Local Hölder exponent estimation for multivariate continuous time processes

D. Blanke

Running head: Local Hölder exponent estimation

Abstract: In continuous time, rates of convergence for nonparametric density estimators are related with sample paths regularity of the underlying process. In this paper, we propose and study two families of estimators for the local Hölder exponent, $\gamma$, of multivariate continuous time processes. For observed sample paths over $[0, T_n]$, $T_n \uparrow \infty$, we used increment-based type estimators and give their almost sure rates of convergence for strongly mixing processes. It is shown that these rates may depend on an extra parameter, $\beta$, corresponding to second order regularity of the observed process. To avoid such a difficulty, a family of preliminary estimators of $\beta$ is also given and studied.

MSC 2000: 62G05, 62G20, 62M99, 60G17

Keywords: Hölder exponent, statistics of processes, continuous time.

1 Introduction

In the special case of real stationary Gaussian processes, there exist many works concerning estimation of fractal dimension of sample paths. They used the simple and well known relationship - see Adler, (1981), Taylor and Taylor, (1991) and references therein - between fractal dimension and fractal index of the covariances to compute increment-based estimators. In such a context, we refer e.g. to works of Constantine and Hall (1994), Istas and Lang (1997), Kent and Wood (1997), Peltier and Lévy-Véhel (1994). Still in the Gaussian case, see also Chan et al. (1995) for a periodogram-type estimator, Feuerverger et al. (1994) for estimation with number of level crossings and finally Philippe and Thilly (2000) who use convex rearrangements of Gaussian processes. In this paper, we want to extend these results to general processes, not necessarily Gaussian.

Our prime interest is to provide results about local smoothness of sample paths of the underlying process. But also, such an information is crucial for nonparametric functional estimation in continuous time. Indeed nowadays special behavior of e.g. density estimators is quite well understood : the rates essentially depend on a coefficient, say $\gamma$,
linked with the Hölderian properties of sample paths: roughly speaking, more they are irregular, best is the rate! More complete discussion on these rates appears below in Section 4.2.

We consider the following general framework: let \( \{X_t, t \in \mathbb{R}\} \) be a \( \mathbb{R}^d \)-valued measurable stochastic process defined on a probability space \((\Omega, \mathcal{A}, P)\), \( \mathbb{R}^d \) equipped with any suitable norm \( \| \cdot \| \). Suppose moreover that for some positive \( \tilde{\gamma} \) and convex function \( \varphi \) increasing over \( \]0, +\infty[ \) there exist a positive \( h \) such that for all \( 0 < s \leq h \), one has

\[
\sup_{t \in \mathbb{R}^+} E \varphi \left( \frac{\|X_{t+s} - X_t\|}{s^{\tilde{\gamma}}} \right) < +\infty. \tag{1.1}
\]

If \( \Gamma \) denotes the set of \( \tilde{\gamma} \) such that condition (1.1) holds then our goal is to estimate \( \gamma \) where \( \gamma = \sup \{ \tilde{\gamma}, \tilde{\gamma} \in \Gamma \} \). Now, if moreover \( X(.) \) is \( P \)-almost surely continuous with respect to \( \| \cdot \| \), then a result of Csáki and Csörgő (1992) implies that if (1.1) is satisfied with \( \varphi(x) = e^{ax} \) and positive constant \( a \) then for any \( 0 \leq t_1 < t_2 < +\infty \) and any \( \tilde{\gamma} \in \Gamma \),

\[
\limsup_{h \downarrow 0} \sup_{t_1 \leq t \leq t_2 - h} \sup_{0 \leq s \leq h} \frac{\|X_{t+s} - X_t\|}{h^{\tilde{\gamma}}(\ln(1/h))} \leq c, \quad \text{a.s.} \tag{1.2}
\]

so our study is linked with moduli of continuity of the process \( X(.) \). Note also that it appears quite difficult to estimate \( \gamma \) directly from (1.2) since this property is only asymptotic and in most cases, constant \( c \) is unknown!

To compute estimators, we will also consider the case when the whole sample path of \( \{X_t\} \) is observed over \( [0, T_n] \) with sequences of times such that \( T_{n+1} - T_n \geq a > 0 \) and \( T_n \uparrow \infty \): in other words \( \{X_t\} \) is displayed at increasing instants, \( T_1, T_2, \ldots \), and during a long period. Note that paper (2002) deals with the same problem but for real-valued processes, only observed at the \( n \) sampling times \( 0, \delta_n, 2\delta_n, \ldots, (n-1)\delta_n \) with \( \delta_n \to 0 \) and \( n\delta_n \to \infty \) as \( n \to \infty \).

The rest of the paper is organized as follows: in Section 2, we give some auxiliary asymptotic convergence theorems needed in further calculations. In Section 3, we present and study the almost sure behavior of two classes of estimators for \( \gamma \). As in previous cited works, asymptotic properties depend on an higher order of regularity \( \beta \) (in general unspecified), thus to complete these results we supply a preliminary estimator of \( \beta \) and obtain its exact rate of convergence. Section 4 is devoted to some discussion and comments: first, we point out the main differences between our results and those obtained in the sampling case: as expected the continuous time framework improves the global behavior of estimators. Next, we give some links with nonparametric density estimation in continuous time. Finally proofs are postponed to Section 5.

2 Preliminary results

Our first results deal with the statistics:

\[
S_{T_n, \ell}(\gamma) = \frac{1}{T_n - \ell\delta_T} \int_0^{T_n - \ell\delta_T} \left\| \frac{X_{t+\ell\delta_T} - X_t}{(\ell\delta_T)^{\gamma}} \right\|^k \, dt,
\]
defined on \([0, +\infty[\] where \(k \in [1, +\infty[, \ell \in \mathbb{N}^* \) are given numbers and \(\delta_{T_n} \to 0 \) as \(T_n \uparrow \infty\).

In Section 3, we use \(S_{T_n, \ell}^{(k)}(\gamma)\) in the construction and the study of our estimators of \(\gamma\).

For the almost sure convergence (a.s.) of \(S_{T_n, \ell}^{(k)}(\gamma)\), we introduce the classical strong mixing coefficient defined by

\[
al_{X}(u) = \sup_{t \in \mathbb{R}} \sup_{A \in \mathcal{F}(X_{s}, s \leq t)} \sup_{B \in \mathcal{F}(X_{s}, s \geq t+u)} \left| P(A \cap B) - P(A)P(B) \right|, \quad u > 0.
\]

Recall that \((X_t)\) is geometrically strongly mixing (G.S.M.) if there exist \(c_1 > 0\) and \(0 \leq \rho < 1\) such that

\[
\alpha_{X}(j) \leq c_1 \rho^j, \quad j > 0; \quad (X_t)\) is arithmetically strongly mixing (A.S.M.) if there exists \(c_2 \geq 0\) and \(b > 0\) such that \(\alpha_{X}(j) \leq c_2 j^{-b}, \quad j > 0.\)

Now we state the a.s. asymptotic behavior of \(S_{T_n, \ell}^{(k)}(\gamma)\) for geometrically strongly mixing processes.

**Theorem 2.1** If \(\{X_t, t \in \mathbb{R}\}\) is a G.S.M. process such that

(a) for a real \(\nu \) with \(\nu > 2\) and a given \(k \geq 1\),

\[
\exists \delta_0 > 0, \forall \delta \leq \delta_0, \quad M(k\nu) := \sup_{t \in \mathbb{R}^+} E \left\| X_{t+\ell \delta} - X_t \right\|^{k\nu} < +\infty \tag{2.1}
\]

one obtains:

\[
\lim_{T_n \uparrow \infty} \left| S_{T_n, \ell}^{(k)}(\gamma) - E S_{T_n, \ell}^{(k)}(\gamma) \right| = 0 \quad \text{a.s.},
\]

(b) if moreover,

\[
\exists \delta_0 > 0, \exists a > 0, \forall \delta \leq \delta_0, \quad M := \sup_{t \in \mathbb{R}^+} E \left( e^{a \left\| X_{t+\ell \delta} - X_t \right\|} \right)^{\gamma} < +\infty, \tag{2.2}
\]

then

\[
\limsup_{T_n \uparrow \infty} \sqrt{\frac{T_n}{\ln T_n}} \left| S_{T_n, \ell}^{(k)}(\gamma) - E S_{T_n, \ell}^{(k)}(\gamma) \right| \leq C_1 \quad \text{a.s.} \tag{2.3}
\]

where \(C_1\) denotes a positive constant.

**Remarks:**

- The constant \(C_1\) is complicated but explicit as shown by the proof (see relation (5.20)). It depends on \(M, \alpha_{X}, k, \nu\) but it is independent of \(\ell\).
- The local Hölder condition (1.2) holds true as soon as assumption (2.2) is satisfied (Csáki and Csörgő, 1992).
- Condition (2.2) is easily checked for real stationary Gaussian with correlation function \(\rho(u)\) such that \(\rho(u) = 1 - c_1 u^{2\gamma} + o(u^{2\gamma})\) when \(u \to 0^+\). As an example, for the Ornstein-Uhlenbeck process, solution of

\[
dX_t = -\theta X_t \, dt + \sigma \, dW_t, \quad \theta > 0, \sigma > 0
\]
one gets $\gamma = 1/2$. For the Wong process (Wong, 1966) given by

$$X_t = \sqrt{3} \exp \left( -\sqrt{3} t \right) \int_0^{\exp(2t/\sqrt{3})} W_s \, ds$$

and with autocorrelation $\rho(u) = (1 - \frac{1}{3} e^{-|u|/\sqrt{3}})$, one obtains $\gamma = 1$.

- Geometric mixing should be replaced by arithmetic mixing:

**Theorem 2.2** If $\{X_t, t \in \mathbb{R}\}$ is an A.S.M. process such that

(a) condition (2.1) holds for a real $\nu$, $\nu > 2$, one obtains for all $b > \frac{\nu+2}{\nu-2}$,

$$
\lim_{T_n \to \infty} \left\| S_{T_n,\ell}^{(k)}(\gamma) - E S_{T_n,\ell}^{(k)}(\gamma) \right\| = 0 \quad \text{a.s.,}
$$

(b) condition (2.2) is satisfied, then for all $b > 3$ one has

$$
\limsup_{T_n \to \infty} \sqrt{\frac{T_n}{\ln T_n}} \left| S_{T_n,\ell}^{(k)}(\gamma) - E S_{T_n,\ell}^{(k)}(\gamma) \right| \leq C_1 \quad \text{a.s.}
$$

where $C_1$ is the same constant as in Theorem 2.1(b).

**Remarks:**

- The rate obtained in Theorems 2.1 and 2.2 is independent of the unknown coefficient $\gamma$.

- In addition if, the sample paths of $T \mapsto Z_T := \frac{T^{1/2}(\ln T)^{-1/2}}{\sqrt{\ell \delta T}} \int_0^{T-\ell \delta T} \left\| X_{t+\ell \delta T} - X_t \right\| k (\ell \delta T)^{-\gamma} \, dt$ are uniformly continuous with probability 1 over $\mathbb{R}^+$, then the Borel-Cantelli type lemma for continuous time processes established by Bosq (1998b) implies that results of Theorems 2.1 and 2.2 hold true with $T_n$ replaced by $T$.

- One may easily show that if $\{X_t, t \in \mathbb{R}\}$ has independent components $\{X_i(t), t \in \mathbb{R}\}, i = 1, \ldots, d$, each with Hölder regularity $\gamma_i$, then condition (2.2) is true if it holds for $\gamma(1) := \inf_{i=1,\ldots,d} \gamma_i$.

- For stationary Gaussian processes, the strong mixing condition is equivalent to the complete regularity condition:

$$
\rho(\tau) = \rho((X_s, s \leq t), (X_s, s \geq t + \tau)) \to 0 \quad \text{as} \quad \tau \to +\infty
$$

where $\rho$ denotes the correlation coefficient. Moreover in this case, complete regularity conditions may be linked with spectral characteristics of processes. Following Theorem 5 in Ibragimov and Rozanov (1978, p. 212), we get this sufficient complete regularity criteria:

If the spectral density $f(\lambda)$ of the process $\{X_t\}$ permits a representation of the form $f(\lambda) = |\Gamma(\lambda)|^2 w(\lambda)$ where $\Gamma$ is a bounded entire function of finite degree $\leq \sigma$, the function $\ln w(\lambda)$ is uniformly continuous on $(-\infty, \infty)$ and $\int_1^{\infty} (w(s)/s^2) \, ds < \infty$; then the process is completely regular with $\rho(\tau) \leq Cw \left( \frac{1}{\tau - 2\sigma} \right), \tau > 2\sigma$.

For the bias, we have the following:
Lemma 2.1 Assume that

\[ \exists \delta_0 > 0, \forall \delta \leq \delta_0, \forall v \geq 0, \; E \| X_{t+\delta} - X_v \|^k = \delta^{k\gamma} m_\delta^{(k)}(v) \]

with \( \frac{1}{T_n} \int_0^{T_n} m_\delta^{(k)}(v) \, dv = m_k + R_k^{(k)}(\ell \delta_{T_n}) \) for a strictly positive \( m_k \),

(a) if \( R_k^{(k)}(\delta_{T_n}) \to 0 \) as \( T_n \uparrow \infty \), then \( \lim_{T_n \uparrow \infty} \left| E S_{T_n,\delta}^{(k)}(\gamma) - m_k \right| = 0 \),

(b) if \( \lim_{T_n \uparrow \infty} \delta^{-\beta} R_k^{(k)}(\delta_{T_n}) = b_k, \; (\beta > 0, b_k > 0) \), then \( \lim_{T_n \uparrow \infty} \delta^{-\beta} \left| E S_{T_n,\delta}^{(k)}(\gamma) - m_k \right| = b_k \ell^\beta \).

Assumptions of Lemma 2.1(a) obviously hold for processes with stationary increments as soon as \( E \| X_u - X_0 \|^k / u^{k\gamma} \to m_k, \; u \to 0 \). Let us complete these results by giving various processes which also fulfill such conditions.

Example 2.1: Recall that if \( Z \) has an \( \mathcal{N}(0, \sigma^2) \) law then for any \( k > 0 \), \( E |Z|^k = \frac{2^{k/2} \Gamma(k+1)}{\sqrt{\pi}} \sigma^k \). So considering real stationary strongly mixing and zero-mean Gaussian processes with correlation function \( \rho(u) = 1 - c_1 u^{2\gamma} + c_2 u^{2\gamma+\beta} + o(u^{2\gamma+\beta}) \) as \( u \to 0 \), one may easily check assumptions of Lemma 2.1 for any \( k \). For example, with the Ornstein-Uhlenbeck process one gets \( \gamma = 1/2, \; \beta = 1 \) while for the Wong process, mean-square differentiable, \( \gamma = 1 \) and \( \beta = 1 \).

Example 2.2: Recall that a real diffusion process \( \{X_t\} \) is defined in particular when the following condition holds:

\[ \lim_{\delta_{T_n} \downarrow 0} \frac{E \left( (X_{t+\delta_{T_n}} - X_t)^2 \mid X_t = x \right)}{\delta_{T_n}^2} = B(t, x). \] (2.4)

If there exists a function \( \psi \) such that \( E \left( (X_t - X_s)^2 \mid X_s = x \right) \leq (t-s) \psi(x) \), then taking expectation of (2.4), one gets condition of Lemma 2.1(a) when \( B(t, x) \equiv b, \; k = 2 \) and \( \gamma = 1/2 \). Note that this last regularity condition is classical for diffusions (see e.g. theorem 4.7.1 p.145 in Kloeden and Platen, 1995).

Example 2.3: Properties of real continuous-time fractional ARMA processes are studied in Viano et al., (1994). These processes are based on the Brownian motion \( W = (W_s)_{s \in \mathbb{R}} \) and defined by \( X_t = \int_{-\infty}^t f(t-s) \, dW(s), \; t \in \mathbb{R} \) where the impulse function \( f \) belongs to \( L^2(\mathbb{R}^+) \) and with Laplace transform \( F \) such that \( F(s) = \prod_{k=1}^K (s - a_k)^{d_k} \) for \( a_k, \; d_k \in \mathbb{C} \) and \( \text{Re}(s) > a := \max(\text{Re}(a_k)/k \in E^+) \), \( E^+ \) denoting the index set of singular points of \( F \). If \( D := \sum_{k=1}^K d_k < -1/2 \) and \( a < 0 \), \( \{X_t\} \) is a zero-mean stationary and strongly mixing Gaussian process. Furthermore the behavior of the covariance function near the origin is
known and in particular (Proposition 7 in Viano et al. 1994): \( \gamma = 1, \beta = 2 \) if \( D < -3/2; \gamma = -D - 1/2, \beta = 2D + 3 \) if \(-3/2 < D < -1/2. \)

Note also that previous examples could be extended to the non Gaussian case, for example \( \{X_t, t \in \mathbb{R}\} \) and \( \{X^2_t, t \in \mathbb{R}\} \) have a similar covariance behavior near the origin since \( \text{Cov}(X_0^2, X_u^2) = 2 \left( \text{Cov}(X_0, X_u) \right)^2. \)

Finally in the multidimensional context, if \( k = 1 \) and each component \( X_t^{(i)} \) has the regularity \( (\gamma_i, \beta_i), i = 1, \ldots, d \) then conditions of Lemma 2.1 are satisfied as soon as \( (\gamma, \beta) \equiv (\gamma^{(1)}, \beta^{(1)}). \)

Combining the previous results, one obtains Corollary 2.1.

**Corollary 2.1**

Under assumptions of Theorem 2.1(a) (or Theorem 2.2(a)) and Lemma 2.1(a), we have

(i) \( \lim_{T_n \to \infty} S_{T_n,\ell}^{(k)}(\gamma) = m_k \) a.s.,

(ii) if \( \lambda > \gamma, \lim_{T_n \to \infty} S_{T_n,\ell}^{(k)}(\lambda) = +\infty \) a.s.,

(iii) if \( \lambda < \gamma, \lim_{T_n \to \infty} S_{T_n,\ell}^{(k)}(\lambda) = 0 \) a.s.

### 3 Estimation of \( \gamma \)

#### 3.1 ‘Genuine estimators’

Corollary 2.1 implies that the function \( A_{T_n}(\lambda) = S_{T_n,\ell}^{(k)}(\lambda) + \left( S_{T_n,\ell}^{(k)}(\lambda) \right)^{-1} \) should be minimal for \( \lambda \) close to \( \gamma \). Define \( \tau_\ell(\ell,k) = \arg \min_{\lambda \in \mathbb{R}} A_{T_n}(\lambda). \)

Using convexity of \( A_{T_n}(\lambda) \), one easily obtains our first family of estimators:

\[
\tilde{\gamma}_{\ell,k} = \ln \left( \frac{1}{T_n - \ell \delta T_n} \int_0^{T_n - \ell \delta T_n} \|X_{t_0 + \ell \delta T_n} - X_{t_0}\|_k^k \, dt \right) / \ln(\ell \delta T_n)
\]

where \( \ell \in \mathbb{N}, k \in [1, +\infty[ \), \( \delta T_n > 0, \delta T_n \to 0 \) are parameters chosen by the statistician. Now results of Section 2 lead to the exact asymptotic behavior of our estimator.

**Theorem 3.1**

Under assumptions of Theorem 2.1(a) (or Theorem 2.2(a)) and Lemma 2.1(a), one has

\[
\lim_{T_n \to \infty} \ln \left( (\delta T_n)^{-k} \left( \gamma - \tilde{\gamma}_{\ell,k} \right) \right) = \ln \left( m_k \right) \quad \text{a.s.}
\]

As noticed at the end of Section 2, in a multidimensional context, \( \gamma \) refers to \( \gamma^{(1)} = \inf_i \gamma_i \) where \( \gamma_i \) is associated to the Hölderian regularity of the \( i \)-th component of \( \{X_t, t \in \mathbb{R}\} \). So \( \tilde{\gamma}_{\ell,1} \) supplies an estimation for the index of the most irregular component of \( \{X_t, t \in \mathbb{R}\} \). Such estimators appear also as quite robust since their convergence does not depend on
extra or unknown parameters. Moreover they seem easy to implement and they do not require existence of an exponential moment as in Assumption (2.2). The drawback is that they will perform well only for very small values of $\delta_{Tn}$. So let us consider some extensions.

### 3.2 General estimators

We introduce for $\ell_2 > \ell_1$:

$$\hat{\gamma}((\ell_1, \ell_2, k)) = \frac{\ln(\ell_2 T_n) \tilde{\gamma}(\ell_2, k) - \ln(\ell_1 T_n) \tilde{\gamma}(\ell_1, k)}{\ln(\ell_2 / \ell_1)}.$$  

Note that for $k = 2$, such an increment-based estimator has been studied in the real Gaussian case by Istas and Lang (1997), Kent and Wood (1997).

**Theorem 3.2**

(a) Under assumptions of Theorem 2.1(a) (or Theorem 2.2(a)), Lemma 2.1(a) and for all $\delta_{Tn} \to 0$, one gets $\hat{\gamma}((\ell_1, \ell_2, k)) \to \gamma$ a.s. as $T_n \uparrow +\infty$.

(b) Under assumptions of Theorem 2.1(b) (or Theorem 2.2(b)), Lemma 2.1(b) and for all $\delta_{Tn} \to 0$ such that $\delta_{Tn} = O\left((\ln T_n / T_n)^{1/2}\right)$,

$$\limsup_{T_n \to \infty} \sqrt{\frac{T_n}{\ln T_n}} (\hat{\gamma}((\ell_1, \ell_2, k)) - \gamma) \leq \frac{2C_1}{kmk} \ln(\ell_2 / \ell_1) \quad \text{a.s.}$$

with same constant $C_1$ as in (2.3).

(c) Under assumptions of Theorem 2.1(b) (or Theorem 2.2(b)), Lemma 2.1(b) and for all $\delta_{Tn} \to 0$ such that $(\ln T_n / T_n)^{1/2} = O\left(\delta_{Tn}^{\beta}\right)$,

$$\limsup_{T_n \to \infty} \delta_{Tn}^{-\beta} (\hat{\gamma}((\ell_1, \ell_2, k)) - \gamma) \leq \frac{b_k(\ell_1^\beta + \ell_2^\beta)}{kmk} \ln(\ell_2 / \ell_1) \quad \text{a.s.}$$

**Remarks**: We see that consistence of these estimators does not involved existence of an exponential moment for $X_t$ or the knowledge of the second order of regularity, $\beta$. But to optimize their behavior, one has to choose $\delta_{Tn}$ below the threshold $\delta_{Tn}(\beta)$ given by $\delta_{Tn}(\beta) = c (\ln T_n / T_n)^{1/2}$, $c > 0$. Note that if $\delta_{Tn}$ and $\delta_{Tn}(\beta)$ are of same order, the value of the constant is increased but the rate of convergence $O\left((\ln T_n / T_n)^{1/2}\right)$ is preserved. Next for choices of $\delta_{Tn}$ exceeding $\delta_{Tn}(\beta)$, the rate degenerates and becomes $O(\delta_{Tn}^\beta)$. This result emphasizes the part played by the coefficient $\beta$ (when it exists) since it determines the threshold $\delta_{Tn}(\beta)$. We propose and study an estimator of $\beta$ in the following section.

### 3.3 Estimation of $\beta$

Examination of the last result shows that for choices of $\delta_{Tn}$ satisfying to $(\ln T_n / T_n)^{1/2} = O(\delta_{Tn}^\beta)$ the rate and the asymptotic constant obtained in Theorem 3.2 depend on $\beta$. To
derive an estimator of $\beta$ an idea is to consider slowly decreasing $\delta_{\tau_n}$ and then construct a properly normalized difference of estimators of $\gamma$:

$$\hat{\beta} = \frac{\ln |\tilde{\gamma}(\ell_1, \ell_2, k) - \tilde{\gamma}(\ell_3, \ell_4, p)|}{\ln \delta_{\tau_n}}$$

with $k, p$ as before and parameters $\ell_i \; i = 1, \ldots, 4$ such that $\ell_1 < \ell_2$, $\ell_3 < \ell_4$ and $(\ell_1, \ell_2, k) \neq (\ell_3, \ell_4, p)$. The asymptotic behavior is given by Theorem 3.3.

**Theorem 3.3** If assumptions of Theorem 2.1(b) (or Theorem 2.2(b)) and Lemma 2.1(b) are fulfilled, then choices of $\delta_{\tau_n}$ such that

$$\left(\frac{\ln T_n}{T_n}\right)^{1/2} = o(\delta_{\tau_n}), \quad (3.1)$$

imply

$$\lim_{T_n \to \infty} \left(\frac{\ln \delta_{\tau_n}^{-1}}{T_n}\right) (\beta - \hat{\beta}) = \ln \left| \frac{b_k}{km_k \ln(\ell_2/\ell_1)} \left(\ell_2^\beta - \ell_1^\beta\right) - \frac{b_p}{pm_p \ln(\ell_4/\ell_3)} \left(\ell_4^\beta - \ell_3^\beta\right) \right| \text{ a.s.}. \quad (3.1)$$

Rates of convergence appear to be poor, however such estimators are only used to get some idea about the maximal choice of $\delta_{\tau_n}$, namely $\delta_{\tau,0}(\beta)$. We infer that by respecting a safety margin below $\delta_{\tau,0}(\hat{\beta})$, one can except a good behavior of estimators $\hat{\gamma}(\ell_1, \ell_2, k)$.

Another drawback of $\hat{\beta}$ is that, due to its construction, condition (3.1) depend on $\beta$ but remark that there exists an omnibus choice, namely $\delta_{\tau_n} = (\ln T_n)^{-1}$, satisfying to (3.1) for all $\beta > 0$!

### 4 Discussion

#### 4.1 Comparison with the sampling scheme case

Here we point out the advantages of the (theoretical) continuous time framework by recalling briefly the results obtained in (2002) where the following sampling scheme is considered:

$$\{X_i, t \in \mathbb{R}\}$$

is a real-valued process observed over $[0, T_n]$, $T_n > 0$ at sampling times $i\delta_n, \; i = 0, \ldots, n - 1$ where $0 < \delta_n \leq 1$, $T_n = (n - 1)\delta_n$, $\delta_n \to 0$ and $T_n \to \infty$ as $n \to \infty$ i.e. the statistician may dispose of frequent observations with ‘high’ frequency and during a long time.

In this paper, analogues of Theorems 2.1-2.2 are established. The main difference appear in the rate of convergence of the discretized version of statistics $S_{\tau_n, s}(\gamma)$ since $O\left(\left(\ln T_n/T_n\right)^{1/2}\right)$ is replaced by $O\left(\left(\ln n/n\delta_n\right)^{1/2}\right)$. So in this case, there exists an additional constraint upon $\delta_n$: $\delta_n$ not too small, namely $\delta_n^{-1} = o(n)$. This fact has a direct impact for the convergence of estimators. First, if we look at the discretized version of the ‘genuine estimator’, it is shown, in an almost sure sense, that its rate is $\left(\ln(\delta_n^{-1})\right)^{-1}$:
so it cannot be better than \((\ln n)^{-1}\). On the other hand, in our Theorem 3.1, since whole a sample path is observed, one may chose \(\delta_{r_n}\) as small as wanted and the rate \(\ln(\delta_{r_n}^{-1})\) can be improved in consequence.

Moreover for our discretized ‘general estimators’, it appears in (2002) that the parameter \(\beta\) plays a major role for their convergence since one has to choose \(\delta_n\) exactly of order \((\ln n/n\delta_n)^{1/2}\) to obtain the rate \(O((\ln n/n\delta_n)^{1/2})\). So conversely to the continuous time framework where \(\beta\) is needed only to compute the maximal threshold \(\delta_{0,T}\), estimators in the discrete sampling scheme will be much more sensitive to \(\beta\) and its preliminary estimation appears very more crucial than in our Theorem 3.2!

Finally, note that a numerical implementation of estimators appears also in the previous cited paper.

### 4.2 Applications to density estimation in continuous time

Let us begin by recalling some major facts on density estimation in this context. First, in 1986, Castellana and Leadbetter (1986) pointed out the surprising result that the variance of \(\delta\)-sequences estimators was of order \(1/T\) for sufficiently irregular sample paths observed on \([0,T]\). This phenomenon yields many works, we refer e.g. to Bosq (1998b) and references therein for results about kernel estimators. For the important particular case of diffusion processes, Kutoyants (1997a, 1997c) gave the exact asymptotic minimax constant, and a discussion about the Castellana and Leadbetter condition may be found in Veretennikov (1999). We also refer to results of Leblanc (1997) for wavelet estimators, Kutoyants (1997b) and, Bosq and Davydov (1999) for estimators based on local time. Finally regression estimation was investigated by Cheze-Payaud (1994), model with errors-in-variables by Blanke (1996) and projection estimators by Comte and Merlevède (2002a).

Lately, Bosq (1998a) shows that for some special processes the kernel estimator exactly reach the minimax rates of the discrete time case, say \(T^{-r}\) (where \(0 < r < 1\) is connected to the regularity of density). Next, existence and optimality of whole a family of intermediate rates (lying between \(T^{-1}\) and \(T^{-r}\)) are established in Blanke and Bosq (1997, 2000), and it is shown in (2000) that all these rates are minimax in some specific sense. In these papers the key assumption, for \(\mathbb{R}^d\)-valued processes, is made on the joint density of \((X_0, X_u)\) near the origin ; written concisely as \(f(x_0, x_u) \leq Mu^{-\gamma_0}\), \(\gamma_0 > 0\). This parameter \(\gamma_0\) is directly linked with Hölderian properties of sample paths since one may equivalently write \(f(x_0, y_u)(x, 0) \leq M\) or \(f(x_0, z_u)(x, 0) \leq M\) where \(y_u\), \(z_u\) stands respectively for vectors \(y_u = (X_u^{(1)} - X_0^{(1)}, \ldots, X_u^{(d)} - X_0^{(d)})\) (then \(\gamma_0 = \sum_{i=1}^{d} \gamma_i\)) and \(z_u = (X_u^{(1)} - X_0^{(1)}/u^{\gamma_1}, \ldots, X_u^{(d)} - X_0^{(d)}/u^{\gamma_d})\) (then \(\gamma_0 = d\gamma\)). Note that exact asymptotic variance for some special kernel estimators had been obtained by Sköld (2001), Sköld and Hössjer (1999) using conditions on the joint density of \((X_0, Z_u)\).

So in continuous time, optimal choice of bandwidths, sampling schemes - see e.g. works of
Bosq (1997), Bosq and Cheze-Payaud (1999), Blanke and Pumo (2003), Comte and Merlevede (2002b) - directly depend on Hölderian properties of sample paths, and namely on the unknown coefficient $\gamma_0$. Results presented in this paper provide preliminary estimators for $\gamma_0$. For example, one may compute estimators for each index $\gamma_i$ and, on a second step, build a plug-in type density estimator (i.e. with a random bandwidth) using $\hat{\gamma}_0 = \sum_{i=1}^d \hat{\gamma}_i$ (this work being actually in preparation). Finally, alternatively to this ‘plug-in’ approach, an adaptive estimation of the density (using an implicit estimator of $\gamma_0$) is proposed in (2003b) and (2003a).

5 Proofs

5.1 Proof of Theorem 2.1

Let us set
\[ Y_{t,T_n} = \frac{1}{(\ell\delta_{T_n})^{k/\gamma}} \left\{ \|X_{t+\ell\delta_{T_n}} - X_t\|^k - E \|X_{t+\ell\delta_{T_n}} - X_t\|^k \right\}. \]

One may write for any positive $\kappa_1$
\[
P\left( \left| S_{T_n,T}^{(k)}(\gamma) - E S_{T_n,T}^{(k)}(\gamma) \right| > \varepsilon \right) = P\left( \left| \int_0^{T_n-\ell\delta_{T_n}} Y_{t,T_n} dt \right| > (T_n - \ell\delta_{T_n})\varepsilon \right)
\leq P\left( \left| \int_0^{T_n-\ell\delta_{T_n}} \bar{Y}_{t,T_n} dt \right| > \frac{(T_n - \ell\delta_{T_n})\varepsilon}{1 + \kappa_1} \right) + P\left( \left| \int_{2qT_n}^{rT_n} \tilde{Y}_{t,T_n} dt \right| > \frac{(T_n - \ell\delta_{T_n})\varepsilon}{1 + \kappa_1} \right)
\]

(5.1)

where
\[ \bar{Y}_{t,T_n} = \frac{1}{(\ell\delta_{T_n})^{k/\gamma}} \left\{ \left( \|X_{t+\ell\delta_{T_n}} - X_t\|^k - E \|X_{t+\ell\delta_{T_n}} - X_t\|^k \right) I_{\|X_{t+\ell\delta_{T_n}} - X_t\|^k \leq L_{T_n}} \right\} \]

and
\[ \tilde{Y}_{t,T_n} = \frac{1}{(\ell\delta_{T_n})^{k/\gamma}} \left\{ \left( \|X_{t+\ell\delta_{T_n}} - X_t\|^k - E \|X_{t+\ell\delta_{T_n}} - X_t\|^k \right) I_{\|X_{t+\ell\delta_{T_n}} - X_t\|^k > L_{T_n}} \right\} \]

with $L_{T_n}$ such that $L_{T_n} \to +\infty$ as $T_n \uparrow \infty$.

5.1.1 Evaluation of term $P\left( \left| \int_0^{T_n-\ell\delta_{T_n}} \bar{Y}_{t,T_n} dt \right| > \frac{(T_n - \ell\delta_{T_n})\varepsilon}{1 + \kappa_1} \right)$

First, define the blocks:
\[
V_T(j) = \int_{jT}^{(j+1)T} \bar{Y}_{t,T_n} dt \quad \text{and} \quad V_T(2qT) := \int_{2qT}^{2qT+T} \bar{Y}_{t,T_n} dt
\]

(5.2)

with $j = 0, \ldots, 2qT - 1$ and $qT$, $r_T$ such that $2qT - 1 < T_n - \ell\delta_{T_n}$.

Next, set:
\[ S_T' = \sum_{j=0}^{[qT]} V_T(2j) \quad \text{and} \quad S_T'' = \sum_{j=1}^{[qT]} V_T(2j - 1)
\]

(5.3)
then $\left| \int_{0}^{T_n - \ell \delta_{T_n}} Y_{t, T_n} \, dt \right| \leq |S'_T| + |S''_T|$ and for any strictly positive $\varepsilon$,

$$P \left( \left| \int_{0}^{T_n - \ell \delta_{T_n}} Y_{t, T_n} \, dt \right| > \frac{(T_n - \ell \delta_{T_n})\varepsilon}{(1 + \kappa_1)} \right) \leq P \left( \left| S'_T \right| > \frac{(T_n - \ell \delta_{T_n})\varepsilon}{2(1 + \kappa_1)} \right) + P \left( \left| S''_T \right| > \frac{(T_n - \ell \delta_{T_n})\varepsilon}{2(1 + \kappa_1)} \right). \quad (5.4)$$

Let us begin by evaluation of $P (|S'_T| > (T_n - \ell \delta_{T_n})\varepsilon/2(1 + \kappa_1))$.

We use the coupling result (Rio, 1995) recursively to approximate $V_T(0), \ldots, V_T(2[q_T])$ by independent variables. Let $(U_j)_{j \geq 1}$ be a sequence of independent random variables with uniform distribution over $[0, 1]$, independent of $(V_T(j))_{j \geq 1}$. Indeed, if $V_T^*(0) = V_T(0)$ then for any $j \geq 1$, there exists a random variable $V_T^*(2j)$, measurable function of $V_T(0), \ldots, V_T(2j), U_{2j}$ such that $V_T^*(2j)$ is independent of $V_T(0), \ldots, V_T(2j - 2)$ and with same law as $V_T(2j)$. Moreover:

$$E \left| V_T^*(2j) - V_T(2j) \right| \leq 4\|V_T(2j)\|_{\infty} \left( \sup |P(AB) - P(A)P(B)| \right)$$

where the supremum is taken over all sets $A$ and $B$ belonging to $\sigma$-algebras of events generated by respectively $\{V_T(0), \ldots, V_T(2j - 2)\}$ and $V_T(2j)$. We obtain for all positive $\kappa_1, \kappa_2$:

$$P \left( |S'_T| > \frac{(T_n - \ell \delta_{T_n})\varepsilon}{2(1 + \kappa_1)} \right) \leq P \left( \left| \sum_{j=0}^{[q_T]} V_T^*(2j) \right| > \frac{(T_n - \ell \delta_{T_n})\varepsilon}{2(1 + \kappa_1)(1 + \kappa_2)} \right) + P \left( \left| \sum_{j=0}^{[q_T]} V_T(2j) - V_T^*(2j) \right| > \frac{\kappa_2(T_n - \ell \delta_{T_n})\varepsilon}{2(1 + \kappa_1)(1 + \kappa_2)} \right) =: I_1 + I_2. \quad (5.5)$$

### Evaluation of $I_1$

First we have $|Y_{t, T_n}| \leq 2L_{T_n}$ so that for $j = 0, \ldots, 2[q_T] - 1$, $|V_T(j)| \leq 2r_T L_{T_n}$ and $|V_T(2[q_T])| \leq 4r_T L_{T_n}$. We deduce

$$E \left| V_T^*(j) \right|^k \leq (2r_T L_{T_n})^{k-2} E \left( V_T^*(j) \right)^2, \quad j = 0, \ldots, 2[q_T] - 1$$

and

$$E \left| V_T^*(2[q_T]) \right|^k \leq (4r_T L_{T_n})^{k-2} E \left( V_T^*(2[q_T]) \right)^2$$

Using independence of the $V_T^*(2j)$, we may used Bernstein’s inequality (written as in Pollard 1984) to obtain

$$I_1 \leq 2 \exp \left( -\frac{(T_n - \ell \delta_{T_n})^2\varepsilon^2/8(1 + \kappa_1)(1 + \kappa_2)}{(1 + \kappa_1)(1 + \kappa_2) \sum_{j=0}^{[q_T]} E \left( V_T^*(2j) \right)^2 + \frac{2}{3} r_T L_{T_n} (T_n - \ell \delta_{T_n})\varepsilon} \right). \quad (5.6)$$
On the other hand, Davydov’s inequality (1968) applied to $V_T^*(2j)$, for $j = 0, \ldots, [q_T] - 1$, Fubini’s theorem implies:

$$E(V_T^*(2j))^2 = 2 \int_{2jr_T} \int_0^{2jr_T} \int_0^{2jr_T} \int_0^{2jr_T} \text{Cov}(\overline{Y}_{u,v}, \overline{Y}_{v,v}, \overline{Y}_{u,v}) \, dv \, du.$$  

Now, Cauchy-Schwarz implies

$$\nu > 2$$

A change of variables leads to

$$E(V_T^*(2j))^2 = 2 \int_{2jr_T} \int_0^{2jr_T} \text{Cov}(\overline{Y}_{u,v}, \overline{Y}_{u+v,v}) \, dv \, du$$

where

$$\nu = \int_0^{2jr_T} \int_0^{2jr_T} \text{Cov}(\overline{Y}_{u,v}, \overline{Y}_{u+v,v}) \, dv \, du$$

and

$$\int_0^{2jr_T} \int_0^{2jr_T} \text{Cov}(\overline{Y}_{u,v}, \overline{Y}_{u+v,v}) \, dv \, du$$

Next, for $j = [q_T]$,

$$E(V_T^*(2[q_T]))^2 = 2 \int_{2[q_T]r_T} \int_0^{2[q_T]r_T} \text{Cov}(\overline{Y}_{u,v}, \overline{Y}_{u+v,v}) \, dv \, du$$

where

$$\nu = \int_0^{2[q_T]r_T} \int_0^{2[q_T]r_T} \text{Cov}(\overline{Y}_{u,v}, \overline{Y}_{u+v,v}) \, dv \, du$$

and

$$\int_0^{2[q_T]r_T} \int_0^{2[q_T]r_T} \text{Cov}(\overline{Y}_{u,v}, \overline{Y}_{u+v,v}) \, dv \, du$$

Now, Cauchy-Schwarz implies

$$J_1 \leq \frac{\int_0^{2[q_T]r_T} \int_0^{2[q_T]r_T} \text{Cov}(\overline{Y}_{u,v}, \overline{Y}_{u+v,v}) \, dv \, du}{\text{Cov}(\overline{Y}_{u,v}, \overline{Y}_{u+v,v})}$$

and

$$J_1' \leq \frac{\int_0^{2[q_T]r_T} \int_0^{2[q_T]r_T} \text{Cov}(\overline{Y}_{u,v}, \overline{Y}_{u+v,v}) \, dv \, du}{\text{Cov}(\overline{Y}_{u,v}, \overline{Y}_{u+v,v})}$$

On the other hand, Davydov’s inequality (1968) applied to $J_2$ and $J_2'$ implies for any $\nu > 2$:

$$J_2 \leq \frac{\nu - 2}{\nu - 2} \left( \frac{\int_0^{2[q_T]r_T} \int_0^{2[q_T]r_T} \text{Cov}(\overline{Y}_{u,v}, \overline{Y}_{u+v,v}) \, dv \, du}{\text{Cov}(\overline{Y}_{u,v}, \overline{Y}_{u+v,v})} \right)^{2/\nu} \times$$

$$\int_0^{(2j+1)r_T} \int_0^{(2j+1)r_T} \left( \sup_{A \in \sigma \{X_{s+u}, \leq \leq \overline{Y}_{u,v}\}} |P(A \cap B) - P(A)P(B)| \right)^{\nu-2} \, dv \, du,$$
so that:

$$J_2 \leq \frac{2^{\nu-2}}{\nu-2} \sup_{t} \left( E \left\| X_{t+\ell \delta T_n} - X_t \right\|^{k \nu} \right)^{2/\nu} r_T \int_0^{\infty} \alpha_x^{\nu-2}(v) \, dv$$

(5.11)

and by the same way

$$J'_2 \leq 2J_2.$$  

(5.12)

Now recall that $M(p) := \sup \left( E \left\| X_{t+\ell \delta T_n} - X_t \right\|^{p} \right)$. There exists $N_1$ such that for all $n \geq N_1$, $\delta T_n \leq \delta_0$ so condition (2.1) and relations (5.6)-(5.12) imply for all $n \geq N_1$:

$$I_1 \leq 2 \exp \left( -\frac{(T - \ell \delta T_n)\varepsilon^2}{C_{k,\nu,M(k\nu)}(1 + R_1(T_n)) + \frac{16}{3}(1 + \kappa_1)(1 + \kappa_2)r_T L_{T_n} \varepsilon} \right)$$

(5.13)

where we have set

$$C_{k,\nu,M(k\nu)} = \frac{2^{\nu-2}}{\nu-2} (1 + \kappa_1)^2(1 + \kappa_2)^2 M^{\frac{v}{2}}(k\nu) \int_0^{\infty} \alpha_x^{\nu-2}(v) \, dv$$

(5.14)

and $R_1(T_n)$ is a non random function such that $\lim_{T_n \to \infty} R_1(T_n) = 0.$

**Evaluation of $I_2$** Using Rio’s inequality we easily obtain,

$$I_2 \leq \frac{16(1 + \kappa_1)(1 + \kappa_2)}{\kappa_2 \varepsilon} L_{T_n} \alpha_X(r_T - \ell \delta T_n).$$

(5.15)

Equations (5.5), (5.13) and (5.15) imply: for all $n \geq N_1$

$$P \left( \left| S'_{T_n} \right| > \frac{(T_n - \ell \delta T_n)\varepsilon}{2(1 + \kappa_1)} \right) \leq 2 \exp \left( -\frac{(T_n - \ell \delta T_n)\varepsilon^2}{C_{k,\nu,M(k\nu)}(1 + R_1(T_n)) + \frac{16}{3}(1 + \kappa_1)(1 + \kappa_2)r_T L_{T_n} \varepsilon} \right) + \frac{16(1 + \kappa_1)(1 + \kappa_2)}{\kappa_2 \varepsilon} L_{T_n} \alpha_X(r_T - \ell \delta T_n).$$

Now, since $P \left( \left| S''_{T_n} \right| > \frac{(T_n - \ell \delta T_n)\varepsilon}{2(1 + \kappa_1)} \right)$ may be handled in a similar manner, one deduces with (5.4)

$$P \left( \left| \int_{0}^{T_n - \ell \delta T_n} Y_{t,T_n} \, dt \right| > \frac{(T_n - \ell \delta T_n)\varepsilon}{1 + \kappa_1} \right) \leq 4 \exp \left( -\frac{(T_n - \ell \delta T_n)\varepsilon^2}{C_{k,\nu,M(k\nu)}(1 + R_1(T_n)) + \frac{16}{3}(1 + \kappa_1)(1 + \kappa_2)r_T L_{T_n} \varepsilon} \right) + \frac{32(1 + \kappa_1)(1 + \kappa_2)}{\kappa_2 \varepsilon} L_{T_n} \alpha_X(r_T - \ell \delta T_n).$$

(5.16)
5.1.2 Evaluation of term $P\left(\left|\int_0^{T_n-\ell\delta T_n} \widetilde{Y}_{t,T_n} \, dt \right| > \frac{\kappa_1(T_n-\ell\delta T_n)\varepsilon}{1+\kappa_1}\right)$

We have:

$$P\left(\left|\int_0^{T_n-\ell\delta T_n} \widetilde{Y}_{t,T_n} \, dt \right| > \frac{\kappa_1(T_n-\ell\delta T_n)\varepsilon}{1+\kappa_1}\right) \leq \frac{1+\kappa_1}{\kappa_1(T_n-\ell\delta T_n)\varepsilon} \int_0^{T_n-\ell\delta T_n} E |\widetilde{Y}_{t,T_n}| \, dt$$

by Markov’s inequality. Using definition of $\widetilde{Y}_{t,T_n}$ and Cauchy-Schwarz inequality, one gets

$$E |\widetilde{Y}_{t,T_n}| \leq \left(E \left\| \frac{X_{t+\ell\delta T_n} - X_t}{(\ell\delta T_n)^{\gamma}} \right\|^{2k}\right)^{1/2} \left(P \left(\left\| \frac{X_{t+\ell\delta T_n} - X_t}{(\ell\delta T_n)^{\gamma}} \right\|^{k} > L_{T_n}\right)\right)^{1/2}$$

and finally,

$$P\left(\left|\int_0^{T_n-\ell\delta T_n} \widetilde{Y}_{t,T_n} \, dt \right| > \frac{\kappa_1(T_n-\ell\delta T_n)\varepsilon}{1+\kappa_1}\right) \leq \left(1+\frac{\kappa_1}{\kappa_1(T_n-\ell\delta T_n)\varepsilon}\right) \int_0^{T_n-\ell\delta T_n} \left(P \left(\left\| \frac{X_{t+\ell\delta T_n} - X_t}{(\ell\delta T_n)^{\gamma}} \right\|^{k} > L_{T_n}\right)\right)^{1/2} \, dt. \quad (5.17)$$

5.1.3 Proof of part (a)

Here we have only to establish that for any $\varepsilon > 0$, $\sum_{n=1}^{+\infty} P \left(\left|S^{(k)}_{t,T_n}(\gamma) - E S^{(k)}_{t,T_n}(\gamma)\right| > \varepsilon\right)$ is convergent. First, in (5.16) set $L_{T_n} = T_n^{-m}$ with $1 > m > 2/\nu$ with $\nu$ given in assumption (2.1), $r_{T_n} = \eta \ln T_n$ with $\eta > \frac{m}{\ln(1/\rho)}$. These choices imply existence of constants $\alpha_0 > 0$, $\beta_0 > 1$ such that for all $n > N_1$

$$P\left(\left|\int_0^{T_n-\ell\delta T_n} Y_{t,T_n} \, dt \right| > \frac{(T_n-\ell\delta T_n)\varepsilon}{1+\kappa_1}\right) \leq \alpha_0 \exp(-\beta_0 \ln T_n).$$

On the other hand, Markov’s inequality applied to (5.17) yields

$$P\left(\left|\int_0^{T_n-\ell\delta T_n} \widetilde{Y}_{t,T_n} \, dt \right| > \frac{\kappa_1(T_n-\ell\delta T_n)\varepsilon}{1+\kappa_1}\right) \leq \frac{1+\kappa_1}{\kappa_1} \varepsilon^{-1} \left(M(2k)M(k\nu)\right)^{1/2} L_{T_n}^{-\nu/2} \leq \alpha_1 \exp(-\beta_1 \ln T_n)$$

with $n > N_1$, $\alpha_0 > 0$ and $\beta_1 > 1$ since $m > 2/\nu$. Now, assumption $T_{n+1} - T_n \geq a > 0$ implies that for all $\beta > 1$, $\sum_{n} T_n^{-\beta} < +\infty$, hence the result holds with the help of Borel Cantelli’s Lemma.

5.1.4 Proof of part (b)

Consider the following choices:

$$\varepsilon = \varepsilon_{r_{T_n}} = \gamma_0 \sqrt{\frac{\ln T_n}{T_n}}, \quad r_{T_n} = \eta \ln T_n, \quad L_{T_n} = \eta_2 (\ln T_n)^k$$
with positive constants $\eta_0, \eta_1, \eta_2$.

First, note that for all positive reals $a$ and $p$

$$\sup_t E \left( e^{\frac{X_{t+\delta T_n} - X_t}{(\delta T_n)^\gamma}} \right) \geq \sup_t \frac{a^p}{\Gamma(p+1)} E \left\| \frac{X_{t+\delta T_n} - X_t}{(\delta T_n)^\gamma} \right\|^p,$$

so for all $n \geq N_1$, $\delta T_n \leq \delta_0$ yields

$$M(p) \leq a^{-p}\Gamma(p+1)M. \tag{5.18}$$

Now constant $C_{k,\nu,M}(k\nu)$ given in (5.14) is bounded by

$$C_{k,\nu,M} = 2 \frac{\nu-2\nu}{\nu} \left( 1 + \kappa_1 \right) \left( 1 + \kappa_2 \right) \left( a^{-k\nu}\Gamma(k\nu + 1)M \right)^{\frac{1}{2}} \int_0^{+\infty} \frac{\nu-2\nu}{\alpha x^\nu} (v) \, dv \tag{5.19}$$

First, choose $\eta_0 > C_{1/2}^{1/2}$ and $\eta_1 > \frac{3}{2\ln(1/\rho)}$, this implies existence of constants $\alpha_0 > 0$ and $\beta_0 > 1$ such that for all $n > N_1$,

$$P \left( \left\| \int_0^{T_n-\ell T_n} \tilde{Y}_{t,T_n} \, dt \right\| > (T_n - \ell T_n)\varepsilon T_n \right) \leq \alpha_0 \exp(-\beta_0 \ln T_n).$$

On the other hand, (5.18) and Markov inequality applied to (5.17) yield

$$P \left( \left\| \int_0^{T_n-\ell T_n} \tilde{Y}_{t,T_n} \, dt \right\| > \frac{\kappa_1(T_n - \ell T_n)\varepsilon T_n}{1+\kappa_1} \right)
\leq \frac{(1+\kappa_1)a^{-k}\Gamma(2k+1)^{1/2}M}{\kappa_1\eta_0} \frac{T_n^{1/2}}{(\ln T_n)^{1/2}} e^{-\frac{a\eta_1}{2\ln T_n}} \leq \alpha_1 \exp(-\beta_1 \ln T_n)$$

with $n > N_1$, $\alpha_1 > 0$ and $\beta_1 > 1$ as soon as $\eta_2 > (3/a)^k$.

Finally Borel-Cantelli’s lemma leads to

$$\limsup_{T_n \to \infty} \sqrt{\frac{T_n}{\ln T_n}} \left| S_{T_n,t,(\gamma)}^{(k)} - E S_{T_n,t,(\gamma)}^{(k)} \right| \leq C_1 \text{ a.s.} \tag{5.20}$$

where $C_1 = \lim_{\kappa \to 0} \lim_{\kappa \to 10} \sqrt{C_{k,\nu,M}}$ with $C_{k,\nu,M}$ given by (5.19).

### 5.2 Proof of Theorem 2.2

Proof is essentially the same as for Theorem 2.1, since upper bounds established in (5.16) and (5.17) are still valid. Since the process is only assumed to be A.S.M. (with coefficient $b$), optimal choices of $L_{T_n}$ and $r_{T_n}$ should be adapted in consequence.
5.2.1 Proof of part (a)

To obtained the desired result, relations (5.16) and (5.17) lead to choices of $L_{\tau_n}$ and $r_{\tau_n}$ satisfying to

$$
\begin{cases}
\frac{r_{\tau_n} L_{\tau_n}}{T_n^{\nu/2}} \to 0 \\
\frac{T_n^{-\nu} L_{\tau_n}}{T_n^{1/2}} = T_n^{-c_1} \quad \text{with } c_1 > 1 \\
L_{\tau_n}^{-\nu/2} = T_n^{-c_2} \quad \text{with } c_2 > 1.
\end{cases}
$$

If \( \{X_t\} \) is A.S.M. with coefficient \( b \) such that \( b > \frac{\nu}{\nu - 2} \) with \( \nu > 2 \), namely \( b = \frac{\nu}{\nu - 2} + \varepsilon \) with \( \varepsilon > 0 \) then adequate choices should be \( L_{\tau_n} = T_n^{\xi_1/(2/\nu)} \) and \( r_{\tau_n} = T_n^{1-\xi_2/(2/\nu)} \) with

$$
\frac{\varepsilon(\nu - 2)^2}{\nu(\nu + 2) + \nu\varepsilon(\nu - 2)} > \xi_2 > \xi_1 > 0.
$$

5.2.2 Proof of part (b)

If we set \( \nu = 4 \) in (5.16), optimal choices of \( \varepsilon_{\tau_n}, L_{\tau_n} \) and \( r_{\tau_n} \) should satisfy:

$$
\begin{cases}
\varepsilon_{\tau_n} = \eta_0 \left( \ln T_n/T_n \right)^{1/2} \quad \text{with } \eta_0 > C_{k,\nu,M}^{1/2} \\
L_{\tau_n} = \eta_1 \left( \ln T_n \right)^k \quad \text{with } \eta_1 > (3/a)^k \\
r_{\tau_n}(\ln T_n)^{k+1/2}T_n^{-1/2} \to 0 \quad \text{as } T_n \to \infty \\
T_n^{1/2}r_{\tau_n}^{-b} = T_n^{-c} \quad \text{with } c > 1.
\end{cases}
$$

These conditions are satisfied if one choose e.g. \( r_{\tau_n} = T_n^m \) with \( \frac{3}{2b} < m < 1/2 \). This implies the condition \( b > 3 \). Remark that \( b > \nu/(\nu - 2) \) is also fulfilled since \( \nu = 4 \).

5.3 Proof of Lemma 2.1

By definition \( E S_{\tau_n}^{(k)}(\gamma) = \frac{1}{T_n - \ell\delta_{\tau_n}} \int_0^{T_n - \ell\delta_{\tau_n}} E \left| \frac{X_{t+\ell\delta_{\tau_n}} - X_t}{(\ell\delta_{\tau_n})^\gamma} \right|^k \, dt = m_k + R_{\tau_n}^{(k)}(\ell\delta_{\tau_n}) \), so Lemma 2.1 easily follows from the assumptions.

5.4 Proof of Corollary 2.1

1) For the first part,

$$
\left| S_{\tau_n,\ell}(\gamma) - m_k \right| \leq \left| S_{\tau_n,\ell}(\gamma) - E S_{\tau_n,\ell}(\gamma) \right| + \left| E S_{\tau_n,\ell}(\gamma) - m_k \right|
$$

and result comes from Theorem 2.1(a) (or Theorem 2.2(a)) and Lemma 2.1(a).

2) For \( \lambda \neq \gamma \), one has just to write \( S_{\tau_n,\ell}(\lambda) = (\ell\delta_{\tau_n})^{k(\gamma-\lambda)} S_{\tau_n,\ell}(\gamma) \),

- if \( \lambda > \gamma \), \( \delta_{\tau_n}^{k(\gamma-\lambda)} \to +\infty \) and \( S_{\tau_n,\ell}(\gamma) \to m_k \) a.s. as \( T_n \uparrow \infty \). Since \( m_k \neq 0 \), \( S_{\tau_n,\ell}(\lambda) \to +\infty \) a.s. as \( T_n \uparrow \infty \).
- if \( \lambda < \gamma \), \( \delta_{\tau_n}^{k(\gamma-\lambda)} \to 0 \) and \( S_{\tau_n,\ell}(\gamma) \to m_k \) a.s. as \( T_n \uparrow \infty \). Since \( m_k \neq +\infty \), \( S_{\tau_n,\ell}(\lambda) \to 0 \) a.s. as \( T_n \uparrow \infty \).


5.5 Proof of Theorem 3.1

$\tilde{\gamma}^{(\ell,k)}$ may be written under its equivalent form:

$$\tilde{\gamma}^{(\ell,k)} = \gamma + \frac{\ln S_{r_n,\ell}(\gamma)}{k\ln(\ell \delta_{r_n})}. \quad (5.21)$$

Corollary 2.1 implies: $\ln S^{(k)}_{r_n,\ell}(\gamma) \rightarrow \ln(m_k)$ a.s. as $T_n \uparrow \infty$ hence,

$$\ln \left( \frac{1}{\delta_{r_n}} \right) (\gamma - \tilde{\gamma}^{(\ell,k)}) \rightarrow \ln(m_k) \text{ a.s. as } T_n \uparrow \infty. \blacksquare$$

5.6 Proof of Theorem 3.2

Using decomposition (5.21) we obtain

$$k \ln(\ell_2/\ell_1) \left( \tilde{\gamma}^{(\ell_1,\ell_2,k)} - \gamma \right) = \ln S_{r_n,\ell_2}(\gamma) - \ln S_{r_n,\ell_1}(\gamma)$$

$$= \left| \ln \left( 1 + \frac{S_{r_n,\ell_2}(\gamma) - m_k}{m_k} \right) - \ln \left( 1 + \frac{S_{r_n,\ell_1}(\gamma) - m_k}{m_k} \right) \right|. \quad (5.22)$$

Theorem 2.1(b) (or Theorem 2.2(b)) together with Lemma 2.1(b) implies $S_{r_n,\ell}(\gamma) - m_k \rightarrow 0$ a.s. as $T_n \uparrow \infty$. So for large enough $T_n$, we observe that

$$k \ln(\ell_2/\ell_1) \left( \tilde{\gamma}^{(\ell_1,\ell_2,k)} - \gamma \right)$$

$$= \frac{S_{r_n,\ell_2}(\gamma) - S_{r_n,\ell_1}(\gamma)}{m_k} + o \left( S_{r_n,\ell_1}(\gamma) - m_k \right) + o \left( S_{r_n,\ell_1}(\gamma) - m_k \right)$$

$$= \frac{1}{m_k} \left\{ \left( S_{r_n,\ell_2}(\gamma) - ES_{r_n,\ell_2}(\gamma) \right) - \left( S_{r_n,\ell_1}(\gamma) - ES_{r_n,\ell_1}(\gamma) \right) + \left( ES_{r_n,\ell_2}(\gamma) - m_k \right) \right\}$$

$$- \left( ES_{r_n,\ell_1}(\gamma) - m_k \right) + o \left( S_{r_n,\ell_1}(\gamma) - m_k \right)$$

$$\leq \frac{1}{m_k} \left\{ \left| S_{r_n,\ell_2}(\gamma) - ES_{r_n,\ell_2}(\gamma) \right| + \left| S_{r_n,\ell_1}(\gamma) - ES_{r_n,\ell_1}(\gamma) \right| + \left| ES_{r_n,\ell_2}(\gamma) - m_k \right| \right\}$$

$$+ o \left( \left| S_{r_n,\ell_1}(\gamma) - m_k \right| \right) + o \left( \left| S_{r_n,\ell_1}(\gamma) - m_k \right| \right) \text{ (a.s.).}$$

Under assumptions of Theorem 2.1(b) or Theorem 2.2(b), and for $\ell = \ell_1$ or $\ell = \ell_2$:

$$\limsup_{T_n \uparrow \infty} \sqrt{\frac{T_n}{\ln T_n}} \left| S_{r_n,\ell}(\gamma) - ES_{r_n,\ell}(\gamma) \right| \leq C_1 \text{ a.s.}$$

In addition, if assumptions of Lemma 2.1(b) hold true,

$$\lim_{T_n \uparrow \infty} \delta_{r_n}^{\beta} \left| ES_{r_n,\ell}(\gamma) - m_k \right| = b_k \ell^\beta$$

and balancing with the stochastic term, we get for all $\delta_{r_n}$ such that $\delta_{r_n} = o \left( \left( \ln T_n/T_n \right)^{2/3} \right)$

$$\limsup_{T_n \uparrow \infty} \sqrt{\frac{T_n}{\ln T_n}} \left( \tilde{\gamma}^{(\ell_1,\ell_2,k)} - \gamma \right) \leq \frac{2C_1}{km_k \ln(\ell_2/\ell_1)} \text{ a.s.} \blacksquare$$
5.7 Proof of Theorem 3.3

We use again decomposition (5.22) to write \( \hat{\beta} \) under its equivalent form:

\[
\hat{\beta} = \left( \ln \left| \frac{\ln S_{T_n, \ell_2}^{(k)}(\gamma) - \ln S_{T_n, \ell_1}^{(k)}(\gamma)}{k \ln(\ell_2/\ell_1)} - \frac{\ln S_{T_n, \ell_3}^{(p)}(\gamma) - \ln S_{T_n, \ell_3}^{(p)}(\gamma)}{p \ln(\ell_4/\ell_3)} \right| \right) / (\ln \delta_{T_n}).
\]

But under assumptions of Theorem 2.1(b) (or Theorem 2.2(b)), \( S_{T_n, \ell}^{(k)}(\gamma) - E S_{T_n, \ell}^{(k)}(\gamma) \to 0 \) a.s. as \( T_n \uparrow \infty \). Moreover if conditions of Lemma 2.1(b) hold true, \( E S_{T_n, \ell}^{(k)}(\gamma) \to m_k \) as \( T_n \uparrow \infty \) with \( 0 < m_k < +\infty \), hence for \( T_n \) large enough:

\[
\ln S_{T_n, \ell}^{(k)}(\gamma) = \ln(m_k) + \frac{S_{T_n, \ell}^{(k)}(\gamma) - E S_{T_n, \ell}^{(k)}(\gamma)}{m_k} + \frac{E S_{T_n, \ell}^{(k)}(\gamma) - m_k}{m_k} + o(S_{T_n, \ell}^{(k)}(\gamma) - m_k).
\]

This implies

\[
(\ln \delta_{T_n}) (\hat{\beta} - \beta) = \left\{ \ln \left| \begin{array}{c}
\delta_{T_n}^{-\beta} \left( \frac{S_{T_n, \ell_2}^{(k)}(\gamma) - E S_{T_n, \ell_2}^{(k)}(\gamma)}{k m_k \ln(\ell_2/\ell_1)} \right) + \delta_{T_n}^{-\beta} \left( \frac{S_{T_n, \ell_3}^{(p)}(\gamma) - E S_{T_n, \ell_3}^{(p)}(\gamma)}{p m_p \ln(\ell_4/\ell_3)} \right)
- \delta_{T_n}^{-\beta} \left( \frac{S_{T_n, \ell_1}^{(k)}(\gamma) - E S_{T_n, \ell_1}^{(k)}(\gamma)}{k m_k \ln(\ell_2/\ell_1)} \right) - \delta_{T_n}^{-\beta} \left( \frac{S_{T_n, \ell_3}^{(p)}(\gamma) - E S_{T_n, \ell_3}^{(p)}(\gamma)}{p m_p \ln(\ell_4/\ell_3)} \right)
+ \delta_{T_n}^{-\beta} \left( \frac{E S_{T_n, \ell_2}^{(k)}(\gamma) - m_k}{k m_k \ln(\ell_2/\ell_1)} \right) + \delta_{T_n}^{-\beta} \left( \frac{E S_{T_n, \ell_3}^{(p)}(\gamma) - m_p}{p m_p \ln(\ell_4/\ell_3)} \right)
- \delta_{T_n}^{-\beta} \left( \frac{E S_{T_n, \ell_1}^{(k)}(\gamma) - m_k}{k m_k \ln(\ell_2/\ell_1)} \right) - \delta_{T_n}^{-\beta} \left( \frac{E S_{T_n, \ell_3}^{(p)}(\gamma) - m_p}{p m_p \ln(\ell_4/\ell_3)} \right)
\end{array} \right| \right\} (a.s.).
\]

Under assumptions of Theorem 2.1(b) (or Theorem 2.2(b)) and with the choice \( \delta_{T_n}^{-\beta} \sqrt{\ln T_n / T_n} \to 0 \), we get, as \( T_n \uparrow \infty \),

\[
\begin{cases}
\delta_{T_n}^{-\beta} \left| S_{T_n, \ell}^{(k)}(\gamma) - E S_{T_n, \ell}^{(k)}(\gamma) \right| \to 0 \text{ a.s.}
\delta_{T_n}^{-\beta} \left( E S_{T_n, \ell}^{(k)}(\gamma) - m_k \right) \to b_k \ell^\beta.
\end{cases}
\]

We conclude by

\[
\lim_{T_n \to \infty} \ln \delta_{T_n}^{-1} (\beta - \hat{\beta}) = \ln \left| \frac{b_k}{k m_k \ln(\ell_2/\ell_1)} (\ell_2^{\beta} - \ell_1^{\beta}) - \frac{b_p}{p m_p \ln(\ell_4/\ell_3)} (\ell_4^{\beta} - \ell_3^{\beta}) \right| \text{ a.s.} \quad \blacksquare
\]

References


Comte, F. and Merlevède, F.: 2002b, Density estimation for a class of continuous time or discretely observed processes. Submitted.


