Estimation d'une rupture en dépendance faible

Patrick Ango Nze[1] et Clémentine Prieur[2]

Résumé – On se place dans un cadre de dépendance faible proposé par Doukhan et Louhichi dans [21]. On estime le saut au point de rupture, connu, de la fonction de régression du modèle $Y_n = m(X_n) + \sigma(X_n)\varepsilon_n$. La suite $(\varepsilon_n)_{n\in\mathbb{N}}$ est indépendante et identiquement distribuée (i.i.d.) et indépendante de la suite faiblement dépendante $(X_n)_{n\in\mathbb{N}}$. On énonce un théorème de limite centrale du saut.

Change point estimation for a weakly dependent sequence

Abstract – Let $(X_n,Y_n)_{n\in\mathbb{N}}$ be a stationary sequence governed by the model $Y_n = m(X_n) + \sigma(X_n)\varepsilon_n$ where $(\varepsilon_n)_{n\in\mathbb{N}}$ is i.i.d. and independent from $(X_n)_{n\in\mathbb{N}}$. The latter sequence satisfy a weak dependence condition proposed by Doukhan and Louhichi in [21]. We provide a Central Limit Theorem for jumps in the regression function. Our method deals with linear local regression described in [27]. We use a variation on Lindeberg-Rio [55].

Key words and phrases. Stationary sequences, Lindeberg Theorem, Central Limit Theorem, nonparametric regression, weakly dependent sequences, local linear regression, change point.

MSC (2000) 60 E 15, 60 F 05, 60 G 10, 60 G 99, 62 G 08, 62 J 05.

3.1 Introduction, notations

Dans cette note, nous nous intéressons au problème d'estimation d'une rupture dans un modèle de régression : $Y_n = m(X_n) + \sigma(X_n)\varepsilon_n$ où la suite $(\varepsilon_n)_{n\in\mathbb{N}}$ est i.i.d. et indépendante de la suite $(X_n)_{n\in\mathbb{N}}$. On reprend l'étude de ce modèle par Grégoire & Hamrouni [27] dans le cadre indépendant, mais dans un contexte de dépendance faible : la suite $(X_n)_{n\in\mathbb{N}}$ est supposée faiblement dépendante au sens de la définition 8. La première partie est consacrée à la présentation du modèle étudié. On suppose que la fonction de régression présente un unique point de discontinuité τ connu. On cherche à estimer l'amplitude du saut. La méthode d'estimation retenue est une méthode non paramétrique. On estime le saut au point τ, $\gamma(\tau) = m_+(\tau) - m_-(\tau)$, en utilisant un estimateur à droite pour $m_+(\tau)$ et un autre à gauche pour $m_-(\tau)$. Les estimateurs à droite et à gauche choisis sont des estimateurs de régression linéaire locale. L'article [27] motive l'utilisation de cette procédure en vue de résoudre des problèmes de rupture. Disons pour résumer que ce choix permet de gérer les effets de bords. En outre, sur le plan pratique, ce type d'estimateurs se comporte relativement bien pour des tailles d'échantillon modestes.

1. Université Lille 3, UFR AES, BP 149, 59653, Villeneuve d'Ascq Cedex, France. Courriel : angonze@univ-lille3.fr

2. Université de Cergy-Pontoise, Laboratoire de mathématiques, Bâtiment A4, Site Saint Martin, 95011, Cergy-Pontoise Cedex, France. Courriel : prieur@math.u-cergy.fr

Nous renvoyons à [27] pour des références complémentaires. Les résultats de convergence pour l'estimateur de l'amplitude du saut sont présentés dans la deuxième partie. On montre en particulier un théorème de limite centrale (théorème 9 et corollaire 1).

3.1.1 Hypothèses de dépendance faible

Commençons par introduire la notion de dépendance faible utilisée dans cette note. Pour un espace mesurable et normé $(E,\|.\|)$ on définit la classe $\mathbb{L}^\infty$ réunion de la famille, $\mathbb{L}^\infty(E^u)$ $(u \in \mathbb{N}^*)$, des fonctions numériques mesurables et bornées sur l'ensemble E^u. Ce dernier est muni de la norme l^1 :

$$\|x\|_1 = \|x_1\| + \cdots + \|x_u\|, \qquad x = (x_1,\ldots,x_u) \in E^u.$$

Le coefficient de Lipschitz d'une fonction $f : E^u \longrightarrow \mathbb{R}$ étant défini par

$$\mathrm{Lip}(f) = \sup_{x \neq y} \frac{|f(x) - f(y)|}{\|x - y\|_1},$$

on s'attache aux fonctions appartenant à la classe

$$\mathcal{L} = \bigcup_{u=1}^{\infty} \{f \in \mathbb{L}^\infty(E^u) : \ \mathrm{Lip}(f) < \infty, \ \|f\|_\infty \leq 1\}.$$

La définition de dépendance faible qui suit reprend l'article [21].

Définition 8 *Une suite $(Z_n)_{n\in\mathbb{N}}$ de v.a. à valeurs dans $E = \mathbb{R}^D$ sera dite $(\theta,\mathcal{L},\psi)$-faiblement dépendante s'il existe une suite $\theta = (\theta_r)_{r\in\mathbb{N}}$ décroissant vers zéro, une fonction numérique ψ définie pour les quadruplets $(f,g,u,v) \in \mathcal{L}^2 \times \mathbb{N}^{*2}$, telles que pour tout triplet $(u,v,r) \in (\mathbb{N}^*)^2 \times \mathbb{N}$*

$$|\mathrm{Cov}(f(Z_{i_1},\ldots,Z_{i_u}),g(Z_{j_1},\ldots,Z_{j_v}))| \leq \psi(f,g,u,v)\theta_r$$

soit vérifiée pour tout u-uplet $(i_1,\ldots,i_u)$ et tout v-uplet $(j_1,\ldots,j_v)$ dès lors que $i_1 \leq \ldots \leq i_u < i_u + r \leq j_1 \leq \ldots \leq j_v$.

Cette définition présente une alternative intéressante à des conditions de dépendance plus classiques (le mélange, par exemple). Parmi les choix les plus courants de $\psi(f,g,u,v)$, on retient

$$\psi_1(f,g,u,v) = u\,\mathrm{Lip}(f) + v\,\mathrm{Lip}(g), \ \psi_1'(f,g,u,v) = v\,\mathrm{Lip}(g).$$

La ψ_1 dépendance faible s'applique à des modèles markoviens généraux ainsi qu'aux schémas de Bernoulli. La fonction ψ_1' renvoie à l'analogue causal de ψ_1. L'article [21] propose une revue plus détaillée de tels modèles applicatifs. Le lecteur intéressé pourra le consulter avec profit.

3.1.2 Hypothèses sur le modèle

Les observations forment une suite $(X_n,Y_n)_{n\in\mathbb{N}}$ de v.a. à valeurs dans $\mathbb{R}^2$. Elles obéissent au modèle

$$Y_i = m(X_i) + \sigma(X_i)\varepsilon_i$$

pour une fonction $m(\cdot)$ régulière partout sauf au point connu τ, une fonction $\sigma(\cdot)$ continûment dérivable, et une suite de v.a. à valeurs réelles $(\varepsilon_n)_{n\in\mathbb{N}}$ i.i.d. centrée réduite. On suppose de plus que les suites $(X_n)_{n\in\mathbb{N}}$ et $(\varepsilon_n)_{n\in\mathbb{N}}$ sont mutuellement indépendantes.

Les hypothèses suivantes seront utilisées par la suite.

1. $\mathbf{E}\,\varepsilon_0 = 0, \mathbf{E}\,(\varepsilon_0^2) = 1, \mathbf{E}\,|\varepsilon_0|^3 < \infty.$

2. La loi de X_0 est concentée sur l'intervalle [0,1]. Elle possède une densité f strictement positive et continue.
3. Les lois des couples (X_0,X_k) possèdent des densités $f_k(x,y)$ uniformément bornées par rapport à k :
$$\sup_{k\geq 1}\sup_{x,\,y}|f_k(x,y)| < \infty.$$
4. La fonction de régression $m(\cdot) = \mathbb{E}(Y_0|X_0 = \cdot)$ est deux fois continûment dérivable dans chaque intervalle $[0,\tau]$ et $[\tau,1]$. En particulier m et ses deux dérivées successives possèdent des limites à gauche et à droite au point τ.
5. La variance conditionnelle $\sigma^2(\cdot) = \mathrm{Var}\,(Y_0|X_0 = \cdot)$ est continûment dérivable.
6. Le point de saut τ de la fonction $m(\cdot)$ est dans l'intervalle $]0,1[$.

On définit, pour $\nu = 0{,}1$ ou 2,

(i) $m_-^{(\nu)}(\tau) = \lim_{t\nearrow\tau} m^{(\nu)}(t)$,

(ii) $m_+^{(\nu)}(\tau) = \lim_{t\searrow\tau} m^{(\nu)}(t)$,

(iii) $\gamma^{(\nu)}(\tau) = m_+^{(\nu)}(\tau) - m_-^{(\nu)}(\tau)$.

3.1.3 Estimateurs linéaires locaux

Les fonctions m et σ sont inconnues. On utilise des estimateurs à noyau pour les estimer. Les définitions qui suivent reprennent [25]. Soit K un noyau, et $h_n = h$ une fenêtre. Les paramètres réels $\hat{\alpha}_x$ et $\hat{\beta}_x$ sont obtenus en résolvant le problème de minimisation de

$$\sum_{i=1}^{n}(Y_i - \alpha_x - \beta_x(x - X_i))^2 K_i\,(x)\,, \tag{1}$$

avec $K_i\,(x) = K\left(\dfrac{x - X_i}{h}\right)$.

Définition 9 *L'estimateur linéaire local de $m(\cdot)$ est*

$$\widehat{m}(x) = \hat{\alpha}_x = \frac{\sum_{i=1}^{n}\omega_i(x)Y_i}{\sum_{i=1}^{n}\omega_i(x)}\ , \tag{2}$$

avec

$$\omega_i(x) = K_i\,(x)\,(S_2(x) - (x - X_i)S_1(x))$$

où, pour $l \in \mathbb{N}$,

$$S_l(x) = \sum_{i=1}^{n}(x - X_i)^l K_i\,(x)\,.$$

3.1.4 Estimateurs à droite et à gauche

On reprend ici des hypothèses de [27] pour faciliter la lecture du texte. On considère un noyau $K_+ : [-1,0] \to \mathbb{R}_+$ continûment différentiable. On pose alors

(i) $K_l^+ = \int_{-1}^{0} x^l K_+(x)dx$, et $L_l^+ = \int_{-1}^{0} x^l K_+^2(x)dx$, $l \in \mathbb{N}$,

(ii) $B_+ = (K_2^+)^2 - K_3^+ K_1^+$,
(iii) $V_+ = \int_{-1}^0 (K_2^+ - xK_1^+)^2 K_+^2(x)dx$.

On définit $K_-(x) = K_+(-x)$, puis les quantités K_l^-,L_l^-,B_- et V_- relativement au noyau K_-. On a alors $V_+ = V_-$, $B_+ = B_-$ et

$$V_+ = (K_2^+)^2 L_0^+ + (K_1^+)^2 L_2^+ - 2K_1^+ K_2^+ L_1^+.$$

D'autre part, il n'y a aucune perte de généralité à imposer la normalisation $K_2K_0 - K_1^2 = 1$. En effet, si K est un noyau quelconque, $K/\sqrt{K_2K_0 - K_1^2}$ vérifie la contrainte de normalisation. Les noyaux K_+ et K_- conduisent alors à des estimateurs linéaires locaux $\widehat{m}_+$ et $\widehat{m}_-$ définis par la formule (2). L'article [27] motive l'utilisation de l'estimateur $\widehat{\gamma} = \widehat{m}_+ - \widehat{m}_-$ en vue de résoudre des problèmes de rupture.

3.2 Résultats

On suppose vérifiées les hypothèses du paragraphe 3.1.2.

Théorème 7 *On suppose que le processus* $\boldsymbol{X}$ *est* $(\theta,\mathcal{L},\psi_1)$*-faiblement dépendant. Si* $\theta_r = r^{-a}$ *avec* $a \geq 3$, *si* $h_n \to 0$, *et* $nh_n \to +\infty$, *alors, pour tout* $l \in \mathbb{N}$, *et pour tout* $x \in]0,1[$

$$\begin{aligned}
S_l(x) &= nh^{l+1} f(x) K_l (1 + o_P(1)), \\
R_l(x) &= \sum_{i=1}^n (x - X_i)^l K_i(x) \Big(m(X_i) - m(x) - (X_i - x) m'(x) \Big) \\
&= \frac{1}{2} nh^{l+3} m''(x) f(x) K_{l+2} (1 + o_P(1)), \qquad x \neq \tau \\
T_l(x) &= \sum_{i=1}^n (x - X_i)^l K_i^2(x) \sigma^2(X_i) \\
&= nh^{l+1} \sigma^2(x) f(x) L_l (1 + o_P(1)), \\
\xi_l(x) &= \sum_{i=1}^n (x - X_i)^l K_i(x) \sigma(X_i) \varepsilon_i \\
&= O_P\left(n^{1/2} h^{l+(1/2)}\right),
\end{aligned}$$

$$\sum_{i=1}^n \omega_i(x) = n^2 h^4 f^2(x)(1 + o_P(1)). \tag{1}$$

Démonstration. – On commence par poser : $S_l(x) = \sum_{i=1}^n Z_i(x)$. Alors, d'après l'inégalité de Bienaymé-Tchebycheff,

$$S_l(x) = n\,\mathbf{E}(Z_1(x)) + \mathcal{O}_P\left(\sqrt{\mathrm{Var}\,(S_l(x))}\right). \tag{2}$$

On a

$$\begin{aligned}
\text{Var}\ (S_l(x)) &= \text{Var}\left[\sum_{i=1}^{n}(x-X_i)^l K_i(x)\right]\\
&= n\text{Var}\ (x-X_0)^l K_0(x) + 2\sum_{i=1}^{n-1}(n-i)\text{Cov}\left((x-X_0)^l K_0(x),(x-X_i)^l K_i(x)\right)\\
&= nh^{2l+1}f(x)\int u^{2l}K(u)du\,(1+\mathrm{o}(1)) + 2\sum_{i=1}^{n-1}(n-i)A_i(x)
\end{aligned}$$

où $A_i(x) = \text{Cov}\left(Z_1(x), Z_{i+1}(x)\right)$.

Grâce à l'hypothèse d'existence et de bornitude des densités jointes f_j, il vient

$$|A_i(x)| \leq \|f_i\|_\infty h^{2l+2}\left\{\int |u^l K(u)|du\right\}^2 = \mathcal{O}\left(h^{2l+2}\right).$$

L'hypothèse de $(\theta,\mathcal{L},\psi_1)$ dépendance faible conduit d'autre part à

$$|A_i(x)| \leq Ch^{2l-1}\theta_i. \tag{3}$$

Un argument de troncature des indices dû a Tran (voir [21]) conduit alors à la majoration

$$\left|\sum_{i=1}^{n-1}(n-i)A_i(x)\right| \leq C\,n\,h^{2l+2-(3/a)}.$$

En équilibrant les termes de covariance avec le terme de variance, on obtient pour $a \geq 3$,

$$\text{Var}\ (S_l(x)) \leq Cnh^{2l+1}.$$

Combinée avec la relation (2) cette majoration conduit au résultat annoncé pour S_l. Les termes R_l, T_l et ξ_l se traitant de manière analogue, le détail de la preuve est omis ici. Enfin, on déduit (1) de la relation $\sum_{i=1}^n \omega_i(x) = S_2(x)S_0(x) - S_1^2(x)$.

Remarque. – Dans le cas où $h = n^{-\delta}$ ($\delta \in]0,1[$), la condition sur le taux de dépendance s'énonce $\theta_r = r^{-a}$ avec $a > 3\,\delta$.

Théorème 8 *On suppose que le processus $\mathbf{X}$ est $(\theta,\mathcal{L},\psi_1)$ faiblement dépendant. Si le paramètre de fenêtre vérifie $h_n \to 0$ et $nh_n \to +\infty$, si de plus $\theta_r = r^{-a}$ avec $a \geq 3$, alors l'erreur quadratique conditionnelle moyenne de $\widehat{\gamma}(\tau)$ est*

$$\mathbb{E}\left[(\widehat{\gamma}(\tau)-\gamma(\tau))^2|X_1,\ldots,X_n\right] = \left[\frac{\gamma^{(2)}(\tau)}{2}h_n^2B_+\right]^2 + \frac{2\sigma^2(\tau)V_+}{nh_nf(\tau)} + \ o_P\left(h_n^4+\frac{1}{nh_n}\right).$$

Démonstration. – La décomposition de l'erreur quadratique conditionnelle moyenne fait ressortir les termes suivants :

$\mathbf{E}\left[(\widehat{m}_+(\tau)-m_+(\tau))^2|X_1,\ldots,X_n\right]$, $\mathbf{E}\left[(\widehat{m}_-(\tau)-m_-(\tau))^2|X_1,\ldots,X_n\right]$, et $\mathbf{E}\left[(\widehat{m}_+(\tau)-m_+(\tau))(\widehat{m}_-(\tau)-m_-(\tau))|X_1,\ldots,X_n\right]$.

Les deux premiers termes se traitent comme dans le cadre indépendant. On suit les décompositions de [27] :

$$
\begin{aligned}
\mathbb{E}\left[(\widehat{m}_+(\tau)-m_+(\tau))^2|X_1,\ldots,X_n\right] &= \left(\frac{\sum_{i=1}^n \omega_i^+(\tau)(m_+(X_i)-m_+(\tau))}{\sum_{i=1}^n \omega_i^+(\tau)}\right)^2 \\
&\quad + \frac{\sum_{i=1}^n (\omega_i^+(\tau))^2\sigma^2(X_i)}{\left(\sum_{i=1}^n \omega_i^+(\tau)\right)^2} \\
&= (B_n^+)^2 + V_n^+.
\end{aligned}
$$

On établit une formule analogue avec m_-. Par contre, on n'a plus ici l'indépendance de $\widehat{m}_+(\tau)-m_+(\tau)$ et de $\widehat{m}_-(\tau)-m_-(\tau)$ conditionnellement à $X_1,\ldots,X_n$. Cependant, remarquons que conditionnellement à $X_1,\ldots,X_n$, $\omega_i^+(\tau)\omega_i^-(\tau)=0$ presque sûrement : c'est une conséquence du fait que les supports de K_+ et de K_- sont respectivement $[-1{,}0]$ et $[0{,}1]$. On a alors, grâce aux propriétés d'indépendance de la suite $(\varepsilon_i)_{i\in\mathbb{N}}$,

$$\mathbb{E}\left[(\widehat{m}_+(\tau)-m_+(\tau))(\widehat{m}_-(\tau)-m_-(\tau))|X_1,\ldots,X_n\right] = B_n^+\,B_n^-.$$

On retrouve alors le même résultat qu'en indépendant [27].

Théorème 9 *On suppose que le processus $\boldsymbol{X}$ est $(\theta,\mathcal{L},\psi_1')$ faiblement dépendant. Si $\theta_r = r^{-a}$ avec $a\geq 3$, et si $nh^5\to 0$, $nh\to+\infty$ alors*

$$\sqrt{nh}\left(\widehat{m}_-(\tau)-m_-(\tau)\,,\widehat{m}_+(\tau)-m_+(\tau)\right)'$$

converge vers la loi normale en dimension 2

$$\mathcal{N}\left(0,\left(V_+\sigma^2(\tau)/f(\tau)\right)I_2\right)$$

où I_2 est la matrice identité en dimension 2.

Corollaire 1 *On suppose que le processus $\boldsymbol{X}$ est $(\theta,\mathcal{L},\psi_1')$ faiblement dépendant. Si $\theta_r = r^{-a}$ avec $a\geq 3$, et si $nh^5\to 0$, $nh\to+\infty$ alors*

$$\sqrt{nh}\left(\widehat{\gamma}(\tau)-\gamma(\tau)\right) \xrightarrow{Loi} \mathcal{N}\left(0{,}2\tfrac{\sigma^2(\tau)}{f(\tau)}V_+\right).$$

Remarque. – Pour montrer le théorème 9, on utilise une technique semblable à celle développée dans [16]. L'hypothèse de causalité est essentielle en raison des termes de covariances, non symétriques, considérés. Le corollaire 1 découle immédiatement du théorème 9.

Démonstration du théorème 9. – Le contrôle du terme de biais suit la ligne de démonstration de [27].

$$
\begin{aligned}
\sqrt{nh}\left(\mathbb{E}\left(\widehat{m}_+(\tau)|X_1,\ldots,X_n\right)-m_+(\tau)\right) &= \sqrt{nh}\frac{\sum_{i=1}^n \omega_i^+(\tau)\left(m_+(X_i)-m_+(\tau)\right)}{\sum_{i=1}^n \omega_i^+(\tau)} \\
&= \sqrt{nh}\frac{h^2 B_+ m_+^{(2)}(\tau)}{2}\left(1+o_P(1)\right) \\
&= \mathcal{O}_P\left(\sqrt{nh^5}\right).
\end{aligned}
$$

On obtient une borne analogue pour $\sqrt{nh}\left(\mathbb{E}\left(\widehat{m}_-(\tau)|X_1,\ldots,X_n\right)-m_-(\tau)\right)$. On déduit de la décomposition

$$\widehat{\gamma}(\tau)-\gamma(\tau)=\left(\widehat{m}_+(\tau)-m_+(\tau)\right)-\left(\widehat{m}_-(\tau)-m_-(\tau)\right),$$

que si $nh^5\to 0$, alors

$$\sqrt{nh}\left(\mathbb{E}\left(\widehat{\gamma}(\tau)|X_1,\ldots,X_n\right)-\gamma(\tau)\right)=\mathrm{o}_P(1). \tag{4}$$

On se propose de prouver le comportement asymptotiquement gaussien de

$$\left(\widehat{m}_+(\tau)-\mathbb{E}\left(\widehat{m}_+(\tau)|X_1,\cdots,X_n\right),\widehat{m}_-(\tau)-\mathbb{E}\left(\widehat{m}_-(\tau)|X_1,\ldots,X_n\right)\right)$$

$$=\left(\frac{\sum_{i=1}^n\omega_i^+(\tau)\sigma\left(X_i\right)\varepsilon_i}{\sum_{i=1}^n\omega_i^+(\tau)},\frac{\sum_{i=1}^n\omega_i^-(\tau)\sigma\left(X_i\right)\varepsilon_i}{\sum_{i=1}^n\omega_i^-(\tau)}\right)$$

en étudiant la v.a.

$$A_{a,b}=a\left(\widehat{m}_+(\tau)-\mathbb{E}\left(\widehat{m}_+(\tau)|X_1,\ldots,X_n\right)\right)+b\left(\widehat{m}_-(\tau)-\mathbb{E}\left(\widehat{m}_-(\tau)|X_1,\ldots,X_n\right)\right),$$

avec $a,b\in\mathbb{R}$. D'après le théorème 7, comme $a\geq 3$, il vient que

$$A_{a,b}=\frac{1}{n^2h^4f^2(\tau)}\left(\sum_{i=1}^n\left(a\omega_i^+(\tau)+b\omega_i^-(\tau)\right)\sigma\left(X_i\right)\varepsilon_i\right)\left(1+\mathrm{o}_P(1)\right).$$

D'autre part, on a la relation

$$\sum_{i=1}^n\omega_i^+(\tau)\sigma\left(X_i\right)\varepsilon_i=S_2^+(\tau)\xi_0^+(\tau)-S_1^+(\tau)\xi_1^+(\tau).$$

On pose

$$\Omega_i^+=\frac{\left(hK_2^+-\left(\tau-X_i\right)K_1^+\right)K_+\left(\frac{\tau-X_i}{h}\right)\sigma\left(X_i\right)\varepsilon_i}{f(\tau)}.$$

On définit la v.a. Ω_i^- par analogie immédiate. Enfin $\Omega_i=a\Omega_i^++b\Omega_i^-$. D'après le théorème 7, il vient alors

$$A_{a,b}=\frac{1}{nh^2}\sum_{i=1}^n\Omega_i+\mathrm{o}_P(n^{-1/2}h^{-1/2}). \tag{5}$$

On applique la variante de la méthode de Lindeberg-Rio décrite dans [16]. Soit une fonction $\varphi:\mathbb{R}\longrightarrow\mathbb{R}$ possédant des dérivées continues et bornées jusqu'à l'ordre trois. Soit η une v.a. réelle gaussienne centrée et réduite. On veut établir

$$\Delta_n\left(\varphi\right)=\mathbb{E}\left(\varphi\left(S_n\right)-\varphi\left(\sigma_n\eta\right)\right)\to 0 \tag{6}$$

avec $S_n=\frac{\sum_{i=1}^n\Omega_i}{\sqrt{nh^3}}$ et $\sigma_n^2=\mathrm{Var}\left(S_n\right)$. Les v.a. Ω_i sont centrées et décorrélées. Alors

$$\begin{aligned}\sigma_n^2 &= \frac{1}{nh^3}\left(\sum_{i=1}^n\mathrm{Var}\left(\Omega_i\right)+2\sum_{i=2}^n(n-i+1)\mathrm{Cov}\left(\Omega_1,\Omega_i\right)\right) &(7)\\ &= \left(a^2+b^2\right)\sigma^2(\tau)\frac{V_+}{f(\tau)}(1+\mathrm{o}(1)). &(8)\end{aligned}$$

On définit $S_{k,n}=\frac{\sum_{i=1}^k\Omega_i}{\sqrt{nh^3}}$, avec $1\leq k\leq n$ et on prend par convention $S_{0,n}=0$.

La propriété de décorrélation décrite plus haut conduit alors à vérifier que

$$\begin{aligned} v_{k,n} &= \text{Var}\,(S_{k,n}) - \text{Var}\,(S_{k-1,n}) \\ &= \frac{1}{nh^3}\text{Var}\,(\Omega_k) + 2\frac{1}{nh^3}\sum_{i=1}^{k-1}\text{Cov}\,(\Omega_i,\Omega_k) \\ &= \frac{(a^2+b^2)\,\sigma^2(\tau)V_+}{n\,f(\tau)} + \text{o}(\frac{1}{n}) \end{aligned}$$

est strictement positif pour n suffisamment grand.

Soit une suite $(\eta_{k,n})_{n\in\mathbb{N}^*,1\leq k\leq n}$ de v.a. gaussiennes indépendantes où la loi de $\eta_{k,n}$ est centrée et de variance $v_{k,n}$. On définit enfin $\varphi_{k,n}(x) = \mathbb{E}\,\varphi\,(x+T_{k,n})$ avec $T_{k,n} = \sum_{i=k+1}^{n}\eta_{i,n}$. Les fonctions $\varphi_{k,n}$ ont les mêmes propriétés de dérivabilité que φ : si pour $0\leq j\leq 3$, $\|\varphi^{(j)}\|_\infty \leq C_j$, alors $\|\varphi^{(j)}_{k,n}\|_\infty \leq C_j$.

Alors il existe $0 < \iota_{k,n} < 1$ tel que

$$\begin{aligned} |\Delta_n(\varphi)| &\leq \sum_{k=1}^{n}\left|\Delta^1_{k,n}(\varphi)\right| + \sum_{k=1}^{n}\left|\Delta^2_{k,n}(\varphi)\right| \\ \Delta^1_{k,n}(\varphi) &= \mathbb{E}\varphi_{k,n}(S_{k-1,n}+\eta_{k,n}) - \mathbb{E}\varphi_{k,n}(S_{k-1,n}) - \frac{v_{k,n}}{2}\mathbb{E}\varphi^{(2)}_{k,n}(S_{k-1,n}), \\ \Delta^2_{k,n}(\varphi) &= \mathbb{E}\varphi_{k,n}(S_{k,n}) - \mathbb{E}\varphi_{k,n}(S_{k-1,n}) - \frac{v_{k,n}}{2}\mathbb{E}\varphi^{(2)}_{k,n}(S_{k-1,n}) \\ &= \mathbb{E}\varphi'_{k,n}(S_{k-1,n})\frac{\Omega_k}{\sqrt{nh^3}} + \frac{1}{2}\mathbb{E}\varphi^{(2)}_{k,n}(S_{k-1,n})\left(\frac{\Omega_k^2}{nh^3} - v_{k,n}\right) \\ &\quad + \frac{1}{6}\mathbb{E}\varphi^{(3)}_{k,n}\left(S_{k-1,n} + \iota_{k,n}\frac{\Omega_k}{\sqrt{nh^3}}\right)\left(\frac{\Omega_k}{\sqrt{nh^3}}\right)^3. \end{aligned} \tag{9}$$

On borne $\Delta^1_{k,n}(\varphi)$ par

$$\begin{aligned} \left|\frac{1}{6}\mathbb{E}\varphi^{(3)}_{k,n}\left((S_{k-1,n}+\iota_{k,n}\eta_{k,n})\,\eta^3_{k,n}\right)\right| &\leq \frac{\|\varphi^{(3)}_{k,n}\|_\infty}{6}\mathbb{E}\left|\eta^3_{k,n}\right| \\ &\leq \frac{2\|\varphi^{(3)}_{k,n}\|_\infty v^{3/2}_{k,n}}{3\sqrt{2\pi}} = \mathcal{O}\left(n^{-3/2}\right). \end{aligned} \tag{10}$$

D'où

$$\sum_{i=1}^{n}\left|\Delta^1_{k,n}(\varphi)\right| = \mathcal{O}\left(n^{-1/2}\right) \to 0. \tag{11}$$

Pour étudier $\Delta^2_{k,n}$ on utilise le développement de Taylor (9).

En utilisant la propriété d'indépendance de la suite centrée $(\varepsilon_i)_{i\in\mathbb{N}}$, il vient

$$\mathbb{E}\varphi'_{k,n}(S_{k-1,n})\frac{\Omega_k}{\sqrt{nh^3}} = 0. \tag{12}$$

D'autre part, comme les Ω_i sont décorrélées, et en notant $\Omega_i = \Psi_n(X_i)\,\varepsilon_i$,

$$\begin{aligned}
\mathbb{E}\varphi^{(2)}_{k,n}\left(S_{k-1,n}\right)\left(\frac{\Omega_k^2}{nh^3} - v_{k,n}\right) &= \mathrm{Cov}\left(\varphi^{(2)}_{k,n}\left(S_{k-1,n}\right), \frac{\Omega_k^2}{nh^3}\right) \\
&= \sum_{i=1}^{k-1} \mathbb{E}\left(\varphi^{(3)}_{k,n}\left(S^*_{i-1,n}\right)\frac{\Omega_i}{\sqrt{nh^3}}\frac{\Psi_n^2(X_k)}{nh^3}\right)\mathbb{E}(\varepsilon_k^2) \\
&\quad - \sum_{i=1}^{k-1}\left(\mathbb{E}\varphi^{(3)}_{k,n}\left(S^*_{i-1,n}\right)\frac{\Omega_i}{\sqrt{nh^3}}\right)\left(\mathbb{E}\frac{\Psi_n^2(X_k)}{nh^3}\mathbb{E}\left(\varepsilon_k^2\right)\right),
\end{aligned}$$

avec $S^*_{i-1,n} = S_{i-1,n} + \iota_i \frac{\Omega_i}{\sqrt{nh^3}}$, $0 < \iota_i < 1$. D'où

$$\begin{aligned}
&\left|\mathbb{E}\varphi^{(2)}_{k,n}\left(S_{k-1,n}\right)\left(\tfrac{\Omega_k^2}{nh^3} - v_{k,n}\right)\right| \\
&\qquad \leq \tfrac{\|\varphi^{(3)}_{k,n}\|_\infty}{\sqrt{n^3h^9}} \textstyle\sum_{i=1}^{k-1} \left(\mathbb{E}\left|\Psi_n(X_i)\varepsilon_i\Psi_n^2(X_k)\right| + \mathbb{E}\left|\Psi_n(X_i)\varepsilon_i\right|\mathbb{E}\left(\Psi_n^2(X_k)\right)\right) \\
&\qquad \leq C \textstyle\sum_{i=1}^{k-1} \dfrac{h^{1/2}}{n^{3/2}}.
\end{aligned}$$

De plus, puisque d'une part ε_k est indépendante de $(\varepsilon_1, \ldots, \varepsilon_{k-1}, \Psi_n(X_1), \ldots, \Psi_n(X_k))$, et d'autre part X_k est indépendante de $(\varepsilon_1, \ldots, \varepsilon_{k-1})$,

$$\begin{aligned}
\mathrm{Cov}\left(\varphi^{(2)}_{k,n}\left(S_{k-1,n}\right), \frac{\Omega_k^2}{nh^3}\right) &= \sum_{i=1}^{k-1}\mathrm{Cov}\left(\varphi^{(3)}_{k,n}\left(S^*_{i-1,n}\right)\frac{\Omega_i}{\sqrt{nh^3}}, \frac{\Psi_n^2(X_k)\varepsilon_k^2}{nh^3}\right) \\
&= \sum_{i=1}^{k-1}\mathbb{E}\left(\mathrm{Cov}\left(\varphi^{(3)}_{k,n}\left(S^*_{i-1,n}\right)\frac{\Omega_i}{\sqrt{nh^3}}, \frac{\Psi_n^2(X_k)}{nh^3}\middle|\varepsilon_1, \cdots, \varepsilon_{k-1}\right)\right).
\end{aligned}$$

En utilisant la dépendance des X_i on obtient :

$$\begin{aligned}
\left|\mathrm{Cov}\left(\varphi^{(2)}_{k,n}\left(S_{k-1,n}\right), \frac{\Omega_k^2}{nh^3}\right)\right| &\leq \sum_{i=1}^{k-1}\frac{\|\varphi^{(3)}_{k,n}\|_\infty\|\Psi_n\|_\infty\mathbb{E}|\varepsilon_i|\,\mathrm{Lip}\left(\Psi_n^2\right)}{\sqrt{nh^3}\,nh^3}\theta_{k-i} \\
&\leq \sum_{i=1}^{k-1}\frac{\theta_i}{n^{3/2}h^{5/2}}. \qquad (13)
\end{aligned}$$

On en tire que

$$\begin{aligned}
\sum_{k=1}^{n}\left|\frac{1}{2}\mathbb{E}\varphi^{(2)}_{k,n}\left(S_{k-1,n}\right)\left(\frac{\Omega_k^2}{nh^3} - v_{k,n}\right)\right| &\leq C\sum_{k=1}^{n}\sum_{i=1}^{k-1}\min\left(\frac{h^{1/2}}{n^{3/2}}, \frac{\theta_i}{n^{3/2}h^{5/2}}\right) \qquad (14) \\
&\leq C\,\frac{1}{\sqrt{nh}}\sum_{i=1}^{n-1}\min\left(h, \frac{\theta_i}{h^2}\right) = \mathcal{O}\left(\frac{h^{1-(3/a)}}{\sqrt{nh}}\right). \qquad (15)
\end{aligned}$$

Comme $nh \to \infty$ et $a \geq 3$, cette borne tend vers 0.

Enfin,

$$\frac{1}{6}\mathbf{E}\left|\varphi_{k,n}^{(3)}\left(S_{k-1,n}+\iota_{k,n}\frac{\Omega_k}{\sqrt{nh^3}}\right)\left(\frac{\Omega_k}{\sqrt{nh^3}}\right)^3\right| \quad\leq\quad C\,\mathbf{E}\left|\frac{\Psi_n^3(X_k)}{\sqrt{n^3h^9}}\right|\mathbf{E}\left|\varepsilon_k^3\right|$$

D'où

$$\sum_{k=1}^{n}\left|\frac{1}{6}\mathbf{E}\varphi_{k,n}^{(3)}\left(S_{k-1,n}+\iota_{k,n}\frac{\Omega_k}{\sqrt{nh^3}}\right)\left(\frac{\Omega_k}{\sqrt{nh^3}}\right)^3\right| \quad=\quad \mathcal{O}\left(\frac{1}{\sqrt{nh}}\right)\to 0. \tag{16}$$

Ceci achève la démonstration.

Remarque. – Si le paramètre de fenêtre vérifie la condition $h=n^{-\delta}$ avec $1/5<\delta<1$, alors le théorème de limite centrale 9 est valable pour $a>6\delta/(1+\delta)$.