

Statistique/Probabilités

# Normalité asymptotique de l'estimateur par ondelettes des composantes d'un modèle additif de régression

Mohammed Debbarh

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

Reçu le 23 août 2005 ; accepté après révision le 2 octobre 2006

Disponible sur Internet le 27 octobre 2006

Présenté par Paul Deheuvels

## Résumé

Dans le cadre des modèles additifs de régression, nous établissons la normalité asymptotique des estimateurs des composantes additives obtenues par la méthode d'intégration marginale d'un estimateur par ondelettes. Pour établir nos résultats nous utilisons la décomposition « grands blocs-petits blocs » et le théorème central limite pour des variables dépendantes. **Pour citer cet article :** *M. Debbarh, C. R. Acad. Sci. Paris, Ser. I 343 (2006).*

© 2006 Académie des sciences. Publié par Elsevier Masson SAS. Tous droits réservés.

## Abstract

**Asymptotic normality for the wavelets estimator of the additive regression components.** In the setting of the additive regression model, we show asymptotic normality of an wavelets estimators of the additive components pertaining with the marginal integration estimation method. Our proof use the usual 'small blocks-big blocks' and the central limit theorem for dependent random variables. **To cite this article:** *M. Debbarh, C. R. Acad. Sci. Paris, Ser. I 343 (2006).*

© 2006 Académie des sciences. Publié par Elsevier Masson SAS. Tous droits réservés.

## Abridged English version

Let  $(\mathbf{X}, Y)$  be an  $\mathbb{R}^d \times \mathbb{R}$ -valued random vector. We consider the regression additive function of  $Y$  given  $\mathbf{X} = (X_1, \dots, X_d) = \mathbf{x}$ , defined by

$$\begin{aligned} m(\mathbf{x}) &= E(Y | \mathbf{X} = \mathbf{x}), \quad \forall \mathbf{x} = (x_1, \dots, x_d) \in \mathbb{R}^d, \\ &= \mu + \sum_{l=1}^d m_l(x_l). \end{aligned}$$

To get round any issue of identifiability, we assume that  $E m_l(X_l) = 0$ ,  $l = 1, \dots, d$ , what ensures that  $\mu = E(Y)$ . Denote by  $q_1, \dots, q_d$ ,  $d$  density functions defined in  $\mathbb{R}$  and set  $q(\mathbf{x}) = \prod_{l=1}^d q_l(x_l)$  and  $q_{-l}(\mathbf{x}_{-l}) = \prod_{j \neq l} q_j(x_j)$  where  $\mathbf{x}_{-l} = (x_1, \dots, x_{l-1}, x_{l+1}, \dots, x_d)$ . Using the marginal integration method (see Newey [14]), we can write

Adresse e-mail : [debbarh@ccr.jussieu.fr](mailto:debbarh@ccr.jussieu.fr) (M. Debbarh).

$$m(\mathbf{x}) = \sum_{l=1}^d \eta_l(x_l) + \int_{\mathbb{R}^d} m(\mathbf{x})q(\mathbf{x}) \, d\mathbf{x} \quad \text{with} \quad \eta_l(x_l) = \int_{\mathbb{R}^{d-1}} m(\mathbf{x})q_{-l}(\mathbf{x}_{-l}) \, d\mathbf{x}_{-l} - \int_{\mathbb{R}^d} m(\mathbf{x})q(\mathbf{x}) \, d\mathbf{x}. \quad (1)$$

Let  $(\mathbf{X}_i, Y_i)_{1 \leq i \leq n}$  be absolutely regular random pairs, taking values in  $\mathbb{R}^d \times \mathbb{R}$ , each having the same distribution as  $(\mathbf{X}, Y)$ . Define the kernel  $K_1$  on  $\mathbb{R}^2$  by  $K_1(u, v) = \sum_{k \in \mathbb{Z}} \varphi(u - k)\varphi(v - k)$ , where the function  $\varphi$  is bounded, left-continuous and compactly supported. Let  $(h(n))_{n \geq 1}$  be a sequence of positive constants such that  $h(n) \rightarrow +\infty$  as  $n \rightarrow \infty$ . In this paper, we propose to establish the asymptotic normality for a wavelets estimator of the additive regression components. From (1), a natural estimate of the first component  $\eta_1$  is given by

$$\hat{\eta}_1(x_1) = \int_{\mathbb{R}^{d-1}} \hat{m}_n(\mathbf{x})q_{-1}(\mathbf{x}_{-1}) \, d\mathbf{x}_{-1} - \int_{\mathbb{R}^d} \hat{m}_n(\mathbf{x})q(\mathbf{x}) \, d\mathbf{x},$$

where  $\hat{m}_n(\mathbf{x})$  is a multivariate wavelets estimate of the regression function  $m$ . Under some assumptions, we have for every  $x_1$  in some compact  $D_1$ ,

$$\left( \frac{n}{2^{h(n)}} \right)^{1/2} \frac{\hat{\eta}_1(x_1) - \eta_1(x_1)}{\sigma_1^2(x_1) \left( \int_{\mathbb{R}} K_1^2(2^{h(n)}x_1, u) \, du \right)^{1/2}} \xrightarrow{\mathcal{L}} \mathcal{N}(0, 1),$$

where  $\sigma_1^2(x_1) = \int_{\mathbb{R}^{d-1}} E(Y^2 | \mathbf{X} = \mathbf{x})q_{-1}(\mathbf{x}_{-1})/f(\mathbf{x}) \, d\mathbf{x}_{-1}$ .

## 1. Introduction

Il est bien connu depuis le travail de Stone [17] que les modèles additifs apportent une solution au problème du fléau de la dimension rencontré en estimation non paramétrique de la fonction de régression multivariée. Ce fléau se caractérise par la perte en vitesse de convergence de l'estimateur de la fonction de régression quand le nombre de covariables augmente. La solution proposée par les modèles additifs permet d'atteindre des vitesses pouvant aller jusqu'à celle obtenue pour le modèle de régression univariée. Dans ce cadre, Camlong-Viot, Sarda et Vieu [3] ont établi des résultats de convergence en moyenne quadratique et en loi des estimateurs des composantes additives (3), en utilisant l'estimateur à noyau. Sperlich, Tjøstheim et Yang [16] ont étendu l'étude de la normalité aux estimateurs des termes d'interactions entre covariables du second ordre. Zhang et Wang [18] ont obtenu la vitesse de convergence en moyenne quadratique de l'estimateur des composantes additives reposant sur la méthode d'ondelettes (voir aussi Amato et Antoniadis [1]). Dans ce travail, en utilisant un estimateur par ondelettes, nous établissons la normalité asymptotique des estimateurs des composantes additives définies en (6).

Soit  $(\mathbf{X}_i, Y_i)_{i \geq 1}$  une suite de couples de vecteurs aléatoires à valeurs dans  $\mathbb{R}^d \times \mathbb{R}$ ,  $d \geq 2$ , de même loi que  $(\mathbf{X}, Y)$ . On considère le modèle additif suivant :

$$Y_i = m(\mathbf{X}_i) + e_i = \mu + \sum_{l=1}^d m_l(X_{i,l}) + e_i, \quad (2)$$

où  $E(e_i) = 0$ ,  $\text{Var}(e_i) = \sigma^2$  et  $e_1, \dots, e_d$  sont des variables aléatoires indépendantes des covariables  $\mathbf{X}_i$ . Nous supposons par la suite que les conditions d'identifiabilité suivantes,  $E(m_l(X_{i,l})) = 0$ , pour  $l = 1, \dots, d$ , sont vérifiées. Soient  $q_1, \dots, q_d$  des densités réelles telles que  $q(\mathbf{x}) = \prod_{l=1}^d q_l(x_l)$ ,  $q_{-l} = \prod_{j \neq l} q_j$  et  $\mathbf{x}_{-l} = (x_1, \dots, x_{l-1}, x_{l+1}, \dots, x_d)$ . Par la méthode d'intégration marginale [14], nous obtenons

$$\eta_l(x_l) = \int_{\mathbb{R}^{d-1}} m(\mathbf{x})q_{-l}(\mathbf{x}_{-l}) \, d\mathbf{x}_{-l} - \int_{\mathbb{R}^d} m(\mathbf{x})q(\mathbf{x}) \, d\mathbf{x}, \quad (3)$$

avec

$$\eta_l(x_l) = m_l(x_l) - \int_{\mathbb{R}} m_l(z_l)q_l(z_l) \, dz_l \quad \text{et} \quad m(\mathbf{x}) = \sum_{l=1}^d \eta_l(x_l) + \int_{\mathbb{R}^d} m(\mathbf{z})q(\mathbf{z}) \, d\mathbf{z}. \quad (4)$$

## 2. Estimation par ondelettes des composantes additives

Pour estimer la fonction de régression  $m$  et les composantes additives  $\eta_l$ , nous utilisons un estimateur par ondelettes. Nous nous référons pour cette méthode d'estimation aux premiers travaux de Doukhan et Léon [5], Kerkycharian et Picard [8], qui ont traité l'estimateur linéaire par ondelettes dans le cas i.i.d. Leblanc [9] a traité l'estimateur linéaire par ondelettes dans le cas d'un processus dépendant. Notons que dans ce travail, nous n'abordons pas les estimateurs non linéaires, dits à « seuil ». Pour ce type d'estimateurs, nous pouvons citer à titre d'exemple Härdle et al. [6] et Picard et Tribouley [15].

Soit  $\mathcal{L}$  la collection des cubes dyadiques

$$\mathbf{Q} = \{ \mathbf{x} \in \mathbb{R}^d; 2^j \mathbf{x} - \mathbf{k} \in [0, 1]^d \}, \quad \mathbf{k} \in \mathbb{Z}^n, \quad j \in \mathbb{Z},$$

de  $\mathbb{R}^d$ . Désignons par  $E$  l'ensemble fini des  $2^d - 1$  éléments  $(\varepsilon_1, \dots, \varepsilon_d)$ , où  $\varepsilon_l = 1$  ou  $0$  avec  $1 \leq l \leq d$ , l'élément  $(0, \dots, 0)$  étant exclu. Nous définissons alors les ondelettes  $d$ -dimensionnelles  $\psi_{\mathbf{Q}}^\varepsilon$  par

$$\psi_{\mathbf{Q}}^\varepsilon(\mathbf{x}) = 2^{nj/2} \psi^{\varepsilon_1}(2^j x_1 - k_1) \cdots \psi^{\varepsilon_d}(2^j x_d - k_d), \quad \varepsilon \in E \text{ et } \mathbf{Q} \in \mathcal{L},$$

avec la convention  $\psi^0 = \varphi$  et  $\psi^1 = \psi$ . Soit  $f$  la fonction densité de la covariable  $\mathbf{X}$ . Nous supposons que les fonctions  $m$  et  $f$  appartiennent à  $L_2(\mathbb{R}^d)$  et que  $|Y_i| \leq M$ , pour tout  $i \geq 1$  avec  $M$  un réel positif. En utilisant l'inégalité de Jensen, nous pouvons facilement montrer que  $g = mf \in L_2(\mathbb{R}^d)$ . Soit  $j_0$  un nombre fixé dans  $\mathbb{Z}$ . Pour toute fonction  $g$  différentiable à support compact, nous avons la décomposition suivante dans la base d'ondelettes constituée des  $\psi_{\mathbf{Q}}^\varepsilon$  avec  $\mathbf{Q} \in \mathcal{L}$ ,  $\varepsilon \in E$  (voir Meyer [13]),

$$g = \sum_{\mathbf{Q} \in \mathcal{L}_{j_0}} (g, \varphi_{\mathbf{Q}}) \varphi_{\mathbf{Q}} + \sum_{j \geq j_0} \sum_{\mathbf{Q} \in \mathcal{L}_j} \sum_{\varepsilon \in E} (g, \psi_{\mathbf{Q}}^\varepsilon) \psi_{\mathbf{Q}}^\varepsilon,$$

où  $\mathcal{L}_j$  désigne la collection des cubes dyadiques dont le côté vaut  $2^{-j}$ ,  $\varphi_{\mathbf{Q}} = 2^{dj/2} \varphi(2^j \cdot - k_1) \cdots \varphi(2^j \cdot - k_d)$  et  $(g, \psi_{\mathbf{Q}}^\varepsilon) = \int g \psi_{\mathbf{Q}}^\varepsilon$ . Soit  $\{X_i = (X_{i,1}, \dots, X_{i,d}), i \geq 1\}$  une suite de vecteurs aléatoires à valeurs dans  $\mathbb{R}^d$  de densité  $f$ .

Les estimateurs  $(\widehat{g, \psi_{\mathbf{Q}}^\varepsilon})$  des coefficients d'ondelettes  $(g, \psi_{\mathbf{Q}}^\varepsilon)$  sont alors définis par

$$(\widehat{g, \psi_{\mathbf{Q}}^\varepsilon}) = \frac{1}{n} \sum_{l=1}^n Y_l \psi_{\mathbf{Q}}^\varepsilon(\mathbf{X}_l).$$

Ceci conduit à considérer l'estimateur « linéaire »  $\hat{g}$  de la fonction  $g$  défini par

$$\hat{g} = \sum_{\mathbf{Q} \in \mathcal{L}_{j_0}} (\widehat{g, \varphi_{\mathbf{Q}}}) \varphi_{\mathbf{Q}} + \sum_{j=j_0}^{h-1} \sum_{\mathbf{Q} \in \mathcal{L}_j} \sum_{\varepsilon \in E} (\widehat{g, \psi_{\mathbf{Q}}^\varepsilon}) \psi_{\mathbf{Q}}^\varepsilon, \tag{5}$$

où  $j_0 \leq j < h$  et  $h := h(n)$  dépend du nombre d'observations  $n$ . Nous pouvons réécrire l'estimateur (5) sous la forme suivante :

$$\hat{g}(\mathbf{x}) = \frac{2^{dh(n)}}{n} \sum_{t=1}^n Y_t K(2^{h(n)} \mathbf{x}, 2^{h(n)} \mathbf{X}_t) \quad \text{où } K(\mathbf{u}, \mathbf{v}) = \sum_{\mathbf{Q} \in \mathcal{L}_0} \varphi_{\mathbf{Q}}(\mathbf{u} - \mathbf{k}) \varphi_{\mathbf{Q}}(\mathbf{v} - \mathbf{k}).$$

Remarquons de plus que  $K$  peut s'écrire sous forme d'un produit  $K := K_1 \times K_{-1}$  où  $K_1(u_1, v_1) = \sum_{k_1 \in \mathbb{Z}} \varphi(u_1 - k_1) \varphi(v_1 - k_1)$ . L'estimateur de la  $l$ -ième composante additive est alors défini par

$$\hat{\eta}_l(x_l) = \int_{\mathbb{R}^{d-1}} \hat{m}_n(\mathbf{x}) q_{-l}(\mathbf{x}_{-l}) d\mathbf{x}_{-l} - \int_{\mathbb{R}^d} \hat{m}_n(\mathbf{x}) q(\mathbf{x}) d\mathbf{x}, \tag{6}$$

où  $\hat{m}_n$  est l'estimateur interne de la fonction de régression  $m$ , lui-même défini par

$$\hat{m}_n(\mathbf{x}) = \frac{2^{dh(n)}}{n} \sum_{t=1}^n \frac{Y_t}{f(\mathbf{X}_t)} K(2^{h(n)} \mathbf{x}, 2^{h(n)} \mathbf{X}_t).$$

Pour traiter le cas où la densité  $f$  des covariables  $\mathbf{X}_i$  est inconnue, nous considérons une fonction  $\Phi$  d'échelle de classe  $\mathcal{C}^r$ . Supposons qu'il existe une constante  $A_{d+1}$ , telle que  $|\Phi(\mathbf{x})| \leq A_{d+1}/(1 + \|\mathbf{x}\|)^{d+1}$ . Posons pour tout  $\mathbf{u}, \mathbf{v}$  de  $\mathbb{R}^d$ ,  $\mathbb{K}(\mathbf{u}, \mathbf{v}) = \sum_{\mathbf{k} \in \mathbb{Z}^d} \Phi(\mathbf{u} - \mathbf{k})\Phi(\mathbf{v} - \mathbf{k})$ . On peut alors estimer  $f$  à partir des observations  $\mathbf{X}_1, \dots, \mathbf{X}_n$  par la fonction  $\hat{f}_n$ , définie pour tout  $\mathbf{x} \in \mathbb{R}^d$  par

$$\hat{f}_n(\mathbf{x}) = \frac{2^{dj(n)}}{n} \sum_{i=1}^n \mathbb{K}(2^{j(n)}\mathbf{x}, 2^{j(n)}\mathbf{X}_i),$$

où  $j(n)$  est un niveau d'analyse multirésolution qui varie avec la taille de l'échantillon  $n$ . Dans le cas où la densité  $f$  est inconnue, les estimateurs des composantes additives  $\eta_l$ , seront notés  $\hat{\eta}_l$ ,  $l = 1, \dots, d$ .

Par ailleurs, introduisons les notations suivantes, en vue de la présentation de nos résultats,

$$\sigma_1^2(u_1) = \frac{\Phi(u_1)}{f_1(u_1)} = \int_{\mathbb{R}^{d-1}} H(\mathbf{u}) \frac{q_{-1}(\mathbf{u}_{-1})}{f(\mathbf{u})} d\mathbf{u}_{-1} \quad \text{et} \quad H(\mathbf{u}) = E(Y^2 | \mathbf{X} = \mathbf{u}).$$

### 3. Normalité asymptotique d'un estimateur par ondelettes

Pour établir nos résultats, les hypothèses suivantes sont suffisantes.

**H<sub>1</sub>** : Les deux suites  $\{h(n), n \geq 1\}$  et  $\{j(n), n \geq 1\}$  vérifient les conditions suivantes

$$h(n) \rightarrow +\infty, \quad n2^{-h(n)} \rightarrow +\infty, \quad n2^{-h(n)/\lambda} \rightarrow 0, \quad 2^{-j(n)} = \left(\frac{\log n}{n}\right)^{1/(d+2(s-d/p))},$$

et  $\sqrt{n}2^{-h(n)/2-j(n)(s-d/p)} \rightarrow 0$  où  $s > d/p$  et  $\lambda > 0$ .

**H<sub>2</sub>** : Les densités  $f$  et  $f_l$  de  $\mathbf{X}$  et  $X_l$ , sont à supports compacts  $D$  et  $D_l$  respectivement. En outre, il existe des réels  $b, B, b_l$  et  $B_l$  tels que

$$0 < b \leq f(\mathbf{x}) \leq B < \infty, \quad \mathbf{x} \in D \quad \text{et} \quad 0 < b_l \leq f_l(x_l) \leq B_l < \infty, \quad x_l \in D_l.$$

**H<sub>3</sub>** : La densité conjointe  $f_{1,i-j}$  des variables aléatoires  $(X_i, X_j)$  existe et vérifie

$$\forall x, y, \quad |f_{1,i-j}(x, y) - f_1(x)f_1(y)| \leq M < \infty.$$

**H<sub>4</sub>** : Les fonctions  $m_l(x_l)$ ,  $1 \leq l \leq d$ , sont lipchitziennes d'ordre  $1/2\lambda$ .

**H<sub>5</sub>** : La suite  $(\mathbf{X}_i, Y_i)_{i \geq 1}$  est  $\beta$ -mélangeante dont le coefficient de mélangeance vérifie la condition  $\beta(s) = \mathcal{O}(s^{-a})$ ,  $a > 4$ .

**H<sub>6</sub>** : La densité  $f$  appartient à l'espace de Besov  $B_{s,p,q}$ ,  $0 < s < r$ ,  $1 \leq p, q \leq \infty$ .

**H<sub>7</sub>** : La densité conjointe des vecteurs  $(\mathbf{X}_0, \mathbf{X}_j)$  existe et est bornée.

Pour présenter nos résultats, il convient de rappeler que si  $\varphi$  est bornée, continue à gauche et à support compact, alors la suite  $\{\int_{-\infty}^{\infty} K_1^2(2^{h(n)}x_1, u_1) du_1, n \geq 1\}$  est bornée (voir par exemple Antoniadis, Gregoire et McKeaque [2]).

**Théorème 3.1.** *On suppose les conditions **H<sub>1</sub>–H<sub>7</sub>** satisfaites. Si  $\varphi$  est continue à gauche, à variation bornée et à support compact, alors*

$$\left(\frac{n}{2^{h(n)}}\right)^{1/2} \frac{\hat{\eta}_1(x_1) - \eta_1(x_1)}{\sigma_1(x_1) \left(\int_{\mathbb{R}} K_1^2(2^{h(n)}x_1, u) du\right)^{1/2}} \xrightarrow{\mathcal{L}} \mathcal{N}(0, 1), \quad \forall x_1 \in D_1.$$

**Remarque 1.** En choisissant  $(h(n))_{n \geq 1}$  telle que  $2^{-h(n)} \sim n^{-1/5}$ ,  $0 < \lambda < 1/5$  et  $s > 2d + d/p$ , on obtient la vitesse de convergence optimale «  $n^{2/5}$  » qui ne dépend pas de la dimension  $d$ .

**Éléments de preuve.** On utilise la décomposition suivante (voir Camlong, Sarda et Vieu [3]),

$$\hat{\eta}_1(x_1) - \eta_1(x_1) = [\hat{\alpha}_1(x_1) - \tilde{m}_n(x_1)] - [\hat{C}_n - C_n - C],$$

où

$$\hat{\alpha}_1(x_1) = \frac{2^{h(n)}}{n} \sum_{i=1}^n \frac{\tilde{Y}_{t,n}}{f_1(X_{t,1})} K_1(2^{h(n)}x_1, 2^{h(n)}X_{t,1}); \quad \tilde{Y}_{t,n} = \frac{Y_t \mathcal{G}(\mathbf{X}_{t,-1})}{f(X_{t,-1}|X_{t,1})};$$

$$\mathcal{G}(\mathbf{X}_{t,-1}) = \int_{\mathbb{R}^{d-1}} 2^{(d-1)h(n)} K_{-1}(2^{h(n)}\mathbf{x}_{-1}, 2^{h(n)}X_{t,-1}) q_1(\mathbf{x}_{-1}) d\mathbf{x}_{-1}; \quad \hat{C}_n = \int_{\mathbb{R}^d} \hat{m}_n(\mathbf{x}) q(\mathbf{x}) d\mathbf{x};$$

$$C_n = \mu + \int_{\mathbb{R}^{d-1}} \sum_{j=2}^d m_j(z_j) \mathcal{G}(\mathbf{z}_{-1}) d\mathbf{z}_{-1}; \quad C = \int_{\mathbb{R}} m_1(x_1) q_1(x_1) dx_1 \quad \text{et} \quad \tilde{m}_n(x_1) = E(Y_{t,n}|X_1 = x_1).$$

La preuve du Théorème 3.1, dans le cas où  $f$  est connue, consiste alors à prouver les quatre résultats suivants

$$\left(\frac{n}{2^{h(n)}}\right)^{1/2} \frac{\hat{\alpha}_1(x_1) - \tilde{m}_n(x_1)}{(\int_{\mathbb{R}} K_1^2(2^{h(n)}x_1, v_1) dv_1)^{1/2}} \rightarrow \mathcal{N}(0, \sigma_1^2(x_1)), \tag{7}$$

$$E(\hat{C}_n - C_n) = C + o\left(\left(\frac{n}{2^{h(n)}}\right)^{-1/2}\right), \tag{8}$$

$$\text{Var}(\hat{C}_n) = o\left(\left(\frac{n}{2^{h(n)}}\right)\right), \tag{9}$$

$$\text{Cov}(\hat{\alpha}_1(x_1), \hat{C}_n) = o\left(\left(\frac{n}{2^{h(n)}}\right)\right). \tag{10}$$

Preuve de (7) : On commence par utiliser la décomposition « grands blocs-petits blocs » (voir Masry [11]) et par l’intermédiaire du lemme classique de Berbee (voir, Doukhan [4], Théorème 1), on peut ramener l’étude d’un processus faiblement dépendant à un processus indépendant. En vérifiant les conditions de stabilité du critère de convergence normale (voir par exemple Loève [10] et Ibragimov [7]) on obtient le résultat (7).

Preuve de (8) : On utilise la condition de Lipschitz sur les composantes  $m_l$  et le fait que  $\varphi$  est à support compact.

Preuve de (9) : Le résultat (9) découle directement de la condition de faible mélangeance.

Preuve de (10) : En combinant les résultats (7) et (9) et en utilisant l’inégalité de Cauchy–Schwartz on déduit aisément le résultat (10).

Supposons maintenant que  $f$  est inconnue. En utilisant la décomposition  $\frac{1}{f_n} = \frac{1}{f} - \frac{\hat{f}_n - f}{f \hat{f}_n}$ , il ressort que

$$\hat{m}_n(\mathbf{x}) = \hat{\hat{m}}_n(\mathbf{x}) + \frac{2^{dh(n)}}{n} \sum_{i=1}^n Y_i K(2^{h(n)}\mathbf{x}, 2^{h(n)}\mathbf{X}_i) \frac{f(\mathbf{X}_i) - \hat{f}_n(\mathbf{X}_i)}{f(\mathbf{X}_i) \hat{f}_n(\mathbf{X}_i)},$$

où  $\hat{\hat{m}}_n(\mathbf{x})$  désigne l’estimateur de la fonction de régression  $m$  dans le cas où  $f$  est connue. En combinant le résultat dû à Masry [12] sur l’estimateur d’une fonction densité  $f$ ,

$$\sup_{\mathbf{x} \in D} |\hat{f}_n(\mathbf{x}) - f(\mathbf{x})| = \mathcal{O}(n^{-j(n)(s-d/p)}) \quad \text{p.s.},$$

et le fait que  $\frac{2^{dh(n)}}{n} \sum_{i=1}^n |K(2^{h(n)}\mathbf{x}, 2^{h(n)}\mathbf{X}_i)|$  est uniformément bornée pour  $n$  assez grand, on obtient

$$\sup_{\mathbf{x} \in D} |\hat{m}_n(\mathbf{x}) - \hat{\hat{m}}_n(\mathbf{x})| = \mathcal{O}(2^{-j(n)(s-d/p)}) \quad \text{p.s.} \tag{11}$$

On sait que  $\hat{\eta}_1(x_1) - \hat{\hat{\eta}}_1(x_1) = \int_{\mathbb{R}^{d-1}} [\hat{m}_n(\mathbf{x}) - \hat{\hat{m}}_n(\mathbf{x})] q_{-1}(\mathbf{x}_{-1}) d\mathbf{x}_{-1} - \int_{\mathbb{R}^d} [\hat{m}_n(\mathbf{x}) - \hat{\hat{m}}_n(\mathbf{x})] q(\mathbf{x}) d\mathbf{x}$ , où  $\hat{\eta}$  désigne l’estimateur de  $\eta_1$  dans le cas où  $f$  est connue. En utilisant (11) il vient immédiatement que  $\sup_{x_1} |\hat{\eta}_1(x_1) - \hat{\hat{\eta}}_1(x_1)| = \mathcal{O}(2^{-j(n)(s-d/p)})$  p.s., on en déduit alors que,

$$\left(\frac{n}{2^{h(n)}}\right)^{1/2} \sup_{x_1 \in D_1} |\hat{\hat{\eta}}_1(x_1) - \hat{\eta}_1(x_1)| = o(1) \quad \text{p.s.}$$

Finalement, en appliquant le théorème de Slutsky on obtient le résultat du Théorème 3.1.  $\square$

## Références

- [1] U. Amato, A. Antoniadis, Adaptive wavelet series estimation in separable nonparametric regression models, *Stat. Comput.* 11 (4) (2001) 373–394.
- [2] A. Antoniadis, G. Gregoire, I.W. Mckeague, Wavelet methods for curve estimation, *J. Amer. Statist. Ass.* 89 (2) (1994) 1340–1353.
- [3] C. Camlong-Viot, P. Sarda, P. Vieu, Additive time series: the kernel integration method, *Math. Methods Statist.* 9 (4) (2000) 358–375.
- [4] P. Doukhan, *Mixing: Properties and Examples*, Lecture Notes in Statistics, vol. 85, Springer-Verlag, New York, 1994.
- [5] P. Doukhan, J.R. León, Déviation quadratique d’estimateurs de densité par projections orthogonales, *C. R. Acad. Sci. Paris Sér. I Math.* 310 (6) (1990) 425–430.
- [6] W. Härdle, G. Kerkycharian, D. Picard, A. Tsybakov, *Wavelets, Approximation, and Statistical Applications*, Lecture Notes in Statistics, vol. 129, Springer-Verlag, New York, 1998.
- [7] I.A. Ibragimov, Some limit theorems for stationary processes, *Teor. Veroyatnost. i Primenen.* 7 (1962) 361–392.
- [8] G. Kerkycharian, D. Picard, Density estimation in Besov spaces, *Statist. Probab. Lett.* 13 (1) (1992) 15–24.
- [9] F. Leblanc, Wavelet linear density estimator for a discrete-time stochastic process:  $L_p$ -losses, *Statist. Probab. Lett.* 27 (1) (1996) 71–84.
- [10] M. Loève, *Probability Theory*, third ed., Van Nostrand, Toronto, 1963.
- [11] E. Masry, Recursive probability density estimation for weakly dependent stationary processes, *IEEE Trans. Inform. Theory* 32 (2) (1986) 254–267.
- [12] E. Masry, Multivariate probability density estimation by wavelet methods: strong consistency and rates for stationary time series, *Stochastic Process. Appl.* 67 (2) (1997) 177–193.
- [13] Y. Meyer, *Wavelets and applications*, in: *Proceedings of the International Congress of Mathematicians*, vols. I, II, Kyoto, 1990, Math. Soc. Japan, Tokyo, 1991, pp. 1619–1626.
- [14] W.K. Newey, Kernel estimation of partial means and a general variance estimator, *Econometric Theory* 10 (2) (1994) 233–253.
- [15] D. Picard, K. Tribouley, Adaptive confidence interval for pointwise curve estimation, *Ann. Statist.* 28 (1) (2000) 298–335.
- [16] S. Sperlich, D. Tjøstheim, L. Yang, Nonparametric estimation and testing of interaction in additive models, *Econometric Theory* 18 (2) (2002) 197–251.
- [17] C.J. Stone, Additive regression and other nonparametric models, *Ann. Statist.* 13 (2) (1985) 689–705.
- [18] S. Zhang, M. Wong, Wavelet threshold estimation for additive regression models, *Ann. Statist.* 31 (1) (2003) 152–173.