relation is extended for non Gaussian processes in Hall and Roy [19]. The recent paper of Gneiting et al. [18] gives a review on estimation of the fractal dimension for times series and spatial data. They also provide a wide range of applications (all for equally spaced observations) in environmental science, e.g. hydrology, topography of sea floor. Still in this framework, we refer especially to Constantine and Hall [13] for estimators based on quadratic variations and their extensions developed by Kent and Wood [24]. For other estimators in the stationary case, we may refer to Chan et al. [12] for a periodogram-type estimator whereas Feuerverger et al. [16] use the number of level crossings.

Our motivation is to extend the previous works in several directions: we consider a Gaussian process X, observed at possibly unevenly spaced sampled points and, not supposed to be stationary or with stationary increments. Also, X has an unknown degree of differentiability, r_0, to be estimated and if $r_0 \geq 1$, the unknown coefficient of smoothness β_0 is related to the unobserved derivative $X^{(r_0)}$. Our methodology is based on an estimator of r_0, say $\widehat{r}_0$, derived from quadratic variations for divided differences of X and consequently, generalize the estimator based on equally spaced data of Blanke and Vial [7]. In a second step, we proceed to the estimation of β_0, with an estimator $\widehat{\beta}_0$ which can be viewed as a simplification of that studied, in the case $r_0 = 0$, by Constantine and Hall [13], Kent and Wood [24]. For equally spaced observations of processes with stationary increments, note that Istas and Lang [22] have proposed and studied an estimator of $H = 2(r_0 + \beta_0)$ by using a linear regression approach. As far as we can judge, our two steps procedure seems to be simpler and more competitive. We obtain an upper bound for $\mathbb{P}(\widehat{r}_0 \neq r_0)$ as well as the mean square error of $\widehat{r}_0$ and almost sure rates of convergence for $\widehat{\beta}_0$. Surprisingly, these rates are comparable to those obtained in the case of r_0 equal to 0: therefore, the preliminary estimation of r_0 does not affect that of β_0, even though $X^{(r_0)}$ is not observed. Next, in Section 4, we derive theoretical and numerical results concerning the approximation and integration problems. We complete this work with an extensive computational study: we compare different estimators of r_0 and $r_0 + \beta_0$ for various processes with different smoothness and, we derive properties of our estimators for finite sample size. Our numerical results show also the importance of well estimating r_0 to get a consistent estimation for β_0. To end this part, we apply our global estimation of (r_0, β_0) to two well-known real data sets: Roller data (Laslett [25]) and Biscuit data (Brown et al. [9]). Finally, proofs of all results are postponed at the end.

2. The framework

2.1. The process and its design

We consider a Gaussian process $X = \{X(t),\, t \in [0,T]\}$ with mean function $\mu(t) := \mathbb{E}X(t)$ and covariance function $\mathbb{K}(s,t) = \operatorname{Cov}(X(s), X(t))$. This process is assumed to be observed at $(n+1)$ instants on $[0,T]$, $T > 0$. We shall assume the following conditions on the regularity of X.

Assumption A2.1. X satisfies the following conditions.

(i) There exists some nonnegative integer r_0, such that X is r_0-times differentiable in quadratic mean, with r_0-th derivative $X^{(r_0)}$.

(ii) The process $X^{(r_0)}$ is assumed to be locally stationary:

$$\lim_{h\to 0} \sup_{s,t\in[0,T],|s-t|\le h, s\neq t} \left| \frac{\mathbb{E}\big(X^{(r_0)}(s) - X^{(r_0)}(t)\big)^2}{|s-t|^{2\beta_0}} - d_0(t) \right| = 0 \tag{2.1}$$

where $\beta_0 \in]0,1[$ and d_0 is a positive continuous function on $[0,T]$.

(iii-p) For either $p = 1$ or $p = 2$, the partial derivative $\mathbb{K}^{(r_0+p, r_0+p)}(s,t)$ exists on $[0,T]^2 \backslash \{s = t\}$ and satisfies for some $D_p > 0$:

$$\left|\mathbb{K}^{(r_0+p,r_0+p)}(s,t)\right| \le D_p \left|s-t\right|^{-(2p-2\beta_0)}.$$

Moreover, we suppose that $\mu \in C^{r_0+1}([0,T])$.

Note that the local stationarity makes reference to Berman [5]'s definition. The condition A2.1-(i) can be translated in terms of the mean and covariance functions. In particular, it implies that the function $\mathbb{K}$ is $2r_0$-times continuously differentiable with derivatives $\mathbb{K}^{(r,r)}(s,t) = \operatorname{Cov}(X^{(r)}(s), X^{(r)}(t))$, for $r = 1, \dots, r_0$ and all $(s,t) \in [0,T]^2$. Also, the mean of the process μ is a r_0-times continuously differentiable function with $\mathbb{E}X^{(r)}(t) = \mu^{(r)}(t)$, $r = 0, \dots, r_0$.

Conditions A2.1-(iii-p) are more technical but classical ones when estimating regularity parameters, see Constantine and Hall [13], Kent and Wood [24].

These assumptions are satisfied by a wide range of examples, e.g. the r_0-fold integrated fractional Brownian motion or the stationary Gaussian process with Matérn covariance, i.e. $\mathbb{K}(t,0) = \frac{\pi^{1/2}\phi}{2^{(\nu)-1}\Gamma(\nu+1/2)}(\alpha|t|)^\nu K_\nu(\alpha|t|)$, where K_ν, is a modified Bessel function of the second kind of order ν. The latter process gets a global smoothness equal to $(\lfloor\nu\rfloor, \nu - \lfloor\nu\rfloor)$, see Stein [40] p. 31. Detailed examples, including other classes of stationary processes, can be found in Blanke and Vial [6, 7]. Note that, for processes with stationary increments, the limit function d_0 of (2.1), is reduced to a constant. Of course, cases with non constant $d_0(\cdot)$ are allowed as well as processes with trend. In particular, for some sufficiently smooth functions a and m on $[0,T]$, the process $Y(t) = a(t)X(t) + m(t)$ will also fulfills Assumption A2.1. More precisely, we get the following result from Seleznjev [37] and straightforward computation.

Lemma 2.1. *Let X be a zero mean process with given regularity (r_0, β_0) and asymptotic function $d_0(t) \equiv C_{r_0,\beta_0}$ that satisfies A2.1(iii-p) ($p = 1$ or 2). For a positive function $a \in C^{r_0+p}([0,T])$ and $m \in C^{r_0+p}([0,T])$, if $Y(t) = a(t)X(t) + m(t)$, then Y has regularity (r_0, β_0) with asymptotic function $D_{r_0,\beta_0}(t) = a^2(t)\, C_{r_0,\beta_0}$ and satisfies A2.1(iii-p).*

Let us turn now to the description of the sampled points. We consider that the process X is observed at $(n+1)$ instants, denoted by

$$0 = t_{0,n} < t_{1,n} < \cdots < t_{n,n} \le T$$

where the $t_k := t_{k,n}$ form a regular sequence design. That is, they are defined as quantiles of a fixed positive and continuous density ψ on $[0,T]$:

$$\int_0^{t_k} \psi(s)\mathrm{d}s = \frac{k\delta_n}{T}, \quad k = 0, \dots, n,$$

for δ_n a positive sequence such that $\delta_n \to 0$ and $n\delta_n \to T^-$. Clearly, if ψ is the uniform density on $[0,T]$, one gets the equidistant case. Some further assumptions on ψ are needed to get some control over the t_k's.

Assumption A2.2. The density ψ satisfies:

(i) $\inf_{t\in[0,T]} \psi(t) > 0$,
(ii) $\forall\, (s,t) \in [0,T]^2$, $|\psi(s) - \psi(t)| \le L|s-t|^\alpha$, for some $\alpha \in]0,1]$.

These hypotheses ensure a controlled spacing between two distinct points of observation, see Lemma 6.1. From a practical point of view, this flexibility may allow to recognize inhomogeneities in the process (e.g. presence of pics in environmental pollution monitoring, see Gilbert [17] and references therein) or else to describe situations where data are collected at equidistant times but become irregularly spaced after some screening (see for example the wolfcamp-aquifer data in Cressie [14] p. 212).

2.2. The methodology

In this part, we give the background ideas about construction of our estimate of the global regularity (r_0, β_0) for unevenly spaced sampled data. The main idea is to introduce divided differences (generalizing the finite differences, studied by Blanke and Vial [7, 8]). Let us first recall that the unique polynomial of degree r that interpolates a function g at $r+1$ points $t_k, \dots, t_{k+ur}$ (for some positive integer u and $k \in \{0, \dots, n-ru\}$) can be written as:

$$g[t_k] + g[t_k, t_{k+u}](t - t_k) + g[t_k, t_{k+u}, t_{k+2u}](t-t_k)(t-t_{k+u}) \\ + \cdots + g[t_k, \dots, t_{k+ru}](t-t_k)\cdots(t-t_{k+(r-1)u}) \quad (2.2)$$

where the divided differences $g[\dots]$ are defined by $g[t_k] = g(t_k)$ and for $j = 1, \dots, r$ (using the Lagrange's representation):

$$g[t_k, \dots, t_{k+ju}] = \sum_{i=0}^{j} \frac{g(t_{k+iu})}{\prod_{m=0, m\neq i}^{j}(t_{k+iu} - t_{k+mu})}.$$

In particular, we write $g[t_k, \dots, t_{k+ru}] = \sum_{i=0}^{r} b_{ikr}^{(u)} g(t_{k+iu})$ with

$$b_{ikr}^{(u)} := \frac{1}{\prod_{m=0, m\neq i}^{r}(t_{k+iu} - t_{k+mu})}, \quad i = 0, \dots, r. \quad (2.3)$$

These coefficients are of particular interest. In fact, their first non-zero moments are of order r. We can also derive an explicit bound and an asymptotic expansion for $b_{irk}^{(u)}$, see lemma 6.2 for details. Then, for positive integers r and u, we consider the u-dilated divided differences of order r for X:

$$D_{r,k}^{(u)} X = \sum_{i=0}^{r} b_{ikr}^{(u)} X(t_{k+iu}), \quad k = 0, \dots, n - ur. \quad (2.4)$$

If $\psi(t) = T^{-1}\mathbb{1}_{[0,T]}(t)$, note that equidistant sampled points are obtained ($t_{k,n} = k\delta_n$), and divided differences become finite differences. More precisely, let define the finite differences $\Delta_{r,k}^{(u)} X = \sum_{i=0}^{r} a_{i,r} X((k+iu)\delta_n)$ with the sequence $a_{i,r} = \binom{r}{i}(-1)^{r-i}$. For equally spaced observations, one gets from (2.3) the relation:

$$b_{ikr}^{(u)} = \frac{(u\delta_n)^{-r}}{\prod_{m=0, m\neq i}^{r}(i-m)} = \frac{a_{i,r}}{r!}(u\delta_n)^{-r},$$

and we deduce that $D_{r,k}^{(u)} X = \frac{(u\delta_n)^{-r}}{r!}\Delta_{r,k}^{(u)} X$.

From now on, we study the u-dilated quadratic variations of X of order r and set: $\overline{(D_r^{(u)} X)^2} = \frac{\sum_{k=0}^{n-ur}(D_{r,k}^{(u)} X)^2}{n-ur+1}$. The construction of our estimators is based on the following asymptotic properties concerning the mean behavior of $\overline{(D_r^{(u)} X)^2}$:

$$(P)\left\{ \begin{array}{l} \text{the quantity } \mathbb{E}\,\overline{(D_r^{(u)} X)^2} \text{ is of order } (u/n)^{-2(p-\beta_0)} \text{ for } r = r_0 + p \\ \text{with } p = 1, 2, \text{ and gets a finite non zero limit when } r \leq r_0 \end{array} \right.$$

(see Proposition 6.1 for a rigourous derivation). These results imply that a good choice of r (namely $r = r_0 + 1$ or $r_0 + 2$) could provide an estimate of β_0, at least with an adequate combination of u-dilated quadratic variations of X. To this end, we propose a two steps procedure:

Step 1: *Estimation of* r_0.
Based on $D^{(1)}_{r,k}X$, $k = 0, \ldots, n-r$, we estimate r_0 with

$$\widehat{r}_0 = \min\left\{ r \in \{2, \ldots, m_n\} : \overline{\left(D^{(1)}_r X\right)^2} \geq n^2 b_n \right\} - 2. \tag{2.5}$$

If the above set is empty, we fix $\widehat{r}_0 = l_0$ for an arbitrary value $l_0 \notin \mathbb{N}_0$. Here, $m_n \to \infty$ but if an upper bound B is known for r_0, one has to choose $m_n = B + 2$. The threshold b_n is a positive sequence chosen such that: $n^{-2(1-\beta_0)} b_n \to 0$ and $n^{2\beta_0} b_n \to \infty$ for all $\beta_0 \in]0, 1[$. For example, omnibus choices are given by $b_n = (\ln n)^\alpha$, $\alpha \in \mathbb{R}$.

Step 2: *Estimation of* β_0.
Next, we derive two families of estimators for β_0, namely $\widehat{\beta}^{(p)}_n$, with either $p = 1$ or $p = 2$ and u, v given integers (with $u < v$):

$$\widehat{\beta}^{(p)}_n := \widehat{\beta}^{(p)}_n(u, v) = p + \frac{1}{2} \frac{\ln\left(\overline{\left(D^{(u)}_{\widehat{r}_0+p}X\right)^2}\right) - \ln\left(\overline{\left(D^{(v)}_{\widehat{r}_0+p}X\right)^2}\right)}{\ln(u/v)}.$$

Remark 2.1 (The case of $r_0 = 0$ with equally spaced observations)**.** Kent and Wood [24] proposed estimators of $2\beta_0 = \alpha$ based on ordinary and generalized least squares on the logarithm of the quadratic variations versus the logarithm of a vector of values u, more precisely

$$\hat{\alpha}^{(p)} = \frac{(\mathbf{1}^\top W \mathbf{1})(\mathbf{u}^\top W \mathbf{Q}^{(p)}) - (\mathbf{1}^\top W \mathbf{u})(\mathbf{1}^\top W \mathbf{Q}^{(p)})}{(\mathbf{1}^\top W \mathbf{1})(\mathbf{u}^\top W \mathbf{u}) - (\mathbf{1}^\top W \mathbf{u})^2}$$

where $\mathbf{1}$ is the m-vector of 1s, $\mathbf{u} = (\ln(u), u = 1, \ldots, m)^\top$, $\mathbf{Q}^{(p)} = (\ln \overline{(\Delta^{(u)}_{p+1}X)^2}$, $u = 1, \ldots, m)^\top$ and W is either the identity matrix I_m of order $m \times m$ or a matrix depending on (n, β_0) which converges to the asymptotic covariance of $n^{1/2}(\overline{(\Delta^{(u)}_{p+1}X)^2} - \mathbb{E}\overline{(\Delta^{(u)}_{p+1}X)^2})$. The ordinary least square estimator – corresponding to $W = I_m$– is denoted by $\widehat{\alpha}^{(p)}_{\mathrm{OLS}}$, where p is adapted to the regularity of the process (supposed to be known in their work). The choice $p = 0$, with the sequence $(-1, 1)$, leads to the estimator studied by Constantine and Hall [13]. For $(u, v) = (1, 2)$, note that one gets $\widehat{\beta}^{(1)}_n = \widehat{\alpha}^{(0)}_{\mathrm{OLS}}$ and $\widehat{\beta}^{(2)}_n = \widehat{\alpha}^{(1)}_{\mathrm{OLS}}$ but, even in this equidistant case, new estimators may be derived with other choices of (u, v) such as $(u, v) = (1, 4)$ (which seems to perform well, see Section 5.2).

3. Asymptotic results

In Blanke and Vial [7], an exponential bound is obtained for $\mathbb{P}(\widehat{r}_0 \neq r_0)$ in the equidistant case, implying that, almost surely for n large enough, $\widehat{r}_0$ is equal to r_0. Here, we generalize this result to regular sequence designs but also, we complete it with the mean square error of $\widehat{r}_0$.

Theorem 3.1. *Under Assumption A2.1 (fulfilled with $p = 1$ or $p = 2$) and A2.2, we have $\mathbb{P}(\widehat{r}_0 \neq r_0) = \mathcal{O}(\exp(-\varphi_n(p)))$ and $\mathbb{E}(\widehat{r}_0 - r_0)^2 = \mathcal{O}(m_n^3 \exp(-\varphi_n(p)))$, where, for some positive constant $C_1(r_0)$, $\varphi_n(p)$ is defined by*

$$\varphi_n(p) = C_1(r_0) \times \begin{cases} n \mathbb{1}_{]0,\frac{1}{2}[}(\beta_0) + n(\ln n)^{-1} \mathbb{1}_{\{\frac{1}{2}\}}(\beta_0) + n^{2-2\beta_0} \mathbb{1}_{]\frac{1}{2},1[}(\beta_0) & \text{if } p = 1 \\ n & \text{if } p = 2. \end{cases}$$

Note that one may choose m_n tending to infinity arbitrary slowly. Indeed, the unique restriction is that r_0 belongs to the grid $\{1, \ldots, m_n\}$ for n large enough. From a practical point of view, one may choose a preliminary fixed bound B, and, in the case where the estimator return the non-integer value l_0, replace B by B' greater than B.

The bias of $\widehat{\beta}_n^{(p)}$ will be controlled by a second-order condition of local stationarity, more specifically we have to strengthen the relation (2.1) into:

$$\lim_{h \to 0} \sup_{\substack{s,t \in [0,T], \\ |s-t| \leq h, \\ s \neq t}} \left| |s-t|^{-\beta_1} \left(\frac{\mathbb{E}\left(X^{(r_0)}(s) - X^{(r_0)}(t)\right)^2}{|s-t|^{2\beta_0}} - d_0(t) \right) - d_1(t) \right| = 0 \quad (3.1)$$

for a positive β_1 and continuous function d_1.

Theorem 3.2. *If relation (3.1), Assumption A2.2 and A2.1 (with either $p = 1$ or $p = 2$) are satisfied, we obtain*

$$\limsup_{n \to \infty} V_n^{(p)} \left| \widehat{\beta}_n^{(p)} - \beta_0 \right| \leq C_1(p) \quad a.s.$$

where $C_1(p)$ is some positive constant and

$$V_n^{(1)} = \min\left(n^{\beta_1}, \sqrt{\frac{n}{\ln n}} \mathbb{1}_{]0,\frac{3}{4}[}(\beta_0) + \frac{\sqrt{n}}{\ln n} \mathbb{1}_{\{\frac{3}{4}\}}(\beta_0) + \frac{n^{2(1-\beta_0)}}{\ln n} \mathbb{1}_{]\frac{3}{4},1[}(\beta_0) \right),$$
$$V_n^{(2)} = \min\left(n^{\beta_1}, \sqrt{\frac{n}{\ln n}} \right).$$

Remark 3.1 (Rates of convergence with equally spaced observations)**.** For stationary Gaussian processes, Kent and Wood [24] give the mean square error and convergence in distribution of their estimator described in Remark 2.1. They obtained, for both families $p = 1$ and $p = 2$, the same rate up to a logarithmic order, due here to almost sure convergence. The asymptotic distribution is either of Gaussian or of Rosenblatt type depending on whether β_0 is less or greater than $3/4$. For stationary increment processes, Istas and Lang [22] introduced an estimator of $H = 2(r_0 + \beta_0)$ by following a global linear regression approach (based on an asymptotic equivalent of the quadratic variation with the choice of some adequate family of sequences). For $r_0 = 0$, their approach matches with the previous one, with $p = 1$, in using dilated sequence of type $a_{jr} = \binom{r}{j}(-1)^{r-j}$. Assuming a known upper bound on r_0, they derived convergence in distribution to a centered Gaussian variable with rate depending on β_0: root-n for $\beta_0 \leq 3/4$ and, $n^{1/2-\alpha(2\beta_0-3/2)}$ for $\beta_0 > 3/4$ and $\delta_n = n^{-\alpha}$ with $0 < \alpha < 1$. In this last case, to obtain a Gaussian limit they have to assume that r_0 is known, the observation interval is no more bounded, and the rate of convergence is lower than root-n.

4. Approximation and integration

4.1. Results for plug-in estimators

A classical and interesting topic in the Gaussian framework is interpolation and/or integration of a sampled process. Let $X = \{X_t,\, t \in \mathbb{R}\}$, be observed at sampled times $t_{0,n}, \ldots, t_{n,n}$ (denoted in the sequel by $t_0, \ldots, t_n$) over some interval $[a, b]$. We want to approximate X (on $[a, b]$) or its (weighted) integral $I_\rho = \int_a^b X(t)\rho(t)\,\mathrm{d}t$ for some known function ρ. An extensive literature exists on these topics, we refer particularly to the recent monograph of Ritter [33] for a detailed overview. Actually, approximation and integration problems are closely linked (see for example, the latter reference p. 19–21). For the specific case of locally stationary derivatives, we may refer to the works of Plaskota et al. [30] and Benhenni [3] for, respectively, the approximation and integration problems. For sake of clarity, we give a brief summary of their obtained results. In the following, we denote by $\mathcal{H}(r_0, \beta_0)$ the family of Gaussian processes having r_0 derivatives in quadratic mean and r_0-th derivative with Hölderian regularity of order $\beta_0 \in]0, 1[$. For measurable $g_i(\cdot)$, we consider the approximation $\mathcal{A}_{n,g}(t) = \sum_{i=0}^n X(t_i)g_i(t)$ and the corresponding, weighted and integrated, L^2-error $e_\rho(\mathcal{A}_{n,g})$ with $e_\rho^2(\mathcal{A}_{n,g}) := \int_a^b \mathbb{E}\left|X(t) - \mathcal{A}_{n,g}(t)\right|^2 \rho(t)\,\mathrm{d}t$ with ρ supposed to be positive and continuous. For $X \in \mathcal{H}(r_0, \beta_0)$ and known (r_0, β_0), Plaskota et al. [30] have shown that

$$\begin{aligned} 0 < c(r_0, \beta_0) &\leq \varliminf_{n\to\infty} n^{r_0+\beta_0} \inf_g e_\rho(\mathcal{A}_{n,g}) \\ &\leq \varlimsup_{n\to\infty} n^{r_0+\beta_0} \inf_g e_\rho(\mathcal{A}_{n,g}) \leq C(r_0, \beta_0) < +\infty \end{aligned}$$

for equidistant sampled times $t_1, \ldots, t_n$ and Gaussian processes defined and observed on the half-line $[0, +\infty[$. Of course, optimal choices of functions g_i, giving a minimal error, depend on the unknown covariance function of X.

For weighted integration, the quadrature is denoted by $\mathcal{Q}_{n,d} = \sum_{i=0}^n X(t_i)d_i$ with well-chosen constants d_i (typically, one may take $d_i = \int_a^b g_i(t)\,\mathrm{d}t$). For known (r_0, β_0), a short list of references could be:

- Sacks and Ylvisaker [34, 35] with $r_0 = 0$ or 1, $\beta_0 = \frac{1}{2}$ and known covariance,
- Benhenni and Cambanis [4] for arbitrary r_0 and $\beta_0 = \frac{1}{2}$,
- Stein [39] for stationary processes and $r_0 + \beta_0 < \frac{1}{2}$,
- Ritter [32] for minimal error, under Sacks and Ylvisaker's conditions, and with arbitrary r_0.

Let us set $e^2_\rho(\mathcal{Q}_{n,d}) := \mathbb{E}\left|I_\rho - \mathcal{Q}_{n,d}\right|^2$, the mean square error of integration. In the stationary case and for known r_0, Benhenni [3] established the following exact behavior: if $\rho \in C^{r_0+3}([a,b])$ then for some given quadrature $\mathcal{Q}_{n,d^*(r_0)}$ on $[a,b]$,

$$n^{r_0+\beta_0+\frac{1}{2}}\, e_\rho(\mathcal{Q}_{n,d^*(r_0)}) \xrightarrow[n\to\infty]{} c_{r_0,\beta_0}\left(\int_a^b \rho^2(t)\psi^{-(2(r_0+\beta_0)+1)}(t)\,\mathrm{d}t\right)^{\frac{1}{2}}$$

where ψ is the density relative to the regular sampling $\{t_1,\dots,t_n\}$. Moreover, following Ritter [32], it appears that this last result is optimal, at least under Sacks and Ylvisaker's conditions. Finally, Istas and Laredo [23] have proposed a quadrature, requiring only an upper bound on r_0, and with also an error of order $\mathcal{O}\big(n^{-(r_0+\beta_0+\frac{1}{2})}\big)$.

All these results show the importance of well estimating r_0 and motivate ourselves to focus on plug-in interpolators, namely those using Lagrange polynomials, with order estimated by $\widehat{r}_0$. More precisely, the Lagrange interpolation of order $r \geq 1$ for the sampled process X is defined by

$$\widetilde{X}_r(t) = \sum_{i=0}^{r} L_{i,k,r}(t) X\big(t_{kr+i}\big), \text{ with } L_{i,k,r}(t) = \prod_{\substack{j=0\\ j\neq i}}^{r} \frac{(t-t_{kr+j})}{t_{kr+i}-t_{kr+j}}, \tag{4.1}$$

for $t \in \mathcal{I}_k := [t_{kr}, t_{kr+r}]$, $k = 0,\dots,\lfloor\frac{n}{r}\rfloor - 2$ and $\mathcal{I}_{\lfloor\frac{n}{r}\rfloor-1} = [\lfloor t_{\lfloor(\frac{n}{r}\rfloor-1)r}, T]$. Our plug-in method consists in the approximation defined for $t \in [0,T]$ with $\mathcal{A}_{n,L}(t) = \widetilde{X}_{\max(\widehat{r}_0,1)}(t)$, and subsequent quadrature $\mathcal{Q}_{n,L} = \int_0^T \widetilde{X}_{\widehat{r}_0+1}(t)\,\rho(t)\,\mathrm{d}t$. Indeed, Lagrange polynomials are of easy implementation, they don't require the knowledge of the covariance structure and they reach the optimal rate of approximation, $n^{-(r_0+\beta_0)}$ (see Plaskota et al. [30]) in the case where r_0 is known. Our following result shows that the associate quadrature has also the expected rate, $n^{-(r_0+\beta_0+\frac{1}{2})}$. In the weighted case and for $T>0$, we obtain the following asymptotic bounds in the case of a regular design.

Theorem 4.1. *Suppose that conditions A2.1(i)–(ii) and A2.2 hold, choose a logarithmic order for m_n in (2.5) and consider a positive and continuous weight function ρ.*

(a) Under condition A2.1(iii-1), we have

$$e_\rho(\mathrm{app}(\widehat{r}_0)) := \Big(\int_0^T \mathbb{E}\left|X(t) - \widetilde{X}_{\max(\widehat{r}_0,1)}(t)\right|^2 \rho(t)\,\mathrm{d}t\Big)^{1/2} = \mathcal{O}\big(n^{-(r_0+\beta_0)}\big),$$

(b) if condition A2.1(iii-2) holds:

$$e_\rho(\mathrm{int}(\widehat{r}_0)) := \left(\mathbb{E}\left|\int_0^T (X(t) - \widetilde{X}_{\widehat{r}_0+1}(t))\rho(t)\,\mathrm{d}t\right|^2\right)^{1/2} = \mathcal{O}\big(n^{-(r_0+\beta_0+\frac{1}{2})}\big).$$

In conclusion, expected rates for approximation and integration are reached by plugging-in Lagrange piecewise polynomials. Of course if r_0 is known, the latter result holds true with $\widehat{r}_0$ replaced by r_0.

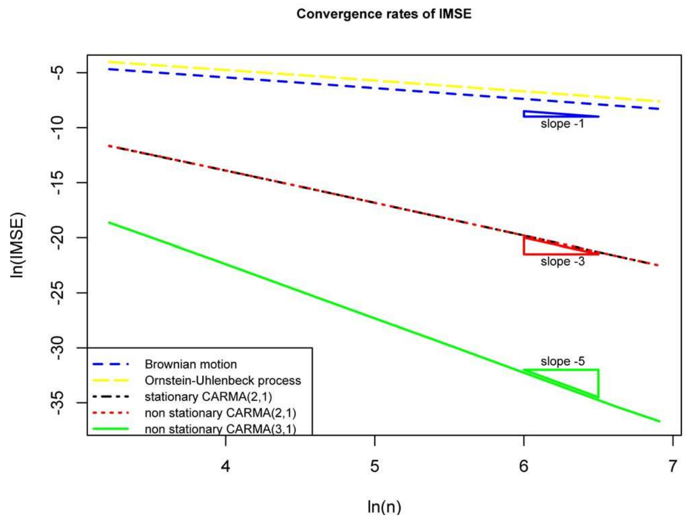


Fig 4.1. *Logarithm of* $e_1^2(\mathrm{app}(\widehat{\mathrm{r}}_0))$*, i.e. the IMSE, in function of* $\ln n$*, for different processes. The dashed line corresponds to Brownian motion, long dashed line to O.U., dashed line to non stationary CARMA(2,1), dotted-dashed line to stationary CARMA(2,1) and solid line to non stationary CARMA(3,1). The small triangles near lines indicate the theoretic slope.*

4.2. Simulation results

We illustrate the results of approximation with a first numerical study. Note that all numerical results of the paper are obtained by simulation of trajectories using two different methods: for stationary processes or with stationary increments we use the procedure described in Wood and Chan [43] and for CARMA (continuous ARMA) processes, we use Tsai and Chan [42]. Each of them consists in n equally spaced sampled points on $[0,1]$ and 1000 simulated sample paths. Also, all computations have been performed with the R software (R Core Team [31]).

The Figure 4.1 illustrates results of approximation for different processes. The logarithm of empirical integrated mean square error (in short IMSE), i.e. $e_1^2(\mathrm{app}(\widehat{\mathrm{r}}_0))$, is drawn as a function of $\ln n$ with a range of sample size from $n = 25$ to 1000. We may notice that we obtain straight lines with slope very near to $-H = -2(r_0+\beta_0)$. Since the Ornstein-Uhlenbeck process is a scaled time-transformed Wiener process, intercepts are different contrary to the stationary versus the non-stationary continuous ARMA processes.

5. Numerical results

In the early part of this section, we restrict ourselves to the equidistant case (with the choice $\psi(t) = \frac{1}{T}\mathbb{1}_{[0,T]}(t)$) in order to numerically compare our estimators with existing ones. As noticed before, we get for $a_{i,r} = \binom{r}{i}(-1)^{r-i}$ and $\Delta_{r,k}^{(u)} =$

TABLE 5.1
Value of the empirical probability that $\widehat{r}_0$ or $\tilde{r}_n$ equals r_0 or r_0+1 with $n=10$ or 25

	Wiener process, $r_0=0$		CARMA(2,1), $r_0=1$		CARMA(3,1), $r_0=2$	
	Number of equally spaced observations n					
event	10	25	10	25	10	25
$\tilde{r}_n=r_0$	0.995	1.000	0.913	1.000	0.585	0.999
$\tilde{r}_n=r_0+1$	0.005	0.000	0.087	0.000	0.415	0.001
$\widehat{r}_0=r_0$	1.000	1.000	1.000	1.000	0.999	1.000
$\widehat{r}_0=r_0+1$	0.000	0.000	0.000	0.000	0.001	0.000

$\sum_{i=0}^{r} a_{i,r}X((k+iu)\delta_n)$, the relation $D_{r,k}^{(u)}X=\frac{(u\delta_n)^{-r}}{r!}\Delta_{r,k}^{(u)}X$. Theorems 3.1 and 3.2 imply in turn that

$$\widehat{H}_n^{(p)}(u,v)=\frac{\ln\left(\overline{\left(\Delta_{\widehat{r}_0+p}^{(u)}X\right)^2}\right)-\ln\left(\overline{\left(\Delta_{\widehat{r}_0+p}^{(v)}X\right)^2}\right)}{\ln(u/v)} \tag{5.1}$$

is a consistent estimator of $H=2(r_0+\beta_0)$, for all (u,v), $u\neq v$. The last part of the section is devoted to the case of unevenly spaced data.

5.1. Results for estimators of r_0

This section is dedicated to the numerical properties of two estimators of r_0. We consider the estimator introduced by Blanke and Vial [7], derived from (2.5) in the equidistant case. An alternative, says $\tilde{r}_n$, based on Lagrange interpolator polynomials was proposed by Blanke and Vial [6]. More precisely, for $\delta_n=n^{-1}$ et $T=1$, $\tilde{r}_n$ is defined by

$$\tilde{r}_n=\min\Big\{r\in\{1,\dots,m_n\}:\frac{1}{r\widetilde{n}_r}\sum_{k=0}^{r\widetilde{n}_r-1}\Big(X\Big(\frac{2k+1}{n}\Big)-\widetilde{X}_r\Big(\frac{2k+1}{n}\Big)\Big)^2 \geq n^{-2r}b_n\Big\}-1$$

where $\widetilde{n}_r=\lfloor\frac{n}{2r}\rfloor$ and $\widetilde{X}_r(t)$ is defined for all $t\in[0,1]$ and each $r\in\{1,\dots,m_n\}$ as follows. There exist $k=0,\dots,\widetilde{n}_r-1$ such that for $t\in\mathcal{I}_{2k}:=[\frac{2kr}{n},\frac{2(k+1)r}{n}]$, the piecewise Lagrange interpolation of $X(t)$, $\widetilde{X}_r(t)$, is given by

$$\widetilde{X}_r(t)=\sum_{i=0}^{r}L_{i,k,r}(t)X\big((kr+i)n^{-1}\big)\text{ with }L_{i,k,r}(t)=\prod_{\substack{j=0\\j\neq i}}^{r}\frac{(t-(kr+j)n^{-1})}{(i-j)n^{-1}}.$$

Both estimators use the critical value b_n which is involved in detection of the jump. Here, due to convergence properties, we make the choice $b_n=(\ln n)^{-1}$. The Table 5.1 illustrates the strong convergence of both estimators and shows that this convergence is valid even for small number of observation points n (up to 10 for the estimator $\widehat{r}_0$). We may noticed that, in the case of bad estimation, our estimators overestimate the number of derivatives. Also for identical sample

Table 5.2
Value of the empirical probability that $\widehat{r}_0$ or $\tilde{r}_n$ equals r_0 for a fractional Brownian motion or an integrated one with fractal index $2\beta_0$

	$\widehat{r}_0 = r_0$					$\tilde{r}_n = r_0$				
	number of equally spaced observations n									
	50	100	500	1000	1200	50	100	500	1000	1200
fBm β_0										
0.90	1.000	1.000	1.000	1.000	1.000	0.655	0.970	1.000	1.000	1.000
0.95	0.969	0.999	1.000	1.000	1.000	0.002	0.002	0.004	0.134	0.331
0.97	0.242	0.521	1.000	1.000	1.000	0.000	0.000	0.000	0.000	0.000
0.98	0.019	0.015	0.0420	0.5258	0.759	0.000	0.000	0.000	0.000	0.000
ifBm β_0										
0.02	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
0.90	1.000	1.000	1.000	1.000	1.000	0.000	0.000	0.645	0.999	1.000
0.95	0.305	0.888	1.000	1.000	1.000	0.000	0.000	0.000	0.000	0.000
0.97	0.000	0.000	0.292	0.993	1.000	0.000	0.000	0.000	0.000	0.000

paths, note that $\widehat{r}_0$ seems to be more robust than $\tilde{r}_n$. This behavior was expected as the latter uses only half of the observations for the detection of the jump in quadratic mean. In these first results, processes have fractal index β_0 equals to $1/2$, but alternative choices of β_0 are of interest, so we consider the fractional Brownian motion (in short fBm) and the integrated fractional Brownian motion (in short ifBm), with respectively $r_0 = 0$ and $r_0 = 1$ and various values of β_0 in $]0, 1[$.

The Table 5.2 shows that $\widehat{r}_0$ succeeds in estimating the true regularity for β_0 up to 0.9. Of course the number of observations must be large enough and, even more important for large values of r_0 in the case where β_0 is greater or equal to 0.95. This latter result is clearly apparent when one compares the errors obtained for an ifBm with $\beta_0 = 0.95$ and a fBm with $\beta_0 = 0.95$. Finally, we can see once more that $\tilde{r}_n$ is less robust against increasing values of β_0, whereas our simulations have shown that, for $n = 2000$ and each simulated path, the estimator $\widehat{r}_0$ is able to distinguish processes with regularity $(0, 0.98)$ and $(1, 0.02)$, an almost imperceptible difference!

5.2. Estimation of H and β_0

This part is dedicated to the estimation of the global regularity $H = 2r_0 + 2\beta_0$ and gives numerical properties of estimators $\widehat{H}_n^{(p)}(u, v)$, defined in (5.1), for $p = 1$ or 2. Here, we present only the results obtained for $H_n^{(p)} = H_n^{(p)}(1, 4)$ since they seem to be the more homogeneous among all the choices we have tested for (u, v). Note that other values for v will be implemented in the case of real data arising in Section 5.3.

5.2.1. Quality of estimation

For this numerical part, we focus on the study of fBm, ifBm and, CARMA(3,1) with $r_0 = 2$, $\beta_0 = 0.5$. The Table 5.3 illustrates the performance of our esti-

TABLE 5.3
Values of mean square error and bias (between brackets) for estimators $\widehat{H}_n^{(p)}$, for $p = 1$ or 2 and $n = 1000$

	fBm β_0				
	0.2	0.5	0.8	0.9	0.95
$\widehat{H}_n^{(1)}$	0.0017 (0.0001)	0.0018 (−0.0023)	0.0034 (−0.0098)	0.0053 (−0.0273)	0.0063 (−0.0426)
$\widehat{H}_n^{(2)}$	0.0030 (0.0017)	0.0039 (0.0005)	0.0040 (−0.0013)	0.0039 (−0.0025)	0.0039 (−0.0040)
	ifBm β_0				
	0.2	0.5	0.8	0.9	0.95
$\widehat{H}_n^{(1)}$	0.0031 (−0.0029)	0.0026 (−0.0004)	0.0040 (−0.0106)	0.0058 (−0.0299)	0.0069 (−0.0444)
$\widehat{H}_n^{(2)}$	0.0054 (−0.0041)	0.0050 (0.0005)	0.0045 (0.0005)	0.0043 (0.0004)	0.0042 (0.0004)

mators for fBm and ifBm: we compute the empirical mean-square error from our 1000 simulated sample paths and $n = 1000$ equally spaced observations are considered. It appears that, contrary to $\widehat{H}_n^{(2)}$, the estimator $\widehat{H}_n^{(1)}$ slightly deteriorates for values of β_0 greater than 0.8. This result is in agreement with the rate of convergence of Theorem 3.2, depending on β_0 for this estimator. Also, we note that the bias seems to be unsensitive to the value of r_0 while, for both estimators, the mean-square error is slightly worsened for r_0 going from 0 to 1. Finally, for β_0 less than 0.8, $\widehat{H}_n^{(1)}$ seems preferable to $\widehat{H}_n^{(2)}$, possibly due to a lower variance of this estimator. Nevertheless, both estimators perform globally well on these numerical experiments. In the case of a CARMA(3,1) process, we compute again 1000 simulated sample paths but, due to algorithmic limitations, with only $n = 950$ equally spaced observations. We obtain for $\widehat{H}_n^{(1)}$, an empirical mean-square error of 0.0032 and an empirical bias equals to −0.0095, and for $\widehat{H}_n^{(2)}$, an empirical mean-square error of 0.0058 and an empirical bias equal to 0.0015. Then, regarding the variance, $\widehat{H}_n^{(1)}$ is also preferable in that case.

5.2.2. Asymptotic properties

In this part, we focus on the asymptotic properties related to Theorem 3.2 where it appears that, contrary to $\widehat{\beta}_n^{(2)}$, the rate of convergence obtained for $\widehat{\beta}_n^{(1)}$ depends on β_0. The Table 5.4 gives quantitative illustration of these results with the estimated slope obtained by the regression of $\ln(\mathbb{E}|\widehat{H}_n^{(p)} - H|)$ on $\ln n$, for n varying in the set $\{400, 600, 700, 800, 900, 1000, 1200\}$ and $\mathbb{E}|\widehat{H}_n^{(p)} - H|$ estimated from our 1000 simulated sample paths. As expected, the slope (corresponding to our polynomial rates of convergence of Theorem 3.2) is constant and approximatively equal to 0.5 for $\widehat{H}_n^{(2)}$ while, for $\widehat{H}_n^{(1)}$, the decrease is apparent for high values of β_0. Finally, the Figure 5.1 illustrates the behavior of the estimators $\widehat{H}_n^{(p)}$ with $p = 1$ or 2, for different values of the regularity parameter β_0. As we can see, the boxplots deteriorate only slightly for $n = 100$ and 250 when β_0 increases from 0.5 to 0.8 but the dispersion for $\widehat{H}_n^{(2)}$ is quite larger. For $\beta_0 = 0.95$, $\widehat{H}_n^{(2)}$ clearly outperforms $\widehat{H}_n^{(1)}$ if $n = 500$. Estimation appears to be

Table 5.4
Results for linear regression of $\ln(\mathbb{E}|\widehat{H}_n^{(p)} - H|)$ *on* $\ln n$ *for* n *in* $\{400, 600, 700, 800, 900, 1000, 1200\}$ *with* $\mathbb{E}|\widehat{H}_n^{(p)} - H|$ *estimated on 1000 simulated sample paths*

	$\widehat{H}_n^{(1)}$			$\widehat{H}_n^{(2)}$		
	theoretical slope	slope	R^2	theoretical slope	slope	R^2
fBm $\beta_0 = 0.5$	-0.5	-0.4754	0.998	-0.5	-0.4923	0.995
$\beta_0 = 0.6$	-0.5	-0.4532	0.999	-0.5	-0.4937	0.995
$\beta_0 = 0.7$	-0.5	-0.4064	0.995	-0.5	-0.494	0.997
$\beta_0 = 0.8$	-0.4	-0.3210	0.987	-0.5	-0.4938	0.998
$\beta_0 = 0.9$	-0.2	-0.2172	0.988	-0.5	-0.4970	0.998
$\beta_0 = 0.95$	-0.1	-0.1725	0.993	-0.5	-0.5027	0.998
ifBm $\beta_0 = 0.9$	-0.2	-0.2592	0.979	-0.5	-0.5480	0.998
$\beta_0 = 0.95$	-0.1	-0.2034	0.968	-0.5	-0.5461	0.998

more difficult for smaller values of n, but it is a quite typical behavior in this framework.

5.2.3. Impact of misspecification of regularity

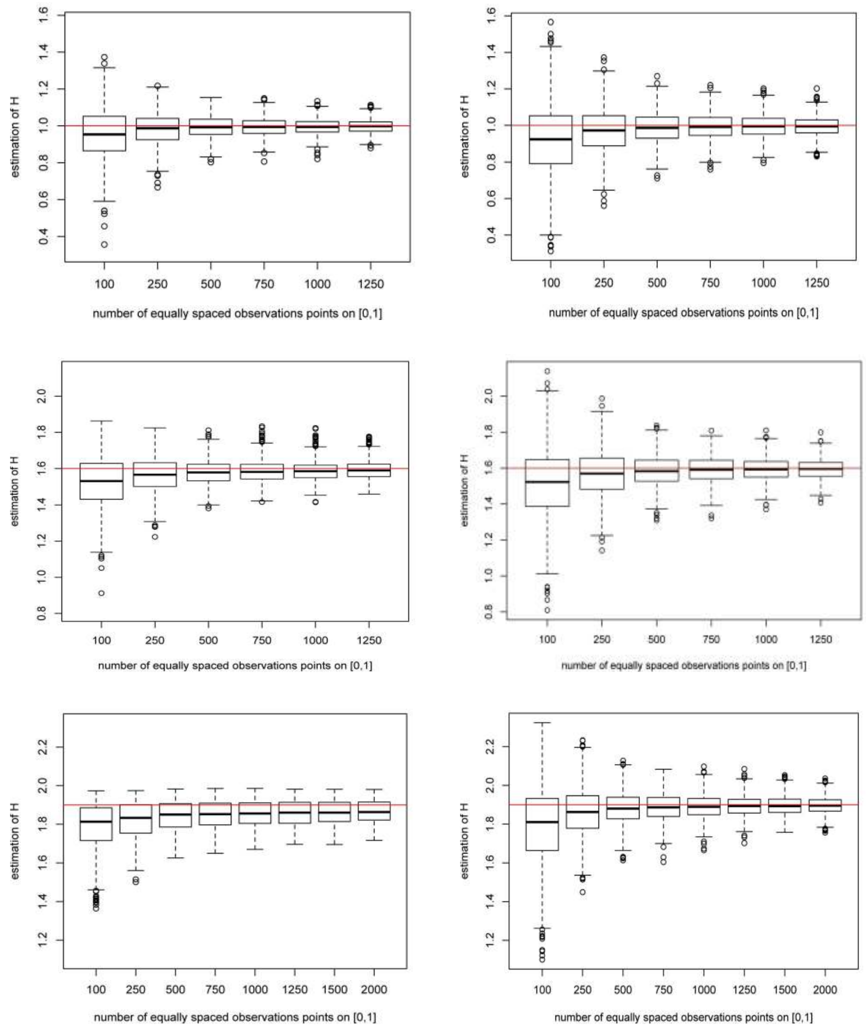
The Table 5.5 illustrates the impact of estimating β_0 when the order r in quadratic variation is misspecified. As we have seen, such an estimation requires the knowledge of r_0 or an upper bound of it. On the other hand, working with a too high value of r_0 may induce artificial variability in estimation. So, a precise estimation of r_0 is important. Recall that $\widehat{\beta}_n^{(p)}$ is defined by

$$\widehat{\beta}_n^{(p)} = p + \frac{1}{2}\frac{\ln\left(\overline{\left(D_{\widehat{r}_0+p}^{(u)}X\right)^2}\right) - \ln\left(\overline{\left(D_{\widehat{r}_0+p}^{(v)}X\right)^2}\right)}{\ln(u/v)},$$

with $p = 1$ or 2. Now, with an under-estimation of r_0, namely if one works with $p + (\ln(\overline{(D_{r+p}^{(u)}X)^2}) - \ln(\overline{(D_{r+p}^{(v)}X)^2}))/(2\ln(u/v))$ for any value of $r + p$ less or equal to r_0, results of Proposition 6.1 suggest that the resulting estimate should be close to p (because $\mathbb{E}(D_r^{(u)}X)^2$ has a finite limit $\ell_\psi(r)$ not depending on u for $r = 1, \ldots, r_0$). Our following numerical results confirm that the estimator of β_0 is no more consistent in this case and gives a value close to p for too small estimates of r_0.

5.2.4. Processes with varying trend or non constant function d_0

All previous examples are locally stationary with a constant function d_0 in the condition (2.1). Processes meeting our conditions but with no stationary increments may be constructed with the Lemma 2.1. As an example, from X a standard Wiener process ($r_0 = 0$, $\beta_0 = 0.5$) or an integrated version ($r_0 = 1$, $\beta_0 = 0.5$), we simulate $Y(t) = (t^{r_0+0.7}+1)\,X(t)$ with regularity $(r_0, 0.5)$ and $d_0(t)$ equal to $(t^{r_0+0.7}+1)^2$. The Figure 5.2 illustrates a Wiener sample path and its obtained transformation. Results are summarized in the Table 5.6: comparing with the Table 5.3 ($\beta_0 = 0.5$), it appears that the estimation is only slightly



$\widehat{H}_n^{(1)}$ $\widehat{H}_n^{(2)}$

FIG 5.1. *Each boxplot corresponds to 1000 estimations of H by $\widehat{H}_n^{(1)}$ on the left, and $\widehat{H}_n^{(2)}$ on the right of the graph. Each realization consists in n equally spaced observations on $[0,1]$ of a fBm with $\beta_0 = 0.5$ (top), $\beta_0 = 0.8$ (middle), $\beta_0 = 0.95$ (bottom), and* $n = 100, 250, 500, 750, 1000, 1250, 1500, 2000$. *The solid line corresponds to the real value of H.*

deteriorated for $r_0 = 1$ but of the same order when $r_0 = 0$. Other non stationary processes may also be obtained by adding some smooth trend. To this aim, we use the same sample paths as for the Table 5.3 and add the trend $m(t) = (1+t)^2$, see the Figure 5.3. We may noticed in the Table 5.7 that we obtain exactly the same results for the estimator $\widehat{H}_n^{(2)}$ and that, again, only a slight loss is observed for $\widehat{H}_n^{(1)}$.

TABLE 5.5
Mean values of the estimator $\widehat{\beta}_n^{(1)}$ (with standard deviation between brackets) based on under-estimates of r_0. The 'good' case $\widehat{r}_0 = 1$ for ifBm is also reported in italics for comparison

	$\widehat{r}_0 = 1$		$\widehat{r}_0 = 0$	
	number n of equidistant observations			
	100	500	100	500
ifBm $\beta_0 = 0.2$	*0.1884* (*0.0924*)	*0.1977* (*0.0423*)	0.9627 (0.0319)	0.9829 (0.0124)
$\beta_0 = 0.5$	*0.4870* (*0.0838*)	*0.4982* (*0.0366*)	0.9883 (0.0173)	0.9980 (0.0030)
$\beta_0 = 0.8$	*0.7682* (*0.0814*)	*0.7933* (*0.0408*)	0.9942 (0.0121)	0.9991 (0.0021)
CARMA(3,1) (r_0=2,β_0=0.5)	0.9829 (0.0272)	0.9968 (0.0053)	0.9967 (0.0069)	0.9992 (0.0015)

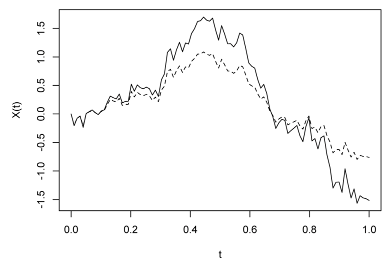


FIG 5.2. *Wiener process (solid) and its locally stationary transformation (dashed) used in the Table 5.6.*

TABLE 5.6
Value of MSE and bias (between brackets) for non constant $d_0(\cdot)$

	Wiener	Integrated Wiener
$\widehat{H}_n^{(1)}$	0.0021 (−0.0010)	0.0033 (0.0104)
$\widehat{H}_n^{(2)}$	0.0043 (−0.0001)	0.0057 (−0.0025)

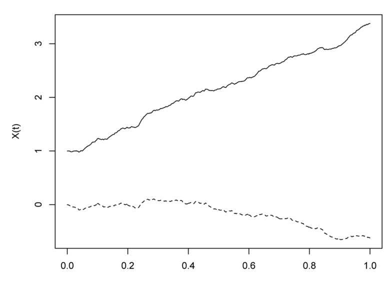


FIG 5.3. *Sample path of a fBm with $\beta_0 = 0.8$ (dashed line) and the same with a trend $m(t) = (1+t)^2$ (solid line).*

Table 5.7
Value of MSE and bias (between brackets) for estimators $\widehat{H}_n^{(p)}$, for $p = 1$ or 2 in the presence of a smooth trend

	fBm		ifBm	
	$\beta_0 = 0.5$	$\beta_0 = 0.8$	$\beta_0 = 0.5$	$\beta_0 = 0.8$
$\widehat{H}_n^{(1)}$	0.0022 (0.0172)	0.0289 (0.1544)	0.0027 (0.0126)	0.0148 (0.0903)
$\widehat{H}_n^{(2)}$	0.0039 (0.0005)	0.0040 (−0.0013)	0.0050 (0.0005)	0.0045 (0.0005)

Table 5.8
Estimates in the roller height example

m	$\widehat{H}_n^{(1)}(1,m)$	$\widehat{\alpha}_{\text{OLS}}^{(0)}$	$\widehat{H}_n^{(2)}(1,m)$	$\widehat{\alpha}_{\text{OLS}}^{(1)}$
2	0.63	0.63	0.77	0.77
4	0.50	0.51	0.64	0.66
6	0.38	0.39	0.49	0.51
8	0.35	0.33	0.44	0.43
10	0.31	0.28	0.39	0.36

5.3. Real data

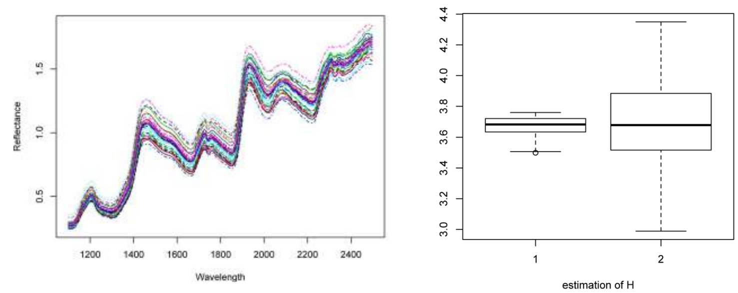
Let us turn to examples based on real data sets. In this part, we compare our estimators of H with those proposed by Constantine and Hall [13], Kent and Wood [24]. We compute estimated values by setting $(u, v) = (1, m)$ in our estimator, see (5.1), with m in $\{2, 4, 6, 8, 10\}$ while for $\widehat{\alpha}_{\text{OLS}}^{(\ell)}$, $\ell = 0, 1, 2$, defined in Remark 2.1, regression is carried out over $\mathbf{u} = (\ln(u), u = 1, \ldots, m)^\top$.

5.3.1. Roller data

We first focus on roller height data introduced by Laslett [25], which consists in $n = 1150$ heights measured at 1 micron intervals along a drum of a roller. This example was already studied in Kent and Wood [24]: they noticed that local self similarity may hold at sufficiently fine scales, so the regularity r_0 was supposed to be zero. Indeed, our estimator $\widehat{r_0}$, directly used on the data with $b_n = 1/\ln(n)$, gives $\widehat{r_0} = 0$ (with a value of $n^{4-2}\overline{(\Delta_2^{(1)} X)^2}$ equal to 1172345). Next, we compute the values obtained for the estimation of H in the Table 5.8, where values of estimates proposed by Constantine and Hall [13], Kent and Wood [24], $\alpha_{\text{OLS}}^{(\ell)}$, are also reported for comparison ($\ell = 0$ corresponds to the choice $(-1, 1)$ for a_{jr} and $\ell = 1$ to the choice $(1, -2, 1)$). It should be observed that our simplified estimators present a similar sensitivity to the choice of m.

5.3.2. Biscuit data

Now, in order to compare the (empirical) variances of these estimators, we consider a second example studied by Brown et al. [9]. The experiment involved varying the composition of 40 biscuit dough pieces and, data consist in near infrared reflectance (NIR) spectra for the same dough. The curves are graphed on the Figure 5.4. Each represents the near-infrared spectrum reflectance mea-



(a) (b)

Fig 5.4. *(a) Curve drawing reflectance in function of wavelength, varying between* 1100 *and* 2498. *(b) Box-plots for both estimators on the left* $\widehat{H}_n^{(1)}$, *on the right* $\widehat{H}_n^{(2)}$ *for the* 39 *curves and* $(u,v)=(1,4)$.

Table 5.9
Biscuit example: Means of estimates (with empirical standard deviations in brackets)

	$m=2$	$m=4$	$m=6$	$m=8$	$m=10$
$\widehat{H}_n^{(1)}(1,m)$	3.61 (0.12)	3.68 (0.07)	3.66 (0.05)	3.63 (0.04)	3.60 (0.04)
$\alpha_{\text{OLS}}^{(1)}$	3.61 (0.12)	3.68 (0.07)	3.67 (0.05)	3.64 (0.04)	3.61 (0.03)
$\widehat{H}_n^{(2)}(1,m)$	2.84 (0.45)	3.69 (0.30)	3.84 (0.24)	3.85 (0.19)	3.85 (0.16)
$\alpha_{\text{OLS}}^{(2)}$	2.84 (0.45)	3.68 (0.31)	3.92 (0.23)	3.99 (0.18)	4.01 (0.14)

sure at each 2 nanometers from 1100 to 2498 nm, so 700 observation points for each biscuit. According to Brown et al. [9], the observation 23 appears as an outlier for their regression setting. We estimate r_0 for each of the left 39 curves, using the threshold $b_n=1$, which gives $\widehat{r}_0=1$ for each curve. Note also that the averaged mean quadratic variation $n^{2r-2}\overline{(D_r^{(1)}X)^2}$ equals to 0.33 when $r=2$ and 122133 when $r=3$. Refereing to the definition of $\widehat{r}_0$ in (2.5), this explosion at $r=3$ confirms us in the choice $\widehat{r}_0=3-2=1$. Now, we turn to the estimation of H and compare our estimators $H_n^{(p)}(1,m)$ together with $\alpha_{\text{OLS}}^{(\ell)}$ (where $\ell=1$ corresponds to the choice $(1,-2,1)$ for a_{jr} and $\ell=2$ to the choice $(-1,3,-3,1)$). The results are summarized in the Table 5.9 where it appears that, our estimator $\widehat{H}_n^{(2)}$ seems to be less sensitive toward high values of m. Also, compared to the $\widehat{\alpha}_{\text{OLS}}^{(p)}$ with $p=1,2$, our simplified estimators present a similar variance. To conclude this part, it should be noticed that for the 23rd curve, the choice $m=4$ gives $\widehat{H}_n^{(1)}=3.64$ and $\widehat{H}_n^{(2)}=3.55$. It appears that these two values belong to the interquartile range calculated with the 39 curves, so we can conclude that the abnormal behavior of this biscuit dough should not be linked with the curve's regularity.

5.4. Irregularly spaced points

In this last part, we focus on the unevenly spaced case in order to emphasize the superiority of divided differences estimators over finite differences ones. First,

Table 5.10

Values of mean square error and bias (between brackets) for estimators $\widehat{H}_n^{(p)}(1,4)$ built with divided (resp. finite) differences (DD) (resp. FD), for $p = 1$ or 2 and $n = 1000$ with 200 points randomly removed. Empirical value of $\mathbb{P}(\widehat{r}_0 = r_0)$

	ifBm, $\beta_0 = 0.2$			ifBm, $\beta_0 = 0.5$		
	$\widehat{H}_n^{(1)}$	$\widehat{H}_n^{(2)}$	$\widehat{\mathbb{P}}(\widehat{r}_0 = 1)$	$\widehat{H}_n^{(1)}$	$\widehat{H}_n^{(2)}$	$\widehat{\mathbb{P}}(\widehat{r}_0 = 1)$
with DD	0.0103 (−0.0749)	0.0431 (−0.1862)	1	0.0052 (−0.0394)	0.0276 (−0.1412)	1
with FD	0.3778 (−0.5978)	1.1021 (−1.0368)	0.305	2.513 (−1.5394)	3.7548 (−1.9341)	0.595

	ifBm, $\beta_0 = 0.8$			CARMA(3,1), $\beta_0 = 0.5$		
	$\widehat{H}_n^{(1)}$	$\widehat{H}_n^{(2)}$	$\widehat{\mathbb{P}}(\widehat{r}_0 = 1)$	$\widehat{H}_n^{(1)}$	$\widehat{H}_n^{(2)}$	$\widehat{\mathbb{P}}(\widehat{r}_0 = 2)$
with DD	0.0055 (−0.0238)	0.0173 (−0.1029)	1	0.0054 (−0.0326)	0.0228 (−0.1207)	1
with FD	5.548 (−2.3189)	6.753 (−2.5963)	0.703	12.1618 (−3.4588)	16.0365 (−4.0036)	0

Table 5.11

Biscuit example: means of estimates (and empirical standard deviations in brackets) in the irregularly spaced case for divided (resp. finite) differences (DD) (resp. FD)

	$m = 2$	$m = 4$	$m = 6$	$m = 8$	$m = 10$
$\widehat{H}_n^{(1)}(1,m)$ with DD	3.65 (0.12)	3.66 (0.06)	3.60 (0.06)	3.57 (0.05)	3.53 (0.05)
$\widehat{H}_n^{(1)}(1,m)$ with FD	1.67 (0.22)	1.82 (0.08)	1.91 (0.10)	1.94 (0.14)	2.01 (0.2)
$\widehat{H}_n^{(2)}(1,m)$ with DD	3.11 (0.59)	3.77 (0.36)	3.82 (0.29)	3.84 (0.22)	3.80 (0.21)
$\widehat{H}_n^{(2)}(1,m)$ with FD	1.26 (0.29)	1.53 (0.19)	1.75 (0.16)	1.98 (0.13)	2.06 (0.10)

note that from a practical point of view, the knowledge of the density ψ is not required to compute our estimations. From a theoretical point of view, the required conditions for the sampling points are given in Lemma 6.1 and are satisfied for ψ fulfilling Assumptions A2.2. In practice, these conditions mean that the sampling sequence should be rather homogeneous and, in particular, too large gaps between successive observations should be avoided (which is a quite natural natural condition in our framework). Finally, the choice of δ_n follows with $\delta_n = T_n/n$ where T_n represents the total time of observation.

To mimic the typical case of missing observations, we consider same trajectories as in Section 5.2.1 but with 200 randomly deleted points (for each) to obtain an irregular design. The Table 5.10 clearly points out that divided differences still behave well and finite differences are no more consistent (with an explosion of their mean-square error). More precisely, it is seen that the key problem is the bad estimation of r_0 in the case $r_0 \geq 1$. For large values of β_0, sample paths with a right estimation of r_0 present again an important error since for them, the joint estimation of β_0 is close to zero. The theoretical explanation for such a bad behavior can be be found in Proposition 6.1: indeed, one may show that the relation (6.11) does not hold (polynomial terms of the Taylor expansion do not vanish anymore when finite differences are combined with unevenly spaced points).

Our last illustration concerns the real data case with the biscuit dough pieces: again we delete randomly 200 observations in each curve and compute finite and divided difference estimators for these new data sets. The conclusion is the same as before: a robust behavior of divided differences (comparing with results given in the Table 5.9) and finite differences again break down, with the regularity r_0 no more consistently estimated.

6. Annexes

6.1. Results for the regular sequence design

We start with useful results for regular sequences defined in Section 2 and satisfying Assumption A2.2.

Lemma 6.1. *Under Assumption A2.2, we get for $k = 0, \ldots, n-1$, and $i = 1, \ldots, n-k$:*

$$t_{k+i} - t_k \geq C_1 i\delta_n, \quad C_1 = (T \sup_{t\in[0,T]} \psi(t))^{-1}, \tag{6.1}$$

$$t_{k+i} - t_k \leq C_2 i\delta_n, \quad C_2 = (T \inf_{t\in[0,T]} \psi(t))^{-1}, \tag{6.2}$$

and if $i = 1, \ldots, i_{\max}$ with $i_{\max}$ not depending on n:

$$t_{k+i} - t_k = \frac{i\delta_n}{T\psi(t_k)}(1 + \mathcal{O}(\delta_n^\alpha)) \tag{6.3}$$

where $\mathcal{O}(\ldots)$ is uniform over i and k.

Proof. Relations (6.1)–(6.2) are obtained from the mean-value theorem that induces, for $k = 0, \ldots, n-1$ and $i = 1, \ldots, n-k$:

$$t_{k+i} - t_k = \frac{i\delta_n}{T\psi(t_k + \theta(t_{k+i} - t_k))}, \quad 0 < \theta < 1.$$

To obtain the equivalence (6.3), one may write $t_{k+i} - t_k = \frac{i\delta_n}{T\psi(t_k)}(1 + R_n)$ with

$$|R_n| = \frac{|\psi(t_k) - \psi(t_k + \theta(t_{k+i} - t_k))|}{\psi(t_k + \theta(t_{k+i} - t_k))} \leq \frac{L\,|t_{k+i} - t_k|^\alpha}{\inf_{t\in[0,T]} \psi(t)} = \mathcal{O}(\delta_n^\alpha)$$

by Assumption A2.2 (uniformly over i, n and k, for $i = 1, \ldots, i_{\max}$). □

6.2. Results on the mean behavior of $\overline{(D_r^{(u)} X)^2}$, $r \geq 1$

We begin this part with results concerning the coefficients $(b_{ikr}^{(u)})$ defined in (2.3) that are useful for the study of $\mathbb{E}(\overline{(D_r^{(u)} X)^2})$.

Lemma 6.2. *We have under Assumption A2.2 and for $r \geq 1$, $i = 0, \ldots, r$,* *1. for $p = 0, \dots, r-1$ and convention $0^0 = 1$*

$$\sum_{i=0}^{r} (t_{k+iu} - t_k)^p b_{ikr}^{(u)} = 0, \tag{6.4}$$

2.

$$\sum_{i=0}^{r} (t_{k+iu} - t_k)^r b_{ikr}^{(u)} = 1, \tag{6.5}$$

3.

$$\left| b_{ikr}^{(u)} \right| \le \frac{C_1^{-r} u^{-r} \delta_n^{-r},}{\prod_{m=0, m\neq i}^{r} |i-m|}, \tag{6.6}$$

with C_1 given by (6.1),

4.

$$b_{ikr}^{(u)} = \frac{u^{-r}\psi^r(t_k) T^r \delta_n^{-r}}{\prod_{m=0, m\neq i}^{r} (i-m)} (1 + \mathcal{O}(\delta_n^{\alpha})) \tag{6.7}$$

with $\mathcal{O}(\dots)$ uniform over i and k.

Proof. The term $g[t_k, \dots, t_{k+ru}] = \sum_{i=0}^{r} b_{ikr}^{(u)} g(t_{k+iu})$ is the leading coefficient in the polynomial approximation of degree r of g, given in the decomposition (2.2). Considering the polynomial $g(t) = (t - t_k)^p$, we may immediately deduce the properties (6.4)–(6.5), from uniqueness of the relation (2.2). Next, (6.6)–(6.7) are direct consequences of Lemma 6.1 and the definition of $b_{ikr}^{(u)}$. □

Next, the construction of our estimator $\widehat{r}_0$ is derived from results of the following proposition.

Proposition 6.1. *Under Assumption A2.1 and A2.2, one obtains:*

(i) for $r = r_0 + p$ with $p = 1, 2$:

$$n^{-2(p-\beta_0)}\, \mathbb{E}\Big(\overline{\big(D_{r_0+p}^{(u)} X\big)^2} \Big) \xrightarrow[n\to\infty]{} u^{-2(p-\beta_0)} \ell(p, r_0, \beta_0)$$

where

$$\ell(p, 0, \beta_0) = -\frac{1}{2} \int_0^T d_0(t) \frac{\psi^{2p+1}(t)}{\psi^{2\beta_0}(t)}\, \mathrm{d}t \sum_{i,j=0}^{p} \frac{|i-j|^{2\beta_0}}{\prod_{\substack{m=0\\ m\neq i}}^{p} (i-m) \prod_{\substack{q=0\\ q\neq j}}^{p} (j-q)}$$

while if $r_0 \ge 1$,

$$\ell(p, r_0, \beta_0) = \frac{(-1)^{r_0+1} \int_0^T d_0(t) \frac{\psi^{2p+1}(t)}{\psi^{2\beta_0}(t)}\, \mathrm{d}t}{2(2\beta_0 + 2r_0) \cdots (2\beta_0 + 1)} \sum_{i,j=0}^{r_0+p} \frac{|i-j|^{2(r_0+\beta_0)}}{\prod_{\substack{m=0\\ m\neq i}}^{r_0+p} (i-m) \prod_{\substack{q=0\\ q\neq j}}^{r_0+p} (j-q)} \tag{6.8}$$

(ii) for $r_0 \ge 1$ and $r = 1, \dots, r_0$:

$$\mathbb{E}\Big(\overline{\big(D_r^{(u)} X\big)^2} \Big) \xrightarrow[n\to\infty]{} \frac{1}{(r!)^2} \int_0^T \mathbb{E}\big(X^{(r)}(t)\big)^2 \psi(t)\, \mathrm{d}t.$$

Proof. **A.** Let us begin with general expressions of $\mathbb{E}(D_{r,k}^{(u)}X D_{r,\ell}^{(u)}X)$ useful for the sequel. First for $\mathbb{L}^{(p,p)}(s,t) = \mathbb{E}(X^{(p)}(s)X^{(p)}(t))$ $(p \geq 0)$, the relation (2.1) is equivalent to

$$\lim_{h\to 0} \sup_{\substack{s,t\in[0,T]\\ |s-t|\leq h, s\neq t}} \left| \frac{\mathbb{L}^{(r_0,r_0)}(s,s) + \mathbb{L}^{(r_0,r_0)}(t,t) - 2\mathbb{L}^{(r_0,r_0)}(s,t)}{|s-t|^{2\beta_0}} - d_0(t) \right| = 0. \quad (6.9)$$

For $(v,w) \in [0,1]^2$, we set $\dot{v}_{ik} = t_k + (t_{k+iu} - t_k)v$ and $\dot{w}_{j\ell} = t_\ell + (t_{\ell+ju} - t_\ell)w$. Next, from the definition of $D_{r,k}^{(u)}X$ given in (2.4), we get

$$\mathbb{E}(D_{r,k}^{(u)}X D_{r,\ell}^{(u)}X) = \sum_{i,j=0}^{r} b_{ikr}^{(u)} b_{j\ell r}^{(u)} \mathbb{L}^{(0,0)}(t_{k+iu}, t_{\ell+ju}).$$

For $r_0 = 0$ and since $\sum_{i=0}^{r} b_{ikr}^{(u)} = 0$, we have:

$$\mathbb{E}(D_{r,k}^{(u)}X D_{r,\ell}^{(u)}X) = \sum_{i,j=0}^{r} b_{ikr}^{(u)} b_{j\ell r}^{(u)} \Big\{ \mathbb{L}^{(0,0)}(t_{k+iu}, t_{\ell+ju}) - \frac{1}{2}\mathbb{L}^{(0,0)}(t_{k+iu}, t_{k+iu}) - \frac{1}{2}\mathbb{L}^{(0,0)}(t_{\ell+ju}, t_{\ell+ju}) \Big\}. \quad (6.10)$$

If $r_0 \geq 1$, we apply multiple Taylor series expansions with integral remainder. Next, the properties $\sum_{i=0}^{r} b_{ikr}^{(u)}(t_{k+iu} - t_k)^p = 0$ for $p = 0, \ldots, r-1$ (and convention $0^0 = 1$) induce:

$$\mathbb{E}(D_{r,k}^{(u)}X D_{r,\ell}^{(u)}X) = \sum_{i,j=0}^{r} b_{ikr}^{(u)} b_{j\ell r}^{(u)} (t_{k+iu} - t_k)^{r^*} (t_{\ell+ju} - t_\ell)^{r^*} \times \iint_{[0,1]^2} \frac{(1-v)^{r^*-1}(1-w)^{r^*-1}}{((r^*-1)!)^2} \mathbb{L}^{(r^*,r^*)}(\dot{v}_{ik}, \dot{w}_{j\ell})\, \mathrm{d}v\mathrm{d}w \quad (6.11)$$

where we have set $r^* = \min(r_0, r) \geq 1$.

B. From expressions (6.10)–(6.11), we are in a position to derive the asymptotic behavior of $\mathbb{E}((\overline{D_r^{(u)}X})^2)$.

Case $r_0 \geq 1$***,*** $r = r_0 + p$***,*** $p = 1$ ***or*** $p = 2$***.*** In this case, $r^* = r_0 \leq r - 1$. From (6.11) and the property $\sum_{i=0}^{r} b_{ikr}^{(u)}(t_{k+iu} - t_k)^{r_0} = 0$, we may write

$$\mathbb{E}(D_{r,k}^{(u)}X)^2 = \sum_{i,j=0}^{r} b_{ikr}^{(u)} b_{jkr}^{(u)} (t_{k+iu} - t_k)^{r_0} (t_{k+ju} - t_k)^{r_0} \iint_{[0,1]^2} \frac{((1-v)(1-w))^{r_0-1}}{((r_0-1)!)^2} \times \Big\{ \mathbb{L}^{(r_0,r_0)}(\dot{v}_{ik}, \dot{w}_{jk}) - \frac{1}{2}\mathbb{L}^{(r_0,r_0)}(\dot{v}_{ik}, \dot{v}_{ik}) - \frac{1}{2}\mathbb{L}^{(r_0,r_0)}(\dot{w}_{jk}, \dot{w}_{jk}) \Big\}\, \mathrm{d}v\mathrm{d}w \quad (6.12)$$

Using the locally stationary condition (6.9), uniform continuity of $d_0(\cdot)$ on $[0,T]$ and the bound: $|\dot{v}_{ik} - \dot{w}_{jk}| \le t_{k+ru} - t_k \le C_1 r u \delta_n$ for $i = 0,\dots,r$ and $j = 0,\dots,r$, we may show that the predominant term for $\mathbb{E}(\overline{(D_r^{(u)} X)^2})$ is given by:

$$\frac{-1}{2(n_r+1)} \sum_{k=0}^{n_r} \sum_{i,j=0}^{r} b_{ikr}^{(u)} b_{jkr}^{(u)} (t_{k+iu} - t_k)^{r_0} (t_{k+ju} - t_k)^{r_0} \\ \times \iint_{[0,1]^2} \frac{((1-v)(1-w))^{r_0-1}}{((r_0-1)!)^2} |\dot{v}_{ik} - \dot{w}_{jk}|^{2\beta_0} \, d_0(t_k) \, \mathrm{d}v \, \mathrm{d}w \quad (6.13)$$

where we have set $n_r = n - ur$. From the equivalents (6.3) and (6.7), we can write the leading term of (6.13) as a Riemann sum on t_k to obtain

$$\delta_n^{2p-2\beta_0} \mathbb{E}\big(\overline{(D_r^{(u)} X)^2}\big) \xrightarrow[n\to\infty]{} -\frac{1}{2}\Big(\frac{u}{T}\Big)^{-2p+2\beta_0} \int_0^T d_0(t) \psi^{2p+1-2\beta_0}(t) \, \mathrm{d}t \\ \times \sum_{i,j=0}^{r} \frac{(ij)^{r_0}}{\prod\limits_{\substack{m=0\\ m\ne i}}^{r} (i-m) \prod\limits_{\substack{q=0\\ q\ne j}}^{r} (j-q)} \iint_{[0,1]^2} \frac{((1-v)(1-w))^{r_0-1}}{((r_0-1)!)^2} |iv - jw|^{2\beta_0} \, \mathrm{d}v \, \mathrm{d}w.$$

Next by performing elementary but tedious multiple integrations by parts, we arrive at the simpler form of $\ell(r, r_0, \beta_0)$ given in (6.8), for $n\delta_n \to T$.

Case $r_0 = 0$**,** $r = r_0 + 1,, r_0 + 2$**.** The proof is the same but starting from (6.10) and $\ell = k$.

Case $r_0 \ge 1$ ***and*** $r = 1,\dots,r_0$**.** In this case, $r^* = r$ and from the relation (6.11), one gets

$$\mathbb{E}(D_{r,k}^{(u)} X)^2 = \sum_{i,j=0}^{r} b_{ikr}^{(u)} b_{jkr}^{(u)} (t_{k+iu} - t_k)^{r} (t_{k+ju} - t_k)^{r} \\ \times \iint_{[0,1]^2} \frac{((1-v)(1-w))^{r-1}}{((r-1)!)^2} \mathbb{L}^{(r,r)}(\dot{v}_{ik}, \dot{w}_{jk}) \, \mathrm{d}v\mathrm{d}w.$$

The result follows after Riemann summation with the help of uniform continuity of $\mathbb{L}^{(r,r)}(\cdot,\cdot)$, $r = 1,\dots,r_0$ and properties (6.3), (6.5). □

6.3. A general exponential bound for $|\overline{(D_r^{(u)} X)^2} - \mathbb{E}(\overline{(D_r^{(u)} X)^2})|$

We start with some auxiliary results. The following lemma gives results on the asymptotic behavior of $\mathbb{C}_r(k,\ell) = \mathrm{Cov}\,(D_{r,k}^{(u)} X, D_{r,\ell}^{(u)} X)$ and $\mathbb{C}_r^2(k,\ell)$ with $n_r = n - ur$ and u a positive integer.

Lemma 6.3. *Suppose that Assumption A2.1 and A2.2 are satisfied.*

(i) Under the condition A2.1-(iii-1) and for $r = r_0 + p$, $p = 1$ or $p = 2$, one obtains

$$\max_{k=0,\dots,n_r} \sum_{\ell=0}^{n_r} |\mathbb{C}_r(k,\ell)| = \begin{cases} \mathcal{O}(n^{2p-2\beta_0}) & \text{if } 0 < \beta_0 < \frac{1}{2}, \\ \mathcal{O}(n^{2p-1}\ln n) & \text{if } \beta_0 = \frac{1}{2}, \\ \mathcal{O}(n^{2p-1}) & \text{if } \frac{1}{2} < \beta_0 < 1 \ ; \end{cases}$$

and

$$\sum_{k=0}^{n_r} \sum_{\ell=0}^{n_r} \mathbb{C}_r^2(k,\ell) = \begin{cases} \mathcal{O}(n^{4p-4\beta_0+1}) & \text{if } 0 < \beta_0 < \frac{3}{4}, \\ \mathcal{O}(n^{4p-2}\ln n) & \text{if } \beta_0 = \frac{3}{4}, \\ \mathcal{O}(n^{4p-2}) & \text{if } \frac{3}{4} < \beta_0 < 1. \end{cases}$$

(ii) Under the condition A2.1-(iii-2) and for $r = r_0 + 2$, one obtains

$$\max_{k=0,\dots,n_r} \sum_{\ell=0}^{n_r} |\mathbb{C}_r(k,\ell)| = \mathcal{O}(n^{4-2\beta_0}) \text{ and } \sum_{k=0}^{n_r} \sum_{\ell=0}^{n_r} \mathbb{C}_r^2(k,\ell) = \mathcal{O}(n^{9-4\beta_0}).$$

(iii) If $r = 1, \dots, r_0$ (with $r_0 \geq 1$), then $\max_{k=0,\dots,n_r} \sum_{\ell=0}^{n_r} |\mathbb{C}_r(k,\ell)| = \mathcal{O}(n)$ and $\sum_{k=0}^{n_r} \sum_{\ell=0}^{n_r} \mathbb{C}_r^2(k,\ell) = \mathcal{O}(n^2)$.

Proof. (i) Setting $\mu(t) = \mathbb{E}\big(X(t)\big)$, μ is r_0-times differentiable and similarly to (6.10)–(6.11), we get the expansion

$$\mathbb{C}_r(k,\ell) = \sum_{i,j=0}^{r} b_{ikr}^{(u)} b_{j\ell r}^{(u)} (t_{k+iu} - t_k)^{r_0} (t_{\ell+ju} - t_\ell)^{r_0} \times \iint_{[0,1]^2} \frac{((1-v)(1-w))^{r_0-1}}{((r_0-1)!)^2} \mathbb{K}^{(r_0,r_0)}(\dot v_{ik}, \dot w_{j\ell}) \,\mathrm{d}v \,\mathrm{d}w$$

for $r_0 \geq 1$ while if $r_0 = 0$, $\mathbb{C}_r(k,\ell) = \sum_{i=0}^{r} \sum_{j=0}^{r} b_{ikr}^{(u)} b_{j\ell r}^{(u)} \mathbb{K}(t_{k+iu}, t_{\ell+ju})$.

***Case** $r = r_0 + 1$ **or** $r_0 + 2$.* For $r_0 \geq 1$, we have the bound:

$$\max_{k=0,\dots,n_r} \sum_{\ell=0}^{n_r} |\mathbb{C}_r(k,\ell)| \leq U_{1n} + U_{2n} + U_{3n}$$

with

$$U_{1n} = \max_{k=ur+1,\dots,n_r} \sum_{\ell=0}^{k-ur-1} |\mathbb{C}_r(k,\ell)|, \qquad U_{2n} = \max_{k=0,\dots,n-2ur-1} \sum_{\ell=k+ur+1}^{n_r} |\mathbb{C}_r(k,\ell)|$$

and $\quad U_{3n} = \max_{k=0,\dots,n_r} \sum_{\ell=\max(0,k-ur)}^{\min(n_r,k+ur)} |\mathbb{C}_r(k,\ell)|.$

First, consider the sum $U_{1n}+U_{2n}$ where $|k-\ell|\geq ur+1$. Since $\sum_{i=0}^{r} b_{ikr}^{(u)}(t_{k+iu}-t_k)^{r_0}=0$ for $r=r_0+1$ or $r=r_0+2$, and $[t_k,\dot v_{ik}]$ is distinct from $[t_\ell,\dot w_{j\ell}]$, we get

$$\mathbb{C}_r(k,\ell)=\sum_{i,j=0}^{r} b_{ikr}^{(u)}b_{jkr}^{(u)}(t_{k+iu}-t_k)^{r_0}(t_{\ell+ju}-t_\ell)^{r_0}$$
$$\times\iint_{[0,1]^2}\frac{((1-v)(1-w))^{r_0-1}}{((r_0-1)!)^2}\int_{t_k}^{\dot v_{ik}}\int_{t_\ell}^{\dot w_{j\ell}}\mathbb{K}^{(r_0+1,r_0+1)}(s,t)\,\mathrm{d}s\mathrm{d}t\mathrm{d}v\mathrm{d}w. \quad (6.14)$$

Condition A2.1-(iii-1), together with (6.2) and (6.6), gives a bound of $\mathcal{O}(n^{2p-2\beta_0}\sum_{i=1}^{n} i^{-2(1-\beta_0)})$ for $|U_{1n}+U_{2n}|$, which is of order $n^{2(p-\beta_0)}$ if $0<\beta_0<\frac{1}{2}$, $n^{2(p-\beta_0)}\ln n$ if $\beta_0=\frac{1}{2}$ and n^{2p-1} if $\beta_0>\frac{1}{2}$. Next, for U_{3n} where $|k-\ell|\leq ur$, we obtain that $U_{3n}=\mathcal{O}(n^{2(p-\beta_0)})$ in a similar way as in the proof of Proposition 6.1, and with the help of Cauchy-Schwarz inequality to control the terms depending on $\mu^{(r_0)}(t)$.

We proceed similarly for the case $r_0=0$, starting from the definition of $\mathbb{C}_r(k,\ell)$ as well as for the study of $\sum_{k=0}^{n_r}\sum_{\ell=0}^{n_r}\mathbb{C}_r^2(k,\ell)$ for which the dominant terms are of order $\mathcal{O}(n^{1+4p-4\beta_0}\sum_{i=1}^{n} i^{-4(1-\beta_0)})$.

(ii) The condition A2.1-(iii-2) and $r=r_0+2$ allow to transform (6.14) into

$$\mathbb{C}_r(k,\ell)=\sum_{i,j=0}^{r} b_{ikr}^{(u)}b_{jkr}^{(u)}(t_{k+iu}-t_k)^{r_0}(t_{\ell+ju}-t_\ell)^{r_0}\iint_{[0,1]^2}\frac{((1-v)(1-w))^{r_0-1}}{((r_0-1)!)^2}$$
$$\times\int_{t_k}^{\dot v_{ik}}\int_{t_\ell}^{\dot w_{j\ell}}\int_{t_k}^{t}\int_{t_\ell}^{s}K^{(r_0+2,r_0+2)}(y,z)\,\mathrm{d}y\,\mathrm{d}z\mathrm{d}s\,\mathrm{d}t\,\mathrm{d}v\,\mathrm{d}w \quad (6.15)$$

which gives that

$$\max_{k=0,\dots,n_r}\sum_{\ell=0}^{n_r}|\mathbb{C}_r(k,\ell)|=\mathcal{O}(n^{2(2-\beta_0)}\sum_{i=1}^{n} i^{-4+2\beta_0})=\mathcal{O}(n^{2(2-\beta_0)})$$

for all $\beta_0\in]0,1[$ and $r_0\geq 1$. From (6.15), we also get that $\sum_{k=0}^{n_r}\sum_{\ell=0}^{n_r}\mathbb{C}_r^2(k,\ell)=\mathcal{O}(n^{9-4\beta_0}\sum_{i=1}^{n} i^{-8+4\beta_0})=\mathcal{O}(n^{9-4\beta_0})$ for all $\beta_0\in]0,1[$.

(iii) Results of this part, where $r_0\geq 1$, are consequences of

$$\mathbb{C}_r(k,\ell)=\sum_{i,j=0}^{r} b_{ikr}^{(u)}b_{j\ell r}^{(u)}(t_{k+iu}-t_k)^{r}(t_{\ell+ju}-t_\ell)^{r}$$
$$\times\iint_{[0,1]^2}\frac{((1-v)(1-w))^{r-1}}{(r-1)!^2}K^{(r,r)}(\dot v_{ik},\dot w_{jl})\,\mathrm{d}v\,\mathrm{d}w=\mathcal{O}(1)$$

with uniform continuity of $K^{(r,r)}(\cdot,\cdot)$ for $r=1,\dots,r_0$ together with the bounds given in (6.2) and (6.6). □

The next proposition gives the general exponential bound which is involved in all of our main results.

Proposition 6.2. *Suppose that Assumption A2.1 and A2.2 are satisfied. Let $\eta_n(r)$ be some given positive sequence and $u \in \mathbb{N}^*$, then*

$$\mathbb{P}\Big(\Big|\overline{(D_r^{(u)}X)^2} - \mathbb{E}\Big(\overline{(D_r^{(u)}X)^2}\Big)\Big| \geq \eta_n(r)\Big)$$

is of order:

$$\mathcal{O}\bigg(\exp\Big(-C(r)n\eta_n(r) \times \min\Big(\big(\max_{0\leq k\leq n_r}\sum_{\ell=0}^{n_r}|\,\mathbb{C}_r(k,\ell)|\,\big)^{-1}, \frac{n\eta_n(r)}{\sum\limits_{k,\ell=0}^{n_r}\mathbb{C}_r^2(k,\ell)}\Big)\Big)\bigg) + \mathcal{O}\bigg(\frac{v_n^{1/2}(r)}{n\eta_n(r)}\exp\Big(-C(r)\frac{n^2\eta_n^2(r)}{v_n(r)}\Big)\bigg)$$

for some positive constant $C(r)$, not depending on $\eta_n(r)$ and

$$v_n(r) := n \max_{k=0,\ldots,n_r}\left(\mathbb{E}(D_{r,k}^{(u)}X)\right)^2 \max_{k=0,\ldots,n_r}\sum_{\ell=0}^{n_r}|\mathbb{C}_r(k,\ell)|\,. \qquad (6.16)$$

Proof. For all $r \geq 1$, we may bound $\mathbb{P}(|\overline{(D_r^{(u)}X)^2} - \mathbb{E}(\overline{(D_r^{(u)}X)^2})| \geq \eta_n(r)\big)$ by $S_1 + S_2$ with

$$S_1 = \mathbb{P}\Big(\Big|\sum_{k=0}^{n_r}(D_{r,k}^{(u)}X - \mathbb{E}(D_{r,k}^{(u)}X))^2 - \mathrm{Var}\,(D_{r,k}^{(u)}X)\Big| > \frac{(n_r+1)\eta_n(r)}{2}\Big)$$

and $S_2 = \mathbb{P}(|\sum_{k=0}^{n_r}(\mathbb{E}(D_{r,k}^{(u)}X))(D_{r,k}^{(u)}X - \mathbb{E}(D_{r,k}^{(u)}X))| > \frac{(n_r+1)\eta_n(r)}{4})$. First, let $\{Y_i\}_{i=1,\ldots,d_n}$ be an orthonormal basis for the linear span of $\{D_{r,k}^{(u)}X\}_{k=0,\ldots,n_r}$ (so that Y_i are i.i.d. with density $\mathcal{N}(0,1)$). We can write $D_{r,k}^{(u)}X - \mathbb{E}(D_{r,k}^{(u)}X) = \sum_{i=1}^{d_n} d_{k,i}Y_i$ with $d_{k,i} = \mathrm{Cov}\,(D_{r,k}^{(u)}X, Y_i)$. Next, if $Y = (Y_1, \ldots, Y_{d_n})^\top$, we obtain

$$\sum_{k=0}^{n_r}(D_{r,k}^{(u)}X - \mathbb{E}D_{r,k}^{(u)}X)^2 = \sum_{i,j=1}^{d_n} c_{i,j}Y_iY_j = Y^\top CY$$

and $\sum_{k=0}^{n_r}\mathrm{Var}\,(D_{r,k}^{(u)}X) = \sum_{i=1}^{d_n} c_{i,i}$ with $c_{i,j} = \sum_{k=0}^{n_r} d_{ki}d_{kj}$. Next, for $C = \big(c_{i,j}\big)_{\substack{i=1,\ldots,d_n\\ j=1,\ldots,d_n}}$ and $D = \big(d_{k,j}\big)_{\substack{k=0,\ldots,n_r\\ j=1,\ldots,d_n}}$, one gets $C = D^\top D$ where C is a real, symmetric and positive semidefinite matrix. There exists an orthogonal matrix P such that $\mathrm{diag}(\lambda_1, \ldots, \lambda_{d_n}) = P^\top CP$, for λ_i eigenvalues of C. Then we can transform the quadratic form as:

$$\sum_{k=0}^{n_r}(D_{r,k}^{(u)}X - \mathbb{E}(D_{r,k}^{(u)}X))^2 = \sum_{i=1}^{d_n}\lambda_i(P^\top Y)_i^2$$

where $(P^\top Y)_i$ denotes the i-th component of the $(d_n \times 1)$ vector $P^\top Y$. Since $\sum_{i=1}^{d_n} c_{i,i} = \sum_{i=1}^{d_n} \lambda_i$, we arrive at

$$S_1 = \mathbb{P}\Big(\Big| \sum_{i=1}^{d_n} \lambda_i \big((P^\top Y)_i^2 - 1\big) \Big| \geq \frac{(n_r+1)\eta_n(r)}{2} \Big).$$

Now, with the exponential bound of Hanson and Wright [20], we obtain for some generic constant c:

$$S_1 \leq 2\exp\Big(-c(n_r+1)\eta_n(r) \times \min\Big(\frac{1}{\max(\lambda_i)}, \frac{(n_r+1)\eta_n(r)}{\sum \lambda_i^2} \Big)\Big).$$

Next, since $D^\top D$ and $DD^\top$ have the same non zero eigenvalues,

$$\max_{i=1,\ldots,d_n} \lambda_i \leq \max_{0\leq k\leq n_r} \sum_{\ell=0}^{n_r} |\,\mathbb{C}_r(k,\ell)|$$

and

$$\sum_{i=1}^{d_n} \lambda_i^2 = \sum_{i=1}^{d_n}\sum_{j=1}^{d_n} c_{ij}c_{ji} = \sum_{k=0}^{n_r}\sum_{\ell=0}^{n_r} \Big(\sum_{i=1}^{d_n} d_{ki} d_{li} \Big)^2 = \sum_{k=0}^{n_r}\sum_{\ell=0}^{n_r} \mathbb{C}_r^2(k,\ell).$$

Finally S_1 is bounded by

$$2\exp\Big(-c(n_r+1)\eta_n(r) \times \min\Big(\big(\max_{0\leq k\leq n_r} \sum_{\ell=0}^{n_r} |\,\mathbb{C}_r(k,\ell)| \big)^{-1}, \frac{(n_r+1)\eta_n(r)}{\sum_{k=0}^{n_r}\sum_{\ell=0}^{n_r} \mathbb{C}_r^2(k,\ell)} \Big)\Big).$$

For S_2, we use the classical exponential bound on a Gaussian variable: $Y \sim \mathcal{N}(0,\sigma^2)$ implies that $\mathbb{P}(|Y| \geq \varepsilon) \leq \min(1, \sqrt{\frac{2\sigma^2}{\pi\varepsilon^2}}) \exp(-\frac{\varepsilon^2}{2\sigma^2})$, $\varepsilon > 0$. Here $Y = \sum_{k=0}^{n_r} (\mathbb{E}D_{r,k}^{(u)}X)(D_{r,k}^{(u)}X - \mathbb{E}D_{r,k}^{(u)}X)$ and we easily get that $\mathrm{Var}\,(Y) \leq v_n(r)$. □

6.4. Proofs of the main results of the Section 3

Proof of Theorem 3.1. Recall that $\widehat{r}_0$ is given by: $\widehat{r}_0 = \min\{r \in \underline{\{2,\ldots,m_n\}}$: $B_n(r)$ holds$\} - 2$ where the event $B_n(r)$ is defined by $B_n(r) = \{(D_r^{(1)}X)^2 \geq n^2 b_n\}$, and $\widehat{r}_0 = l_0$ if $\cap_{r=2}^{m_n} B_n^c(r)$ holds. The condition $m_n \to \infty$ guarantees that for n large enough, $r_0 + 2 \in \{2,\ldots,m_n\}$. From this definition, we write

$$\mathbb{E}\big(\widehat{r}_0 - r_0\big)^2 = \sum_{r=0}^{m_n-2} (r-r_0)^2 \mathbb{P}\big(\widehat{r}_0 = r\big) + (l_0 - r_0)^2 \mathbb{P}\big(\widehat{r}_0 = l_0\big)$$

where $\mathbb{P}\big(\widehat{r}_0 = 0\big) = \mathbb{P}\big(B_n(2)\big)$, $\mathbb{P}\big(\widehat{r}_0 = r\big) = \mathbb{P}\big(B_n^c(2)\cap\cdots\cap B_n^c(r+1)\cap B_n(r+2)\big)$ if $r = 1,\ldots,m_n-2$, and $\mathbb{P}\big(\widehat{r}_0 = l_0\big) \leq \mathbb{P}\big(B_n^c(r_0+2)\big)$. Then, for all $r_0 \in \mathbb{N}_0$:

$\mathbb{E}\big(\widehat{r}_0 - r_0\big)^2 = \mathcal{O}\big(T_{1n}(r_0)\big) + \mathcal{O}\big(m_n^3 T_{2n}(r_0)\big)$ where we have set $T_{1n}(0) = 0$, $T_{1n}(r_0) = \sum_{r=2}^{r_0+1} \mathbb{P}\big(B_n(r)\big)$ (for $r_0 \geq 1$) and $T_{2n}(r_0) = \mathbb{P}\big(B_n^c(r_0+2)\big)$. Now, the study of T_{1n} and T_{2n} is derived from results of Lemma 6.1, Lemma 6.2, Proposition 6.1 and Lemma 6.3. In particular, since $\mu \in C^{r_0+1}([0,T])$ we get:

$$\mathbb{E}(D_{r,k}^{(u)}X) = \sum_{i=0}^{r} b_{ikr}(u)(t_{k+iu} - t_k)^{r^*} \int_0^1 \frac{(1-v)^{r^*-1}}{(r^*-1)!} \mu^{(r^*)}(t_k + (t_{k+iu} - t_k)v)\, \mathrm{d}v$$

which is $\mathcal{O}(n^{r-r^*})$ for $r^* = \min(r, r_0+1)$ implying that $\mathbb{E}(D_{r,k}^{(u)}X) = \mathcal{O}(1)$ for $r = 1, \ldots, r_0+1$, and $\mathbb{E}(D_{r,k}^{(u)}X) = \mathcal{O}(n)$ for $r = r_0+2$. Then one may bound $v_n(r)$ given in equation (6.16) by $\mathcal{O}(n^2)$ if $r = 1, \ldots, r_0$, $\mathcal{O}\big(n^{3-2\beta_0}\mathbb{1}_{]0,\frac{1}{2}[}(\beta_0) + n^2 \ln n \mathbb{1}_{\{\frac{1}{2}\}}(\beta_0) + n^2 \mathbb{1}_{]\frac{1}{2},1[}(\beta_0)\big)$ if $r = r_0+1$ with A2.1-(iii-1), $\mathcal{O}\big(n^{7-2\beta_0}\mathbb{1}_{]0,\frac{1}{2}[}(\beta_0) + n^6 \ln n \mathbb{1}_{\{\frac{1}{2}\}}(\beta_0) + n^6 \mathbb{1}_{]\frac{1}{2},1[}(\beta_0)\big)$ if $r = r_0+2$ with A2.1-(iii-1), and $\mathcal{O}(n^{7-2\beta_0})$ if $r = r_0+2$ and A2.1-(iii-2) holds. After some calculations based on the properties $n^{2\beta_0} b_n \to \infty$ and $n^{-2(1-\beta_0)} b_n \to 0$, one may derive from Proposition 6.2 that:

$$T_{1n}(r_0) = \mathcal{O}\Big(\exp\Big(-D(r_0) b_n \big(n^{2\beta_0+1}\mathbb{1}_{]0,\frac{1}{2}[}(\beta_0) + \big(\frac{n^2}{\ln n}\big)\mathbb{1}_{\{\frac{1}{2}\}}(\beta_0) + n^2 \mathbb{1}_{]\frac{1}{2},1[}(\beta_0)\big)\Big)\Big).$$

Next, if A2.1-(iii-1) holds

$$T_{2n}(r_0) = O\Big(\exp\Big(-D(r_0)\big(n\mathbb{1}_{]0,\frac{1}{2}[}(\beta_0) + \big(\frac{n}{\ln n}\big)\mathbb{1}_{\{\frac{1}{2}\}}(\beta_0) + n^{2(1-\beta_0)}\mathbb{1}_{]\frac{1}{2},1[}(\beta_0)\big)\Big)\Big)$$

while, under A2.1-(iii-2) and for all $\beta_0 \in]0,1[$, $T_{2n}(r_0) = O\big(\exp(-D(r_0)n)\big)$. For $p = 1, 2$, we get that $T_{1n}(r_0) = o\big(T_{2n}(r_0)\big)$ and the mean square error follows. Finally, to obtain a bound for $\mathbb{P}(\widehat{r}_0 \neq r_0)$, it suffices to notice that $\{\widehat{r}_0 = 0\} = B_n(2)$ for $r_0 = 0$ and $\{\widehat{r}_0 = r_0\} = B_n^c(2) \cap \cdots \cap B_n^c(r_0+1) \cap B_n(r_0+2)$ for $r_0 \geq 1$, therefore $\mathbb{P}(\widehat{r}_0 \neq r_0) = T_{1n}(r_0) + T_{2n}(r_0) = T_{2n}(r_0)(1 + o(1))$. □

Proof of Theorem 3.2. We start the proof, with either $p = 1$ or $p = 2$, and thus denote by $\widehat{r}_p$ (resp. r_p) the quantity $\widehat{r}_0 + p$ (resp. $r_0 + p$). We set

$$l_n(p, r_0, \beta_0) = -\frac{1}{2n} \sum_{k=0}^{n} d_0(t_k) \psi^{2(p-\beta_0)}(t_k) \sum_{i,j=0}^{r_p} \frac{(ij)^{r_0}}{\prod_{\substack{m=0\\ m\neq i}}^{r_p}(i-m) \prod_{\substack{q=0\\ q\neq j}}^{r_p}(j-q)} \times \iint_{[0,1]^2} \frac{((1-v)(1-w))^{r_0-1}}{(r_0-1)!^2} |iv - jw|^{2\beta_0}\, \mathrm{d}v\mathrm{d}w, \quad (6.17)$$

for all $r_0 \geq 1$ while if $r_0 = 0$,

$$l_n(p, 0, \beta_0) = -\frac{1}{2n} \sum_{k=0}^{n} d_0(t_k) \psi^{2(p-\beta_0)}(t_k) \sum_{i,j=0}^{r_p} \frac{|i-j|^{2\beta_0}}{\prod_{\substack{m=0\\ m\neq i}}^{p}(i-m) \prod_{\substack{q=0\\ q\neq j}}^{p}(j-q)}. \quad (6.18)$$

We study the convergence of $\widehat{\alpha}_p = 2(\widehat{\beta}_n^{(p)} - p)$ toward $\alpha_p = 2(\beta_0 - p)$, so that

$$\widehat{\alpha}_p = \frac{\ln\big(\overline{(D_{\widehat{r}_p}^{(u)}X)^2}\big) - \ln\big(\overline{(D_{\widehat{r}_p}^{(v)}X)^2}\big)}{\ln(u/v)}.$$

We consider the following decomposition of $\ln(u/v)\widehat{\alpha}_p$:

$$\ln\Big(\frac{n^{\alpha_p}}{n-u\widehat{r}_p+1}\sum_{k=0}^{n-u\widehat{r}_p}\big(D_{\widehat{r}_p,k}^{(u)}X\big)^2 - u^{\alpha_p}l_n(p,r_0,\beta_0) + u^{\alpha_p}l_n(p,r_0,\beta_0)\Big)$$
$$- \ln\Big(\frac{n^{\alpha_p}}{n-v\widehat{r}_p+1}\sum_{k=0}^{n-v\widehat{r}_p}\big(D_{\widehat{r}_p,k}^{(v)}X\big)^2 - v^{\alpha_p}l_n(p,r_0,\beta_0) + v^{\alpha_p}l_n(p,r_0,\beta_0)\Big)$$

Hence $\ln(u/v)(\widehat{\alpha}_p - \alpha_p) = F_n(u) - F_n(v) + o(F_n(u) + F_n(v))$ where $o(\cdot) \xrightarrow[n\to\infty]{a.s.} 0$ as soon as $F_n(\cdot) \xrightarrow[n\to\infty]{a.s.} 0$ with

$$F_n(u) = \frac{n^{\alpha_p}\overline{\big(D_{\widehat{r}_p}^{(u)}X\big)^2} - u^{\alpha_p}l_n(p,r_0,\beta_0)}{u^{\alpha_p}l_n(p,r_0,\beta_0)} = \frac{F_{1,n,p}(u) + F_{2,n,p}(u) + F_{3,n,p}(u)}{u^{\alpha_p}l_n(p,r_0,\beta_0)}$$

for $F_{1,n,p}(u) = n^{\alpha_p}\big(\overline{(D_{\widehat{r}_p}^{(u)}X)^2} - \overline{(D_{r_p}^{(u)}X)^2}\big)$, $F_{2,n,p}(u) = n^{\alpha_p}\big(\overline{(D_{r_p}^{(u)}X)^2} - \mathbb{E}\overline{(D_{r_p}^{(u)}X)^2}\big)$ and $F_{3,n,p}(u) = n^{\alpha_p}\mathbb{E}\big(\overline{(D_{r_p}^{(u)}X)^2}\big) - u^{\alpha_p}\,l_n(p,r_0,\beta_0)$.

(i) *Study of* $F_{1,n,p}(u)$**.** From Theorem 3.1, we get that $\sum_n \mathbb{P}(\widehat{r}_0 \neq r_0) < \infty$, so, a.s. for n large enough, $\widehat{r}_0 = r_0$ and $F_{1,n,p}(u) \equiv 0$, $p = 1$ or $p = 2$.

(ii) *Study of* $F_{2,n,p}(u)$**.** We study

$$\mathbb{P}\Big(\Big|\overline{\big(D_{r_p}^{(u)}X\big)^2} - \mathbb{E}\overline{\big(D_{r_p}^{(u)}X\big)^2}\Big| > c_p n^{2(p-\beta_0)}\psi_{np}^{-1}(\beta_0)\Big)$$

for c_p a positive constant, $\psi_{n2}(\beta_0) \equiv \big(\frac{n}{\ln n}\big)^{\frac{1}{2}}$ and

$$\psi_{n1}(\beta_0) = \Big(\frac{n}{\ln n}\Big)^{\frac{1}{2}}\mathbb{1}_{]0,\frac{3}{4}[}(\beta_0) + \Big(\frac{n^{1/2}}{\ln n}\Big)\mathbb{1}_{\{\frac{3}{4}\}}(\beta_0) + \Big(\frac{n^{2(1-\beta_0)}}{\ln n}\Big)\mathbb{1}_{]\frac{3}{4},1[}(\beta_0).$$

We apply Lemma 6.3 and Proposition 6.2 with $p = 1$ or $p = 2$. After some calculations and the application of Borel Cantelli's lemma with c_p chosen large enough, we obtain that almost surely, $\overline{\lim}_{n\to\infty}\,\psi_{np}(\beta_0)\,|F_{2,n,p}(u)| < +\infty$ under the condition A2.1-(iii-p), with $p = 1$ or 2.

(iii) *Study of* $F_{3,n,p}(u)$**.** From (6.11) and proceeding similarly as in (6.12), we get for $r_0 \geq 1$, that $n^{\beta_1}(n^{\alpha_p}\mathbb{E}\overline{(D_{r_p}^{(u)}X)^2} - u^{\alpha_p}\,l_n(p,r_0,\beta_0))$ could be decomposed into $B_{n1} + B_{n2} + B_{n3}$ with

$$B_{n1} = -\frac{n^{\alpha_p+\beta_1}}{2(n-ur_p+1)} \sum_{k=0}^{n-ur_p} \sum_{i,j=0}^{r_p} b_{ikr}^{(u)} b_{jkr}^{(u)} (t_{k+iu}-t_k)^{r_0} (t_{k+ju}-t_k)^{r_0} \times \iint_{[0,1]^2} \frac{((1-v)(1-w))^{r_0-1}}{(r_0-1)!^2} |\dot{v}_{ik}-\dot{w}_{jk}|^{2\beta_0+\beta_1} \times \left\{ \frac{\frac{\mathbb{L}^{(r_0,r_0)}(\dot{v}_{ik},\dot{w}_{jk})-\frac{1}{2}\mathbb{L}^{(r_0,r_0)}(\dot{v}_{ik},\dot{v}_{ik})-\frac{1}{2}\mathbb{L}^{(r_0,r_0)}(\dot{w}_{jk},\dot{w}_{jk})}{|\dot{v}_{ik}-\dot{w}_{jk}|^{2\beta_0}} - d_0(\dot{w}_{jk})}{|\dot{v}_{ik}-\dot{w}_{jk}|^{\beta_1}} - d_1(\dot{w}_{jk}) \right\} \mathrm{d}v\,\mathrm{d}w$$

$$B_{n2} = -\frac{n^{\alpha_p+\beta_1}}{2(n-ur_p+1)} \sum_{k=0}^{n-ur_p} \sum_{i,j=0}^{r_p} b_{ikr}^{(u)} b_{jkr}^{(u)} (t_{k+iu}-t_k)^{r_0} (t_{k+ju}-t_k)^{r_0} \times \iint_{[0,1]^2} \frac{((1-v)(1-w))^{r_0-1}}{(r_0-1)!^2} |\dot{v}_{ik}-\dot{w}_{jk}|^{2\beta_0+\beta_1} d_1(\dot{w}_{jk}),$$

$$B_{n3} = n^{\beta_1}\Big(\frac{-n^{\alpha_p}}{2(n-ur_p+1)} \sum_{k=0}^{n-ur_p} \sum_{i,j=0}^{r_p} b_{ikr}^{(u)} b_{jkr}^{(u)} (t_{k+iu}-t_k)^{r_0} (t_{k+ju}-t_k)^{r_0} \times \iint_{[0,1]^2} \frac{((1-v)(1-w))^{r_0-1}}{(r_0-1)!^2} |\dot{v}_{ik}-\dot{w}_{jk}|^{2\beta_0} d_0(\dot{w}_{jk})\,\mathrm{d}v\,\mathrm{d}w - u^{\alpha_p} l_n(p,r_0,\beta_0)\Big)$$

and $l_n(p,r_0,\beta_0)$ given by (6.17). Next, using Lemma 6.1 and 6.2 and the condition (3.1) with uniform continuity of $d_1(\cdot)$, we get that $B_{n1}=o(1)$ and B_{n2} has the limit:

$$-\frac{u^{\alpha_p+\beta_1}}{2} \sum_{i,j=0}^{r_p} \frac{(ij)^{r_0} \int_0^T d_1(t)\psi^{1-\alpha_p-\beta_1}(t)\,\mathrm{d}t}{\prod_{\substack{m=0\\m\neq i}}^{r_p}(i-m)\prod_{\substack{q=0\\q\neq j}}^{r_p}(j-q)} \times \iint_{[0,1]^2} \frac{\big((1-v)(1-w)\big)^{r_0-1}}{((r_0-1)!)^2} |iv-jw|^{2\beta_0+\beta_1}\,\mathrm{d}v\mathrm{d}w.$$

For the last term B_{n3}, one may show that it is of order $\mathcal{O}(n^{\beta_1-1})$. Finally, the case $r_0=0$ is treated similarly from (6.10).

Conclusion. One may note that the deterministic term, $l_n(p,r_0,\beta_0)$, defined in (6.17)–(6.18), converges to the nonzero term:

$$-\frac{1}{2}\sum_{i,j=0}^{r_p} \frac{(ij)^{r_0}\int_0^T d_0(t)\psi^{-\alpha_p+1}(t)\,\mathrm{d}t}{\prod\limits_{\substack{m=0\\m\neq i}}^{r_p}(i-m)\prod\limits_{\substack{q=0\\q\neq j}}^{r_p}(j-q)} \iint_{[0,1]^2} \frac{((1-v)(1-w))^{r_0-1}}{(r_0-1)!^2} |iv-jw|^{2\beta_0}\mathrm{d}v\mathrm{d}w$$

for $r_0 \geq 1$ while if $r_0=0$, the limit is

$$-\frac{1}{2}\sum_{i,j=0}^{p} \frac{|i-j|^{2\beta_0}\int_0^T d_0(t)\psi^{-\alpha_p+1}(t)\,\mathrm{d}t}{\prod_{\substack{m=0\\m\neq i}}^{p}(i-m)\prod_{\substack{q=0\\q\neq j}}^{p}(j-q)}.$$

□

6.5. Proof of the results on approximation and integration

Proof of Theorem 4.1. We set $\widetilde{r_0} = \max(\widehat{r}_0, 1)$ and, for $\widehat{r}_0$ and $\widetilde{X}_r(\cdot)$ respectively defined in (2.5) and (4.1). For $\widehat{r}_0 = l_0$, we make use of the convention: $\widetilde{X}_{\widetilde{r_0}}(\cdot) = \widetilde{X}_{m_n-1}(\cdot)$ and $\widetilde{X}_{\widehat{r}_0+1}(\cdot) = \widetilde{X}_{m_n}(\cdot)$.

(a) If $\overline{r} = \max(r, 1)$ and $\overline{r_0} = \max(r_0, 1)$, we get, for n large enough such that $r_0 \le m_n - 2$,

$$\begin{aligned}\big(X(t) - \widetilde{X}_{\widetilde{r_0}}(t)\big)^2 &= \sum_{r=0}^{m_n-2} \big(X(t) - \widetilde{X}_{\overline{r}}(t)\big)^2 \mathbb{1}_{\{\widehat{r}_0 = r\}} + \big(X(t) - \widetilde{X}_{m_n-1}(t)\big)^2 \mathbb{1}_{\{\widehat{r}_0 = l_0\}} \\ &\le \big(X(t) - \widetilde{X}_{\overline{r_0}}(t)\big)^2 + \mathbb{1}_{\{\widehat{r}_0 \ne r_0\}} \sum_{r=0, r\ne r_0}^{m_n-1} \big(X(t) - \widetilde{X}_{\overline{r}}(t)\big)^2.\end{aligned}$$

Therefore, $e_\rho^2\big(\mathrm{app}(\widehat{r}_0)\big)$ could be bounded by

$$\int_0^T \mathbb{E}\big(X(t) - \widetilde{X}_{\overline{r_0}}(t)\big)^2 \rho(t)\,\mathrm{d}t + \big(\mathbb{P}(\widehat{r}_0 \ne r_0)\big)^{\frac{1}{2}} \sum_{r=0, r\ne r_0}^{m_n-1} \int_0^T \Big(\mathbb{E}\big(X(t) - \widetilde{X}_{\overline{r}}(t)\big)^4\Big)^{\frac{1}{2}} \rho(t)\,\mathrm{d}t$$

We use the exponential bound established for $\mathbb{P}(\widehat{r}_0 \ne r_0)$ in Theorem 3.1 as well as the property $\mathbb{E}(Y^4) \le 3(\mathbb{E}(Y^2))^2$ for a Gaussian r.v. Y. Moreover, $\sup_{t\in[0,T]}(\mathbb{E}(X(t) - \widetilde{X}_r(t))^2) = \max_{k=0,\dots,\lfloor \frac{n}{r} \rfloor -1} \sup_{t\in\mathcal{I}_k}(\mathbb{E}(X(t) - \widetilde{X}_r(t))^2)$. If $r_0 \ge 1$, we start from the decomposition established in Blanke and Vial [6, Lemma 4.1] to obtain, for $t \in \mathcal{I}_k$ and $r^* = \min(r, r_0)$:

$$\begin{aligned}\mathbb{E}\big(X(t) - \widetilde{X}_r(t)\big)^2 &= \sum_{i,j=0}^{r} L_{i,k,r}(t) L_{j,k,r}(t) \frac{(t_{kr+i} - t_{kr})^{r^*}(t_{kr+j} - t_{kr})^{r^*}}{((r^*-1)!)^2} \\ \times \iint_{[0,1]^2} &\big((1-v)(1-w)\big)^{r^*-1} \Big\{ \mathbb{L}^{(r^*,r^*)}(t_{kr} + (t - t_{kr})v, t_{kr} + (t - t_{kr})w) \\ &- \mathbb{L}^{(r^*,r^*)}(t_{kr} + (t - t_{kr})v, t_{kr} + (t_{kr+j} - t_{kr})w) \\ &- \mathbb{L}^{(r^*,r^*)}(t_{kr} + (t_{kr+i} - t_{kr})v, t_{kr} + (t - t_{kr})w) \\ &+ \mathbb{L}^{(r^*,r^*)}(t_{kr} + (t_{kr+i} - t_{kr})v, t_{kr} + (t_{kr+j} - t_{kr})w) \Big\}\,\mathrm{d}v\mathrm{d}w.\end{aligned}$$

If $r = 1, \dots, r_0 - 1$ ($r_0 \ge 2$), we obtain the uniform bound $\mathcal{O}(\delta_n^{2r+2})$ by uniform continuity of $\mathbb{L}^{(r+1,r+1)}(\cdot,\cdot)$ and results of Lemma 6.1. For $r = r_0, \dots, m_n$, we have $r^* = r_0$ so we apply the Hölderian regularity condition (6.9). Since $L_{i,k,r}(t) \le r^r$, we arrive at $\sup_{t\in[0,T]} \mathbb{E}(X(t) - \widetilde{X}_{\overline{r_0}}(t))^2 = \mathcal{O}(\delta_n^{2(r_0+\beta_0)})$ for $r = r_0$ while if $r = r_0 + 1, \dots, m_n$, $\sup_{t\in[0,T]} \mathbb{E}(X(t) - \widetilde{X}_{\overline{r}}(t))^2 = \mathcal{O}(m_n^{2(m_n+r_0+\beta_0)} \delta_n^{2(r_0+\beta_0)})$. The logarithmic order of m_n yields the final result. In the case $r_0 = 0$, note that the above results hold true starting from

$$\mathbb{E}\big(X(t)-\widetilde{X}_{\overline{r}}(t)\big)^2=\sum_{i,j=0}^{\overline{r}} L_{i,k,\overline{r}}(t)L_{j,k,\overline{r}}(t)\Big\{\mathbb{L}(t,t)-\mathbb{L}(t,t_{k\overline{r}+j}) - \mathbb{L}(t_{k\overline{r}+i},t)+\mathbb{L}(t_{k\overline{r}+i},t_{k\overline{r}+j})\Big\}.$$

(b) For $e_\rho^2\big(\mathrm{int}\big(\widehat{r}_0\big)\big)$, $\int_0^T \big(X(t)-\widetilde{X}_{r+1}\big)\rho(t)\,\mathrm{d}t$ is again a Gaussian variable, so in a similar way as for approximation, we get the following bound for this term:

$$\sqrt{3}\big(\mathbb{P}(\widehat{r}_0\neq r_0)\big)^{\frac12}\sum_{r=0}^{m_n}\Big(\sup_{t\in[0,T]}\Big(\mathbb{E}\big(X(t)-\widetilde{X}_{r+1}(t)\big)^2\Big)^{\frac12}\Big)^2\Big(\int_0^T\rho(t)\,\mathrm{d}t\Big)^2$$
$$+\sum_{k=0}^{\lfloor\frac{n}{r_0+1}\rfloor-1}\sum_{\ell=0}^{\lfloor\frac{n}{r_0+1}\rfloor-1}\int_{\mathcal{I}_k}\int_{\mathcal{I}_\ell}\mathbb{E}\big(X(t)-\widetilde{X}_{r_0+1}(t)\big)\big(X(s)-\widetilde{X}_{r_0+1}(s)\big)\rho(t)\rho(s)\,\mathrm{d}s\mathrm{d}t.$$

Study of the term $\mathbb{E}\big(X(t)-\widetilde{X}_{r_0+1}(t)\big)\big(X(s)-\widetilde{X}_{r_0+1}(s)\big)$, $(s,t)\in\mathcal{I}_\ell\times\mathcal{I}_k$. Denoting $\overline{r}=r_0+1$ we get again from lemma 4.1 of Blanke and Vial [6] that $\mathbb{E}\big(X(t)-\widetilde{X}_{\overline{r}}(t)\big)\big(X(s)-\widetilde{X}_{\overline{r}}(s)\big)$ is equal to:

$$\sum_{i,j=0}^{\overline{r}} L_{i,k,\overline{r}}(t)L_{j,\ell,\overline{r}}(s)\frac{\big((t_{k\overline{r}+i}-t_{k\overline{r}})(t_{\ell\overline{r}+j}-t_{\ell\overline{r}})\big)^{r_0}}{((r_0-1)!)^2}\iint_{[0,1]^2}\mathrm{d}v\mathrm{d}w\,\big((1-v)(1-w)\big)^{r_0-1}$$
$$\times\big\{\mathbb{L}^{(r_0,r_0)}(t_{k\overline{r}}+(t-t_{k\overline{r}})v,t_{\ell\overline{r}}+(t-t_{\ell\overline{r}})w)-\mathbb{L}^{(r_0,r_0)}(t_{k\overline{r}}+(t-t_{k\overline{r}})v,t_{\ell\overline{r}}+(t_{\ell\overline{r}+j}-t_{\ell\overline{r}})w)$$
$$-\mathbb{L}^{(r_0,r_0)}(t_{k\overline{r}}+(t_{k\overline{r}+i}-t_{k\overline{r}})v,t_{\ell\overline{r}}+(t-t_{\ell\overline{r}})w)+\mathbb{L}^{(r_0,r_0)}(t_{k\overline{r}}+(t_{k\overline{r}+i}-t_{k\overline{r}})v,t_{\ell\overline{r}}+(t-t_{\ell\overline{r}})w)\big\}.$$

For non-overlapping intervals $\mathcal{I}_k$ and $\mathcal{I}_\ell$, that is $|k-l|\geq 2$, we make use of Condition A2.2(2) four times, by adding and subtracting the necessary terms, and noting that

$$\sum_{i,j=0}^{\overline{r}} L_{i,k,\overline{r}}(t)L_{j,\ell,\overline{r}}(s)(t_{k\overline{r}+i}-t_{k\overline{r}})^{r_1}(t_{\ell\overline{r}+j}-t_{\ell\overline{r}})^{r_2}=(t-t_{k\overline{r}})^{r_1}(s-t_{\ell\overline{r}})^{r_2}$$

with either $r_i=\overline{r}-1$ or $r_i=\overline{r}$ for $i=1,2$. Therefore, we get

$$\sum_{\substack{k,\ell=0\\|k-\ell|\geq 2}}^{\lfloor\frac{n}{\overline{r}}\rfloor-1}\int_{\mathcal{I}_k}\int_{\mathcal{I}_\ell}\mathbb{E}\big(X(t)-\widetilde{X}_{\overline{r}}(t)\big)\big(X(s)-\widetilde{X}_{\overline{r}}(s)\big)\rho(t)\rho(s)\,\mathrm{d}s\mathrm{d}t$$
$$=\mathcal{O}\Big(\delta_n^{2(r_0+\beta_0+1)}\sum_{\substack{k,\ell=0\\|k-\ell|\geq 2}}^{\lfloor\frac{n}{\overline{r}}\rfloor-1}\big|\,|k-\ell|-1\big|^{-2(2-\beta_0)}\Big)$$

which is a $\mathcal{O}(\delta_n^{2(r_0+\beta_0)+1})$. For overlapping intervals $\mathcal{I}_k$ and $\mathcal{I}_\ell$, namely in the case where $|k-l|\leq 1$, we make use of Cauchy-Schwarz inequality to obtain the same bound as above. Since the second part of $e_\rho^2(\mathrm{int}(\widehat{\mathrm{r}}_0))$ is negligible, we obtain the claimed result. □

Acknowledgements

We thank the Editor and Referee for their careful reading and insightful comments.

References

[1] Adler, R. J. (1981). *The Geometry of Random Fields*. Wiley, New-York. MR0611857

[2] Adler, R. J. (1990). *An Introduction to Continuity, Extrema, and Related Topics for General Gaussian Processes*. I.M.S. Lecture Notes—Monograph Series, 12. Institute of Mathematical Statistics, Hayward, CA. MR1088478

[3] Benhenni, K. (1998). Approximating integrals of stochastic processes: Extensions. *J. Appl. Probab. 35*(4), 843–855. MR1671235

[4] Benhenni, K. and Cambanis, S. (1992). Sampling designs for estimating integrals of stochastic processes. *Ann. Statist. 20*(1), 161–194. MR1150339

[5] Berman, S. (1974). Sojourns and extremes of Gaussian processes. *Ann. Probab. 2*, 999–1026 (Corrections (1980), 8, 999 and (1984) 12, 281).

[6] Blanke, D. and Vial, C. (2008). Assessing the number of mean-square derivatives of a Gaussian process. *Stochastic Process. Appl. 118*(10), 1852–1869. MR2454467

[7] Blanke, D. and Vial, C. (2011). Estimating the order of mean-square derivatives with quadratic variations. *Stat. Inference Stoch. Process. 14*(1), 85–99. MR2780650

[8] Blanke, D. and Vial, C. (2012). On estimation of regularity for Gaussian processes. Preprint arXiv:1211.2763 (November), 34 pages. `http://arxiv.org/pdf/1211.2763v1`.

[9] Brown, P. J., Fearn, T., and Vannucci, M. (2001). Bayesian wavelet regression on curves with application to a spectroscopic calibration problem. *J. Amer. Statist. Assoc. 96*(454), 398–408. MR1939343

[10] Bucklew, J. A. (1985). A note on the prediction error for small time lags into the future. *IEEE Trans. Inform. Theory 31*(5), 677–679. MR0808241

[11] Cambanis, S. (1985). Sampling designs for time series. In *Time Series in the Time Domain*. Handbook of Statist., 5, pp. 337–362. North-Holland, Amsterdam. MR0831755

[12] Chan, G., Hall, P., and Poskitt, D. (1995). Periodogram-based estimators of fractal properties. *Ann. Statist. 23*(5), 1684–1711. MR1370303

[13] Constantine, A. G. and Hall, P. (1994). Characterizing surface smoothness via estimation of effective fractal dimension. *J. Roy. Statist. Soc., Ser. B 56*(1), 97–113. MR1257799

[14] Cressie, N. A. C. (1993). *Statistics for Spatial Data*. Wiley, New-York. MR1239641

[15] Cuzick, J. (1977). A lower bound for the prediction error of stationary Gaussian processes. *Indiana Univ. Math. J. 26*(3), 577–584. MR0438452

[16] Feuerverger, A., Hall, P., and Wood, A. (1994). Estimation of fractal index and fractal dimension of a Gaussian process by counting the number of level crossing. *J. Time Ser. Anal. 15*(6), 587–606. MR1312323

[17] Gilbert, R. O. (1987). *Statistical Methods for Environmental Pollution Monitoring*. Van Nostrand-Reinhold, New York.

[18] Gneiting, T., Ševčíková, and Percival, D. B. (2012). Estimators of fractal dimension: assessing the roughness of time series and spatial data. *Statist. Sci. 27*(2), 247–277. MR2963995

[19] Hall, P. and Roy, R. (1994). On the relationship between fractal dimension and fractal index for stationary stochastic processes. *Ann. Appl. Probab. 4*(1), 241–253. MR1258183

[20] Hanson, D. L. and Wright, F. T. (1971). A bound on tail probabilities for quadratic forms in independent random variables. *Ann. Math. Statist. 42*(3), 1079–1083. MR0279864

[21] Istas, J. (1992). Wavelet coefficients of a Gaussian process and applications. *Ann. Inst. H. Poincaré Probab. Statist. 28*(4), 537–556. MR1193084

[22] Istas, J. and Lang, G. (1997). Quadratic variations and estimation of the local hölder index of a Gaussian process. *Ann. Inst. H. Poincaré, Probab. Statist. 33*(4), 407–436. MR1465796

[23] Istas, J. and Laredo, C. (1997). Estimating functionals of a stochastic process. *Adv. in Appl. Probab. 29*(1), 249–270. MR1432939

[24] Kent, J. T. and Wood, A. T. (1997). Estimating the fractal dimension of a locally self-similar Gaussian process by using increments. *J. Roy. Statist. Soc., Ser. B 59*(3), 679–699. MR1452033

[25] Laslett, G. M. (1994). Kriging and splines: An empirical comparison of their predictive performance in some applications. *J. Amer. Statist. Assoc. 89*(426), 391–409. With comments and a rejoinder by the author. MR1294071

[26] Lindgren, G. (1979). Prediction of level crossings for normal processes containing deterministic components. *Adv. in Appl. Probab. 11*(1), 93–117. MR0517553

[27] Müller-Gronbach, T. (1996). Optimal designs for approximating the path of a stochastic process. *J. Statist. Plann. Inference 49*(3), 371–385. MR1381165

[28] Müller-Gronbach, T. and Ritter, K. (1997). Uniform reconstruction of Gaussian processes. *Stochastic Process. Appl. 69*(1), 55–70. MR1464174

[29] Müller-Gronbach, T. and Ritter, K. (1998). Spatial adaption for predicting random functions. *Ann. Statist. 26*(6), 2264–2288. MR1700231

[30] Plaskota, L., Ritter, K., and Wasilkowski, G. (2004). Optimal designs for weighted approximation and integration of stochastic processes on $[0,\infty($. *J. Complexity 20*(1), 108–131. MR2031561

[31] R Core Team (2014). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria.

[32] Ritter, K. (1996). Asymptotic optimality of regular sequence designs. *Ann. Statist. 24*(5), 2081–2096. MR1421162

[33] Ritter, K. (2000). *Average-case analysis of numerical problems.* Lecture Notes in Mathematics, 1733. Springer. MR1763973

[34] Sacks, J. and Ylvisaker, D. (1968). Designs for regression problems with correlated errors; many parameters. *Ann. Math. Statist. 39*, 49–69. MR0220424

[35] Sacks, J. and Ylvisaker, D. (1970). Designs for regression problems with correlated errors. III. *Ann. Math. Statist. 41*, 2057–2074. MR0270530

[36] Seleznjev, O. (1996). Large deviations in the piecewise linear approximation of Gaussian processes with stationary increments. *Adv. in Appl. Probab. 28*(2), 481–499. MR1387887

[37] Seleznjev, O. (2000). Spline approximation of random processes and design problems. *J. Statist. Plann. Inference 84*(1–2), 249–262. MR1747507

[38] Seleznjev, O. and Buslaev, A. (1998). Best approximation for classes of random processes. Technical Report 13, 14 p., Univ. Lund Research Report.

[39] Stein, M. L. (1995). Predicting integrals of stochastic processes. *Ann. Appl. Probab. 5*(1), 158–170. MR1325047

[40] Stein, M. L. (1999). *Interpolation of Spatial Data.* Springer Series in Statistics. Springer-Verlag, New York. Some theory for Kriging. MR1697409

[41] Taylor, C. C. and Taylor, S. J. (1991). Estimating the dimension of a fractal. *J. Roy. Statist. Soc. Ser. B 53*(2), 353–364. MR1108332

[42] Tsai, H. and Chan, K. S. (2000). A note on the covariance structure of a continuous-time process. *Statist. Sinica 10*, 989–998. MR1789804

[43] Wood, A. T. and Chan, G. (1994). Simulation of stationary Gaussian processes in $[0,1]^d$. *J. Comput. Graph. Statist. 3*(4), 409–432. MR1323050