Available online at www.sciencedirect.com
ScienceDirect
C. R. Acad. Sci. Paris, Ser. I 346 (2008) 901–906
COMPTES RENDUS
MATHEMATIQUE
http://france.elsevier.com/direct/CRASS1/

Statistics

# Asymptotic normality of the additive regression components for continuous time processes

Mohammed Debbarh, Bertrand Maillot

*L.S.T.A., Université de Paris 6, 175, rue du Chevaleret, 75013 Paris, France*

Received 29 August 2007; accepted after revision 27 June 2008

Available online 5 August 2008

Presented by Paul Deheuvels

**Abstract**

In multivariate regression estimation, the rate of convergence depends on the dimension of the regressor. This fact, known as the *curse of the dimensionality*, motivated several works. The additive model, introduced by Stone [C.J. Stone, Additive regression and other nonparametric models, Ann. Statist. 13 (2) (1985) 689–705. [9]], offers an efficient response to this problem. In the setting of continuous time processes, using the marginal integration method, we obtain the quadratic convergence rate and the asymptotic normality of the components of the additive model. ***To cite this article: M. Debbarh, B. Maillot, C. R. Acad. Sci. Paris, Ser. I 346 (2008).***


**Résumé**

**Normalité asymptotique des composantes d'un modèle additif de régression dans le cas de processus en temps continu.** Dans l'estimation de la régression multivariée, la vitesse de convergence dépend de la dimension du régresseur. Ce phénomène, connu sous le nom de *fléau de la dimension*, a motivé plusieurs travaux. Le modèle additif, introduit par Stone [C.J. Stone, Additive regression and other nonparametric models, Ann. Statist. 13 (2) (1985) 689–705. [9]], propose une réponse à ce problème. Dans le cadre des processus à temps continu, nous utilisons la méthode d'intégration marginale pour obtenir la vitesse de convergence quadratique et la normalité asymptotique des composantes additives. ***Pour citer cet article : M. Debbarh, B. Maillot, C. R. Acad. Sci. Paris, Ser. I 346 (2008).***


## Version française abrégée

Soit $(\mathbf{X}_t, Y_t)_{(t\in\mathbb{R})}$ un processus stationnaire à temps continu défini dans l'espace probabilisé $(\Omega, \mathcal{A}, P)$, à valeurs dans $\mathbb{R}^d \times \mathbb{R}$ et observé pour $t \in [0, T]$. Soit, par ailleurs, $\psi$ une fonction mesurable à valeurs dans $\mathbb{R}$. Nous nous intéressons à la fonction de régression additive définie par

$$m_\psi(\mathbf{x}) = E\big(\psi(Y) \mid \mathbf{X} = \mathbf{x}\big) := \mu + \sum_{l=1}^{d} m_l(x_l), \quad \forall \mathbf{x} = (x_1, \dots, x_d) \in \mathcal{C}^\delta,$$

*E-mail addresses:* mohammed.debbarh@upmc.fr (M. Debbarh), bertrand.maillot@upmc.fr (B. Maillot).

1631-073X/$ – see front matter 
doi:10.1016/j.crma.2008.06.012

où $\mathcal{C}^\delta = \{\mathbf{x}\colon \inf_{\mathbf{z}\in\mathcal{C}} \|\mathbf{x}-\mathbf{z}\|_{\mathbb{R}^d} < \delta\}$, est le $\delta$-voisinage du compact $\mathcal{C}$ de $\mathbb{R}^d$ et $\|\cdot\|_{\mathbb{R}^d}$ est la norme euclidienne sur $\mathbb{R}^d$. Pour $1 \leqslant l \leqslant d$, posons $\mathbf{x}_{-l} = (x_1, \ldots, x_{l-1}, x_{l+1}, \ldots, x_d)$ où $\mathbf{x} = (x_1, \ldots, x_d) \in \mathcal{C}^\delta$. Pour estimer les composantes additives, nous utilisons la méthode d'intégration marginale (voir Linton et Nielsen [7] et Newey [8]). Pour cela, introduisons $d$ densités $q_1, \ldots, q_d$, $k$-fois dérivables sur $\mathbb{R}$ et posons $q(\mathbf{x}) = \prod_{l=1}^d q_l(x_l)$ et $q_{-l}(\mathbf{x}_{-l}) = \prod_{j\neq l} q_j(x_j)$, $\{l = 1, \ldots, d\}$. Nous pouvons alors écrire

$$m_\psi(\mathbf{x}) = \sum_{l=1}^d \eta_l(x_l) + \int_{\mathbb{R}^d} m_\psi(\mathbf{z}) q(\mathbf{z})\,\mathrm{d}\mathbf{z},$$

avec $\eta_l(x_l) := \int_{\mathbb{R}^{d-1}} m_\psi(\mathbf{x}) q_{-l}(\mathbf{x}_{-l})\,\mathrm{d}\mathbf{x}_{-l} - \int_{\mathbb{R}^d} m_\psi(\mathbf{x}) q(\mathbf{x})\,\mathrm{d}\mathbf{x} = m_l(x_l) - \int_{\mathbb{R}} m_l(z) q_l(z)\,\mathrm{d}z$, $l = 1, \ldots, d$.

La normalité asymptotique de l'estimateur de la régression multivariée pour les processus à temps continu a été obtenue par Cheze-Payaud [5], avec une normalisation dépendant de la dimension $d$ de la covariable $\mathbf{X}$. Cela justifie notre étude dans le cadre des modèles additifs. Nous traitons alors dans cette Note la convergence en moyenne quadratique puis établissons la normalité asymptotique des estimateurs des composantes de la fonction de régression additive, comme cela fut fait par Camlong et al. [4] pour les processus à temps discret faiblement mélageants. Soit $\hat{\eta}_{l,T}(x_l)$, défini en (7), l'estimateur de $\eta_l(x_l)$, et $Q_\alpha$ le quantile d'ordre $\alpha$ d'une loi normale centrée réduite. Nous obtenons aussi sous des conditions moins contraignantes le résultat suivant, pour tout $(\alpha, \beta) \in ]0; 0,5[ \times ]0,5; 1[$,

$$\liminf_{T\to\infty} P\left(\left\{T^{\frac{k}{2k+1}}\left(\hat{\eta}_{l,T}(x_l) - \eta_l(x_l)\right)\right\} - b_l(x_l) \in [AQ_\alpha; AQ_\beta]\right) \geqslant \beta - \alpha, \tag{1}$$

où

$$A^2 = \limsup_{T\to+\infty} T^{\frac{2k}{2k+1}} \operatorname{Var}\left(\hat{\eta}_{l,T}(x_l)\right), \qquad b_l(x_l) = \frac{c_1^k}{k!} \int_{\mathbb{R}} u^k K_l(u)\,\mathrm{d}u \left((-1)^k m_l^{(k)}(x_l) - \int_{\mathbb{R}} m_l(z) q_l^{(k)}(z)\,\mathrm{d}z\right),$$

avec $c_1$ une constante et $K_l$ un noyau à valeur dans $\mathbb{R}$.

## 1. Introduction

Let $\mathbf{Z}_t = (\mathbf{X}_t, Y_t)_{(t\in\mathbb{R})}$ be a $\mathbb{R}^d \times \mathbb{R}$-valued measurable stationary process defined on a probability space $(\Omega, \mathcal{A}, P)$ and observed for $t \in [0, T]$. Let $\mathcal{C}_1, \ldots, \mathcal{C}_d$, be $d$ compact intervals of $\mathbb{R}$ and set $\mathcal{C} = \mathcal{C}_1 \times \cdots \times \mathcal{C}_d$. Set now $\delta > 0$ and introduce the $\delta$-neighborhood $\mathcal{C}^\delta$ of $\mathcal{C}$, namely $\mathcal{C}^\delta = \{\mathbf{x}\colon \inf_{\mathbf{z}\in\mathcal{C}} \|\mathbf{x}-\mathbf{z}\|_{\mathbb{R}^d} < \delta\}$, with $\|\cdot\|_{\mathbb{R}^d}$ standing for the Euclidean norm on $\mathbb{R}^d$. Let $\psi$ be a real valued measurable function. Consider the regression function $m_\psi$ defined by,

$$m_\psi(\mathbf{x}) = E\left(\psi(Y) \mid \mathbf{X} = \mathbf{x}\right), \quad \forall \mathbf{x} = (x_1, \ldots, x_d) \in \mathcal{C}^\delta. \tag{2}$$

Let $K$ be a kernel defined on $\mathbb{R}^d$ and having a compact support. Let $\hat{f}_T$ be the estimate of $f$, the density function of the covariable $\mathbf{X}$ (see Banon [1]), defined by,

$$\hat{f}_T(\mathbf{x}) = \frac{1}{T h_T^d} \int_0^T K\left(\frac{\mathbf{x} - \mathbf{X}_s}{h_T}\right) \mathrm{d}s,$$

where $h_T$ is a positive parameter. In the sequel, to estimate the regression function defined in (2), we use the following estimator (see, for example, Bosq [3] and Jones et al. [6])

$$\tilde{m}_{\psi,T}(\mathbf{x}) = \int_0^T W_{T,t}(\mathbf{x}) \psi(\mathbf{Y}_t)\,\mathrm{d}t \quad \text{with } W_{T,t}(\mathbf{x}) = \frac{\prod_{l=1}^d \frac{1}{h_{l,T}} K_l\left(\frac{\mathbf{x}_l - \mathbf{X}_t}{h_{l,T}}\right)}{T \hat{f}_T(\mathbf{X}_t)}, \tag{3}$$

where $(h_{j,T})_{1\leqslant j\leqslant d}$ are positive parameters and $(K_l)_{1\leqslant j\leqslant d}$ are $d$ kernels defined on $\mathbb{R}$ with compact supports. Consider now that the nonparametric regression function (2) may be written as a sum of univariate functions, i.e.

$$m_\psi(\mathbf{x}) \equiv \mu + \sum_{l=1}^d m_l(x_l) =: m_{\psi,\mathrm{add}}(\mathbf{x}), \quad \forall \mathbf{x} = (x_1, \ldots, x_d) \in \mathcal{C}^\delta, \tag{4}$$

where, for $1 \leqslant l \leqslant d$, $Em_l(X_l) = 0$. For $1 \leqslant l \leqslant d$ and any $\mathbf{x} = (x_1, \dots, x_d) \in \mathcal{C}^\delta$ set $\mathbf{x}_{-l} = (x_1, \dots, x_{l-1}, x_{l+1}, \dots, x_d)$. To estimate the additive components, we use the marginal integration method (see Linton and Nielsen [7] and Newey [8]). To this aim, we introduce $d$ densities $q_1, \dots, q_d$ defined on $\mathbb{R}$ and set $q(\mathbf{x}) = \prod_{l=1}^d q_l(x_l)$ and $q_{-l}(\mathbf{x}_{-l}) = \prod_{j \neq l} q_j(x_j)$, with $l = 1, \dots, d$. We can then write

$$m_\psi(\mathbf{x}) = \sum_{l=1}^d \eta_l(x_l) + \int_{\mathbb{R}^d} m_\psi(\mathbf{z}) q(\mathbf{z}) \, \mathrm{d}\mathbf{z}, \tag{5}$$

with

$$\eta_l(x_l) := \int_{\mathbb{R}^{d-1}} m_\psi(\mathbf{x}) q_{-l}(\mathbf{x}_{-l}) \, \mathrm{d}\mathbf{x}_{-l} - \int_{\mathbb{R}^d} m_\psi(\mathbf{x}) q(\mathbf{x}) \, \mathrm{d}\mathbf{x} = m_l(x_l) - \int_{\mathbb{R}} m_l(z) q_l(z) \, \mathrm{d}z, \quad 1 \leqslant l \leqslant d. \tag{6}$$

Making use of the statements (3) and (6), it follows that a natural estimate of the $l$th component is given by

$$\hat{\eta}_{l,T}(x_l) = \int_{\mathbb{R}^{d-1}} \tilde{m}_{\psi,T}(\mathbf{x}) q_{-l}(\mathbf{x}_{-l}) \, \mathrm{d}\mathbf{x}_{-l} - \int_{\mathbb{R}^d} \tilde{m}_{\psi,T}(\mathbf{x}) q(\mathbf{x}) \, \mathrm{d}\mathbf{x}, \quad 1 \leqslant l \leqslant d. \tag{7}$$

## 2. Hypotheses and notations

In order to state our results, we introduce some assumptions and additional notations:

(C.1) There exists a positive constant $M$ such that, for any $y \in \mathbb{R}$, $|\psi(y)| \leqslant M < \infty$,

(C.2) $m_\psi$ is a $k$-times continuously differentiable function, $k \geqslant 1$, and $\sup_{\mathbf{x}} |\frac{\partial^k m_\psi}{\partial x_l^k}(\mathbf{x})| < \infty$, $\forall l \in \{1, \dots, d\}$.

For $1 \leqslant l \leqslant d$, we denote by $f_l$, the density function of $X_l$ and we suppose that the functions $f$ and $f_l$ are continuous and bounded. We need the additional conditions:

(F.1) $\forall \mathbf{x} \in \mathcal{C}^\delta$, $f(\mathbf{x}) > 0$ and $f_l(x_l) > 0$, $l = 1, \dots, d$,

(F.2) $f$ is $k'$-times continuously differentiable on $\mathcal{C}^\delta$, $k' > kd$,

(F.3) for some $0 < \lambda \leqslant 1$, $|\frac{\partial f^{(k')}}{\partial x_1^{j_1} \cdots \partial_d^{j_d}}(\mathbf{x}') - \frac{\partial f^{(k')}}{\partial x_1^{j_1} \cdots \partial_d^{j_d}}(\mathbf{x})| \leqslant L \|\mathbf{x}' - \mathbf{x}\|_{\mathbb{R}^d}^\lambda$ with $j_1 + \dots + j_d = k'$, and $L$ a positive constant. We use the notation $r := k' + \lambda$.

The kernels $K$ and $K_j$, $1 \leqslant j \leqslant d$ are assumed to fulfill the following conditions:

(K.1) For $1 \leqslant j \leqslant d$, $K$ and $K_j$ are continuous on compact supports $\mathcal{S}$ and $\mathcal{S}_j \subset \mathcal{C}_j$, respectively,

(K.2) $\int K = 1$ and $\int K_j = 1$, $1 \leqslant j \leqslant d$,

(K.3) $\prod_{j=1}^d K_j$ is of order $k$,

(K.4) $K$ is of order $k'$ and is Lipschitzian.

The known integration density functions $q_l$, $1 \leqslant l \leqslant d$, satisfy the following assumption:

(Q.1) $q_l$ has $k$ continuous and bounded derivatives, with compact support included in $\mathcal{C}_l$, $1 \leqslant l \leqslant d$.

There exists $\Gamma \in \mathcal{B}_{\mathbb{R}^2}$ containing $D = \{(s,t) \in \mathbb{R}^2 \colon s = t\}$ such that:

(D.1) $f_{(\mathbf{X}_s, Y_s),(\mathbf{X}_t, Y_t)} - f_{(\mathbf{X}_s, Y_s)} \otimes f_{(\mathbf{X}_t, Y_t)}$ exists everywhere for $(s,t) \in \Gamma^C$, where $\Gamma^C$ is the complement of $\Gamma$,

(D.2) $A_\Gamma := \sup_{(s,t) \in \Gamma^C} \sup_{\mathbf{x},\mathbf{y} \in \mathcal{C}^\delta \times \mathcal{C}^\delta} \int_{u,v \in \mathbb{R}^2} |f_{(\mathbf{X}_s, Y_s),(\mathbf{X}_t, Y_t)}(\mathbf{x}, u, \mathbf{y}, v) - f_{(\mathbf{X}_s, Y_s)}(\mathbf{x}, u) f_{(\mathbf{X}_t, Y_t)}(\mathbf{y}, v)| \, \mathrm{d}u \, \mathrm{d}v < \infty$,

(D.3) there exists $\ell_\Gamma < \infty$ and $T_0$ such that, $\forall T > T_0$, $\frac{1}{T} \int_{[0,T]^2 \cap \Gamma} \mathrm{d}s \, \mathrm{d}t \leqslant \ell_\Gamma$.

We will work under the following conditions on the smoothing parameters $h_T$ and $h_{j,T}$, $j = 1, \dots, d$:

(H.1) $h_T = c'(\frac{\log T}{T})^{1/(2k'+d)}$, for a fixed $0 < c' < \infty$,
(H.2) $h_{j,T} = c_1 T^{-1/(2k+1)}$, for fixed $0 < c_1 < \infty$.

We will use the $\alpha$-mixing coefficient defined in [2, p. 17]. For all Borel set $I \subset \mathbb{R}^+$ the $\sigma$-algebra defined by $(Z_t, t \in I)$ will be denoted by $\sigma(Z_t, t \in I)$. Writing $\alpha(u) = \sup_{t\in\mathbb{R}_+} \alpha(\sigma(Z_v, v \leqslant t), \sigma(Z_v, v \geqslant t+u))$, we will use the condition

(A.1) $\alpha(t) = \mathcal{O}(t^{-b})$ with $b > \frac{7r+5d}{2r}$.

We denote by $\hat{\hat{\eta}}_{l,T}$ and $\tilde{\tilde{m}}_{\psi,T}(\mathbf{x})$ the versions of $\hat{\eta}_{l,T}$ and $\tilde{m}_{\psi,T}(\mathbf{x})$ corresponding to a known density $f$. Introduce now the following quantities (see, for the discrete case, Camlong et al. [4]):

$$\tilde{Y}_{\psi,T,t,l} = \psi(Y_t) \int_{\mathbb{R}^{d-1}} \prod_{j\neq l}^{d} \left( \frac{1}{h_{j,T}} K_j\left(\frac{\mathbf{x}_j - \mathbf{X}_{t,j}}{h_{j,T}}\right) \frac{q_{-l}(\mathbf{x}_{-l})}{f(X_{t,-l}|X_{t,l})}\right) \mathrm{d}\mathbf{x}_{-l}; \quad \tilde{m}^T_{\psi,l}(x_l) = E(\tilde{Y}_{\psi,T,t,l}|X_{t,l} = x_l);$$

$$\hat{a}_l(x_l) = \frac{1}{Th_{l,T}} \int_0^T \frac{\tilde{Y}_{\psi,T,t}}{f_1(X_{t,l})} K_l\left(\frac{x_l - X_{t,l}}{h_{l,T}}\right) \mathrm{d}t; \quad \mathcal{G}_l(\mathbf{u}_{-l}) = \int_{\mathbb{R}^{d-1}} \prod_{j\neq l}^{d} \left(\frac{1}{h_{j,T}} K_j\left(\frac{\mathbf{x}_j - \mathbf{u}_j}{h_{j,T}}\right) q_{-l}(\mathbf{x}_{-l})\right) \mathrm{d}\mathbf{x}_{-l};$$

$$C_{T,l} = \mu + \int_{\mathbb{R}^{d-1}} \sum_{j\neq l} m_j(u_j)\mathcal{G}_l(\mathbf{u}_{-l})\,\mathrm{d}\mathbf{u}_{-l}; \quad \hat{C}_T = \int_{\mathbb{R}^d} \tilde{\tilde{m}}_{\psi,T}(\mathbf{x}) q(\mathbf{x})\,\mathrm{d}\mathbf{x}; \quad C_l = \int_{\mathbb{R}} m_l(x_l) q_l(x_l)\,\mathrm{d}x_l;$$

$$b_l(x_l) = \frac{1}{k!} \int_{\mathbb{R}} u^k K_l(u)\,\mathrm{d}u \left((-1)^k m_l^{(k)}(x_l) + \int_{\mathbb{R}} m_l(z) q_l^{(k)}(z)\,\mathrm{d}z\right).$$

## 3. Results

**Theorem 3.1.** *Under assumptions* (C.1)–(C.2), (F.1)–(F.3), (K.1)–(K.4), (Q.1), (D.1)–(D.3), (H.1)–(H.2) *and* (A.1) *we have*

$$E\big(\hat{\eta}_{l,T}(x_l) - \eta_l(x_l)\big)^2 = \mathcal{O}\big(T^{-2k/(2k+1)}\big).$$

The next theorem needs the following additional hypothesis:

($\mathcal{V}$) $\liminf_{T\to\infty} Th_{l,T}\mathrm{Var}(\hat{\eta}_{l,T}(x_l)) > 0$ where $(\log(T)/T)^{k'/(2k'+d)} = o(h^k_{l,T})$.

**Theorem 3.2.** *Under the hypotheses of Theorem* 3.1 *and* ($\mathcal{V}$) *we have, for every* $l \in \{1, \ldots, d\}$ *and for any* $x_l \in \mathcal{C}_l$,

$$\frac{\hat{\eta}_{l,T}(x_l) - \eta_l(x_l) - h^k_{l,T} b_l(x_l)}{\sqrt{\mathrm{Var}(\hat{\eta}_{l,T}(x_l))}} \xrightarrow{\mathcal{L}} \mathcal{N}(0, 1).$$

**Proposition 3.3.** *Under the conditions* (C.1)–(C.4), (F.1)–(F.2), (K.1), (Q.1)–(Q.2) *and* (H.1)–(H.2), *we have, for every* $l \in \{1, \ldots, d\}$, *for any* $x_l \in \mathcal{C}_l$ *and every* $(\alpha, \beta) \in ]0; 0,5[ \times ]0,5; 1[$,

$$\liminf_{T\to\infty} P\big(T^{\frac{k}{2k+1}}\big\{\hat{\eta}_{l,T}(x_l) - \eta_l(x_l) - h^k_{l,T} b_l(x_l)\big\} \in [AQ_\alpha; AQ_\beta]\big) \geqslant \beta - \alpha, \tag{8}$$

*where* $A := (\limsup_{T\to+\infty} T^{\frac{2k}{2k+1}} \mathrm{Var}(\hat{\eta}_{l,T}(x_l)))^{1/2}$ *and* $Q_u$ *is such that* $P(\mathcal{N}(0,1) < Q_u) = u$.

The proofs of our theorems are split into two steps. We first consider the density as known, and then treat the general case where $f$ is unknown by using the decomposition $1/f = 1/\hat{f}_T - (f - \hat{f}_T)/f\hat{f}_T$ and the following lemma:

**Lemma 3.4.** *Under the assumptions* (F.1)–(F.3), (K.1), (K.2), (K.4), (D.1)–(D.3), (H.1) *and* (A.1), *we have*

$$\sup_{\mathbf{x}\in\mathcal{C}}\big|\hat{f}_T(\mathbf{x})-f(\mathbf{x})\big|=\mathcal{O}\bigg(\bigg(\frac{\log T}{T}\bigg)^{k'/(2k'+d)}\bigg)\quad a.s. \tag{9}$$

**Proof of Lemma.** It is easily seen that under our assumptions, the result follows by using the arguments used in the demonstration of Theorem 4.9. in [2, p. 112] and by replacing $\log_m$ by 1. □

**Sketch of the proof of Theorem 3.1.** Observe that

$$\hat{\eta}_{l,T}(x_l)-\eta_l(x_l)=\big\{\hat{\eta}_{l,T}(x_l)-\hat{\hat{\eta}}_{l,T}(x_l)\big\}+\big\{\hat{a}_l(x_l)-E\hat{a}_l(x_l)\big\}+\big\{E\hat{a}_l(x_l)-\tilde{m}^T_{\psi,l}(x_l)\big\}+E\{\hat{C}_T-C_{T,l}-C_l\}.$$

It follows that

$$E\big\{\hat{\eta}_{l,T}(x_l)-\eta_l(x_l)\big\}^2\leqslant 4E\big\{\hat{\eta}_{l,T}(x_l)-\hat{\hat{\eta}}_{l,T}(x_l)\big\}^2+4E\big\{\hat{a}_l(x_l)-E\hat{a}_l(x_l)\big\}^2+4\big\{E\hat{a}_l(x_l)-\tilde{m}^T_{\psi,l}(x_l)\big\}^2+4E^2\{\hat{C}_T-C_{T,l}-C_l\}.$$

To prove the Theorem 3.1, it suffices to establish the following statements

$$E\big(\hat{\eta}_{l,T}(x_l)-\hat{\hat{\eta}}_{l,T}(x_l)\big)^2=\mathcal{O}\big(T^{-2k/(2k+1)}\big), \tag{10}$$

$$\operatorname{Var}\big(\hat{a}_l(x_l)\big)=\mathcal{O}\big(T^{-2k/(2k+1)}\big), \tag{11}$$

$$E\hat{a}_l(x_l)-\tilde{m}^T_{\psi,l}(x_l)=\mathcal{O}\big(T^{-k/(2k+1)}\big), \tag{12}$$

$$E(\hat{C}_T-C_{T,l}+C_l)=\mathcal{O}\big(T^{-k/(2k+1)}\big). \tag{13}$$

**Proof of (10).** By combining the definitions of $\hat{\eta}_{1,T}$ and $\hat{\hat{\eta}}_{1,T}$ and the result of Lemma 3.4, we easily obtain, under the conditions on the kernel, the statement (10). □

**Proof of (11).** Set $\phi(t,s)=\operatorname{Cov}(\frac{\tilde{Y}_{\psi,T,t}}{f_1(X_{t,1})h_{1,T}}K_1(\frac{x_1-X_{t,1}}{h_{1,T}}),\frac{\tilde{Y}_{\psi,T,s}}{f_1(X_{s,1})h_{1,T}}K_1(\frac{x_1-X_{s,1}}{h_{1,T}}))$ and $S=\{(s,t)\in\mathbb{R}^2;\,|t-s|\leqslant h_T^{-1}\}$. We use the following decomposition

$$\operatorname{Var}\big(\hat{a}_1(x_1)\big)=\int\limits_{[0,T]^2\cap\Gamma}\phi(t,s)\,\mathrm{d}t\,\mathrm{d}s+\int\limits_{[0,T]^2\cap\Gamma^c\cap S}\phi(t,s)\,\mathrm{d}t\,\mathrm{d}s+\int\limits_{[0,T]^2\cap\Gamma^c\cap S^c}\phi(t,s)\,\mathrm{d}t\,\mathrm{d}s:=A+E+F.$$

Under (C.1), (F.1), (K.1)–(K.2) and (Q.1), we have, for $T$ large enough,

$$A=\mathcal{O}(1/Th_{1,T})\quad\text{and}\quad E=\mathcal{O}\big(h_T^{-1}\|K_1\|^2_{\mathbb{L}_1}A_\Gamma/T\big). \tag{14}$$

Using the Billingsley's inequality, it follows that

$$F=\mathcal{O}(1/Th_{1,T}). \tag{15}$$

Combining (14) and (15), we obtain (11). To prove the statements (12) and (13), we use similar arguments as in the discrete case (see Camlong et al. [4]). □

**Sketch of the proof of Theorem 3.2.** To obtain our theorem it suffices to show that

$$\sup_{x_l\in\mathcal{C}_l}\big|\hat{\eta}_{l,T}(x_l)-\hat{\hat{\eta}}_{l,T}(x_l)\big|=\mathcal{O}\Big(\sup_{\mathbf{x}\in\mathcal{C}}\big|\hat{f}_T(\mathbf{x})-f(\mathbf{x})\big|\Big)\quad\text{a.s.}, \tag{16}$$

$$\frac{\{\hat{a}_l(x_l)-E(\hat{a}_l(x_l))\}}{\sqrt{\operatorname{Var}(\hat{a}_l(x_l))}}\to\mathcal{N}(0,1), \tag{17}$$

$$E\hat{a}_l(x_l)-\tilde{m}^T_{\psi,l}(x_l)=\frac{(-h_{l,T})^k}{k!}m_l^{(k)}(x_l)\int\limits_{\mathbb{R}}v_l^kK_l(v_l)\,\mathrm{d}v_l+\mathrm{o}\big(h^k_{l,T}\big), \tag{18}$$

$$\text{and}\quad E\{\hat{C}_T-C_{T,l}+C_l\}=\frac{h^k_{l,T}}{k!}\int\limits_{\mathbb{R}}q_l^{(k)}(x_l)m_l(x_l)\,\mathrm{d}x_l\int\limits_{\mathbb{R}}v_l^kK_l(v_l)\,\mathrm{d}v_l+\mathrm{o}\big(h^k_{l,T}\big). \tag{19}$$

**Proof of (16).** The result arises directly from the definitions of estimates of $\eta_l$ and the conditions on the kernels $K_l, 1 \leqslant l \leqslant d$. □

**Proof of (17).** Set $\frac{\{\hat{a}_l(x_l)-E(\hat{a}_l(x_l))\}}{\sqrt{\operatorname{Var}(\hat{a}_l(x_l))}} = \int_0^T Z_t \,\mathrm{d}t =: S_T$. We employ then the big block–small block procedure. Indeed setting, $S_T = \sum_{j=1}^{k-1}(\nu_j + \xi_j) =: S'_T + S''_T$ where $\nu_j = \int_{j(p+q)}^{j(p+q)+p} Z_t \,\mathrm{d}t$ and $\xi_j = \int_{j(p+q)+p}^{(j+1)(p+q)} Z_t \,\mathrm{d}t$. Now, it suffices to prove the following statements,

$$ES_T''^2 \to 0 \quad \text{as} \quad T \to +\infty, \tag{20}$$

$$\left| E\left(\mathrm{e}^{\mathrm{i}tS'_T}\right) - \prod_{j=0}^{k-1} E\left(\mathrm{e}^{\mathrm{i}t\nu_j}\right) \right| \to 0 \quad \text{as } T \to +\infty, \tag{21}$$

$$\sum_{j=0}^{k-1} E\left[\nu_j^2\right] \to 1 \quad \text{as } T \to +\infty, \tag{22}$$

$$\text{and} \quad \sum_{j=0}^{k-1} E\left[\nu_j^2 \mathbb{I}_{\{\nu_j^2 > \epsilon\}}\right] \to 0 \quad \text{as } T \to +\infty. \tag{23}$$

To show (22) and (23), we use the same arguments as those deployed in the discrete case. □

## References


[1] G. Banon, Nonparametric identification for diffusion processes, SIAM J. Control Optim. 16 (3) (1978) 380–395.

[2] D. Bosq, Nonparametric Statistics for Stochastic Processes, Estimation and Prediction, Lecture Notes in Statistics, vol. 110, Springer-Verlag, New York, 1998.

[3] D. Bosq, Vitesses optimales et superoptimales des estimateurs fonctionnels pour les processus à temps continu, C. R. Acad. Sci. Paris Sér. I Math. 317 (11) (1993) 1075–1078.

[4] C. Camlong-Viot, P. Sarda, P. Vieu, Additive time series: The kernel integration method, Math. Methods Statist. 9 (4) (2000) 358–375.

[5] N. Cheze-Payaud, Nonparametric regression and prediction for continuous-time processes, Publ. Inst. Statist. Univ. Paris 38 (2) (1994) 37–58.

[6] M.C. Jones, S.J. Davies, B.U. Park, Versions of kernel-type regression estimators, J. Amer. Statist. Assoc. 89 (427) (1994) 825–832.

[7] O. Linton, J.P. Nielsen, A kernel method of estimating structured nonparametric regression based on marginal integration, Biometrika 82 (1) (1995) 93–100.

[8] W.K. Newey, Kernel estimation of partial means and a general variance estimator, Econometric Theory 10 (2) (1994) 233–253.

[9] C.J. Stone, Additive regression and other nonparametric models, Ann. Statist. 13 (2) (1985) 689–705.