

### Inférence statistique par lissage linéaire local pour une fonction de régression présentant des discontinuités Zouhir Hamrouni

### ▶ To cite this version:

Zouhir Hamrouni. Inférence statistique par lissage linéaire local pour une fonction de régression présentant des discontinuités. Modélisation et simulation. Université Joseph-Fourier - Grenoble I, 1999. Français. NNT: . tel-00004840

### HAL Id: tel-00004840 https://theses.hal.science/tel-00004840

Submitted on 18 Feb 2004

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers. L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# THÈSE

présentée par

### Zouhir HAMROUNI

pour obtenir le titre de Docteur de l'Université de Joseph Fourier - Grenoble I (Arrêtés ministériels du 5 juillet 1984 et du 30 Mars 1992)

Spécialité : Mathématiques Appliquées

# INFÉRENCE STATISTIQUE PAR LISSAGE LINÉAIRE LOCAL POUR UNE FONCTION DE RÉGRESSION PRÉSENTANT DES DISCONTINUITÉS

Thèse soutenue le 20 janvier 1999 devant la Commission d'Examen

A. Le BRETON	Président du jury
I. GIJBELS	Rapporteur
M. LEJEUNE	Rapporteur
A. ANTONIADIS	Examinateur
J. DIEBOLT	Examinateur
M. LAVIELLE	Examinateur
G. GRÉGOIRE	Directeur de thèse

Thèse préparée au sein du Laboratoire de  $\mathbf{M}$ odélisation et Calcul

# Table des matières

Introduction 3				
	А	A L'analyse de ruptures, présentation		3
		A.1	Motivations	3
		A.2	Exemples	4
		A.3	Etudes ON-LINE et OFF-LINE	6
	В	L'anal	yse de ruptures, position du problème	7
		<b>B.</b> 1	Modèle	7
		B.2	Notations	8
		B.3	Les problèmes traités : détection, estimation et segmentation $\ldots \ldots \ldots$	8
	С	Une re	evue sélective de travaux concernant l'analyse de rupture	9
	D	Un pa	rcours de notre approche par "régression linéaire locale"	12
1	$\mathbf{Esti}$	imatio	n des ruptures	17
	1.1	Introd	uction	17
	1.2	1.2 Modèle et Notations		
	1.3	3 Estimation de l'amplitude lorsque la localisation est connue		
		1.3.1	Construction de l'estimateur	20
		1.3.2	Erreur quadratique moyenne	21
		1.3.3	Normalité Asymptotique	26
	1.4 La situation où la localisation de la discontinuité est inconnue			30
		1.4.1	Approximation par la somme des éléments d'une ligne d'un tableau trian-	
			gulaire	31
		1.4.2	La situation où le noyau s'annule en zéro	42
		1.4.3	La situation où le noyau est strictement positif en zéro	48
	1.5	Simula	ations	57
	1.6	Annex	се А	84

<b>2</b>	2 Tests de rupture					
	2.1	Introduction	87			
	2.2	Test Strictement Local	38			
		2.2.1 Hypothèses	39			
		2.2.2 Résultats asymptotiques	39			
	2.3	Test Local	92			
		2.3.1 Résultats asymptotiques	93			
	2.4	Test Global $\ldots \ldots \ldots$	08			
	2.5	Simulations	33			
3	Est	stimateur du nombre de ruptures 141				
	3.1	Introduction	41			
	3.2	Construction de l'estimateur du nombre de ruptures	41			
	3.3	Hypothèses $\ldots \ldots \ldots$	44			
	3.4	Consistance des estimateurs	44			
	3.5	Vitesses de convergence	53			
	3.6	Simulations	71			
4	mentation 17	79				
	4.1	Introduction	79			
	4.2	Effet de la segmentation en deux intervalles 18	81			
		4.2.1 Notations et hypothèses	81			
		4.2.2 Résultats asymptotiques	83			
	4.3	Segmentation basée sur l'estimateur de la localisation de la rupture 20	04			
	4.4	Plusieurs points de rupture	06			
		4.4.1 Notations et hypothèses	06			
		4.4.2 Résultats asymptotiques	06			
	4.5	Choix des estimateurs du nombre et des localisations des ruptures 20	09			
	4.6 Simulations					
	4.7	Données réelles : Exemple du Nil	24			

### Bibliographie

à mon papa et à ma maman.

### Remerciements

Mes premiers remerciements vont à Monsieur Alain LE BRETON, professeur à l'Université Joseph Fourier de Grenoble, qui me fait l'honneur de présider le jury.

Je suis très reconnaissant envers Madame Irène GIJBELS, Professeur à l'Université Louvain-Le-Neuve, et Monsieur Michel LEJEUNE Professeur associé au CNAM, qui ont voulu montrer de l'intérêt à ce travail en acceptant de rapporter sur cette thèse.

Je remercie également Monsieur Anestis ANTONIADIS, Professeur à l'Université Joseph Fourier de Grenoble, Monsieur Jean DIEBOLT, Directeur de recherche au CNRS, et Monsieur Marc LAVIELLE, Professeur à l'Université Paris-Sud d'Orsay, d'avoir accepté de faire partie du jury.

Enfin, je tiens à remercier particulièrement Monsieur Gérard GREGOIRE, Professeur à l'Université Pierre-Mendès France de Grenoble, qui est à l'origine de cette thèse et grâce à qui j'ai pu mener ce travail à son terme. Il m'est difficile de résumer en quelques mots tout ce qu'il m'a apporté par son implication tout au long de cette thèse, par son soutien et ses encouragements, ainsi que par la sympathie qu'il m'a témoigner tout au long de ce travail.

## Introduction

### A L'analyse de ruptures, présentation

### A.1 Motivations

Lorsqu'on observe un phénomène sur une longue période, on peut se demander s'il est stable sur toute la période étudiée. Il peut arriver que certaines circonstances provoquent une altération du modèle à partir d'une certaine date  $\tau$ . Il peut s'agir pour une fonction de régression d'un saut de cette fonction ou d'une rupture de pente. Lorsqu'on s'intéresse à un processus aléatoire, cela peut être un changement de valeurs des paramètres décrivant la loi de la variable d'état, ou un changement dans les paramètres définissant la covariance, etc. La façon dont l'altération se traduit dépend de l'objet étudié et de la modélisation adoptée.

L'un des premiers problèmes de rupture étudié concernait le contrôle d'une chaîne de production, où il est en effet essentiel de détecter une baisse du régime de production pour pouvoir, par exemple, remplacer une machine défectueuse. Les premières méthodes utilisées étaient principalement graphiques : on estimait, régulièrement, la moyenne de production et on en faisait la représentation graphique. Si celle-ci passait au dessous d'une certaine valeur limite, on concluait alors à une baisse du régime de production et donc à une rupture. Un autre secteur où très tôt des chercheurs se sont intéressés à des problèmes de rupture est celui de la qualité. Page [PAG 57], par exemple, s'intéressait aux ruptures dans la qualité de production d'une machine. Le dérèglement d'une machine entraîne une augmentation de la proportion d'objets défectueux dans la production. La rupture ici est synonyme de baisse notable de la qualité du produit industriel concerné.

De manière générale nous appelons rupture tout changement dans les paramètres du système qui survient instantanément ou très vite à l'échelle de la période d'échantillonnage. Remarquons qu'il ne s'agit pas nécessairement de changements de forte amplitude. Un changement même de faible amplitude peut être révélateur d'un dysfonctionnement du système étudié, et il peut être fondamental à plus d'un titre (économique, sécuritaire, etc) de diagnostiquer ce dysfonctionnement. Le terme de rupture est souvent entendu au sens de "date" à laquelle les propriétés du système ont subitement changé, étant entendu que ces propriétés restent stables pendant un certain temps avant et après la rupture. C'est cette idée qui est à la base de la formalisation mathématique de la notion de rupture. L'expression "détection de ruptures" renvoie aux outils d'aide à la décision concernant le fait qu'un changement est intervenu ou pas dans les caractéristiques de l'objet considéré.

Nous évoquons ci-dessous quelques exemples de situations où des problèmes de ruptures sont étudiés.

### A.2 Exemples

Dans le domaine de la statistique mathématique, les modèles de ruptures connaissent depuis quelques années un développement important. Le champ d'application est large : données médicales, traitement du signal, seismographie, finance, écologie, etc. Les conséquences d'une rupture peuvent être primordiales : contamination par un virus (ruptures du taux de certains marqueurs dans le sang), tremblement de terre (changement de régularité des courbes fournies par les traceurs d'activité du sol), déséquilibre dans le système écologique (forte variation subite de la densité de certaines espèces), changement de tendance sur un marché financier, crash boursier, (saut dans le prix de certaines actions type) ... Nous évoquons ci-dessous de manière plus détaillée quelques exemples de situations où des problèmes de ruptures sont étudiés.

### Contrôle de qualité

Une des applications les plus connues de la détection de rupture est le problème de contrôle de qualité, ou la surveillance en continu de la production. Considérons un procédé de production qui peut être sous contrôle (fonctionnement correspondant aux normes) ou hors contrôle (fonctionnement hors normes). Les situations où ce procédé est hors contrôle sont considérées comme des situations de dysfonctionnement. Pour de nombreuses raisons, il est souhaitable de détecter le dysfonctionnement et d'estimer sa date d'incidence. Cela peut être une question de sécurité du procédé industriel, de contraintes imposées par le marché sur le produit , ou de classification de l'article produit, ... Il est donc souvent important de détecter le plus rapidement possible un dysfonctionnement. Par ailleurs les fausses alarmes, situations où la procédure de suivi annonce à tort un dysfonctionnement, ont aussi un coût important puisqu'elles conduisent à un arrêt de la production et peuvent génèrer chez le personnel une démobilisation conséquence de la perte de confiance dans la procédure de contrôle. Il faut donc trouver un compromis entre une détection rapide et un minimum de fausses alarmes en utilisant un critère tenant compte des pertes impliquées par ces deux événements.

#### Traitement du signal

Le problème de détection de changement dans les caractéristiques d'un signal est courant en traitement de la parole, en médecine, en géophysique ou encore en sismologie. Un exemple classique est celui rencontré sur les plates-formes pétrolières : connaissant les vibrations intrinsèques d'une plate-forme émises en réponse à la houle, grâce à de simples sismographes, la question après passage de la vague du siècle est de décider si oui ou non la structure nécessite des réparations. C'est à dire si oui ou non il est nécessaire de ramener la plate-forme à terre pour remise en état standard. L'idée est de répondre à ce genre de question à l'aide de tests, dits de rupture.

#### Santé

Dans le domaine de la santé, on rencontre aussi de nombreux problèmes de rupture, principalement dans des situations de surveillance au cours du temps. Dans le diagnostic précoce du Sida chez un sujet à partir de l'observation de l'évolution dans le temps de données biologiques chez ce sujet, on sait que le cycle naturel de l'évolution de ce syndrome se présente ainsi :

- contamination du sujet : infection par le virus HIV,
- période de HIV-positivité ou de séro-positivité : l'organisme développe des anticorps anti-HIV dont on peut déceler la présence dans le sang,
- période de Sida avéré: le sujet présente l'un des symptômes entrant dans la définition du syndrome du Sida, ce qui conduit à une dégradation de son état de santé, et, à moyen terme au décès.

Le processus décrit comporte donc deux ruptures. La première, la rencontre avec le virus, dépend du comportement du sujet. La seconde, le passage de la séropositivité à la maladie avérée, présente un intérêt médical évident, et sa détection rapide, voire sa prédiction sont d'un intérêt primordial pour le comportement thérapeutique.

#### Écologie : problème du Nil

La figure 0.1 représente le débit annuel du Nil entre 1871 et 1970. Ces données sont fournies par Cobb [COB78] qui les a analysées en utilisant une approche paramétrique. La question étant : oui ou non y a t-il eu , dans la période fin du siècle dernier-début de ce siècle, une chute brutale du niveau du Nil; et si oui, quand s'est-elle produite? Cobb diagnostique une rupture en 1898 et suggére qu'elle est due à un changement dans le climat à cette époque, se traduisant par une baisse de la pluviosité. La question présentait un certain intérêt pour ceux qui voulaient voir un lien entre la baisse du débit du Nil et la réalisation du barrage d'Assouan par les Anglais dans la période fin du siècle dernier-début de ce siècle. En réalité le barrage d'Assouan a commencé à être opérationnel en 1902.



Figure 0.1 : Le problème du Nil.

### A.3 Etudes ON-LINE et OFF-LINE

Pour des raisons historiques et techniques, on peut décomposer la problématique des modèles de ruptures en deux grandes catégories :

Etude ON-LINE: Les données sont supposées arriver au fur et à mesure et elles sont analysées en direct (analyse séquentielle). On cherche à détecter une rupture sur le modèle le plus rapidement possible, avec un pourcentage fixé de fausses alarmes avant l'instant de rupture τ. On cherche surtout à minimiser l'intervalle entre τ et l'instant de déclenchement de l'alarme τ<sub>a</sub>. Les procédures ON-LINE s'écrivent sous forme de temps d'arrêt markoviens:

$$\tau_a = \inf \{k \ge 1, g_k(Y_1, \dots, Y_k) \ge S_0\}$$

• Etude OFF-LINE : Au contraire, dans les algorithmes de détection OFF-LINE, les données sont collectées et ensuite analysées en différé. Le but peut être de tester l'existence de

ruptures dans le modèle ou d'estimer leur nombre, leurs localisations et leurs amplitudes.

Dans toute la suite de ce document nous ne nous intéressons qu'à cette dernière problématique. Nous la formalisons plus précisément ci-dessous.

### **B** L'analyse de ruptures, position du problème

Les formalisations du problème de rupture étudié peuvent différer par :

- Le type de modèle adopté : modèles de régression, modèles associés à la densité de probabilité ou la densité spectrale ou à la fonction de hasard, ...
- Le type d'observations : discrètes ou continues.
- Les hypothèses sur les ruptures :
  - type des discontinuités (sauts, points de rebroussement, ... ),
  - points de rupture simples ou multiples,
  - le nombre, les amplitudes des ruptures sont supposés connus ou non.
- Cadre de travail : paramétrique, semi-paramétrique ou non-paramétrique.

Le travail que nous présentons dans ce document se situe dans le cadre non-paramétrique. Nous précisons ce cadre dans le paragraphe suivant.

### B.1 Modèle

Soit  $(X_1, Y_1), \ldots, (X_n, Y_n)$  un échantillon de variables aléatoires bivariées. Nous désignons par  $f_{X,Y}(\cdot, \cdot)$  la densité conjointe du couple (X, Y) dans l'espace  $(\mathbb{R}^2, \mathcal{B}(\mathbb{R}^2))$ . Nous supposons que ces observations sont régies par le modèle de régression classique :

$$Y_i = m(X_i) + \sigma(X_i)\varepsilon_i, \qquad i = 1, \dots, n,$$

où  $m(\cdot)$  est la fonction de régression et  $\sigma^2(\cdot)$  la fonction de variance. Les  $\varepsilon_i$  sont des variables aléatoires indépendantes centrées réduites. Les fonctions  $m(\cdot)$  et  $\sigma(\cdot)$  sont inconnues et astreintes uniquement à satisfaire de simples conditions de régularité. Ce cadre très général, motive le recours à une méthode non-paramétrique.

Les notations définies ci-dessous seront valables tout au long du document.

### **B.2** Notations

- $f_{X,Y}(\cdot, \cdot)$ : la densité conjointe du couple (X, Y);
- $f_X(\cdot)$ : la densité marginale de X;
- $m(\cdot) = \mathbb{E}[Y|X = \cdot]$ : la fonction de régression de Y en X;
- $\sigma^2(\cdot) = Var(Y \mid X = \cdot)$ : la fonction de variance conditionnelle de Y sachant X;
- $m_{-}^{(\nu)}(\tau) = \lim_{t \nearrow \tau} m^{(\nu)}(t), \ \nu = 0, 1, 2;$
- $m_+^{(\nu)}(\tau) = \lim_{t \searrow \tau} m^{(\nu)}(t), \ \nu = 0, 1, 2;$

• 
$$\gamma^{(\nu)}(\tau) = m_+^{(\nu)}(\tau) - m_-^{(\nu)}(\tau), \ \nu = 0, 1, 2;$$

- p: le nombre de points de discontinuité de la fonction de régression  $m(\cdot)$ ;
- $\tau_1, \ldots, \tau_p$ : les p points de discontinuité de la fonction de régression;
- Pour tout 1 ≤ i ≤ p, γ(τ<sub>i</sub>) = γ<sup>(0)</sup>(τ<sub>i</sub>): l'amplitude du saut de la fonction de régression au point τ<sub>i</sub>.

### B.3 Les problèmes traités : détection, estimation et segmentation

L'analyse de ruptures regroupe trois problèmes différents et selon les cas on s'intéresse à tout ou partie de ces problèmes.



Figure 0.2 : L'analyse de ruptures : Algorithme

- 1. **Détection** : on cherche à détecter une rupture sur le modèle i.e. prendre un décision sur le fait qu'il y a ou non une rupture dans le modèle.
- 2. Estimation : La présence de rupture(s) est un fait connu. On cherche, au vu des observations, à estimer les paramètres des ruptures dans le modèle :
  - *p* : le nombre de ruptures,
  - $\tau_1, \ldots, \tau_p$ : les instants de rupture,
  - $\gamma(\tau_1), \ldots, \gamma(\tau_p)$ : les amplitudes des ruptures.
- 3. Segmentation: on cherche à identifier la fonction de régression entre les points de ruptures, c'est à dire pour tout point  $x \in [\tau_i, \tau_{i+1}]$ , on veut estimer m(x).

### C Une revue sélective de travaux concernant l'analyse de rupture dans un cadre non-paramétrique

Comme nous l'avons dit plus haut, nous avons sélectionné pour notre travail une approche *off-line* non paramétrique pour traiter en tout premier lieu des problèmes de discontinuités d'ordre zéro dans le cadre du modèle défini par

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

De manière précise, le choix que nous avons fait est celui d'une approche basée sur le lissage à droite et à gauche par régression linéaire locale. La motivation pour un tel choix est double. Tout d'abord, partant de l'idée que les discontinuités d'ordre zéro peuvent être considérées comme des points frontières il s'agissait de tirer parti de la propriété d'absence d'effet de bord de la méthode de régression polynomiale locale : près des bords, la fonction de régression peut être estimée avec la même vitesse de convergence que dans la partie centrale. L'autre motivation relève plus de l'expérience de la pratique : il semble que la méthode de régression polynomiale locale se comporte de façon très satisfaisante pour des tailles d'échantillon relativement modestes.

Les méthodes développées dans le cadre non paramétrique pour traiter les problèmes d'inférence statistique sur les ruptures de la fonction  $m(\cdot)$  peuvent être majoritairement rattachées à deux familles d'approches.

Tout d'abord les méthodes qui modélisent la fonction  $m(\cdot)$  sur [0, 1] sous la forme

$$m(x) = m_0(x) + \sum_{j=1}^k \gamma_j p_{\tau_j}^{(j)}(x), \qquad (0.1)$$

où  $p_{\tau_j}^{(j)}(x) = x^j I_{[\tau_j,1]}(x)$  définit une discontinuité d'ordre j au point  $\tau_j$  et  $m_0(\cdot)$  est une fonction lisse. Cette approche est souvent qualifiée de semi-paramétrique. Dans le cadre du modèle (0.1), l'inférence concernant la fonction lisse  $m_0(\cdot)$ , les paramètres de localisation  $\tau_1, \ldots, \tau_k$  et les paramètres d'amplitude  $\gamma_1, \ldots, \gamma_k$  est conduite en utilisant des méthodes de lissage diverses. L'idée la plus naturelle consiste à estimer les paramètres par un critère de moindres carrés avec une pénalisation portant sur la lissitude de la fonction  $m_0(\cdot)$ . D. Girard [GIR90], Laurent et Utreras [LU86], Shiau [SH185], Cline & al [ECS95] ont eu recours à cette approche. Une approche basée sur la méthode du noyau a été développée dans le cadre de l'estimation d'une fonction de régression par Eubank et Speckman [ES94], Speckman [SPE95] et dans celui de l'estimation d'une fonction d'intensité de processus ponctuel par A. Antoniadis et G. Grégoire [AG93]. Lorsque les localisations sont connues, la méthode consiste à lisser d'abord les observations et les fonctions de discontinuité, puis à projeter les résidus de lissage des observations sur ceux de chaque fonction de discontinuité pour obtenir les estimateur  $\hat{\gamma}_j$  des amplitudes. On déduit alors facilement l'estimateur  $\hat{m}_0(\cdot)$  de la partie lisse. Cette méthode, élégante au plan mathématique conduit à un taux de convergence en  $n^{-4/5}$  pour les estimateurs  $\hat{\gamma}_j$ .

Nous rangeons dans le deuxième groupe de méthodes, les procédures basées sur l'estimation du saut éventuel au point t,  $\gamma(t) = m_{+}(t) - m_{-}(t)$ , obtenue en utilisant un estimateur à droite pour  $m_{+}(t)$  et un autre à gauche pour  $m_{-}(t)$ . Parmi les premiers auteurs à proposer une méthode basée sur ce principe, il faut citer McDonald et Owen [MO86], qui définissaient en fait trois estimateurs en utilisant séparément les observations à droite pour l'un, celles à gauche pour l'autre et enfin des observations de part et d'autre pour le troisième. Par la suite, Müller [MÜL87], Yin [YIN88], Wu et Chu [WC93A], Qiu [QIU94] ont exploré les problèmes de rupture à l'aide de méthodes basées sur ce principe. Dans son travail Yin faisait une hypothèse gaussienne sur le bruit et utilisait des estimateurs à droite et à gauche qui étaient de simples moyennes arithmétiques. Müller [MÜL87] et Wu et Chu [WC93A] ont recours à un lissage par la méthode du noyau. Le premier se place dans le cadre d'un plan d'expérience déterministe et ses estimateurs à droite et à gauche sont construits en utilisant l'estimateur de Gasser et Müller. Dans la situation d'une fonction de régression admettant une discontinuité d'ordre zéro localisée en  $\tau$ , il obtient comme conséquence d'un résultat de normalité asymptotique, la convergence à la vitesse  $a_n/n$  de son estimateur  $\hat{\tau}$ , où  $a_n$  est une suite de réels positifs tendant vers l'infini avec une vitesse qui peut être choisie arbitrairement lente. Il s'ensuit que, en norme  $L_p$ , l'estimateur de la fonction de régression présente la même vitesse de convergence que si la localisation de la rupture était connue. Le travail de Wu et Chu se situe dans le cadre d'un plan d'expérience déterministe et régulier. Les auteurs utilisent des tailles de fenêtre différentes pour estimer la localisation et l'amplitude. Sous un nombre important de conditions, ils obtiennent la convergence p.s. des estimateurs de localisation et d'amplitude ainsi que la normalité asymptotique des couples  $(\hat{\tau}_i, \hat{\gamma}_i)$ . Loader [LOA96], pour construire des estimateurs à droite et à gauche, a utilisé la méthode des polynômes locaux dans le cadre d'un plan d'expérience déterministe et régulier, avec un bruit gaussien de variance unité. Le recours à un noyau ne s'annulant pas en zéro lui permet d'obtenir une vitesse de convergence de  $\hat{\tau}$ , estimateur de la localisation, exactement égale à  $n^{-1}$ .

D'un certain point de vue, les travaux de Gijbels, Hall et Kneip [HGK96, HGK95], et de Müller et Song [MS97] peuvent être considérés comme relevant de l'approche "différences des lissages à droite et à gauche". Les premiers auteurs utilisent d'abord une étape préliminaire qui consiste à estimer le point de plus forte pente de la fonction de régression, méthode que l'on retrouve aussi chez Müller et Wang [WM90] pour l'estimation des ruptures d'une fonction de hasard. Ils recherchent ensuite la discontinuité dans le voisinage de ce point. Cette dernière recherche s'effectue par comparaison des résidus obtenus pour chaque point  $x_{i_0}$  d'une grille de points dans le voisinage, lorsqu'on estime la fonction de régression à droite (respectivement à gauche) par la moyenne des observations à droite (respectivement à gauche). Les auteurs proposent aussi la construction d'une bande de confiance basée sur le boostrap des résidus. Müller et Song ont aussi recours à une localisation préliminaire et proposent comme estimateur  $\hat{\tau}$  le point où le processus de saut estimé, défini comme différence normalisée de moyennes locales à droite et à gauche, réalise son maximum.

Parallèlement aux deux types d'approches dont nous venons de parler, sont apparues aussi des procédures qui ne font appel ni à la représentation semi-paramétrique, ni à l'idée d'une estimation du saut éventuel par différence de lissages à droite et à gauche. Qiu et Yandell  $[\mathbf{Q}\mathbf{Y}98]$ ajustent localement au voisinage de chaque point  $t_i$  du plan d'expérience déterministe, un polynôme  $\hat{Y}^{(i)}(t) = \hat{\beta}_0^{(i)} + \hat{\beta}_1^{(i)}t + \ldots + \hat{\beta}_{p+1}^{(i)}t^{p+1}$ , et batissent une procédure de détection en exploitant le fait que si  $m(\cdot)$  admet une discontinuité en  $t_i$  alors  $\{\hat{\beta}_1^{(j)}\}_{j=l+1}^{n-l}$  montre une forte priorier e  $\hat{\beta}_1^{(i)}$  cut forte variation en  $\hat{\beta}_1^{(i)}$ . Citons aussi les approches par méthodes d'ondelettes développées en particulier par Wang [WAN95], Antoniadis et Gijbels [AG96], Raimondo [RAI96] ... L'idée de base de Wang est que pour certaines résolutions, les coefficients d'ondelettes d'un signal différentiable et ceux du processus de bruit  $\sigma \cdot W$ , W désignant le processus de Wiener, sont du même ordre de grandeur. A cette résolution, au voisinage d'une discontinuité, les coefficients d'ondelette du signal vont subir une forte perturbation qui se retrouvera dans les coefficients d'ondelette du processus observé. La méthode adoptée par Antoniadis et Gijbels consiste à localiser les discontinuités en calculant la transformée en ondelette continue sur une grille de points convenablement choisie et en sélectionnant les points où ces transformées dépassent un seuil adapté. La procédure conduit à une partition en p+1 intervalles sur lesquels le signal est estimé ensuite séparément en utilisant l'analyse multirésolution sur intervalle. Cette revue

des travaux développés dans un cadre non-paramétrique n'est certainement pas exhaustive, mais les contributions évoquées nous semblent assez bien représenter les approches proposées principalement durant les deux dernières décennies.

### D Un parcours de notre approche par "régression linéaire locale"

Etant donné les observations  $(X_1, Y_1), \ldots, (X_n, Y_n)$  provenant du modèle de régression

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

où les  $\varepsilon_i$  sont i.i.d. de moyenne nulle et de variance unité, nous définissons des estimateurs à droite et à gauche de la fonction  $m(\cdot)$  de la manière suivante.  $K_+$  étant un noyau à support  $[-1, 0], K_-(x) = K_+(-x)$ , nous posons  $\hat{m}_+(t) = b(t)$  où b(t) est obtenu en résolvant le problème

suivant

$$\min_{a,b} \sum_{i=1}^{n} [Y_i - (a(t - X_i) - b)]^2 K_+ \left(\frac{t - X_i}{h_n}\right).$$

De même en remplaçant ci-dessus  $K_+$  par  $K_-$ , nous définissons un estimateur  $\hat{m}_-(t)$  de la limite à gauche  $m_-(t)$ . L'utilisation des poids  $K_+\left(\frac{t-X_i}{h_n}\right)$  fait que seules les observations à droite de tsont utilisées pour estimer  $m_+(t)$ , de même seules les observations à gauche sont utilisées pour estimer  $m_-(t)$ . L'estimateur du saut éventuel est donc

$$\hat{\gamma}(t) = \hat{m}_{+}(t) - \hat{m}_{-}(t)$$

Le processus  $t \mapsto \hat{\gamma}(t)$ , avec changement d'échelle et normalisation éventuels, joue un rôle fondamental tout au long de notre travail.

Nous commençons au chapitre I par étudier le problème d'estimation dans la situation d'une seule rupture localisée en  $\tau$ , de signe connu que nous prenons positif. Lorsque  $\tau$  est connue, il est facile d'obtenir l'expression de l'erreur quadratique asymptotique, et de montrer que  $\hat{\gamma}(\tau)$ est asymptotiquement normal. Dans la situation où  $\tau$  n'est pas connue, nous définissons un estimateur  $\hat{\tau}$  et estimons l'amplitude par  $\hat{\gamma}(\hat{\tau}) = \hat{m}_{+}(\hat{\tau}) - \hat{m}_{-}(\hat{\tau})$ . L'estimateur  $\hat{\tau}$  est obtenu comme point où le processus de déviation locale  $\mathcal{Z}_n(z) = \alpha(n, h_n) \left( \hat{\gamma} \left( \tau + \delta(n, h_n) z \right) - \hat{\gamma}(\tau) \right)$ , réalise son maximum. Pour des choix convenables de  $\alpha$  et  $\delta$ , nous montrons que le processus  ${\mathcal Z}_n(z)$  converge vers un processus  ${\mathcal Z}(z)$  qui peut être soit un processus gaussien, soit un processus de Poisson composé. Ces deux classes de processus limites sont en fait associées à des hypothèses différentes sur le noyau:  $K_{+}(0) = 0$  dans le premier cas,  $K_{+}(0) > 0$  dans le second. Nous montrons que ces lois limites de  $\mathcal{Z}_n(z)$  conduisent dans le premier cas à la normalité asymptotique de  $\hat{\tau}$ , et dans le second cas à une convergence vers une loi dépendant de la distribution du bruit  $\varepsilon$ . La convergence de  $\hat{\tau}$  vers  $\tau$  est alors dans le premier cas en  $\mathcal{O}_P(a_n/n)$ , où  $(a_n)$  est une suite tendant vers l'infini avec une vitesse qui peut être arbitrairement lente. Dans le second cas, elle est en  $\mathcal{O}_P(n^{-1})$ . Ces résultats confirment ceux qu'avait obtenu Loader [LOA96] dans un cadre très simplifié. Nous montrons que l'estimateur de l'amplitude,  $\hat{\gamma}(\hat{\tau})$ , est asymptotiquement normal aussi bien dans un cas que dans l'autre.

Le chapitre II est consacré à l'étude des tests d'existence d'une rupture. Trois tests sont présentés. Le premier que nous qualifions de test strictement local, s'intéresse à l'existence d'un saut en un point  $\tau$  quelconque fixé. Il est basé sur la statistique  $\hat{\gamma}(\tau)$  : ses propriétés découlent naturellement des résultats obtenus au chapitre précédent. Nous construisons ensuite un test d'existence d'un saut au point  $\tau$ , que nous appelons test local, en étudiant le comportement du processus  $\hat{\gamma}(t)$ , dans un voisinage du point  $\tau$ . Le principe de ce test repose sur le fait que, en l'absence de rupture, le processus  $\mathcal{Z}_n(z) = \sqrt{nh_n}\hat{\gamma} (\tau + h_n z)$  converge vers un processus  $\mathcal{Z}(z)$ gaussien dont la fonction de covariance satisfait les conditions sous lesquelles les résultats de convergence des valeurs extrêmes de Leadbetter et al. peuvent être appliqués. Il s'ensuit que la statistique sup  $\mathcal{Z}_n(z)/\sqrt{\hat{C}(0)}$ , où  $\hat{C}(0)$  est un estimateur de la variance de  $\mathcal{Z}(z)$ , que nous utilisons comme statistique de test de l'hypothèse d'absence de rupture, est asymptotiquement de loi de Gumbel. Le troisième test que nous appelons test global, est un test de l'existence d'une rupture sur ]0, 1[. Ce test tire profit du fait que le processus  $\xi_n(t) = \sqrt{nh_n}V(t)\hat{\gamma}(t)$ où V(t) dépend de la densité  $f_X(t)$  des  $X_i$ , et de la variance du bruit, est asymptotiquement gaussien sous l'hypothèse nulle. Comme la fonction de covariance du processus limite satisfait aussi les conditions requises pour appliquer les résultats de Leadbetter et al., une procédure de test découle de manière analogue à la situation du test local. Nous montrons que ces trois tests sont asymptotiquement de puissance unité.

Au chapitre III, nous nous plaçons dans la situation d'un nombre p inconnu de discontinuités. Pour estimer p et les localisations  $\tau_1, \ldots, \tau_p$ , nous découpons l'intervalle [0, 1] en intervalles de longueur  $2h_n$ , où  $h_n$  est la taille de la fenêtre utilisée par le noyau. Dénombrant le nombre d'intervalles où  $\hat{\gamma}$  dépasse un certain seuil  $S_0$  en agrégeant les intervalles contigüs, nous obtenons un estimateur naturel  $\hat{p}$  du nombre de discontinuités. Nous montrons que cet estimateur est convergent et que les localisations de dépassement du seuil convergent vers les points de rupture  $\tau_1, \ldots, \tau_p$ . Pour être en mesure d'appliquer une telle procédure, il faut pouvoir choisir  $S_0 <$  $\min(\gamma(\tau_1), \ldots, \gamma(\tau_p))$ , ce qui nécessite une information a priori sur les amplitudes. Pour pallier cette difficulté, une solution consiste à utiliser une suite de seuils  $(S_n)$  convergents vers 0. Sous des conditions liant la vitesse de décroissance de  $(h_n)$  à celle de  $(S_n)$  nous obtenons la convergence p.s. de  $\hat{p}$  vers p et des  $\hat{\tau}_i$  vers les  $\tau_i$ .

Le problème de l'étude de la reconstitution de la fonction de régression dans une situation de rupture fait l'objet du chapitre IV. Une fois les discontinuités localisées aux points  $\hat{\tau}_1, \ldots, \hat{\tau}_{\hat{p}}$ , la fonction de régression est reconstituée sur chaque intervalle  $[\hat{\tau}_i, \hat{\tau}_{i+1}]$  par régression linéaire locale en situation de fonction de régression lisse, en ayant recours à un noyau à support [-1, 1]. En premier lieu nous analysons les conséquences sur  $\hat{m}(\cdot)$  d'une segmentation arbitraire en fonction des localisations réelles des discontinuités. Nous montrons ensuite qu'en utilisant la procédure d'estimation du nombre de ruptures et des localisations développée au chapitre précédent, la segmentation ne conduit à aucune perte de vitesse en ce qui concerne l'erreur quadratique moyenne intégrée sur [0, 1] de l'estimateur  $\hat{m}(\cdot)$ . Il s'agit là d'une conséquence d'une part de l'absence d'effet de bord de la régression linéaire locale, et d'autre part de la forte vitesse de convergence des estimateurs des localisations.

Chacun des chapitres I à IV contient les résultats de simulations effectués la plupart du temps sur des exemples étudiés par d'autres auteurs avec des méthodes concurrentes (Loader [LOA96], Gijbels, Hall et Kneip [HGK96, HGK95]). Ces simulations démontrent l'efficacité de la méthode de régression linéaire locale, et mettent en évidence la sensibilité aux divers paramètres: choix du noyau, taille de l'échantillon, rapport signal sur bruit, et taille de la fenêtre. Le fait que les performances sont meilleures pour un noyau satisfaisant  $K_{\pm}(0) > 0$ que pour un noyau tel que  $K_+(0) = 0$ , est confirmé par les expérimentations. Par ailleurs le simulations montrent l'importance du choix de la taille de la fenêtre, mais fournissent aussi des guides sous la forme d'indications de plages de valeurs raisonnables en fonction de la taille de l'échantillon. Nous n'avons pas traité dans ce document la question de l'élaboration d'une procédure de choix de la taille de fenêtre. Une procédure idéale mais peu réaliste consisterait à chercher une fenêtre  $h_n$  réalisant le minimum du ISE ou du MISE sur [0, 1]. Une estimation d'un tel  $h_n$  pourrait être obtenue par validation croisée. Une autre proposition pourrait être de n'utiliser comme critère de choix de  $h_n$  que la qualité de reconstruction de la fonction au voisinage des discontinuités. Nous étudions actuellement les diverses procédures envisageables pour résoudre cette question.

### Chapitre 1

## Estimation des ruptures

### 1.1 Introduction

Considérons le modèle de régression où les observations  $(X_i, Y_i)$  satisfont :

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

où  $m(\cdot)$  est la fonction de régression et  $\sigma(\cdot)$  est l'écart-type conditionnel. Les  $\varepsilon_i$  sont des variables aléatoires i.i.d centrées et réduites. Nous supposons dans ce chapitre que  $m(\cdot)$  est "lisse", sauf en un point  $\tau$  où elle admet une discontinuité d'amplitude  $\gamma(\tau) = m_{+}(\tau) - m_{-}(\tau)$ . Dans la première partie, nous nous plaçons dans la situation où la localisation au de la discontinuité est connue, nous construisons un estimateur  $\hat{\gamma}(\tau)$  de  $\gamma(\tau)$ , différence des estimateurs  $\hat{m}_{+}(\tau)$  et  $\hat{m}_{-}(\tau)$  des lissages à droite et à gauche de  $m(\cdot)$  au point  $\tau$  obtenus par RLL (Régression Linéaire Locale), Nous montrons que cet estimateur est consistant et asymptotiquement normal. Dans la deuxième partie la localisation  $\tau$  n'est plus supposée connue, nous estimons  $\tau$  en recourant à un processus de déviation locale. L'estimateur  $\hat{\tau}$  ainsi construit conduit à un estimateur  $\hat{\gamma}(\hat{\tau})$ de l'amplitude. Pour étudier les propriétés de  $\hat{\tau}$ , nous envisageons la situation où le noyau unilatéral utilisé  $K_{+}(\cdot)$  est nul à l'origine, et celle où  $K_{+}(0) > 0$ . Loader [LOA96] a montré dans la situation où le plan d'expérience est déterministe et régulier, et en supposant le bruit gaussien, qu'on améliore la vitesse de convergence de  $\hat{\tau}$  lorsque  $K_+(0) > 0$ . Nous observons que le comportement du processus de déviation locale utilisé pour estimer au est différent dans les deux cas  $K_{+}(0) = 0$  et  $K_{+}(0) > 0$ , et que la vitesse de convergence de  $\hat{\tau}$  est meilleure dans ce dernier cas.

### 1.2 Modèle et Notations

Soit  $(X_1, Y_1), \ldots, (X_n, Y_n)$  un échantillon de variables aléatoires bivariées. Nous désignons par  $f_{X,Y}(\cdot, \cdot)$  la densité conjointe du couple (X, Y) dans l'espace  $(\mathbb{R}^2, \mathcal{B}(\mathbb{R}^2))$ . Nous supposons que ces observations sont régies par le modèle de régression classique suivant :

$$Y_i = m(X_i) + \sigma(X_i)\varepsilon_i, \qquad i = 1, \dots, n, \qquad (1.1)$$

où  $m(\cdot)$  est la fonction de régression et  $\sigma(\cdot)$  est l'écart-type conditionnel. Les  $\varepsilon_i$  sont des variables aléatoires **i**.**i**.**d** centrées et réduites. Les fonctions  $m(\cdot)$  et  $\sigma(\cdot)$  sont supposées inconnues, et nous ne leur imposons a priori aucune contrainte supplémentaire. Les contraintes que nous serons amenés à imposer par la suite sont simplement des hypothèses de régularité mathématiques. Ceci justifie, pour les problèmes d'estimation considérés, l'emploi d'une approche non-paramétrique.

### Notations et Hypothèses

- A1.  $f_X(\cdot)$ : la densité marginale de X est supposée continue et strictement positive sur l'intervalle [0, 1].
- A2.  $\sigma^2(x) = Var(Y \mid X = x)$  est supposée aussi continue sur l'intervalle [0, 1].
- A3. La fonction de régression,  $m(\cdot)$ , est supposée de classe  $\mathcal{C}^2$  séparément sur chacun des intervalles  $[0, \tau]$  et  $[\tau, 1]$ .

**Remarque 1.1.** Les hypothèses A1 et A2 peuvent, en fait, être affaiblies. Pour simplifier les notations dans les démonstrations, nos résultats sont énoncés sous les hypothèses A1 et A2. Les modifications à apporter à certains résultats, pour prendre en compte le fait que les fonctions  $f_X(\cdot)$  et/ou  $\sigma(\cdot)$  peuvent être discontinues au point  $\tau$ , se trouvent à la fin du paragraphe (1.3).

On pose

$$\begin{array}{ll} \text{i)} & m_{-}^{(\nu)}(\tau) = \lim_{t \nearrow \tau} m^{(\nu)}(t), \ \nu = 0, 1, 2. \\ \\ \text{ii)} & m_{+}^{(\nu)}(\tau) = \lim_{t \searrow \tau} m^{(\nu)}(t), \ \nu = 0, 1, 2. \\ \\ \\ \text{iii)} & \gamma^{(\nu)}(\tau) = m_{+}^{(\nu)}(\tau) - m_{-}^{(\nu)}(\tau), \ \nu = 0, 1, 2. \end{array}$$

 $\gamma(\tau) = m_+(\tau) - m_-(\tau)$  est l'amplitude du saut de la fonction de régression  $m(\cdot)$  au point  $\tau$ . Sans perte de généralité, dans ce chapitre, nous supposons que  $\gamma(\tau) > 0$ . Le cas  $\gamma(\tau) < 0$  peut être abordé de manière similaire. En effet lorsque  $\gamma(\tau) < 0$ , il suffit de remplacer dans ce qui suit les maximisations par des minimisations, et il est facile d'obtenir les résultats analogues à ceux de ce chapitre dans le cas où  $\gamma(\tau)$  est négatif.

### Choix des noyaux de lissage

**Définition 1.1.** On appelle noyau de lissage à droite  $K_{+}(\cdot)$  (respectivement à gauche  $K_{-}(\cdot)$ ) toute fonction  $K_{+}(\cdot)$  (resp.  $K_{-}(\cdot)$ ) bornée à support [-1,0] (resp. [0,1]). Pour la suite nous allons adopter une notation commune pour désigner l'un des deux noyaux,  $K_{\pm}(\cdot)$ , et nous notons pour tout l = 0, 1, 2, ...

i)  $K_l^{\pm} = \int_{-1}^1 x^l K_{\pm}(x) dx$ ,

ii) 
$$L_l^{\pm} = \int_{-1}^1 x^l K_{\pm}^2(x) dx$$

iii) 
$$B_{\pm} = B(K_{\pm}) = \frac{(K_{2}^{\pm})^{2} - K_{1}^{\pm} K_{3}^{\pm}}{K_{0}^{\pm} K_{2}^{\pm} - (K_{1}^{\pm})^{2}},$$

iv) 
$$V_{\pm}^2 = V^2(K_{\pm}) = \frac{1}{\left(K_0^{\pm}K_2^{\pm} - \left(K_1^{\pm}\right)^2\right)^2} \int_{-1}^1 \left(K_2^{\pm} - K_1^{\pm}x\right)^2 K_{\pm}^2(x) dx$$

Nous supposons également que :

$$K_{-}(x) = K_{+}(-x), \qquad \forall x \in [0, 1].$$
 (1.2)

Tout au long de notre travail, les noyaux  $K_+(\cdot)$  et  $K_-(\cdot)$  sont supposés positifs, ceci afin de conserver l'interprétation de l'estimateur en termes de moindres carrés pondérés (voir 1.4).

Sans perte de généralité, on imposera une sorte de normalisation pour  $K_{\pm}(\cdot)$ :

$$|K_2^{\pm}K_0^{\pm} - (K_1^{\pm})^2| = 1.$$
(1.3)

En effet si  $K_{\pm}(\cdot)$  est un noyau quelconque,  $\frac{K_{\pm}(\cdot)}{\sqrt{|K_2^{\pm}K_0^{\pm}-(K_1^{\pm})^2|}}$  vérifie la contrainte de normalisation (1.3). En pratique puisque  $K_{\pm}(\cdot)$  est positif on  $a(K_1^{\pm})^2 \leq K_2^{\pm}K_0^{\pm}$ .

**Remarque 1.2.** Tout au long de ce travail les noyaux sont supposés à support compact mais les mêmes résultats pourraient être obtenus, au prix de difficultés techniques supplémentaires, pour des noyaux de Parzen-Rosenblat  $(\lim_{|x|\to+\infty} x^l K(x) = 0, l \ge 0).$ 

### 1.3 Estimation de l'amplitude lorsque la localisation est connue

Nous aborderons dans ce paragraphe l'estimation de l'amplitude lorsqu'on connaît la localisation  $\tau$  du point de discontinuité. Les motivations pour étudier le cas où les localisations des singularités sont connues se situent aussi bien au plan de la modélisation qu'à celui de la théorie :

- Dans un certain nombre de situations réelles, nous disposons d'informations a priori sur les localisations des ruptures suite à des études antérieures, les opinions de certains experts ou parce que les données elles-mêmes montrent clairement les localisations des ruptures, ou encore parce que la rupture a été provoquée par une intervention extérieure (en médecine, cette intervention peut être une opération chirurgicale; en économie, il peut s'agir de décisions gouvernementales; etc ...);
- ii) Pour pouvoir étudier les performances asymptotiques des procédures statistiques que nous proposons dans la situation où nous ne disposons pas d'information ni sur le nombre des ruptures ni sur leurs localisations, nous avons besoin de contrôler leurs performances dans le cas où les localisations des singularités sont connues.

Nous allons tout d'abord étudier les propriétés asymptotiques de  $\hat{m}_+(\tau)$  et  $\hat{m}_-(\tau)$ , puisque nous pouvons déduire facilement des résultats analogues pour  $\hat{\gamma}(\tau)$ .

### 1.3.1 Construction de l'estimateur

Rappelons (cf. Fan [FAN92]), que  $K(\cdot)$  étant un noyau, l'estimateur de régression linéaire locale  $\hat{m}(x)$  de la fonction m(x) est défini comme la valeur de b = b(x) obtenue lorsqu'on résout le problème de minimisation :

$$\min_{(a,b)} \sum_{i=1}^{n} \left[ Y_i - (a(x - X_i) + b) \right]^2 K(\frac{x - X_i}{h_n}).$$
(1.4)

 $K_{+}(\cdot)$  étant un noyau à support [-1, 0], et  $K_{-}(\cdot)$  défini par  $K_{-}(x) = K_{+}(-x)$ , nous définissons en remplaçant dans (1.4)  $K(\cdot)$  par  $K_{+}(\cdot)$  et  $K_{-}(\cdot)$  les estimateurs  $\hat{m}_{+}(\cdot)$  et  $\hat{m}_{-}(\cdot)$ , qui peuvent être qualifiés respectivement d'estimateur à droite et d'estimateur à gauche de la fonction de régression  $m(\cdot)$ .

La figure ci-dessous illustre le lissage par RLL à droite et à gauche aux points x = 0.25 et x = 0.75 lorsque la fenêtre de lissage  $h_n$  est égale à 0.10.



Figure 1.1 : Les lissages linéaires locaux à gauche et à droite.

**Définition 1.2.** L'estimateur linéaire local de  $m_{\pm}(\tau)$  est défini comme suit

$$\hat{m}_{\pm}(\tau) = \frac{\sum_{i=1}^{n} w_i^{\pm}(\tau) Y_i}{\sum_{i=1}^{n} w_i^{\pm}(\tau)},\tag{1.5}$$

avec pour tout  $i = 1, \ldots, n$ ,

$$w_i^{\pm}(\tau) = K_{\pm}\left(\frac{\tau - X_i}{h_n}\right) \left(S_2^{\pm}(\tau) - (\tau - X_i)S_1^{\pm}(\tau)\right), \qquad (1.6)$$

 $et \ o \dot{u}$ 

$$S_l^{\pm}(\tau) = \sum_{i=1}^n (\tau - X_i)^l K_{\pm}(\frac{\tau - X_i}{h_n}), \qquad l = 0, 1, 2.$$
(1.7)

Ainsi l'estimateur de  $\gamma(\tau) = m_+(\tau) - m_-(\tau)$  sera  $\hat{\gamma}(\tau) = \hat{m}_+(\tau) - \hat{m}_-(\tau)$ .

### 1.3.2 Erreur quadratique moyenne

Pour étudier les propriétés asymptotiques de  $\hat{\gamma}(\tau)$ , on va commencer par étudier celles de  $\hat{m}_{+}(\tau)$  et  $\hat{m}_{-}(\tau)$ . Les propriétés asymptotiques de  $\hat{m}_{-}(\tau)$  sont analogues à celles de  $\hat{m}_{+}(\tau)$ , c'est

pourquoi dans les démonstrations on va s'intéresser seulement aux propriétés asymptotiques de  $\hat{m}_{+}(\tau)$ . Celles de  $\hat{m}_{-}(\tau)$ , seront déduites en changeant les indices "+" en "-".

**Théorème 1.1.** Supposons que les hypothèses A1-A3 sont satisfaites et que  $h_n \rightarrow 0$  et  $nh_n \rightarrow +\infty$ . L'erreur quadratique moyenne conditionnelle admet le développement asymptotique suivant :

$$\mathbb{E}\left[\left(\hat{m}_{\pm}(\tau) - m_{\pm}(\tau)\right)^{2} \mid X_{1}, \dots, X_{n}\right] = \left[\frac{m_{\pm}^{(2)}(\tau)}{2}h_{n}^{2}B_{+}\right]^{2} + \frac{\sigma^{2}(\tau)}{nh_{n}f_{X}(\tau)}V_{+}^{2} + o_{P}\left(h_{n}^{4} + \frac{1}{nh_{n}}\right).$$

### Preuve du théorème 1.1

L'erreur quadratique moyenne conditionnelle de  $\hat{m}_+(\tau)$  est donnée par :

$$\mathbb{E}\left[\left(\hat{m}_{+}(\tau) - m_{+}(\tau)\right)^{2} \mid X_{1}, \dots, X_{n}\right] \\
= \left[\frac{\sum_{i=1}^{n} w_{i}^{+}(\tau) \left(m(X_{i}) - m_{+}(\tau)\right)}{\sum_{i=1}^{n} w_{i}^{+}(\tau)}\right]^{2} + \frac{\sum_{i=1}^{n} \left(w_{i}^{+}(\tau)\right)^{2} \sigma^{2}(X_{i})}{\left(\sum_{i=1}^{n} w_{i}^{+}(\tau)\right)^{2}}, \\
= \left(B_{n}^{+}\right)^{2} + V_{n}^{+}.$$
(1.8)

 $B_n^+$  représente le biais de  $\hat{m}_+(\tau)$  et  $V_n^+$  sa variance.

$$B_n^+ = \frac{\sum_{i=1}^n w_i^+(\tau) \left(m(X_i) - m_+(\tau)\right)}{\sum_{i=1}^n w_i^+(\tau)},$$
  
=  $\frac{\sum_{i=1}^n w_i^+(\tau) \left(m(X_i) - m_+(\tau) - (X_i - \tau)m_+'(\tau)\right)}{\sum_{i=1}^n w_i^+(\tau)},$ 

La dernière égalité est due au fait que

$$\sum_{i=1}^{n} w_i^+(\tau) (X_i - \tau) = 0.$$

En remplaçant les  $w_i^+(\tau)$  par leurs valeurs,  $B_n^+$  s'écrit sous la forme suivante :

$$B_n^+ = \frac{S_2^+(\tau)R_0^+(\tau) - S_1^+(\tau)R_1^+(\tau)}{S_2^+(\tau)S_0^+(\tau) - (S_1^+(\tau))^2},$$
(1.9)

où pour tout  $l = 0, 1, \ldots$ 

$$S_{l}^{+}(\tau) = \sum_{i=1}^{n} (\tau - X_{i})^{l} K_{+}(\frac{\tau - X_{i}}{h_{n}}),$$

 $\operatorname{et}$ 

$$R_l^+(\tau) = \sum_{i=1}^n (\tau - X_i)^l K_+(\frac{\tau - X_i}{h_n}) \left( m(X_i) - m_+(\tau) - (X_i - \tau)m_+'(\tau) \right)$$

Cette deuxième forme nous sera utile par la suite.

$$V_{n}^{+} = \frac{\sum_{i=1}^{n} (w_{i}^{+}(\tau))^{2} \sigma^{2}(X_{i})}{\left(\sum_{i=1}^{n} w_{i}^{+}(\tau)\right)^{2}},$$
  
$$= \frac{(S_{2}^{+}(\tau))^{2} T_{0}^{+}(\tau) - 2S_{2}^{+}(\tau) S_{1}^{+}(\tau) T_{1}^{+}(\tau) + (S_{1}^{+}(\tau))^{2} T_{2}^{+}(\tau)}{\left(S_{2}^{+}(\tau) S_{0}^{+}(\tau) - (S_{1}^{+}(\tau))^{2}\right)^{2}},$$
 (1.10)

avec

$$T_l^+(\tau) = \sum_{i=1}^n (\tau - X_i)^l K_+^2(\frac{\tau - X_i}{h_n}) \sigma^2(X_i).$$

Il découle alors de (1.9) et du lemme 1.1 (a et b) ci-dessous que :

$$B_n^+ = \frac{m_+^{(2)}(\tau)}{2}h_n^2 B_+ + o_P(h_n^2).$$

et de (1.10) et du lemme 1.1 (**a** et **c**) que

$$V_n^+ = \frac{\sigma^2(\tau)}{nh_n f_X(\tau)} V_+^2 + o_P\left(\frac{1}{nh_n}\right).$$

**Lemme 1.1.** Soit  $x \in (0,1)$  et  $K(\cdot)$  une fonction à support compact inclus dans [-1,1]. Sous les hypothèses que  $f_X(\cdot)$  et  $\sigma^2(\cdot)$  sont continues en x et que  $m(\cdot)$  admet une dérivée seconde continue en ce même point x, alors nous avons pour tout l = 0, 1, ...

a) 
$$S_l(x) = nh_n^{(l+1)} f_X(x) (K_l + o_P(1)),$$
  
b)  $R_l(x) = \frac{1}{2}m^{(2)}(x)nh_n^{(l+3)} f_X(x) (K_{l+2} + o_P(1)), \quad x \neq \tau$   
c)  $T_l(x) = nh_n^{(l+1)} f_X(x)\sigma^2(x) (L_l + o_P(1)),$   
d)  $\xi_l(x) = \mathcal{O}_P\left(n^{1/2}h_n^{l+1/2}\right),$   
e)  $\sum_{i=1}^n w_i(x) = n^2 h_n^4 f_X^2(x) (K_2 K_0 - (K_1)^2) (1 + o_P(1)).$ 

Si les fonctions sont supposées continues sur [0, 1], alors les fonctions  $o_p(1)$  peuvent être choisies de manière indépendante de x.

**Remarque 1.3.** Lorsque nous utilisons  $K_+(\cdot)$  ou  $K_-(\cdot)$  et  $x = \tau$ , la partie b) doit être remplacée par la formule suivante:

$$R_l^+(\tau) = \frac{1}{2}nh_n^{l+3}m_+^{(2)}(\tau)f_X(\tau)K_{l+2}^+ + o_P(nh_n^{l+3}),$$

nous obtenons une formule analogue en changeant le "+" par un "-".

### Preuve du lemme 1.1

- Les démonstrations de certains de ces résultats peuvent être trouvés dans l'article de Fan et Gijbels [FG92]. Nous donnons néanmoins les démonstrations de a) et b). Celles de c) et d) sont analogues à celle de a).
- 2) La partie e) est une conséquence immédiate de la partie a), puisque  $\sum_{i=1}^{n} w_i(x) = S_2(x)S_0(x) S_1^2(x)$ .

On pose

$$Z_{i,1} = (x - X_i)^l K(\frac{x - X_i}{h_n}),$$

 $\operatorname{et}$ 

$$Z_{i,2} = (\tau - X_i)^l K(\frac{x - X_i}{h_n}) \left( m(X_i) - m(x) - (X_i - x)m'(x) \right).$$

D'après l'inégalité de Bienaymé-Tchebycheff nous avons

$$\sum_{i=1}^{n} Z_{i,j} = \mathbb{E}\left(\sum_{i=1}^{n} Z_{i,j}\right) + \mathcal{O}_P\left(\sqrt{\operatorname{Var}\left(\sum_{i=1}^{n} Z_{i,j}\right)}\right),$$
$$= n\mathbb{E}Z_{1,j} + \mathcal{O}_P\left(\sqrt{n\mathbb{E}Z_{1,j}^2}\right),$$

avec

$$\mathbb{E}Z_{1,1} = h_n^{(l+1)} \int u^l K(u) f_X(x - h_n u) du ,$$
  
=  $h_n^{(l+1)} f_X(x) (K_l + o(1)) .$ 

Cette dernière égalité résulte de la continuité de la fonction  $f_X(\cdot)$  au point x. Si  $f_X(\cdot)$  est continue sur le compact [0, 1], cette majoration devient uniforme par rapport à x.

$$\mathbb{E}Z_{1,1}^2 = h_n^{(2l+1)} \int u^{2l} K^2(u) f_X(x - h_n u) du ,$$
  
=  $\mathcal{O}\left(h_n^{(2l+1)}\right),$ 

d'où

$$S_{l}(x) = n \mathbb{E} Z_{1,1} + \mathcal{O}_{P} \left( \sqrt{n \mathbb{E} Z_{1,1}^{2}} \right)$$
  
=  $n h_{n}^{(l+1)} f_{X}(x) \left( K_{l} + o_{P}(1) \right) .$ 

Pour  $\mathbb{E}Z_{1,2}$  et  $\mathbb{E}Z_{1,2}^2$ , en développant  $m(x - h_n u)$  à l'ordre 2, nous obtenons :

$$\mathbb{E}Z_{1,2} = h_n^{(l+1)} \int u^l K_+(u) \left( m(x - h_n u) - m(x) + h_n u m'(x) \right) f_X(x - h_n u) du ,$$
  
$$= h_n^{(l+1)} \int u^l K_+(u) \left( \frac{1}{2} u^2 m^{(2)}(x) h_n^2 + o(h_n^2) \right) f_X(x - h_n u) du ,$$
  
$$= \frac{1}{2} m^{(2)}(x) h_n^{(l+3)} f_X(x) \left( K_{l+2} + o(1) \right) ,$$

$$\begin{split} \mathbb{E}Z_{1,2}^2 &= h_n^{(2l+1)} \int u^l K_+^2(u) \left( m(x - h_n u) - m(x) + h_n u m'(x) \right)^2 f_X(x - h_n u) du \,, \\ &= h_n^{(2l+1)} \int u^l K_+^2(u) \left( \frac{1}{2} u^2 m^{(2)}(x) h_n^2 + o(h_n^2) \right)^2 f_X(x - h_n u) du \,, \\ &= \mathcal{O}\left( h_n^{2l+5} \right). \end{split}$$

Nous en déduisons :

$$R_{l}(x) = n \mathbb{E} Z_{1,2} + \mathcal{O}_{P} \left( \sqrt{n \mathbb{E} Z_{1,2}^{2}} \right),$$
  
$$= \frac{1}{2} m^{(2)}(x) n h_{n}^{(l+3)} f_{X}(x) \left( K_{l+2} + o_{P}(1) \right).$$

Ceci achève la démonstration du lemme 1.1.

Corollaire 1.1. Sous les mêmes hypothèses que celles du théorème 1.1, nous avons

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

### Preuve du corollaire 1.1

Remarquons tout d'abord que  $\hat{m}_+(\tau)$  et  $\hat{m}_-(\tau)$  sont indépendants conditionnellement à  $X_1, \ldots, X_n$ . Puisque  $\hat{\gamma}(\tau) - \gamma(\tau) = (\hat{m}_+(\tau) - m_+(\tau)) - (\hat{m}_-(\tau) - m_-(\tau))$  on a alors:

1) 
$$B_n = (\mathbb{E}[\hat{\gamma}(\tau)|X_1, \dots, X_n] - \gamma(\tau))^2 = (B_n^+ - B_n^-)^2,$$

2) 
$$V_n = Var[\hat{\gamma}(\tau)|X_1, \dots, X_n] = V_n^+ + V_n^-$$

Utilisons les arguments de la fin de la démonstrations du théorème 1.1 nous en déduisons le corollaire 1.1.

### 1.3.3 Normalité Asymptotique

### Théorème 1.2.

a) Supposons que les hypothèses du théorème 1.1 sont satisfaites et qu'il existe  $\delta > 0$  tel que  $\mathbb{E}(|\varepsilon|^{2+\delta}) < \infty$ , alors:

$$\sqrt{nh_n} \left( \begin{array}{c} \hat{m}_{-}(\tau) - \mathbb{E}[\hat{m}_{-}(\tau) \mid X_1, \dots, X_n] \\ \hat{m}_{+}(\tau) - \mathbb{E}[\hat{m}_{+}(\tau) \mid X_1, \dots, X_n] \end{array} \right) \xrightarrow{\mathcal{L}} \mathcal{N} \left( 0, \left( \frac{\sigma^2(\tau)}{f_X(\tau)} V_+^2 \right) I_2 \right).$$

b) Si de plus nous avons  $nh_n^5 \to 0$ , alors

$$\sqrt{nh_n} \left( \begin{array}{c} \hat{m}_-(\tau) - m_-(\tau) \\ \hat{m}_+(\tau) - m_+(\tau) \end{array} \right) \xrightarrow{\mathcal{L}} \mathcal{N} \left( 0, \left( \frac{\sigma^2(\tau)}{f_X(\tau)} V_+^2 \right) I_2 \right).$$

où  $I_2$  est la matrice identité d'ordre 2.

### Preuve du théorème 1.2

Il suffit de montrer (théorème de Cramer-Wold) que pour tout a et b,  $a\hat{m}_{+}(\tau) + b\hat{m}_{-}(\tau)$  est asymptotiquement normal.

Tout d'abord notons que grâce au lemme 1.1, nous pouvons écrire

$$a\left(\hat{m}_{-}(\tau) - \mathbb{E}[\hat{m}_{-}(\tau) \mid X_{1}, ..., X_{n}]\right) + b\left(\hat{m}_{+}(\tau) - \mathbb{E}[\hat{m}_{+}(\tau) \mid X_{1}, ..., X_{n}]\right)$$
  
$$= \frac{\sum_{i=1}^{n} (aw_{i}^{-}(\tau) + bw_{i}^{+}(\tau))\sigma(X_{i})\varepsilon_{i}}{n^{2}h_{n}^{4}f_{X}^{2}(\tau)}(1 + o_{P}(1)).$$
(1.11)

En utilisant a) et d) du lemme 1.1 nous obtenons

$$\sum_{i=1}^{n} w_{i}^{+}(\tau) \sigma(X_{i}) \varepsilon_{i} = S_{2}^{+}(\tau) \xi_{0}^{+}(\tau) - S_{1}^{+}(\tau) \xi_{1}^{+}(\tau),$$
  
$$= n h_{n}^{3} f_{X}(\tau) K_{2}^{+} \xi_{0}^{+}(\tau) - n h_{n}^{2} f_{X}(\tau) K_{1}^{+} \xi_{1}^{+}(\tau) + o_{P} \left( n^{3/2} h_{n}^{7/2} \right),$$
  
$$= n h_{n}^{2} f_{X}(\tau) \sum_{i=1}^{n} \Omega_{i}^{+} + o_{P} \left( n^{3/2} h_{n}^{7/2} \right),$$

оù

$$\Omega_i^+ = \left(h_n K_2^+ - (\tau - X_i) K_1^+\right) K_+ \left(\frac{\tau - X_i}{h_n}\right) \sigma(X_i) \varepsilon_i$$

En traitant de la même manière  $\sum_{i=1}^{n} w_i^-(\tau) \sigma(X_i) \varepsilon_i$ , nous en déduisons

$$\begin{split} \sum_{i=1}^{n} \left( a w_i^-(\tau) + b w_i^+(\tau) \right) \right) \sigma(X_i) \varepsilon_i &= n h_n^2 f_X(\tau) \sum_{i=1}^{n} \left( a \Omega_i^- + b \Omega_i^+ \right) + o_P \left( n^{3/2} h_n^{7/2} \right), \\ &= n h_n^2 f_X(\tau) \sum_{i=1}^{n} \Omega_i + o_P \left( n^{3/2} h_n^{7/2} \right), \end{split}$$

où  $\Omega_i$  est défini par  $\Omega_i = a\Omega_i^- + b\Omega_i^+$ .

Remarquons que  $\mathbb{E}(\Omega_i^- \Omega_i^+) = 0$ , ce qui implique que  $\Omega_i^+$  et  $\Omega_i^-$  sont des variables aléatoires centrées non corrélées. Sachant que  $X_i$  et  $\varepsilon_i$  sont indépendants, avec des arguments similaires à ceux employés dans la démonstration du lemme 1.1, nous obtenons

$$\begin{aligned} Var(\Omega_i^+) &= \mathbb{E}\left[ \left( h_n K_2^+ - (\tau - X_i) K_1^+ \right)^2 K_+^2 \left( \frac{\tau - X_i}{h_n} \right) \sigma^2(X_i) \right], \\ &= h_n^3 \sigma^2(\tau) f_X(\tau) \left( \int_{-1}^0 (K_2^+ - u K_1^+)^2 K_+^2(u) \, du \right) (1 + o(1)), \\ &= h_n^3 \sigma^2(\tau) f_X(\tau) V_+^2 + o(h_n^3), \end{aligned}$$

et idem pour  $\Omega_i^-$  quitte à changer  $K_+(\cdot)$  en  $K_-(\cdot)$ .

D'où

$$Var(\Omega_i) = h_n^3 \sigma^2(\tau) f_X(\tau) (a^2 V_-^2 + b^2 V_+^2) + o(h_n^3).$$
(1.12)

Nous allons maintenant prouver que le tableau triangulaire  $(\Omega_i)$  satisfait les conditions de Lindeberg dont nous avons besoin pour appliquer le théorème central limite de Liapounov.

$$\sum_{i=1}^{n} \operatorname{I\!E}(|\Omega_{i}|^{2+\delta}) = n \operatorname{I\!E}(|Z_{1}|^{2+\delta}) \operatorname{I\!E}(|\varepsilon_{1}|^{2+\delta}), \qquad (1.13)$$

où  $Z_1 = aZ_1^- + bZ_1^+$ , avec

$$Z_1^+ = \left(h_n K_2^+ - (\tau - X_1) K_1^+\right) K_+ \left(\frac{\tau - X_1}{h_n}\right) \sigma(X_1),$$

 $\operatorname{et}$ 

$$Z_1^- = \left(h_n K_2^- - (\tau - X_1) K_1^-\right) K_-\left(\frac{\tau - X_1}{h_n}\right) \sigma(X_1)$$

Un calcul analogue conduit à

$$\begin{split} \mathbb{E}(|\Omega_{i}|^{2+\delta}) &\leq C\Big(a^{2+\delta}h_{n}^{3+\delta}\sigma^{2+\delta}(\tau)f_{X}(\tau)\int_{-1}^{0}|K_{2}^{-}-uK_{1}^{-}|^{2+\delta}K_{-}^{2+\delta}(u)\,du \\ &+ b^{2+\delta}h_{n}^{3+\delta}\sigma^{2+\delta}(\tau)f_{X}(\tau)\int_{-1}^{0}|K_{2}^{+}-uK_{1}^{+}|^{2+\delta}K_{+}^{2+\delta}(u)\,du\Big) \\ &+ o\left(h_{n}^{3+\delta}\right), \\ &= \mathcal{O}\left(h_{n}^{3+\delta}\right). \end{split}$$

Donc  $\sum_{i=1}^{n} \mathbb{E}(|\Omega_i|^{2+\delta}) = \mathcal{O}\left(nh_n^{3+\delta}\right).$ 

Revenons maintenant au calcul de la variance : (1.12) implique

$$Var(\sum_{i=1}^{n} \Omega_i) = nh_n^3 \sigma^2(\tau) f_X(\tau) \left(a^2 V_-^2 + b^2 V_+^2\right) + o(nh_n^3).$$
(1.14)

D'où il vient que  $\sum_{i=1}^{n} \mathbb{E}(|\Omega_i|^{2+\delta}) / (Var \sum_{i=1}^{n} \Omega_i)^{(2+\delta)/2} = \mathcal{O}((nh_n)^{-\delta/2}).$ 

Pour terminer la démonstration, nous écrivons le terme dominant de l'équation (1.11) comme suit :

$$\frac{\sum_{i=1}^{n} (aw_i^-(\tau) + bw_i^+(\tau))\sigma(X_i)\varepsilon_i}{n^2 h_n^4 f_X^2(\tau)} = \frac{(1 + o_P(1))\left(\sum_{i=1}^{n} \Omega_i\right)}{\sqrt{nh_n}\left(Var(\sum_{i=1}^{n} \Omega_i)\right)^{1/2}\left(f_X(\tau)/(a^2V_- + b^2V_+)\sigma^2(\tau)\right)^{1/2}}$$

Nous obtenons ainsi la partie **a**) du théorème comme conséquence du théorème central limite de Liapounov.

D'après le théorème 1.1 nous avons ( $\mathbb{E}[\hat{m}_+(\tau) \mid X_1, \ldots, X_n] - m_+(\tau)$ ) =  $\mathcal{O}_P(h_n^2)$ , nous en déduisons immédiatement la partie **b**) du théorème.

#### Corollaire 1.2.

a) Supposons que les hypothèses du théorème 1.1 sont satisfaites et qu'il existe  $\delta > 0$  tel que  $\mathbb{E}(|\varepsilon|^{2+\delta}) < \infty$ , alors nous avons

$$\sqrt{nh_n}\left(\hat{\gamma}(\tau) - \mathbb{E}\left[\hat{\gamma}(\tau) \mid X_1, \dots, X_n\right]\right) \xrightarrow{\mathcal{L}} \mathcal{N}\left(0, \frac{2\sigma^2(\tau)}{f_X(\tau)}V_+^2\right).$$

b) Si de plus nous avons  $nh_n^5 \to 0$ , alors

$$\sqrt{nh_n}\left(\hat{\gamma}(\tau) - \gamma(\tau)\right) \xrightarrow{\mathcal{L}} \mathcal{N}\left(0, \frac{2\sigma^2(\tau)}{f_X(\tau)}V_+^2\right).$$

#### Preuve du corollaire 1.2

Puisque  $\hat{\gamma}(\tau) = \hat{m}_{+}(\tau) - \hat{m}_{-}(\tau)$ , la partie **a**) est donc une conséquence immédiate de la partie **a**) du théorème 1.2. Sachant aussi que le biais de  $\hat{\gamma}(\tau)$  a le même comportement que celui de  $\hat{m}_{+}(\tau)$  ou celui de  $\hat{m}_{-}(\tau)$ , nous en déduisons donc la partie **b**).

Dans la pratique, les quantités  $f_X(\tau)$  et  $\sigma(\tau)$  sont en général inconnues. Les résultats du corollaire 1.2 ne sont donc pas directement exploitables. En revanche, on peut remplacer  $f_X(\tau)$  et  $\sigma^2(\tau)$  par des estimateurs consistants, d'où le corollaire suivant. Notons que, par exemple, on peut utiliser la méthode du noyau pour obtenir un estimateur consistant du premier paramètre, et l'estimateur proposé par Müller et Stadmüller [MS87], pour le deuxième paramètre.

**Corollaire 1.3.** Soient  $\hat{f}_X(\tau)$  et  $\hat{\sigma}^2(\tau)$  deux estimateurs consistants de  $f_X(\tau)$  et  $\sigma^2(\tau)$ . L'estimateur  $\hat{f}_X(\tau)$  est supposé positif. Sous les mêmes hypothèses que celles du corollaire 1.2, nous avons

$$\sqrt{nh_n} \frac{\sqrt{\hat{f}_X(\tau)} \left(\hat{\gamma}(\tau) - \gamma(\tau)\right)}{(2\hat{\sigma}^2(\tau))^{1/2}} \xrightarrow{\mathcal{L}} \mathcal{N}\left(0, V_+^2\right).$$
(1.15)

Comme conséquence immédiate de ce corollaire, un intervalle de confiance, approché, au seuil  $\alpha$  de  $\gamma(\tau)$  est donné par la formule suivante :

$$\hat{\gamma}(\tau) \pm \frac{q_{\alpha}}{\sqrt{nh_n}} \left(\frac{2\hat{\sigma}^2(\tau)}{\hat{f}_X(\tau)}V_+^2\right)^{1/2}.$$

où  $q_{\alpha}$  est le  $(1 - \alpha/2)$ -quantile de la loi normale standard.
Comme nous l'avons mentionné dans la remarque 1.1, il est possible de prendre en compte le fait que les fonctions  $f_X(\cdot)$  et/ou  $\sigma(\cdot)$  sont discontinues au point  $\tau$ . En général, ceci n'est pas une hypothèse réaliste pour la fonction  $f_X(\cdot)$ . En revanche, dans la pratique, il arrive fréquemment que le point  $\tau$  soit aussi un point de rupture pour la fonction  $\sigma(\cdot)$ . Les modifications à apporter pour traiter le cas où les fonctions  $f_X(\cdot)$  et  $\sigma(\cdot)$  sont discontinues au point  $\tau$ , peuvent être déduites facilement en lisant les démonstrations. Supposons que  $\tau$  est un point de rupture des fonctions  $f_X(\cdot)$  et  $\sigma(\cdot)$ , alors l'erreur quadratique moyenne conditionnelle de l'amplitude de la rupture  $\hat{\gamma}(\tau)$  est donnée par (comparer avec corollaire 1.1) :

$$\mathbb{E}\left[\left(\hat{\gamma}(\tau) - \gamma(\tau)\right)^2 \mid X_1, \dots, X_n\right] \\ = \left[\frac{\gamma^{(2)}(\tau)}{2}h_n^2 B_+\right]^2 + \frac{1}{nh_n}\left(\frac{\sigma_+^2(\tau)}{f_+(\tau)} + \frac{\sigma_-^2(\tau)}{f_-(\tau)}\right)V_+^2 + o_P(h_n^4 + \frac{1}{nh_n}).$$

Dans le théorème 1.2 la matrice de covariance asymptotique reste diagonale, mais avec  $V_+^2 \sigma_-^2(\tau)/f_-(\tau)$  et  $V_+^2 \sigma_+^2(\tau)/f_+(\tau)$  comme termes de la diagonale. Par conséquent la variance asymptotique dans le corollaire 1.2 devient  $\left(\frac{\sigma_+^2(\tau)}{f_+(\tau)} + \frac{\sigma_-^2(\tau)}{f_-(\tau)}\right)V_+^2$ .

# 1.4 La situation où la localisation de la discontinuité est inconnue

Pour tout  $t \in [0, 1]$ , posons

$$\hat{\gamma}(t) = \hat{m}_{+}(t) - \hat{m}_{-}(t).$$

Intuitivement, la localisation du maximum de ces différences peut être un bon estimateur du point de la discontinuité de la fonction de régression  $m(\cdot)$ . Ainsi nous proposons comme estimateur de  $\tau$ :

$$\hat{\tau} = \inf\{t \in \kappa; \hat{\gamma}(t) = \sup_{x \in \kappa} \hat{\gamma}(x)\},\tag{1.16}$$

où  $\kappa$  un compact inclus dans ]0, 1[.

Comme estimateur de l'amplitude de la discontinuité, il est naturel de poser :

$$\hat{\gamma}(\hat{\tau}) = \hat{m}_{+}(\hat{\tau}) - \hat{m}_{-}(\hat{\tau}). \tag{1.17}$$

La restriction à  $\kappa$  a pour but d'éliminer les extrémités de l'intervalle car elles sont considérées comme des points de rupture. Dans toute la suite, nous supposons que  $\tau \in ]0, 1[$ .

La figure 1.2 illustre l'action de la RLL séparément sur la partie signal et sur la partie bruit. On note sur l'exemple une amélioration du rapport signal sur bruit suite au lissage effectué. Dans ce travail nous désignons par rapport signal sur bruit le rapport entre l'amplitude de la rupture et l'écart-type du bruit :

$$RSB = \frac{|\gamma(\tau)|}{\sigma(\tau)}.$$

En pratique ni le signal ni le bruit ne sont accessibles et on a interêt à ce que l'amélioration dûe au lissage soit la plus forte possible. La partie droite de la figure 1.2 donne les représentations du processus du saut calculé sur le signal échantillonné aux points  $X_i$  (en haut à droite,  $\hat{\gamma}_m(\cdot)$ ) et sur le bruit aux points  $X_i$  (en bas à droite,  $\hat{\gamma}_{\varepsilon}(\cdot)$ ). Le processus de saut  $\hat{\gamma}(\cdot) = \hat{\gamma}_m(\cdot) + \hat{\gamma}_{\varepsilon}(\cdot)$ est reconstitué dans la partie centrale droite.

Pour étudier la distribution asymptotique de  $\hat{\tau}$  et  $\hat{\gamma}(\hat{\tau})$ , nous introduisons le processus de déviation locale suivant:

$$\begin{aligned} \mathcal{Z}_n(z) &= \alpha(n, h_n) \left( \hat{\gamma} \left( \tau + \frac{h_n}{\beta(n, h_n)} z \right) - \hat{\gamma}(\tau) \right), \quad \forall z \in [-M, M] \\ &= \alpha(n, h_n) \left( \hat{\gamma} \left( \tau + h_n y \right) - \hat{\gamma}(\tau) \right), \end{aligned}$$

où nous posons  $y = z/\beta(n, h_n)$ , M étant une constante positive quelconque.

L'idée du recours à ce processus de déviation locale semble être due à Eddy [EDD82] qui l'a utilisée pour estimer le mode d'une densité. Avec des choix convenables pour  $\alpha(n,h_n)$  et  $\beta(n,h_n)$ , nous montrons que  $\mathcal{Z}_n(z)$  converge vers un processus  $\mathcal{Z}(z)$ . Il s'ensuit que  $z_{max}(n) =$  $\arg \sup_{z} \mathcal{Z}_n(z)$  converge vers  $z_{max} = \arg \sup_{z} \mathcal{Z}(z)$ . De ces convergences découlent des informations sur la vitesse de convergence de  $\hat{\tau} = \tau + \frac{h_n}{\beta(n,h_n)} z_{max}(n)$  que nous explicitons plus loin.

# 1.4.1 Approximation par la somme des éléments d'une ligne d'un tableau triangulaire

Nous avons

$$\mathcal{Z}_n(z) = \mathcal{Z}_n^+(z) - \mathcal{Z}_n^-(z),$$



Figure 1.2 : Localisation d'une rupture dans un modèle de régression additif par R.L.L : Démonstration

où on peut écrire

$$\begin{aligned} \mathcal{Z}_{n}^{+}(z) &= \alpha(n,h_{n}) \left[ \frac{\sum_{i=1}^{n} w_{i}^{+}(\tau+h_{n}y)y_{i}}{\sum_{i=1}^{n} w_{i}^{+}(\tau+h_{n}y)} - \frac{\sum_{i=1}^{n} w_{i}^{+}(\tau)y_{i}}{\sum_{i=1}^{n} w_{i}^{+}(\tau)} \right], \\ &= \alpha(n,h_{n}) \frac{\sum_{i=1}^{n} (w_{i}^{+}(\tau+h_{n}y) - w_{i}^{+}(\tau))(y_{i} - m_{+}(\tau))}{\sum_{i=1}^{n} w_{i}^{+}(\tau+h_{n}y)} \\ &- \alpha(n,h_{n}) \left( \frac{\sum_{i=1}^{n} w_{i}^{+}(\tau)(y_{i} - m_{+}(\tau))}{\sum_{i=1}^{n} w_{i}^{+}(\tau)} \right) \left( \frac{\sum_{i=1}^{n} w_{i}^{+}(\tau+h_{n}y) - w_{i}^{+}(\tau)}{\sum_{i=1}^{n} w_{i}^{+}(\tau+h_{n}y)} \right), \\ &= A_{n} - B_{n} \times C_{n} \end{aligned}$$
(1.18)

Or nous avons

$$\begin{split} w_i^+(\tau + h_n y) &- w_i^+(\tau) \\ &= \left[ S_2^+(\tau + h_n y) - S_2^+(\tau) \right] K_+(\frac{\tau - X_i}{h_n}) - \left[ S_1^+(\tau + h_n y) - S_1^+(\tau) \right] (\tau - X_i) K_+(\frac{\tau - X_i}{h_n}) \\ &+ S_2^+(\tau + h_n y) \left[ K_+(\frac{\tau + h_n y - X_i}{h_n}) - K_+(\frac{\tau - X_i}{h_n}) \right] \\ &- S_1^+(\tau + h_n y) \left[ (\tau + h_n y - X_i) K_+(\frac{\tau + h_n y - X_i}{h_n}) - (\tau - X_i) K_+(\frac{\tau - X_i}{h_n}) \right]. \end{split}$$

Il vient donc, pour les numérateurs de  $A_n$  et  $C_n$ , les écritures suivantes qui nous seront utiles par la suite :

$$\sum_{i=1}^{n} w_{i}^{+}(\tau + h_{n}y) - w_{i}^{+}(\tau)$$

$$= \left[S_{2}^{+}(\tau + h_{n}y) - S_{2}^{+}(\tau)\right]S_{0}^{+}(\tau) - \left[S_{1}^{+}(\tau + h_{n}y) - S_{1}^{+}(\tau)\right]S_{1}^{+}(\tau) \qquad (1.19)$$

$$+ \left[S_{0}^{+}(\tau + h_{n}y) - S_{0}^{+}(\tau)\right]S_{2}^{+}(\tau + h_{n}y) - \left[S_{1}^{+}(\tau + h_{n}y) - S_{1}^{+}(\tau)\right]S_{1}^{+}(\tau + h_{n}y),$$

 $\operatorname{et}$ 

$$\sum_{i=1}^{n} (w_{i}^{+}(\tau + h_{n}y) - w_{i}^{+}(\tau))(y_{i} - m_{+}(\tau))$$

$$= \left[S_{2}^{+}(\tau + h_{n}y) - S_{2}^{+}(\tau)\right] D_{0}^{+}(\tau) - \left[S_{1}^{+}(\tau + h_{n}y) - S_{1}^{+}(\tau)\right] D_{1}^{+}(\tau) \qquad (1.20)$$

$$+ \left[D_{0}^{+}(\tau + h_{n}y) - D_{0}^{+}(\tau)\right] S_{2}^{+}(\tau + h_{n}y) - \left[D_{1}^{+}(\tau + h_{n}y) - D_{1}^{+}(\tau)\right] S_{1}^{+}(\tau + h_{n}y),$$

оù

$$D_l^+(x) = \sum_{i=1}^n (x - X_i)^l K_+(\frac{x - X_i}{h_n})(y_i - m_+(\tau)).$$

Les résultats qui vont suivre sont énoncés sous certaines des hypothèses suivantes :

 $\begin{array}{lll} \text{H1.} & \lim_{n \to \infty} h_n = 0 & \text{et} & \lim_{n \to \infty} nh_n = \infty. \\ \\ \text{H2.} & \lim_{n \to \infty} \frac{h_n}{\beta(n,h_n)} = 0 & \text{et} & \lim_{n \to \infty} \alpha(n,h_n) = \infty. \\ \\ \\ \text{H3-0.} & K_+(0) = K_+(-1) = 0 & \text{et} & K_+^{(1)}(0) < 0. \\ \\ \\ \\ \text{H4-0.} & K_+(0) > 0 & \text{et} & K_+(-1) = 0. \end{array}$ 

#### Lemme 1.2.

a) Sous les hypothèses H1 et H2,

$$S_{l}^{+}(\tau + h_{n}y) = nh_{n}^{l+1}f_{X}(\tau)(K_{l}^{+} + o_{p}(1)).$$

b) Sous les hypothèses H1, H2 et H3-0,

$$S_l^+(\tau + h_n y) - S_l^+(\tau) = \mathcal{O}\left(\frac{nh_n^{l+1}}{\beta^2(n,h_n)}\right) + \mathcal{O}_P\left(\frac{\sqrt{nh_n^{2l+1}}}{\beta(n,h_n)}\right).$$

c) Sous les hypothèses H1, H2 et H4-0,

$$S_l^+(\tau + h_n y) - S_l^+(\tau) = \mathcal{O}\left(\frac{nh_n^{l+1}}{\beta(n,h_n)}\right) + \mathcal{O}_P\left(\sqrt{\frac{nh_n^{2l+1}}{\beta(n,h_n)}}\right).$$

d) Sous l'hypothèse H1,

$$D_l^+(\tau) = \mathcal{O}\left(nh_n^{l+2}\right) + \mathcal{O}_P\left(\sqrt{nh_n^{2l+1}}\right)$$

#### Preuve du lemme 1.2

- a) Voir la démonstration du lemme 1.1.
- b) Nous avons

$$\mathbb{E}\left[S_{l}^{+}(\tau+h_{n}y)-S_{l}^{+}(\tau)\right] = nh_{n}^{l+1}\int\left[(x+y)^{l}K_{+}(x+y)-x^{l}K_{+}(x)\right]f_{X}(\tau-h_{n}x)dx , \\ = nh_{n}^{l+1}\int_{-1-y}^{-1}(x+y)^{l}K_{+}(x+y)f_{X}(\tau-h_{n}x)dx \\ + nh_{n}^{l+1}\int_{-1}^{-y}\left[(x+y)^{l}K_{+}(x+y)-x^{l}K_{+}(x)\right]f_{X}(\tau-h_{n}x)dx \\ - nh_{n}^{l+1}\int_{-y}^{0}x^{l}K_{+}(x)f_{X}(\tau-h_{n}x)dx .$$

Or sous H1, H2 et H3-0, nous avons

$$\begin{split} &\int_{-1-y}^{-1} (x+y)^l K_+(x+y) f_X(\tau - h_n x) dx \\ &= \int_0^y (x-1-y)^l K_+(x-1-y) f_X(\tau - h_n(x-1-y)) dx , \\ &= \int_0^y \left[ (-1)^l K_+(-1) + \mathcal{O}(y) \right] \left[ f_X(\tau) + o(1) \right] dx , \\ &= \mathcal{O} \left( y^2 \right) , \\ &= \mathcal{O} \left( 1/\beta^2(n,h_n) \right) . \end{split}$$

Un calcul similaire conduit à

$$\int_{-y}^{0} x^{l} K_{+}(x) f_{X}(\tau - h_{n}x) dx = \mathcal{O}\left(1/\beta^{2}(n,h_{n})\right).$$

Un développement de Taylor à l'ordre 2 au voisinage de x implique que

$$\int_{-1}^{-y} \left[ (x+y)^{l} K_{+}(x+y) - x^{l} K_{+}(x) \right] f_{X}(\tau - h_{n}x) dx$$
  
= 
$$\int_{-1}^{-y} \left[ y \left( x^{l} K_{+}(x) \right)' + \mathcal{O}(y^{2}) \right] \left[ f_{X}(\tau) + o(1) \right] dx ,$$
  
= 
$$\mathcal{O} \left( y^{2} \right) ,$$
  
= 
$$\mathcal{O} \left( 1/\beta^{2}(n,h_{n}) \right) .$$

Il en résulte que sous H1, H2 et H3-0

$$\mathbb{E}\left[S_l^+(\tau+h_n y) - S_l^+(\tau)\right] = \mathcal{O}\left(\frac{nh_n^{l+1}}{\beta^2(n,h_n)}\right).$$

D'autre part en utilisant les mêmes arguments que ceux utilisés ci-dessus, nous obtenons

$$Var\left[S_l^+(\tau+h_n y) - S_l^+(\tau)\right] = \mathcal{O}\left(\frac{nh_n^{2l+1}}{\beta^2(n,h_n)}\right).$$

D'où sous H1, H2 et H3-0,

$$S_l^+(\tau + h_n y) - S_l^+(\tau) = \mathcal{O}\left(\frac{nh_n^{l+1}}{\beta^2(n,h_n)}\right) + \mathcal{O}_P\left(\frac{\sqrt{nh_n^{2l+1}}}{\beta(n,h_n)}\right).$$

c) Sous H1, H2 et H4-0, nous avons:

1) 
$$\int_{-1-y}^{-1} (x+y)^{l} K_{+}(x+y) f_{X}(\tau - h_{n}x) dx = \mathcal{O}\left(1/\beta^{2}(n,h_{n})\right),$$
  
2) 
$$\int_{-y}^{0} x^{l} K_{+}(x) f_{X}(\tau - h_{n}x) dx = \mathcal{O}\left(1/\beta(n,h_{n})\right),$$

3) 
$$\int_{-1}^{-y} \left[ (x+y)^l K_+(x+y) - x^l K_+(x) \right] f_X(\tau - h_n x) dx = \mathcal{O}\left( 1/\beta(n,h_n) \right).$$

Nous en déduisons que sous H1, H2 et H4-0

$$\mathbb{E}\left[S_l^+(\tau+h_n y) - S_l^+(\tau)\right] = \mathcal{O}\left(\frac{nh_n^{l+1}}{\beta(n,h_n)}\right).$$

De la même manière, nous obtenons que sous H1, H2 et H4-O

$$Var\left[S_l^+(\tau+h_n y) - S_l^+(\tau)\right] = \mathcal{O}\left(\frac{nh_n^{2l+1}}{\beta(n,h_n)}\right).$$

Donc sous H1, H2 et H4-0,

$$S_l^+(\tau + h_n y) - S_l^+(\tau) = \mathcal{O}\left(\frac{nh_n^{l+2}}{\beta(n,h_n)}\right) + \mathcal{O}_P\left(\sqrt{\frac{nh_n^{2l+1}}{\beta(n,h_n)}}\right).$$

d) Puisque nous avons,

$$\mathbb{E} \left[ D_l^+(\tau) \right] = n h_n^{l+1} \int x^l K_+(x) \left( m(\tau - h_n x) - m_+(\tau) \right) f_X(\tau - h_n x) dx ,$$
  
=  $\mathcal{O} \left( n h_n^{l+2} \right) ,$ 

 $\operatorname{et}$ 

$$Var\left[D_{l}^{+}(\tau)\right] \leq nh_{n}^{2l+1} \int x^{2l} K_{+}^{2}(x) \left[\left(m(\tau - h_{n}x) - m_{+}(\tau)\right)^{2} + \sigma^{2}(\tau - xh_{n})\right] f_{X}(\tau - h_{n}x) dx,$$
  
$$= \mathcal{O}\left(nh_{n}^{2l+1}\right),$$

alors

$$D_l^+(\tau) = \mathcal{O}\left(nh_n^{l+2}\right) + \mathcal{O}_P\left(\sqrt{nh_n^{2l+1}}\right).$$

Le lemme suivant précise le comportement des termes intervenant dans  $A_n$ ,  $B_n$ ,  $C_n$  définis en (1.18). Le noyau  $M_+(\cdot)$  qui intervient dans la définition du processus  $\mathcal{M}_n(z)$  est défini par :

$$M_{+}(x) = (K_{2}^{+} - xK_{1}^{+})K_{+}(x).$$

 $M_{+}(\cdot)$  s'appelle le noyau équivalent. Il a été introduit par Lejeune [LEJ85]. On définit de la même façon le noyau équivalent pour  $K_{-}(\cdot)$ :

$$M_{-}(x) = (K_{2}^{-} - xK_{1}^{-})K_{-}(x) = M_{+}(-x),$$

et on pose

$$M(x) = M_{+}(x) - M_{-}(x).$$



**Figure 1.3 :** Exemples de noyaux équivalents utilisés dans nos simulations. On montre  $K_+(\cdot)$ (à gauche) et  $M_+(\cdot)$  (à droite). Il s'agit dans le premier exemple (en haut à gauche) du noyau de Epanechnikov translaté ( $K_+(x) = -12\sqrt{5}x(1+x) \ I\!I_{[-1,0]}(x)$ ) et dans le deuxième exemple du noyau d'Epanechnikov ( $K_+(x) = 12\sqrt{5/19}(1-x^2) \ I\!I_{[-1,0]}(x)$ ).

**Table 1.1 :** Exemples de noyaux équivalents utilisés dans nos simulations. Les noyaux  $K_+(\cdot)$  sont normalisés de manière à vérifier la contrainte (1.3).

$K_{\pm}(\cdot)$	$M_{\pm}(\cdot)$	$B_+$	$V_{+}^{2}$
$-12\sqrt{5}x(x+1) \prod_{[-1,0]}(x)$	$-12x(3+5x)(1+x) \prod_{[-1,0]}(x)$	$-\frac{1}{5}$	$\frac{192}{35}$
$12\sqrt{5/19}(1-x^2) \prod_{[-1,0]}(x)$	$\frac{4}{19}(24+45x)(1-x^2) 1\!\!\!\mathrm{I}_{[-1,0]}(x)$	$-\frac{11}{95}$	$\frac{56832}{12635}$

Lemme 1.3.

a) Sous les hypothèses H1 et H2,  $\sum_{i=1}^{n} w_{i}^{+}(\tau + h_{n}y) = n^{2}h_{n}^{4}f_{X}^{2}(\tau)(1 + o_{p}(1)).$  b) Sous les hypothèses H1, H2 et H3-0:

1) 
$$\frac{\sum_{i=1}^{n} w_{i}^{+}(\tau+h_{n}y) - w_{i}^{+}(\tau)}{\sum_{i=1}^{n} w_{i}^{+}(\tau+h_{n}y)} = \mathcal{O}_{P}\left(\frac{1}{\beta^{2}(n,h_{n})} + \frac{1}{\sqrt{nh_{n}\beta}(n,h_{n})}\right),$$
  
2)  $\alpha(n,h_{n})\frac{\sum_{i=1}^{n} (w_{i}^{+}(\tau+h_{n}y) - w_{i}^{+}(\tau))(y_{i} - m_{+}(\tau))}{\sum_{i=1}^{n} w_{i}^{+}(\tau+h_{n}y)} = \frac{1 + o_{P}(1)}{f_{X}(\tau)}\mathcal{M}_{n}^{+}(z)$   
 $+ \mathcal{O}_{P}\left(\left[\frac{\alpha(n,h_{n})}{\beta^{2}(n,h_{n})} + \frac{\alpha(n,h_{n})}{\sqrt{nh_{n}\beta}(n,h_{n})}\right]\left[h_{n} + \frac{1}{\sqrt{nh_{n}}}\right]\right).$ 

c) Sous les hypothèses H1, H2 et H4-0:

1) 
$$\frac{\sum_{i=1}^{n} w_{i}^{+}(\tau+h_{n}y) - w_{i}^{+}(\tau)}{\sum_{i=1}^{n} w_{i}^{+}(\tau+h_{n}y)} = \mathcal{O}_{P}\left(\frac{1}{\beta(n,h_{n})} + \frac{1}{\sqrt{nh_{n}\beta(n,h_{n})}}\right),$$
  
2)  $\alpha(n,h_{n})\frac{\sum_{i=1}^{n}(w_{i}^{+}(\tau+h_{n}y) - w_{i}^{+}(\tau))(y_{i} - m_{+}(\tau))}{\sum_{i=1}^{n} w_{i}^{+}(\tau+h_{n}y)} = \frac{1 + o_{P}(1)}{f_{X}(\tau)}\mathcal{M}_{n}^{+}(z)$   
 $+ \mathcal{O}_{P}\left(\left[\frac{\alpha(n,h_{n})}{\beta(n,h_{n})} + \frac{\alpha(n,h_{n})}{\sqrt{nh_{n}\beta(n,h_{n})}}\right]\left[h_{n} + \frac{1}{\sqrt{nh_{n}}}\right]\right).$ 

d) Sous l'hypothèse H1,  $\frac{\sum_{i=1}^{n} w_i^+(\tau)(y_i - m_+(\tau))}{\sum_{i=1}^{n} w_i^+(\tau)} = \mathcal{O}_P\left(h_n^2 + \frac{1}{\sqrt{nh_n}}\right).$ 

$$\mathcal{M}_n^+(z) \text{ est } d\acute{e}fini \text{ par :} \\ \mathcal{M}_n^+(z) = \frac{\alpha(n,h_n)}{nh_n} \sum_{i=1}^n \left( M_+(\frac{\tau+h_ny-X_i}{h_n}) - M_+(\frac{\tau-X_i}{h_n}) \right) \left( Y_i - m_+(\tau) \right)$$

#### Preuve du lemme 1.3

a) Puisque  $\sum_{i=1}^{n} w_i^+(\tau + h_n y) = S_2^+(\tau + h_n y)S_0^+(\tau + h_n y) - (S_1^+(\tau + h_n y))^2$ , la partie a) du lemme 1.3 découle directement de la partie a) du lemme 1.2.

b) La partie b) du lemme 1.3 est une conséquence immédiate de (1.19), (1.20) et des parties a),
b) et d) du lemme 1.2.

En effet sous les hypothèses H1, H2 et H3-0, nous avons

$$\begin{split} &\sum_{i=1}^{n} w_{i}^{+}(\tau + h_{n}y) - w_{i}^{+}(\tau) \\ &= \left[S_{2}^{+}(\tau + h_{n}y) - S_{2}^{+}(\tau)\right] S_{0}^{+}(\tau) - \left[S_{1}^{+}(\tau + h_{n}y) - S_{1}^{+}(\tau)\right] S_{1}^{+}(\tau) \\ &+ \left[S_{0}^{+}(\tau + h_{n}y) - S_{0}^{+}(\tau)\right] S_{2}^{+}(\tau + h_{n}y) - \left[S_{1}^{+}(\tau + h_{n}y) - S_{1}^{+}(\tau)\right] S_{1}^{+}(\tau + h_{n}y), \\ &= \mathcal{O}_{P}\left(\left[\frac{nh_{n}^{3}}{\beta^{2}(n,h_{n})} + \frac{\sqrt{nh_{n}^{5}}}{\beta(n,h_{n})}\right] nh_{n}\right), \end{split}$$

 $\operatorname{et}$ 

$$\sum_{i=1}^{n} w_i^+(\tau + h_n y) = n^2 h_n^4 f_X^2(\tau) \left(1 + o_p(1)\right).$$

Il en résulte que sous H1, H2 et H3-0,

$$\frac{\sum_{i=1}^{n} w_i^+(\tau + h_n y) - w_i^+(\tau)}{\sum_{i=1}^{n} w_i^+(\tau + h_n y)} = \mathcal{O}_P\left(\frac{1}{\beta^2(n, h_n)} + \frac{1}{\sqrt{nh_n}\beta(n, h_n)}\right)$$

D'autre part

$$\begin{split} &\sum_{i=1}^{n} (w_{i}^{+}(\tau+h_{n}y)-w_{i}^{+}(\tau))(y_{i}-m_{+}(\tau)) \\ &= \left[S_{2}^{+}(\tau+h_{n}y)-S_{2}^{+}(\tau)\right] D_{0}^{+}(\tau) - \left[S_{1}^{+}(\tau+h_{n}y)-S_{1}^{+}(\tau)\right] D_{1}^{+}(\tau) \\ &+ \left[D_{0}^{+}(\tau+h_{n}y)-D_{0}^{+}(\tau)\right] S_{2}^{+}(\tau+h_{n}y) - \left[D_{1}^{+}(\tau+h_{n}y)-D_{1}^{+}(\tau)\right] S_{1}^{+}(\tau+h_{n}y), \\ &= \mathcal{O}_{P}\left(\frac{nh_{n}^{3}}{\beta^{2}(n,h_{n})} + \frac{\sqrt{nh_{n}^{5}}}{\beta(n,h_{n})}\right) \times \left[\mathcal{O}\left(nh_{n}^{2}\right) + \mathcal{O}_{P}\left(\sqrt{nh_{n}}\right)\right] \\ &+ \left[D_{0}^{+}(\tau+h_{n}y) - D_{0}^{+}(\tau)\right] nh_{n}^{3}f_{X}(\tau)(K_{2}^{+}+o_{p}(1)) \\ &- \left[D_{1}^{+}(\tau+h_{n}y) - D_{1}^{+}(\tau)\right] nh_{n}^{2}f_{X}(\tau)(K_{1}^{+}+o_{p}(1)), \\ &= nh_{n}^{3}f_{X}(\tau)\left(1+o_{p}(1)\right) \mathcal{M}_{n}^{+}(z) + \mathcal{O}_{P}\left(\frac{n^{2}h_{n}^{5}}{\beta^{2}(n,h_{n})}\right) \\ &+ \mathcal{O}_{P}\left(\frac{n^{3/2}h_{n}^{9/2}}{\beta(n,h_{n})}\right) + \mathcal{O}_{P}\left(\frac{n^{2}h_{n}^{4}}{\sqrt{nh_{n}\beta^{2}(n,h_{n})}}\right) + \mathcal{O}_{P}\left(\frac{nh_{n}^{3}}{\beta(n,h_{n})}\right). \end{split}$$

D'où sous les hypothèses H1, H2 et H3-0,

$$\begin{aligned} &\alpha(n,h_n) \frac{\sum_{i=1}^n (w_i^+(\tau+h_n y) - w_i^+(\tau))(y_i - m_+(\tau)))}{\sum_{i=1}^n w_i^+(\tau+h_n y)} \\ &= \frac{1 + o_P(1)}{f_X(\tau)} \mathcal{M}_n^+(z) + \mathcal{O}_P\left(\left[\frac{\alpha(n,h_n)}{\beta^2(n,h_n)} + \frac{\alpha(n,h_n)}{\sqrt{nh_n}\beta(n,h_n)}\right] \left[h_n + \frac{1}{\sqrt{nh_n}}\right]\right). \end{aligned}$$

c) Un calcul similaire conduit à la partie c).

d) La partie d) provient du théorème 1.1.

Ceci termine la démonstration du lemme.

Afin d'établir le résultat énoncé dans le lemme 1.4, nous aurons recours à l'un ou l'autre des groupes d'hypothèses suivants :

1) H3-0. 
$$K_{+}(0) = K_{+}(-1) = 0$$
 et  $K_{+}^{(1)}(0) < 0$ .  
H3-1.  $\lim_{n \to \infty} \frac{\alpha(n,h_n)h_n}{\beta(n,h_n)} = L_1 < \infty$ .  
H3-2.  $0 < \lim_{n \to \infty} \frac{\alpha(n,h_n)}{\beta^2(n,h_n)} = L_2 < \infty$ .  
H3-3.  $\lim_{n \to \infty} \frac{\alpha^2(n,h_n)}{nh_n\beta^2(n,h_n)} = L_3 < \infty$ .

2) H4-0.  $K_+(0) > 0$  et  $K_+(-1) = 0$ .

H4-1. 
$$0 < \lim_{n \to \infty} \frac{\alpha(n,h_n)}{\beta(n,h_n)} = L_4 < \infty.$$
  
H4-2.  $\lim_{n \to \infty} \frac{\alpha(n,h_n)}{nh_n} = L_5 < \infty.$ 

Dorénavant nous utiliserons H3 pour désigner les hypothèses H3-0, H3-1, H3-2 et H3-3 et H4 pour les hypothèses H4-0, H4-1 et H4-2.

Si H4 est satisfaite, nous posons également

$$\lim_{n \to \infty} \frac{\alpha^2(n, h_n)}{n h_n \beta(n, h_n)} = L_4 \times L_5 = L_6 < \infty.$$

Lemme 1.4. Sous les hypothèses {H1, H2 et H3} ou {H1, H2 et H4},

$$\begin{aligned} \mathcal{Z}_{n}(z) &= \frac{(1+o_{P}(1))}{f_{X}(\tau)} \left( \mathcal{M}_{n}^{+}(z) - \mathcal{M}_{n}^{-}(z) \right) + o_{P}(1) , \\ &= \frac{(1+o_{P}(1))}{f_{X}(\tau)} \mathcal{M}_{n}(z) + o_{P}(1) , \end{aligned}$$

où on a posé  $\mathcal{M}_n(z) = \mathcal{M}_n^+(z) - \mathcal{M}_n^-(z)$ .

Dans la suite nous montrons la convergence de  $\mathcal{M}_n(z)$  vers un processus qui sera soit un processus gaussien dégénéré, soit un processus de Poisson composé bilatéral. Pour cela nous montrons tout d'abord la convergence cylindrique en dimension 2:  $a\mathcal{M}_n(z_1) + b\mathcal{M}_n(z_2)$  converge vers une loi normale ou une loi de Poisson composée, et les paramètres de ces lois permettront d'identifier les paramètres des processus limites. Finalement la démonstration de la convergence est terminée en montrant que la suite des processus ( $\mathcal{M}_n(z)$ ) est tendue.

Pour démontrer la convergence en loi de  $a\mathcal{M}_n(z_1) + b\mathcal{M}_n(z_2)$ , nous observons qu'il s'agit de la somme des termes d'une ligne d'un tableau triangulaire. En effet

$$\mathcal{M}_n(z) = \mathcal{M}_n^+(z) - \mathcal{M}_n^-(z),$$

оù

$$\mathcal{M}_{n}^{\pm}(z) = \frac{\alpha(n,h_{n})}{nh_{n}} \sum_{i=1}^{n} \left( M_{\pm}(\frac{\tau+h_{n}y-X_{i}}{h_{n}}) - M_{\pm}(\frac{\tau-X_{i}}{h_{n}}) \right) (Y_{i} - m_{\pm}(\tau)),$$
  
$$= \sum_{i=1}^{n} M_{i}^{\pm}(z).$$

Il en résulte que

$$\mathcal{M}_{n}(z) = \sum_{i=1}^{n} \left( M_{i}^{+}(z) - M_{i}^{-}(z) \right),$$
$$= \sum_{i=1}^{n} M_{i}(z),$$

 $\operatorname{et}$ 

$$a\mathcal{M}_n(z_1) + b\mathcal{M}_n(z_2) = \sum_{i=1}^n (aM_i(z_1) + bM_i(z_2))$$

où nous avons noté  $M_i(z) = M_i^+(z) - M_i^-(z)$ . Les variables  $M_i(z)$ ,  $M_i^+(z)$  et  $M_i^-(z)$  dépendent en réalité de n, mais pour alléger les écritures nous omettons l'indice n. Posons également

$$T_i^n = aM_i(z_1) + bM_i(z_2)$$

Nous montrons la convergence de  $\sum_{i=1}^{n} T_i^n$ , en recourant au résultat classique de convergence de Gnedenko [GK68] vers des lois indéfiniment divisibles que nous rappelons ci-dessous :

#### Théorème 1.3 (Gnedenko [GK68, page 100])

Soit,  $_{n}(u) = \sum_{i=1}^{n} \int_{-\infty}^{u} x^{2} d\tilde{F}_{ni}(x)$  où  $\tilde{F}_{ni}(\cdot)$  est la fonction de répartition de  $\tilde{T}_{i}^{n} = T_{i}^{n} - \mathbb{E}(T_{i}^{n})$ . Supposons que le tableau  $(T_{i}^{n})$  satisfait la propriété :

$$\forall \, \epsilon > 0 \,, \quad \sup_{1 \le i \le n} \mathbb{P}\left( |\tilde{T}_i^n| > \epsilon \right) \to 0 \ (n \to \infty) \,.$$

Alors ,  $_n(u)$  tend vers une fonction , (u) croissante, à variation bornée, si et seulement si  $Var\left(\sum_{i=1}^n T_i^n\right)$  converge et  $\sum_{i=1}^n T_i^n$  converge en loi. La loi limite est alors indéfiniment divisible, de fonction caractéristique associée  $\phi(\cdot)$  vérifiant:

$$\log \phi(t) = it\gamma + \int \left(\exp(itu) - 1 - itu\right) \frac{1}{u^2}, \ (du) \ ,$$

où  $\gamma$  est la moyenne, (et  $\gamma = \mathbb{E}\left(\sum_{i=1}^{n} T_{i}^{n}\right)$ ).

La convergence de ,  $_n(\cdot)$  sera étudiée séparément dans les deux cas  $K_+(0) = 0$  et  $K_+(0) > 0$ . Pour étudier cette convergence, nous introduisons les notations:

1)  $g_{\pm}(x,z) = M_{\pm}(x+y) - M_{\pm}(x) = M_{\pm}(x + \frac{z}{\beta(n,h_n)}) - M_{\pm}(x),$ 

2) 
$$g_1^{\pm}(x,z) = g_{\pm}(x,z) \left( m(\tau - xh_n) - m_{\pm}(\tau) \right),$$

3) 
$$g_1(x,z) = g_1^+(x,z) - g_1^-(x,z),$$

4)  $g_{2}^{\pm}(x,z) = g_{\pm}(x,z) \left(m(\tau - xh_{n}) - m_{\pm}(\tau) + \sigma(\tau - xh_{n})\varepsilon\right),$ 5)  $g_{2}(x,z) = g_{2}^{+}(x,z) - g_{2}^{-}(x,z),$ 6)  $A_{n} = \sum_{i=1}^{n} \mathbb{E}\left[T_{i}^{n}\right],$ 7)  $, n(u) = \sum_{i=1}^{n} \mathbb{E}\left[(T_{i}^{n} - \mathbb{E}(T_{i}^{n}))^{2} \prod_{\{T_{i}^{n} - \mathbb{E}(T_{i}^{n}) \leq u\}}\right],$ 

où  $1_A(\cdot)$  est la fonction indicatrice de l'ensemble A.

Suite aux définitions ci-dessus, nous avons

$$A_n = \alpha(n, h_n) \int (ag_1(x, z_1) + bg_1(x, z_2)) f_X(\tau - xh_n) dx ,$$

 $\operatorname{et}$ 

$$, {}_{n}(u) = \frac{\alpha^{2}(n,h_{n})}{nh_{n}} \int \mathbb{E}\left[\left(ag_{2}(x,z_{1})+bg_{2}(x,z_{2})-\frac{h_{n}}{\alpha(n,h_{n})}A_{n}\right)^{2} \times 1\!\!1_{\left\{\frac{\alpha(n,h_{n})}{nh_{n}}(ag_{2}(x,z_{1})+bg_{2}(x,z_{2}))-\frac{1}{n}A_{n}\leq u\right\}}\right] f_{X}(\tau-xh_{n})dx .$$

Nous commençons par étudier dans le paragraphe suivant la situation où le noyau s'annule en zéro  $(K_+(0) = 0)$ .

#### 1.4.2 La situation où le noyau s'annule en zéro

Sous les hypothèses H1, H2 et H3, par un développement de Taylor, nous obtenons:

a) Si x < -y < 0 ou x < 0 < -y,

1) 
$$g_1(x,z)f_X(\tau - xh_n) = -\frac{h_n z}{\beta(n,h_n)} x M_+^{(1)}(x) m_+^{(1)}(\tau) f_X(\tau) + o(\frac{h_n}{\beta(n,h_n)}),$$

2) 
$$g_2(x,z) - \frac{h_n}{\alpha(n,h_n)} A_n = \frac{z}{\beta(n,h_n)} \sigma(\tau) \varepsilon M_+^{(1)}(x) + o_P(\frac{1}{\beta(n,h_n)}).$$

b) Si -y < 0 < x ou 0 < -y < x,

1) 
$$g_1(x,z)f_X(\tau - xh_n) = \frac{h_n z}{\beta(n,h_n)} x M_{-}^{(1)}(x) m_{-}^{(1)}(\tau) f_X(\tau) + o(\frac{h_n}{\beta(n,h_n)}),$$

2) 
$$g_2(x,z) - \frac{h_n}{\alpha(n,h_n)} A_n = -\frac{z}{\beta(n,h_n)} \sigma(\tau) \varepsilon M_-^{(1)}(x) + o_P(\frac{1}{\beta(n,h_n)}).$$

c) Si -y < x < 0 ou 0 < x < -y,

1) 
$$g_1(x,z)f_X(\tau - xh_n) = -\gamma(\tau)f_X(\tau)(x+y)M_{-}^{(1)}(0) + o(\frac{1}{\beta(n,h_n)}),$$

2) 
$$g_2(x,z) - \frac{h_n}{\alpha(n,h_n)} A_n = \mathcal{O}_P(\frac{1}{\beta(n,h_n)}).$$

d'où

$$\begin{aligned} A_n &= \frac{\alpha(n,h_n)h_n(az_1+bz_2)}{\beta(n,h_n)}\gamma^{(1)}(\tau)f_X(\tau) \\ &+ \frac{\alpha(n,h_n)(az_1^2+bz_2^2)}{2\beta^2(n,h_n)}M_+^{(1)}(0)\gamma(\tau)f_X(\tau) + o\left(\frac{\alpha(n,h_n)h_n}{\beta(n,h_n)}\right) + o\left(\frac{\alpha(n,h_n)}{\beta^2(n,h_n)}\right). \end{aligned}$$

Nous en déduisons

Lemme 1.5. Sous les hypothèses H1, H2 et H3, alors, quand  $n \to \infty$ , nous avons: a)  $A_n \to L_1(az_1 + bz_2)\gamma^{(1)}(\tau)f_X(\tau) + \frac{L_2(az_1^2 + bz_2^2)}{2}M_+^{(1)}(0)\gamma(\tau)f_X(\tau),$ b)  $, n(u) \to , (u),$ où

$$, (u) = \begin{cases} 0 & si \ u < 0 \\ 2L_3(az_1 + bz_2)^2 f_X(\tau)\sigma^2(\tau) \int (M^{(1)}_+(x))^2 dx & si \ u \ge 0 \end{cases}$$

Nous avons également

$$\begin{split} \sup_{1 \le i \le n} & \mathbb{P}\left[ |aM_i(z_1) + bM_i(z_2) - \mathbb{E}\left[ aM_i(z_1) + bM_i(z_2) \right] | \ge \epsilon \right] \\ &= & \mathbb{P}\left[ |aM_1(z_1) + bM_1(z_2) - \mathbb{E}\left[ aM_1(z_1) + bM_1(z_2) \right] | \ge \epsilon \right], \\ &\le \frac{& \mathbb{E}\left[ \left( aM_1(z_1) + bM_1(z_2) - \mathbb{E}\left[ aM_1(z_1) + bM_1(z_2) \right] \right)^2 \right]}{\epsilon^2}, \\ &\le \frac{, n(\infty)}{n\epsilon^2} \to 0 \quad (n \to \infty). \end{split}$$

D'où

**Lemme 1.6.** Sous les hypothèses H1, H2, H3 et pour tout  $\epsilon > 0$ , nous avons

$$\sup_{1\leq i\leq n} \mathbb{P}\left[\left|aM_i(z_1) + bM_i(z_2) - \mathbb{E}\left[aM_i(z_1) + bM_i(z_2)\right]\right| \geq \epsilon\right] \to 0,$$

quand  $n \to \infty$ 

D'après Gnedenko [GK68], les lemmes 1.5 et 1.6 montrent que  $a\mathcal{M}_n(z_1) + b\mathcal{M}_n(z_2)$  admet une limite en loi. La loi limite est la loi normale de moyenne et de variance donnée par le lemme 1.5. D'où nous avons démontré la convergence cylindrique, sous les hypothèses H1, H2 et H3, du processus  $\mathcal{M}_n$  vers un processus gaussien  $\mathcal{M}$  défini par :

i) 
$$\mathbb{E}\left[\mathcal{M}(z)\right] = L_1 z \gamma^{(1)}(\tau) f_X(\tau) + \frac{L_2 z^2}{2} M_+^{(1)}(0) \gamma(\tau) f_X(\tau),$$
  
ii)  $Cov\left(\mathcal{M}(z_1), \mathcal{M}(z_2)\right) = 2L_3 z_1 z_2 f_X(\tau) \sigma^2(\tau) \left(\int_{-1}^0 \left(M_+'(x)\right)^2 dx\right)$ 

Puisque

$$\mathcal{Z}_n(z) = \frac{1 + o_P(1)}{f_X(\tau)} \mathcal{M}_n(z),$$

alors

$$\mathcal{Z}_n(z) \Rightarrow \mathcal{Z}(z),$$

où  ${\mathcal Z}$  est un processus gaussien, avec

i) 
$$\mathbb{E}\left[\mathcal{Z}(z)\right] = L_1 z \gamma^{(1)}(\tau) + \frac{L_2 z^2}{2} M_+^{(1)}(0) \gamma(\tau),$$
  
ii)  $Cov\left(\mathcal{Z}(z_1), \mathcal{Z}(z_2)\right) = 2L_3 z_1 z_2 \frac{\sigma^2(\tau)}{f_X(\tau)} \left(\int_{-1}^0 (M_+'(x))^2 dx\right),$ 

Lemme 1.7. Sous les hypothèses H1, H2 et H3, le processus  $\mathcal{M}_n$  est tendu.

#### Preuve du lemme 1.7

Sous les hypothèses H1, H2 et H3, le processus  $\mathcal{Z}_n(z)$  vit dans l'espace  $\mathcal{C}[-\mathcal{M}, \mathcal{M}]$ . Pour montrer que la suite de processus  $(\mathcal{M}_n(z))$  est tendue, il suffit (cf Billingsley [BIL68]), de v 'erifier que pour tout  $z_1 \leq z_2$ :

$$\mathbb{E}\left[\left(\mathcal{M}_n(z_2)-\mathcal{M}_n(z_1)\right)^2\right] \leq \lambda(z_2-z_1)^2.$$

D'après l'inégalité de Cauchy-Schwartz, nous obtenons

$$\begin{split} & \mathbb{E}\left[\left(\mathcal{M}_{n}(z_{2})-\mathcal{M}_{n}(z_{1})\right)^{2}\right] \\ &\leq \mathbb{E}\left[\left(\mathcal{M}_{n}^{+}(z_{2})-\mathcal{M}_{n}^{+}(z_{1})\right)^{2}\right]+\mathbb{E}\left[\left(\mathcal{M}_{n}^{-}(z_{2})-\mathcal{M}_{n}^{-}(z_{1})\right)^{2}\right] \\ &+ 2\mathbb{E}\left[|\mathcal{M}_{n}^{+}(z_{2})-\mathcal{M}_{n}^{+}(z_{1})||\mathcal{M}_{n}^{-}(z_{2})-\mathcal{M}_{n}^{-}(z_{1})|\right], \\ &\leq \mathbb{E}\left[\left(\mathcal{M}_{n}^{+}(z_{2})-\mathcal{M}_{n}^{+}(z_{1})\right)^{2}\right]+\mathbb{E}\left[\left(\mathcal{M}_{n}^{-}(z_{2})-\mathcal{M}_{n}^{-}(z_{1})\right)^{2}\right] \\ &+ 2\left[\mathbb{E}\left[\left(\mathcal{M}_{n}^{+}(z_{2})-\mathcal{M}_{n}^{+}(z_{1})\right)^{2}\right]\mathbb{E}\left[\left(\mathcal{M}_{n}^{-}(z_{2})-\mathcal{M}_{n}^{-}(z_{1})\right)^{2}\right]\right]^{1/2}. \end{split}$$

Or

$$\mathbb{E}\left[\left(\mathcal{M}_{n}^{\pm}(z_{2})-\mathcal{M}_{n}^{\pm}(z_{1})\right)^{2}\right] = \mathbb{E}\left[\left(\sum_{i=1}^{n}M_{i}^{\pm}(z_{2})-M_{i}^{\pm}(z_{1})\right)^{2}\right],\\ = n\mathbb{E}\left[\left(M_{1}^{\pm}(z_{2})-M_{1}^{\pm}(z_{1})\right)^{2}\right]+n(n-1)\left[\mathbb{E}\left(M_{1}^{\pm}(z_{2})-M_{1}^{\pm}(z_{1})\right)\right]^{2}.$$

Puisque

$$\begin{split} & \mathbb{E}\left[\left(M_{1}^{\pm}(z_{2})-M_{1}^{\pm}(z_{1})\right)^{2}\right] \\ &= \frac{\alpha^{2}(n,h_{n})}{n^{2}h_{n}}\int\left[\left(m(\tau+h_{n}(y_{1}-x))-m_{\pm}(\tau)\right)^{2}+\sigma^{2}(\tau+h_{n}(y_{1}-x))\right] \\ &\qquad \times(g_{\pm}(x,z_{2}-z_{1}))^{2}f_{X}(\tau+h_{n}(y_{1}-x))dx \\ &\leq c\frac{\alpha^{2}(n,h_{n})}{n^{2}h_{n}}\int\left(g_{\pm}(x,z_{2}-z_{1})\right)^{2}dx \end{split}$$

 $\operatorname{et}$ 

$$\begin{split} &\left[\mathbb{E}\left(M_{1}^{\pm}(z_{2})-M_{1}^{\pm}(z_{1})\right)\right]^{2} \\ &=\frac{\left[\int\left(m(\tau+h_{n}(y_{1}-x))-m_{\pm}(\tau)\right)g_{\pm}(x,z_{2}-z_{1})f_{X}(\tau+h_{n}(y_{1}-x))dx\right]^{2}}{n^{2}/\alpha^{2}(n,h_{n})} \\ &\leq c\frac{\alpha^{2}(n,h_{n})}{n^{2}}\int\left(g_{\pm}(x,z_{2}-z_{1})\right)^{2}dx \end{split}$$

 $\operatorname{alors}$ 

$$\mathbb{E}\left[\left(\mathcal{M}_{n}^{\pm}(z_{2})-\mathcal{M}_{n}^{\pm}(z_{1})\right)^{2}\right] \leq c_{1}\frac{\alpha^{2}(n,h_{n})}{nh_{n}}\left[1+nh_{n}\right]\int \left(g_{\pm}(x,z_{2}-z_{1})\right)^{2}dx \\ \leq c_{2}\frac{\alpha^{2}(n,h_{n})}{nh_{n}\beta^{2}(n,h_{n})}\left[1+nh_{n}\right]\left(z_{2}-z_{1}\right)^{2},$$

où nous avons appliqué le résultat b) de la proposition 1.1. D'où sous les hypothèses H1, H2 et H3 et pour tout  $z_1 \leq z_2$ ,

$$\mathbb{E}\left[\left(\mathcal{M}_n(z_2) - \mathcal{M}_n(z_1)\right)^2\right] \leq \lambda(z_2 - z_1)^2.$$
(1.21)

Nous en déduisons, d'après l'inégalité (1.21) que la suite de processus  $(\mathcal{M}_n)$  est tendue.  $\Box$ 

Enfin, ayant montré la convergence cylindrique et la tension du processus  $\mathcal{M}_n$ , nous en déduisons la convergence du processus  $\mathcal{M}_n$  vers un processus limite  $\mathcal{M}$ .

#### Théorème 1.4.

Sous les hypothèses H1, H2 et H3 nous avons,

$$\mathcal{Z}_n \Rightarrow \mathcal{Z}, \qquad sur \ \mathcal{C}[-M, M]$$

$$(1.22)$$

оù

$$\mathcal{Z}(z) = L_2 \frac{\gamma(\tau)}{2} M_+^{(1)}(0) z^2 + (L_1 \gamma^{(1)}(\tau) + U) z, \qquad (1.23)$$

avec

$$U \sim \mathcal{N}\left(0, \left(\frac{2L_{3}\sigma^{2}(\tau)}{f_{X}(\tau)}\int_{-1}^{0} [M_{+}^{(1)}(x)]^{2}dx\right)\right).$$

Corollaire 1.4. Sous les hypothèses H1, H2 et H3,

$$\frac{\beta(n,h_n)}{h_n} \left( \hat{\tau} - \tau \right) \xrightarrow{\mathcal{D}} \mathcal{N} \left( B, V^2 \right).$$

 $o\dot{u}$ 

$$V^{2} = \frac{2L_{3}\sigma^{2}(\tau)}{\left(L_{2}\gamma(\tau)M_{+}^{(1)}(0)\right)^{2}f_{X}(\tau)}\int_{-1}^{0}[M_{+}^{(1)}(x)]^{2}dx,$$

et

$$B = -\frac{L_1 \gamma^{(1)}(\tau)}{L_2 \gamma(\tau) M_+^{(1)}(0)}.$$

#### Preuve du corollaire 1.4

Il découle de Whitt [WHI70] que nous pouvons étendre la convergence du processus  $\mathcal{Z}_n$  vers le processus  $\mathcal{Z}$  dans l'espace  $\mathcal{C}([-M, M])$  à la convergence dans l'espace  $\mathcal{C}(\mathbb{R})$ . Ainsi nous en déduisons que

$$\sup_{z} \mathcal{Z}_{n}(z) \xrightarrow{\mathcal{D}} \sup_{z} \mathcal{Z}(z),$$

 $\operatorname{et}$ 

$$\arg\sup_{z} \mathcal{Z}_{n}(z) \xrightarrow{\mathcal{D}} \arg\sup_{z} \mathcal{Z}(z).$$

lorsque

 $n \to \infty$ .

Notons  $z_{max} = \arg \sup_{z} \mathcal{Z}(z)$ , alors

$$z_{max} = \frac{L_1 \gamma^{(1)}(\tau) + U}{L_2 \gamma(\tau) M_+^{(1)}(0)},$$

оù

$$U \sim \mathcal{N}\left(0, \left(\frac{2L_{3}\sigma^{2}(\tau)}{f_{X}(\tau)}\int_{-1}^{0} [M_{+}^{(1)}(x)]^{2}dx\right)\right).$$

Posons  $z_{max}(n)$  le point où le processus  $\mathcal{Z}_n$  atteint son maximum. Par construction, nous avons

$$\hat{\tau} = \tau + \frac{h_n}{\beta(n, h_n)} z_{max}(n), \qquad (1.24)$$

Nous en déduisons que

$$\frac{\beta(n,h_n)}{h_n} \left( \hat{\tau} - \tau \right) \xrightarrow{\mathcal{D}} z_{max}.$$

Ceci termine la démonstration du corollaire 1.4 .

**Corollaire 1.5.** Supposons que les hypothèses H1, H2 et H3,  $nh_n^5 \to 0$  et  $\mathbb{E}\left(|\varepsilon|^{2+\delta}\right) < \infty$  où  $\delta > 0$ . Si  $\sqrt{nh_n}/\alpha(n,h_n) \to 0$  alors

$$\sqrt{nh_n}\left(\hat{\gamma}(\hat{\tau}) - \gamma(\tau)\right) \xrightarrow{\mathcal{D}} \mathcal{N}\left(0, \frac{2\sigma^2(\tau)}{f_X(\tau)}V_+^2\right).$$

#### Preuve du corollaire 1.5

Nous avons

$$\begin{split} \sqrt{nh_n} \left( \hat{\gamma}(\hat{\tau}) - \gamma(\tau) \right) &= \sqrt{nh_n} \left( \hat{\gamma}(\hat{\tau}) - \hat{\gamma}(\tau) \right) + \sqrt{nh_n} \left( \hat{\gamma}(\tau) - \gamma(\tau) \right), \\ &= \frac{\sqrt{nh_n}}{\alpha(n,h_n)} \mathcal{Z}_n \left( z_{max}(n) \right) + \sqrt{nh_n} \left( \hat{\gamma}(\tau) - \gamma(\tau) \right), \\ &= o_P(1) + \sqrt{nh_n} \left( \hat{\gamma}(\tau) - \gamma(\tau) \right), \\ &\stackrel{\mathcal{D}}{\longrightarrow} \quad \mathcal{N} \left( 0, \frac{2\sigma^2(\tau)}{f_X(\tau)} V_+^2 \right). \end{split}$$

En remplaçant  $f_X(\tau)$  et  $\sigma^2(\tau)$  par des estimateurs consistants, nous obtenons un corollaire analogue au corollaire 1.3.

**Corollaire 1.6.** Soient  $\hat{f}_X(\tau)$  et  $\hat{\sigma}^2(\tau)$  deux estimateurs consistants de  $f_X(\tau)$  et  $\sigma^2(\tau)$ . L'estimateur  $\hat{f}_X(\tau)$  est supposé positif. Sous les mêmes hypothèses que celles du corollaire 1.6, nous avons

$$\sqrt{nh_n} \frac{\sqrt{\hat{f}_X(\tau)} \left(\hat{\gamma}(\hat{\tau}) - \gamma(\tau)\right)}{(2\hat{\sigma}^2(\tau))^{1/2}} \xrightarrow{\mathcal{L}} \mathcal{N}\left(0, V_+^2\right).$$
(1.25)

Vitesse de convergence de  $\hat{\tau}$  et  $\hat{\gamma}(\hat{\tau})$ 

Étant donné deux suites réelles positives  $(a_n)$  et  $(b_n)$  nous notons  $a_n \sim b_n$  lorsque  $a_n/b_n \to L$ , où  $0 < L < \infty$ 

- 1) Si  $h_n \sim n^{-1/3}$  prenons  $\beta(n, h_n) \sim n^{1/3}$  et  $\alpha(n, h_n) \sim n^{2/3}$ . Ceci implique :
  - i)  $n^{2/3}(\hat{\tau} \tau) \xrightarrow{\mathcal{D}} \mathcal{N}(B, V^2)$ , où les constantes  $L_1$  et  $L_3$  qui interviennent dans les définitions de B et  $V^2$  sont telles que  $L_1L_3 > 0$ .

ii) 
$$n^{1/3} \left( \hat{\gamma}(\hat{\tau}) - \gamma(\tau) \right) \xrightarrow{\mathcal{D}} \mathcal{N} \left( 0, \frac{2\sigma^2(\tau)}{f_X(\tau)} V_+^2 \right).$$

Dans ce premier cas on est loin des vitesses de convergence optimales pour  $\hat{\tau}$  et  $\hat{\gamma}(\hat{\tau})$ .

- 2) Si  $h_n = o(n^{-1/3})$  prenons  $\beta(n, h_n) \sim \sqrt{nh_n}$  et  $\alpha(n, h_n) \sim nh_n$ . Ceci implique :
  - i)  $\sqrt{\frac{n}{h_n}} (\hat{\tau} \tau) \xrightarrow{\mathcal{D}} \mathcal{N} (0, V^2).$ ii)  $\sqrt{nh_n} (\hat{\gamma}(\hat{\tau}) - \gamma(\tau)) \xrightarrow{\mathcal{D}} \mathcal{N} \left(0, \frac{2\sigma^2(\tau)}{f_X(\tau)}V_+^2\right).$

Ici on a  $L_1 = 0$  et  $L_3 > 0$ . Dans ce deuxième cas on peut s'approcher des vitesses de convergence optimales pour  $\hat{\tau}$  et  $\hat{\gamma}(\hat{\tau})$  mais avec des choix de fenêtre de lissage qui ne sont pas compatibles. Une façon pour pallier ce défaut consiste tout d'abord à localiser la rupture en utilisant des fenêtre de lissage de petites tailles (par exemple  $h_n^1 \sim \frac{\log n}{n}$ ) et ensuite à estimer  $\hat{\gamma}(\hat{\tau}_{h_n^1})$  en utilisant des fenêtres de lissage plus grandes (par exemple  $h_n^2 \sim \frac{1}{(n \log n)^{1/5}}$ ).

3) Si  $nh_n^3 \to \infty$  prenons  $\beta(n, h_n) \sim \frac{1}{h_n}$  et  $\alpha(n, h_n) \sim \frac{1}{h_n^2}$ . Il en résulte que  $L_1 > 0$  et  $L_3 = 0$  et par conséquent  $\mathcal{Z}(z) = L_2 \frac{\gamma(\tau)}{2} M_+^{(1)}(0) z^2 + L_1 \gamma^{(1)}(\tau) z$ . Ceci implique :

i) 
$$\frac{(\hat{\tau}-\tau)}{h_n^2} \to -\frac{L_1\gamma^{(1)}(\tau)}{L_2\gamma(\tau)M_+^{(1)}(0)}.$$
  
ii) Si  $h_n = o\left(n^{-1/5}\right)$ , alors  $\sqrt{nh_n}\left(\hat{\gamma}(\hat{\tau}) - \gamma(\tau)\right) \xrightarrow{\mathcal{D}} \mathcal{N}\left(0, \frac{2\sigma^2(\tau)}{f_X(\tau)}V_+^2\right).$ 

Dans ce dernier cas la vitesse de convergence de  $\hat{\tau}$  est  $h_n^2$ . Dans le meilleur des cas cette vitesse ne dépasse pas  $n^{-2/3}$  (comparer avec la vitesse optimale qui est  $n^{-1}$ ).

#### 1.4.3 La situation où le noyau est strictement positif en zéro

Sous les hypothèses H1, H2 et H4

a) Si x < -y < 0 ou x < 0 < -y ou -y < 0 < x ou 0 < -y < x,

1) 
$$g_1(x,z)f_X(\tau - xh_n) = \mathcal{O}(\frac{h_n}{\beta(n,h_n)}),$$

2) 
$$g_2(x,z) - \frac{h_n}{\alpha(n,h_n)} A_n = \mathcal{O}_P(\frac{1}{\beta(n,h_n)})$$

b) Si -y < x < 0 ou 0 < x < -y,

1) 
$$g_1(x,z)f_X(\tau - xh_n) = -\gamma(\tau)f_X(\tau)M_+(0) + o(1),$$

2) 
$$g_2(x,z) - \frac{h_n}{\alpha(n,h_n)} A_n = (-\gamma(\tau) - 2sign(z)\sigma(\tau)\varepsilon) M_+(0) + o_P(1).$$

Donc

$$A_{n} = -\frac{\alpha(n,h_{n})}{\beta(n,h_{n})}\gamma(\tau)f_{X}(\tau)M_{+}(0)(a|z_{1}|+b|z_{2}|) + o(\frac{\alpha(n,h_{n})}{\beta(n,h_{n})}).$$

Nous en déduisons

**Lemme 1.8.** Sous les hypothèses H1, H2 et H4 alors, quand  $n \to \infty$ , nous avons

a)  $A_n \to A = -L_4 \gamma(\tau) f_X(\tau) M_+(0) (a|z_1| + b|z_2|),$ 

b) , 
$$_n(u) \rightarrow$$
 ,  $(u)$  ,

 $o\dot{u}$ 

1) Si 
$$0 \le z_1 \le z_2$$
 alors,  
,  $(u) = 4L_6 M_+^2(0) f_X(\tau) \left[ z_1(a+b)^2 \nu_{a+b}^+(u) + (z_2 - z_1) b^2 \nu_b^+(u) \right].$ 

2) Si 
$$z_2 \le z_1 \le 0$$
 alors,  
,  $(u) = 4L_6 M_+^2(0) f_X(\tau) \left[ -z_1(a+b)^2 \nu_{a+b}^-(u) + (z_1 - z_2) b^2 \nu_b^-(u) \right].$ 

3) Si 
$$z_1 \le 0 \le z_2$$
 alors,  
,  $(u) = 4L_6 M_+^2(0) f_X(\tau) \left[ -z_1 a^2 \nu_a^-(u) + z_2 b^2 \nu_b^+(u) \right].$ 

avec

$$\nu_a^{\pm}(u) = \mathbb{E}\left[ \left( -\gamma(\tau) \mp 2\sigma(\tau)\varepsilon \right)^2 \mathcal{I}_{\{2L_5 a M_+(0)(-\gamma(\tau)+\sigma(\tau)\varepsilon) \le u\}} \right].$$

Nous avons également

$$\begin{split} \sup_{1 \le i \le n} & \mathbb{P}\left[ |aM_i(z_1) + bM_i(z_2) - \mathbb{E}\left[ aM_i(z_1) + bM_i(z_2) \right] | \ge \epsilon \right] \\ &= & \mathbb{P}\left[ |aM_1(z_1) + bM_1(z_2) - \mathbb{E}\left[ aM_1(z_1) + bM_1(z_2) \right] | \ge \epsilon \right], \\ &\le \frac{& \mathbb{E}\left[ \left( aM_1(z_1) + bM_1(z_2) - \mathbb{E}\left[ aM_1(z_1) + bM_1(z_2) \right] \right)^2 \right]}{\epsilon^2}, \\ &\le \frac{& , n(\infty)}{n\epsilon^2} \to 0 \quad (n \to \infty). \end{split}$$

D'où

**Lemme 1.9.** Sous les hypothèses H1, H2 et H4 et pour tout  $\epsilon > 0$ , nous avons

$$\sup_{1\leq i\leq n} \mathbb{P}\left[\left|aM_i(z_1) + bM_i(z_2) - \mathbb{E}\left[aM_i(z_1) + bM_i(z_2)\right]\right| \geq \epsilon\right] \to 0,$$

quand  $n \to \infty$ 

Posons

1)  $\lambda_1 = 2L_5 M_+(0),$ 2) si  $L_5 > 0, \lambda_2 = \frac{L_6}{L_5^2} f_X(\tau),$ 3)  $\lambda_3 = L_4 M_+(0) f_X(\tau) = \frac{\lambda_1 \lambda_2}{2},$ 4)  $\xi^{\pm} = -\gamma(\tau) \mp 2\sigma(\tau)\varepsilon.$ 

Nous avons alors,

a) Si  $0 \leq z_1 \leq z_2$ , alors

$$\frac{d, (u)}{u^2} = \lambda_2 \left[ z_1 dF_{\lambda_1(a+b)\xi^+}(u) + (z_2 - z_1) dF_{\lambda_1 b\xi^+}(u) \right].$$

b) Si  $z_2 \leq z_1 \leq 0$ , alors

$$\frac{d, (u)}{u^2} = \lambda_2 \left[ -z_1 dF_{\lambda_1(a+b)\xi^-}(u) + (z_1 - z_2) dF_{\lambda_1 b\xi^-}(u) \right].$$

c) Si  $z_1 \leq 0 \leq z_2$ , alors

$$\frac{d, (u)}{u^2} = \lambda_2 \left[ -z_1 dF_{\lambda_1 a \xi^-}(u) + z_2 dF_{\lambda_1 b \xi^+}(u) \right].$$

où  $F_X(\cdot)$  est la fonction de répartition de X.

D'après Gnedenko [GK68], les lemmes 1.8 et 1.9 montrent que  $a\mathcal{M}_n(z_1) + b\mathcal{M}_n(z_2)$  admet une limite en loi. Le logarithme de la fonction caractéristique de la loi limite est donné par la formule de Kolmogorov (cf théorème 1.3) avec la constante égale à A et la fonction , (u) est définie comme dans le lemme 1.8.

Notons  $\phi(\cdot)$  la fonction caractéristique de la loi limite et  $\phi_X(\cdot)$  celle de X. D'après le lemme 1.8, nous avons

a) Si  $0 \le z_1 \le z_2$ , alors

$$\begin{split} \phi(t) &= & \exp\left[i\lambda_3\gamma(\tau)(a|z_1|+b|z_2|)t\right] \\ &\times \exp\left[\lambda_2z_1\left(\phi_{\lambda_1(a+b)\xi^+}(t)-1\right)\right] \\ &\times \exp\left[\lambda_2(z_2-z_1)\left(\phi_{\lambda_1b\xi^+}(t)-1\right)\right]. \end{split}$$

b) Si  $z_2 \leq z_1 \leq 0$ , alors

$$\begin{split} \phi(t) &= & \exp\left[i\lambda_{3}\gamma(\tau)\left(a|z_{1}|+b|z_{2}|\right)t\right] \\ &\times \exp\left[-\lambda_{2}z_{1}\left(\phi_{\lambda_{1}\left(a+b\right)\xi^{-}}\left(t\right)-1\right)\right] \\ &\times \exp\left[\lambda_{2}\left(z_{1}-z_{2}\right)\left(\phi_{\lambda_{1}b\xi^{-}}\left(t\right)-1\right)\right]. \end{split}$$

c) Si  $z_1 \leq 0 \leq z_2$ , alors

$$\begin{split} \phi(t) &= &\exp\left[i\lambda_{3}\gamma(\tau)(a|z_{1}|+b|z_{2}|)t\right] \\ &\times &\exp\left[-\lambda_{2}z_{1}\left(\phi_{\lambda_{1}a\xi^{-}}(t)-1\right)\right] \\ &\times &\exp\left[\lambda_{2}z_{2}\left(\phi_{\lambda_{1}b\xi^{+}}(t)-1\right)\right]. \end{split}$$

D'où nous avons démontré la convergence cylindrique, sous les hypothèses H1, H2 et H4, du processus  $\mathcal{M}_n$  vers le processus de Poisson composé bilatéral  $\mathcal{M}$  défini comme suit :

$$\mathcal{M}(z) = \lambda_3 \gamma(\tau) |z| + \lambda_1 \mathcal{N}(z),$$

оù

$$\mathcal{N}(z) = \begin{cases} 0 & \text{si } z = 0, \\ \sum_{i=1}^{N_z^+} -\gamma(\tau) - 2\sigma(\tau)\varepsilon_i^+ & \text{si } z > 0, \\ \sum_{i=1}^{N_{-z}^-} -\gamma(\tau) + 2\sigma(\tau)\varepsilon_i^- & \text{si } z < 0, \end{cases}$$
(1.26)

avec  $N_z^+$  et  $N_z^-$  sont des processus de Poisson homogènes et indépendants, de paramètre  $\lambda_2$ . Les  $\varepsilon_i^+$ 's et  $\varepsilon_i^-$  sont des variables aléatoires, indépendantes, identiquement distribuées et de même loi que  $\varepsilon$ . Elles sont aussi indépendantes de  $N_z^+$  et  $N_z^-$ .

On vérifie aisément

i) 
$$\mathbb{E}\left[\mathcal{M}(z)\right] = -\lambda_3 \gamma(\tau)|z|,$$

ii) 
$$Cov(\mathcal{M}(z_1), \mathcal{M}(z_2)) = \lambda_1^2 \lambda_2 \left(\gamma^2(\tau) + \sigma^2(\tau)\right) \left(\min(|z_1|, |z_2|) \coprod_{\{z_1, z_2 > 0\}}\right).$$

Puisque

$$\mathcal{Z}_n(z) = \frac{1 + o_P(1)}{f_X(\tau)} \mathcal{M}_n(z),$$

 $\operatorname{alors}$ 

$$\mathcal{Z}_n(z) \Rightarrow \mathcal{Z}(z),$$

оù

a)

$$\mathcal{Z}(z) = \frac{\lambda_3}{f_X(\tau)} \gamma(\tau) |z| + \frac{\lambda_1}{f_X(\tau)} \mathcal{N}(z).$$

Lemme 1.10. Sous les hypothèses H1, H2 et H4, la suite de processus  $(\mathcal{M}_n)$  est tendue.

#### Preuve du lemme 1.10

Sous les hypothèses H1, H2 et H4, le processus  $\mathcal{Z}_n(z)$  vit dans l'espace  $\mathcal{D}[-\mathcal{M}, \mathcal{M}]$ . Pour montrer que la suite de processus  $(\mathcal{M}_n(z))$  est tendue, il suffit (cf Billingsley [BIL68]), de vérifier que pour tout  $z_1 \leq z \leq z_2$ :

$$\mathbb{E}\left[\left|\mathcal{M}_{n}(z_{2})-\mathcal{M}_{n}(z)\right|\right|\mathcal{M}_{n}(z_{1})-\mathcal{M}_{n}(z)\right] \leq \lambda(z_{2}-z_{1})^{2}.$$

En utilisant l'inégalité triangulaire, le terme à gauche de l'inégalité ci-dessus peut être majoré comme suit :

$$\begin{split} & \mathbb{E}\left[|\mathcal{M}_{n}(z_{2}) - \mathcal{M}_{n}(z)||\mathcal{M}_{n}(z_{1}) - \mathcal{M}_{n}(z)|\right] \\ & \leq \mathbb{E}\left[|\mathcal{M}_{n}^{+}(z_{2}) - \mathcal{M}_{n}^{+}(z)||\mathcal{M}_{n}^{+}(z_{1}) - \mathcal{M}_{n}^{+}(z)|\right] + \mathbb{E}\left[|\mathcal{M}_{n}^{+}(z_{2}) - \mathcal{M}_{n}^{+}(z)||\mathcal{M}_{n}^{-}(z_{1}) - \mathcal{M}_{n}^{-}(z)|\right] \\ & + \mathbb{E}\left[|\mathcal{M}_{n}^{-}(z_{2}) - \mathcal{M}_{n}^{-}(z)||\mathcal{M}_{n}^{+}(z_{1}) - \mathcal{M}_{n}^{+}(z)|\right] + \mathbb{E}\left[|\mathcal{M}_{n}^{-}(z_{2}) - \mathcal{M}_{n}^{-}(z)||\mathcal{M}_{n}^{-}(z_{1}) - \mathcal{M}_{n}^{-}(z)|\right]. \end{split}$$

Or

$$\begin{split} &\mathbb{E}\left[|\mathcal{M}_{n}^{\pm}(z_{2}) - \mathcal{M}_{n}^{\pm}(z)||\mathcal{M}_{n}^{\pm}(z_{1}) - \mathcal{M}_{n}^{\pm}(z)|\right] \\ &= \mathbb{E}\left[\left|\sum_{i=1}^{n} \left(M_{i}^{\pm}(z_{2}) - M_{i}^{\pm}(z)\right)\right|\right|\sum_{i=1}^{n} \left(M_{i}^{\pm}(z_{1}) - M_{i}^{\pm}(z)\right)\right|\right], \\ &\leq \sum_{1 \leq i, j \leq n} \mathbb{E}\left[\left|M_{i}^{\pm}(z_{2}) - M_{i}^{\pm}(z)\right|\right|M_{j}^{\pm}(z_{1}) - M_{j}^{\pm}(z)\right|\right], \\ &= n\mathbb{E}\left[\left|M_{1}^{\pm}(z_{2}) - M_{1}^{\pm}(z)\right|\left|M_{1}^{\pm}(z_{1}) - M_{1}^{\pm}(z)\right|\right] \\ &+ n(n-1)\mathbb{E}\left[\left|M_{1}^{\pm}(z_{2}) - M_{1}^{\pm}(z)\right|\right] \times \mathbb{E}\left[\left|M_{1}^{\pm}(z_{1}) - M_{1}^{\pm}(z)\right|\right] \end{split}$$

Puisque

$$\begin{split} & \mathbb{E}\left[\left|M_{1}^{\pm}(z_{2})-M_{1}^{\pm}(z)\right|\left|M_{1}^{\pm}(z_{1})-M_{1}^{\pm}(z)\right|\right] \\ & \leq \frac{\alpha^{2}(n,h_{n})}{n^{2}h_{n}}\int \mathbb{E}[|Y-m_{\pm}(\tau)||Y-m_{\pm}(\tau)||X=\tau+h_{n}(y-x)] \\ & \times |g_{\pm}(x,z_{2}-z)||g_{\pm}(x,z_{1}-z)|f_{X}(\tau+h_{n}(y-x))dx \\ & \leq c_{1}\frac{\alpha^{2}(n,h_{n})}{n^{2}h_{n}}\int |g_{\pm}(x,z_{2}-z)||g_{\pm}(x,z_{1}-z)|dx , \\ & \leq c_{2}\frac{\alpha^{2}(n,h_{n})}{n^{2}h_{n}\beta^{2}(n,h_{n})}|z_{2}-z||z_{1}-z| , \end{split}$$

où nous avons appliqué le résultat c) de la proposition 1.1.

Nous avons également

$$\begin{split} & \operatorname{I\!E} \left[ \left| M_{1}^{\pm}(z_{1}) - M_{1}^{\pm}(z) \right| \right] \\ & \leq \frac{\alpha(n,h_{n})}{n} \int \operatorname{I\!E} \left[ \left| Y - m_{\pm}(\tau) \right| \left| X = \tau + h_{n}(y-x) \right] \\ & \times \left| g_{\pm}(x,z_{1}-z) \right| f_{X}(\tau + h_{n}(y-x)) dx \right] \\ & \leq c_{2} \frac{\alpha(n,h_{n})}{n} \int \left| g_{\pm}(x,z_{1}-z) \right| dx , \\ & \leq c_{3} \frac{\alpha(n,h_{n})}{n\beta(n,h_{n})} |z_{1}-z| . \end{split}$$

D'où sous les hypothèses H1, H2 et H4 et pour tout  $z_1 \leq z \leq z_2$ ,

$$\mathbb{E}\left[\left|\mathcal{M}_{n}(z_{2})-\mathcal{M}_{n}(z)\right|\left|\mathcal{M}_{n}(z_{1})-\mathcal{M}_{n}(z)\right|\right] \leq \lambda(z_{2}-z_{1})^{2}.$$
(1.27)

Nous en déduisons, d'après (1.27) que le processus  $\mathcal{M}_n$  est tendu.

Enfin, en montrant la convergence cylindrique et la tension du processus  $\mathcal{M}_n$ , nous établissons la convergence du processus  $\mathcal{M}_n$  vers un processus limite  $\mathcal{M}$ . Nous en déduisons le théorème suivant

Théorème 1.5. Sous les hypothèses H1, H2 et H4 nous avons,

$$\mathcal{Z}_n \Rightarrow \mathcal{Z}, \qquad sur \quad \mathcal{D}[-M, M]$$

$$(1.28)$$

оù

$$\mathcal{Z}(z) = \frac{\lambda_3}{f_X(\tau)} \gamma(\tau) |z| + \frac{\lambda_1}{f_X(\tau)} \mathcal{N}(z), \qquad (1.29)$$

avec  $\mathcal{N}(z)$  est un processus de Poisson composé bilatéral, (voir (1.26)).

Lorsque  $L_5 = 0$  le processus  $\mathcal{N}(z)$  n'est plus défini  $(\lambda_2 = +\infty)$ . Néanmoins le processus limite  $\mathcal{Z}(z)$  existe et il est dégénéré  $(\mathcal{Z}(z) = -\frac{\lambda_3}{f_X(\tau)}\gamma(\tau)|z|)$ .

Corollaire 1.7. Sous les hypothèses H1, H2 et H4,

$$\frac{\beta(n,h_n)}{h_n} \left( \hat{\tau} - \tau \right) \xrightarrow{\mathcal{D}} \mathcal{M}$$

оù

$$\mathcal{M} = \arg \sup_{z} \frac{\lambda_3}{f_X(\tau)} \gamma(\tau) |z| + \frac{\lambda_1}{f_X(\tau)} \mathcal{N}(z).$$
(1.30)

#### Preuve du corollaire 1.7

D'après Whitt [WHI80], nous pouvons étendre la convergence du processus  $\mathcal{Z}_n$  vers le processus  $\mathcal{Z}$  dans l'espace  $\mathcal{D}([-M, M])$  à la convergence dans l'espace  $\mathcal{D}(\mathbb{R})$ . Il en résulte que

$$\sup_{z} \mathcal{Z}_{n}(z) \xrightarrow{\mathcal{D}} \sup_{z} \mathcal{Z}(z),$$

 $\operatorname{et}$ 

$$\arg\sup_{z} \mathcal{Z}_{n}(z) \xrightarrow{\mathcal{D}} \arg\sup_{z} \mathcal{Z}(z),$$

lorsque

 $n \to \infty$ .

Lemme 1.11. Sous les hypothèses H1, H2 et H4, alors

$$\mathbb{IP}\left[\sup_{z\in\mathbb{IR}}\mathcal{Z}(z)<\infty\right]=1.$$

#### Preuve du lemme 1.11

Il est clair que

$$\sup_{z \in \mathrm{IR}} \mathcal{Z}(z) = \max\left(\sup_{z \le 0} \mathcal{Z}(z), \sup_{z \ge 0} \mathcal{Z}(z)\right).$$

Pour tout  $z \ge 0$ , nous avons

$$\begin{aligned} \mathcal{Z}(z) &= \frac{\lambda_3}{f_X(\tau)} \gamma(\tau) z + \frac{\lambda_1}{f_X(\tau)} \mathcal{N}(z), \\ &= \frac{\lambda_3}{f_X(\tau)} \gamma(\tau) z + \frac{\lambda_1}{f_X(\tau)} \sum_{i=1}^{N_z^+} \left( -\gamma(\tau) - 2\sigma(\tau)\varepsilon_i^+ \right). \end{aligned}$$

1. Si  $L_5 = 0$ , alors  $\mathcal{Z}(z) = -\frac{\lambda_3}{f_X(\tau)}\gamma(\tau)|z|$ , et par conséquent

$$\mathbb{P}\left[\sup_{z\in\mathrm{IR}}\mathcal{Z}(z)<\infty\right]=1.$$

2. Si  $L_5 > 0$ , alors

$$\sup_{z \ge 0} \mathcal{Z}(z) = \sup_{z \ge 0} \frac{\lambda_3}{f_X(\tau)} \gamma(\tau) z + \frac{\lambda_1}{f_X(\tau)} \sum_{i=1}^{N_z^+} \left( -\gamma(\tau) - 2\sigma(\tau)\varepsilon_i^+ \right),$$
$$= \sup_{z \ge 0} \frac{\lambda_2}{2} z - \sum_{i=1}^{N_z^+} \left( 1 + \frac{2\sigma(\tau)}{\gamma(\tau)} \varepsilon_i^+ \right).$$

Notons

- 1)  $T_1^+, T_2^+, \ldots$  : les dates d'occurrence du processus de Poisson  $\mathcal{N}_z^+$ .
- 2)  $E_i^+ = T_{i+1}^+ T_i^+$ .



Figure 1.4 : Une trajectoire du processus  $\mathcal{Z}(z)$  sous les hypothèses H1, H2 et H4.

Le maximum du processus  $\mathcal{Z}(z)$  ne peut avoir lieu qu'en une extrémité droite de segment. D'où

$$\begin{split} \sup_{z \ge 0} \mathcal{Z}(z) &= \max_{j \ge 1} \lim_{z \nearrow T_j^+} \mathcal{Z}(z), \\ &= \max_{j \ge 1} \frac{\lambda_2}{2} T_j^+ - \sum_{i=1}^{j-1} \left( 1 + \frac{2\sigma(\tau)}{\gamma(\tau)} \varepsilon_i^+ \right), \\ &= \max \left[ \frac{\lambda_2}{2} T_1^+, \max_{j \ge 2} \left( \frac{\lambda_2}{2} T_1^+ + \sum_{i=1}^{j-1} \Theta_i^+ \right) \right], \\ \frac{\lambda_2}{2} E_i^+ - 1 + \frac{2\sigma(\tau)}{\gamma(\tau)} \varepsilon_i^+. \end{split}$$

On définit la marche aléatoire

où  $\Theta_i^+ =$ 

$$W_i^+ = \begin{cases} \frac{\lambda_2}{2} T_1^+ & \text{si } i = 1, \\ \\ W_{i-1}^+ + \Theta_i^+ & \text{si } i > 1. \end{cases}$$

Puisque  $E_i \sim \mathcal{EXP}(\lambda_2)$  (loi exponentielle de paramètre  $\lambda_2$ ) et  $\mathbb{E}\varepsilon_i = 0$ , il est facile de vérifier que

$$I\!E\Theta_i^+ = -1/2.$$

Or toute marche aléatoire  $\sum_{i=1}^{n} \Theta_{i}^{+}$  vérifiant  $\mathbb{E}\Theta_{i}^{+} < 0$  admet un maximum presque sûrement (voir par exemple Feller [FEL66, chapitre XII]). Il vient, donc, que sous les hypothèses H1, H2 et H4, le processus  $\mathcal{Z}$  admet un maximum avec une probabilité égale à 1.

Posons  $z_{max}(n)$  le point où le processus  $\mathcal{Z}_n$  atteint son maximum. Par construction, nous avons

$$\hat{\tau} = \tau + \frac{h_n}{\beta(n, h_n)} z_{max}(n).$$
(1.31)

Nous en déduisons que

$$\frac{\beta(n,h_n)}{h_n} \left( \hat{\tau} - \tau \right) \xrightarrow{\mathcal{D}} z_{max},$$
$$\frac{\lambda_1}{\lambda_1} \mathcal{N}(z)$$

où  $z_{max} = \arg \sup_{z} \frac{\lambda_3}{f_X(\tau)} \gamma(\tau) |z| + \frac{\lambda_1}{f_X(\tau)} \mathcal{N}(z).$ 

Ceci termine la démonstration du corollaire 1.7 .

Les corollaires 1.8 et 1.9 ci-dessous sont obtenus de la même façon que leurs analogues dans la situation  $K_+(0) = 0$  (cf les corollaires 1.5 et 1.6).

**Corollaire 1.8.** Supposents que les hypothèses H1, H2 et H4 sont satisfaites,  $nh_n^5 \to 0$ , et  $\mathbb{E}(|\varepsilon|^{2+\delta})$  où  $\delta > 0$ , alors

$$\sqrt{nh_n}\left(\hat{\gamma}(\hat{\tau}) - \gamma(\tau)\right) \xrightarrow{\mathcal{D}} \mathcal{N}\left(0, \frac{2\sigma^2(\tau)}{f_X(\tau)}V_+^2\right).$$

En remplaçant  $f(\tau)$  et  $\sigma^2(\tau)$  par des estimateurs consistants, nous obtenons un corollaire analogue au corollaire 1.3.

**Corollaire 1.9.** Soient  $\hat{f}_X(\tau)$  et  $\hat{\sigma}^2(\tau)$  deux estimateurs consistants de  $f(\tau)$  et  $\sigma^2(\tau)$ . L'estimateur  $\hat{f}_X(\tau)$  est supposé positif. Sous les mêmes hypothèses que celles du corollaire 1.3, nous avons

$$\sqrt{nh_n} \frac{\sqrt{\hat{f}_X(\tau) \left(\hat{\gamma}(\hat{\tau}) - \gamma(\tau)\right)}}{(2\hat{\sigma}^2(\tau))^{1/2}} \xrightarrow{\mathcal{L}} \mathcal{N}\left(0, V_+^2\right).$$
(1.32)

Vitesses de convergence de  $\hat{\tau}$  et  $\hat{\gamma}(\hat{\tau})$ 

La vitesse de convergence de τ̂ est indépendante du choix de la fenêtre de lissage h<sub>n</sub>. Elle est égale à n<sup>-1</sup>. Il s'agit là de la vitesse optimale pour ce type de problème. En effet si on choisit β(n,h<sub>n</sub>) ~ nh<sub>n</sub> et α(n,h<sub>n</sub>) ~ nh<sub>n</sub>. Ceci implique :

$$n\left(\hat{\tau} - \tau\right) = \mathcal{O}_P(1).$$

2) Si de plus  $h_n = o(n^{-1/5})$ , alors  $\sqrt{nh_n} \left(\hat{\gamma}(\hat{\tau}) - \gamma(\tau)\right) \xrightarrow{\mathcal{D}} \mathcal{N}\left(0, \frac{2\sigma^2(\tau)}{f_X(\tau)}V_+^2\right)$ .

## 1.5 Simulations

Nous avons étudié dans les paragraphes précédentes quelques propriétés asymptotiques des paramètres de rupture  $\tau$  et  $\gamma(\tau)$ . Nous présentons maintenant des simulations que nous avons effectuées pour explorer le comportement de nos estimateurs dans un cadre non asymptotique. Il s'agit seulement d'une partie d'un ensemble plus étendu de simulations que nous ne pouvons reproduire ici de manière exhaustive. Dans ces simulations nous étudions les comportements des estimateurs des ruptures en fonction des quatre paramètres suivants :

- 1) Le noyau de lissage.
- 2) La taille de l'échantillon.
- 3) Le rapport signal sur bruit (RSB).
- 4) La fenêtre de lissage.

Soit x le paramètre à estimer (amplitude ou localisation de la discontinuité) et  $\hat{X}$  son estimateur. On note p le nombre d'échantillons simulés et  $\hat{x}_{(i)}$  l'estimation de x à partir du ième échantillon. Dans ce paragraphe, nous adoptons les notations suivantes :

- $biais = \frac{1}{p} \sum_{i=1}^{p} (\hat{x}_{(i)} x).$
- $variance = \frac{1}{p} \sum_{i=1}^{p} (\hat{x}_{(i)} \bar{x}_p)^2$ , où  $\bar{x}_p = \frac{1}{p} \sum_{i=1}^{p} \hat{x}_{(i)}$ .

• 
$$MSE = \frac{1}{p} \sum_{i=1}^{p} (\hat{x}_{(i)} - x)^2 = biais^2 + variance.$$

Les noyaux utilisés dans les simulations sont définis par :

- 1)  $K0(x) = x^2(1+x)^2 \prod_{[-1,0]}(x).$
- 2)  $K1(x) = (1 x^2) 1 \mathbb{I}_{[-1,0]}(x).$
- 3)  $K2(x) = -x(1+x) \coprod_{[-1,0]}(x).$

Les simulations présentées ici sont toutes faites à partir de l'un des modèles ci-dessous et sont répétées 200 fois pour chaque exemple :

M1

• 
$$m(x) = \prod_{[0.4,1]} (x),$$

- $x_i = \frac{i}{n}$ ,
- $\varepsilon$ : variable gaussienne centrée réduite.

М2

- $m(x) = 4\sin(5x) + 3x + \prod_{[0.7,1]} (x),$
- $X_i \sim \mathcal{U}[0,1],$
- $\varepsilon$ : variable gaussienne centrée réduite.

# Noyau de Lissage

Nous avons simulé les modèles M1 et M2 pour deux combinaisons de la taille de l'échantillon et de la variance du bruit, et étudié le comportement des estimateurs en fonction du choix du noyau. Nous constatons que l'utilisation d'un noyau qui ne s'annule pas en zéro conduit en général à une réduction du biais, de la variance et donc de la MSE de chacun des deux estimateurs. Il semble donc que de tels noyaux conduisent à une amélioration de la vitesse de convergence, mais aussi à de meilleures performances dans un cadre non asymptotique. Pour illustrer graphiquement ces différentes réductions de biais , variance et MSE, nous représentons dans les figures 1.5, 1.6, 1.7 et 1.8, le comportement du biais au carré , de la variance et de la MSE en fonction des trois noyaux K0, K1 et K2.



**Figure 1.5 :** Estimations du biais au carré, de la variance et de la MSE à partir de 200 échantillons de taille 200 simulés selon le modèle M1 et pour un niveau de bruit  $\sigma = 0.5$ . On montre à gauche ces différentes estimations pour l'estimateur de l'amplitude de la rupture  $\hat{\gamma}(\hat{\tau})$ et à droite pour l'estimateur de la localisation  $\hat{\tau}$ .



**Figure 1.6 :** Estimations du biais au carré, de la variance et de la MSE à partir de 200 échantillons de taille 1000 simulés selon le modèle M1 et pour un niveau de bruit  $\sigma = 1$ . On montre à gauche ces différentes estimations pour l'estimateur de l'amplitude de la rupture  $\hat{\gamma}(\hat{\tau})$ et à droite pour l'estimateur de la localisation  $\hat{\tau}$ .



**Figure 1.7 :** Estimations du biais au carré, de la variance et de la MSE à partir de 200 échantillons de taille 200 simulés selon le modèle M2 et pour un niveau de bruit  $\sigma = 0.5$ . On montre à gauche ces différentes estimations pour l'estimateur de l'amplitude de la rupture  $\hat{\gamma}(\hat{\tau})$ et à droite pour l'estimateur de la localisation  $\hat{\tau}$ .



**Figure 1.8 :** Estimations du biais au carré, de la variance et de la MSE à partir de 200 échantillons de taille 1000 simulés selon le modèle M2 et pour un niveau de bruit  $\sigma = 1$ . On montre à gauche ces différentes estimations pour l'estimateur de l'amplitude de la rupture  $\hat{\gamma}(\hat{\tau})$ et à droite pour l'estimateur de la localisation  $\hat{\tau}$ .

## Taille de l'échantillon

Les figures 1.9, 1.10, 1.11 et 1.12 illustrent les gains en biais, variance et MSE des deux estimateurs ( $\hat{\tau}$  et  $\hat{\gamma}(\hat{\tau})$ ) obtenus lorsqu'on augmente le nombre d'observations. Il est clair que l'augmentation du nombre d'observations s'accompagne d'une meilleure précision dans les estimations de la localisation et de l'amplitude de la rupture. Dans les situations où le rapport signal sur bruit n'est pas favorable (égale à 1 par exemple), on peut s'attendre à ce que le nombre d'observations ait de l'importance dans la localisation de la rupture et dans l'estimation de l'amplitude du saut. C'est effectivement le cas. Nous observons qu'en augmentant la taille de l'échantillon, les performances s'améliorent, mais il faut noter que le gain est inférieur à celui obtenu en situation de rapport signal sur bruit plus favorable. Un rapport signal sur bruit médiocre nécessite une taille d'échantillon importante.



**Figure 1.9 :** Estimations du biais au carré, de la variance et de la MSE à partir de 200 échantillons de taille n (n = 200,500,1000) simulés selon le modèle M1 et pour un niveau de bruit  $\sigma = 0.5$ . On montre à gauche ces différentes estimations pour l'estimateur de l'amplitude de la rupture  $\hat{\gamma}(\hat{\tau})$  et à droite pour l'estimateur de la localisation  $\hat{\tau}$ . Le noyau utilisé est K1.



**Figure 1.10 :** Estimations du biais au carré, de la variance et de la MSE à partir de 200 échantillons de taille n (n = 200,500,1000) simulés selon le modèle M1 et pour un niveau de bruit  $\sigma = 1$ . On montre à gauche ces différentes estimations pour l'estimateur de l'amplitude de la rupture  $\hat{\gamma}(\hat{\tau})$  et à droite pour l'estimateur de la localisation  $\hat{\tau}$ . Le noyau utilisé est K1.


**Figure 1.11 :** Estimations du biais au carré, de la variance et de la MSE à partir de 200 échantillons de taille n (n = 200,500,1000) simulés selon le modèle M2 et pour un niveau de bruit  $\sigma = 0.5$ . On montre à gauche ces différentes estimations pour l'estimateur de l'amplitude de la rupture  $\hat{\gamma}(\hat{\tau})$  et à droite pour l'estimateur de la localisation  $\hat{\tau}$ . Le noyau utilisé est K1.



**Figure 1.12 :** Estimations du biais au carré, de la variance et de la MSE à partir de 200 échantillons de taille n (n = 200,500,1000) simulés selon le modèle M2 et pour un niveau de bruit  $\sigma = 1$ . On montre à gauche ces différentes estimations pour l'estimateur de l'amplitude de la rupture  $\hat{\gamma}(\hat{\tau})$  et à droite pour l'estimateur de la localisation  $\hat{\tau}$ . Le noyau utilisé est K1.

# Rapport signal sur bruit (RSB)

Les figures 1.13, 1.14, 1.15 et 1.16 illustrent les gains en biais, variance et MSE des deux estimateurs ( $\hat{\tau}$  et  $\hat{\gamma}(\hat{\tau})$ ) obtenus lorsqu'on augmente le rapport signal sur bruit. Le RSB n'a pas d'influence sur l'ordre de grandeur des vitesses de convergence de ces estimateurs mais les constantes multiplicatives de ces vitesses dépendent de ce rapport. Les simulations montrent clairement qu'en situation non asymptotique, les performances des d'estimations sont fortement dépendantes du rapport signal sur bruit.



**Figure 1.13 :** Estimations du biais au carré, de la variance et de la MSE à partir de 200 échantillons de taille 200 simulés selon le modèle M1. On montre à gauche ces différentes estimations pour l'estimateur de l'amplitude de la rupture  $\hat{\gamma}(\hat{\tau})$  et à droite pour l'estimateur de la localisation  $\hat{\tau}$ .



**Figure 1.14 :** Estimations du biais au carré, de la variance et de la MSE à partir de 200 échantillons de taille 1000 simulés selon le modèle M1. On montre à gauche ces différentes estimations pour l'estimateur de l'amplitude de la rupture  $\hat{\gamma}(\hat{\tau})$  et à droite pour l'estimateur de la localisation  $\hat{\tau}$ .



**Figure 1.15 :** Estimations du biais au carré, de la variance et de la MSE à partir de 200 échantillons de taille 200 simulés selon le modèle M2. On montre à gauche ces différentes estimations pour l'estimateur de l'amplitude de la rupture  $\hat{\gamma}(\hat{\tau})$  et à droite pour l'estimateur de la localisation  $\hat{\tau}$ .



**Figure 1.16 :** Estimations du biais au carré, de la variance et de la MSE à partir de 200 échantillons de taille 1000 simulés selon le modèle M2. On montre à gauche ces différentes estimations pour l'estimateur de l'amplitude de la rupture  $\hat{\gamma}(\hat{\tau})$  et à droite pour l'estimateur de la localisation  $\hat{\tau}$ .

## Fenêtre de lissage

Comme dans le cadre d'estimation d'une fonction de régression lisse, le choix de la fenêtre de lissage  $h_n$  est crucial dans les procédures d'estimations des paramètres de localisation et d'amplitude. Nous avons simulé les modèles M1 et M2 pour deux combinaisons de la taille d'échantillon et de la variance du bruit et étudié la sensibilité des estimateurs  $\hat{\tau}$  et  $\hat{\gamma}(\hat{\tau})$  au choix de la taille de fenêtre. Nous constatons les faits suivants:

- Les simulations montrent que τ̂ est moins sensible au choix de la fenêtre de lissage que γ̂(τ̂). Ceci semble cohérent avec les résultats asymptotiques puisque lorsque le noyau de lissage est strictement positif à l'origine, la vitesse de convergence de τ̂ est indépendante de h<sub>n</sub>.
- 2) L'utilisation de fenêtres de lissage de petites tailles est accompagnée d'une augmentation de la variance du processus \(\hat{\gamma}(t)\), ce qui conduit \(\hat{a}\) une surestimation de l'amplitude de la rupture. Les distributions empiriques de \(\hat{\gamma}(\tau)\) pour des choix de fenêtres de lissage de petites tailles sont d\(\epsilon\) de vers la droite par rapport \(\hat{a}\) 1 (la vraie amplitude de la rupture).
- 3) Le choix de fenêtres de lissage de grandes tailles conduit à une sous-estimation de l'amplitude de la rupture. Les distributions empiriques de  $\hat{\gamma}(\hat{\tau})$ , pour de telles fenêtres, sont décalées vers la gauche par rapport à 1.

Les constatations 2) et 3) sont cohérentes avec l'idée qu'un sous-lissage produit un estimateur très perturbé et peut créer des discontinuités artificielles de forte amplitude alors qu'un surlissage élimine les irrégularités et atténue les discontinuités.



**Figure 1.17 :** Estimations de la MSE à partir de 200 échantillons de taille n simulés selon le modèle M1. On montre à gauche ces différentes estimations pour l'estimateur de l'amplitude de la rupture  $\hat{\gamma}(\hat{\tau})$  et à droite pour l'estimateur de la localisation  $\hat{\tau}$ .



**Figure 1.18 :** Estimations de la MSE à partir de 200 échantillons de taille n simulés selon le modèle M2. On montre à gauche ces différentes estimations pour l'estimateur de l'amplitude de la rupture  $\hat{\gamma}(\hat{\tau})$  et à droite pour l'estimateur de la localisation  $\hat{\tau}$ .



**Figure 1.19 :** La distribution empirique de  $\hat{\tau}$  calculée à partir de 200 échantillons de taille 200 simulés selon le modèle M1 et pour un niveau de bruit  $\sigma = 0.5$ .



**Figure 1.20 :** La distribution empirique de  $\hat{\gamma}(\hat{\tau})$  calculée à partir de 200 échantillons de taille 200 simulés selon le modèle M1 et pour un niveau de bruit  $\sigma = 0.5$ .



**Figure 1.21 :** La distribution empirique de  $\hat{\tau}$  calculée à partir de 200 échantillons de taille 1000 simulés selon le modèle M1 et pour un niveau de bruit  $\sigma = 1$ .



**Figure 1.22 :** La distribution empirique de  $\hat{\gamma}(\hat{\tau})$  calculée à partir de 200 échantillons de taille 1000 simulés selon le modèle M1 et pour un niveau de bruit  $\sigma = 1$ .



**Figure 1.23 :** La distribution empirique de  $\hat{\tau}$  calculée à partir de 200 échantillons de taille 200 simulés selon le modèle M2 et pour un niveau de bruit  $\sigma = 0.5$ .



**Figure 1.24 :** La distribution empirique de  $\hat{\gamma}(\hat{\tau})$  calculée à partir de 200 échantillons de taille 200 simulés selon le modèle M2 et pour un niveau de bruit  $\sigma = 0.5$ .



**Figure 1.25 :** La distribution empirique de  $\hat{\tau}$  calculée à partir de 200 échantillons de taille 1000 simulés selon le modèle M2 et pour un niveau de bruit  $\sigma = 1$ .



**Figure 1.26 :** La distribution empirique de  $\hat{\gamma}(\hat{\tau})$  calculée à partir de 200 échantillons de taille 1000 simulés selon le modèle M2 et pour un niveau de bruit  $\sigma = 1$ .

## 1.6 Annexe A

Proposition 1.1.

a)  $\forall u = \int |M_{\pm}(x+u) - M_{\pm}(x)| dx \le c |u|.$ 

b) Supposons H3-0 satisfaite, alors

$$\forall u \quad \int \left(M_{\pm}(x+u) - M_{\pm}(x)\right)^2 dx \le cu^2.$$

c) Supposons H4-0 satisfaite, alors

$$\forall u < 0 < v \quad \int |M_{\pm}(x+v) - M_{\pm}(x)| |M_{\pm}(x+u) - M_{\pm}(x)| dx \le c |uv|$$

Preuve de la proposition 1.1.



Figure 1.27 : Les supports des fonctions  $M_{+}(x+u)$ ,  $M_{+}(x)$ ,  $M_{+}(x+v)$ ,  $M_{-}(x+u)$ ,  $M_{-}(x)$  et  $M_{-}(x+v)$ , où u < 0 < v.

a) En remarquant que  $\int |M_+(x+u) - M_+(x)| dx = \int |M_+(x-u) - M_+(x)| dx$ , on peut donc supposer pour la démonstration que u < 0. La démonstration pour le noyau  $M_-(\cdot)$  est analogue en tout point à celle pour le noyau  $M_+(\cdot)$  puisque :

$$\int |M_{-}(x+u) - M_{-}(x)| dx = \int |M_{+}(x-u) - M_{+}(x)| dx .$$

$$\int |M_{+}(x+u) - M_{+}(x)| dx \leq \int_{-1}^{-1-u} |M_{+}(x)| dx + \int_{-1-u}^{0} |M_{+}(x+u) - M_{+}(x)| dx$$

$$+ \int_{0}^{-u} |M_{+}(x+u)| dx ,$$

$$\leq c_{1} |u|.$$

#### b) Nous avons

$$\int (M_{+}(x+u) - M_{+}(x))^{2} dx \leq \int_{-1}^{-1-u} (M_{+}(x))^{2} dx + \int_{-1-u}^{0} (M_{+}(x+u) - M_{+}(x))^{2} dx + \int_{0}^{-u} (M_{+}(x+u)) dx,$$
  
 
$$\leq c_{2}|u|.$$

c) Il y a quatre possibilités pour le produit  $|M_{\pm}(x+v) - M_{\pm}(x)||M_{\pm}(x+u) - M_{\pm}(x)|$ :

$$|M_{\pm}(x+v) - M_{\pm}(x)||M_{\pm}(x+u) - M_{\pm}(x)| = \begin{cases} |M_{+}(x+v) - M_{+}(x)||M_{+}(x+u) - M_{+}(x)|, \\ |M_{+}(x+v) - M_{+}(x)||M_{-}(x+u) - M_{-}(x)|, \\ |M_{-}(x+v) - M_{-}(x)||M_{+}(x+u) - M_{+}(x)|, \\ |M_{-}(x+v) - M_{-}(x)||M_{-}(x+u) - M_{-}(x)|. \end{cases}$$

c-1)

$$\begin{split} &\int |M_{+}(x+v) - M_{+}(x)| |M_{+}(x+u) - M_{+}(x)| dx \\ &\leq \int_{-1}^{-1-u} |M_{+}(x+v) - M_{+}(x)| |M_{+}(x)| dx \\ &+ \int_{-1-u}^{-v} |M_{+}(x+v) - M_{+}(x)| |M_{+}(x+u) - M_{+}(x)| dx \\ &+ \int_{-v}^{0} |M_{+}(x)| |M_{+}(x+u) - M_{+}(x)| dx , \\ &\leq c_{3} \int_{-1}^{-1-u} |M_{+}(x+v) - M_{+}(x)| |M_{+}(x+u) - M_{+}(x)| dx \\ &+ \int_{-1-u}^{-v} |M_{+}(x+v) - M_{+}(x)| |M_{+}(x+u) - M_{+}(x)| dx \\ &+ c_{4} \int_{-v}^{0} |M_{+}(x+u) - M_{+}(x)| dx , \\ &\leq c_{5} |uv|. \end{split}$$

c-2) Il est facile de voir que

$$\int |M_{+}(x+v) - M_{+}(x)||M_{-}(x+u) - M_{-}(x)|dx = 0.$$

c-3)

$$\int |M_{-}(x+v) - M_{-}(x)| |M_{+}(x+u) - M_{+}(x)| dx$$

$$\leq \int_{-v}^{0} |M_{-}(x+v)| |M_{+}(x+u) - M_{+}(x)| dx$$

$$+ \int_{0}^{-u} |M_{-}(x+v) - M_{-}(x)| |M_{+}(x+u)| dx ,$$

$$\leq c_{6} \int_{-v}^{0} |M_{+}(x+u) - M_{+}(x)| dx$$

$$+ c_{7} \int_{0}^{-u} |M_{-}(x+v) - M_{-}(x)| dx ,$$

$$\leq c_{8} |uv|.$$

c-4)

$$\begin{split} &\int |M_{-}(x+v) - M_{-}(x)| |M_{-}(x+u) - M_{-}(x)| dx \\ &\leq \int_{0}^{-u} |M_{-}(x+v) - M_{-}(x)| |M_{-}(x)| dx \\ &+ \int_{-u}^{1-v} |M_{-}(x+v) - M_{-}(x)| |M_{-}(x+u) - M_{-}(x)| dx \\ &+ \int_{1-v}^{1} |M_{-}(x+u) - M_{-}(x)| |M_{-}(x)| dx \\ &+ \int_{-u}^{1-v} |M_{-}(x+v) - M_{-}(x)| |M_{-}(x+u) - M_{-}(x)| dx \\ &+ \int_{-u}^{1-v} |M_{-}(x+v) - M_{-}(x)| |M_{-}(x+u) - M_{-}(x)| dx \\ &+ c_{10} \int_{1-v}^{1} |M_{-}(x+u) - M_{-}(x)| dx , \\ &\leq c_{11} |uv|. \end{split}$$

Ceci termine la démonstration de la proposition 1.1.

# Chapitre 2

# Tests de rupture

## 2.1 Introduction

Nous nous intéressons dans ce chapitre aux tests de rupture dans la fonction  $m(\cdot)$  du modèle de régression qui fait l'objet de ce travail. Rappelons que dans ce modèle les observations  $(X_i, Y_i)$ i = 1, ..., n sont i.i.d. et satisfont

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

Trois types de tests sont présentés dans ce chapitre.

Le premier, qui peut être qualifié de strictement local, a pour but de décider s'il existe ou pas un saut de la fonction de régression en un point  $\tau$  fixé. Il est basé sur un estimateur de l'amplitude du saut éventuel au point  $\tau$ . Sous l'hypothèse nulle d'absence de discontinuité, cet estimateur est asymptotiquement gaussien, ce qui conduit à un test naturel de rupture au point  $\tau$ . Il s'agit ici d'exploiter certains résultats obtenus au chapitre précédent, mais aussi de les étendre puisque nous considérons la situation d'une suite de modèles où l'amplitude de la rupture décroît vers 0 avec la taille de l'échantillon. Nous montrons que notre test est asymptotiquement de puissance unité.

Un deuxième test local est basé sur l'étude de l'estimateur de saut dans le voisinage du point  $\tau$ . En l'absence de saut le processus  $\mathcal{Z}_n(z) = \sqrt{nh_n}\hat{\gamma}(\tau + h_n z)$  converge vers un processus gaussien. Le supremum de  $\mathcal{Z}_n(z)$  convenablement normalisé converge alors vers une loi de Gumbel et nous basons notre test sur ce résultat. Comme le test strictement local il est asymptotiquement de puissance unité.

Le dernier test que nous proposons a pour but de décider s'il existe un point de rupture sur ]0, 1[. Nous utilisons pour le construire le processus défini en tout t, à une normalisation près, par

la valeur de l'estimateur du saut en t,  $\xi_n(t) = \sqrt{nh_n}V(t)\hat{\gamma}(t)$  où  $V(t) = (f_X(t)/\sigma^2(t))^{1/2}/\sqrt{2}V_+$ . Le processus  $\xi_n(t)$  peut être approché par une suite de processus  $(\zeta_n(t))_n$  convergeant vers un processus gaussien. On en déduit des résultats de convergence vers des lois de Gumbel pour le supremum de  $\xi_n(t)$  ou de  $|\xi_n(t)|$  convenablement normalisés. C'est sur ce résultat qu'est basé notre test d'existence d'une rupture. Il est à remarquer ici que, comme on pouvait s'y attendre, les deux situations  $K_+(0) = 0$  et  $K_+(0) > 0$  conduisent à un traitement différent.

## 2.2 Test Strictement Local

Soit  $\tau \in ]0,1[$ . Étant donné les observations  $(X_1,Y_1),\ldots,(X_n,Y_n)$ , nous souhaitons tester l'hypothèse nulle

HO:  $m_+(\tau) = m_-(\tau)$ ,

l'hypothèse alternative étant

H1:  $m_{+}(\tau) \neq m_{-}(\tau)$  (il y a un point de rupture en  $\tau$  dans la fonction de régression  $m(\cdot)$ ).

Nous introduisons la quantité  $\hat{\gamma}(\tau)$  définie par :

$$\hat{\gamma}(\tau) = \hat{m}_{+}(\tau) - \hat{m}_{-}(\tau).$$
(2.1)

Cette statistique est basée sur l'estimateur de la régression linéaire locale,  $\hat{\gamma}(\tau)$ , du saut éventuel de la fonction de régression au point  $\tau$ . Clairement, sous l'hypothèse nulle HO, la valeur de cette statistique, à une normalisation près, doit être faible. Dans le cas contraire, sous H1, elle doit être d'autant plus grande que l'amplitude de la discontinuité,  $\gamma(\tau)$ , est grande. Pour pouvoir construire un test sur la base de cette statistique, nous établissons la loi asymptotique sous HO de sa version normalisée  $\sqrt{nh_n}\hat{\gamma}(\tau)$ . Nous rappelons tout d'abord quelques notations:

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

ii) 
$$V_{+}^{2} = \int_{-1}^{0} \left(K_{2}^{+} - xK_{1}^{+}\right)^{2} K_{+}^{2}(x) dx$$
,

iii)  $K_{-}(x) = K_{+}(-x),$ 

iv) 
$$M_+(x) = (K_2^+ - xK_1^+)K_+(x)$$
 et  $M(x) = M_+(x) - M_+(-x)$ ,

où  $K_{+}(\cdot)$  est une fonction a support [-1,0]. Nous supposons de plus que  $K_{+}(\cdot)$  est normalisé de telle sorte que  $|K_{2}^{+}K_{0}^{+} - (K_{1}^{+})^{2}| = 1$ .

#### 2.2.1 Hypothèses

- B1. La fonction  $m(\cdot)$  ainsi que ses dérivées première et seconde sont supposées càdlàg en  $\tau$  (continue à droite et limitée à gauche).
- B2.  $\sigma^2(x) = Var(Y \mid X = x)$  est supposée continue en  $\tau$ .
- B3. La densité marginale de X,  $f_X(\cdot)$ , est supposée continue et strictement positive en  $\tau$ .
- B4.  $(h_n)_{n\geq 1}$  est une suite réelle positive vérifiant:  $\lim_{n\to+\infty} h_n = 0$ ,  $\lim_{n\to+\infty} nh_n = +\infty$  et  $\lim_{n\to+\infty} nh_n^5 = 0$ .
- B5. On suppose qu'il existe un  $\delta > 0$  tel que  $\mathbb{E}(|\varepsilon|^{2+\delta}) < +\infty$ .

#### 2.2.2 Résultats asymptotiques

Le théorème qui suit est une généralisation du théorème de normalité asymptotique de  $\hat{\gamma}(\tau)$ , (cf le corollaire 1.2), au sens où l'amplitude de la rupture au point  $\tau$  est autorisée à dépendre du nombre d'observations n. Nous supposons que pour tout n, la fonction de régression s'écrit :

$$m_n(x) = c(x) + \gamma_n(\tau) \operatorname{II}_{[\tau,1]}(x),$$

où  $c(\cdot)$  est une fonction continue au point  $\tau$  qui ne dépend pas de n, et nous notons

$$\gamma_n(\tau) = m_{n,+}(\tau) - m_{n,-}(\tau).$$
(2.2)

Théorème 2.1. Supposons que les hypothèses B1-B5 soient vérifiées, alors :

$$\sqrt{nh_n}\left(\hat{\gamma}_n(\tau) - \gamma_n(\tau)\right) \xrightarrow{\mathcal{L}} \mathcal{N}\left(0, \left(\frac{2\sigma^2(\tau)}{f(\tau)}V_+^2\right)\right).$$
(2.3)

#### Preuve du théorème 2.1.

L'estimateur du saut admet la décomposition suivante :

$$\begin{split} \hat{\gamma}_{n}(\tau) &= \hat{m}_{n,+}(\tau) - \hat{m}_{n,-}(\tau), \\ &= \frac{\sum_{i=1}^{n} w_{i}^{+}(\tau) y_{i}}{\sum_{i=1}^{n} w_{i}^{+}(\tau)} - \frac{\sum_{i=1}^{n} w_{i}^{-}(\tau) y_{i}}{\sum_{i=1}^{n} w_{i}^{-}(\tau)}, \\ &= \frac{\sum_{i=1}^{n} w_{i}^{+}(\tau) c(X_{i})}{\sum_{i=1}^{n} w_{i}^{+}(\tau)} - \frac{\sum_{i=1}^{n} w_{i}^{-}(\tau) c(X_{i})}{\sum_{i=1}^{n} w_{i}^{-}(\tau)} \\ &+ \frac{\sum_{i=1}^{n} w_{i}^{+}(\tau) \gamma_{n}(\tau) \mathbb{1}_{[\tau,1]}(X_{i})}{\sum_{i=1}^{n} w_{i}^{+}(\tau)} - \frac{\sum_{i=1}^{n} w_{i}^{-}(\tau) \gamma_{n}(\tau) \mathbb{1}_{[\tau,1]}(X_{i})}{\sum_{i=1}^{n} w_{i}^{-}(\tau)} \\ &+ \frac{\sum_{i=1}^{n} w_{i}^{+}(\tau) \sigma(X_{i}) \varepsilon_{i}}{\sum_{i=1}^{n} w_{i}^{-}(\tau)} - \frac{\sum_{i=1}^{n} w_{i}^{-}(\tau) \sigma(X_{i}) \varepsilon_{i}}{\sum_{i=1}^{n} w_{i}^{-}(\tau)}, \\ &= (C) + (D) + (B), \end{split}$$

d'où

$$\hat{\gamma}_n(\tau) - \mathbb{E}\left[\hat{\gamma}_n(\tau) \mid X_1, \dots, X_n\right] = (B).$$

Le comportement de (B) est indépendant de  $\gamma_n(\tau)$  et la démonstration de sa normalité asymptotique est une modification immédiate de celle du corollaire 1.2 (voir la démonstration dans le chapitre 1). Il s'ensuit

$$\sqrt{nh_n}(B) \xrightarrow{\mathcal{L}} \mathcal{N}\left(0, \left(\frac{2\sigma^2(\tau)}{f(\tau)}V_+^2\right)\right).$$
(2.4)

Nous avons

$$\operatorname{IE}\left[\hat{\gamma}_n(\tau) \mid X_1, \dots, X_n\right] = (C) + (D).$$

Compte tenu de la définition des  $w_i^{\pm}(\tau)$ , il vient que pour tout  $i = 1, \ldots, n$ ,

$$w_i^+(\tau) \coprod_{[\tau,1]}(X_i) = w_i^+(\tau),$$

 $\operatorname{et}$ 

$$w_i^-(\tau) 1\!\!1_{[\tau,1]}(X_i) = 0,$$

ce qui entraîne:

$$\mathbb{E}\left[\hat{\gamma}_n(\tau) \mid X_1, \dots, X_n\right] - \gamma_n(\tau) = (C).$$

Or, d'après l'étude de l'erreur quadratique moyenne de  $\hat{\gamma}(\tau)$  faite au chapitre 1 (cf le théorème 1.1), nous avons

$$(C) = \frac{\sum_{i=1}^{n} w_{i}^{+}(\tau) c(X_{i})}{\sum_{i=1}^{n} w_{i}^{+}(\tau)} - \frac{\sum_{i=1}^{n} w_{i}^{-}(\tau) c(X_{i})}{\sum_{i=1}^{n} w_{i}^{-}(\tau)},$$
  
$$= \frac{\sum_{i=1}^{n} w_{i}^{+}(\tau) (c(X_{i}) - c_{+}(\tau))}{\sum_{i=1}^{n} w_{i}^{+}(\tau)} - \frac{\sum_{i=1}^{n} w_{i}^{-}(\tau) (c(X_{i}) - c_{-}(\tau))}{\sum_{i=1}^{n} w_{i}^{-}(\tau)},$$
  
$$= \mathcal{O}_{P}(h_{n}^{2}).$$

Il en découle que, si nous supposons que  $nh_n^5 \to 0$ , alors

$$\sqrt{nh_n} \left( \mathbb{E} \left[ \hat{\gamma}_n(\tau) \mid X_1, \dots, X_n \right] - \gamma_n(\tau) \right) \xrightarrow{\text{IP}} 0.$$
(2.5)

Pour finir, il suffit de combiner (2.4) et (2.5) et nous obtenons

$$\sqrt{nh_n}\left(\hat{\gamma}_n(\tau) - \gamma_n(\tau)\right) \xrightarrow{\mathcal{L}} \mathcal{N}\left(0, \left(\frac{2\sigma^2(\tau)}{f(\tau)}V_+^2\right)\right).$$

Ce qui achève la démonstration du théorème.

Sous l'hypothèse nulle, HO,  $\gamma(\tau) = 0$ , nous avons immédiatement :

**Corollaire 2.1.** Supposons que les hypothèses du théorème 2.1 soient satisfaites, alors sous l'hypothèse nulle HO

$$\sqrt{nh_n}\hat{\gamma}_n(\tau) \xrightarrow{\mathcal{L}} \mathcal{N}\left(0, \left(\frac{2\sigma^2(\tau)}{f(\tau)}V_+^2\right)\right).$$
(2.6)

Sous l'hypothèse nulle HO, la loi limite de la statistique définie dans (2.1) dépend des quantités  $f_X(\tau)$  et  $\sigma^2(\tau)$  qui sont souvent, en pratique, inconnues. Pour que le test de rupture basé sur cette statistique soit opérationnel, il faut avoir un résultat analogue à celui du corollaire 2.1 avec  $f_X(\tau)$  et  $\sigma^2(\tau)$  remplacées par leurs estimateurs. Comme estimateur consistant de  $f_X(\tau)$ , on peut utiliser un estimateur à noyau de la fonction de densité:

$$\hat{f}_X(\tau) = \frac{1}{nh_n} S_0(\tau),$$

où  $S_0(\tau)$  est définie par (1.1). Avec quelques modifications, l'estimateur proposé par Müller et Stadtmüller [MS87] peut être utilisé pour estimer  $\sigma^2(\tau)$ .

**Corollaire 2.2.** Soient  $\hat{f}_X(\tau)$  et  $\hat{\sigma}^2(\tau)$  deux estimateurs consistants respectivement de  $f_X(\tau)$  et  $\sigma^2(\tau)$ . L' estimateur  $\hat{f}_X(\tau)$  est supposé positif. Supposons les conditions du corollaire 2.1 satisfaites. Sous l'hypothèse nulle, HO, :

$$\sqrt{nh_n} \frac{\sqrt{\hat{f}_X(\tau)} \hat{\gamma}_n(\tau)}{\hat{\sigma}(\tau)} \xrightarrow{\mathcal{L}} \mathcal{N}\left(0, 2V_+^2\right).$$
(2.7)

**Corollaire 2.3.** Supposons les conditions du corollaire 2.2 satisfaites. Sous l'hypothèse nulle, HO, :

$$\lim_{n \to +\infty} \operatorname{IP}_{HO} \left[ \sqrt{nh_n} | \frac{\sqrt{\hat{f}_X(\tau)} \hat{\gamma}_n(\tau)}{\sqrt{2} V_+ \hat{\sigma}(\tau)} | \ge u_\alpha \right] = \alpha,$$
(2.8)

avec  $u_{\alpha}$  est le  $\alpha$ -quantile de la loi normale centrée réduite.

Nous considérons maintenant la puissance du test pour une suite d'alternatives locales. De manière précise, soit la suite de contre-hypothèses définies par

$$\operatorname{H1}(n) : \gamma_n(\tau) = \frac{\delta_n}{\sqrt{nh_n}},$$

avec  $|\delta_n| \to \infty$ . Le théorème suivant montre que le test basé sur  $\hat{\gamma}_n(\tau)$  a une puissance locale asymptotique égale à 1.

**Théorème 2.2.** Soient  $(\delta_n)$  une suite réelle telle que  $|\delta_n| \to \infty$  et  $\frac{|\delta_n|}{\sqrt{nh_n}} \to 0$ , et  $(\mathtt{H1}(n))$  la suite des alternatives locales associées définies par  $\gamma_n(\tau) = \frac{\delta_n}{\sqrt{nh_n}}$ . Pour tout u > 0:

$$\lim_{n \to +\infty} \mathbb{P}_{\mathbf{H1}(n)} \left[ \sqrt{nh_n} \left| \frac{\sqrt{\hat{f}_X(\tau)} \hat{\gamma}_n(\tau)}{\sqrt{2}V_+ \hat{\sigma}(\tau)} \right| \ge u \right] = 1.$$
(2.9)

En particulier, pour toute suite d'alternatives où  $\gamma_n(\tau)$  tend vers 0 moins vite que  $1/\sqrt{nh_n}$ , la puissance locale tend vers 1.

#### Preuve du théorème 2.2.

Pour tout u > 0, notons

$$E_n^- = \left[\sqrt{nh_n}\left(\hat{\gamma}_n(\tau) - \gamma_n(\tau)\right) \le \frac{\sqrt{2}V_+\hat{\sigma}(\tau)}{\sqrt{\hat{f}_X(\tau)}}u - \delta_n\right],$$

 $\operatorname{et}$ 

$$E_n^+ = \left[\sqrt{nh_n}\left(\hat{\gamma}_n(\tau) - \gamma_n(\tau)\right) \ge -\frac{\sqrt{2}V_+\hat{\sigma}(\tau)}{\sqrt{\hat{f}_X(\tau)}}u - \delta_n\right].$$

Nous avons alors

$$\lim_{n \to +\infty} \operatorname{IP}_{\mathtt{H1}(n)} \left[ \sqrt{nh_n} \left| \frac{\sqrt{\hat{f}_X(\tau)} \hat{\gamma}_n(\tau)}{\sqrt{2} V_+ \hat{\sigma}(\tau)} \right| \le u \right] = \lim_{n \to +\infty} \operatorname{IP}_{\mathtt{H1}(n)} \left[ E_n^- \cap E_n^+ \right],$$
$$= 0,$$

d'où

$$\lim_{n \to +\infty} \operatorname{IP}_{\mathtt{H1}(n)} \left[ \sqrt{nh_n} \left| \frac{\sqrt{\hat{f}_X(\tau)} \hat{\gamma}_n(\tau)}{\sqrt{2}V_+ \hat{\sigma}(\tau)} \right| \ge u \right] = 1.$$

	-	
	. 1	
	. 1	

## 2.3 Test Local

Le test que nous avons construit au paragraphe précédent est défini en terme de la statistique  $\hat{\gamma}_n(\tau)$ , estimateur du saut éventuel au point  $\tau$ . Nous proposons maintenant un test basé sur la valeur de la statistique  $\hat{\gamma}_n(t)$  dans un voisinage de  $\tau$ . Le problème de test est identique à celui du paragraphe précédent, à savoir : tester l'hypothèse

HO: 
$$m_+(\tau) = m_-(\tau)$$
,

contre

H1:  $m_{+}(\tau) \neq m_{-}(\tau)$  (il y a un point de rupture en  $\tau$  dans la fonction de régression  $m(\cdot)$ ).

Nous introduisons le processus de variation locale  $\mathcal{Z}_n$  défini, pour  $z \in [-M, M]$ , M étant une constante positive quelconque, par:

$$\mathcal{Z}_n(z) = \sqrt{nh_n} \left( \hat{m}_+(\tau + h_n z) - \hat{m}_-(\tau + h_n z) \right) = \sqrt{nh_n} \hat{\gamma}(\tau + h_n z).$$
(2.10)

Dans certaines des énoncés de ce paragraphe B4 et B5 seront remplacés par B4' et B5' :

- B4'.  $(h_n)_{n\geq 1}$  est une suite réelle positive vérifiant:  $\lim_{n\to +\infty} h_n = 0$ ,  $\lim_{n\to +\infty} nh_n = +\infty$  et  $\lim_{n\to +\infty} nh_n^3 = 0$ .
- B5'. On suppose que  $\mathbb{E}(|\varepsilon|^4) < +\infty$ .

Nous supposons également l'une ou l'autre des deux hypothèses ci-dessous vérifiée :

B6-1. 
$$K_{+}(\cdot) \in \mathcal{C}^{1}[-1, 0]$$
 avec  $K_{+}(0) = K_{+}(-1) = 0$ ,  
B6-2.  $K_{+}(\cdot) \in \mathcal{C}^{1}[-1, 0]$  avec  $K_{+}(0) \neq 0$  ou  $K_{+}(-1) \neq 0$ .

#### 2.3.1 Résultats asymptotiques

Le résultat clé sur lequel est basé ce test, est le suivant :

**Théorème 2.3.** Supposons que les hypothèses B1-B3, B4' et B5' soient satisfaites. Sous l'hypothèse nulle H0,  $(\mathcal{Z}_n)$  converge faiblement vers un processus gaussien centré stationnaire  $(\mathcal{Z})$  avec pour fonction de covariance:

$$Cov\left(\mathcal{Z}(z_1), \mathcal{Z}(z_2)\right) = \frac{\sigma^2(\tau)}{f_X(\tau)} \left(\int M(x+z_1-z_2)M(x)dx\right).$$
(2.11)

#### Preuve du théorème 2.3.

La démonstration de la convergence faible du processus  $\mathcal{Z}_n$  est décomposée en trois parties :

1) Démonstration de la convergence cylindrique du processus centré  $\overline{\mathcal{Z}}_n$ ,

$$\overline{\mathcal{Z}}_n = \mathcal{Z}_n - \mathbb{E}\left[\mathcal{Z}_n | X_1, \dots, X_n\right], \qquad (2.12)$$

vers un processus gaussien stationnaire centré de fonction de covariance définie par (2.11).

- 2) Vérification du critère de tension pour le processus  $\overline{\mathcal{Z}}_n$ .
- 3) Démonstration, sous l'hypothèse HO, de la convergence en probabilité de  $\mathbb{E} \left[ \mathcal{Z}_n | X_1, \dots, X_n \right]$  vers zéro.

Les étapes 1 et 2 montrent la convergence faible du processus  $\overline{\mathcal{Z}}_n$  (cf Billingsley [BIL68]).

1. Convergence cylindrique. Pour montrer la convergence cylindrique du processus  $\overline{\mathcal{Z}}_n$  il suffit (théorème de Cramer-Wold) de montrer la convergence en distribution de toute combinaison linéaire de variables d'état du processus. Pour des raisons de simplicité, nous monterons ici la convergence cylindrique seulement pour la dimension deux. La transposition à une dimension supérieure est immédiate. Une étape préliminaire à la démonstration de la convergence cylindrique du processus  $\overline{\mathcal{Z}}_n$ , consiste à montrer qu'il est, à un facteur près, asymptotiquement équivalent à un processus  $\widetilde{\mathcal{Z}}_n$ , qui s'écrit comme la somme des termes d'une ligne d'un tableau triangulaire, où tous les termes d'une même ligne sont  $\mathbf{i} \cdot \mathbf{i} \cdot \mathbf{d}$ . C'est ce qu'énonce la proposition suivante :

Proposition 2.1. Sous les hypothèses B1-B5, nous avons

$$\overline{\mathcal{Z}}_n(z) = \frac{(1+o_P(1))}{f_X(\tau)} \tilde{\mathcal{Z}}_n(z) + o_P(1),$$

avec

$$\tilde{\mathcal{Z}}_n(z) = \frac{1}{\sqrt{nh_n}} \left[ \sum_{i=1}^n M(\frac{\tau + h_n z - X_i}{h_n}) \sigma(X_i) \varepsilon_i \right].$$

#### Preuve de la proposition 2.1.

Remarquons d'abord que nous avons

$$\overline{\mathcal{Z}}_n(z) = \overline{\mathcal{Z}}_n^+(z) - \overline{\mathcal{Z}}_n^-(z),$$

оù

$$\overline{\mathcal{Z}}_n^{\pm}(z) = \mathcal{Z}_n^{\pm}(z) - \mathbb{E}\left[\mathcal{Z}_n^{\pm}(z) | X_1, \dots, X_n\right],$$

 $\operatorname{et}$ 

$$\mathcal{Z}_n^{\pm}(z) = \sqrt{nh_n} \hat{m}_{\pm}(\tau + zh_n).$$

D'où

$$\overline{Z}_{n}(z) = \sqrt{nh_{n}} \left[ \frac{\sum_{i=1}^{n} w_{i}^{+}(\tau + h_{n}z)\sigma(X_{i})\varepsilon_{i}}{\sum_{i=1}^{n} w_{i}^{+}(\tau + h_{n}z)} - \frac{\sum_{i=1}^{n} w_{i}^{-}(\tau + h_{n}z)\sigma(X_{i})\varepsilon_{i}}{\sum_{i=1}^{n} w_{i}^{-}(\tau + h_{n}z)} \right],$$

$$= \sqrt{nh_{n}} \left( \frac{\sum_{i=1}^{n} (w_{i}^{+}(\tau + h_{n}z) - w_{i}^{-}(\tau + h_{n}z))\sigma(X_{i})\varepsilon_{i}}{\sum_{i=1}^{n} w_{i}^{+}(\tau + h_{n}z)} \right)$$

$$- \sqrt{nh_{n}} \left( \frac{\sum_{i=1}^{n} w_{i}^{-}(\tau + h_{n}z)\sigma(X_{i})\varepsilon_{i}}{\sum_{i=1}^{n} w_{i}^{-}(\tau + h_{n}z)} \right) \left( \frac{\sum_{i=1}^{n} w_{i}^{+}(\tau + h_{n}z) - w_{i}^{-}(\tau + h_{n}z)}{\sum_{i=1}^{n} w_{i}^{-}(\tau + h_{n}z)} \right),$$

$$= A_{n} - B_{n} \times C_{n}.$$
(2.13)

Le produit  $B_n \times C_n$  est asymptotiquement négligeable par rapport à  $A_n$ , et  $A_n$  se comporte comme  $\frac{(1+o_P(1))}{f_X(\tau)}\tilde{\mathcal{Z}}_n(z)$ . En effet

$$\begin{split} w_i^+(\tau + h_n z) &- w_i^-(\tau + h_n z) \\ &= \left[ S_2^+(\tau + h_n z) - S_2^-(\tau + h_n z) \right] K_+(\frac{\tau + h_n z - X_i}{h_n}) \\ &- \left[ S_1^+(\tau + h_n z) + S_1^-(\tau + h_n z) \right] (\tau + h_n z - X_i) K_+(\frac{\tau + h_n z - X_i}{h_n}) \\ &+ S_2^-(\tau + h_n z) \left[ K_+(\frac{\tau + h_n z - X_i}{h_n}) - K_-(\frac{\tau + h_n z - X_i}{h_n}) \right] \\ &+ S_1^-(\tau + h_n z) (\tau + h_n z - X_i) \left[ K_+(\frac{\tau + h_n z - X_i}{h_n}) + K_-(\frac{\tau + h_n z - X_i}{h_n}) \right]. \end{split}$$

Les numérateurs de  $C_n$  et  $B_n$  s'écrivent donc

$$\sum_{i=1}^{n} w_{i}^{+}(\tau + h_{n}z) - w_{i}^{-}(\tau + h_{n}z)$$

$$= \left[S_{2}^{+}(\tau + h_{n}z) - S_{2}^{-}(\tau + h_{n}z)\right] S_{0}^{+}(\tau + h_{n}z)$$

$$- \left[S_{1}^{+}(\tau + h_{n}z) + S_{1}^{-}(\tau + h_{n}z)\right] S_{1}^{+}(\tau + h_{n}z)$$

$$+ \left[S_{0}^{+}(\tau + h_{n}z) - S_{0}^{-}(\tau + h_{n}z)\right] S_{2}^{-}(\tau + h_{n}z)$$

$$+ \left[S_{1}^{+}(\tau + h_{n}z) + S_{1}^{-}(\tau + h_{n}z)\right] S_{1}^{-}(\tau + h_{n}z), \qquad (2.14)$$

 $\operatorname{et}$ 

$$\sum_{i=1}^{n} (w_i^+(\tau + h_n z) - w_i^-(\tau + h_n z))\sigma(X_i)\varepsilon_i$$
  
=  $[S_2^+(\tau + h_n z) - S_2^-(\tau + h_n z)]\xi_0^+(\tau + h_n z)$   
-  $[S_1^+(\tau + h_n z) + S_1^-(\tau + h_n z)]\xi_1^+(\tau + h_n z)$   
+  $[\xi_0^+(\tau + h_n z) - \xi_0^-(\tau + h_n z)]S_2^-(\tau + h_n z)$   
+  $[\xi_1^+(\tau + h_n z) + \xi_1^-(\tau + h_n z)]S_1^-(\tau + h_n z),$  (2.15)

avec

$$\xi_l^{\pm}(x) = \sum_{i=1}^n (x - X_i)^l K_{\pm}(\frac{x - X_i}{h_n}) \sigma(X_i) \varepsilon_i.$$

La fin de la démonstration de la proposition est une conséquence immédiate des résultats énoncés dans le lemme suivant :

Lemme 2.1. Sous les hypothèses B1-B5, nous avons:

a) 
$$\sum_{i=1}^{n} w_{i}^{\pm}(\tau + h_{n}z) = n^{2}h_{n}^{4}f_{X}^{2}(\tau)(K_{2}^{+}K_{0}^{+} - (K_{1}^{+})^{2} + o_{p}(1)),$$
  
b) 
$$\frac{\sum_{i=1}^{n} w_{i}^{+}(\tau + h_{n}z) - w_{i}^{-}(\tau + h_{n}z)}{\sum_{i=1}^{n} w_{i}^{+}(\tau + h_{n}z)} = o_{P}(1),$$

c) 
$$\frac{\sum_{i=1}^{n} w_{i}^{i} (\tau + h_{n}z) \sigma(X_{i}) \varepsilon_{i}}{\sum_{i=1}^{n} w_{i}^{\pm} (\tau + h_{n}z)} = \mathcal{O}_{P} \left(\frac{1}{\sqrt{nh_{n}}}\right),$$
  
d) 
$$\frac{\sum_{i=1}^{n} (w_{i}^{+} (\tau + h_{n}z) - w_{i}^{-} (\tau + h_{n}z)) \sigma(X_{i}) \varepsilon_{i}}{\sum_{i=1}^{n} w_{i}^{+} (\tau + h_{n}z)} = \frac{(1 + o_{P}(1))}{\sqrt{nh_{n}f_{X}(\tau)}} \tilde{\mathcal{Z}}_{n}(z) + o_{P} \left(\frac{1}{\sqrt{nh_{n}}}\right).$$

Pour démontrer ce lemme, nous devons contrôler le comportement des termes  $S_l^{\pm}(\tau + h_n z)$ et celui des différences  $S_l^{+}(\tau + h_n z) - (-1)^l S_l^{-}(\tau + h_n z)$ . Rappelons que

$$S_l^{\pm}(\tau + h_n z) = n h_n^{l+1} f_X(\tau) (K_l^{\pm} + o_p(1)).$$
(2.16)

Le comportement des différences et celui des termes  $\xi_l^{\pm}(\tau + h_n z)$  est précisé dans le lemme suivant.

Lemme 2.2. Sous les hypothèses B1-B5, nous avons

a) 
$$S_l^+(\tau + h_n z) - (-1)^l S_l^-(\tau + h_n z) = o_P(nh_n^{l+1}),$$
  
b)  $\xi_l^{\pm}(\tau + h_n z) = \mathcal{O}_P\left(\sqrt{nh_n^{2l+1}}\right).$ 

### Preuve du lemme 2.2.

a) Remarquant que  $S_l^{\pm}(\tau + h_n z)$  s'écrit sous la forme  $\sum_{i=1}^n Z_{i,1}^{\pm}$  avec

$$Z_{i,1}^{\pm} = (\tau + h_n z - X_i)^l K_{\pm} (\frac{\tau + h_n z - X_i}{h_n}),$$

il vient:

$$\begin{split} & \operatorname{I\!E}\left[S_{l}^{+}(\tau+h_{n}z)-(-1)^{l}S_{l}^{-}(\tau+h_{n}z)\right] \\ &= nh_{n}^{l+1}\int x^{l}K_{+}(x)\left[f_{X}(\tau+h_{n}(z-x))-f_{X}(\tau+h_{n}(z+x))\right]dx \,, \\ &= o\left(nh_{n}^{l+1}\right), \end{split}$$

 $\operatorname{et}$ 

$$Var \left[ S_{l}^{+}(\tau + h_{n}z) - (-1)^{l}S_{l}^{-}(\tau + h_{n}z) \right]$$
  

$$\leq nh_{n}^{2l+1} \int \left( x^{l}K_{+}(x) - x^{l}K_{-}(x) \right)^{2} f_{X}(\tau + h_{n}(z - x)) dx ,$$
  

$$= \mathcal{O} \left( nh_{n}^{2l+1} \right) ,$$

 $\operatorname{donc}$ 

$$S_{l}^{+}(\tau + h_{n}z) - (-1)^{l}S_{l}^{-}(\tau + h_{n}z) = o\left(nh_{n}^{l+1}\right) + \mathcal{O}_{P}\left(\sqrt{nh_{n}^{2l+1}}\right),$$
$$= o_{P}\left(nh_{n}^{l+1}\right).$$

b) Comme

$$\operatorname{IE}\left[\xi_{l}^{\pm}(\tau+h_{n}z)\right] = 0,$$

 $\operatorname{et}$ 

$$\begin{aligned} Var\left[\xi_{l}^{\pm}(\tau+h_{n}z)\right] &\leq nh_{n}^{2l+1}\int x^{2l}K_{\pm}^{2}(x)\sigma^{2}(\tau+h_{n}(z-x))f_{X}(\tau+h_{n}(z-x))dx ,\\ &= \mathcal{O}\left(nh_{n}^{2l+1}\right), \end{aligned}$$

alors

$$\xi_l^{\pm}(\tau + h_n z) = \mathcal{O}_P\left(\sqrt{nh_n^{2l+1}}\right).$$

_	
- L	_

### Preuve du lemme 2.1.

a) Puisque  $\sum_{i=1}^{n} w_i^{\pm}(\tau + h_n z) = S_2^{\pm}(\tau + h_n z)S_0^{\pm}(\tau + h_n z) - (S_1^{\pm}(\tau + h_n z))^2$ , la partie a) du lemme 2.1 peut être déduite facilement de (2.16).

b) Les parties b) et c) du lemme 2.1 sont des conséquences immédiates de (2.14), (2.15), des parties a) et b) du lemme 2.2 et de la partie a) du lemme 2.1.

En effet nous avons

$$\sum_{i=1}^{n} w_{i}^{+}(\tau + h_{n}z) - w_{i}^{-}(\tau + h_{n}z) = \left[S_{2}^{+}(\tau + h_{n}z) - S_{2}^{-}(\tau + h_{n}z)\right]S_{0}^{+}(\tau + h_{n}z) - \left[S_{1}^{+}(\tau + h_{n}z) + S_{1}^{-}(\tau + h_{n}z)\right]S_{1}^{+}(\tau + h_{n}z) + \left[S_{0}^{+}(\tau + h_{n}z) - S_{0}^{-}(\tau + h_{n}z)\right]S_{2}^{-}(\tau + h_{n}z) + \left[S_{1}^{+}(\tau + h_{n}z) + S_{1}^{-}(\tau + h_{n}z)\right]S_{1}^{-}(\tau + h_{n}z) = o_{P}\left(n^{2}h_{n}^{4}\right),$$

et

$$\sum_{i=1}^{n} w_i^+(\tau + h_n z) = n^2 h_n^4 f_X^2(\tau) (K_2^+ K_0^+ - (K_1^+)^2 + o_p(1)),$$

 $\operatorname{donc}$ 

$$\frac{\sum_{i=1}^{n} w_i^+(\tau + h_n z) - w_i^-(\tau + h_n z)}{\sum_{i=1}^{n} w_i^+(\tau + h_n z)} = o_P(1) .$$

c)

$$\sum_{i=1}^{n} w_i^{\pm}(\tau + h_n z) \sigma(X_i) \varepsilon_i = S_2^{\pm}(\tau + h_n z) \xi_0^{\pm}(\tau + h_n z) - S_1^{\pm}(\tau + h_n z) \xi_1^{\pm}(\tau + h_n z),$$
$$= \mathcal{O}_P\left(nh_n^3\right) \mathcal{O}_P\left(\sqrt{nh_n}\right),$$
$$= \mathcal{O}_P\left(n\frac{3}{2}h_n^{\frac{7}{2}}\right).$$

d'où

$$\frac{\sum_{i=1}^{n} w_i^{\pm}(\tau + h_n z) \sigma(X_i) \varepsilon_i}{\sum_{i=1}^{n} w_i^{\pm}(\tau + h_n z)} = \mathcal{O}_P\left(\frac{1}{\sqrt{nh_n}}\right).$$

d)

$$\begin{split} &\sum_{i=1}^{n} (w_{i}^{+}(\tau+h_{n}z)-w_{i}^{-}(\tau+h_{n}z))\sigma(X_{i})\varepsilon_{i} \\ &= \left[S_{2}^{+}(\tau+h_{n}z)-S_{2}^{-}(\tau+h_{n}z)\right]\xi_{0}^{+}(\tau+h_{n}z) \\ &- \left[S_{1}^{+}(\tau+h_{n}z)+S_{1}^{-}(\tau+h_{n}z)\right]\xi_{1}^{+}(\tau+h_{n}z) \\ &+ \left[\xi_{0}^{+}(\tau+h_{n}z)-\xi_{0}^{-}(\tau+h_{n}z)\right]S_{2}^{-}(\tau+h_{n}z) \\ &+ \left[\xi_{1}^{+}(\tau+h_{n}z)+\xi_{1}^{-}(\tau+h_{n}z)\right]S_{1}^{-}(\tau+h_{n}z), \\ &= o_{P}\left(nh_{n}^{3}\right)\mathcal{O}_{P}\left(\sqrt{nh_{n}}\right) \\ &+ \left[\xi_{0}^{+}(\tau+h_{n}z)-\xi_{0}^{-}(\tau+h_{n}z)\right]nh_{n}^{3}f_{X}(\tau)(K_{2}^{-}+o_{p}(1)) \\ &+ \left[\xi_{1}^{+}(\tau+h_{n}z)+\xi_{1}^{-}(\tau+h_{n}z)\right]nh_{n}^{2}f_{X}(\tau)(K_{1}^{-}+o_{p}(1)), \\ &= nh_{n}^{3}f_{X}(\tau)\left(1+o_{p}(1)\right)\left[K_{2}^{-}\left(\xi_{0}^{+}(\tau+h_{n}z)-\xi_{0}^{-}(\tau+h_{n}z)\right) \\ &+ \frac{K_{1}}{h_{n}}\left(\xi_{1}^{+}(\tau+h_{n}z)+\xi_{1}^{-}(\tau+h_{n}z)\right)\right] + o_{P}\left(\sqrt{n^{3}h_{n}^{7}}\right), \end{split}$$

d'où

$$\frac{\sum_{i=1}^{n} (w_i^+(\tau + h_n z) - w_i^-(\tau + h_n z))\sigma(X_i)\varepsilon_i}{\sum_{i=1}^{n} w_i^+(\tau + h_n z)} = \frac{1 + o_P(1)}{\sqrt{nh_n} f_X(\tau)} \tilde{\mathcal{Z}}_n(z) + o_P\left(\frac{1}{\sqrt{nh_n}}\right).$$

Il résulte des énoncés b) et c) du lemme 2.1 que le produit  $B_n \times C_n$  dans l'écriture (2.13) tend en probabilité vers zéro, et, de la partie d) de ce même lemme que  $A_n$  a le comportement annoncé. Ceci termine la démonstration de la proposition 2.1.

Montrons maintenant la convergence cylindrique de  $\left( \tilde{\mathcal{Z}} \right)$ . Pour cela considérons :

$$\begin{split} a\tilde{\mathcal{Z}}_n(z_1) + b\tilde{\mathcal{Z}}_n(z_2) &= \frac{\left[\sum_{i=1}^n \left(aM(\frac{\tau + h_n z_1 - X_i}{h_n}) + bM(\frac{\tau + h_n z_2 - X_i}{h_n})\right)\sigma(X_i)\varepsilon_i\right]}{\sqrt{nh_n}f_X(\tau)},\\ &= \sum_{i=1}^n a\tilde{\mathcal{Z}}_i^{(n)}(z_1) + b\tilde{\mathcal{Z}}_i^{(n)}(z_2). \end{split}$$

Lemme 2.3. Sous les mêmes hypothèses que celles du théorème 2.3, nous avons :

a) 
$$\mathbb{E}\left[a\tilde{\mathcal{Z}}_{n}(z_{1})+b\tilde{\mathcal{Z}}_{n}(z_{2})\right]=0,$$
  
b)  $Var\left[a\tilde{\mathcal{Z}}_{n}(z_{1})+b\tilde{\mathcal{Z}}_{n}(z_{2})\right]=V_{a,b}^{2}(z_{1}-z_{2})(1+o(1)),$   
c)  $\sum_{i=1}^{n}\mathbb{E}\left[\left|a\tilde{\mathcal{Z}}_{i}^{(n)}(z_{1})+b\tilde{\mathcal{Z}}_{i}^{(n)}(z_{2})\right|^{2+\delta}\right]=\mathcal{O}((nh)^{-\delta/2}),$ 

avec

$$V_{a,b}^2(z) = \frac{\sigma^2(\tau)}{f_X(\tau)} \left( \int \left[ aM(x+z) + bM(x) \right]^2 dx \right).$$

Le comportement de la covariance du processus  $(\tilde{\mathcal{Z}}_n)$  se déduit immédiatement de la partie b) :

$$Cov\left(\tilde{\mathcal{Z}}_n(z_1),\tilde{\mathcal{Z}}_n(z_2)\right) = \frac{\sigma^2(\tau)}{f_X(\tau)} \left(\int M(x+z_1-z_2)M(x)dx\right) (1+o(1)).$$

#### Preuve du lemme 2.3.

Les  $\varepsilon_i$  sont des variables aléatoires centrées et indépendantes des  $X_i$ , alors il est facile de montrer que :

a)

$$\operatorname{I\!E}\left[a\tilde{\mathcal{Z}}_n(z_1)+b\tilde{\mathcal{Z}}_n(z_2)\right]=0.$$

b)

$$\begin{split} &Var\left[a\tilde{\mathcal{Z}}_{n}(z_{1})+b\tilde{\mathcal{Z}}_{n}(z_{2})\right] \\ &=\frac{1}{f_{X}^{2}(\tau)h_{n}}\mathbb{E}\left[\left(aM(\frac{\tau+h_{n}z_{1}-X}{h_{n}})+bM(\frac{\tau+h_{n}z_{2}-X}{h_{n}})\right)^{2}\sigma^{2}(X)\varepsilon^{2}\right], \\ &=\frac{1}{f_{X}^{2}(\tau)}\int\left(aM(x+z_{1}-z_{2})+bM(x)\right)^{2}\sigma^{2}(\tau+h_{n}(z_{2}-x))f_{X}(\tau+h_{n}(z_{2}-x))dx, \\ &=\frac{\sigma^{2}(\tau)}{f_{X}(\tau)}\left(\int\left[aM(x+z_{1}-z_{2})+bM(x)\right]^{2}dx\right)(1+o(1)), \\ &=V_{a,b}^{2}(z_{1}-z_{2})(1+o(1)). \end{split}$$

c)

$$\sum_{i=1}^{n} \mathbb{E}\left[\left|a\tilde{\mathcal{Z}}_{i}^{(n)}(z_{1})+b\tilde{\mathcal{Z}}_{i}^{(n)}(z_{2})\right|^{2+\delta}\right]$$
$$=\frac{n\mathbb{E}\left[\left|aM(\frac{\tau+h_{n}z_{1}-X}{h_{n}})+bM(\frac{\tau+h_{n}z_{2}-X}{h_{n}})\right|^{2+\delta}\sigma^{2+\delta}(X)|\varepsilon|^{2+\delta}\right]}{f_{X}^{2+\delta}(\tau)(nh_{n})^{1+\delta/2}},$$
$$=\mathcal{O}\left((nh_{n})^{-\delta/2}\right).$$

Ceci achève la démonstration du lemme 2.3

Comme conséquence du lemme précédent nous avons

$$\frac{\sum_{i=1}^{n} \mathbb{E}\left[\left|a\tilde{\mathcal{Z}}_{i}^{(n)}(z_{1})+b\tilde{\mathcal{Z}}_{i}^{(n)}(z_{2})\right|^{2+\delta}\right]}{\left(Var\left[a\tilde{\mathcal{Z}}_{n}(z_{1})+b\tilde{\mathcal{Z}}_{n}(z_{2})\right]\right)^{1+\delta/2}}=\mathcal{O}((nh)^{-\delta/2})\to 0.$$

D'après le théorème central limite de Liapounov nous pouvons conclure que :

$$a\tilde{\mathcal{Z}}_n(z_1) + b\tilde{\mathcal{Z}}_n(z_2) \xrightarrow{\mathcal{L}} \mathcal{N}\left(0, V^2_{a,b}(z_2 - z_1)\right).$$

Ceci établit la convergence cylindrique du processus  $\overline{\mathcal{Z}}_n$ .

2. Critère de Tension. Il est facile aussi d'établir que :

a) Sous l'hypothèse B6-1, nous avons pour tout  $z_1 \leq z_2$ ,

$$\mathbb{E}\left[\left(\tilde{\mathcal{Z}}_n(z_2) - \tilde{\mathcal{Z}}_n(z_1)\right)^2\right] \le \alpha(z_2 - z_1)^2.$$
(2.17)

b) Sous l'hypothèse B6-2, nous avons pour tout  $z_1 \leq z \leq z_2$ ,

$$\mathbb{E}\left[\left(\tilde{\mathcal{Z}}_n(z_2) - \tilde{\mathcal{Z}}_n(z)\right)^2 \left(\tilde{\mathcal{Z}}_n(z_1) - \tilde{\mathcal{Z}}_n(z)\right)^2\right] \le \alpha(z_2 - z_1)^2.$$
(2.18)

En effet

a)

$$\begin{split} & \mathbb{E}\Big[\tilde{\mathcal{Z}}_{n}(z_{2}) - \tilde{\mathcal{Z}}_{n}(z_{1})\Big]^{2} \\ &= \frac{1}{f_{X}^{2}(\tau)h_{n}} \mathbb{E}\left[\left(M(\frac{\tau + h_{n}z_{2} - X}{h_{n}}) - M(\frac{\tau + h_{n}z_{1} - X}{h_{n}})\right)^{2} \sigma^{2}(X)\varepsilon^{2}\right], \\ &= \frac{1}{f_{X}^{2}(\tau)} \int \left(M(x + z_{1} - z_{2}) - M(x)\right)^{2} \sigma^{2}(\tau + h_{n}(z_{2} - x)) f_{X}(\tau + h_{n}(z_{2} - x)) dx, \\ &\leq \lambda \int \left[M(x + z_{2} - z_{1}) - M(x)\right]^{2} dx, \\ &\leq \alpha(z_{2} - z_{1})^{2}. \end{split}$$

b) Nous avons

$$\mathbb{E}\left[\left(\tilde{\mathcal{Z}}_n(z_2) - \tilde{\mathcal{Z}}_n(z)\right)^2 \left(\tilde{\mathcal{Z}}_n(z_1) - \tilde{\mathcal{Z}}_n(z)\right)^2\right] = \frac{\mathbb{E}\left[\left(\sum_{1 \le i, j \le n} U_i V_j\right)^2\right]}{f_X^4(\tau) n^2 h_n^2},$$

où

i) 
$$U_{i} = \left(M\left(\frac{\tau + h_{n}z_{1} - X_{i}}{h_{n}}\right) - M\left(\frac{\tau + h_{n}z - X_{i}}{h_{n}}\right)\right)\sigma(X_{i})\varepsilon_{i},$$
  
ii) 
$$V_{i} = \left(M\left(\frac{\tau + h_{n}z_{2} - X_{i}}{h_{n}}\right) - M\left(\frac{\tau + h_{n}z - X_{i}}{h_{n}}\right)\right)\sigma(X_{i})\varepsilon_{i}.$$

D'après l'inégalité de Burkholder-Rosenthal, nous obtenons

$$\mathbb{E}\left[\left(\tilde{\mathcal{Z}}_{n}(z_{2}) - \tilde{\mathcal{Z}}_{n}(z)\right)^{2}\left(\tilde{\mathcal{Z}}_{n}(z_{1}) - \tilde{\mathcal{Z}}_{n}(z)\right)^{2}\right] \leq \frac{2\mathbb{E}\left[\sum_{1 \leq i, j \leq n} U_{i}^{2}V_{j}^{2}\right]}{f_{X}^{4}(\tau)n^{2}h_{n}^{2}}, \\ \leq \frac{2n\mathbb{E}\left[U_{1}^{2}V_{1}^{2}\right] + 2n(n-1)\mathbb{E}\left[U_{1}^{2}\right]\mathbb{E}\left[V_{1}^{2}\right]}{f_{X}^{4}(\tau)n^{2}h_{n}^{2}}.$$

Or

i)

$$\begin{split} \operatorname{I\!E} \begin{bmatrix} U_1^2 \end{bmatrix} &= h_n \int (M(x+z_1-z) - M(x))^2 \sigma^2 (\tau + h_n(z-x)) f_X(\tau + h_n(z-x)) dx \,, \\ &\leq \alpha_1 h_n |z_1 - z|, \\ &\leq \alpha_1 h_n |z_2 - z_1|. \end{split}$$
ii) Un calcul similaire conduit à

$$\mathbb{E}\left[V_1^2\right] \leq \alpha_1 h_n |z_2 - z_1|.$$

iii) Si nous supposons que  ${\rm I\!E}\left[\varepsilon^4\right]<\infty,$  alors

$$\mathbb{E} \left[ U_1^2 V_1^2 \right] = h_n \int \left( M(x + z_2 - z) - M(x) \right)^2 (M(x + z_1 - z) - M(x))^2 \\ \times \sigma^4 (\tau + h_n(z - x)) \mathbb{E} \left[ \varepsilon^4 \right] f_X (\tau + h_n(z - x)) dx , \\ \leq \alpha_2 h_n (z_2 - z_1)^2.$$

d'où

$$\mathbb{E}\left[\left(\tilde{\mathcal{Z}}_n(z_2) - \tilde{\mathcal{Z}}_n(z)\right)^2 \left(\tilde{\mathcal{Z}}_n(z_1) - \tilde{\mathcal{Z}}_n(z)\right)^2\right] \leq \alpha(z_2 - z_1)^2.$$

Enfin d'après Billingsley [BIL68], les inégalités (2.17) et (2.18) impliquent que le processus  $\tilde{Z}_n$  est tendu.

3. Limite de  $\mathbb{E}\left[\mathcal{Z}_n(z) \mid X_1, \ldots, X_n\right]$ . Sous l'hypothèse HO, nous avons :

$$\mathbb{E}\left[\mathcal{Z}_{n}(z) \mid X_{1}, \dots, X_{n}\right] = \mathbb{E}\left[\mathcal{Z}_{n}^{+}(z) \mid X_{1}, \dots, X_{n}\right] - \mathbb{E}\left[\mathcal{Z}_{n}^{-}(z) \mid X_{1}, \dots, X_{n}\right],$$

$$= \left(\mathbb{E}\left[\mathcal{Z}_{n}^{+}(z) \mid X_{1}, \dots, X_{n}\right] - \sqrt{nh_{n}}m(\tau + h_{n}z)\right)$$

$$- \left(\mathbb{E}\left[\mathcal{Z}_{n}^{-}(z) \mid X_{1}, \dots, X_{n}\right] - \sqrt{nh_{n}}m(\tau + h_{n}z)\right),$$

$$= \left(\mathbb{E}\left[\mathcal{Z}_{n}^{+}(z) \mid X_{1}, \dots, X_{n}\right] - \sqrt{nh_{n}}m_{+}(\tau + h_{n}z)\right)$$

$$- \left(\mathbb{E}\left[\mathcal{Z}_{n}^{-}(z) \mid X_{1}, \dots, X_{n}\right] - \sqrt{nh_{n}}m_{-}(\tau + h_{n}z)\right).$$

Or

$$\frac{\mathbb{E}\left[\mathcal{Z}_{n}^{\pm}(z) \mid X_{1}, \dots, X_{n}\right]}{\sqrt{nh_{n}}} - m_{\pm}(\tau + h_{n}z) 
= \frac{\sum_{i=1}^{n} w_{i}^{\pm}(\tau + h_{n}z) \left(m(X_{i}) - m_{\pm}(\tau + h_{n}z)\right)}{\sum_{i=1}^{n} w_{i}^{\pm}(\tau + h_{n}z)}, 
= \frac{S_{2}^{\pm}(\tau + h_{n}z) M_{0}^{\pm}(\tau + h_{n}z) - S_{1}^{\pm}(\tau + h_{n}z) M_{1}^{\pm}(\tau + h_{n}z)}{S_{2}^{\pm}(\tau + h_{n}z) S_{0}^{\pm}(\tau + h_{n}z) - S_{1}^{\pm}(\tau + h_{n}z) S_{1}^{\pm}(\tau + h_{n}z)},$$

où

$$M_l^{\pm}(x) = \sum_{i=1}^n (x - X_i)^l K_{\pm}(\frac{x - X_i}{h_n}) \left( m(X_i) - m_{\pm}(x) \right).$$

Lemme 2.4. Sous l'hypothèse nulle HO, nous avons

$$M_l^{\pm}(\tau + h_n z) = \mathcal{O}_P(nh_n^{l+2}).$$

**Preuve du lemme 2.4.** Elle est en tous points semblable à celle du lemme 2.2.

Appliquons maintenant le lemme 2.4 et l'égalité (2.16). Nous en déduisons alors,

$$\mathbb{E}\left[\mathcal{Z}_n^{\pm}(z) \mid X_1, \dots, X_n\right] - \sqrt{nh_n}m_{\pm}(\tau + h_n z) = \mathcal{O}_P\left(\sqrt{nh_n^3}\right).$$

Si nous supposons que  $nh_n^3$  tend vers zéro, alors nous en déduisons que  $\mathbb{E}\left[\mathcal{Z}_n(z) \mid X_1, \ldots, X_n\right]$ converge en probabilité vers zéro.

4. Conclusion. Sous l'hypothèse nulle, le processus  $\mathcal{Z}_n$  converge faiblement vers le processus gaussien centré stationnaire  $\mathcal{Z}$  avec pour fonction de covariance :

$$Cov\left(\mathcal{Z}(z_1), \mathcal{Z}(z_2)\right) = \frac{\sigma^2(\tau)}{f_X(\tau)} \left(\int M(x+z_1-z_2)M(x)dx\right)$$

Ceci termine la démonstration du théorème 2.3.

Le résultat de Leadbetter & al[LLR83] que nous rappelons ci-dessous est fondamental pour la construction de la statistique du test local que nous définissons plus loin.

**Théorème 2.4 (Leadbetter & al**[LLR83]) Soit  $\{\zeta(t); t \ge 0\}$  un processus gaussien stationnaire standard ( $\mathbb{E}[\zeta(t)] = 0$  et  $\mathbb{E}[\zeta^2(t)] = 1$ ) de fonction de covariance  $\rho(\cdot)$ . Supposons que:

C1. la fonction de corrélation,  $\rho(\cdot)$ , admette au voisinage de zéro  $(h \to 0)$  un développement limité de la forme suivante :

$$\rho(h) = 1 - \lambda |h|^{\alpha} + o(|h|^{\alpha}), \qquad (2.19)$$

 $o\dot{u}\ 0<\alpha\leq 2,$ 

C2. lorsque  $h \to \infty$ , nous ayons

$$\rho(h)\log(h) \to 0, \tag{2.20}$$

alors, quand  $T \to \infty$  nous avons:

a) 
$$\mathbb{P}\left[a_T\left(\sup_{t\in[0,T]}\zeta(t)-b_T\right)\leq x\right] \to \exp\left(-\exp(-x)\right)$$
  
b)  $\mathbb{P}\left[a_T\left(\sup_{t\in[0,T]}|\zeta(t)|-b_T\right)\leq x\right] \to \exp\left(-2\exp(-x)\right),$ 

 $o\dot{u}$ 

$$a_T = \sqrt{2\log(T)}, \tag{2.21}$$

et

$$b_T = a_T + \frac{1}{a_T} \left[ \frac{2-\alpha}{2\alpha} \log(\log(T)) + \log(\frac{\lambda^{1/\alpha} H_{\alpha 2}(2-\alpha)/2\alpha}{\sqrt{2\pi}}) \right],$$
(2.22)

avec  $H_{\alpha}$  une constante strictement positive,  $(H_1 = 1, H_2 = \frac{1}{\sqrt{\pi}})$ .

De manière précise notons :

i)  $C(h) = Cov(\mathcal{Z}(z+h), \mathcal{Z}(z)),$ 

ii) 
$$\rho(h) = \frac{C(h)}{C(0)}$$

Lemme 2.5. Sous les conditions B1-B6, le processus Z vérifie les hypothèses C1 et C2.

## Preuve du lemme 2.5.

Vérifions C1. Nous avons

$$\rho(h) = \frac{\int M(x+h)M(x)dx}{\int (M(x))^2 dx}.$$

**Proposition 2.2.** Si la fonction  $K_+(\cdot)$  est de classe  $C^1[-1,0]$ , alors

$$\rho(h) = 1 - c_1(K_+)|h| - c_2(K_+)h^2 + o(h^2),$$

 $o\dot{u}$ 

$$c_1(K_+) = \frac{2M_+^2(0) + M_+^2(-1)}{\int (M(x))^2 dx},$$

et

$$c_{2}(K_{+}) = \frac{M_{+}'(-1)M_{+}(-1) - 2M_{+}'(0)M_{+}(0) + \int (M^{(1)}(x))^{2}dx}{\int (M(x))^{2}dx}$$

Il est clair que  $\int M(x+h)M(x)dx = \int M(x-h)M(x)dx$ . D'où pour prouver la proposition 2.2, quitte à changer h en -h, on peut supposer h > 0.

Puisque nous cherchons un développement limité de  $\rho(\cdot)$  au voisinage de zéro, on peut également supposer que h < 1.  $\int M(x+y)M(x)dx$  se décompose comme suit :

$$\int M(x+h)M(x)dx = \int_{-1}^{-h} M_{+}(x+h)M_{+}(x)dx - \int_{-h}^{0} M_{-}(x+h)M_{+}(x)dx + \int_{0}^{1-h} M_{-}(x+h)M_{-}(x)dx.$$

Il est facile de voir en faisant un changement de variable u = -x - h que

$$\int_{0}^{1-h} M_{-}(x+h)M_{-}(x)dx = \int_{-1}^{-h} M_{-}(-x)M_{-}(-x-h)dx,$$
$$= \int_{-1}^{-h} M_{+}(x+h)M_{+}(x)dx.$$

Comme  $K_+(\cdot) \in \mathcal{C}^1[-1,0]$ , alors  $M_+(\cdot) \in \mathcal{C}^1[-1,0]$  et  $M_-(\cdot) \in \mathcal{C}^1[0,1]$ . Par le théorème de Taylor, nous obtenons:

1)

$$\begin{split} &\int_{-1}^{-h} M_{+}(x+h)M_{+}(x)dx \\ &= \int_{-1}^{0} M_{+}^{2}(x)dx - \frac{h}{2} \left( M_{+}^{2}(0) + M_{+}^{2}(-1) \right) \\ &- \frac{h^{2}}{2} \left( M_{+}'(-1)M_{+}(-1) - M_{+}'(0)M_{+}(0) + \int_{-1}^{0} \left( M_{+}'(x) \right)^{2} dx \right) + o(h^{2}). \end{split}$$

2)

$$\int_{-h}^{0} M_{-}(x+h)M_{+}(x)dx = hM_{+}^{2}(0) - h^{2}M_{+}(0)M_{+}'(0) + o(h^{2}).$$

Il en résulte que

$$\int M(x+h)M(x)dx$$
  
=  $\int_{-1}^{1} M^{2}(x)dx - h\left(2M_{+}^{2}(0) + M_{+}^{2}(-1)\right)$   
-  $h^{2}\left(M_{+}'(-1)M_{+}(-1) - 2M_{+}'(0)M_{+}(0) + \int_{-1}^{0}\left(M_{+}'(x)\right)^{2}dx\right) + o(h^{2}).$ 

ceci implique qu'au voisinage de zéro  $(h \rightarrow 0)$ , nous avons

$$\begin{split} \rho(h) &= 1 - |h| \frac{2M_+^2(0) + M_+^2(-1)}{\int (M(x))^2 dx} \\ &- h^2 \frac{M_+'(-1)M_+(-1) - 2M_+'(0)M_+(0) + \int (M^{(1)}(x))^2 dx}{\int (M(x))^2 dx} + o(h^2) \end{split}$$

Vérifions C2. La condition C2 est clairement réalisée puisque  $\rho(h) = 0$  dès que  $|h| \ge 2$ .

D'où le résultat énoncé dans le lemme 2.5.

De manière standard, nous en déduisons que le supremum de  $\mathcal{Z}_n$  convenablement normalisé admet pour loi asymptotique une loi du type loi de Gumbel. De manière précise notons :

$$\begin{array}{l} \text{i)} \ R_{n,M} = \sup_{z \in [-M,M]} \frac{\mathcal{Z}_n(z)}{\sqrt{C(0)}}, \\ \\ \text{ii)} \ , \ _{n,M} = \sup_{z \in [-M,M]} \frac{|\mathcal{Z}_n(z)|}{\sqrt{C(0)}}, \\ \\ \\ \text{iii)} \ a_M = \sqrt{2\log(M)}, \\ \\ \\ \text{iv)} \ b_M^1 = a_M + \frac{1}{a_M} \left[ \frac{1}{2} \log(\log(M)) + \log(\frac{c_1(K_+)}{\sqrt{\pi}}) \right], \\ \\ \\ \text{v)} \ b_M^2 = a_M + \frac{1}{a_M} \log(\frac{\sqrt{c_2(K_+)}}{\sqrt{2\pi}}). \end{array}$$

Nous avons alors le résultat suivant et son corollaire immédiat :

**Théorème 2.5.** Supposons les conditions du théorème 2.3 satisfaites. Sous l'hypothèse nulle HO :

1) Si B6-1 est vérifiée, alors

a) 
$$\lim_{M \to +\infty} \lim_{n \to +\infty} \mathbb{P}\left[a_M(R_{n,M} - b_M^1) \le x\right] = \exp(-\exp(-x)).$$
  
b) 
$$\lim_{M \to +\infty} \lim_{n \to +\infty} \mathbb{P}\left[a_M(r_{n,M} - b_M^1) \le x\right] = \exp(-2\exp(-x)).$$

- 2) Si B6-2 est vérifiée, alors
  - a)  $\lim_{M \to +\infty} \lim_{n \to +\infty} \operatorname{IP} \left[ a_M (R_{n,M} b_M^2) \le x \right] = \exp(-\exp(-x)).$
  - b)  $\lim_{M \to +\infty} \lim_{n \to +\infty} \operatorname{IP} \left[ a_M(, _{n,M} b_M^2) \le x \right] = \exp(-2 \exp(-x)).$

Il est immédiat de montrer que lorsqu'on remplace C(0) par un estimateur consistant dans ,  $_{n,M}$  et  $R_{n,M}$ , les résultats ci-dessus restent valables. Soit  $\hat{C}(0)$  un estimateur consistant de C(0). Notons

i) 
$$\hat{R}_{n,M} = \sup_{z \in [-M,M]} \frac{\mathcal{Z}_n(z)}{\sqrt{\hat{C}(0)}},$$
  
ii)  $\hat{P}_{n,M} = \sup_{z \in [-M,M]} \frac{|\mathcal{Z}_n(z)|}{\sqrt{\hat{C}(0)}}.$ 

**Corollaire 2.4.** Supposons les conditions du théorème 2.3 satisfaites. Sous l'hypothèse nulle H0 :

1) Si B6-1 est vérifiée, alors

a) 
$$\lim_{M \to +\infty} \lim_{n \to +\infty} \mathbb{P}\left[a_M(\hat{R}_{n,M} - b_M^1) \le x\right] = \exp(-\exp(-x)).$$
  
b) 
$$\lim_{M \to +\infty} \lim_{n \to +\infty} \mathbb{P}\left[a_M(\hat{r}_{n,M} - b_M^1) \le x\right] = \exp(-2\exp(-x)).$$

2) Si B6-2 est vérifiée, alors

a) 
$$\lim_{M \to +\infty} \lim_{n \to +\infty} \mathbb{P}\left[a_M(\hat{R}_{n,M} - b_M^2) \le x\right] = \exp(-\exp(-x)).$$
  
b) 
$$\lim_{M \to +\infty} \lim_{n \to +\infty} \mathbb{P}\left[a_M(\hat{r}_{n,M} - b_M^2) \le x\right] = \exp(-2\exp(-x)).$$

Il est facile de montrer que les tests de l'hypothèse d'absence de rupture, basés naturellement sur les statistiques,  $_{n,M}$  ou  $, _{n,M}$  ont un comportement asymptotique satisfaisant. Ici également on considère des alternatives locales de la forme:

$$\mathrm{H1}(n) \ : \ \gamma_n(\tau) = \frac{\delta_n}{\sqrt{nh_n}},$$

avec  $|\delta_n| \to \infty$ . Le théorème suivant montre que les tests basés sur ,  $_{n,M}$  ou ,  $_{n,M}$  ont une puissance locale asymptotique égale à 1.

**Théorème 2.6.** Supposons B1-B5 et B6-1 ou B6-2 satisfaites. Sous la suite des alternatives locales (H1(n)), nous avons pour tout u, pour tout M > 0 et pour tout  $i = 1, 2, si |\delta_n| \to \infty$ :

a)  $\lim_{n \to +\infty} \operatorname{IP}_{\operatorname{H1}(n)} \left[ a_M(, _{n,M} - b_M^i) \ge u \right] = 1.$ b)  $\lim_{n \to +\infty} \operatorname{IP}_{\operatorname{H1}(n)} \left[ a_M(, _{n,M} - b_M^i) \ge u \right] = 1.$ 

#### Preuve du théorème 2.6.

Pour tout u, pour tout M > 0 et pour tout i = 1, 2 on note

$$E_{n,i}^{+} = \left[ \sqrt{\frac{nh_n}{\hat{C}(0)}} \left( \hat{\gamma}_n(\tau) - \gamma_n(\tau) \right) \ge - \left( \frac{u}{a_M} + b_M^i \right) - \frac{\delta_n(\tau)}{\sqrt{\hat{C}(0)}} \right],$$

 $\operatorname{et}$ 

$$\begin{split} E_{n,i}^{-} &= \left[ \sqrt{\frac{nh_n}{\hat{C}(0)}} \left( \hat{\gamma}_n(\tau) - \gamma_n(\tau) \right) \le \left( \frac{u}{a_M} + b_M^i \right) - \frac{\delta_n(\tau)}{\sqrt{\hat{C}(0)}} \right].\\ \lim_{n \to +\infty} \mathbb{P}_{\mathtt{H1}(n)} \left[ a_M(\hat{,}_{n,M} - b_M^i) \le u \right] &\le \lim_{n \to +\infty} \mathbb{P}_{\mathtt{H1}(n)} \left[ E_{n,i}^+ \cap E_{n,i}^- \right],\\ &= 0. \end{split}$$

Dans plusieurs cas pratiques, on connaît a priori le signe de la rupture, si elle a eu lieu. Pour tester ce type d'hypothèse, un test unilatéral est mieux adapté. Dans le cas où ce signe est positif, les statistiques  $R_{n,M}$  et  $\hat{R}_{n,M}$  fournissent un moyen pour tester l'hypothèse

 $\operatorname{HO}: \gamma(\tau) = 0,$ 

contre

H1:  $\gamma(\tau) > 0$  (il y a un saut positif en  $\tau$  dans la fonction de régression  $m(\cdot)$ ).

Le théorème suivant montre la consistance de ces tests.

**Théorème 2.7.** Supposons B1-B5 et B6-1 ou B6-2 satisfaites. Sous la suite des alternatives locales (H1(n)), nous avons pour tout u, pour tout M > 0 et pour tout  $i = 1, 2, si \delta_n \to \infty$ :

a) 
$$\lim_{n \to +\infty} \operatorname{IP}_{\mathrm{H1}(n)} \left[ a_M (R_{n,M} - b_M^i) \ge u \right] = 1.$$

b)  $\lim_{n \to +\infty} \operatorname{IP}_{\operatorname{H1}(n)} \left[ a_M(\hat{R}_{n,M} - b_M^i) \ge u \right] = 1.$ 

# 2.4 Test Global

À partir des observations  $(X_1, Y_1), \ldots, (X_n, Y_n)$ , nous souhaitons tester l'hypothèse nulle

$$HO: \forall t \in ]0, 1[, \qquad m_{+}(t) = m_{-}(t),$$

l'hypothèse alternative étant

H1: ∃  $\tau \in ]0, 1[, m_+(\tau) \neq m_-(\tau)$  (il y a un point de rupture dans la fonction de régression  $m(\cdot)$ ).

Nous introduisons le processus  $\xi_n(t)$  défini par :

$$\xi_n(t) = \frac{\sqrt{nh_n}}{\sqrt{2}V_+} \left(\frac{f_X(t)}{\sigma^2(t)}\right)^{1/2} \left(\hat{m}_+(t) - \hat{m}_-(t)\right), \qquad (2.23)$$

et nous posons:

$$V(t) = \frac{1}{\sqrt{2}V_{+}} \left(\frac{f_X(t)}{\sigma^2(t)}\right)^{1/2}.$$
(2.24)

Pour faciliter la lecture de cette partie du chapitre, nous allons présenter l'essentiel de notre démarche et énoncer les résultats sans démonstration. L'énoncé précis des hypothèses ainsi que les preuves seront donnés plus loin.

La fonction de régression  $m(\cdot)$  est supposée n'avoir que des discontinuités de première espèce, et on peut donc considérer une version c adl a g (voir hypothèse B7 plus loin). La principale restriction par rapport à la situation du test local et à celle du test strictement local concerne le bruit  $\varepsilon$ . Nous sommes conduits à supposer que la distribution du bruit est suffisamment concentrée (voir hypothèses B10-1 et B10-2). De plus le comportement de la taille de fenêtre  $h_n$  est lié à celui de la queue de la distribution de  $\varepsilon$ . Les conditions concernant  $\varepsilon$  sont a fortiori satisfaites dans le cas borné; elles le sont aussi dans le cas gaussien. Elles sont destinées à nous permettre d'utiliser un résultat de Johnston [JOH82] montrant que le processus  $\zeta_n(t)$ , approximation de  $\xi_n(t)$ , que nous définissons de manière précise ci-dessous peut être approché par un processus gaussien  $G_n(t)$  obtenu par lissage du processus de Wiener bilatéral. Comme la fonction de covariance de ce processus satisfait les hypothèses du théorème de Leadbetter et al. (conditions C1 et C2 du du théorème 2.4), nous obtenons la convergence du supremum de  $\xi_n(t)$  ou de  $|\xi_n(t)|$  vers des lois de Gumbel.

Nous commençons par décomposer  $\xi_n(t)$  en somme de sa moyenne  $E_n(t)$  et de la partie aléatoire centrée  $B_n(t)$ :

$$\xi_n(t) = E_n(t) + B_n(t),$$

оù

$$E_n(t) = E_n^+(t) - E_n^-(t),$$
  
=  $\sqrt{nh_n}V(t) \left[\frac{\sum_{i=1}^n \tilde{w}_i^+(t)m(X_i)}{\sum_{i=1}^n \tilde{w}_i^+(t)} - \frac{\sum_{i=1}^n \tilde{w}_i^-(t)m(X_i)}{\sum_{i=1}^n \tilde{w}_i^-(t)}\right],$ 

 $\operatorname{et}$ 

$$B_n(t) = B_n^+(t) - B_n^-(t),$$
  
=  $\sqrt{nh_n}V(t) \left[ \frac{\sum_{i=1}^n \tilde{w}_i^+(t)\sigma(X_i)\varepsilon_i}{\sum_{i=1}^n \tilde{w}_i^+(t)} - \frac{\sum_{i=1}^n \tilde{w}_i^-(t)\sigma(X_i)\varepsilon_i}{\sum_{i=1}^n \tilde{w}_i^-(t)} \right].$ 

Les quantités  $\tilde{w}_i(t)$  sont définies comme suit :

$$\tilde{w}_i(\cdot) = \frac{1}{n^2 h_n^4} w_i(\cdot).$$

Posons également

$$\tilde{M}_{\pm}(\cdot) = \frac{1}{nh_n} M_{\pm}(\cdot),$$

 $\operatorname{et}$ 

$$\zeta_n(t) = \zeta_n^+(t) - \zeta_n^-(t),$$

où

$$\begin{aligned} \zeta_n^{\pm}(t) &= \frac{1}{\sqrt{2}V_{K_+}\sqrt{nh_nf_X(t)}\sigma(t)}\sum_{i=1}^n M_{\pm}(\frac{t-X_i}{h_n})\sigma(X_i)\varepsilon_i, \\ &= \frac{\sqrt{nh_n}V(t)}{f_X(t)}\sum_{i=1}^n \tilde{M}_{\pm}(\frac{t-X_i}{h_n})\sigma(X_i)\varepsilon_i, \end{aligned}$$

Le lemme suivant montre que  $\xi_n(t)$  peut être approché par  $\zeta_n(t)$ 

**Lemme 2.6.** Sous les hypothèses B7-B12 et pour tout  $h_n = n^{-\delta}$  avec  $1/5 < \delta < 1/3$ , alors il existe  $c_1 > 0$  et  $c_2 > 0$  telles que :

a) 
$$\mathbb{P}\left[\sup_{t\in[0,1]}|B_n(t)-\zeta_n(t)|\geq c_1(nh_n^5)^{1/4}\right]\leq \exp(-c_2nh_n^3)$$
  
b)  $\mathbb{P}_{HO}\left[\sup_{t\in[0,1]}|E_n(t)|\geq c_1(nh_n^5)^{1/2}\right]\leq \exp(-c_2nh_n^3).$ 

La démonstration de ce résultat est donnée plus loin.

On peut maintenant écrire sous l'hypothèse d'absence de discontinuités (HO):

$$\begin{split} & \mathbb{P}_{\mathrm{HO}}\left[a_{n}\left|\sup_{t\in[0,1]}|\xi_{n}(t)|-\sup_{t\in[0,1]}|\zeta_{n}(t)|\right|\geq 2c_{1}a_{n}(nh_{n}^{5})^{1/4}\right],\\ &\leq \mathbb{P}_{\mathrm{HO}}\left[a_{n}\sup_{t\in[0,1]}|\xi_{n}(t)-\zeta_{n}(t)|\geq 2c_{1}a_{n}(nh_{n}^{5})^{1/4}\right],\\ &\leq \mathbb{P}_{\mathrm{HO}}\left[a_{n}\left(\sup_{t\in[0,1]}|B_{n}(t)-\zeta_{n}(t)|+\sup_{t\in[0,1]}|E_{n}(t)|\right)\geq 2c_{1}a_{n}(nh_{n}^{5})^{1/4}\right],\\ &\leq \mathbb{P}_{\mathrm{HO}}\left[a_{n}\sup_{t\in[0,1]}|B_{n}(t)-\zeta_{n}(t)|\geq c_{1}a_{n}(nh_{n}^{5})^{1/4}\right]+\mathbb{P}_{\mathrm{HO}}\left[a_{n}\sup_{t\in[0,1]}|E_{n}(t)|\geq c_{1}a_{n}(nh_{n}^{5})^{1/4}\right],\\ &\leq 2\exp(-c_{2}nh_{n}^{3}). \end{split}$$

Il s'ensuit alors

$$\lim_{n \to +\infty} \mathbb{P}\left[a_n \left(\sup_{t \in [0,1]} \xi_n(t) - b_n\right) \le x\right] = \lim_{n \to +\infty} \mathbb{P}\left[a_n \left(\sup_{t \in [0,1]} \zeta_n(t) - b_n\right) \le x\right].$$

Comme il a été dit plus haut le processus  $\zeta_n(t)$  peut être approché par le processus  $G_n(t) = 1/\sqrt{2V_+^2h_n}\int M(\frac{t-x}{h_n}) dW(x)$  où W est le processus de Wiener bilatéral, au sens suivant

$$\sup_{t \in [0,1]} |\zeta_n(t) - G_n(t)| = o_P\left( (\log n)^{-1/2} \right).$$

Il vient alors:

$$\lim_{n \to +\infty} \mathbb{P}\left[a_n \left(\sup_{t \in [0,1]} \zeta_n(t) - b_n\right) \le x\right] = \lim_{n \to +\infty} \mathbb{P}\left[a_n \left(\sup_{t \in [0,1]} G_n(t) - b_n\right) \le x\right].$$

On vérifie facilement par changement de variable que

$$\mathcal{L}\left\{\sup G_n(t)\right\} = \mathcal{L}\left\{\sup_{t\in[0,T_n]}\frac{1}{\sqrt{2}V_+}\int M(t-x)\,dW(x)\right\}$$

où  $T_n = 1/h_n$ . Le processus  $\int M(t-x) dW(x)$  est un processus gaussien centré de fonction de covariance définie par

$$C(h) = \int M(x+h)M(x) \, dx$$

dont on a vu plus haut qu'elle satisfait les conditions du théorème 2.4. Avec la normalisation habituelle sup  $G_n(t)$  est donc asymptotiquement de loi de Gumbel. Nous avons donné les grandes lignes de la démonstration du résultat suivant:

Théorème 2.8. Notons

$$a_n = \sqrt{-2\log(h_n)}, \qquad (2.25)$$

1) Si  $K_{+}(-1) \neq 0$  ou  $K_{+}(0) \neq 0$ ,

$$b_n = a_n + \frac{1}{a_n} \left[ \log \left( \frac{c_1(K_+)}{\sqrt{\pi}} \right) + \frac{1}{2} \log \left( \log \left( \frac{1}{h_n} \right) \right) \right], \qquad (2.26)$$

2) Si 
$$K_{+}(-1) = K_{+}(0) = 0$$
,

$$b_n = a_n + \frac{1}{a_n} \log\left(\frac{\sqrt{c_2(K_+)}}{\sqrt{2\pi}}\right),$$
 (2.27)

оù

$$c_1(K_+) = \frac{2M_+^2(0) + M_+^2(-1)}{\int M^2(u) du},$$

et

$$c_2(K_+) = \frac{\int (M'(u))^2 du}{2 \int M^2(u) du}.$$

Supposons les conditions B7-B12 satisfaites et  $h_n = n^{-\delta}$  avec  $\frac{1}{5} < \delta < \frac{1}{3}$ . Sous l'hypothèse nulle H0 :

a) 
$$\lim_{n \to +\infty} \mathbb{P}\left[a_n \left(\sup_{t \in [0,1]} \xi_n(t) - b_n\right) \le x\right] = \exp(-\exp(-x)).$$
  
b) 
$$\lim_{n \to +\infty} \mathbb{P}\left[a_n \left(\sup_{t \in [0,1]} |\xi_n(t)| - b_n\right) \le x\right] = \exp(-2\exp(-x)).$$

Les arguments précédant l'énoncé seront développés plus loin de manière plus précise.

Il est facile de montrer que le test global de l'hypothèse d'absence de rupture, basé sur le théorème 2.8 a un comportement asymptotique satisfaisant. Ici également on considère des alternatives locales de la forme :

$$\mathrm{H1}(n) : \gamma_n(\tau) = \frac{\delta_n}{\sqrt{nh_n}},$$

avec  $\frac{|\delta_n|}{\sqrt{\log(n)}} \to \infty$ . Le théorème suivant montre que le test global, lui aussi, a une puissance locale asymptotique égale à 1.

**Théorème 2.9.** Supposons les hypothèses B1-B5 satisfaites. Soient  $(a_n)$  et  $(b_n)$  les suites définies au théorème 2.8.

1) Si 
$$\frac{\delta_n}{\sqrt{\log(n)}} \to +\infty$$
, sous la suite d'alternatives locales  $H1(n)$ , nous avons pour tout  $u > 0$ 

$$\lim_{n \to +\infty} \mathbb{P}_{\mathrm{H1}(n)} \left[ a_n \left( \sup_{t \in [0,1]} \xi_n(t) - b_n \right) \ge u \right] = 1.$$
(2.28)

2) Si  $\frac{|\delta_n|}{\sqrt{\log(n)}} \to \infty$ , sous la suite d'alternatives locales H1(n), nous avons pour tout u > 0

$$\lim_{n \to +\infty} \mathbb{P}_{\mathtt{H1}(n)} \left[ a_n \left( \sup_{t \in [0,1]} |\xi_n(t)| - b_n \right) \ge u \right] = 1.$$
(2.29)

### Hypothèses

- B7. La fonction  $m(\cdot)$  ainsi que ses dérivées première et seconde sont supposées càdlàg sur [0, 1].
- B8.  $f_X(\cdot)$ : la densité marginale de X est supposée lipschitzienne et strictement positive sur [0, 1].
- B9.  $\sigma^2(\cdot)$  est supposée bornée sur l'intervalle [0,1].

**B10-1.** 
$$\exists C > 0$$
 tel que  $\mathbb{E}\left[|\varepsilon|^k\right] < C^{k-2}k!, k = 3, 4, \dots$ 

**B10-2.** Il existe une suite  $(\alpha_n)_{n\geq 1}$  tel que

1) 
$$\alpha_n \to +\infty$$
,  
2)  $\operatorname{E}\left[|\varepsilon|^2 \operatorname{II}(|\varepsilon| > \alpha_n)\right] = \mathcal{O}(1/n)$ 

B11. Soit  $(h_n)_{n\geq 1}$  une suite réelle positive vérifiant:

1) 
$$\lim_{n \to +\infty} h_n = 0$$
 et  $\lim_{n \to +\infty} nh_n = +\infty$ ,  
2)  $\alpha_n h_n^{-1/2} n^{-1/6} (\log n)^2 \to 0.$ 

B12.  $K_+(\cdot)$  est une fonction dérivable à support sur [-1, 0].

## Remarque 2.1.

- Si les ε sont bornés, pour toute suite (α<sub>n</sub>) qui tend vers l'infini, les conditions B10-1 et B10-2 sont vérifiées.
- Si ε → N(0, 1) et en choisissant α<sub>n</sub> = n<sup>β</sup> (β > 0, β proche de zéro), les conditions B10-1 et B10-2 sont vérifiées.
- 3) Si  $\alpha_n h_n^{-1/2} n^{-1/6} (\log n)^2 \to 0$  alors  $nh_n \to \infty$ .

On peut écrire

$$\xi_n(t) = E_n(t) + B_n(t),$$

où

$$E_n(t) = E_n^+(t) - E_n^-(t),$$
  
=  $\sqrt{nh_n}V(t) \left[\frac{\sum_{i=1}^n \tilde{w}_i^+(t)m(X_i)}{\sum_{i=1}^n \tilde{w}_i^+(t)} - \frac{\sum_{i=1}^n \tilde{w}_i^-(t)m(X_i)}{\sum_{i=1}^n \tilde{w}_i^-(t)}\right],$ 

 $\operatorname{et}$ 

$$B_n(t) = B_n^+(t) - B_n^-(t),$$
  
=  $\sqrt{nh_n}V(t) \left[\frac{\sum_{i=1}^n \tilde{w}_i^+(t)\sigma(X_i)\varepsilon_i}{\sum_{i=1}^n \tilde{w}_i^+(t)} - \frac{\sum_{i=1}^n \tilde{w}_i^-(t)\sigma(X_i)\varepsilon_i}{\sum_{i=1}^n \tilde{w}_i^-(t)}\right].$ 

Les quantités  $\tilde{w}_i(t)$  sont définies comme suit :

$$\tilde{w}_i(\cdot) = \frac{1}{n^2 h_n^4} w_i(\cdot).$$

Posons également

$$\tilde{M}_{\pm}(\cdot) = \frac{1}{nh_n} M_{\pm}(\cdot),$$

 $\operatorname{et}$ 

$$\zeta_n(t) = \zeta_n^+(t) - \zeta_n^-(t),$$

оù

$$\begin{aligned} \zeta_n^{\pm}(t) &= \frac{1}{\sqrt{2}V_{K_{\pm}}\sqrt{nh_nf_X(t)}\sigma(t)} \sum_{i=1}^n M_{\pm}(\frac{t-X_i}{h_n})\sigma(X_i)\varepsilon_i, \\ &= \frac{\sqrt{nh_n}V(t)}{f_X(t)} \sum_{i=1}^n \tilde{M}_{\pm}(\frac{t-X_i}{h_n})\sigma(X_i)\varepsilon_i, \end{aligned}$$

alors nous avons le lemme suivant

**Lemme 2.7.** Sous les hypothèses B7-B12 et pour tout  $h_n = n^{-\delta}$  avec  $1/5 < \delta < 1/3$ , alors il existe  $c_1 > 0$  et  $c_2 > 0$  telles que :

a) IP 
$$\left[\sup_{t \in [0,1]} |B_n(t) - \zeta_n(t)| \ge c_1 (nh_n^5)^{1/4}\right] \le \exp(-c_2 nh_n^3).$$
  
b) IP<sub>HO</sub>  $\left[\sup_{t \in [0,1]} |E_n(t)| \ge c_1 (nh_n^5)^{1/2}\right] \le \exp(-c_2 nh_n^3).$ 

# Preuve du lemme 2.7.

a) D'après l'inégalité triangulaire, nous avons

$$\sup_{t \in [0,1]} |B_n(t) - \zeta_n(t)| \le \sup_{t \in [0,1]} |B_n^+(t) - \zeta_n^+(t)| + \sup_{t \in [0,1]} |B_n^-(t) - \zeta_n^-(t)|.$$

D'autre part

$$B_{n}^{\pm}(t) - \zeta_{n}^{\pm}(t) = \sqrt{nh_{n}}V(t)\frac{\sum_{i=1}^{n} \left(\tilde{w}_{i}^{\pm}(t) - f_{X}(t)\tilde{M}_{\pm}(\frac{t-X_{i}}{h_{n}})\right)\sigma(X_{i})\varepsilon_{i}}{\sum_{i=1}^{n}\tilde{w}_{i}^{\pm}(t)} + \frac{(nh_{n})^{1/4}V(t)}{f_{X}(t)}\left[\sum_{i=1}^{n}\tilde{M}_{\pm}(\frac{t-X_{i}}{h_{n}})\sigma(X_{i})\varepsilon_{i}\right]\frac{(nh_{n})^{1/4}\left(f_{X}^{2}(t) - \sum_{i=1}^{n}\tilde{w}_{i}^{\pm}(t)\right)}{\sum_{i=1}^{n}\tilde{w}_{i}^{\pm}(t)}.$$

Nous en déduisons, qu'il existe  $a_1,a_2>0$  telles que

$$\begin{split} \sup_{t \in [0,1]} &|B_n^{\pm}(t) - \zeta_n^{\pm}(t)| \\ \leq a_1 \left( \sup_{t \in [0,1]} \sqrt{nh_n} \left| \sum_{i=1}^n \left( \tilde{w}_i^{\pm}(t) - f_X(t) \tilde{M}_{\pm}(\frac{t - X_i}{h_n}) \right) \sigma(X_i) \varepsilon_i \right| \right) \left( \sup_{t \in [0,1]} \frac{1}{|\sum_{i=1}^n \tilde{w}_i^{\pm}(t)|} \right) \\ &+ a_2 \left( \sup_{t \in [0,1]} (nh_n)^{1/4} \left| \sum_{i=1}^n \tilde{M}_{\pm}(\frac{t - X_i}{h_n}) \sigma(X_i) \varepsilon_i \right| \right) \\ &\times \left( \sup_{t \in [0,1]} (nh_n)^{1/4} \left| \sum_{i=1}^n \tilde{w}_i^{\pm}(t) - f_X^2(t) \right| \right) \left( \sup_{t \in [0,1]} \frac{1}{|\sum_{i=1}^n \tilde{w}_i^{\pm}(t)|} \right), \\ \leq \max(a_1, a_2) A \times (B + C \times D), \end{split}$$

оù

$$\begin{aligned} \mathbf{i)} \ \ A &= \sup_{t \in [0,1]} \frac{1}{|\sum_{i=1}^{n} \tilde{w}_{i}^{\pm}(t)|}, \\ \mathbf{ii)} \ \ B &= \sqrt{nh_{n}} \left( \sup_{t \in [0,1]} \left| \sum_{i=1}^{n} \left( \tilde{w}_{i}^{\pm}(t) - f_{X}(t) \tilde{M}_{\pm}(\frac{t-X_{i}}{h_{n}}) \right) \sigma(X_{i}) \varepsilon_{i} \right| \right), \\ \mathbf{iii)} \ \ C &= (nh_{n})^{1/4} \left( \sup_{t \in [0,1]} \left| \sum_{i=1}^{n} \tilde{M}_{\pm}(\frac{t-X_{i}}{h_{n}}) \sigma(X_{i}) \varepsilon_{i} \right| \right), \\ \mathbf{iv}) \ \ D &= (nh_{n})^{1/4} \left( \sup_{t \in [0,1]} \left| \sum_{i=1}^{n} \tilde{w}_{i}^{\pm}(t) - f_{X}^{2}(t) \right| \right). \end{aligned}$$

Nous avons

$$A \leq \frac{1}{\left| \inf_{t \in [0,1]} f_X^2(t) - \sup_{t \in [0,1]} \left| \sum_{i=1}^n \tilde{w}_i^{\pm}(t) - f_X^2(t) \right| \right|},$$

En posant

$$\inf_{t \in [0,1]} f_X^2(t) = f_0^2,$$

$$\begin{split} & \mathbb{P}\left[A \times (B + C \times D) \geq \frac{2}{f_0^2} x^2\right] \\ & \leq \mathbb{P}\left[A \geq \frac{2}{f_0^2}\right] + \mathbb{P}\left[B + C \times D \geq x^2\right], \\ & \leq \mathbb{P}\left[A \geq \frac{2}{f_0^2}\right] + \mathbb{P}\left[B \geq \frac{x^2}{2}\right] + \mathbb{P}\left[C \times D \geq \frac{x^2}{2}\right], \\ & \leq \mathbb{P}\left[A \geq \frac{2}{f_0^2}\right] + \mathbb{P}\left[B \geq \frac{x^2}{2}\right] + \mathbb{P}\left[C \geq \frac{|x|}{\sqrt{2}}\right] + \mathbb{P}\left[D \geq \frac{|x|}{\sqrt{2}}\right] \end{split}$$

Comme si  $|x| \leq 1,$  alors  $x^2 \leq |x|$  nous pouvons donc écrire

$$\begin{split} & \mathbb{P}\left[A \times (B + C \times D) \geq \frac{2}{f_0^2} x^2\right] \\ & \leq \mathbb{P}\left[A \geq \frac{2}{f_0^2}\right] + \mathbb{P}\left[B \geq \frac{x^2}{2}\right] + \mathbb{P}\left[C \geq \frac{x^2}{2}\right] + \mathbb{P}\left[D \geq \frac{x^2}{2}\right], \end{split}$$

or

$$\begin{split} \mathbb{P}\left[A \geq \frac{2}{f_0^2}\right] &\leq \mathbb{P}\left[\frac{1}{\left|\inf_{t \in [0,1]} f_X^2(t) - \sup_{t \in [0,1]} \left|\sum_{i=1}^n \tilde{w}_i^{\pm}(t) - f_X^2(t)\right|\right|} \geq \frac{2}{f_0^2}\right],\\ &\leq \mathbb{P}\left[\sup_{t \in [0,1]} \left|\sum_{i=1}^n \tilde{w}_i^{\pm}(t) - f_X^2(t)\right| \geq \frac{f_0^2}{2}\right]. \end{split}$$

La majoration de ces quatre termes est assurée par le lemme 2.8.

b) Sous l'hypothèse nulle, nous avons

$$\sup_{t \in [0,1]} |E_n(t)| \leq \sup_{t \in [0,1]} |E_n^+(t) - \sqrt{nh_n} V(t)m_+(t)| + \sup_{t \in [0,1]} |E_n^-(t) - \sqrt{nh_n} V(t)m_-(t)|.$$

Or

$$\begin{split} E_n^{\pm}(t) &- \sqrt{nh_n} V(t) m_{\pm}(t) \\ &= \sqrt{nh_n} V(t) \frac{\sum_{i=1}^n \tilde{w}_i^{\pm}(t) \left[ m(X_i) - m_{\pm}(t) - (X_i - t) m_{\pm}^{(1)}(t) \right]}{\sum_{i=1}^n \tilde{w}_i^{\pm}(t)}, \\ &= \sqrt{nh_n} V(t) \frac{\tilde{S}_2^{\pm}(t) \tilde{R}_0^{\pm}(t) - \tilde{S}_1^{\pm}(t) \tilde{R}_1^{\pm}(t)}{\sum_{i=1}^n \tilde{w}_i^{\pm}(t)}, \\ &= \frac{\sqrt{nh_n} V(t)}{\left(\sum_{i=1}^n \tilde{w}_i^{\pm}(t) - f_X^2(t)\right) + f_X^2(t)} \left[ \left( \tilde{S}_2^{\pm}(t) - K_2^{\pm} f_X(t) \right) \tilde{R}_0^{\pm}(t) + K_2^{\pm} f_X(t) \tilde{R}_0^{\pm}(t) \right. \\ &- \left( \tilde{S}_1^{\pm}(t) - K_1^{\pm} f_X(t) \right) \tilde{R}_1^{\pm}(t) - K_1^{\pm} f_X(t) \tilde{R}_1^{\pm}(t) \right], \end{split}$$
où  $\tilde{S}_l^{\pm}(\cdot) = \frac{1}{nh_n^{l+1}} S_l^{\pm}(\cdot) \text{ (idem pour } \tilde{R}_l^{\pm}(\cdot) \text{ et } \tilde{\xi}_l^{\pm}(\cdot)). \end{split}$ 

Il en résulte que

$$\begin{split} \sup_{t \in [0,1]} &|E_n^{\pm}(t) - \sqrt{nh_n} V(t) m_{\pm}(t)| \\ &\leq 2 \sup_{t \in [0,1]} V(t) \sqrt{nh_n} \max_{l=0,1} \sup_{t \in [0,1]} \left| \tilde{R}_l^{\pm}(t) \right| \\ &\times \frac{\left[ \max_{l=1,2} \sup_{t \in [0,1]} \left| \tilde{S}_l^{\pm}(t) - K_l^{\pm} f_X(t) \right| + \max_{l=1,2} \sup_{t \in [0,1]} \left| K_l^{\pm} f_X(t) \right| \right]}{\left| \inf_{t \in [0,1]} f_X^2(t) - \sup_{t \in [0,1]} \left| \sum_{i=1}^n \tilde{w}_i^{\pm}(t) - f_X^2(t) \right| \right|}, \\ &\leq \alpha_1 \sqrt{nh_n} \max_{l=0,1} \sup_{t \in [0,1]} \left| \tilde{R}_l^{\pm}(t) \right| \frac{\left[ \max_{l=1,2} \sup_{t \in [0,1]} \left| \tilde{S}_l^{\pm}(t) - K_l^{\pm} f_X(t) \right| + \alpha_2 \right]}{\left| \alpha_3 - \sup_{t \in [0,1]} \left| \sum_{i=1}^n \tilde{w}_i^{\pm}(t) - f_X^2(t) \right| \right|}, \end{split}$$

où

i) 
$$\alpha_1 = 2 \sup_{t \in [0,1]} V(t),$$
  
ii)  $\alpha_2 = \max_{l=1,2} \sup_{t \in [0,1]} |K_l^{\pm} f_X(t)|,$   
iii)  $\alpha_3 = \inf_{t \in [0,1]} f_X^2(t).$ 

Posons

$$A_{n} = \alpha_{1} \frac{\left[ \max_{l=1,2} \sup_{t \in [0,1]} \left| \tilde{S}_{l}^{\pm}(t) - K_{l}^{\pm} f_{X}(t) \right| + \alpha_{2} \right]}{\left| \alpha_{3} - \sup_{t \in [0,1]} \left| \sum_{i=1}^{n} \tilde{w}_{i}^{\pm}(t) - f_{X}^{2}(t) \right| \right|},$$

 $\operatorname{alors}$ 

$$\begin{split} & \mathbb{P}_{\mathrm{H0}} \left[ \sup_{t \in [0,1]} \left| E_{n}^{\pm}(t) - \sqrt{nh_{n}} V(t) m_{\pm}(t) \right| \geq c_{5} (nh_{n}^{5})^{1/2} \right] \\ & \leq \mathbb{P}_{\mathrm{H0}} \left[ A_{n} \left( \max_{l=0,1} \sup_{t \in [0,1]} \left| \tilde{R}_{l}^{\pm}(t) \right| \right) \geq c_{5} h_{n}^{2} \right], \\ & \leq \mathbb{P}_{\mathrm{H0}} \left[ A_{n} \left( \max_{l=0,1} \sup_{t \in [0,1]} \left| \tilde{R}_{l}^{\pm}(t) \right| \right) \geq c_{5} h_{n}^{2} ; A_{n} \leq \frac{2\alpha_{1}\alpha_{2}}{\alpha_{3}} \right] \\ & + \mathbb{P}_{\mathrm{H0}} \left[ A_{n} \left( \max_{l=0,1} \sup_{t \in [0,1]} \left| \tilde{R}_{l}^{\pm}(t) \right| \right) \geq c_{5} h_{n}^{2} ; A_{n} > \frac{2\alpha_{1}\alpha_{2}}{\alpha_{3}} \right], \\ & \leq \mathbb{P}_{\mathrm{H0}} \left[ \max_{l=0,1} \sup_{t \in [0,1]} \left| \tilde{R}_{l}^{\pm}(t) \right| \geq \frac{\alpha_{3}c_{5}}{2\alpha_{1}\alpha_{2}} h_{n}^{2} \right] + \mathbb{P}_{\mathrm{H0}} \left[ A_{n} > \frac{2\alpha_{1}\alpha_{2}}{\alpha_{3}} \right]. \end{split}$$

Or

$$\begin{split} & \mathbb{P}_{\mathsf{H0}}\left[A_n > \frac{2\alpha_1\alpha_2}{\alpha_3}\right] \\ & \leq \mathbb{P}\left[\max_{l=1,2}\sup_{t\in[0,1]} \left|\tilde{S}_l^{\pm}(t) - K_l^{\pm}f_X(t)\right| > \frac{\alpha_2}{2}\right] + \mathbb{P}\left[\sup_{t\in[0,1]} \left|\sum_{i=1}^n \tilde{w}_i^{\pm}(t) - f_X^2(t)\right| > \frac{\alpha_3}{4}\right], \\ & \leq \exp(-c_6nh_n^3). \end{split}$$

Ceci implique qu'il existe  $c_1>0$  et  $c_2>0$  telles que

$$\operatorname{IP}_{HO}\left[\sup_{t\in[0,1]}|E_n(t)|\geq c_1(nh_n^5)^{1/2}\right]\leq \exp(-c_2nh_n^3).$$

**Lemme 2.8.** Sous les mêmes hypothèses que celles du lemme 2.7, il existe  $c_3, c_4 > 0$  telles que

a) 
$$\mathbb{P}\left[ (nh_{n})^{1/2} \sup_{t \in [0,1]} \left| \sum_{i=1}^{n} \left( \tilde{w}_{i}^{\pm}(t) - f_{X}(t) \tilde{M}_{\pm}(\frac{t-X_{i}}{h_{n}}) \right) \sigma(X_{i}) \varepsilon_{i} \right| \ge c_{3}(nh_{n}^{5})^{1/4} \right] \le \exp(-c_{4}nh_{n}^{3}).$$
  
b) 
$$\mathbb{P}\left[ \sup_{t \in [0,1]} \left| \sum_{i=1}^{n} \tilde{M}_{\pm}(\frac{t-X_{i}}{h_{n}}) \sigma(X_{i}) \varepsilon_{i} \right| \ge c_{3}h_{n} \right] \le \exp(-c_{4}nh_{n}^{3}).$$
  
c) 
$$\mathbb{P}\left[ \sup_{t \in [0,1]} \left| \sum_{i=1}^{n} \tilde{w}_{i}^{\pm}(t) - (K_{2}^{\pm}K_{0}^{\pm} - (K_{1}^{\pm})^{2})f_{X}^{2}(t) \right| \ge c_{3}h_{n} \right] \le \exp(-c_{4}nh_{n}^{3}).$$

Preuve du lemme 2.8.

$$(nh_n)^{1/2} \sup_{t \in [0,1]} \left| \sum_{i=1}^n \left( \tilde{w}_i^{\pm}(t) - f_X(t) \tilde{M}_{\pm}(\frac{t - X_i}{h_n}) \right) \sigma(X_i) \varepsilon_i \right|,$$

$$= (nh_n)^{1/2} \sup_{t \in [0,1]} \left| \sum_{i=1}^n \tilde{w}_i^{\pm}(t) \sigma(X_i) \varepsilon_i - K_2^{\pm} f_X(t) \tilde{\xi}_0^{\pm}(t) + K_1^{\pm} f_X(t) \tilde{\xi}_1^{\pm}(t) \right|,$$

$$\le (nh_n)^{1/2} \left( \sup_{t \in [0,1]} \left| \tilde{S}_2^{\pm}(t) - K_2^{\pm} f_X(t) \right| \right) \left( \sup_{t \in [0,1]} \left| \tilde{\xi}_0^{\pm}(t) \right| \right)$$

$$+ \left( \sup_{t \in [0,1]} \left| \tilde{S}_1^{\pm}(t) - K_1^{\pm} f_X(t) \right| \right) \left( \sup_{t \in [0,1]} \left| \tilde{\xi}_1^{\pm}(t) \right| \right),$$

$$\le 2(nh_n)^{1/2} \left( \max_{l=1,2} \sup_{t \in [0,1]} \left| \tilde{S}_l^{\pm}(t) - K_l^{\pm} f_X(t) \right| \right) \left( \max_{l=0,1} \sup_{t \in [0,1]} \left| \tilde{\xi}_l^{\pm}(t) \right| \right).$$

**Lemme 2.9.** Soient X et Y deux variables aléatoires. Supposons qu'il existe  $0 \le a \le 1$  et  $b \ge 0$  telles que nous ayons :

- 1)  $\mathbb{P}\left[|X| \geq a\right] \leq b$ ,
- 2)  $\mathbb{P}[|Y| \ge a] \le b$ ,

a lors,

$$\mathbb{P}\left[|XY| \geq a\right] \quad \leq \quad 2b$$

## Preuve du lemme 2.9.

Nous avons

$$\begin{split} \mathbb{P}\left[|XY| \ge a\right] &= \mathbb{P}\left[|XY| \ge a \cap |X| > 1\right] + \mathbb{P}\left[|XY| \ge a \cap |X| \le 1\right], \\ &\leq \mathbb{P}\left[|X| > 1\right] + \mathbb{P}\left[|Y| \ge a\right], \\ &\leq 2b. \end{split}$$

Pour terminer la démonstration du lemme 2.8, nous nous basons sur le lemme 2.9 en posant :

$$\begin{aligned} \text{i)} \quad X &= 2(nh_n)^{1/4} \left( \max_{l=1,2} \sup_{t \in [0,1]} |\tilde{S}_l^{\pm}(t) - K_l^{\pm} f_X(t)| \right), \\ \text{ii)} \quad Y &= (nh_n)^{1/4} | \left( \max_{l=0,1} \sup_{t \in [0,1]} \tilde{\xi}_l^{\pm}(t)| \right), \\ \text{iii)} \quad a &= c_5 (nh_n^5)^{1/4}, \end{aligned}$$

iv) 
$$b = \exp(-c_7 n h_n^3)$$
,

et en utilisanta) et b) du lemme 2.10 démontré ci-dessous. Nous en déduisons donc qu'il existe  $c_3 > 0$  et  $c_5 > 0$  telles que

$$\mathbb{P}\left[\sup_{t\in[0,1]}\sqrt{nh_n}\left|\sum_{i=1}^n\left(\tilde{w}_i^{\pm}(t)-f_X(t)\tilde{M}_{\pm}(\frac{t-X_i}{h_n})\right)\sigma(X_i)\varepsilon_i\right|\geq c_3(nh_n^5)^{1/4}\right]\\\leq \exp(-c_4nh_n^3).$$

b)

$$\begin{aligned} \sup_{t\in[0,1]} \left| \sum_{i=1}^{n} \tilde{M}_{\pm}\left(\frac{t-X_{i}}{h_{n}}\right) \sigma(X_{i}) \varepsilon_{i} \right| &\leq |K_{2}^{\pm} \left| \sup_{t\in[0,1]} \left| \tilde{\xi}_{0}^{\pm} \right| + |K_{1}^{\pm} \left| \sup_{t\in[0,1]} \left| \tilde{\xi}_{1}^{\pm} \right| \right|, \\ &\leq c_{8} \left( \max_{l=0,1} \sup_{t\in[0,1]} \left| \tilde{\xi}_{l}^{\pm} \right| \right), \end{aligned}$$

or d'après le e) du lemme 2.10 ci-dessous nous avons

$$\mathbb{P}\left[\max_{l=0,1}\sup_{t\in[0,1]}\left|\tilde{\xi}_{l}^{\pm}(t)\right|\geq c_{5}h_{n}\right] \leq \exp\left(-c_{7}nh_{n}^{3}\right),$$

ceci termine la démonstration de la deuxième partie du lemme 2.8.

c) En écrivant

$$\begin{split} &\sum_{i=1}^{n} \tilde{w}_{i}^{\pm}(t) - \left(K_{2}^{\pm}K_{0}^{\pm} - \left(K_{1}^{\pm}\right)^{2}\right)f_{X}^{2}(t) \\ &= \tilde{S}_{2}^{\pm}(t)\tilde{S}_{0}^{\pm}(t) - \tilde{S}_{1}^{\pm}(t)\tilde{S}_{1}^{\pm}(t) - \left(K_{2}^{\pm}K_{0}^{\pm} - \left(K_{1}^{\pm}\right)^{2}\right)f_{X}^{2}(t), \\ &= \left(\tilde{S}_{2}^{\pm}(t) - K_{2}^{\pm}f_{X}(t)\right)\left(\tilde{S}_{0}^{\pm}(t) - K_{0}^{\pm}f_{X}(t)\right) - \left(\tilde{S}_{1}^{\pm}(t) - K_{1}^{\pm}f_{X}(t)\right)^{2} \\ &+ \left(\tilde{S}_{2}^{\pm}(t) - K_{2}^{\pm}f_{X}(t)\right)K_{0}^{\pm}f_{X}(t) + \left(\tilde{S}_{0}^{\pm}(t) - K_{0}^{\pm}f_{X}(t)\right)K_{2}^{\pm}f_{X}(t) \\ &- 2K_{1}^{\pm}f_{X}(t)\left(\tilde{S}_{1}^{\pm}(t) - K_{1}^{\pm}f_{X}(t)\right). \end{split}$$

Nous en déduisons

$$\begin{split} \sup_{t\in[0,1]} \left| \sum_{i=1}^{n} \tilde{w}_{i}^{\pm}(t) - (K_{2}^{\pm}K_{0}^{\pm} - (K_{1}^{\pm})^{2})f_{X}^{2}(t) \right| \\ &\leq \left( \sup_{t\in[0,1]} \left| \tilde{S}_{2}^{\pm}(t) - K_{2}^{\pm}f_{X}(t) \right| \right) \left( \sup_{t\in[0,1]} \left| \tilde{S}_{0}^{\pm}(t) - K_{0}^{\pm}f_{X}(t) \right| \right) + \sup_{t\in[0,1]} \left( \tilde{S}_{1}^{\pm}(t) - K_{1}^{\pm}f_{X}(t) \right)^{2} \\ &+ \left( \sup_{t\in[0,1]} \left| \tilde{S}_{2}^{\pm}(t) - K_{2}^{\pm}f_{X}(t) \right| \right) \left( \sup_{t\in[0,1]} |K_{0}^{\pm}|f_{X}(t) \right) \\ &+ \left( \sup_{t\in[0,1]} \left| \tilde{S}_{0}^{\pm}(t) - K_{0}^{\pm}f_{X}(t) \right| \right) \left( \sup_{t\in[0,1]} |K_{2}^{\pm}|f_{X}(t) \right) \\ &+ 2 \left( \sup_{t\in[0,1]} |K_{1}^{\pm}|f_{X}(t) \right) \left( \sup_{t\in[0,1]} \left| \tilde{S}_{1}^{\pm}(t) - K_{1}^{\pm}f_{X}(t) \right| \right) \\ &\leq \left( \max_{l=0,1,2} \sup_{t\in[0,1]} \left| \tilde{S}_{l}^{\pm}(t) - K_{l}^{\pm}f_{X}(t) \right| \right) \\ &\times \left[ 2 \max_{l=0,1,2} \sup_{t\in[0,1]} \left| \tilde{S}_{l}^{\pm}(t) - K_{l}^{\pm}f_{X}(t) \right| + 4 \max_{l=0,1,2} \sup_{t\in[0,1]} \left| K_{l}^{\pm}f_{X}(t) \right| \right]. \end{split}$$

Utilisant à nouveau le lemme 2.10, on voit qu'il existe  $c_3 > 0$  et  $c_4 > 0$  tels que

$$\mathbb{P}\left[\sup_{t\in[0,1]}\left|\sum_{i=1}^{n}\tilde{w}_{i}^{\pm}(t)-(K_{2}^{\pm}K_{0}^{\pm}-(K_{1}^{\pm})^{2})f_{X}^{2}(t)\right|\geq c_{3}h_{n}\right]\leq\exp(-c_{4}nh_{n}^{3}).$$

Ceci achève la démonstration du lemme 2.8.

Rappelons que

i) 
$$\tilde{S}_{l}^{\pm}(t) = \frac{1}{nh_{n}^{l+1}} \sum_{i=1}^{n} (t - X_{i})^{l} K_{\pm}(\frac{t - X_{i}}{h_{n}}),$$
  
ii)  $\tilde{\xi}_{l}^{\pm}(t) = \frac{1}{nh_{n}^{l+1}} \sum_{i=1}^{n} (t - X_{i})^{l} K_{\pm}(\frac{t - X_{i}}{h_{n}}) \sigma(X_{i}) \varepsilon_{i},$   
iii)  $\tilde{R}_{l}^{\pm}(t) = \frac{1}{nh_{n}^{l+1}} \sum_{i=1}^{n} (t - X_{i})^{l} K_{\pm}(\frac{t - X_{i}}{h_{n}}) \left[ m(X_{i}) - m_{\pm}(t) - (X_{i} - t)m_{\pm}^{(1)}(t) \right].$ 

**Lemme 2.10.** Supposons que les hypothèses B7-B12 soient satisfaites, alors il existe  $c_5 > 0$  et  $c_6 > 0$  telles que pour  $h_n = n^{-\delta}$  avec  $1/5 < \delta < 1/3$  nous avons

a) 
$$\mathbb{P}\left[\sup_{t\in[0,1]} |\tilde{S}_{l}^{\pm}(t) - K_{l}^{\pm}f_{X}(t)| \ge c_{5}h_{n}\right] \le \exp(-c_{6}nh_{n}^{3}).$$
  
b)  $\mathbb{P}\left[\sup_{t\in[0,1]} |\tilde{\xi}_{l}^{\pm}(t)| \ge c_{5}h_{n}\right] \le \exp(-c_{6}nh_{n}^{3}).$ 

c) 
$$\operatorname{IP}_{\operatorname{HO}}\left[\sup_{t\in[0,1]} |\tilde{R}_l^{\pm}(t)| \ge c_5 h_n^2\right] \le \exp(-c_6 n h_n)$$

Nous en déduisons que si les hypothèses B7-B12 soient satisfaites, alors il existe  $c_5 > 0$  et  $c_6 > 0$ telles que pour  $h_n = n^{-\delta}$  avec  $1/5 < \delta < 1/3$  nous avons pour tout entier L fixé

$$\begin{aligned} \mathbf{d}) \ & \mathbb{P}\left[\max_{0 \le l \le L} \sup_{t \in [0,1]} |\tilde{S}_l^{\pm}(t) - K_l^{\pm} f_X(t)| \ge c_5 h_n\right] \le \exp(-c_6 n h_n^3). \\ \mathbf{e}) \ & \mathbb{P}\left[\max_{0 \le l \le L} \sup_{t \in [0,1]} |\tilde{\xi}_l^{\pm}(t)| \ge c_5 h_n\right] \le \exp(-c_6 n h_n^3). \\ \mathbf{f}) \ & \mathbb{P}_{\mathrm{HO}}\left[\max_{0 \le l \le L} \sup_{t \in [0,1]} |\tilde{R}_l^{\pm}(t)| \ge c_5 h_n^2\right] \le \exp(-c_6 n h_n). \end{aligned}$$

Avant de donner la preuve de ce lemme, nous rappelons l'énoncé de l'inégalité de Bernstein pour préciser la version que nous utilisons.

Lemme 2.11 (Inégalité de Bernstein) Soient  $\xi_1, \ldots, \xi_n$  des variables aléatoires réelles indépendantes définies sur l'espace probabilisé  $(\Omega, \mathcal{A}, P)$ .

1) S'il existe C > 0 tel que pour tout  $k \ge 3$  et pour tout  $1 \le i \le n$ , nous avons  $\mathbb{E}\left[\left|\xi_i - \mathbb{E}\left(\xi_i\right)\right|^k\right] \le C^{k-2}k!Var(\xi_i)$ , alors pour tout  $x \ge 0$  et  $v \ge 2\sum_{i=1}^n Var(\xi_i)$  nous avons

$$\mathbb{IP}\left[\left|\sum_{i=1}^{n} \xi_{i} - \mathbb{IE}(\xi_{i})\right| \ge x\right] \le 2 \exp\left(-\frac{x^{2}}{2(v+Cx)}\right).$$

2) S'il existe C > 0 tel que pour tout 1 ≤ i ≤ n, nous avons IP [|ξ<sub>i</sub>| ≤ C] = 1, alors, pour tout x ≥ 0 et v ≥ ∑<sup>n</sup><sub>i=1</sub> Var(ξ<sub>i</sub>), nous avons

$$\mathbb{P}\left[\left|\sum_{i=1}^{n} \xi_{i} - \mathbb{E}(\xi_{i})\right| \ge x\right] \le 2 \exp\left(-\frac{x^{2}}{2(v+Cx/3)}\right).$$

### Preuve du lemme 2.10.

a) Cherchons des majorants de  $\frac{1}{nh_n^{l+1}}(t-X_i)^l K_{\pm}(\frac{t-X_i}{h_n})$ , et  $Var(\tilde{S}_l^{\pm}(t))$  permettant d'appliquer l'inégalité de Bernstein :

1)

$$\frac{1}{nh_n^{l+1}}(t-X_i)^l K_{\pm}(\frac{t-X_i}{h_n})| \le \frac{3v_1}{2nh_n}$$

2)

$$Var(\tilde{S}_l^{\pm}(t)) \leq \frac{1}{nh_n} \int x^{2l} K_{\pm}^2(x) f_X(t-h_n x) dx,$$
  
$$\leq \frac{v_2}{2nh_n}.$$

d'après l'inégalité de Bernstein (la deuxième partie du lemme), nous avons donc

$$\operatorname{IP}\left[|\tilde{S}_l^{\pm}(t) - \operatorname{IE}(\tilde{S}_l^{\pm}(t))| \ge x\right] \le 2 \exp\left(-\frac{nh_n x^2}{v_2 + v_1 x}\right),$$

or

$$\mathbb{E}(\tilde{S}_l^{\pm}(t)) = \int x^l K_{\pm}(x) f_X(t - h_n x) dx.$$

Sachant que  $f_X(\cdot)$  est lipschitzienne (hypothèse B8), nous en déduisons que

$$|\mathbb{E}(\tilde{S}_{l}^{\pm}(t)) - K_{l}^{\pm}f_{X}(t)| \leq \int |x^{l}K_{\pm}(x)||f_{X}(t - h_{n}x) - f_{X}(t)|dx, \\ \leq v_{3}h_{n}.$$

D'où, comme application de l'inégalité triangulaire, nous avons

$$\begin{aligned} |\tilde{S}_{l}^{\pm}(t) - K_{l}^{\pm}f_{X}(t)| &\leq |\tilde{S}_{l}^{\pm}(t) - \mathbb{E}(\tilde{S}_{l}^{\pm}(t))| + |\mathbb{E}(\tilde{S}_{l}^{\pm}(t)) - K_{l}^{\pm}f_{X}(t)|, \\ &\leq |\tilde{S}_{l}^{\pm}(t) - \mathbb{E}(\tilde{S}_{l}^{\pm}(t))| + v_{3}h_{n}. \end{aligned}$$

il en résulte que pour tout  $x \geq v_3 h_n$  nous avons,

$$\mathbb{P}\left[|\tilde{S}_{l}^{\pm}(t) - K_{l}^{\pm}f_{X}(t)| \ge x\right] \le 2\exp\left(-\frac{nh_{n}(x - v_{3}h_{n})^{2}}{v_{2} + v_{1}(x - v_{3}h_{n})}\right).$$

En choisissant  $x = v_4 h_n > v_3 h_n$  et  $h_n = n^{-\delta}$  avec  $0 < \delta < 1/3$ , l'inégalité précédente devient, pour tout t fixé,

$$\mathbb{P}\left[ |\tilde{S}_{l}^{\pm}(t) - K_{l}^{\pm}f_{X}(t)| \ge v_{4}h_{n} \right] \le 2 \exp\left(-\frac{nh_{n}^{3}(v_{4} - v_{3})^{2}}{v_{2} + v_{1}(v_{4} - v_{3})h_{n}}\right), \\
\le \exp(-v_{5}nh_{n}^{3}).$$

Nous avons

$$[0,1] \subset \bigcup_{k=1}^{\lceil 1/h_n^3 \rceil} [(k-1)h_n^3, kh_n^3],$$

où [x] désigne la partie entière supérieure de x.

Pour  $t \in [(k-1)h_n^3, kh_n^3]$ , nous avons

$$\begin{split} &|\tilde{S}_{l}^{\pm}(t) - K_{l}^{\pm}f_{X}(t) - \tilde{S}_{l}^{\pm}(kh_{n}^{3}) + K_{l}^{\pm}f_{X}(kh_{n}^{3})| \\ &\leq |K_{l}^{\pm}||f_{X}(t) - f_{X}(kh_{n}^{3})| + |\tilde{S}_{l}^{\pm}(t) - \tilde{S}_{l}^{\pm}(kh_{n}^{3})|, \\ &\leq v_{6}|t - kh_{n}^{3}| + \frac{\sum_{i=1}^{n} \left| \left(\frac{t - X_{i}}{h_{n}}\right)^{l} K_{\pm} \left(\frac{t - X_{i}}{h_{n}}\right) - \left(\frac{kh_{n}^{3} - X_{i}}{h_{n}}\right)^{l} K_{\pm} \left(\frac{kh_{n}^{3} - X_{i}}{h_{n}}\right) \right|}{nh_{n}}, \\ &\leq v_{6}|t - kh_{n}^{3}| + v_{7} \frac{|t - kh_{n}^{3}|}{h_{n}^{2}}, \\ &\leq v_{8}h_{n}. \end{split}$$

Pour obtenir cette majoration, nous avons utilisé le fait que  $K_+(\cdot)$  et  $f_X(\cdot)$  sont lipschitziennes (B8 et B12).

D'où

$$\sup_{t \in [0,1]} |\tilde{S}_l^{\pm}(t) - K_l^{\pm} f_X(t)| \le \max_{1 \le k \le \lfloor 1/h_n^3 \rfloor} |\tilde{S}_l^{\pm}(kh_n^3) - K_l^{\pm} f_X(kh_n^3)| + v_8h_n,$$

et par conséquent

$$\begin{split} & \mathbb{P}\left[\sup_{t\in[0,1]} |\tilde{S}_{l}^{\pm}(t) - K_{l}^{\pm}f_{X}(t)| \geq v_{5}h_{n}\right] \\ & \leq \mathbb{P}\left[\max_{1\leq k\leq \lfloor 1/h_{n}^{3}\rfloor} |\tilde{S}_{l}^{\pm}(kh_{n}^{3}) - K_{l}^{\pm}f_{X}(kh_{n}^{3})| \geq (v_{5} - v_{8})h_{n}\right], \\ & \leq \sum_{i=1}^{\lfloor 1/h_{n}^{3}\rfloor} \mathbb{P}\left[|\tilde{S}_{l}^{\pm}(kh_{n}^{3}) - K_{l}^{\pm}f_{X}(kh_{n}^{3})| \geq (v_{5} - v_{8})h_{n}\right], \\ & \leq \frac{1}{h_{n}^{3}} \exp(-v_{9}nh_{n}^{3}), \\ & \leq \exp(-c_{6}nh_{n}^{3}), \end{split}$$

donc il existe  $c_5 > 0$  et  $c_6 > 0$  telles que

$$\operatorname{IP}\left[\sup_{t\in[0,1]}|\tilde{S}_l^{\pm}(t)-K_l^{\pm}f_X(t)|\geq c_5h_n\right] \leq \exp(-c_6nh_n^3).$$

b) Nous avons:

1)

$$\begin{split} & \mathbb{E}\left[\left|\frac{1}{nh_{n}^{l+1}}(t-X_{i})^{l}K_{\pm}(\frac{t-X_{i}}{h_{n}})\right|^{k}\sigma^{k}(X_{i})|\varepsilon_{i}|^{k}\right] \\ & \leq \frac{1}{(nh_{n})^{k}}h_{n}\int|x^{l}K_{\pm}(x)|^{k}\sigma^{k}(t-xh_{n})\mathbb{E}|\varepsilon|^{k}f_{X}(t-xh_{n})dx \,, \\ & \leq \frac{1}{(nh_{n})^{k}}h_{n}v_{10}^{k}\mathbb{E}|\varepsilon|^{k}, \\ & \leq \frac{1}{(nh_{n})^{k}}h_{n}v_{10}^{k}C^{k-2}k!, \\ & \leq \frac{v_{11}^{k-2}}{(nh_{n})^{k-2}}k!, \end{split}$$

2)

$$\begin{aligned} \operatorname{Var}(\tilde{\xi}_{l}^{\pm}(t)) &\leq \frac{1}{nh_{n}} \int x^{2l} K_{\pm}^{2}(x) \sigma^{2}(t-xh_{n}) f_{X}(t-h_{n}x) dx, \\ &\leq \frac{1/2v_{12}}{nh_{n}}, \end{aligned}$$

d'après l'inégalité de Bernstein (la première partie du lemme) nous avons donc

$$\mathbb{P}\left[|\tilde{\xi}_l^{\pm}(t)| \ge x\right] \le 2 \exp\left(-\frac{nh_n x^2}{2(v_{12}+v_{11}x)}\right).$$

En utilisant de nouveau le fait que

$$[0,1] \subset \bigcup_{k=1}^{\lceil 1/h_n^3 \rceil} [(k-1)h_n^3, kh_n^3],$$

nous avons alors pour  $t \in [(k-1)h_n^3, kh_n^3]$ ,

$$\begin{split} &|\tilde{\xi}_{l}^{\pm}(t) - \tilde{\xi}_{l}^{\pm}(kh_{n}^{3})| \\ &\leq \sum_{i=1}^{n} \left| \left( \frac{t - Xi}{h_{n}} \right)^{l} K_{\pm} \left( \frac{t - X_{i}}{h_{n}} \right) - \left( \frac{kh_{n}^{3} - Xi}{h_{n}} \right)^{l} K_{\pm} \left( \frac{kh_{n}^{3} - X_{i}}{h_{n}} \right) \right| \frac{\sigma(X_{i})|\varepsilon_{i}|}{nh_{n}}, \\ &\leq v_{13} \frac{|t - kh_{n}^{3}|}{nh_{n}^{2}} \sum_{i=1}^{n} |\varepsilon_{i}|, \\ &\leq v_{13}h_{n} \frac{\sum_{i=1}^{n} |\varepsilon_{i}|}{n}, \end{split}$$

 $\operatorname{donc}$ 

$$\sup_{t \in [0,1]} |\tilde{\xi}_l^{\pm}(t)| \le \max_{1 \le k \le \lceil 1/h_n^3 \rceil} |\tilde{\xi}_l^{\pm}(kh_n^3)| + v_{13}h_n \frac{\sum_{i=1}^n |\varepsilon_i|}{n},$$

et par conséquent

$$\begin{split} & \mathbb{P}\left[\sup_{t\in[0,1]}|\tilde{\xi}_{l}^{\pm}(t)|\geq x\right] \\ &\leq \mathbb{P}\left[\max_{1\leq k\leq \lceil 1/h_{n}^{3}\rceil}|\tilde{\xi}_{l}^{\pm}(kh_{n}^{3})|+v_{13}h_{n}\frac{\sum_{i=1}^{n}|\varepsilon_{i}|}{n}\geq x\right], \\ &\leq \mathbb{P}\left[\max_{1\leq k\leq \lceil 1/h_{n}^{3}\rceil}|\tilde{\xi}_{l}^{\pm}(kh_{n}^{3})|+v_{13}h_{n}\frac{\sum_{i=1}^{n}|\varepsilon_{i}|}{n}\geq x\ ;\ \frac{\sum_{i=1}^{n}|\varepsilon_{i}|}{n}\leq 2\mathbb{E}|\varepsilon|\right] \\ &+ \mathbb{P}\left[\max_{1\leq k\leq \lceil 1/h_{n}^{3}\rceil}|\tilde{\xi}_{l}^{\pm}(kh_{n}^{3})|+v_{13}h_{n}\frac{\sum_{i=1}^{n}|\varepsilon_{i}|}{n}\geq x\ ;\ \frac{\sum_{i=1}^{n}|\varepsilon_{i}|}{n}>2\mathbb{E}|\varepsilon|\right], \\ &\leq \mathbb{P}\left[\max_{1\leq k\leq \lceil 1/h_{n}^{3}\rceil}|\tilde{\xi}_{l}^{\pm}(kh_{n}^{3})|\geq x-2v_{13}\mathbb{E}|\varepsilon|h_{n}\right]+\mathbb{P}\left[\frac{\sum_{i=1}^{n}(|\varepsilon_{i}|-\mathbb{E}|\varepsilon_{i}|)}{n}>\mathbb{E}|\varepsilon|\right]. \end{split}$$

Puisque

$$\begin{split} \mathbb{E}||\varepsilon| - \mathbb{E}|\varepsilon||^k &\leq 2^{k-1} \left[ \mathbb{E}|\varepsilon|^k + (\mathbb{E}|\varepsilon|)^k \right], \\ &\leq 2^k \mathbb{E}|\varepsilon|^k. \end{split}$$

Alors d'après l'inégalité de Bernstein appliquée aux  $\varepsilon_i$  nous avons

$$\mathbb{P}\left[\frac{\sum_{i=1}^{n} |\varepsilon_{i}| - \mathbb{E}|\varepsilon_{i}|}{n} > \mathbb{E}|\varepsilon|\right] \leq \mathbb{P}\left[\left|\sum_{i=1}^{n} |\varepsilon_{i}| - \mathbb{E}|\varepsilon_{i}|\right| > n\mathbb{E}|\varepsilon|\right], \\ \leq \exp(-v_{14}n).$$

En posant  $x = v_{15}h_n \leq (2v_{13}\mathbb{E}|\varepsilon|) h_n$ , nous obtenons ainsi

$$\begin{split} & \mathbb{P}\left[\sup_{t\in[0,1]}|\tilde{\xi}_{l}^{\pm}(t)|\geq v_{15}h_{n}\right] \\ &\leq \mathbb{P}\left[\max_{1\leq k\leq \lceil 1/h_{n}^{3}\rceil}|\tilde{\xi}_{l}^{\pm}(kh_{n}^{3})|\geq (v_{15}-2v_{13}\mathbb{E}|\varepsilon|)h_{n}\right] + \mathbb{P}\left[\frac{\sum_{i=1}^{n}|\varepsilon_{i}|-\mathbb{E}|\varepsilon_{i}|}{n}>\mathbb{E}|\varepsilon|\right], \\ &\leq \sum_{i=1}^{\lceil 1/h_{n}^{3}\rceil}\mathbb{P}\left[|\tilde{\xi}_{l}^{\pm}(kh_{n}^{3})|\geq (v_{15}-2v_{13}\mathbb{E}|\varepsilon|)h_{n}\right] + \exp(-v_{14}n), \\ &\leq \exp(-c_{6}nh_{n}^{3}), \end{split}$$

donc il existe  $c_5 > 0$  et  $c_6 > 0$  telles que

$$\mathbb{P}\left[\sup_{t\in[0,1]}|\tilde{\xi}_l^{\pm}(t)| \ge c_5 h_n\right] \le \exp(-c_6 n h_n^3).$$

c) Nous avons

1)

$$\left|\frac{(t-X_i)^l}{nh_n^{l+1}}K_{\pm}(\frac{t-X_i}{h_n})\left[m(X_i)-m_{\pm}(t)-(X_i-t)m_{\pm}^{(1)}(t)\right]\right| \leq \frac{3v_1h_n}{2n},$$

2)

$$\begin{aligned} Var(\tilde{S}_{l}^{\pm}(t)) &\leq \frac{1}{nh_{n}} \int x^{2l} K_{\pm}^{2}(x) \left[ m(t-xh_{n}) - m_{\pm}(t) + xh_{n} m_{\pm}^{(1)}(t) \right]^{2} f_{X}(t-h_{n}x) dx \,, \\ &\leq \frac{v_{2}h_{n}^{3}}{2n}. \end{aligned}$$

Appliquons l'inégalité de Bernstein (la deuxième partie du lemme), nous avons donc

$$\mathbb{P}\left[|\tilde{R}_l^{\pm}(t) - \mathbb{E}(\tilde{R}_l^{\pm}(t))| \ge x\right] \le 2 \exp\left(-\frac{nx^2}{v_2 h_n^3 + v_1 x h_n}\right),$$

or nous avons

$$\mathbb{E}(\tilde{R}_{l}^{\pm}(t)) = \int x^{l} K_{\pm}(x) \left[ m(t - xh_{n}) - m_{\pm}(t) + xh_{n}m_{\pm}^{(1)}(t) \right] f_{X}(t - h_{n}x) dx,$$

donc sous l'hypothèse nulle

$$|\mathbb{E}(\tilde{R}_l^{\pm}(t))| \leq v_3 h_n^2.$$

Comme

$$\begin{aligned} |\tilde{R}_{l}^{\pm}(t)| &\leq |\tilde{R}_{l}^{\pm}(t) - \mathbb{E}(\tilde{R}_{l}^{\pm}(t))| + |\mathbb{E}(\tilde{R}_{l}^{\pm}(t))|, \\ &\leq |\tilde{R}_{l}^{\pm}(t) - \mathbb{E}(\tilde{R}_{l}^{\pm}(t))| + v_{3}h_{n}^{2}. \end{aligned}$$

Il en résulte que pour tout  $x \ge v_3 h_n^2$  nous avons

$$\mathbb{P}\left[|\tilde{R}_{l}^{\pm}(t)| \ge x\right] \le 2\exp\left(-\frac{n(x-v_{3}h_{n}^{2})^{2}}{v_{2}h_{n}^{3}+v_{1}h_{n}(x-v_{3}h_{n}^{2})}\right)$$

Si on prend  $x = v_4 h_n^2 > v_3 h_n^2$  et  $h_n = n^{-\delta}$  avec  $0 < \delta < 1/3$ , alors pour tout t fixé nous avons

$$\mathbb{P}\left[|\tilde{R}_{l}^{\pm}(t)| \geq v_{4}h_{n}^{2}\right] \leq 2\exp\left(-\frac{nh_{n}(v_{4}-v_{3})^{2}}{v_{2}+v_{1}(v_{4}-v_{3})}\right), \\ \leq \exp(-v_{5}nh_{n}).$$

En outre

$$[0,1] \subset \bigcup_{k=1}^{\lceil 1/h_n^3 \rceil} [(k-1)h_n^3, kh_n^3]$$

où  $\lceil x \rceil$  désigne la partie entière supérieure de x. Pour  $t \in [(k-1)h_n^3, kh_n^3]$ , nous avons

$$\begin{split} \tilde{R}_{l}^{\pm}(t) &- \tilde{R}_{l}^{\pm}(kh_{n}^{3}) \\ &= \frac{1}{nh_{n}} \sum_{i=1}^{n} \left(\frac{t-Xi}{h_{n}}\right)^{l} K_{\pm} \left(\frac{t-X_{i}}{h_{n}}\right) \left[m(X_{i}) - m_{\pm}(t) - (X_{i}-t)m_{\pm}^{(1)}(t)\right] \\ &- \frac{1}{nh_{n}} \sum_{i=1}^{n} \left(\frac{kh_{n}^{3} - Xi}{h_{n}}\right)^{l} K_{\pm} \left(\frac{kh_{n}^{3} - X_{i}}{h_{n}}\right) \left[m(X_{i}) - m_{\pm}(kh_{n}^{3}) - (X_{i}-kh_{n}^{3})m_{\pm}^{(1)}(kh_{n}^{3})\right]. \end{split}$$

Or

$$m(X_i) - m_{\pm}(kh_n^3) - (X_i - kh_n^3)m_{\pm}^{(1)}(kh_n^3)$$
  
=  $\left[m(X_i) - m_{\pm}(t) - (X_i - t)m_{\pm}^{(1)}(t)\right]$   
+  $\left[m_{\pm}(t) - m_{\pm}(kh_n^3)\right] + (X_i - t)\left[m_{\pm}^{(1)}(t) - m_{\pm}^{(1)}(kh_n^3)\right]$   
+  $(kh_n^3 - t)m_{\pm}^{(1)}(kh_n^3),$   
=  $\left[m(X_i) - m_{\pm}(t) - (X_i - t)m_{\pm}^{(1)}(t)\right] + \mathcal{O}(h_n^3),$ 

d'où

$$\begin{split} \tilde{R}_{l}^{\pm}(t) &- \tilde{R}_{l}^{\pm}(kh_{n}^{3}) \\ &= \frac{1}{nh_{n}} \sum_{i=1}^{n} \left[ \left( \frac{t - Xi}{h_{n}} \right)^{l} K_{\pm} \left( \frac{t - X_{i}}{h_{n}} \right) - \left( \frac{kh_{n}^{3} - Xi}{h_{n}} \right)^{l} K_{\pm} \left( \frac{kh_{n}^{3} - X_{i}}{h_{n}} \right) \right] \\ &\times \left[ m(X_{i}) - m_{\pm}(t) - (X_{i} - t)m_{\pm}^{(1)}(t) \right] + \mathcal{O}(h_{n}^{2}), \\ &= \mathcal{O}(h_{n}^{2}), \end{split}$$

 $\operatorname{donc}$ 

$$\sup_{t \in [0,1]} |\tilde{R}_l^{\pm}(t)| \le \max_{1 \le k \le \lceil 1/h_n^3 \rceil} |\tilde{R}_l^{\pm}(kh_n^3)| + v_8h_n^2,$$

et par suite

$$\mathbb{P}\left[\sup_{t\in[0,1]} |\tilde{R}_{l}^{\pm}(t)| \geq v_{5}h_{n}^{2}\right] \leq \mathbb{P}\left[\max_{1\leq k\leq \lceil 1/h_{n}^{3}\rceil} |\tilde{R}_{l}^{\pm}(kh_{n}^{3})| \geq (v_{5}-v_{8})h_{n}^{2}\right], \\
\leq \sum_{i=1}^{\lceil 1/h_{n}^{3}\rceil} \mathbb{P}\left[|\tilde{R}_{l}^{\pm}(kh_{n}^{3})| \geq (v_{5}-v_{8})h_{n}^{2}\right], \\
\leq \frac{1}{h_{n}^{3}} \exp(-v_{9}nh_{n}), \\
\leq \exp(-c_{6}nh_{n}),$$

d'où le résultat : il existe  $c_5 > 0$  et  $c_6 > 0$  telles que

$$\mathbb{P}\left[\sup_{t\in[0,1]}|\tilde{R}_l^{\pm}(t)| \ge c_5 h_n\right] \le \exp(-c_6 n h_n).$$

d) Les démonstrations des parties d), e) et f) sont de simples déductions des parties a), b) et
c). On fournit ici seulement la démonstration de la partie d), les autres sont similaires en tous points à celle-ci.

$$\mathbb{P}\left[\max_{0\leq l\leq L}\sup_{t\in[0,1]}|\tilde{S}_{l}^{\pm}(t)-K_{l}^{\pm}f_{X}(t)|\geq c_{5}h_{n}\right] \leq \sum_{l=0}^{L}\mathbb{P}\left[\sup_{t\in[0,1]}|\tilde{S}_{l}^{\pm}(t)-K_{2}^{\pm}f_{X}(t)|\geq c_{5}h_{n}\right], \\ \leq (L+1)\exp(-c_{6}nh_{n}^{3}), \\ \leq \exp(-c_{7}nh_{n}^{3}).$$

### Preuve du théorème 2.8.

La démonstration du théorème 2.8 est fondée principalement sur le résultat établi par Johnston [JOH82](cf aussi Bickel et Rosenblatt [BR73]) pour la loi limite de la déviation maximale de l'estimateur de Nadaraya-Watson de la fonction de régression.

$$\begin{split} & \mathbb{P}_{\mathrm{H0}} \left[ a_n \left| \sup_{t \in [0,1]} |\xi_n(t)| - \sup_{t \in [0,1]} |\zeta_n(t)| \right| \ge 2c_1 a_n (nh_n^5)^{1/4} \right], \\ & \le \mathbb{P}_{\mathrm{H0}} \left[ a_n \sup_{t \in [0,1]} |\xi_n(t) - \zeta_n(t)| \ge 2c_1 a_n (nh_n^5)^{1/4} \right], \\ & \le \mathbb{P}_{\mathrm{H0}} \left[ a_n \left( \sup_{t \in [0,1]} |B_n(t) - \zeta_n(t)| + \sup_{t \in [0,1]} |E_n(t)| \right) \ge 2c_1 a_n (nh_n^5)^{1/4} \right], \\ & \le \mathbb{P}_{\mathrm{H0}} \left[ a_n \sup_{t \in [0,1]} |B_n(t) - \zeta_n(t)| \ge c_1 a_n (nh_n^5)^{1/4} \right] + \mathbb{P}_{\mathrm{H0}} \left[ a_n \sup_{t \in [0,1]} |E_n(t)| \ge c_1 a_n (nh_n^5)^{1/4} \right], \\ & \le 2 \exp(-c_2 nh_n^3). \end{split}$$

Lemme 2.12 (Konakov & Piterbarg [PK84]) Soient X et Y deux variables aléatoires réelles telles qu'il existe  $\epsilon_1 > 0$  et  $\epsilon_2 > 0$  pour lesquels nous avons:

$$\mathbb{P}\left[|X - Y| \ge \epsilon_1\right] \le \epsilon_2,$$

alors pour tout x

$$\mathbb{P}\left[Y \le x - \epsilon_1\right] - \epsilon_2 \le \mathbb{P}\left[X \le x\right] \le \mathbb{P}\left[Y \le x + \epsilon_1\right] + \epsilon_2.$$

Posons:

i) 
$$\epsilon_{n,1} = 2c_1 a_n (nh_n^5)^{1/4} \to 0,$$
  
ii)  $\epsilon_{n,2} = 2 \exp(-c_2 nh_n^3) \to 0.$ 

D'après le lemme 2.12 nous avons donc pour tout x

$$\operatorname{IP}_{\operatorname{HO}}\left[a_n\left(\sup_{t\in[0,1]}|\xi_n|-b_n\right)\leq x\right] \geq \operatorname{IP}_{\operatorname{HO}}\left[a_n\left(\sup_{t\in[0,1]}|\zeta_n|-bn\right)\leq x-\epsilon_{n,1}\right]-\epsilon_{n,2},$$

 $\operatorname{et}$ 

$$\mathbb{P}_{\mathrm{HO}}\left[a_n\left(\sup_{t\in[0,1]}|\xi_n|-b_n\right)\leq x\right] \leq \mathbb{P}_{\mathrm{HO}}\left[a_n\left(\sup_{t\in[0,1]}|\zeta_n|-b_n\right)\leq x+\epsilon_{n,1}\right]+\epsilon_{n,2}$$

Pour la même raison nous obtenons également

$$\mathbb{P}_{\mathrm{HO}}\left[a_n\left(\sup_{t\in[0,1]}\xi_n-b_n\right)\leq x\right] \geq \mathbb{P}_{\mathrm{HO}}\left[a_n\left(\sup_{t\in[0,1]}\zeta_n-bn\right)\leq x-\epsilon_{n,1}\right]-\epsilon_{n,2},$$

 $\operatorname{et}$ 

$$\mathbb{P}_{\mathrm{HO}}\left[a_n\left(\sup_{t\in[0,1]}\xi_n-b_n\right)\leq x\right] \leq \mathbb{P}_{\mathrm{HO}}\left[a_n\left(\sup_{t\in[0,1]}\zeta_n-b_n\right)\leq x+\epsilon_{n,1}\right]+\epsilon_{n,2}.$$

Or d'après Johnston [JOH82], sous les hypothèses B7-B12, on peut approcher le processus  $\zeta_n$  par un processus gaussien  $G_n(t) = \frac{1}{\sqrt{2V_+^2h_n}} \int M\left(\frac{t-x}{h_n}\right) dW(x)$ , où W est le processus de Wiener bilatéral, tel que

$$\sup_{t \in [0,1]} |\zeta_n(t) - G_n(t)| = o_P\left( (\log n)^{-1/2} \right).$$

Nous avons aussi grâce à la propriété d'invariance en loi par changement d'échelle du processus de Wiener :

$$\mathcal{L}\left\{G_{n}(t) ; 0 \leq t \leq 1\right\} = \mathcal{L}\left\{\frac{1}{\sqrt{2V_{+}^{2}h_{n}}}\int M\left(\frac{t-x}{h_{n}}\right)dW(x) ; 0 \leq t \leq 1\right\},\$$
$$= \mathcal{L}\left\{\frac{1}{\sqrt{2}V_{+}}\int M\left(\frac{t}{h_{n}}-x\right)dW(x) ; 0 \leq t \leq 1\right\},\$$

où  $\mathcal{L}(X)$  désigne la loi de X.

Posons  $T_n = \frac{1}{h_n}$ , d'où

1)

$$\mathcal{L}\left\{\sup_{t\in[0,1]}G_n(t)\right\} = \mathcal{L}\left\{\sup_{t\in[0,1]}\frac{1}{\sqrt{2}V_+}\int M\left(\frac{t}{h_n}-x\right)dW(x)\right\},\$$
$$= \mathcal{L}\left\{\sup_{t\in[0,T_n]}\frac{1}{\sqrt{2}V_+}\int M\left(t-x\right)dW(x)\right\}.$$

2)

$$\mathcal{L}\left\{\sup_{t\in[0,1]}|G_n(t)|\right\} = \mathcal{L}\left\{\sup_{t\in[0,1]}\left|\frac{1}{\sqrt{2}V_+}\int M\left(\frac{t}{h_n}-x\right)dW(x)\right|\right\},\$$
$$= \mathcal{L}\left\{\sup_{t\in[0,T_n]}\left|\frac{1}{\sqrt{2}V_+}\int M\left(t-x\right)dW(x)\right|\right\}.$$

 $\int M(t-x) dW(x)$  est un processus gaussien stationnaire centré de fonction de covariance  $C(\cdot)$  définie par :

$$C(h) = \int M(x+h)M(x)dx \; .$$

Or nous avons montré dans le paragraphe précédent que la fonction de covariance  $C(\cdot)$  vérifie bien les hypothèses du théorème 2.4, nous en déduisons donc :

1) 
$$\lim_{n \to +\infty} \operatorname{IP}\left[a_n \left(\sup_{t \in [0,1]} G_n(t) - b_n\right) \le x\right] = \exp(-\exp(-x)),$$
  
2) 
$$\lim_{n \to +\infty} \operatorname{IP}\left[a_n \left(\sup_{t \in [0,1]} |G_n(t)| - b_n\right) \le x\right] = \exp(-2\exp(-x))$$

où  $a_n = a_{T_n}$  et  $b_n = b_{T_n}$ .

Par conséquent, sous l'hypothèse nulle, nous avons

a)

$$\lim_{n \to +\infty} \mathbb{P}\left[a_n \left(\sup_{t \in [0,1]} \xi_n(t) - b_n\right) \le x\right] = \lim_{n \to +\infty} \mathbb{P}\left[a_n \left(\sup_{t \in [0,1]} \zeta_n(t) - b_n\right) \le x\right],$$
$$= \lim_{n \to +\infty} \mathbb{P}\left[a_n \left(\sup_{t \in [0,1]} G_n(t) - b_n\right) \le x\right],$$
$$= \exp(-\exp(-x)).$$

b)

$$\lim_{n \to +\infty} \mathbb{P}\left[a_n \left(\sup_{t \in [0,1]} |\xi_n(t)| - b_n\right) \le x\right] = \lim_{n \to +\infty} \mathbb{P}\left[a_n \left(\sup_{t \in [0,1]} |\zeta_n(t)| - b_n\right) \le x\right],$$
$$= \lim_{n \to +\infty} \mathbb{P}\left[a_n \left(\sup_{t \in [0,1]} |G_n(t)| - b_n\right) \le x\right],$$
$$= \exp(-2\exp(-x)).$$

Ceci achève la preuve du théorème 2.8.

Il est facile de montrer que le test global de l'hypothèse d'absence de rupture, basé sur le théorème 2.8 a un comportement asymptotique satisfaisant. Ici également on considère des alternatives locales de la forme :

$$\operatorname{H1}(n) : \gamma_n(\tau) = \frac{\delta_n}{\sqrt{nh_n}},$$

avec  $\frac{|\delta_n|}{\sqrt{\log(n)}} \to \infty$ . Le théorème suivant montre que le test global, lui aussi, a une puissance locale asymptotique égale à 1.

### Preuve du théorème 2.9.

Pour tout u > 0,

$$\mathbb{P}_{\mathtt{H1}(n)}\left[a_n\left(\sup_{t\in[0,1]}|\xi_n(t)|-b_n\right)\leq u\right] \leq \mathbb{P}_{\mathtt{H1}(n)}\left[a_n\left(|\xi_n(\tau)|-b_n\right)\leq u\right].$$

Notons

$$E_{n,i}^+ = \left[\sqrt{nh_n}V(\tau)\left(\hat{\gamma}_n(\tau) - \gamma_n(\tau)\right) \ge -\left(\frac{u}{a_n} + b_n\right) - \delta_n(\tau)V(\tau)\right],$$

 $\operatorname{et}$ 

$$E_{n,i}^{-} = \left[\sqrt{nh_n}V(\tau)\left(\hat{\gamma}_n(\tau) - \gamma_n(\tau)\right) \le \left(\frac{u}{a_n} + b_n\right) - \delta_n(\tau)V(\tau)\right],$$

d'où

$$\lim_{n \to +\infty} \mathbb{P}_{\mathtt{H1}(n)} \left[ a_n \left( \sup_{t \in [0,1]} |\xi_n(t)| - b_n \right) \le u \right] \le \lim_{n \to +\infty} \mathbb{P}_{\mathtt{H1}(n)} \left[ E_{n,i}^+ \cap E_{n,i}^- \right],$$
$$= 0.$$

# 2.5 Simulations

Dans cette section, nous effectuons des simulations pour mettre en évidence le comportement des trois tests pour des tailles échantillons raisonnables (n = 250, n = 500). Les facteurs déterminants dans ces simulations sont le choix du noyau, le choix de la fenêtre de lissage, le nombre d'observations et le rapport signal sur bruit. Pour nos comparaisons, nous avons retenu quatre modèles MO<sub>1</sub>, MO<sub>2</sub> (modèles sans rupture), M1<sub>1</sub> et M1<sub>2</sub> (modèles avec rupture):

$$\mathsf{MO}_1 \ m(x) = 4x^2,$$

- $M1_1 \ m(x) = 4x^2 + 1 I_{[0.5,1]},$
- $MO_2 \ m(x) = \exp(2|x 0.35|),$
- M1<sub>2</sub>  $m(x) = \exp(2|x 0.35|) + \prod_{[0.35,1]}$

Les noyaux K0, K1 et K2 utilisés sont définis par :

- 1)  $K0(x) = x^2(1+x)^2 \prod_{[-1,0]}(x).$
- 2)  $K1(x) = (1 x^2) \prod_{[-1,0]} (x).$
- 3)  $K2(x) = -x(1+x) \coprod_{[-1,0]}(x).$

K0 et K2 s'annulent en zéro .K1>0 en zéro.

Lors de ces simulations, le plan d'expérience est aléatoire de loi uniforme et le bruit est supposé gaussien. Dans chaque test et pour chaque valeur du paramètre concerné (taille de la fenêtre de lissage, taille d'échantillon, rapport signal sur bruit) nous avons effectué 100 simulations, et relevé la fréquence avec laquelle l'hypothèse nulle est acceptée obtenant ainsi une estimation de l'erreur de première espèce (pour les modèles  $MO_1$  et  $MO_2$ ) et l'erreur de deuxième espèce (pour les modèles  $M1_1$  et  $M1_2$ ). Les tests sont construits de manière à assurer une erreur de première espèce théorique asymptotique  $\alpha = 5\%$ . Pour le test strictement local et le test local, la discontinuité est localisée en  $\tau = 0.5$  pour les modèles  $MO_1$  et  $M1_1$  et en  $\tau = 0.35$  pour les modèles  $MO_2$  et  $M1_2$ .

**Choix du noyau.** Les figures 2.1 et 2.3 montrent que les trois tests sont légèrement sensibles au choix du noyau en particulier en ce qui concerne leur puissance. Rappelons que la situation où le noyau s'annule en zéro et celle où le noyau est strictement positif à l'origine conduisent à deux traitements différents pour le test local et le test global (deux normalisations différentes). Dans la suite, nous avons choisi seulement de présenter les résultats pour le noyau K2; les simulations

ont été effectuées pour les trois types de noyau et nous avons constaté la même sensibilité de ces trois tests vis à vis du choix du noyau.

Choix de la fenêtre de lissage. Tout au long de nos simulations, nous avons choisi de faire varier  $h_n$  entre 0.07 et 0.25 avec un pas de 0.02 (10 valeurs de  $h_n$  en tout). Nous remarquons que les trois tests sont plus sensibles au choix de la fenêtre de lissage qu'à celui du noyau. Cette sensibilité est quasi nulle en ce qui concerne l'erreur de première espèce (2.1, 2.2 et 2.5)mais elle devient cependant plus importante lorsque la puissance des tests est considérée. Le test local semble être le plus conservateur des trois.

Nombre d'observations et RSB. Les simulations dont les résultats sont présentés par les figures 2.5 et 2.6) montrent la sensibilité des trois procédures de test à la taille de l'échantillon et au rapport signal sur bruit. En absence de discontinuité les procédures apparaissent peu sensibles à la taille de l'échantillons et au rapport signal sur bruit. Par contre en situation de discontinuité, comme on s'y attend la discontinuité est d'autant mieux détectée que le bruit est faible par rapport à l'amplitude de la discontinuité, et que la taille de l'échantillon est grande.

**Étude comparative des trois tests (TSL, TL et TG).** Sur les exemples considérés ici ainsi que sur d'autres exemples, nous avons constaté que pour des tailles déchantillons raisonnables, le test strictement local (TSL) est plus performant que le test local (TL) ou le test global (TG). Dans les situations où il y a une rupture dans le modèle et où le rapport signal sur bruit est faible, le TSL semble avoir un comportement acceptable alors que les autres tests ont tendance, dans ce cas, à accepter souvent l'hypothèse nulle d'absence de ruptures. Ceci semble cohérent avec l'étude asymptotique qui est faite dans ce chapitre et qui montre que le test strictement local est plus puissant que les deux autres tests.



**Figure 2.1 :** Evolution du pourcentage d'acceptation de l'hypothèse nulle (absence de discontinuité) en fonction de la taille de la fenêtre de lissage pour le modèle  $MO_1$  dans le cas des trois noyaux K0, K1, K2 (n = 250,  $\sigma^2 = 0.1$ ).



**Figure 2.2 :** Evolution du pourcentage d'acceptation de l'hypothèse nulle (absence de discontinuité) en fonction de la taille de la fenêtre de lissage pour le modèle  $MO_1$  dans le cas des trois noyaux K0, K1, K2 (n = 250,  $\sigma^2 = 0.1$ ).



**Figure 2.3 :** Evolution du pourcentage d'acceptation de l'hypothèse nulle (absence de discontinuité) en fonction de la taille de la fenêtre de lissage pour le modèle M1<sub>1</sub> dans le cas des trois noyaux K0, K1, K2 (n = 250,  $\sigma^2 = 0.1$ ).


**Figure 2.4 :** Evolution du pourcentage d'acceptation de l'hypothèse nulle (absence de discontinuité) en fonction de la taille de la fenêtre de lissage pour le modèle M1<sub>1</sub> dans le cas des trois noyaux K0, K1, K2 (n = 250,  $\sigma^2 = 0.1$ ).



**Figure 2.5 :** Evolution du pourcentage d'acceptation de l'hypothèse nulle (absence de discontinuité) en fonction de la taille de la fenêtre de lissage pour le modèle MO<sub>2</sub>. Le noyau utilisé est K2.



**Figure 2.6 :** Evolution du pourcentage d'acceptation de l'hypothèse nulle (absence de discontinuité) en fonction de la taille de la fenêtre de lissage pour le modèle M1<sub>2</sub>. Le noyau utilisé est K2.

# Chapitre 3

# Estimateur du nombre de ruptures

# 3.1 Introduction

L'objet de ce chapitre est de proposer une procédure permettant d'estimer par des estimateurs convergents, à la fois le nombre p de points de rupture et les localisations de ces ruptures. La procédure est basée sur le comportement du processus de saut estimé  $\hat{\gamma}(t)$  sur l'intervalle [0, 1] découpé en intervalles de longueur  $2h_n$ . Considérant le nombre d'intervalles où  $\hat{\gamma}(\cdot)$  dépasse un seuil  $S_0$ , en convenant de compter comme un seul intervalle une famille de tels intervalles dès lors qu'ils sont contigüs, nous obtenons un estimateur naturel de p pour lequel nous montrons qu'il y a bien convergence. De plus, les localisations des dépassements du seuil convergent vers les points de rupture  $\tau_1, \ldots, \tau_p$ . Cette procédure présente l'inconvénient de nécessiter une information a priori pour choisir  $S_0 < \min(\gamma(\tau_1), \ldots, \gamma(\tau_p))$ . Une méthode pour contourner cette difficulté consiste à avoir recours à une suite de seuils  $(S_n)$  convergeant vers 0. Avec des hypothèses un peu plus fortes que dans le situation du seuil  $S_0$  fixé, nous obtenons la convergence p.s. de  $\hat{p}$  vers p et de  $\hat{\tau}_i$  vers  $\tau_i$ , i = 1, ..., p. Pour avoir des résultats sur la vitesse de convergence nous nous plaçons ensuite dans un cadre d'hypothèses un peu différent qui nous permet de montrer que les probabilités d'erreur sur l'estimation de p et d'écarts  $|\hat{\tau}_i - \tau_i|$  supérieurs à  $h_n$ tendent exponentiellement vers 0. Ces résultats conduisent aussi à des énoncés de convergence presque sûre.

## 3.2 Construction de l'estimateur du nombre de ruptures

Soit *p* le nombre de ruptures de la fonction de régression  $m(\cdot)$  dans l'intervalle [0, 1]. On note  $0 < \tau_1 < \ldots < \tau_p < 1$ , les localisations de ces ruptures, et  $\{\gamma(\tau_i) = m_+(\tau_i) - m_-(\tau_i)\}_{i=1,\ldots,p}$  les

amplitudes des sauts.  $m_{+}(t)$  étant la limite à droite et  $m_{-}(t)$  la limite à gauche définies par :

$$m_+(t) = \lim_{x \searrow t, x > t} m(x) \text{ et } m_-(t) = \lim_{x \nearrow t, x < t} m(x).$$

Dans un premier temps on va supposer que  $\min_{i=1,...,p} |\gamma(\tau_i)| > S_0$ . Cette hypothèse sera affaiblie par la suite. Par convention, nous posons

$$\tau_0 = 0$$
 et  $\tau_{p+1} = 1$ .

Introduisons la quantité  $\hat{\gamma}(t)$  définie par :

$$\hat{\gamma}(t) = \hat{m}_{+}(t) - \hat{m}_{-}(t), \quad h_n \le t \le 1 - h_n, \tag{3.1}$$

 $\hat{\gamma}(t)$  est un estimateur du saut éventuel de la fonction de régression au point t. Cette quantité est proche de zéro dans les régions où la fonction  $m(\cdot)$  est lisse. Par contre, elle est grande en module lorsqu'on est au voisinage d'un point de rupture. De ce fait,  $\hat{\gamma}(\cdot)$  est un bon critère de détection de points de rupture.

Notons pour la suite

$$g_j = \frac{j}{G}, \qquad j = 1, 2, \dots, G-1,$$

 $\operatorname{et}$ 

$$G = \left[\frac{1}{2h_n}\right] + 1,$$

(

où [m] est la partie entière de m.

On pose

$$J = \{g_j, \ j = 1, \dots, G - 2 \ | \ || \ \hat{\gamma}(\cdot) \ ||_{[g_j, g_{j+1}]} > S_0\},$$
(3.2)

avec

$$||h(\cdot)||_{[a,b]} = \sup_{t \in [a,b]} |h(t)|.$$

Les  $g_j \in J$  sont des candidats potentiels pour être des points de rupture de la fonction de régression  $m(\cdot)$ . Lorsque  $g_{j_0}$  appartient à J, les  $g_j$  qui sont au voisinage de  $g_{j_0}$  vont vraisemblablement eux aussi appartenir à J. Par conséquent, si on prend comme estimateur de p le cardinal de J, le nombre de ruptures de  $m(\cdot)$  sera probablement surestimé. Afin de parer à cet inconvénient, nous proposons de regrouper les éléments de J dans des classes (une classe regroupe les points de J qui sont proches) et de prendre le nombre de ces classes comme estimateur du nombre de ruptures.

Notons  $\mathcal{C}$  l'ensemble des chaînes maximales de J, c'est-à-dire

$$C = \left\{ C = \{g_j\}_{j=r,\dots,r+k} \mid g_j \in J, j = r,\dots,r+k, \ g_{r-1} \notin J \text{ et } g_{r+k+1} \notin J \right\}$$

C est une partition de J  $(J = \cup_{i=1}^{\hat{p}} C_i)$ . L'estimateur du nombre de discontinuités est alors défini par :

$$\hat{p} = \# \mathcal{C}$$

On définit aussi les estimateurs de localisation des ruptures :

$$\hat{\tau}_i = \arg \max_{\widehat{ID}_i} |\hat{\gamma}(t)|,$$

avec  $i = 1, \ldots, \hat{p}$  et  $\widehat{ID}_i = [\min C_i, \max C_i + \frac{1}{G}].$ 

La figure 3.1 illustre, dans la situation d'un signal sans bruit avec deux ruptures la procédure que nous avons définie. Au voisinage de la première rupture, l'ensemble des points t où  $\hat{\gamma}(t) \geq S_n$ se répartit sur deux intervalles consécutifs  $[g_3, g_4[$  et  $[g_4, g_5[$ . On a donc  $C_1 = \{g_4, g_5\}$ . Au voisinage de la deuxième rupture, ces points sont situés dans un seul intervalle  $[g_{i_0}, g_{i_0+1}[$ , on a donc  $C_2 = \{g_{i_0}\}$ .



Figure 3.1 : Estimation du nombre de ruptures : Algorithme.

Pour continuer, nous aurons besoin de quelques hypothèses de travail.

# 3.3 Hypothèses

- 1) Les hypothèses sur les modèles sont les suivantes :
  - M1. il existe s > 2 tel que  $\mathbb{E}|Y|^s < \infty$ ,
  - M2.  $f_X(\cdot)$  est supposée continue et strictement positive sur l'intervalle [0, 1],
  - M3. les restrictions de la fonction de régression  $m(\cdot)$  aux intervalles  $[\tau_i, \tau_{i+1}]$  sont supposées continues et  $\min_{i=1,\dots,p} |\gamma(\tau_i)| > S_0$ ,
  - M4.  $\min_{i=1,...,p} |\gamma(\tau_i)| > S_0,$
  - M5. les restrictions de la fonction de régression  $m(\cdot)$  aux intervalles  $[\tau_i, \tau_{i+1}]$  sont supposées deux fois continûment dérivables,

Naturellement l'hypothèse M5 est plus forte que M3: Nous préférons garder la distinction afin de préserver dans la mesure du possible la plus grande généralité dans les énoncés. Nous avons recours à M3 dans la situation d'un seuil  $S_0$  fixe, et à M5 lorsqu'on a une suite de seuils  $(S_n)_n$ .

- 2) La fenêtre de lissage  $h_n$  est supposée satisfaire les conditions suivantes :
  - H1.  $h_n \to 0$  et il existe  $\beta < 1 s^{-1}$  tel que  $n^{\beta}h_n \to \infty$ ,
  - H2.  $h_n \to 0$  et il existe  $\beta < 1 s^{-1}$  tel que  $n^{2\beta 1}h_n \to \infty$ ,
  - H3. il existe  $\lambda > 0$  tel que  $\sum_n h_n^\lambda < \infty$ ,
  - H4. Soient  $r_n = \sqrt{\frac{1}{nh_n}\log(\frac{1}{h_n})}$  et  $(S_n)_{n\geq 0}$  une suite de réels positifs telle que  $S_n \to 0$  et  $\max(h_n^2, r_n) = o(S_n)$ ,
- 3) les noyaux de lissage vérifient l'hypothèse suivante
  - K1.  $K_{+}(\cdot)$  est une fonction continue sur [-1,0].

# 3.4 Consistance des estimateurs

Le résultat suivant est à la base des propriétés de convergence de notre estimateur du nombre de ruptures.

Théorème 3.1. Supposons que les hypothèses M1-M4, H1, H3 et K1 soient satisfaites. Alors:

- a)  $||\hat{\gamma}(\cdot)||_{\cup_{i=1}^{p}[\tau_{i}+h_{n},\tau_{i+1}-h_{n}]} = o(1), p.s.$
- b)  $\inf_{1 \le i \le p} \liminf_{n} ||\hat{\gamma}(\cdot)||_{]\tau_i h_n, \tau_i + h_n[} > S_0, \quad p.s.$

D'après la première partie du théorème 3.1, nécessairement  $\hat{\gamma}(\cdot)$  est proche de zéro dans la partie où la fonction de régression  $m(\cdot)$  est lisse. Par contre la deuxième partie montre que  $\hat{\gamma}(\cdot)$  est grand, en module, au voisinage de chaque point de rupture de  $m(\cdot)$ .

Théorème 3.2. Sous les mêmes hypothèses que le théorème 3.1, nous avons :

- a)  $\hat{p} \longrightarrow p$ , p.s.
- b)  $\forall 1 \leq i \leq p, \quad \hat{\tau}_i \longrightarrow \tau_i, \quad p.s.$

En remplaçant dans (3.2)  $S_0$  par une suite réelle positive  $(S_n)_{n\geq 0}$  qui vérifie l'hypothèse H4, on peut obtenir deux théorèmes analogues aux théorèmes 3.1 et 3.2.

**Théorème 3.3.** Supposons que les hypothèses M1, M2, M5, H2-H4 et K1 soient satisfaites. Alors:

- a)  $\|\hat{\gamma}(\cdot)\|_{\bigcup_{i=1}^{p}[\tau_{i+h_{n},\tau_{i+1}-h_{n}]}} = o(S_{n}), \quad p.s.$
- b)  $\inf_{1 \le i \le p} \liminf_{n} ||\hat{\gamma}(\cdot)||_{]\tau_i h_n, \tau_i + h_n[} > S_n, \quad p.s.$

**Théorème 3.4.** Supposons que les hypothèses du théorème 3.3 soient satisfaites. Alors:

- a)  $\hat{p} \longrightarrow p$ , p.s.
- b)  $\forall 1 \leq i \leq p, \qquad \hat{\tau}_i \longrightarrow \tau_i, \quad p.s.$

Comme  $(\hat{p}) = (\hat{p}_n)_{n \ge 0}$  est une suite entière qui converge presque sûrement vers une limite entière, p, alors le théorème 3.2 peut s'énoncer : il existe presque sûrement un entier  $n_0$  tel que pour tout  $n > n_0$ ,

$$\hat{p} = p$$
.

#### Preuve des théorèmes 3.1 et 3.3

Pour simplifier les écritures, nous allons noter :

i)  $IC = \bigcup_{i=0}^{p} [\tau_i + h_n, \tau_{i+1} - h_n],$ 

ii) 
$$||\cdot||_{[h_n,1-h_n]} = ||\cdot||.$$

a) Nous avons

$$||\hat{\gamma}(\cdot)||_{IC} \le ||\hat{m}_{+}(\cdot) - m_{+}(\cdot)||_{IC} + ||\hat{m}_{-}(\cdot) - m_{-}(\cdot)||_{IC}, \qquad (3.3)$$

or

$$\hat{m}_{\pm}(t) = \frac{\sum_{i=1}^{n} \tilde{w}_{i}^{\pm}(t) Y_{i}}{\sum_{i=1}^{n} \tilde{w}_{i}^{\pm}(t)},$$
(3.4)

avec

$$\tilde{w}_{i}^{\pm}(t) = \frac{1}{nh_{n}} K_{\pm}(\frac{t - X_{i}}{h_{n}}) \left( \tilde{S}_{2}^{\pm}(t) - \frac{t - X_{i}}{h_{n}} \tilde{S}_{1}^{\pm}(t) \right) = \frac{1}{n^{2}h_{n}^{4}} w_{i}^{\pm}(t),$$
(3.5)

et où

$$\tilde{S}_{l}^{\pm}(t) = \frac{1}{nh_{n}^{l+1}} \sum_{i=1}^{n} (t - X_{i})^{l} K_{\pm}(\frac{t - X_{i}}{h_{n}}) = \frac{1}{nh_{n}^{l+1}} S_{l}^{\pm}(t), \qquad (3.6)$$

 $\operatorname{donc}$ 

$$\begin{aligned} &||\hat{m}_{\pm}(\cdot) - m_{\pm}(\cdot)||_{IC} \\ &= ||\frac{\sum_{i=1}^{n} \tilde{w}_{i}^{\pm}(\cdot)(Y_{i} - m_{\pm}(\cdot))}{\sum_{i=1}^{n} \tilde{w}_{i}^{\pm}(\cdot)}||_{IC}, \\ &\leq ||\frac{1}{\sum_{i=1}^{n} \tilde{w}_{i}^{\pm}(\cdot)}|| \left[ ||\sum_{i=1}^{n} \tilde{w}_{i}^{\pm}(\cdot)(m(X_{i}) - m_{\pm}(\cdot))||_{IC} + ||\sum_{i=1}^{n} \tilde{w}_{i}^{\pm}(\cdot)\sigma(X_{i})\varepsilon_{i}|| \right]. \end{aligned}$$

Tout d'abord, notons:

1) Sous M3,

$$\tilde{R}_{l}^{\pm}(t) = \frac{\sum_{i=1}^{n} (t - X_{i})^{l} K_{\pm}(\frac{t - X_{i}}{h_{n}}) (m(X_{i}) - m_{\pm}(t)}{n h_{n}^{l+1}}, \\
= \frac{1}{n h_{n}^{l+1}} D_{l}^{\pm}(t).$$

2) Sous M5,  $\,$ 

$$\tilde{R}_{l}^{\pm}(t) = \frac{\sum_{i=1}^{n} (t - X_{i})^{l} K_{\pm}(\frac{t - X_{i}}{h_{n}}) (m(X_{i}) - m_{\pm}(t) - (X_{i} - t) m_{\pm}^{(1)}(t))}{n h_{n}^{l+1}},$$
  
$$= \frac{1}{n h_{n}^{l+1}} R_{l}^{\pm}(t).$$

3)

$$\tilde{\xi}_l^{\pm}(t) = \frac{1}{nh_n^{l+1}} \sum_{i=1}^n (t - X_i)^l K_{\pm}(\frac{t - X_i}{h_n}) \sigma(X_i) \varepsilon_i,$$
  
$$= \frac{1}{nh_n^{l+1}} \xi_l^{\pm}(t).$$

$$\sum_{i=1}^{n} \tilde{w}_{i}^{\pm}(t) (m(X_{i}) - m_{\pm}(t)) = \sum_{i=1}^{n} \tilde{w}_{i}^{\pm}(t) (m(X_{i}) - m_{\pm}(t) - (X_{i} - t)m_{\pm}^{(1)}(t)),$$
  
$$= \tilde{S}_{2}^{\pm}(t) \tilde{R}_{0}^{\pm}(t) - \tilde{S}_{1}^{\pm}(t) \tilde{R}_{1}^{\pm}(t),$$

 $\operatorname{donc}$ 

$$\begin{aligned} &||\sum_{i=1}^{n} \tilde{w}_{i}^{\pm}(\cdot)(m(X_{i}) - m_{\pm}(\cdot))||_{IC} \\ &\leq \left[||\tilde{S}_{2}^{\pm}(\cdot) - K_{2}^{\pm}f_{X}(\cdot)|| + ||K_{2}^{\pm}f_{X}(\cdot)||\right]||\tilde{R}_{0}^{\pm}(\cdot)||_{IC} \\ &+ \left[||\tilde{S}_{1}^{\pm}(\cdot) - K_{1}^{\pm}f_{X}(\cdot)|| + ||K_{1}^{\pm}f_{X}(\cdot)||\right]||\tilde{R}_{1}^{\pm}(\cdot)||_{IC}, \end{aligned}$$
(3.7)

de la même manière nous obtenons

$$\sum_{i=1}^{n} \tilde{w}_{i}^{\pm}(t) \sigma(X_{i}) \varepsilon_{i} = \tilde{S}_{2}^{\pm}(t) \tilde{\xi}_{0}^{\pm}(t) - \tilde{S}_{1}^{\pm}(t) \tilde{\xi}_{1}^{\pm}(t),$$

d'où

$$\begin{aligned} &||\sum_{i=1}^{n} \tilde{w}_{i}^{\pm}(\cdot)\sigma(X_{i})\varepsilon_{i}|| \\ &\leq \left[ ||\tilde{S}_{2}^{\pm}(\cdot) - K_{2}^{\pm}f_{X}(\cdot)|| + ||K_{2}^{\pm}f_{X}(\cdot)|| \right] ||\tilde{\xi}_{0}^{\pm}(\cdot)|| \\ &+ \left[ ||\tilde{S}_{1}^{\pm}(\cdot) - K_{1}^{\pm}f_{X}(\cdot)|| + ||K_{1}^{\pm}f_{X}(\cdot)|| \right] ||\tilde{\xi}_{1}^{\pm}(\cdot)||. \end{aligned}$$
(3.8)

Nous avons également

$$\inf_{t} \left| \sum_{i=1}^{n} \tilde{w}_{i}^{\pm}(\cdot) \right| \geq \left| f_{X}^{2}(\cdot) - \left| \left| \sum_{i=1}^{n} \tilde{w}_{i}^{\pm}(\cdot) - f_{X}^{2}(\cdot) \right| \right| \right|,$$

or

$$\begin{split} \sum_{i=1}^{n} \tilde{w}_{i}^{\pm}(t) &- f_{X}^{2}(t) &= \tilde{S}_{2}^{\pm}(t)\tilde{S}_{0}^{\pm}(t) - \left(\tilde{S}_{1}^{\pm}(t)\right)^{2} - \left(K_{2}^{\pm}K_{0}^{\pm} - \left(K_{1}^{\pm}\right)^{2}\right)f_{X}^{2}(t), \\ &= \left[\tilde{S}_{2}^{\pm}(t) - K_{2}^{\pm}f_{X}(t)\right]\tilde{S}_{0}^{\pm}(t) \\ &+ \left[\tilde{S}_{0}^{\pm}(t) - K_{0}^{\pm}f_{X}(t)\right]K_{2}^{\pm}f_{X}(t) \\ &- \left[\tilde{S}_{1}^{\pm}(t) - K_{1}^{\pm}f_{X}(t)\right]\left[\tilde{S}_{1}^{\pm}(t) + K_{1}^{\pm}f_{X}(t)\right], \end{split}$$

 $\operatorname{donc}$ 

$$\begin{split} &||\sum_{i=1}^{n} \tilde{w}_{i}^{\pm}(\cdot) - f_{X}^{2}(\cdot)|| \\ \leq ||\tilde{S}_{2}^{\pm}(\cdot) - K_{2}^{\pm}f_{X}(\cdot)|| \times ||\tilde{S}_{0}^{\pm}(\cdot)|| \\ &+ ||\tilde{S}_{0}^{\pm}(\cdot) - K_{0}^{\pm}f_{X}(\cdot)|| \times ||K_{2}^{\pm}f_{X}(\cdot)|| \\ &+ ||\tilde{S}_{1}^{\pm}(\cdot) - K_{1}^{\pm}f_{X}(\cdot)|| \times ||\tilde{S}_{1}^{\pm}(\cdot) + K_{1}^{\pm}f_{X}(\cdot)||, \\ \leq ||\tilde{S}_{2}^{\pm}(\cdot) - K_{2}^{\pm}f_{X}(\cdot)|| \times \left[ ||\tilde{S}_{0}^{\pm}(\cdot) - K_{0}^{\pm}f_{X}(\cdot)|| + ||K_{0}^{\pm}f_{X}(\cdot)|| \right] \\ &+ ||\tilde{S}_{0}^{\pm}(\cdot) - K_{0}^{\pm}f_{X}(\cdot)|| \times ||K_{2}^{\pm}f_{X}(\cdot)|| + ||\tilde{S}_{1}^{\pm}(\cdot) - K_{1}^{\pm}f_{X}(\cdot)|| \\ &\times \left[ ||\tilde{S}_{1}^{\pm}(\cdot) - K_{1}^{\pm}f_{X}(\cdot)|| + 2||K_{1}^{\pm}f_{X}(\cdot)||. \right] \end{split}$$
(3.9)

Pour finir la démonstration des deux premières parties des théorèmes 3.1 et 3.3, nous utilisons les résultats du lemme suivant.

# Lemme 3.1 (Silverman [SIL78], Mack & Silverman [MS82])

- 1) Sous les mêmes hypothèses que le théorème 3.1 nous avons
  - a)  $||\tilde{S}_{l}^{\pm}(\cdot) K_{l}^{\pm}f_{X}(\cdot)|| = o(1) \quad p.s.$ b)  $||\tilde{\xi}_{l}^{\pm}(\cdot)|| = o(1) \quad p.s.$ c)  $||\tilde{R}_{l}^{\pm}(\cdot)||_{IC} = o(1) \quad p.s.$ c') Pour tout  $j = 1, \dots, p, \quad |\tilde{R}_{l}^{\pm}(\tau_{j})| = o(1) \quad p.s.$

2) Sous les mêmes hypothèses que le théorème 3.3 nous avons

d) 
$$||\tilde{\xi}_{l}^{\pm}(\cdot)|| = \mathcal{O}(r_{n}) \quad p.s.$$
  
e)  $||\tilde{R}_{l}^{\pm}(\cdot)||_{IC} = \mathcal{O}(h_{n}^{2}) \quad p.s.$   
e') Pour tout  $j = 1, ..., p, \qquad |\tilde{R}_{l}^{\pm}(\tau_{j})| = \mathcal{O}(h_{n}^{2}) \quad p.s.$ 

#### Preuve du lemme 3.1

- a) Voir théorème A de Silverman [SIL78].
- b) et d) Voir Proposition 4 de Mack et Silverman [MS82].

c) Pour tout  $0 \le i \le p$ ,  $m(\cdot)$  est supposée continue sur l'intervalle  $[\tau_i, \tau_{i+1}]$ . Alors, pour tout  $0 \le i \le p$  et pour tout  $t \in [\tau_i + h_n, \tau_{i+1} - h_n]$ ,

$$\begin{split} |\tilde{R}_{l}^{\pm}(t)| &\leq \frac{\sum_{i=1}^{n} \left| (t - X_{i})^{l} K_{\pm}(\frac{t - X_{i}}{h_{n}}) \right| \times |m(X_{i}) - m_{\pm}(t)|}{n h_{n}^{l+1}}, \\ &= o(1) \frac{\sum_{i=1}^{n} \left| (t - X_{i})^{l} K_{\pm}(\frac{t - X_{i}}{h_{n}}) \right|}{n h_{n}^{l+1}} \quad p.s, \\ &= o(1) \mathcal{O}(1) \quad p.s, \\ &= o(1) \quad p.s. \end{split}$$

d'où

$$||\tilde{R}_l^{\pm}(\cdot)||_{IC} = o(1) \quad p.s$$

e) Puisque pour tout  $0 \le i \le p$ ,  $m(\cdot)$  est supposée deux fois dérivable sur l'intervalle  $[\tau_i, \tau_{i+1}]$ , nous avons donc tout  $0 \le i \le p$  et pour tout  $t \in [\tau_i + h_n, \tau_{i+1} - h_n]$ 

$$\begin{split} |\tilde{R}_{l}^{\pm}(t)| &\leq \frac{\sum_{i=1}^{n} \left| (t - X_{i})^{l} K_{\pm}(\frac{t - X_{i}}{h_{n}}) \right| \times \left| m(X_{i}) - m_{\pm}(t) - (X_{i} - t) m_{\pm}^{(1)}(t) \right|}{n h_{n}^{l+1}}, \\ &\leq \lambda h_{n}^{2} \frac{\sum_{i=1}^{n} \left| (t - X_{i})^{l} K_{\pm}(\frac{t - X_{i}}{h_{n}}) \right|}{n h_{n}^{l+1}} \quad p.s, \\ &\leq \lambda h_{n}^{2} \mathcal{O}(1) \quad p.s, \\ &= \mathcal{O}(h_{n}^{2}) \quad p.s, \end{split}$$

d'où

$$||\tilde{R}_l^{\pm}(\cdot)||_{IC} = \mathcal{O}(h_n^2) \quad p.s.$$

Les démonstrations de c) et e) restent valables si nous remplaçons t par  $\tau_i$ . D'où les résultats énoncés dans c') et e').

Ceci termine la preuve du lemme 3.1.

#### Lemme 3.2.

- 1) Sous les hypothèses du théorème 3.1 nous avons
  - a)  $||\sum_{i=1}^{n} \tilde{w}_{i}^{\pm}(\cdot) f_{X}^{2}(\cdot)|| = o(1) \quad p.s.$

p.s.

- b)  $||\sum_{i=1}^{n} \tilde{w}_{i}^{\pm}(\cdot)\sigma(X_{i})\varepsilon_{i}|| = o(1) \quad p.s.$ c)  $||\sum_{i=1}^{n} \tilde{w}_{i}^{\pm}(\cdot)(m(X_{i}) - m_{\pm}(\cdot))||_{IC} = o(1) \quad p.s.$ c') Pour tout  $j = 1, \ldots, p,$  $|\sum_{i=1}^{n} \tilde{w}_{i}^{\pm}(\tau_{j})(m(X_{i}) - m_{\pm}(\tau_{j}))| = o(1) \quad p.s.$
- 2) Sous les mêmes hypothèses que le théorème 3.3 nous avons

d) 
$$||\sum_{i=1}^{n} \tilde{w}_{i}^{\pm}(\cdot)\sigma(X_{i})\varepsilon_{i}|| = \mathcal{O}(r_{n}) \quad p.s.$$
  
e)  $||\sum_{i=1}^{n} \tilde{w}_{i}^{\pm}(\cdot)(m(X_{i}) - m_{\pm}(\cdot))||_{IC} = \mathcal{O}(h_{n}^{2}) \quad p.s.$   
e') Pour tout  $j = 1, \dots, p,$   
 $|\sum_{i=1}^{n} \tilde{w}_{i}^{\pm}(\tau_{j})(m(X_{i}) - m_{\pm}(\tau_{j}))| = \mathcal{O}(h_{n}^{2})$ 

#### Preuve du lemme 3.2.

a) La partie a) du lemme 3.1 et (3.9) nous permettent de conclure.

b) Ce résultat est une conséquence de la partie b) du lemme 3.1 et de (3.8).

c) Les parties a) et c) du lemme 3.1 et (3.7) permettent de conclure la démonstration de c) du lemme 3.2.

d) Cette partie est une conséquence de la partie d) du lemme 3.1 et de (3.8).

e) Enfin le e) du lemme 3.1 et (3.7) permettent de conclure la démonstration de la partie e) du lemme 3.2.

Les démonstrations de c') et e') sont similaires à celles de c) et e) respectivement.

Revenons maintenant à la démonstration de la partie 1 des théorèmes 3.1 et 3.3. Nous avons

$$\begin{aligned} ||\tilde{m}_{\pm}(\cdot) - m_{\pm}(\cdot)||_{IC} \\ &\leq ||\frac{1}{\sum_{i=1}^{n} \tilde{w}_{i}^{\pm}(\cdot)}|| \left[ ||\sum_{i=1}^{n} \tilde{w}_{i}^{\pm}(\cdot)(m(X_{i}) - m_{\pm}(\cdot))||_{IC} + ||\sum_{i=1}^{n} \tilde{w}_{i}^{\pm}(\cdot)\sigma(X_{i})\varepsilon_{i}|| \right], \\ &\leq \frac{\left[ ||\sum_{i=1}^{n} \tilde{w}_{i}^{\pm}(\cdot)(m(X_{i}) - m_{\pm}(\cdot))||_{IC} + ||\sum_{i=1}^{n} \tilde{w}_{i}^{\pm}(\cdot)\sigma(X_{i})\varepsilon_{i}|| \right]}{\left| \inf_{t} f_{X}^{2}(t) - \left| \left| \sum_{i=1}^{n} \tilde{w}_{i}^{\pm}(\cdot) - f_{X}^{2}(\cdot) \right| \right| \right|. \end{aligned}$$

Pour finir nous utilisons les résultats du lemme 3.2 :

6 5 11

• Sous les hypothèses du théorème 3.1, nous obtenons

$$||\hat{m}_{\pm}(\cdot) - m_{\pm}(\cdot)||_{IC} = o(1) \quad p.s.$$

Nous en déduisons donc

$$||\hat{\gamma}(\cdot)||_{IC} = o(1) \quad p.s.$$

• Sous les hypothèses du théorème 3.3, nous avons

$$||\hat{m}_{\pm}(\cdot) - m_{\pm}(\cdot)||_{IC} = \mathcal{O}(h_n^2) + \mathcal{O}(r_n) = o(S_n) \quad p.s.$$

Ce qui implique

$$||\hat{\gamma}(\cdot)||_{IC} = o(S_n) \quad p.s$$

b) Nous avons

$$\begin{aligned} ||\hat{\gamma}(\cdot)||_{[\tau_i - h_n, \tau_i + h_n]} &\geq |\gamma(\tau_i)| - |\hat{\gamma}(\tau_i) - \gamma(\tau_i)|, \\ &\geq |\gamma(\tau_i)| - |\hat{m}_+(\tau_i) - m_+(\tau_i)| - |\hat{m}_-(\tau_i) - m_-(\tau_i)|. \end{aligned}$$

Or

$$\hat{m}_{\pm}(\tau_j) - m_{\pm}(\tau_j) = \frac{\sum_{i=1}^n \tilde{w}_i^{\pm}(\tau_j) (m(X_i) - m_{\pm}(\tau_j))}{\sum_{i=1}^n \tilde{w}_i^{\pm}(\tau_j)} + \frac{\sum_{i=1}^n \tilde{w}_i^{\pm}(\tau_j) \sigma(X_i) \varepsilon_i}{\sum_{i=1}^n \tilde{w}_i^{\pm}(\tau_j)},$$

d'où

$$\begin{aligned} &|\hat{m}_{\pm}(\tau_{j}) - m_{\pm}(\tau_{j})| \\ &\leq \frac{|\sum_{i=1}^{n} \tilde{w}_{i}^{\pm}(\tau_{j})(m(X_{i}) - m_{\pm}(\tau_{j}))| + ||\sum_{i=1}^{n} \tilde{w}_{i}^{\pm}(\cdot)\sigma(X_{i})\varepsilon_{i}||}{|\inf f_{X}^{2}(t) - ||\sum_{i=1}^{n} \tilde{w}_{i}^{\pm}(\cdot) - f_{X}^{2}(\cdot)|||}, \\ &= o(1), \quad p.s. \end{aligned}$$

Par conséquent, soit  $\epsilon > 0$  tel que  $|\gamma(\tau_i)| - S_0 > \epsilon$ . Pour *n* assez grand nous avons

$$|\hat{m}_{\pm}(\tau_i) - m_{\pm}(\tau_i)| < \frac{\epsilon}{2} \qquad p.s,$$

d'où

$$\begin{aligned} ||\hat{\gamma}(\cdot)||_{[\tau_i - h_n, \tau_i + h_n]} &\geq |\gamma(\tau_i)| - \epsilon, \\ &\geq S_0 \qquad p.s, \\ &\geq S_n \qquad p.s, \end{aligned}$$

donc pour tout  $1 \leq i \leq p$ , on a

$$\begin{split} \liminf_{n} ||\hat{\gamma}(\cdot)||_{[\tau_i - h_n, \tau_i + h_n]} &\geq S_0 \qquad p.s, \\ &\geq S_n \qquad p.s, \end{split}$$

d'où

$$\inf_{1 \le i \le p} \liminf_{n} ||\hat{\gamma}(\cdot)||_{[\tau_i - h_n, \tau_i + h_n]} \ge S_0 \qquad p.s,$$
$$\ge S_n \qquad p.s.$$

Ceci achève les démonstrations des théorèmes 3.1 et 3.3.

#### Preuve des théorèmes 3.2 et 3.4.

Les démonstrations de ces deux théorèmes sont analogues. Pour cette raison, nous allons prouver seulement le théorème 3.4. D'après le théorème 3.3, pour tout  $1 \le i \le p$  et pour *n* assez grand on a

$$\left|\left|\hat{\gamma}(\cdot)\right|\right|_{]\tau_i - h_n, \tau_i + h_n[} > S_n \qquad p.s,$$

 $\operatorname{et}$ 

$$\left\| \hat{\gamma}(\cdot) \right\|_{[\tau_i + h_n, \tau_{i+1} - h_n]} < S_n \qquad p.s,$$

 $\operatorname{donc}$ 

• pour tout  $1 \leq j \leq \hat{p}, \forall \ 1 \leq i \leq p$  on a

$$\widehat{ID}_j \not\subset [\tau_i + h_n, \tau_{i+1} - h_n],$$

d'où, il existe un  $1 \le i \le p$  tel que  $\widehat{ID}_j \bigcap ]\tau_i - h_n, \tau_i + h_n \neq \emptyset$ .

Or  $]\tau_i - h_n, \tau_i + h_n[$  ne peut croiser qu'un seul intervalle  $\widehat{ID}_j$ . D'où, pour *n* assez grand,

$$\hat{p} \le p \qquad p.s. \tag{3.10}$$

Nous avons également

$$|\hat{\tau}_j - \tau_i| \le h_n \qquad p.s. \tag{3.11}$$

∀i = 1,..., p, il existe un 1 ≤ j ≤ p̂ tel que ÎD<sub>j</sub> ∩ ]τ<sub>i</sub> − h<sub>n</sub>, τ<sub>i</sub> + h<sub>n</sub>[ ≠ Ø, sinon nous aurions ||Ŷ(·)||<sub>]τi</sub>−h<sub>n</sub>, τ<sub>i</sub>+h<sub>n</sub>[ ≤ S<sub>n</sub>. Or asymptotiquement un ÎD<sub>j</sub> ne peut intersecter qu'un seul ]τ<sub>i</sub> − h<sub>n</sub>, τ<sub>i</sub> + h<sub>n</sub>[, sinon il existe un 1 ≤ i ≤ p tel que pour n assez grand, ||Ŷ(·)||<sub>[τi+hn,τi+1</sub>−h<sub>n</sub>[ > S<sub>n</sub>. Il en résulte que

$$\hat{p} \ge p, \quad p.s. \tag{3.12}$$

En combinant (3.10) et (3.12), nous montrons que

$$\hat{p} \to p$$
  $p.s.$ 

D'après (3.11) et quitte à faire une permutation sur l'indice i, nous montrons que pour tout  $i = 1, \ldots, p$ ,

$$\hat{\tau}_i \rightarrow \tau_i \qquad p.s.$$

# 3.5 Vitesses de convergence

Rappelons que le processus  $\hat{\gamma}(t)$  est défini par :

$$\hat{\gamma}(t) = \hat{m}_{+}(t) - \hat{m}_{-}(t). \tag{3.13}$$

Supposons que la fonction de régression  $m(\cdot)$  admet p points de rupture dans l'intervalle [a, b] $(p = 0 \text{ si } m(\cdot) \text{ est continue sur l'intervalle } [a, b])$ . Soient  $\tau_1, \ldots, \tau_p$  les p points de discontinuité de la fonction de régression. Par convention, nous notons  $\tau_0 = a$  et  $\tau_{p+1} = b$ .

Posons

i) Pour tout  $0 \le i \le p$ ,  $IC_i = [\tau_i + h_n, \tau_{i+1} - h_n]$ . ii) Pour tout  $1 \le i \le p$ ,  $ID_i = ]\tau_i - h_n, \tau_i + h_n[$ . iii)  $IC = \bigcup_{i=0}^p [\tau_i + h_n, \tau_{i+1} - h_n] = \bigcup_{i=0}^p IC_i$ .

## Hypothèses

- C1. La fonction  $m(\cdot)$  est deux fois continûment dérivable sur chaque  $[\tau_i, \tau_{i+1}, i = 0, 1, \dots, P]$ .
- C2. La densité marginale de X,  $f_X(\cdot)$ , est supposée strictement positive sur [0,1] et telle qu'il existe F > 0 pour laquelle

$$|f_X(y) - f_X(x)| \le F |y - x|$$
,  $\forall (x, y) \in [0, 1] \times [0, 1].$ 

- C3.  $\sigma^2(\cdot)$  est supposée bornée sur l'intervalle [0,1].
- C4.  $\exists C > 0$  tel que  $\operatorname{I\!E}\left[\left|\varepsilon\right|^{k}\right] < C^{k-2}k!, k = 3, 4, \dots$

C5. Soit  $(h_n)_{n>1}$  une suite réelle positive vérifiant:

$$\lim_{n \to +\infty} h_n = 0 \text{ et } \lim_{n \to +\infty} nh_n = +\infty.$$

C6.  $K_{+}(\cdot)$  une fonction à support [-1,0] lipschitzienne de rapport K:

$$\forall (x,y) \in [-1,0] \times [-1,0], |K_+(y) - K_+(x)| \le K |y-x|.$$

Lemme 3.3. Sous les hypothèses C1-C6, on a:

a) Il existe  $c_1, c_2, c_3, c_4 > 0$  tels que pour tout  $x > c_1 h_n$ ,

$$\operatorname{IP}\left[\sup_{t\in IC}|\hat{\gamma}(t)| \ge x\right] \le \exp\left[-c_4 \frac{nh_n(x-c_1h_n)^2}{c_3+c_2(x-c_1h_n)}\right]$$

b) Si  $|\gamma(\tau_i)| > 0$ , alors pour tout  $1 \le i \le p$  il existe  $c_1, c_2, c_3, c_4, c_5 > 0$  tels que pour tout  $c_1h_n < x < |\gamma(\tau_i)| - c_5h_n^2$ ,

$$\operatorname{IP}\left[\sup_{t\in ID_{i}}|\hat{\gamma}(t)| < x\right] \leq \exp\left[-c_{4}\frac{nh_{n}(x-c_{1}h_{n})^{2}}{c_{3}+c_{2}(x-c_{1}h_{n})}\right].$$

#### Preuve du lemme 3.3.

Pour montrer la première partie du lemme 3.3, il suffit de prouver que pour tout  $0 \le i \le p$ ,

$$\mathbb{P}\left[\sup_{t\in IC_i}|\hat{\gamma}(t)|\geq x\right] \leq \exp\left[-c_4\frac{nh_n(x-c_1h_n)^2}{c_3+c_2(x-c_1h_n)}\right],$$

en remarquant que

$$\mathbb{IP}\left[\sup_{t\in IC}|\hat{\gamma}(t)|\geq x\right] \leq \sum_{i=0}^{p}\mathbb{IP}\left[\sup_{t\in IC_{i}}|\hat{\gamma}(t)|\geq x\right].$$

Soit I un intervalle inclus dans [a, b]. Notons HO(I) et H1(I) les hypothèses suivantes :

HO(I): la fonction de régression  $m(\cdot)$  est deux fois continûment dérivable sur l'intervalle I,

H1(I): la fonction  $m(\cdot)$  admet un point de rupture dans l'intervalle I.

Ainsi pour prouver le lemme 3.3, il suffit de montrer que

1) 
$$\mathbb{P}_{\text{HO}(I)}\left[\sup_{t\in I} |\hat{\gamma}(t)| \ge x\right] \le \exp\left[-c_4 \frac{nh_n(x-c_1h_n)^2}{c_3+c_2(x-c_1h_n)}\right].$$
  
2)  $\mathbb{P}_{\text{H1}(I)}\left[\sup_{t\in I} |\hat{\gamma}(t)| < x\right] \le \exp\left[-c_4 \frac{nh_n(x-c_1h_n)^2}{c_3+c_2(x-c_1h_n)}\right].$ 

Sans perdre de généralité, nous allons supposer que I = [0, 1] et nous notons HO([0, 1]) = HO et H1([0, 1]) = H1.

Décomposons en partie moyenne et partie aléatoire centrée la variable de saut estimé :

$$\hat{\gamma}(t) = E_n(t) + B_n(t),$$

оù

$$\tilde{E}_{n}(t) = \tilde{E}_{n}^{+}(t) - \tilde{E}_{n}^{-}(t), 
= \frac{\sum_{i=1}^{n} \tilde{w}_{i}^{+}(t) m(X_{i})}{\sum_{i=1}^{n} \tilde{w}_{i}^{+}(t)} - \frac{\sum_{i=1}^{n} \tilde{w}_{i}^{-}(t) m(X_{i})}{\sum_{i=1}^{n} \tilde{w}_{i}^{-}(t)},$$

 $\operatorname{et}$ 

$$\begin{split} \tilde{B}_n(t) &= \tilde{B}_n^+(t) - \tilde{B}_n^-(t), \\ &= \frac{\sum_{i=1}^n \tilde{w}_i^+(t)\sigma(X_i)\varepsilon_i}{\sum_{i=1}^n \tilde{w}_i^+(t)} - \frac{\sum_{i=1}^n \tilde{w}_i^-(t)\sigma(X_i)\varepsilon_i}{\sum_{i=1}^n \tilde{w}_i^-(t)}. \end{split}$$

Nous avons le lemme suivant

Lemme 3.4. Sous les hypothèses C1-C6,

a) il existe  $c_1, c_2, c_3, c_4 > 0$  telles que pour tout  $x > c_1 h_n$  nous avons

$$\operatorname{IP}\left[\sup_{t\in[0,1]}|\tilde{B}_{n}(t)| \ge x\right] \le \exp\left[-c_{4}\frac{nh_{n}(x-c_{1}h_{n})^{2}}{c_{3}+c_{2}(x-c_{1}h_{n})}\right].$$

b) il existe  $c_1, c_2, c_3, c_4 > 0$  telles que pour tout  $x > c_1 h_n^2$  nous avons

$$\mathbb{P}_{\text{HO}}\left[\sup_{t\in[0,1]}|\tilde{E}_n(t)| \ge x\right] \le \exp\left[-c_4 \frac{n(x-c_1h_n^2)^2}{c_3h_n^3 + c_2h_n(x-c_1h_n^2)}\right].$$

c) Soit  $\tau$  un point de rupture de  $m(\cdot)$ , alors il existe  $c_1, c_2, c_3, c_4 > 0$  telles que pour tout  $|\gamma(\tau)| - c_1 h_n^2 > x$ ,

$$\mathbb{P}_{\mathtt{H1}}\left[\sup_{t\in[0,1]}|\tilde{E}_n(t)| < x\right] \le \exp\left[-c_4 \frac{n(|\gamma(\tau)| - x - c_1 h_n^2)^2}{c_3 h_n^3 + c_2 h_n(|\gamma(\tau)| - x - c_1 h_n^2)}\right].$$

Nous en déduisons

e) qu'il existe  $c_1, c_2, c_3, c_4 > 0$  telles que pour tout  $x > c_1 h_n$ ,

$$\mathbb{IP}_{HO}\left[\sup_{t\in[0,1]}|\hat{\gamma}(t)| \ge x\right] \le \exp\left[-c_4 \frac{nh_n(x-c_1h_n)^2}{c_3+c_2(x-c_1h_n)}\right].$$

f) Soit  $\tau$  un point de rupture de  $m(\cdot)$ , alors il existe  $c_1, c_2, c_3, c_4, c_5 > 0$  telles que pour tout  $c_1h_n < x < |\gamma(\tau)| - c_5h_n^2$ ,

$$\mathbb{P}_{\mathbf{H1}}\left[\sup_{t\in[0,1]}|\hat{\gamma}(t)| < x\right] \leq \exp\left[-c_4 \frac{nh_n(x-c_1h_n)^2}{c_3+c_2(x-c_1h_n)}\right].$$

## Preuve du lemme 3.4.

Il est facile de voir que

$$\begin{split} \begin{split} \mathbb{P}_{\mathrm{HO}} \begin{bmatrix} \sup_{t \in [0,1]} |\hat{\gamma}(t)| \ge x \end{bmatrix} &\leq & \mathbb{P}_{\mathrm{HO}} \begin{bmatrix} \sup_{t \in [0,1]} |\tilde{E}_n(t)| + \sup_{t \in [0,1]} |\tilde{B}_n(t)| \ge x \end{bmatrix}, \\ &\leq & \mathbb{P}_{\mathrm{HO}} \begin{bmatrix} \sup_{t \in [0,1]} |\tilde{E}_n(t)| \ge \frac{x}{2} \end{bmatrix} + \mathbb{P} \begin{bmatrix} \sup_{t \in [0,1]} |\tilde{B}_n(t)| \ge \frac{x}{2} \end{bmatrix}, \end{split}$$

 $\operatorname{et}$ 

$$\begin{split} \mathbb{P}_{\mathtt{H1}} \begin{bmatrix} \sup_{t \in [0,1]} |\hat{\gamma}(t)| < x \end{bmatrix} &\leq \mathbb{P}_{\mathtt{H1}} \begin{bmatrix} \sup_{t \in [0,1]} |\tilde{E}_n(t)| - \sup_{t \in [0,1]} |\tilde{B}_n(t)| < x \end{bmatrix}, \\ &\leq \mathbb{P}_{\mathtt{H1}} \begin{bmatrix} \sup_{t \in [0,1]} |\tilde{E}_n(\tau)| < 2x \end{bmatrix} + \mathbb{P} \begin{bmatrix} \sup_{t \in [0,1]} |\tilde{B}_n(t)| \ge x \end{bmatrix}, \end{split}$$

par conséquent les parties d) et e) du lemme 3.4 sont des conséquences immédiates des parties a) et b) et des parties a) et c) respectivement. a) D'après l'inégalité triangulaire,

$$\sup_{t \in [0,1]} |\tilde{B}_n(t)| \le \sup_{t \in [0,1]} |\tilde{B}_n^+(t)| + \sup_{t \in [0,1]} |\tilde{B}_n^-(t)|.$$

On peut écrire:

$$\tilde{B}_{n}^{\pm}(t) = \frac{\tilde{S}_{2}^{\pm}(t)\tilde{\xi}_{0}^{\pm}(t) - \tilde{S}_{1}^{\pm}(t)\tilde{\xi}_{2}^{\pm}(t)}{\sum_{i=1}^{n}\tilde{w}_{i}^{\pm}(t)}.$$

Nous en déduisons que

$$\begin{split} &\sup_{t\in[0,1]} |\tilde{B}_{n}^{\pm}(t)| \\ &\leq \frac{\left[ \left(\sup_{t\in[0,1]} |\tilde{S}_{2}^{\pm}(t)|\right) \left(\sup_{t\in[0,1]} |\tilde{\xi}_{0}^{\pm}(t)|\right) + \left(\sup_{t\in[0,1]} |\tilde{S}_{1}^{\pm}(t)|\right) \left(\sup_{t\in[0,1]} |\tilde{\xi}_{1}^{\pm}(t)|\right) \right]}{\inf_{t\in[0,1]} |\sum_{i=1}^{n} \tilde{w}_{i}^{\pm}(t)|}, \\ &\leq 2 \left( \max_{l=0,1} \sup_{t\in[0,1]} \left| \tilde{\xi}_{l}^{\pm}(t) \right| \right) \frac{\left[ \max_{l=1,2} \sup_{t\in[0,1]} \left| \tilde{S}_{l}^{\pm}(t) - K_{l}^{\pm}f_{X}(t) \right| + \max_{l=1,2} \sup_{t\in[0,1]} \left| K_{l}^{\pm}f_{X}(t) \right| \right]}{\left| \inf_{t\in[0,1]} f_{X}^{2}(t) - \sup_{t\in[0,1]} \left| \sum_{i=1}^{n} \tilde{w}_{i}^{\pm}(t) - f_{X}^{2}(t) \right| \right|}, \\ &\leq \frac{2 \left( \max_{l=0,1} \sup_{t\in[0,1]} \left| \tilde{\xi}_{l}^{\pm}(t) \right| \right) \left[ \max_{l=1,2} \sup_{t\in[0,1]} \left| \tilde{S}_{l}^{\pm}(t) - K_{l}^{\pm}f_{X}(t) \right| + \alpha_{1} \right]}{\left| \alpha_{2} - \sup_{t\in[0,1]} \left| \sum_{i=1}^{n} \tilde{w}_{i}^{\pm}(t) - f_{X}^{2}(t) \right| \right|}, \end{split}$$

avec

i) 
$$\alpha_1 = \max_{l=1,2} \sup_{t \in [0,1]} |K_l^{\pm} f_X(t)|,$$
  
i)  $\alpha_2 = \inf_{t \in [0,1]} f_X^2(t).$ 

 $\operatorname{Posons}$ 

$$A_{n} = 2 \frac{\left[ \max_{l=1,2} \sup_{t \in [0,1]} \left| \tilde{S}_{l}^{\pm}(t) - K_{l}^{\pm} f_{X}(t) \right| + \alpha_{1} \right]}{\left| \alpha_{2} - \sup_{t \in [0,1]} \left| \sum_{i=1}^{n} \tilde{w}_{i}^{\pm}(t) - f_{X}^{2}(t) \right| \right|},$$

 $\operatorname{alors}$ 

$$\begin{split} \mathbb{P}\left[\sup_{t\in[0,1]}|\tilde{B}_{n}^{\pm}(t)|\geq x\right] &\leq \mathbb{P}\left[A_{n}\left(\max_{l=0,1}\sup_{t\in[0,1]}\left|\tilde{\xi}_{l}^{\pm}(t)\right|\right)\geq x\right],\\ &\leq \mathbb{P}\left[A_{n}\left(\max_{l=0,1}\sup_{t\in[0,1]}\left|\tilde{\xi}_{l}^{\pm}(t)\right|\right)\geq x\ ;\ A_{n}\leq\frac{4\alpha_{1}}{\alpha_{2}}\right]\\ &+\mathbb{P}\left[A_{n}\left(\max_{l=0,1}\sup_{t\in[0,1]}\left|\tilde{\xi}_{l}^{\pm}(t)\right|\right)\geq x\ ;\ A_{n}>\frac{4\alpha_{1}}{\alpha_{2}}\right],\\ &\leq \mathbb{P}\left[\max_{l=0,1}\sup_{t\in[0,1]}\left|\tilde{\xi}_{l}^{\pm}(t)\right|\geq\frac{\alpha_{2}}{4\alpha_{1}}x\right]+\mathbb{P}\left[A_{n}>\frac{4\alpha_{1}}{\alpha_{2}}\right].\end{split}$$

Or

$$\mathbb{P}\left[A_n > \frac{4\alpha_1}{\alpha_2}\right] \leq \mathbb{P}\left[\max_{l=1,2} \sup_{t \in [0,1]} \left|\tilde{S}_l^{\pm}(t) - K_l^{\pm} f_X(t)\right| > \frac{\alpha_1}{2}\right] + \mathbb{P}\left[\sup_{t \in [0,1]} \left|\sum_{i=1}^n \tilde{w}_i^{\pm}(t) - f_X^2(t)\right| > \frac{\alpha_2}{4}\right].$$

Ceci implique d'après le lemme 3.5 qu'il existe  $c_1, c_2, c_3, c_4 > 0$  telles que pour tout  $x > c_1 h_n$ ,

$$\mathbb{P}\left[\sup_{t\in[0,1]}|\tilde{B}_n(t)| \ge x\right] \le \exp\left[-c_4 \frac{nh_n(x-c_1h_n)^2}{c_3+c_2(x-c_1h_n)}\right].$$

b) Sous l'hypothèse nulle HO, nous avons

$$\sup_{t \in [0,1]} |\tilde{E}_n(t)| \leq \sup_{t \in [0,1]} |\tilde{E}_n^+(t) - m_+(t)| + \sup_{t \in [0,1]} |\tilde{E}_n^-(t) - m_-(t)|.$$

Or

$$\begin{split} \tilde{E}_{n}^{\pm}(t) &- m_{\pm}(t) \\ &= \frac{\sum_{i=1}^{n} \tilde{w}_{i}^{\pm}(t) \left[ m(X_{i}) - m_{\pm}(t) - (X_{i} - t) m_{\pm}^{(1)}(t) \right]}{\sum_{i=1}^{n} \tilde{w}_{i}^{\pm}(t)}, \\ &= \frac{\tilde{S}_{2}^{\pm}(t) \tilde{R}_{0}^{\pm}(t) - \tilde{S}_{1}^{\pm}(t) \tilde{R}_{1}^{\pm}(t)}{\sum_{i=1}^{n} \tilde{w}_{i}^{\pm}(t)}, \\ &= \frac{1}{\left(\sum_{i=1}^{n} \tilde{w}_{i}^{\pm}(t) - f_{X}^{2}(t)\right) + f_{X}^{2}(t)} \\ &\times \left[ \left( \tilde{S}_{2}^{\pm}(t) - K_{2}^{\pm}f_{X}(t) \right) \tilde{R}_{0}^{\pm}(t) + K_{2}^{\pm}f_{X}(t) \tilde{R}_{0}^{\pm}(t) \right. \\ &- \left( \tilde{S}_{1}^{\pm}(t) - K_{1}^{\pm}f_{X}(t) \right) \tilde{R}_{1}^{\pm}(t) - K_{1}^{\pm}f_{X}(t) \tilde{R}_{1}^{\pm}(t) \right], \end{split}$$

 ${\rm donc}$ 

$$\begin{split} \sup_{t \in [0,1]} &|\tilde{E}_{n}^{\pm}(t) - m_{\pm}(t)| \\ \leq 2 \left( \max_{l=0,1} \sup_{t \in [0,1]} \left| \tilde{R}_{l}^{\pm}(t) \right| \right) \\ &\times \frac{\left[ \left( \max_{l=1,2} \sup_{t \in [0,1]} \left| \tilde{S}_{l}^{\pm}(t) - K_{l}^{\pm} f_{X}(t) \right| \right) + \left( \max_{l=1,2} \sup_{t \in [0,1]} \left| K_{l}^{\pm} f_{X}(t) \right| \right) \right]}{\left| \inf_{t \in [0,1]} f_{X}^{2}(t) - \sup_{t \in [0,1]} \left| \sum_{i=1}^{n} \tilde{w}_{i}^{\pm}(t) - f_{X}^{2}(t) \right| \right|}, \\ \leq \frac{2 \left( \max_{l=0,1} \sup_{t \in [0,1]} \left| \tilde{R}_{l}^{\pm}(t) \right| \right) \left[ \max_{l=1,2} \sup_{t \in [0,1]} \left| \tilde{S}_{l}^{\pm}(t) - K_{l}^{\pm} f_{X}(t) \right| + \alpha_{1} \right]}{\left| \alpha_{2} - \sup_{t \in [0,1]} \left| \sum_{i=1}^{n} \tilde{w}_{i}^{\pm}(t) - f_{X}^{2}(t) \right| \right|}. \end{split}$$

La suite de la démonstration est identique à celle de a). Il en résulte qu'il existe  $c_1, c_2, c_3, c_4 > 0$ telles que pour tout  $x > c_1 h_n^2$ ,

$$\mathbb{P}_{HO}\left[\sup_{t\in[0,1]}|\tilde{E}_n(t)| \ge x\right] \le \exp\left[-c_4 \frac{n(x-c_1h_n^2)^2}{c_3h_n^3 + c_2h_n(x-c_1h_n^2)}\right].$$

c) Sous l'hypothèse H1, il existe  $\tau \in [0, 1]$  tel que  $\gamma(\tau) \neq 0$ . Alors:

$$\sup_{t \in [0,1]} |\tilde{E}_n(t)| \geq |\gamma(\tau)| - |\tilde{E}_n^+(\tau) - m_+(\tau)| - |\tilde{E}_n^-(\tau) - m_-(\tau)|.$$

D'où

$$\begin{split} & \mathbb{P}_{\mathrm{H1}} \left[ \sup_{t \in [0,1]} |\tilde{E}_{n}^{\pm}(t)| < x \right] \\ & \leq \mathbb{P}_{\mathrm{H1}} \left[ |\tilde{E}_{n}^{+}(\tau) - m_{+}(\tau)| + |\tilde{E}_{n}^{-}(\tau) - m_{-}(\tau)| \geq |\gamma(\tau)| - x \right], \\ & \leq \mathbb{P}_{\mathrm{H1}} \left[ |\tilde{E}_{n}^{+}(\tau) - m_{+}(\tau)| \geq \frac{|\gamma(\tau)| - x}{2} \right] + \mathbb{P}_{\mathrm{H1}} \left[ |\tilde{E}_{n}^{-}(\tau) - m_{-}(\tau)| \geq \frac{|\gamma(\tau)| - x}{2} \right]. \end{split}$$

Or

$$\begin{split} |\tilde{E}_{n}^{\pm}(\tau) - m_{\pm}(\tau)| &\leq \frac{2\left(\max_{l=0,1} \left|\tilde{R}_{l}^{\pm}(\tau)\right|\right) \left[\max_{l=1,2} \left|\tilde{S}_{l}^{\pm}(\tau) - K_{l}^{\pm}f_{X}(\tau)\right| + \max_{l=1,2} \left|K_{l}^{\pm}f_{X}(\tau)\right|\right]}{\left|\inf_{t\in[0,1]} f_{X}^{2}(t) - \sup_{t\in[0,1]} \left|\sum_{i=1}^{n} \tilde{w}_{i}^{\pm}(t) - f_{X}^{2}(t)\right|\right|}, \\ &\leq \frac{2\left(\max_{l=0,1} \left|\tilde{R}_{l}^{\pm}(\tau)\right|\right) \left[\max_{l=1,2} \sup_{t\in[0,1]} \left|\tilde{S}_{l}^{\pm}(t) - K_{l}^{\pm}f_{X}(t)\right| + \alpha_{1}\right]}{\left|\alpha_{2} - \sup_{t\in[0,1]} \left|\sum_{i=1}^{n} \tilde{w}_{i}^{\pm}(t) - f_{X}^{2}(t)\right|\right|}. \end{split}$$

Utilisons maintenant les mêmes arguments que ceux qui ont permis de terminer la démonstrations de **a**). Nous en déduisons qu'il existe  $c_1$ ,  $c_2$ ,  $c_3$ ,  $c_4 > 0$  telles que pour tout  $|\gamma(\tau)| - x > c_1 h_n^2$ ,

$$\mathbb{P}_{\mathtt{H1}}\left[|\tilde{E}_{n}^{\pm}(\tau) - m_{\pm}(\tau)| \ge x\right] \le \exp\left[-c_{4}\frac{n(|\gamma(\tau)| - x - c_{1}h_{n}^{2})^{2}}{c_{3}h_{n}^{3} + c_{2}h_{n}(|\gamma(\tau)| - x - c_{1}h_{n}^{2})}\right].$$

Ceci termine la démonstrations du lemme 3.4.

Lemme 3.5. Soit L un entier fixé. Supposons les hypothèses C1-C6 sont satisfaites, alors :

a) Il existe  $c_1, c_2, c_3, c_4 > 0$  telles que pour tout  $x > c_1h_n$ ,

$$\mathbb{IP}\left[\max_{0 \le l \le L} \sup_{t \in [0,1]} |\tilde{S}_l^{\pm}(t) - K_l^{\pm} f_X(t)| \ge x\right] \le \exp\left[-c_4 \frac{nh_n(x - c_1h_n)^2}{c_3 + c_2(x - c_1h_n)}\right].$$

b) Il existe  $c_1, c_2, c_3, c_4 > 0$  telles que pour tout  $x > c_1 h_n$ ,

$$\mathbb{P}\left[\max_{0 \le l \le L} \sup_{t \in [0,1]} |\tilde{\xi}_l^{\pm}(t)| \ge x\right] \le \exp\left[-c_4 \frac{nh_n(x-c_1h_n)^2}{c_3 + c_2(x-c_1h_n)}\right].$$

c) Il existe  $c_1, c_2, c_3, c_4 > 0$  telles que pour tout  $x > c_1 h_n^2$ ,

$$\mathbb{IP}_{\mathrm{HO}} \left[ \max_{0 \le l \le L} \sup_{t \in [0,1]} |\tilde{R}_l^{\pm}(t)| \ge x \right] \le \exp \left[ -c_4 \frac{n(x-c_1 h_n^2)^2}{c_3 h_n^3 + c_2 h_n (x-c_1 h_n^2)} \right].$$

d) Soit  $\tau$  un point de rupture de  $m(\cdot)$ . Alors il existe  $c_1, c_2, c_3, c_4 > 0$  telles que pour tout  $x > c_1 h_n^2$ ,

$$\mathbb{P}_{\mathtt{H1}}\left[\max_{0 \le l \le L} |\tilde{R}_l^{\pm}(\tau)| \ge x\right] \le \exp\left[-c_4 \frac{n(x-c_1h_n^2)^2}{c_3h_n^3 + c_2h_n(x-c_1h_n^2)}\right].$$

e) Il existe  $c_1, c_2, c_3, c_4 > 0$  telles que pour tout  $x > c_1 h_n$ ,

$$\mathbb{P}\left[\sup_{t\in[0,1]}|\sum_{i=1}^{n}\tilde{w}_{i}^{\pm}(t) - f_{X}^{2}(t)| \ge x\right] \le \exp\left[-c_{4}\frac{nh_{n}(x-c_{1}h_{n})^{2}}{c_{3}+c_{2}(x-c_{1}h_{n})}\right].$$

### Preuve du lemme 3.5.

Les démonstrations des parties a), b), c) et d) sont triviales et identiques en tous points. Pour cela nous donnons ici seulement la démonstration de la partie a).

a) Nous avons

$$\mathbb{P}\left[\max_{0\leq l\leq L}\sup_{t\in[0,1]}\left|\tilde{S}_{l}^{\pm}(t)-K_{l}^{\pm}f_{X}(t)\right|\geq x\right] \leq \sum_{l=0}^{L}\mathbb{P}\left[\sup_{t\in[0,1]}\left|\tilde{S}_{l}^{\pm}(t)-K_{l}^{\pm}f_{X}(t)\right|\geq x\right].$$

Appliquons maintenant le résultata) du lemme 3.6. Nous en déduisons le résultat énoncé dans le a) du lemme 3.5.

e) En écrivant

$$\begin{split} &\sum_{i=1}^{n} \tilde{w}_{i}^{\pm}(t) - \left(K_{2}^{\pm}K_{0}^{\pm} - \left(K_{1}^{\pm}\right)^{2}\right)f_{X}^{2}(t) \\ &= \tilde{S}_{2}^{\pm}(t)\tilde{S}_{0}^{\pm}(t) - \tilde{S}_{1}^{\pm}(t)\tilde{S}_{1}^{\pm}(t) - \left(K_{2}^{\pm}K_{0}^{\pm} - \left(K_{1}^{\pm}\right)^{2}\right)f_{X}^{2}(t), \\ &= \left(\tilde{S}_{2}^{\pm}(t) - K_{2}^{\pm}f_{X}(t)\right)\left(\tilde{S}_{0}^{\pm}(t) - K_{0}^{\pm}f_{X}(t)\right) - \left(\tilde{S}_{1}^{\pm}(t) - K_{1}^{\pm}f_{X}(t)\right)^{2} \\ &+ \left(\tilde{S}_{2}^{\pm}(t) - K_{2}^{\pm}f_{X}(t)\right)K_{0}^{\pm}f_{X}(t) + \left(\tilde{S}_{0}^{\pm}(t) - K_{0}^{\pm}f_{X}(t)\right)K_{2}^{\pm}f_{X}(t) \\ &- 2K_{1}^{\pm}f_{X}(t)\left(\tilde{S}_{1}^{\pm}(t) - K_{1}^{\pm}f_{X}(t)\right). \end{split}$$

Nous en déduisons

$$\begin{split} \sup_{t\in[0,1]} &|\sum_{i=1}^{n} \tilde{w}_{i}^{\pm}(t) - f_{X}^{2}(t)| \\ = \sup_{t\in[0,1]} &|\sum_{i=1}^{n} \tilde{w}_{i}^{\pm}(t) - (K_{2}^{\pm}K_{0}^{\pm} - (K_{1}^{\pm})^{2})f_{X}^{2}(t)| \\ \leq &\left(\sup_{t\in[0,1]} \left|\tilde{S}_{2}^{\pm}(t) - K_{2}^{\pm}f_{X}(t)\right|\right) \left(\sup_{t\in[0,1]} \left|\tilde{S}_{0}^{\pm}(t) - K_{0}^{\pm}f_{X}(t)\right|\right) \\ &+ &\left(\sup_{t\in[0,1]} \left(\tilde{S}_{1}^{\pm}(t) - K_{1}^{\pm}f_{X}(t)\right)^{2}\right) \\ &+ &\left(\sup_{t\in[0,1]} \left|\tilde{S}_{2}^{\pm}(t) - K_{2}^{\pm}f_{X}(t)\right|\right) \left(\sup_{t\in[0,1]} |K_{0}^{\pm}|f_{X}(t)\right) \\ &+ &\left(\sup_{t\in[0,1]} \left|\tilde{S}_{0}^{\pm}(t) - K_{0}^{\pm}f_{X}(t)\right|\right) \left(\sup_{t\in[0,1]} |K_{2}^{\pm}|f_{X}(t)\right) \\ &+ &2 \left(\sup_{t\in[0,1]} |K_{1}^{\pm}|f_{X}(t)\right) \left(\sup_{t\in[0,1]} \left|\tilde{S}_{1}^{\pm}(t) - K_{1}^{\pm}f_{X}(t)\right|\right) \\ &\leq &\left(\max_{l=0,1,2} \sup_{t\in[0,1]} \left|\tilde{S}_{l}^{\pm}(t) - K_{l}^{\pm}f_{X}(t)\right|\right) \\ &\times &\left[2 \max_{l=0,1,2} \sup_{t\in[0,1]} \left|\tilde{S}_{l}^{\pm}(t) - K_{l}^{\pm}f_{X}(t)\right| + 4 \max_{l=0,1,2} \sup_{t\in[0,1]} |K_{l}^{\pm}f_{X}(t)|\right]\right]. \end{split}$$

Pour terminer la démonstration du e) du lemme 3.5, nous utilisons le a) du même lemme. Il existe  $c_1, c_2, c_3, c_4 > 0$  telles que pour tout  $x > c_1 h_n$ ,

$$\mathbb{P}\left[\sup_{t\in[0,1]}|\sum_{i=1}^{n}\tilde{w}_{i}^{\pm}(t) - f_{X}^{2}(t)| \ge x\right] \le \exp\left[-c_{4}\frac{nh_{n}(x-c_{1}h_{n})^{2}}{c_{3}+c_{2}(x-c_{1}h_{n})}\right].$$

et ceci achève la démonstration du lemme 3.5.

Rappelons que

$$\begin{aligned} &\text{i)} \quad \tilde{S}_{l}^{\pm}(t) = \frac{1}{nh_{n}^{l+1}} \sum_{i=1}^{n} (t - X_{i})^{l} K_{\pm}(\frac{t - X_{i}}{h_{n}}), \\ &\text{iii)} \quad \tilde{\xi}_{l}^{\pm}(t) = \frac{1}{nh_{n}^{l+1}} \sum_{i=1}^{n} (t - X_{i})^{l} K_{\pm}(\frac{t - X_{i}}{h_{n}}) \sigma(X_{i}) \varepsilon_{i}, \\ &\text{iii)} \quad \tilde{R}_{l}^{\pm}(t) = \frac{\sum_{i=1}^{n} (t - X_{i})^{l} K_{\pm}(\frac{t - X_{i}}{h_{n}}) \left[ m(X_{i}) - m_{\pm}(t) - (X_{i} - t) m_{\pm}^{(1)}(t) \right]}{nh_{n}^{l+1}}. \end{aligned}$$

Lemme 3.6. Supposons que les hypothèses C1-C6 soient satisfaites, alors

a) Il existe  $c_1, c_2, c_3, c_4 > 0$  telles que pour tout  $x > c_1 h_n$ ,

$$\mathbb{P}\left[\sup_{t\in[0,1]} |\tilde{S}_l^{\pm}(t) - K_l^{\pm} f_X(t)| \ge x\right] \le \exp\left[-c_4 \frac{nh_n(x - c_1h_n)^2}{c_3 + c_2(x - c_1h_n)}\right].$$

b) Il existe  $c_1, c_2, c_3, c_4 > 0$  telles que pour tout  $x > c_1h_n$ ,

$$\mathbb{P}\left[\sup_{t\in[0,1]}|\tilde{\xi}_{l}^{\pm}(t)| \ge x\right] \le \exp\left[-c_{4}\frac{nh_{n}(x-c_{1}h_{n})^{2}}{c_{3}+c_{2}(x-c_{1}h_{n})}\right].$$

c) Il existe  $c_1, c_2, c_3, c_4 > 0$  telles que pour tout  $x > c_1 h_n^2$ ,

$$\mathbb{P}_{\text{HO}}\left[\sup_{t\in[0,1]}|\tilde{R}_{l}^{\pm}(t)| \geq x\right] \leq \exp\left[-c_{4}\frac{n(x-c_{1}h_{n}^{2})^{2}}{c_{3}h_{n}^{3}+c_{2}h_{n}(x-c_{1}h_{n}^{2})}\right].$$

d) Soit  $\tau$  un point de rupture de  $m(\cdot)$ . Alors il existe  $c_1, c_2, c_3, c_4 > 0$  telles que pour tout  $x > c_1 h_n^2$ ,

$$\mathbb{P}_{\mathtt{H1}}\left[|\tilde{R}_{l}^{\pm}(\tau)| \ge x\right] \le \exp\left[-c_{4}\frac{n(x-c_{1}h_{n}^{2})^{2}}{c_{3}h_{n}^{3}+c_{2}h_{n}(x-c_{1}h_{n}^{2})}\right]$$

#### Preuve du lemme 3.6.

a) Cherchons des majorants de  $\frac{1}{nh_n^{l+1}}(t-X_i)^l K_{\pm}(\frac{t-X_i}{h_n})$  et  $Var(\tilde{S}_l^{\pm}(t))$  permettant d'appliquer l'inégalité de Bernstein 2.11:

d'une part nous avons

$$\left|\frac{1}{nh_n^{l+1}}(t-X_i)^l K_{\pm}(\frac{t-X_i}{h_n})\right| \le \frac{3c_2}{2nh_n}.$$

d'autre part

$$Var(\tilde{S}_l^{\pm}(t)) \leq \frac{1}{nh_n} \int x^{2l} K_{\pm}^2(x) f_X(t-h_n x) dx,$$
  
$$\leq \frac{c_3}{2nh_n},$$

d'après l'inégalité de Bernstein (la deuxième partie du lemme) nous avons donc

$$\operatorname{IP}\left[|\tilde{S}_l^{\pm}(t) - \operatorname{IE}(\tilde{S}_l^{\pm}(t))| \ge x\right] \le 2 \exp\left(-\frac{nh_n x^2}{c_3 + c_2 x}\right),$$

or

$$\mathbb{E}(\tilde{S}_l^{\pm}(t)) = \int x^l K_{\pm}(x) f_X(t - h_n x) dx ,$$

donc si nous supposons que  $f_X(\cdot)$  vérifie C2, nous obtenons

$$|\mathbb{E}(\tilde{S}_{l}^{\pm}(t)) - K_{l}^{\pm}f_{X}(t)| \leq \int |x^{l}K_{\pm}(x)||f_{X}(t - h_{n}x) - f_{X}(t)|dx,$$
  
$$\leq v_{1}h_{n}.$$

En appliquant l'inégalité triangulaire nous avons

$$\begin{aligned} |\tilde{S}_{l}^{\pm}(t) - K_{l}^{\pm}f_{X}(t)| &\leq |\tilde{S}_{l}^{\pm}(t) - \mathbb{E}(\tilde{S}_{l}^{\pm}(t))| + |\mathbb{E}(\tilde{S}_{l}^{\pm}(t)) - K_{l}^{\pm}f_{X}(t)|, \\ &\leq |\tilde{S}_{l}^{\pm}(t) - \mathbb{E}(\tilde{S}_{l}^{\pm}(t))| + v_{1}h_{n}, \end{aligned}$$

d'où pour tout  $x \ge v_1 h_n$ ,

$$\mathbb{P}\left[|\tilde{S}_{l}^{\pm}(t) - K_{l}^{\pm}f_{X}(t)| \ge x\right] \le 2 \exp\left[-\frac{nh_{n}(x - v_{1}h_{n})^{2}}{c_{3} + c_{2}(x - v_{1}h_{n})}\right].$$

Nous avons

$$[0,1] \subset \bigcup_{k=1}^{\lceil 1/h_n^3 \rceil} [(k-1)h_n^3, kh_n^3],$$

où  $\lceil x \rceil$  désigne la partie entière supérieure de x.

Pour  $t \in [(k-1)h_n^3, kh_n^3]$ ,

$$\begin{split} |\tilde{S}_{l}^{\pm}(t) - K_{l}^{\pm}f_{X}(t) - \tilde{S}_{l}^{\pm}(kh_{n}^{3}) + K_{l}^{\pm}f_{X}(kh_{n}^{3})| \\ \leq |K_{l}^{\pm}||f_{X}(t) - f_{X}(kh_{n}^{3})| + |\tilde{S}_{l}^{\pm}(t) - \tilde{S}_{l}^{\pm}(kh_{n}^{3})|, \\ \leq v_{2}|t - kh_{n}^{3}| + \frac{\sum_{i=1}^{n} \left| \left(\frac{t - X_{i}}{h_{n}}\right)^{l} K_{\pm} \left(\frac{t - X_{i}}{h_{n}}\right) - \left(\frac{kh_{n}^{3} - X_{i}}{h_{n}}\right)^{l} K_{\pm} \left(\frac{kh_{n}^{3} - X_{i}}{h_{n}}\right) \right|}{nh_{n}}, \\ \leq v_{2}|t - kh_{n}^{3}| + v_{3}\frac{|t - kh_{n}^{3}|}{h_{n}^{2}}, \\ \leq v_{4}h_{n}, \end{split}$$

où nous avons utilisé le fait que les fonctions  $f_X(\cdot)$  et  $K_+(\cdot)$  sont lipschitziennes (C2 et C6).

 $\operatorname{Donc}$ 

$$\sup_{t \in [0,1]} |\tilde{S}_l^{\pm}(t) - K_l^{\pm} f_X(t)| \le \max_{1 \le k \le \lceil 1/h_n^3 \rceil} |\tilde{S}_l^{\pm}(kh_n^3) - K_l^{\pm} f_X(kh_n^3)| + v_4 h_n,$$

et par suite, pour tout  $x > (v_1 + v_4)h_n$ ,

$$\begin{split} \begin{split} \Pr\left[\sup_{t\in[0,1]}|\tilde{S}_{l}^{\pm}(t)-K_{l}^{\pm}f_{X}(t)| \geq x\right] &\leq \Pr\left[\max_{1\leq k\leq \lceil 1/h_{n}^{3}\rceil}|\tilde{S}_{l}^{\pm}(kh_{n}^{3})-K_{l}^{\pm}f_{X}(kh_{n}^{3})| \geq x-v_{4}h_{n}\right],\\ &\leq \sum_{k=1}^{\lceil 1/h_{n}^{3}\rceil}\Pr\left[|\tilde{S}_{l}^{\pm}(kh_{n}^{3})-K_{l}^{\pm}f_{X}(kh_{n}^{3})| \geq x-v_{4}h_{n}\right],\\ &\leq \frac{2}{h_{n}^{3}}\exp\left[-\frac{nh_{n}(x-(v_{1}+v_{4})h_{n})^{2}}{c_{3}+c_{2}(x-(v_{1}+v_{4})h_{n})}\right],\end{split}$$

donc il existe  $c_1 > 0$  et  $c_4 > 0$  telles que pour tout  $x > c_1 h_n$ ,

$$\mathbb{P}\left[\sup_{t\in[0,1]}|\tilde{S}_{l}^{\pm}(t)-K_{l}^{\pm}f_{X}(t)| \ge x\right] \le \exp\left[-c_{4}\frac{nh_{n}(x-c_{1}h_{n})^{2}}{c_{3}+c_{2}(x-c_{1}h_{n})}\right].$$

b) Nous avons

$$\mathbb{E}\left[\left|\frac{1}{nh_n^{l+1}}(t-X_i)^l K_{\pm}(\frac{t-X_i}{h_n})\right|^k \sigma^k(X_i)|\varepsilon_i|^k\right] \leq \frac{1}{(nh_n)^k} v_5^k \mathbb{E}|\varepsilon|^k,$$
$$\leq \frac{c_2^{k-2}}{(nh_n)^{k-2}} k!.$$

$$\begin{aligned} Var(\tilde{\xi}_l^{\pm}(t)) &\leq \frac{1}{nh_n} \int x^{2l} K_{\pm}^2(x) \sigma^2(t-xh_n) f_X(t-h_n x) dx ,\\ &\leq \frac{c_3}{2nh_n}, \end{aligned}$$

d'après l'inégalité de Bernstein (la première partie du lemme) nous avons donc

$$\mathbb{P}\left[|\tilde{\xi}_l^{\pm}(t)| \ge x\right] \le 2 \exp\left[-\frac{nh_n x^2}{2(c_3 + c_2 x)}\right].$$

Pour  $t \in [(k-1)h_n^3, kh_n^3],$ 

$$\begin{split} |\tilde{\xi}_{l}^{\pm}(t) - \tilde{\xi}_{l}^{\pm}(kh_{n}^{3})| &\leq \frac{\sum_{i=1}^{n} \left| \left( \frac{t-Xi}{h_{n}} \right)^{l} K_{\pm} \left( \frac{t-X_{i}}{h_{n}} \right) - \left( \frac{kh_{n}^{3}-Xi}{h_{n}} \right)^{l} K_{\pm} \left( \frac{kh_{n}^{3}-X_{i}}{h_{n}} \right) \right| \sigma(X_{i})|\varepsilon_{i}|}{nh_{n}}, \\ &\leq v_{6} \frac{\left| t-kh_{n}^{3} \right|}{nh_{n}^{2}} \sum_{i=1}^{n} |\varepsilon_{i}|, \\ &\leq v_{6}h_{n} \frac{\sum_{i=1}^{n} |\varepsilon_{i}|}{n}, \end{split}$$

 $\operatorname{donc}$ 

$$\sup_{t \in [0,1]} |\tilde{\xi}_l^{\pm}(t)| \le \max_{1 \le k \le \lceil 1/h_n^3 \rceil} |\tilde{\xi}_l^{\pm}(kh_n^3)| + v_6 h_n \frac{\sum_{i=1}^n |\varepsilon_i|}{n},$$

et par conséquent

$$\begin{split} & \mathbb{P}\left[\sup_{t\in[0,1]}|\tilde{\xi}_{l}^{\pm}(t)|\geq x\right] \\ &\leq \mathbb{P}\left[\max_{1\leq k\leq \lceil 1/h_{n}^{3}\rceil}|\tilde{\xi}_{l}^{\pm}(kh_{n}^{3})|+v_{6}h_{n}\frac{\sum_{i=1}^{n}|\varepsilon_{i}|}{n}\geq x\right], \\ &\leq \mathbb{P}\left[\max_{1\leq k\leq \lceil 1/h_{n}^{3}\rceil}|\tilde{\xi}_{l}^{\pm}(kh_{n}^{3})|+v_{6}h_{n}\frac{\sum_{i=1}^{n}|\varepsilon_{i}|}{n}\geq x\ ;\ \frac{\sum_{i=1}^{n}|\varepsilon_{i}|}{n}\leq 2\mathbb{E}|\varepsilon|\right] \\ &+ \mathbb{P}\left[\max_{1\leq k\leq \lceil 1/h_{n}^{3}\rceil}|\tilde{\xi}_{l}^{\pm}(kh_{n}^{3})|+v_{6}h_{n}\frac{\sum_{i=1}^{n}|\varepsilon_{i}|}{n}\geq x\ ;\ \frac{\sum_{i=1}^{n}|\varepsilon_{i}|}{n}>2\mathbb{E}|\varepsilon|\right], \\ &\leq \mathbb{P}\left[\max_{1\leq k\leq \lceil 1/h_{n}^{3}\rceil}|\tilde{\xi}_{l}^{\pm}(kh_{n}^{3})|\geq x-2v_{6}\mathbb{E}|\varepsilon|h_{n}\right]+\mathbb{P}\left[\frac{\sum_{i=1}^{n}(|\varepsilon_{i}|-\mathbb{E}|\varepsilon_{i}|)}{n}>\mathbb{E}|\varepsilon|\right]. \end{split}$$

Nous avons

$$\begin{split} \operatorname{I\!E} \begin{bmatrix} ||\varepsilon| - \operatorname{I\!E}|\varepsilon||^k \end{bmatrix} &\leq 2^{k-1} \begin{bmatrix} \operatorname{I\!E}|\varepsilon|^k + (\operatorname{I\!E}|\varepsilon|)^k \end{bmatrix}, \\ &\leq 2^k \operatorname{I\!E}|\varepsilon|^k, \end{split}$$

d'après l'inégalité de Bernstein appliquée aux  $|\varepsilon_i|,$ 

$$\mathbb{P}\left[\frac{\sum_{i=1}^{n} |\varepsilon_i| - \mathbb{E}|\varepsilon_i|}{n} > \mathbb{E}|\varepsilon|\right] \leq \mathbb{P}\left[\left|\sum_{i=1}^{n} |\varepsilon_i| - \mathbb{E}|\varepsilon_i|\right| > n\mathbb{E}|\varepsilon|\right], \\ \leq \exp(-v_7 n)$$

d'où, en prenant  $x \ge v_8 h_n \ge 2v_6 \mathbb{E}|\varepsilon|h_n,$ 

$$\begin{split} & \mathbb{P}\left[\sup_{t\in[0,1]}|\tilde{\xi}_{l}^{\pm}(t)|\geq x\right] \\ &\leq \mathbb{P}\left[\max_{k}|\tilde{\xi}_{l}^{\pm}(kh_{n}^{3})|\geq x-2v_{6}\mathbb{E}|\varepsilon|h_{n}\right]+\mathbb{P}\left[\frac{\sum_{i=1}^{n}|\varepsilon_{i}|-\mathbb{E}|\varepsilon_{i}|}{n}>\mathbb{E}|\varepsilon|\right], \\ &\leq \sum_{k=1}^{\lceil 1/h_{n}^{3}\rceil}\mathbb{P}\left[|\tilde{\xi}_{l}^{\pm}(kh_{n}^{3})|\geq x-v_{8}h_{n}\right]+\exp(-v_{7}n), \\ &\leq \frac{2}{h_{n}^{3}}\exp\left[-\frac{nh_{n}(x-v_{8}h_{n})^{2}}{2(c_{3}+c_{2}(x-v_{8}h_{n}))}\right]+\exp(-v_{7}n), \end{split}$$

donc il existe  $c_1 > 0$  et  $c_4 > 0$  telles que pour tout  $x > c_1 h_n$ ,

$$\mathbb{P}\left[\sup_{t\in[0,1]}|\tilde{\xi}_{l}^{\pm}(t)| \ge x\right] \le \exp\left[-c_{4}\frac{nh_{n}(x-c_{1}h_{n})^{2}}{c_{3}+c_{2}(x-c_{1}h_{n})}\right].$$

c) Nous avons

$$\begin{aligned} \left| \frac{(t-X_i)^l}{nh_n^{l+1}} K_{\pm}(\frac{t-X_i}{h_n}) \left[ m(X_i) - m_{\pm}(t) - (X_i - t)m_{\pm}^{(1)}(t) \right] \right| &\leq \frac{3c_2h_n}{2n}. \end{aligned}$$

$$Var(\tilde{R}_l^{\pm}(t)) &\leq \frac{1}{nh_n} \int x^{2l} K_{\pm}^2(x) \left[ m(t-xh_n) - m_{\pm}(t) + xh_n m_{\pm}^{(1)}(t) \right]^2 f_X(t-h_n x) dx ,$$

$$&\leq \frac{c_3h_n^3}{2n}, \end{aligned}$$

d'après l'inégalité de Bernstein (la deuxième partie du lemme) nous obtenons donc

$$\mathbb{P}\left[|\tilde{R}_l^{\pm}(t) - \mathbb{E}(\tilde{R}_l^{\pm}(t))| \ge x\right] \le 2 \exp\left(-\frac{nx^2}{c_3h_n^3 + c_2xh_n}\right),$$

or

$$\mathbb{E}(\tilde{R}_{l}^{\pm}(t)) = \int x^{l} K_{\pm}(x) \left[ m(t - xh_{n}) - m_{\pm}(t) + xh_{n} m_{\pm}^{(1)}(t) \right] f_{X}(t - h_{n}x) dx ,$$

donc sous l'hypothèse nulle nous aurions

$$|\mathbb{E}(\tilde{R}_l^{\pm}(t))| \leq v_9 h_n^2.$$

En appliquant de l'inégalité triangulaire nous obtenons

$$\begin{aligned} |\tilde{R}_l^{\pm}(t)| &\leq |\tilde{R}_l^{\pm}(t) - \mathbb{E}(\tilde{R}_l^{\pm}(t))| + |\mathbb{E}(\tilde{R}_l^{\pm}(t))| \\ &\leq |\tilde{R}_l^{\pm}(t) - \mathbb{E}(\tilde{R}_l^{\pm}(t))| + v_9 h_n^2, \end{aligned}$$

d'où pour tout  $x \ge v_9 h_n^2$ ,

$$\mathbb{P}\left[ |\tilde{R}_{l}^{\pm}(t)| \ge x \right] \le 2 \exp\left[ -\frac{n(x-v_{9}h_{n}^{2})^{2}}{c_{2}h_{n}^{3}+c_{3}h_{n}(x-v_{9}h_{n}^{2})} \right]. 
 (3.14)$$

Pour  $t \in [(k-1)h_n^3, kh_n^3]$ , nous avons

$$\begin{split} & \left[ m(X_i) - m_{\pm}(kh_n^3) - (X_i - kh_n^3)m_{\pm}^{(1)}(kh_n^3) \right] \\ &= \left[ m(X_i) - m_{\pm}(t) - (X_i - t)m_{\pm}^{(1)}(t) \right] \\ &+ \left[ m_{\pm}(t) - m_{\pm}(kh_n^3) \right] + (X_i - t) \left[ m_{\pm}^{(1)}(t) - m_{\pm}^{(1)}(kh_n^3) \right] + (kh_n^3 - t)m_{\pm}^{(1)}(kh_n^3), \\ &= \left[ m(X_i) - m_{\pm}(t) - (X_i - t)m_{\pm}^{(1)}(t) \right] + \mathcal{O}(h_n^3), \end{split}$$

 $\operatorname{donc}$ 

$$\begin{split} \tilde{R}_{l}^{\pm}(t) &= \tilde{R}_{l}^{\pm}(kh_{n}^{3}) \\ &= \frac{\sum_{i=1}^{n} \left(\frac{t-X_{i}}{h_{n}}\right)^{l} K_{\pm} \left(\frac{t-X_{i}}{h_{n}}\right) r_{i}^{\pm}(t)}{nh_{n}} - \frac{\sum_{i=1}^{n} \left(\frac{kh_{n}^{3}-X_{i}}{h_{n}}\right)^{l} K_{\pm} \left(\frac{kh_{n}^{3}}{h_{n}}\right) r_{i}^{\pm}(kh_{n}^{3})}{nh_{n}}, \\ &= \frac{\sum_{i=1}^{n} \left(\frac{t-X_{i}}{h_{n}}\right)^{l} K_{\pm} \left(\frac{t-X_{i}}{h_{n}}\right) r_{i}^{\pm}(t)}{nh_{n}} - \frac{\sum_{i=1}^{n} \left(\frac{kh_{n}^{3}-X_{i}}{h_{n}}\right)^{l} K_{\pm} \left(\frac{kh_{n}^{3}-X_{i}}{h_{n}}\right) \left[r_{i}^{\pm}(kh_{n}^{3}) + \mathcal{O}(h_{n}^{3})\right]}{nh_{n}}, \\ &= \frac{\sum_{i=1}^{n} \left[\left(\frac{t-X_{i}}{h_{n}}\right)^{l} K_{\pm} \left(\frac{t-X_{i}}{h_{n}}\right) - \left(\frac{kh_{n}^{3}-X_{i}}{h_{n}}\right)^{l} K_{\pm} \left(\frac{kh_{n}^{3}-X_{i}}{h_{n}}\right)\right] r_{i}^{\pm}(t)}{nh_{n}}, \\ &= \mathcal{O}(h_{n}^{2}), \end{split}$$

où  $r_i^{\pm}(t) = \left[ m(X_i) - m_{\pm}(t) - (X_i - t)m_{\pm}^{(1)}(t) \right].$ 

 $\operatorname{Donc}$ 

$$\sup_{t \in [0,1]} |\tilde{R}_l^{\pm}(t)| \le \max_{1 \le k \le \lceil 1/h_n^3 \rceil} |\tilde{R}_l^{\pm}(kh_n^3)| + v_{10}h_n^2,$$

et par suite pour tout  $x > (v_9 + v_{10})h_n^2$ ,

$$\mathbb{P}\left[\sup_{t\in[0,1]} |\tilde{R}_{l}^{\pm}(t)| \geq x\right] \leq \mathbb{P}\left[\max_{1\leq k\leq \lceil 1/h_{n}^{3}\rceil} |\tilde{R}_{l}^{\pm}(kh_{n}^{3})| \geq x - v_{10}h_{n}^{2}\right], \\
\leq \sum_{k=1}^{\lceil 1/h_{n}^{3}\rceil} \mathbb{P}\left[|\tilde{R}_{l}^{\pm}(kh_{n}^{3})| \geq x - v_{10}h_{n}^{2}\right], \\
\leq \frac{2}{h_{n}^{3}} \exp\left[-\frac{n(x - (v_{9} + v_{10})h_{n}^{2})^{2}}{c_{2}h_{n}^{3} + c_{3}h_{n}(x - (v_{9} + v_{10})h_{n}^{2})}\right],$$

donc il existe  $c_1 > 0$  et  $c_4 > 0$  telles que pour tout  $x > c_1 h_n^2$ ,

$$\operatorname{IP}\left[\sup_{t\in[0,1]} |\tilde{R}_l^{\pm}(t)| \ge x\right] \le \exp\left[-c_4 \frac{n(x-c_1h_n^2)^2}{c_3h_n^3 + c_2h_n(x-c_1h_n^2)}\right].$$

d) En remplaçant t par  $\tau$ , (3.14) reste valable, ce qui prouve la partie 4 du lemme 3.6.

**Théorème 3.5.** Supposons que les hypothèses C1-C6 sont satisfaites et que le seuil  $S_n$  est tel que  $S_n = \alpha_n h_n$ ,  $S_n \to 0$  et  $\alpha_n \to \infty$ . Alors il existe  $\theta > 0$  tel que :

- a)  $\operatorname{IP}[|\hat{p} p| > 0] = \mathcal{O}\left(\exp\left(-\theta n h_n^3 \alpha_n^2\right)\right),$
- b)  $1 \mathbb{P}\left[\forall 1 \le i \le p, \ \hat{\tau}_i \in ID_i \ et \ \hat{p} = p\right] = \mathcal{O}\left(\exp\left(-\theta n h_n^3 \alpha_n^2\right)\right),$

 $o\dot{u} ID_i = ]\tau_i - h_n, \tau_i + h_n[.$ 

#### Preuve du théorème 3.5.

a)

$$\mathbb{P}[\hat{p} - p > 0] \leq \mathbb{P}[\sup_{\bigcup_{i=1}^{p}[\tau_i + h_n, \tau_{i+1} - h_n]} |\hat{\gamma}(t)| \geq S_n],$$

$$= \mathbb{P}[\sup_{IC} |\hat{\gamma}(t)| \geq S_n].$$

$$\begin{split} \mathbb{P}[\hat{p} - p < 0] &\leq \mathbb{P}[\exists i = 1, \dots, p, | \sup_{\substack{[\tau_i - h_n, \tau_i + h_n] \\ [\tau_i - h_n, \tau_i + h_n]}} |\hat{\gamma}(t)| < S_n] \\ &+ \mathbb{P}[\exists i = 1, \dots, p, | \sup_{\substack{[\tau_i + h_n, \tau_{i+1} - h_n] \\ [\tau_i + h_n, \tau_{i+1} - h_n]}} |\hat{\gamma}(t)| \geq S_n], \\ &\leq \sum_{i=1}^p \mathbb{P}[\sup_{ID_i} |\hat{\gamma}(t)| < S_n] + \mathbb{P}[\sup_{IC} |\hat{\gamma}(t)| \geq S_n]. \end{split}$$

Or d'après lemme 3.3 il existe  $c_1, c_2, c_3, c_4$  et  $c_5 > 0$  telles que si  $c_1h_n < S_n < \min_{1 \le i \le p} |\gamma(\tau_i)| - c_5h_n^2$ , alors :

1) 
$$\mathbb{P}\left[\sup_{t\in IC} |\hat{\gamma}(t)| \ge S_n\right] \le \exp\left[-c_4 \frac{nh_n(S_n-c_1h_n)^2}{c_3+c_2(S_n-c_1h_n)}\right],$$
  
2)  $\mathbb{P}\left[\sup_{t\in ID_i} |\hat{\gamma}(t)| < S_n\right] \le \exp\left[-c_4 \frac{nh_n(S_n-c_1h_n)^2}{c_3+c_2(S_n-c_1h_n)}\right].$ 

En choisissant le seuil  $S_n$  tel que  $S_n \to 0$  et  $S_n = \alpha_n h_n$  où  $\alpha_n \to \infty$ , nous en déduisons qu'il existe une constante A > 0 telle que

1) 
$$\mathbb{P}\left[\sup_{t\in IC} |\hat{\gamma}(t)| \ge S_n\right] \le \exp\left(-Anh_n^3\alpha_n^2\right),$$
  
2) 
$$\mathbb{P}\left[\sup_{t\in ID_i} |\hat{\gamma}(t)| < S_n\right] \le \exp\left(-Anh_n^3\alpha_n^2\right).$$

Donc il existe  $\theta > 0$  tel que

$$\mathbb{P}[|\hat{p} - p| > 0] = \mathcal{O}\left(\exp\left(-\theta n h_n^3 \alpha_n^2\right)\right).$$
(3.15)

b)

$$\mathbb{P}\left[\exists 1 \leq i \leq \hat{p}, 1 \leq j \leq p, \quad \hat{\tau}_i \in [\tau_j + h_n, \tau_{j+1} - h_n] \text{ ou } \hat{p} \neq p\right] \\
\leq \mathbb{P}\left[\sup_{\bigcup_{i=1}^p [\tau_i + h_n, \tau_{i+1} - h_n]} |\hat{\gamma}(t)| \geq S_n\right] + \mathbb{P}\left[\hat{p} \neq p\right], \\
= \mathcal{O}\left(\exp\left(-\theta n h_n^3 \alpha_n^2\right)\right),$$

 $\operatorname{donc}$ 

$$1 - \mathbb{P}\left[\{\hat{\tau}_1, \dots, \hat{\tau}_p\} \subset \bigcup_{i=1}^p ]\tau_i - h_n, \tau_i + h_n [\text{ et } \hat{p} = p] = \mathcal{O}\left(\exp\left(-\theta n h_n^3 \alpha_n^2\right)\right)$$

Comme chaque intervalle  $]\tau_i - h_n, \tau_i + h_n[$  ne peut contenir qu'un seul  $\hat{\tau}_i$ , quitte à faire une permutation sur l'indice *i* on peut affirmer que

$$1 - \mathbb{P}\left[\forall 1 \le i \le p, \ \hat{\tau}_i \in ]\tau_i - h_n, \tau_i + h_n\left[ \text{ et } \hat{p} = p \right] = \mathcal{O}\left(\exp\left(-\theta n h_n^3 \alpha_n^2\right)\right).$$

**Théorème 3.6.** Supposons que les hypothèses C1-C6 sont satisfaites et que  $\min_{1 \le i \le p} |\gamma(\tau_i)| > S_0$ . Alors il existe  $\theta > 0$  tel que :

- a)  $IP[|\hat{p} p| > 0] = O(\exp(-\theta n h_n)),$
- b)  $1 \mathbb{P}\left[\forall 1 \le i \le p, \ \hat{\tau}_i \in ID_i \ et \ \hat{p} = p\right] = \mathcal{O}\left(\exp\left(-\theta nh_n\right)\right),$

$$o\dot{u} ID_i = ]\tau_i - h_n, \tau_i + h_n[.$$

**Preuve du théorème 3.6.** Reprendre la démonstration du théorème 3.5 en remplaçant  $S_n$  par  $S_0$ .

**Corollaire 3.1.** Supposents que les hypothèses C1-C6 sont satisfaites et que  $nh_nS_n^2/\log n \to \infty$ . Alors:

- a)  $\hat{p} \rightarrow p$ , p.s.
- b)  $\hat{\tau}_i \rightarrow \tau_i$ , p.s,  $i = 1, \dots, p$ .

La variance de  $\hat{\gamma}(t)$  étant de l'ordre de  $1/nh_n$ . D'après le théorème 3.1, pour que la procédure que nous avons définie pour estimer le nombre de ruptures soit consistante (fortement), il suffit que la suite des seuils  $(S_n)_n$  soit d'un ordre de grandeur plus grand que celui l'écart-type, i.e.  $\sigma(\hat{\gamma}(t)) = o(S_n)$  (cf la figure 1.2).

#### Preuve du corollaire 3.1.

Sous les hypothèses du corollaire 3.1, il est facile de voir que nous avons :

- 1)  $\operatorname{IP}[|\hat{p} p| > 0] = \mathcal{O}\left(\exp\left(-\theta n h_n S_n^2\right)\right),$
- 2)  $1 \operatorname{IP}\left[\forall 1 \le i \le p, \ \hat{\tau}_i \in ID_i \text{ et } \hat{p} = p\right] = \mathcal{O}\left(\exp\left(-\theta n h_n S_n^2\right)\right).$

Or  $nh_nS_n^2/\log n \to \infty$ . Il en résulte que la série  $\sum \exp(-\theta nh_nS_n^2)$  converge. Nous en déduisons, donc, que :

- a)  $\hat{p} \rightarrow p$ , pco,
- b)  $\hat{\tau}_i \to \tau_i$ , pco,  $i = 1, \dots, p$ .

Ceci termine la démonstration du corollaire.

Il est facile de vérifier le corollaire suivant qui nous sera utile dans le chapitre suivant.

**Corollaire 3.2.** Sous les hypothèses du théorème 3.5 et si nous supposons que

$$S_n = \frac{1}{\log\left(n\right)}$$

et  $h_n = n^{-\delta}$  avec  $0 < \delta < 1$ , alors:

$$1 - \mathbb{P}\left[\forall 1 \le i \le p, \ \hat{\tau}_i \in ]\tau_i - h_n, \tau_i + h_n[\ et \ \hat{p} = p] \right] = o(h_n^4).$$

# 3.6 Simulations

Nous explorons dans ce paragraphe, à l'aide de simulations, les performances de notre procédure d'estimation du nombre de ruptures pour des tailles d'échantillon raisonnables (n = 200et n = 500). Quatre modèles sont considérés: MO (modèle sans rupture), M1 (modèle avec une seule rupture), M2 et M3 (modèles avec 2 ruptures).

#### Exemple 1.(Loader, C.R.)

Dans cet exemple nous nous sommes intéressé seulement à la qualité de l'estimation du nombre de ruptures pour chacun des trois modèles, MO, M1 et M2. Le modèle MO est défini par une fonction continue. Les modèles M1 et M2 sont obtenus en rajoutant une (M1), puis deux (M2) discontinuités. 1000 échantillons de taille n (200 ou 500) sont simulés pour chacun des trois modèles suivants:

(MO)

$$Y_i = 4\sin\left(5X_i\right) + 3X_i + \varepsilon_i$$

(M1)

$$Y_i = 4\sin(5X_i) + 3X_i + 1_{[0.3,1]}(X_i) + \varepsilon_i$$

(M2)

$$Y_i = 4\sin(5X_i) + 3X_i + \mathbf{1}_{[0,3,1]}(X_i) - \mathbf{1}_{[0,7,1]}(X_i) + \varepsilon_i$$

avec  $\varepsilon_i \sim 0.5 \mathcal{N}(0, 1)$  et  $X_i \sim \mathcal{U}([0, 1])$ .

Nous avons regroupé dans les tableaux 3.1 et 3.2 les distributions observées des estimateurs du nombre de ruptures pour chaque modèle. La taille des échantillons considérés dans le tableau 3.1 est de 200 et celle de ceux considérés dans le tableau 3.2 est de 500. Compte tenu du rapport signal sur bruit (RSB = 2), les taux d'exacte découverte du nombre de ruptures apparaissent satisfaisants. Les résultats s'améliorent lorsqu'on passe de n = 200 à n = 500. Dans la figure 3.2 nous avons visualisé un échantillon de taille n = 200 de chacun des modèles MO, M1 et M2 ainsi que les fonctions de régression qui leur sont associées.



Figure 3.2 : Un échantillon de taille n = 200 pour les modèles MO, M1 et M2.

$\hat{p} \setminus p$	МО	M1	M2
0	778	239	6
1	217	737	380
2	5	24	614
3	0	0	0

**Table 3.1 :** Estimation du nombre de ruptures :  $h_n = 0.2$  et n = 200.

**Table 3.2** : Estimation du nombre de ruptures :  $h_n = 0.15$  et n = 500.

$\hat{p} \setminus p$	МО	M1	M2
0	913	23	0
1	84	960	38
2	3	17	962
3	0	0	0

#### Exemple 2.(Loader, C.R.)

Dans ce deuxième exemple nous avons étudié la qualité des estimations des localisations ainsi que des amplitudes de rupture lorsque leur nombre est correctement estimé. De façon précise, 1000 échantillons de taille n (n = 200 ou n = 500) sont simulés à partir du modèle suivant :

(M2)

$$Y_i = 4\sin(5X_i) + 3X_i + \mathbf{1}_{[0,3,1]}(X_i) - \mathbf{1}_{[0,7,1]}(X_i) + \varepsilon_i$$

avec  $\varepsilon_i \sim 0.5 \mathcal{N}(0, 1)$  et  $X_i \sim \mathcal{U}([0, 1])$ .

Pour chacun des échantillons pour lequels le nombre de ruptures est bien estimé (c'est-à-dire  $\hat{p} = 2$ ) nous avons représenté dans les figures 3.3 (pour les échantillons de taille n = 200) et 3.5 (pour les échantillons de taille n = 500) les localisations estimées de ces deux ruptures. Dans les figures 3.4 et 3.6, nous avons rapporté les estimations des amplitudes de ces ruptures. Nous remarquons qu'une bonne estimation du nombre de ruptures est accompagnée la plupart du temps d'une bonne localisation des ruptures ainsi que d'une bonne estimation de leurs amplitudes. Ceci est cohérent avec les résultats asymptotiques établis dans ce chapitre (voir théorème 3.5).


Figure 3.3 : Estimations de  $\hat{\tau}_1$  et  $\hat{\tau}_2$  sachant  $\hat{p} = 2$ . n = 200 et  $h_n = 0.2$ .



**Figure 3.4** : Estimations de  $\hat{\gamma}(\hat{\tau}_1)$  et  $\hat{\gamma}(\hat{\tau}_2)$  sachant  $\hat{p} = 2$ . n = 200 et  $h_n = 0.2$ .



**Figure 3.5** : Estimations de  $\hat{\tau}_1$  et  $\hat{\tau}_2$  sachant  $\hat{p} = 2$ . n = 500 et  $h_n = 0.15$ .



**Figure 3.6** : Estimations de  $\hat{\gamma}(\hat{\tau}_1)$  et  $\hat{\gamma}(\hat{\tau}_2)$  sachant  $\hat{p} = 2$ . n = 500 et  $h_n = 0.15$ .

## Exemple 3.(Gijbels & al)

Dans ce troisième exemple, nous analysons le comportement de notre procédure d'estimation du nombre de ruptures en fonction du choix de la fenêtre de lissage  $h_n$ . ce que nous observons est cohérent avec les constatations faites dans les chapitres précédents :

- 1) L'utilisation de faibles tailles de fenêtre de lissage se traduit par une augmentation de la variance du processus  $\hat{\gamma}(t)$ . Les perturbations conduisent à des extrema artificiels et à une surestimation du nombre de discontinuités.
- 2) Le choix de fenêtres de lissage de grandes tailles conduit à gommer certaines propriétés locales de  $\hat{\gamma}(t)$  comme les maxima relatifs, ce qui conduit à une sous-estimation du nombre de ruptures dans la fonction de régression.

200 échantillons de taille n = 250 sont simulés à partir du modèle M3 suivant :

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

où

$$m(x) = \begin{cases} \exp(-2(x-0.35)) - 1 & \text{si } x \in [0, 0.35[, \\ \exp(-2(x-0.35)) & \text{si } x \in [0.35, 0.65[, \\ \exp(2(x-0.65)) + \exp(-0.6) - 2 & \text{si } x \in [0.65, 1], \end{cases}$$

avec  $\varepsilon_i \sim \mathcal{N}(0,1)$  et  $X_i \sim \mathcal{U}([0,1])$ .

Les résultats de ces simulations sont présentés dans le tableau 3.3 ci dessous.

$\hat{p} \setminus h_n$	0.06	0.09	0.12	0.15	0.18
0	0	0	0	0	0
1	4	6	18	25	155
2	55	169	182	175	55
3	88	25	0	0	0
4	40	0	0	0	0
5	13	0	0	0	0

Table 3.3 : Estimation du nombre de rupture : modèle M3 avec  $\sigma^2 = 0.1$  et n = 250.

# Chapitre 4

# Segmentation

# 4.1 Introduction

Le problème abordé dans ce chapitre est celui de la reconstitution de la fonction de régression lorsque celle-ci présente un nombre fini de singularités. Dans la littérature, on distingue deux approches pour estimer une fonction de régression présentant des points de rupture. D'une part les méthodes dites directes qui n'utilisent pas d'estimation préalable des points de rupture ( cf [MO86] pour un exemple de cette façon de procéder). D'autre part les approches indirectes ou de segmentation. Ce dernier type d'approches procède en deux étapes : une première étape qui consiste à estimer le nombre de points de rupture et leurs localisations en utilisant une méthode quelconque de détection de ruptures. Ensuite, sur chaque intervalle défini par deux points de rupture consécutifs, la fonction de régression est estimée en utilisant les observations de ce seul intervalle. Ainsi les méthodes de segmentation suivent en général le schéma suivant :

- i) Estimer le nombre de ruptures de la fonction de régression dans l'intervalle d'étude [a,b]:  $\hat{p}$ ;
- ii) Localiser les  $\hat{p}$  ruptures dans le signal:  $\hat{\tau}_1 < \ldots < \hat{\tau}_{\hat{p}}$ . Par convention, on note

$$\hat{\tau}_0 = a \quad \text{et} \quad \hat{\tau}_{\hat{p}+1} = b.$$

iii) Segmenter les données: Supposons qu'on dispose de *n* observations  $D = \{(x_i, y_i); 1 \le i \le n\}$ . On construit une partition  $(D_j)_{0 \le j \le \hat{p}}$  de l'ensemble des observations définie par :

$$D_j = \{(x_i, y_i); 1 \le i \le n \text{ et } \hat{\tau}_j \le x_i < \hat{\tau}_{j+1}\}.$$

iv) Pour tout  $\hat{\tau}_j \leq x < \hat{\tau}_{j+1}$ , estimer m(x) par une méthode standard pour les fonctions lisses en utilisant seulement les données  $D_j$ .



Figure 4.1 : Les méthodes de segmentation: Algorithme.

Ici nous allons estimer le signal en segmentant les données toujours dans le cadre d'estimation par régression linéaire locale. Dans un premier temps, nous estimons le nombre de ruptures et leurs localisations  $\tau_1, \ldots, \tau_p$  par régression linéaire locale avec une taille de fenêtre  $g_n$  en utilisant des noyaux dissymétriques  $K_+(\cdot)$  et  $K_-(\cdot)$ . Dans un deuxième temps, nous effectuons sur chaque intervalle  $[\hat{\tau}_i, \hat{\tau}_{i+1}]$  l'estimation de la fonction de régression en utilisant une deuxième taille de fenêtre  $h_n$  pour un noyau  $K(\cdot)$  standard à support [-1, 1]. Nous montrons que, dans les conditions habituelles, avec un choix convenable pour  $g_n$ , on obtient un estimateur  $\hat{m}(\cdot)$  qui a la même vitesse de convergence que dans le cas où  $m(\cdot)$  ne présente pas de ruptures. Il s'agit bien entendu d'une conséquence de la propriété d'absence d'effets de bords de la régression polynômiale locale. Les estimateurs de type Nadaraya-Watson ou Gasser-Müller ne peuvent conduire au même résultat qu'en modifiant le noyau au voisinage des points de rupture.

# 4.2 Un seul point de rupture. Effet de la segmentation en deux intervalles

Nous commençons par étudier, dans ce paragraphe, la situation d'un seul point de rupture  $\tau$ et un point de segmentation. Nous explicitons le comportement asymptotique de l'erreur quadratique moyenne (MSE(x)) de  $\hat{m}(x)$  d'une part au voisinage du point de segmentation, d'autre part au voisinage du point de rupture. On observe une détérioration de la qualité de l'estimation uniquement au voisinage du point de rupture  $\tau$ . De plus, lorsque le point de segmentation est "suffisamment proche" de  $\tau$ , cette détérioration n'a pas de répercussions au niveau du comportement asymptotique du MISE. En particulier, dans le cas où  $\tau$  est connu, la segmentation en deux intervalles  $[0, \tau[$  et  $[\tau, 1]$  ne modifie pas la qualité (au sens du MISE) de l'estimation de  $m(\cdot)$  par rapport à la situation où  $m(\cdot)$  ne présente pas de rupture.

Nous supposons, dans ce paragraphe, que nous savons que la fonction de régression admet un seul point de rupture,  $\tau$ . L'intervalle d'étude [a, b] est égal à [0, 1]. Nous montrons au paragraphe 4.4 que, lorsque le nombre de ruptures et leurs localisations sont estimés, les vitesses de convergence obtenues dans ce paragraphe sont conservées.

#### 4.2.1 Notations et hypothèses

Nous présentons ici les notations et les hypothèses considérées dans ce chapitre :

- D1. La fonction de densité marginale  $f_X(\cdot)$  est supposée continue et strictement positive sur l'intervalle [0, 1].
- D2. La fonction de régression  $m(\cdot) = \mathbb{E}(Y \mid X = \cdot)$  est deux fois continûment dérivable en tout point différent de  $\tau$  et est *càdlàq* au point  $\tau$  ainsi que ses deux premières dérivées.
- D3. La variance conditionnelle  $\sigma^2(x) = Var(Y \mid X = x)$  est continue sur l'intervalle [0,1].
- D4. Le point de discontinuité de la fonction de régression,  $\tau$ , appartient à l'intervalle ]0, 1[.

Comme dans les chapitres précédents, nous posons pour  $\nu = 0, 1, 2$ 

i) 
$$m_{-}^{(\nu)}(\tau) = \lim_{t \nearrow \tau} m^{(\nu)}(t),$$
  
ii)  $m_{+}^{(\nu)}(\tau) = \lim_{t \searrow \tau} m^{(\nu)}(t),$   
iii)  $\gamma^{(\nu)}(\tau) = m_{+}^{(\nu)}(\tau) - m_{-}^{(\nu)}(\tau)$ 

Ainsi  $\gamma(\tau) = m_+(\tau) - m_-(\tau)$  est l'amplitude de la discontinuité de la fonction de régression au point  $\tau$ .

Pour tout v, nous posons

i) 
$$K_l(v) = \int_v^1 x^l K(x) dx$$
,  
ii)  $B_K(v) = \frac{K_2^2(v) - K_3(v)K_1(v)}{K_2(v)K_0(v) - K_1^2(v)}$ ,  
iii)  $L_l(v) = \int_v^1 x^l K^2(x) dx$ ,

iv) 
$$V_K^2(v) = \frac{\int_v^1 (K_2(v) - xK_1(v))^2 K^2(x) dx}{(K_2(v)K_0(v) - K_1^2(v))^2}.$$

Dans la suite, le noyau  $K(\cdot)$  est supposé symétrique, i.e :

$$K(-x) = K(x).$$

**Remarque 4.1.** Comme  $K(\cdot)$  a pour support l'intervalle [-1, 1], pour tout  $v \leq -1$ , on a  $K_l(v) = K_l(-1) = K_l$ . Cette propriété est également vraie pour les quantités  $B_K(v)$ ,  $L_l(v)$  et  $V_K(v)$ .

Nous notons  $\hat{m}(\cdot, t)$  l'estimateur de  $m(\cdot)$ , où t désigne le point de segmentation. L'estimateur  $\hat{m}(\cdot, t)$  est obtenu de la façon suivante :

- Lorsque  $x \in [0, t[$ , on ajuste localement une droite  $\alpha_x + \beta_x(x u)$  par la régression linéaire locale (RLL) pour obtenir  $\hat{m}(x, t) = \hat{\beta}_x$  en utilisant seulement les observations  $X_i < t$ .
- Lorsque  $x \in [t, 1]$ , on effectue la même construction en utilisant les observations  $X_i \ge t$ .

En d'autres termes l'estimateur de m(x) est défini par :

$$\hat{m}(x,t) = \frac{\sum_{i=1}^{n} w_i(x,t) Y_i}{\sum_{i=1}^{n} w_i(x,t)},$$
(4.1)

avec

$$w_{i}(x,t) = \begin{cases} 0 & \text{si } i \in B(x,t), \\ K(\frac{x-X_{i}}{h_{n}})(S_{2}(x,t) - (x-X_{i})S_{1}(x,t)) & \text{sinon}, \end{cases}$$
(4.2)

où pour l = 0, 1, 2, ...

$$S_l(x,t) = \begin{cases} \sum_{X_i < t} (x - X_i)^l K(\frac{x - X_i}{h_n}) & \text{si } x < t, \\ \\ \sum_{X_i \ge t} (x - X_i)^l K(\frac{x - X_i}{h_n}) & \text{si } x \ge t, \end{cases}$$
(4.3)

 $\operatorname{et}$ 

$$B(x,t) = \{ 1 \le i \le n; x < t \ge X_i \text{ ou } X_i < t \ge x \}$$



**Figure 4.2** : Construction de l'estimateur  $\hat{m}(\cdot, t)$  : t = 0.7 et x = 0.65, 0.75.

**Remarque 4.2.** Il est important de noter que l'estimateur  $\hat{m}(x,t)$  ainsi obtenu peut être défini comme la valeur de b = b(x) solution du problème de minimisation :

$$\min_{(a,b)} \sum_{X_i \in D(x,t)}^n [Y_i - (a(x - X_i) + b)]^2 K(\frac{x - X_i}{h_n}),$$
(4.4)

où D(x,t) désigne l'intervalle  $[0,t[si \ x < t \ et \ [t,1] \ sinon.$ 

# 4.2.2 Résultats asymptotiques

Nous examinons d'abord la situation d'un point de segmentation loin du point de rupture  $\tau$ . La première partie du théorème suivant est une reformulation de l'expression de l'erreur

quadratique moyenne pour une fonction lisse, que ce soit aux bords ou loin des bords. Nous démontrons ces résultats, pour mettre plus clairement en évidence les éléments qui sont modifiés lorsque la fonction n'est plus lisse. Dans ce dernier cas, au voisinage du point de rupture, le biais est de l'ordre de l'amplitude de cette rupture c'est à dire  $\mathcal{O}_P(1)$ . L'ordre de grandeur de la variance n'est pas modifié par la présence de la rupture. Il en résulte que l'erreur quadratique moyenne au voisinage du point de rupture est  $\mathcal{O}_P(1)$ .

**Théorème 4.1.** Supposons que les hypothèses D1-D4 sont vérifiées et que  $t = \tau + rh_n$ , avec |r| > 1 et  $x = t + ch_n$ , alors

- 1)  $si |x \tau| > h_n$ , a)  $\operatorname{IE}[\hat{m}(x,t) \mid X_1, \dots, X_n] - m(x) = \frac{1}{2}m^{(2)}(x)h_n^2 B_K(-|c|) + o_P(h_n^2),$ b  $Var[\hat{m}(x,t) \mid X_1, \dots, X_n] = \frac{\sigma^2(x)}{nh_n f(x)}V_K^2(-|c|) + o_P(\frac{1}{nh_n}),$
- 2)  $si x = \tau + h_n z, |z| \le 1$

a) 
$$\mathbb{E}[\hat{m}(x,t) \mid X_1, \dots, X_n] - m(x) = -sign(z)c_{1,r}(z)\gamma(\tau) + o_P(1),$$
  
b)  $Var[\hat{m}(x,t) \mid X_1, \dots, X_n] = \frac{\sigma^2(x)}{nh_n f(x)}V_K^2(-|z-r|) + o_P(\frac{1}{nh_n})$ 

avec

et

$$c_{1,r}(z) = \frac{K_2(-|z-r|)K_0(|z|) - K_1(-|z-r|)K_1(|z|)}{K_2(-|z-r|)K_0(-|z-r|) - K_1^2(-|z-r|)},$$

$$sign(z) = \begin{cases} -1 & si \ z < 0, \\ 1 & si \ z \ge 0. \end{cases}$$

**Remarque 4.3.** Nous constatons qu'avec une segmentation basée sur un point de segmentation loin de  $\tau$ , la vitesse de convergence de l'erreur quadratique moyenne se détériore au voisinage de  $\tau$  de la même façon que si l'on ne faisait pas de segmentation. Par contre cette vitesse reste optimale au voisinage de t, alors qu'elle se dégrade avec des estimateurs tels que ceux de Nadaraya-Watson ou Gasser-Müller. Néanmoins, dans cette situation, la vitesse de convergence de l'erreur quadratique moyenne intégrée est la même pour la régression linéaire locale et les estimateurs de Nadaraya-Watson et Gasser-Müller (cf corollaire 4.1 et la remarque qui suit).

Nous présentons ci-dessous les graphiques des fonctions  $B_K^2(-|z|)$ ,  $V_K^2(-|z|)$  et  $c_{1,r}(z)$  pour le noyau d'Epanechnikov :

$$K(x) = \begin{cases} \frac{3}{4}(1-x^2) & \text{si } |x| \le 1, \\ \\ 0 & \text{sinon.} \end{cases}$$



Figure 4.3 :  $B_K^2(-|z|)$ .



Figure 4.4 :  $V_K^2(-|z|)$ .



**Figure 4.5** :  $c_{1,r}(z)$  pour 9 valeurs de  $r = 1, 1 + \frac{1}{9}, \ldots, 2$ .

## Preuve théorème 4.1.

L'erreur quadratique moyenne conditionnelle (MSE) de  $\hat{m}(x,t)$  est donnée par

$$\begin{split} & \mathbb{E}\left[\left(\hat{m}_{+}(x,t)-m(x)\right)^{2}\mid X_{1},\ldots,X_{n}\right] \\ &=\left[\frac{\sum_{i=1}^{n}w_{i}(x,t)\left(m(X_{i})-m(x)\right)}{\sum_{i=1}^{n}w_{i}(x,t)}\right]^{2}+\frac{\sum_{i=1}^{n}\left(w_{i}(x,t)\right)^{2}\sigma^{2}(X_{i})}{\left(\sum_{i=1}^{n}w_{i}(x,t)\right)^{2}}, \\ &=B_{n}^{2}+V_{n}. \end{split}$$

 $B_n$  représente le biais conditionnel de  $\hat{m}(x,t)$  et  $V_n$  sa variance conditionnelle. En remarquant que  $\sum_{i=1}^n w_i(x,t)(X_i-x) = 0$ , nous obtenons

$$B_n = \frac{\sum_{i=1}^n w_i(x,t) (m(X_i) - m(x))}{\sum_{i=1}^n w_i(x,t)},$$
  
=  $\frac{\sum_{i=1}^n w_i(x,t) (m(X_i) - m(x) - (X_i - x)m'(x))}{\sum_{i=1}^n w_i(x,t)}.$ 

En remplaçant les quantités  $w_i(x,t)$  par leurs expressions (4.2),  $B_n$  et  $V_n$  s'écrivent de la manière suivante :

$$B_n = \frac{S_2(x,t)R_0(x,t) - S_1(x,t)R_1(x,t)}{S_2(x,t)S_0(x,t) - (S_1(x,t))^2},$$
(4.5)

 $\operatorname{et}$ 

$$V_{n} = \frac{\sum_{i=1}^{n} (w_{i}(x,t))^{2} \sigma^{2}(X_{i})}{(\sum_{i=1}^{n} w_{i}(x,t))^{2}},$$
  

$$= \frac{(S_{2}(x,t))^{2} T_{0}(x,t) - 2S_{2}(x,t)S_{1}(x,t)T_{1}(x,t)}{(S_{2}(x,t)S_{0}(x,t) - (S_{1}(x,t))^{2})^{2}} + \frac{(S_{1}(x,t))^{2} T_{2}(x,t)}{(S_{2}(x,t)S_{0}(x,t) - (S_{1}(x,t))^{2})^{2}},$$
(4.6)

avec pour tout x < t

- a)  $S_l(x,t) = \sum_{X_i < t} Z_{i,1},$
- b)  $T_l(x,t) = \sum_{X_i < t} Z_{i,2},$

c) 
$$R_l(x,t) = \sum_{X_i < t} Z_{i,3}.$$

pour tout  $x \ge t$ 

- d)  $S_l(x,t) = \sum_{X_i \ge t} Z_{i,1},$
- e)  $T_l(x,t) = \sum_{X_i \ge t} Z_{i,2},$

f) 
$$R_l(x,t) = \sum_{X_i \ge t} Z_{i,3}$$

et où

i) 
$$Z_{i,1} = (x - X_i)^l K(\frac{x - X_i}{h_n}),$$

ii) 
$$Z_{i,2} = (x - X_i)^l K^2(\frac{x - X_i}{h_n}) \sigma^2(X_i),$$

iii) 
$$Z_{i,3} = (x - X_i)^l K(\frac{x - X_i}{h_n}) (m(X_i) - m(x) - (X_i - x)m'(x)).$$

Cette deuxième façon d'écrire le biais et la variance va nous être utile par la suite.

1) Supposons que  $|x - \tau| > h_n$ 

**Lemme 4.1.** Soit  $t = \tau + rh_n$ , |r| > 1 et  $x = t + ch_n$ . Sous les hypothèses D1-D4 nous avons :

a) 
$$S_l(x,t) = nh_n^{l+1} f_X(x) \left( (-sign(c))^l K_l(-|c|) + o_P(1) \right),$$
  
b)  $T_l(x,t) = nh_n^{l+1} f_X(x) \sigma^2(x) \left( (-sign(c))^l L_l(-|c|) + o_P(1) \right),$   
c)  $R_l(x,t) = \frac{1}{2} m^{(2)}(x) nh_n^{l+3} f_X(x) \left( (-sign(c))^l K_{l+2}(-|c|) + o_P(1) \right).$ 

### Preuve du lemme 4.1.

Posons  $x = t + ch_n$ . Nous allons montrer le lemme 4.1 seulement pour x < t, c'est à dire pour c < 0. La démonstration pour  $x \ge t$  est analogue.

D'après l'inégalité de Bienaymé-Tchebycheff, nous avons

$$\sum_{X_i < t} Z_{i,j} = \mathbb{E} \sum_{X_i < t} Z_{i,j} + \mathcal{O}_P \left( \sqrt{Var\left(\sum_{X_i < t} Z_{i,j}\right)} \right),$$
  
$$= n \mathbb{E} \left[ Z_{1,j} \mathbb{I}_{\{X_1 < t\}} \right] + \mathcal{O}_P \left( \sqrt{n \mathbb{E} \left[ Z_{1,j}^2 \mathbb{I}_{\{X_1 < t\}} \right]} \right),$$

avec

$$\mathbb{E}\left[Z_{1,1} 1\!\!1_{\{X_1 < t\}}\right] = h_n^{l+1} \int_{\frac{x-t}{h_n}}^1 u^l K(u) f_X(x - h_n u) du,$$
  
=  $h_n^{l+1} f_X(x) \left(K_l(c) + o(1)\right).$ 

La fonction o(1) peut être choisie indépendamment de x puisque  $f_X(\cdot)$  est uniformément continue. De la même manière, nous obtenons

$$\mathbb{E}\left[Z_{1,1}^{2} \prod_{\{X_{1} < t\}}\right] = h_{n}^{2l+1} \int_{\frac{x-t}{h_{n}}}^{1} u^{2l} K^{2}(u) f_{X}(x - h_{n}u) du,$$
  
$$= \mathcal{O}\left(h_{n}^{2l+1}\right).$$

Donc pour tout  $x = t + ch_n$  avec c < 0, on a

$$S_{l}(x,t) = n \mathbb{E} \left[ Z_{1,1} 1\!\!1_{\{X_{1} < t\}} \right] + \mathcal{O}_{P} \left( \sqrt{n \mathbb{E} \left[ Z_{1,1}^{2} 1\!\!1_{\{X_{1} < t\}} \right]} \right),$$
  
$$= n h_{n}^{l+1} f_{X}(x) \left( K_{l}(c) + o_{P}(1) \right).$$

Les approximations asymptotiques de  $\mathbb{E}[Z_{1,2}]$  et  $\mathbb{E}[Z_{1,2}^2]$  sont obtenues de façon similaire. En effet

$$\mathbb{E}\left[Z_{1,2} \prod_{\{X_1 < t\}}\right] = h_n^{l+1} \int_{\frac{x-t}{h_n}}^1 u^l K^2(u) f_X(x - h_n u) \sigma^2(x - h_n u) du,$$
  
$$= h_n^{l+1} f_X(x) \sigma^2(x) \left(L_l(c) + o(1)\right),$$

$$\mathbb{E}\left[Z_{1,2}^{2} 1\!\!1_{\{X_{1} < t\}}\right] = h_{n}^{2l+1} \int_{\frac{x-t}{h_{n}}}^{1} u^{2l} K^{4}(u) f_{X}(x-h_{n}u) \sigma^{4}(x-h_{n}u) du,$$
  
$$= \mathcal{O}\left(h_{n}^{2l+1}\right).$$

 $\operatorname{Ainsi}$ 

$$T_{l}(x,t) = n\mathbb{E}\left[Z_{1,2}\mathbb{1}_{\{X_{1} < t\}}\right] + \mathcal{O}_{P}\left(\sqrt{n\mathbb{E}\left[Z_{1,2}^{2}\mathbb{1}_{\{X_{1} < t\}}\right]}\right),$$
  
$$= nh_{n}^{l+1}f_{X}(x)\sigma^{2}(x)\left(L_{l}(c) + o_{P}(1)\right).$$

En utilisant la continuité uniforme de  $f_X(\cdot)$ , le fait que pour tout  $|x - \tau| > h_n$ ,

$$m(x - h_n u) - m(x) + h_n u m'(x) = \frac{m^{(2)}(x)}{2} h_n^2 + o(h_n^2),$$

et la continuité uniforme de  $m^{(2)}(\cdot)$  sur chacun des intervalles  $[0, \tau]$  et  $[\tau, 1]$ , nous obtenons

$$\begin{split} & \mathbb{E}\left[Z_{1,3}\Pi_{\{X_1 < t\}}\right] \\ &= h_n^{l+1} \int_{\frac{x-t}{h_n}}^1 u^l K(u) f_X(x - h_n u) \left(m(x - h_n u) - m(x) + h_n u m'(x)\right) du \,, \\ &= \frac{1}{2} m^{(2)}(x) h_n^{l+3} f_X(x) \left(K_{l+2}(c) + o(1)\right) \,. \end{split}$$

Un calcul similaire conduit à

$$\begin{split} & \mathbb{E}\left[Z_{1,3}^{2} \mathbf{1}_{\{X_{1} < t\}}\right] \\ &= h_{n}^{2l+1} \int_{\frac{x-t}{h_{n}}}^{1} u^{l} K^{2}(u) f_{X}(x - h_{n}u) \left(m(x - h_{n}u) - m(x) + h_{n}um'(x)\right)^{2} du \,, \\ &= \mathcal{O}\left(h_{n}^{2l+5}\right) \,, \end{split}$$

d'où

$$\begin{aligned} R_l(x,t) &= n \mathbb{E}\left[Z_{1,3} 1\!\!1_{\{X_1 < t\}}\right] + \mathcal{O}_P\left(\sqrt{n \mathbb{E}\left[Z_{1,3}^2 1\!\!1_{\{X_1 < t\}}\right]}\right), \\ &= \frac{1}{2} m^{(2)}(x) n h_n^{l+3} f_X(x) \left(K_{l+2}(c) + o_P(1)\right). \end{aligned}$$

Pour obtenir le résultat sous la forme énoncée dans le lemme, il suffit de remarquer que pour  $c \leq 0$ 

$$\int_{c}^{1} u^{l} K(u) du = (-sign(c))^{l} K_{l}(-|c|),$$

que pour  $c \ge 0$  et l pair

$$\int_{-1}^{c} u^{l} K(u) du = \int_{-c}^{1} u^{l} K(u) du = (-sign(c))^{l} K_{l}(-|c|),$$

et enfin que pour  $c \ge 0$  et l impair

$$\int_{-1}^{c} u^{l} K(u) du = -\int_{-c}^{1} u^{l} K(u) du,$$
  
=  $-K_{l}(-c),$   
=  $(-sign(c))^{l} K_{l}(-|c|).$ 

Ces égalités, qui sont encore satisfaites lorsqu'on remplace K par L, seront aussi utiles dans les démonstrations d'autres résultats de ce chapitre. Ceci achève la démonstration du lemme 4.1.

Enfin, pour finir la démonstration de la première partie du théorème 4.1, il suffit de combiner le lemme 4.1 et les égalités (4.5) et (4.6).

2) Supposents  $x = x_n = \tau + h_n z$ ,  $|z| \le 1$  (avec toujours  $|t - \tau| > h_n$ )

**Lemme 4.2.** Soit  $t = \tau + rh_n$ , |r| > 1 et  $x = t + ch_n = \tau + zh_n$ . Sous les hypothèses D1-D4, nous avons:

a) 
$$S_l(x_n, t) = nh_n^{l+1} f_X(\tau) \left( (sign(c))^l K_l(-|c|) + o_P(1) \right),$$
  
b)  $T_l(x_n, t) = nh_n^{l+1} f_X(\tau) \sigma^2(\tau) \left( (sign(c))^l L_l(-|c|) + o_P(1) \right),$ 

c) 
$$R_l(x_n,t) = -(sign(z))^{l+1}nh_n^{l+1}f_X(\tau)\gamma(\tau) (K_l(|z|) + o_P(1)).$$

#### Preuve du lemme 4.2.

La démonstration des résultats concernant  $S_l(\cdot, \cdot)$  et  $T_l(\cdot, \cdot)$  est identique à celle donnée dans le lemme précédent, puisque  $S_l(\cdot, \cdot)$  et  $T_l(\cdot, \cdot)$  ne font intervenir que les  $X_i$ . Montrons le résultat pour  $R_l(\cdot, \cdot)$ .

$$\begin{split} &\frac{1}{h_n^{l+1}} \mathbb{E}\left[ Z_{1,3} \coprod_{\{X_1 < t\}} \right] \\ &= \int_{\frac{x_n - t}{h_n}}^1 u^l K(u) f_X(x_n - h_n u) \left( m(x_n - h_n u) - m(x_n) \right) du + o(1), \\ &= \int_{\frac{x_n - t}{h_n}}^z u^l K(u) f_X(\tau + h_n(z - u)) \left( m(\tau + h_n(z - u)) - m(\tau + h_n z) \right) du \\ &+ \int_z^1 u^l K(u) f_X(\tau + h_n(z - u)) \left( m(\tau + h_n(z - u)) - m(\tau + h_n z) \right) du \\ &+ o(1). \end{split}$$

En utilisant un développement de Taylor à l'ordre 1, il est facile de déduire que :

• Si  $z \ge 0$ , alors

$$\begin{aligned} &\frac{1}{h_n^{l+1}} \mathbb{E}\left[Z_{1,3} \prod_{\{X_1 < t\}}\right] \\ &= \int_z^1 u^l K(u) f_X(\tau + h_n(z - u)) \left(m(\tau + h_n(z - u)) - m(\tau + h_n z)\right) du \\ &+ o(1), \\ &= f_X(\tau) \gamma(\tau) K_l(z) + o(1). \end{aligned}$$

• Si z < 0, alors

$$\begin{aligned} &\frac{1}{h_n^{l+1}} \mathbb{E}\left[Z_{1,3} \prod_{\{X_1 < t\}}\right] \\ &= \int_{\frac{x_n - t}{h_n}}^z u^l K(u) f_X(\tau + h_n(z - u)) \left(m(\tau + h_n(z - u)) - m(\tau + h_n z)\right) du \\ &+ o(1), \\ &= \int_{-1}^z u^l K(u) f_X(\tau + h_n(z - u)) \left(m(\tau + h_n(z - u)) - m(\tau + h_n z)\right) du \\ &+ o(1), \\ &= f_X(\tau) \gamma(\tau) (sign(z))^l K_l(-z) + o(1). \end{aligned}$$

Par un calcul analogue, nous obtenons également

$$\begin{split} & \mathbb{E}\left[Z_{1,3}^{2}\Pi_{\{X_{1} < t\}}\right] \\ &= h_{n}^{2l+1} \int_{\frac{x_{n}-t}{h_{n}}}^{1} u^{l} K^{2}(u) f_{X}(x_{n} - h_{n}u) \left(m(x_{n} - h_{n}u) - m(x_{n})\right)^{2} du \,, \\ &= \mathcal{O}\left(h_{n}^{2l+1}\right). \end{split}$$

D'où pour tout  $z \ge 0$ ,

$$\begin{aligned} R_{l}(x_{n},t) &= n \mathbb{E}\left[Z_{1,3} 1\!\!1_{\{X_{1} \ge t\}}\right] + \mathcal{O}_{P}\left(\sqrt{n \mathbb{E}\left[Z_{1,3}^{2} 1\!\!1_{\{X_{1} \ge t\}}\right]}\right), \\ &= -nh_{n}^{l+1} f_{X}(\tau) \gamma(\tau) \left(K_{l}(z) + o_{P}(1)\right), \\ &= -(sign(z))^{l+1} nh_{n}^{l+1} f_{X}(\tau) \gamma(\tau) \left(K_{l}(|z|) + o_{P}(1)\right), \end{aligned}$$

et pour tout z < 0,

$$\begin{aligned} R_{l}(x_{n},t) &= n \mathbb{E} \left[ Z_{1,3} \mathrm{I\!I}_{\{X_{1} < t\}} \right] + \mathcal{O}_{P} \left( \sqrt{n \mathbb{E} \left[ Z_{1,3}^{2} \mathrm{I\!I}_{\{X_{1} < t\}} \right]} \right), \\ &= n h_{n}^{l+1} f_{X}(\tau) \gamma(\tau) \left( (-1)^{l} K(-z) + o_{P}(1) \right), \\ &= -(sign(z))^{l+1} n h_{n}^{l+1} f_{X}(\tau) \gamma(\tau) \left( K_{l}(|z|) + o_{P}(1) \right). \end{aligned}$$

Refaisant les calculs pour  $x_n > t$  et utilisant les relations données à la fin de la démonstration du lemme 4.1, nous obtenons la forme annoncée pour  $R_l(x_n, t)$ .

Comme pour la première partie, pour terminer la démonstration de cette deuxième partie, il suffit de combiner le lemme 4.2 et les égalités (4.5) et (4.6). Ceci achève la démonstration du théorème 4.1.

**Corollaire 4.1.** Sous les hypothèses du théorème 4.1, si  $t = \tau + rh_n$ , avec |r| > 1, alors

$$\int_{0}^{1} \mathbb{E} \left[ \left( \hat{m}(x,t) - m(x) \right)^{2} \mid X_{1}, \dots, X_{n} \right] dx$$
  
=  $\left( \gamma^{2}(\tau) \int_{-1}^{1} c_{1,r}^{2}(z) dz \right) h_{n} + \frac{1}{nh} \left( V_{K}^{2} \int_{0}^{1} \frac{\sigma^{2}(x)}{f(x)} dx \right) + o_{P} \left( h_{n} + \frac{1}{nh_{n}} \right).$ 

**Remarque 4.4.** Le lissage des données par la régression linéaire locale sans la segmentation (a fortiori pour les estimateurs de Nadaraya-Watson ou Gasser-Müller) engendrera une erreur

quadratique moyenne intégrée de l'ordre de  $\mathcal{O}_P\left(h_n + \frac{1}{nh_n}\right)$ . En choisissant une fenêtre de lissage  $h_n$  proportionnelle à  $n^{-\frac{1}{2}}$ , nous obtiendrons dans le meilleur des cas

$$\int_0^1 \mathbb{E} \left[ (\hat{m}(x,t) - m(x))^2 \mid X_1, \dots, X_n \right] dx = \mathcal{O}_P \left( n^{-\frac{1}{2}} \right).$$

L'absence de segmentation ou une segmentation trop loin du point de rupture conduisent à des vitesses de convergence du MISE identiques.

#### Preuve du corollaire 4.1.

Ce corollaire découle immédiatement du théorème 4.1. En effet

$$\begin{split} &\int_{0}^{1} \mathbb{E} \left[ (\hat{m}(x,t) - m(x))^{2} \mid X_{1}, \dots, X_{n} \right] dx \\ &= \int_{0}^{\tau - h_{n}} \mathbb{E} \left[ (\hat{m}(x,t) - m(x))^{2} \mid X_{1}, \dots, X_{n} \right] dx \\ &+ \int_{\tau - h_{n}}^{\tau + h_{n}} \mathbb{E} \left[ (\hat{m}(x,t) - m(x))^{2} \mid X_{1}, \dots, X_{n} \right] dx \\ &+ \int_{\tau + h_{n}}^{1} \mathbb{E} \left[ (\hat{m}(x,t) - m(x))^{2} \mid X_{1}, \dots, X_{n} \right] dx , \\ &= \left( \gamma^{2}(\tau) \int_{-1}^{1} c_{1,r}^{2}(z) dz \right) h_{n} + \frac{1}{nh} \left( V_{K}^{2} \int_{0}^{1} \frac{\sigma^{2}(x)}{f(x)} dx \right) + o_{P}(h_{n} + \frac{1}{nh_{n}}). \end{split}$$

	Le théorème ci-après décrit le comportement de l'erreur quadratique moyenne au	voisinage
du	point de rupture lorsque la segmentation est faite au point $ au_n$ qui est proche du	point de
	nture – Par la quite – na core autre que l'estimateur de la logalisation du point é	lo nunturo

rupture au. Par la suite,  $au_n$  ne sera autre que l'estimateur de la localisation du point de rupture et  $|\tau_n - \tau|$  convergera vers 0 plus vite que  $h_n$ . Si nous supposons que  $\tau_n \leq \tau$ , nous constatons qu'il y a trois régions où l'erreur quadratique moyenne se comporte différemment:

i) Dans la première région  $R1(x \leq \tau_n \text{ ou } x \geq \tau + h_n)$ : le biais est de l'ordre  $\mathcal{O}_P(h_n^2)$ . C'est l'ordre de grandeur optimal que l'on peut espérer obtenir. Ceci s'explique par le fait que pour les points x qui appartiennent à la région R1, nous utilisons pour estimer m(x)uniquement les observations qui se trouvent avant le point de discontinuité.



Figure 4.6 : Région R1.

ii) Dans la deuxième région R2 (x ∈ [τ<sub>n</sub>,τ]), le biais a le même ordre de grandeur que l'amplitude de la rupture. Ceci est dû au fait que pour les points τ<sub>n</sub> < x < τ qui se trouvent avant la discontinuité, nous utilisons pour estimer m(x) des observations qui, en grande partie, se situent après le point de rupture. L'importance de ces points dans l'erreur quadratique moyenne intégrée (MISE) est proportionnelle à la distance entre τ<sub>n</sub> et τ c'est-à-dire |τ<sub>n</sub> − τ|. Plus cette distance est grande et plus l'effet de ces points est dominant dans le MISE.



Figure 4.7 : Région R2.

iii) Dans cette dernière région  $R3(x \in [\tau, \tau+h_n])$ , le biais se comporte comme  $\frac{|\tau_n-\tau|}{h_n}$ . Pour ces points qui se trouvent après la discontinuité, nous utilisons, pour estimer m(x), quelques observations qui se trouvent avant le point de rupture. Le nombre de ces observations est proportionnel à  $\frac{|\tau_n-\tau|}{h_n}$ .



Figure 4.8 : Région R3.

Théorème 4.2. Supposons que

- i) Les hypothèses D1-D4 sont satisfaites,
- ii)  $\tau_n = \tau + \frac{h_n}{a_n} r \text{ avec } |r| \leq 1 \text{ et } a_n \to \infty.$
- x<sub>n</sub> = τ + h<sub>n</sub>z, si r ≤ 0, z ≤ 0 et x<sub>n</sub> < τ<sub>n</sub> (r > 0, z > 0 et x<sub>n</sub> ≥ τ<sub>n</sub>), alors

   a E [m̂(x<sub>n</sub>, τ<sub>n</sub>) m(x<sub>n</sub>)|X<sub>1</sub>,...,X<sub>n</sub>] = ½m<sup>(2)</sup><sub>±</sub>(τ)h<sup>2</sup><sub>n</sub>B<sub>K</sub>(-|z|) + o<sub>P</sub>(h<sup>2</sup><sub>n</sub>),
   b) Var [m̂(x<sub>n</sub>, τ<sub>n</sub>)|X<sub>1</sub>,...,X<sub>n</sub>] = σ<sup>2</sup>(τ)/nh<sub>n</sub>f(τ)V<sup>2</sup><sub>K</sub>(-|z|) + o<sub>P</sub>(1/nh<sub>n</sub>),

   x<sub>n</sub> = τ + h<sub>n</sub>z, si r ≤ 0 et τ<sub>n</sub> ≤ x<sub>n</sub> < τ (r > 0 et τ ≤ x<sub>n</sub> < τ<sub>n</sub>), alors

   a) E [m̂(x<sub>n</sub>, τ<sub>n</sub>) m(x<sub>n</sub>)|X<sub>1</sub>,...,X<sub>n</sub>] = -sign(r)γ(τ) + o<sub>P</sub>(1),
   b) Var [m̂(x<sub>n</sub>, τ<sub>n</sub>)|X<sub>1</sub>,...,X<sub>n</sub>] = σ<sup>2</sup>(τ)/nh<sub>n</sub>f(τ)V<sup>2</sup><sub>K</sub>(0) + o<sub>P</sub>(1),
   b) Var [m̂(x<sub>n</sub>, τ<sub>n</sub>)|X<sub>1</sub>,...,X<sub>n</sub>] = σ<sup>2</sup>(τ)/nh<sub>n</sub>f(τ)V<sup>2</sup><sub>K</sub>(0) + o<sub>P</sub>(1/nh<sub>n</sub>),

a) 
$$\mathbb{E}\left[\hat{m}(x_n, \tau_n) - m(x_n)|X_1, \dots, X_n\right] = \frac{\tau_n - \tau}{h_n} \gamma(\tau) c_2(z) + o_P(\frac{|\tau_n - \tau|}{h_n}),$$
  
b)  $Var\left[\hat{m}(x_n, \tau_n)|X_1, \dots, X_n\right] = \frac{\sigma^2(\tau)}{nh_n f(\tau)} V_K^2(-|z|) + o_P(\frac{1}{nh_n}),$ 

оù

$$c_2(z) = \frac{(K_2(-|z|) + |z|K_1(-|z|))K(-|z|)}{K_2(-|z|)K_0(-|z|) - K_1^2(-|z|)}.$$

**Remarque 4.5.** L'utilisation d'un estimateur tel que celui de Nadaraya-Watson ou Gasser-Müller, aurait conduit à un biais de l'ordre de  $\mathcal{O}_P(h_n)$  pour les points de la première région (R1).

Nous présentons ci-dessous le graphique de la fonction  $c_2(z)$  pour le noyau d'Epanechnikov:



Figure 4.9 :  $c_2(z)$ .

# Preuve du théorème 4.2.

Supposons  $x = x_n = \tau + zh_n$  et  $t = \tau_n = \tau + \frac{h_n}{a_n}r$ . L'erreur quadratique moyenne conditionnelle (MSE) de  $\hat{m}(x_n, \tau_n)$  est toujours donnée par

$$\mathbb{E}\left[\left(\hat{m}_{+}(x_{n},\tau_{n})-m(x_{n})\right)^{2}\mid X_{1},\ldots,X_{n}\right] \\ = \left[\frac{\sum_{i=1}^{n}w_{i}(x_{n},\tau_{n})\left(m(X_{i})-m(x_{n})\right)}{\sum_{i=1}^{n}w_{i}(x_{n},\tau_{n})}\right]^{2} + \frac{\sum_{i=1}^{n}\left(w_{i}(x_{n},\tau_{n})\right)^{2}\sigma^{2}(X_{i})}{\left(\sum_{i=1}^{n}w_{i}(x_{n},\tau_{n})\right)^{2}} \\ = B_{n}^{2} + V_{n},$$

où

$$B_n = \frac{S_2(x_n, \tau_n) R_0(x_n, \tau_n) - S_1(x_n, \tau_n) R_1(x_n, \tau_n)}{S_2(x_n, \tau_n) S_0(x_n, \tau_n) - (S_1(x_n, \tau_n))^2},$$
(4.7)

 $\operatorname{et}$ 

$$V_{n} = \frac{\sum_{i=1}^{n} (w_{i}(x_{n},\tau_{n}))^{2} \sigma^{2}(X_{i})}{(\sum_{i=1}^{n} w_{i}(x_{n},\tau_{n}))^{2}},$$
  

$$= \frac{(S_{2}(x_{n},\tau_{n}))^{2} T_{0}(x_{n},\tau_{n}) - 2S_{2}(x_{n},\tau_{n})S_{1}(x_{n},\tau_{n})T_{1}(x_{n},\tau_{n})}{(S_{2}(x_{n},\tau_{n})S_{0}(x_{n},\tau_{n}) - (S_{1}(x_{n},\tau_{n}))^{2})^{2}} + \frac{(S_{1}(x_{n},\tau_{n}))^{2} T_{2}(x_{n},\tau_{n})}{(S_{2}(x_{n},\tau_{n})S_{0}(x_{n},\tau_{n}) - (S_{1}(x_{n},\tau_{n}))^{2})^{2}}.$$
(4.8)

1)  $z < 0, \, r < 0$  et  $x_n < \tau_n \; (z \ge 0, \, r \ge 0$  et  $x_n \ge \tau_n$  )

**Lemme 4.3.** Soient  $\tau_n = \tau + r \frac{h_n}{a_n}$  et  $x_n = \tau + zh_n$  telles que z < 0, r < 0 et  $x_n < \tau_n$  ( $z \ge 0$ ,  $r \ge 0$ ,  $x_n \ge \tau_n$ ) et  $a_n \to \infty$ . Sous les hypothèses D1-D4, nous avons:

a) 
$$S_l(x_n, \tau_n) = nh_n^{l+1} f_X(\tau) \left( (-sign(z))^l K_l(-|z|) + o_P(1) \right),$$
  
b)  $T_l(x_n, \tau_n) = nh_n^{l+1} f_X(\tau) \sigma^2(\tau) \left( (-sign(z))^l L_l(-|z|) + o_P(1) \right),$   
c)  $R_l(x_n, \tau_n) = \frac{1}{2} m_{\pm}^{(2)}(\tau) nh_n^{l+3} f_X(\tau) \left( (-sign(z))^l K_{l+2}(-|z|) + o_P(1) \right).$ 

### Preuve du lemme 4.3.

Afin de démontrer ce lemme nous considérons le cas où z < 0, r < 0 et  $x_n < \tau_n$ . La démonstration du cas symétrique  $(z \ge 0, r \ge 0$  et  $x_n \ge \tau_n)$  s'en déduit facilement.

Nous avons par l'inégalité de Bienaymé-Tchebycheff

$$\sum_{X_i < \tau_n} Z_{i,j} = \mathbb{E} \sum_{X_i < \tau_n} Z_{i,j} + \mathcal{O}_P\left(\sqrt{Var\left(\sum_{X_i < \tau_n} Z_{i,j}\right)}\right),$$
$$= n\mathbb{E} \left[Z_{1,j} 1\!\!1_{\{X_1 < \tau_n\}}\right] + \mathcal{O}_P\left(\sqrt{n\mathbb{E} \left[Z_{1,j}^2 1\!\!1_{\{X_1 < \tau_n\}}\right]}\right),$$

$$\mathbb{E}\left[Z_{1,1} 1\!\!1_{\{X_1 < \tau_n\}}\right] = h_n^{l+1} \int_{\frac{x_n - \tau_n}{h_n}}^1 u^l K_+(u) f_X(x_n - h_n u) du ,$$
  
=  $h_n^{l+1} f_X(\tau) \left(K_l(z) + o(1)\right).$ 

De même

$$\mathbb{E}\left[Z_{1,1}^{2} \prod_{\{X_{1} < \tau_{n}\}}\right] = h_{n}^{2l+1} \int_{\frac{x_{n} - \tau_{n}}{h_{n}}}^{1} u^{2l} K^{2}(u) f_{X}(x_{n} - h_{n}u) du ,$$
  
$$= \mathcal{O}\left(h_{n}^{2l+1}\right),$$

d'où

$$S_{l}(x_{n},\tau_{n}) = n \mathbb{E} \left[ Z_{1,1} \mathbb{1}_{\{X_{1} < \tau_{n}\}} \right] + \mathcal{O}_{P} \left( \sqrt{n \mathbb{E} \left[ Z_{1,1}^{2} \mathbb{1}_{\{X_{1} < \tau_{n}\}} \right]} \right),$$
  
$$= n h_{n}^{l+1} f_{X}(\tau) \left( K_{l}(z) + o_{P}(1) \right).$$

De la même façon, nous obtenons pour  $\mathbb{E}[Z_{1,2}]$  et  $\mathbb{E}[Z_{1,2}^2]$ ,

$$\mathbb{E}\left[Z_{1,2}\mathbf{1}_{\{X_{1}<\tau_{n}\}}\right] = h_{n}^{l+1} \int_{\frac{x_{N}-\tau_{n}}{h_{n}}}^{1} u^{l} K^{2}(u) f_{X}(x_{n}-h_{n}u) \sigma^{2}(x_{n}-h_{n}u) du,$$
  
$$= h_{n}^{l+1} f_{X}(\tau) \sigma^{2}(\tau) \left(L_{l}(z)+o(1)\right),$$

$$\mathbb{E}\left[Z_{1,2}^{2} 1\!\!1_{\{X_{1} < \tau_{n}\}}\right] = h_{n}^{2l+1} \int_{\frac{x_{n}-\tau_{n}}{h_{n}}}^{1} u^{2l} K^{4}(u) f_{X}(x_{n}-h_{n}u) \sigma^{4}(x_{n}-h_{n}u) du ,$$
  
$$= \mathcal{O}\left(h_{n}^{2l+1}\right).$$

Nous en déduisons

$$T_{l}(x_{n},\tau_{n}) = n\mathbb{E}\left[Z_{1,2}\mathbb{1}_{\{X_{1}<\tau_{n}\}}\right] + \mathcal{O}_{P}\left(\sqrt{n\mathbb{E}\left[Z_{1,2}^{2}\mathbb{1}_{\{X_{1}<\tau_{n}\}}\right]}\right),$$
  
$$= nh_{n}^{l+1}f_{X}(\tau)\sigma^{2}(\tau)\left(L_{l}(z) + o_{P}(1)\right).$$

$$\begin{split} & \mathbb{E}\left[Z_{1,3}1\!\!\!1_{\{X_1<\tau_n\}}\right] \\ &= h_n^{l+1} \int_{\frac{x_n-\tau_n}{h_n}}^1 u^l K(u) \left(m(x_n-h_n u) - m(x_n) + h_n u m'(x_n)\right) f_X(x_n-h_n u) du \,, \\ &= \frac{1}{2} m_{-}^{(2)}(\tau) h_n^{l+3} f_X(\tau) \left(K_{l+2}(z) + o(1)\right) \,. \end{split}$$

Un même calcul conduit à

$$\begin{split} & \mathbb{E}\left[Z_{1,3}^{2}\mathbb{I}_{\{X_{1}<\tau_{n}\}}\right] \\ &= h_{n}^{2l+1}\int_{\frac{x_{n}-\tau_{n}}{h_{n}}}^{1}u^{l}K^{2}(u)\left(m(x_{n}-h_{n}u)-m(x_{n})+h_{n}um'(x_{n})\right)^{2} \\ &\times f_{X}(x_{n}-h_{n}u)du, \\ &= \mathcal{O}\left(h_{n}^{2l+5}\right). \end{split}$$

Ainsi

$$R_{l}(x_{n},\tau_{n}) = n\mathbb{E}\left[Z_{1,3}\mathbb{1}_{\{X_{1}<\tau_{n}\}}\right] + \mathcal{O}_{P}\left(\sqrt{n\mathbb{E}\left[Z_{1,3}^{2}\mathbb{1}_{\{X_{1}<\tau_{n}\}}\right]}\right),$$
  
$$= \frac{1}{2}m_{-}^{(2)}(\tau)nh_{n}^{l+3}f_{X}(\tau)\left(K_{l+2}(z)+o_{P}(1)\right).$$

Ceci termine la preuve du lemme 4.3.

En combinant le lemme 4.3 et les égalités (4.7) et (4.8), nous prouvons la première partie du théorème 4.2.

2) 
$$r < 0$$
 et  $au_n \leq x_n < au$   $(r \geq 0$  et  $au \leq x_n < au_n$  )

**Lemme 4.4.** Soit  $\tau_n = \tau + r \frac{h_n}{a_n}$  tel que r < 0 et  $x_n \ge \tau_n$   $(r \ge 0, x_n < \tau_n)$  et  $a_n \to \infty$ . Sous les hypothèses D1-D4 nous avons :

a) 
$$S_l(x_n, \tau_n) = nh_n^{l+1} f_X(\tau) \left( (sign(r))^l K_l(0) + o_P(1) \right),$$
  
b)  $T_l(x_n, \tau_n) = nh_n^{l+1} f_X(\tau) \sigma^2(\tau) \left( (sign(r))^l L_l(0) + o_P(1) \right),$   
c)  $R_l(x_n, \tau_n) = -sign(r)\gamma(\tau)nh_n^{l+1} f_X(\tau) \left( (sign(r))^l K_l(0) + o_P(1) \right).$ 

# Preuve du lemme 4.4.

Nous avons

$$\mathbb{E}\left[Z_{1,1} 1\!\!1_{\{X_1 \ge \tau_n\}}\right] = h_n^{l+1} \int_{-1}^{\frac{x_n - \tau_n}{h_n}} u^l K_+(u) f_X(x_n - h_n u) du,$$
  
$$= h_n^{l+1} f_X(\tau) \left(K_l(-1) - K_l(0) + o(1)\right),$$

de même,

$$\mathbb{E}\left[Z_{1,1}^{2} 1\!\!1_{\{X_{1} \ge \tau_{n}\}}\right] = h_{n}^{2l+1} \int_{-1}^{\frac{x_{n}-\tau_{n}}{h_{n}}} u^{2l} K^{2}(u) f_{X}(x_{n}-h_{n}u) du$$
$$= \mathcal{O}\left(h_{n}^{2l+1}\right),$$

d'où

$$S_{l}(x_{n},\tau_{n}) = n \mathbb{E} \left[ Z_{1,1} \mathbb{1}_{\{X_{1} \geq \tau_{n}\}} \right] + \mathcal{O}_{P} \left( \sqrt{n \mathbb{E} \left[ Z_{1,1}^{2} \mathbb{1}_{\{X_{1} \geq \tau_{n}\}} \right]} \right),$$
  
$$= n h_{n}^{l+1} f_{X}(\tau) \left( K_{l}(-1) - K_{l}(0) + o_{P}(1) \right),$$
  
$$= n h_{n}^{l+1} f_{X}(\tau) \left( (sign(r))^{l} K_{l}(0) + o_{P}(1) \right).$$

Les approximations de  $\mathbb{E}[Z_{1,2}]$  et  $\mathbb{E}[Z_{1,2}^2]$  sont obtenues de manière analogue.

$$\begin{split} \mathbb{E}\left[Z_{1,2} \, \mathrm{I\!I}_{\{X_1 \ge \tau_n\}}\right] &= h_n^{l+1} \int_{-1}^{\frac{x_n - \tau_n}{h_n}} u^l K^2(u) f_X(x_n - h_n u) \sigma^2(x_n - h_n u) du \,, \\ &= h_n^{l+1} f_X(\tau) \sigma^2(\tau) \left(L_l(-1) - L_l(0) + o(1)\right), \end{split}$$

$$\mathbb{E}\left[Z_{1,2}^{2} 1\!\!1_{\{X_{1} \ge \tau_{n}\}}\right] = h_{n}^{2l+1} \int_{-1}^{\frac{x_{n} - \tau_{n}}{h_{n}}} u^{2l} K^{4}(u) f_{X}(x_{n} - h_{n}u) \sigma^{4}(x_{n} - h_{n}u) du,$$
  
$$= \mathcal{O}\left(h_{n}^{2l+1}\right).$$

Nous en déduisons que

$$\begin{split} T_{l}(x_{n},\tau_{n}) &= n \mathbb{E}\left[Z_{1,2} \mathrm{I\!I}_{\{X_{1} \geq \tau_{n}\}}\right] + \mathcal{O}_{P}\left(\sqrt{n \mathbb{E}\left[Z_{1,2}^{2} \mathrm{I\!I}_{\{X_{1} \geq \tau_{n}\}}\right]}\right), \\ &= nh_{n}^{l+1} f_{X}(\tau) \sigma^{2}(\tau) \left(L_{l}(-1) - L_{l}(0) + o_{P}(1)\right), \\ &= nh_{n}^{l+1} f_{X}(\tau) \sigma^{2}(\tau) \left((sign(r))^{l} L_{l}(0) + o_{P}(1)\right). \end{split}$$

$$\begin{split} & \mathbb{E}\left[Z_{1,3} \prod_{\{X_1 \ge \tau_n\}}\right] \\ &= h_n^{l+1} \int_{-1}^{\frac{x_n - \tau_n}{h_n}} u^l K(u) \left(m(x_n - h_n u) - m(x_n) + h_n u m'(x_n)\right) \\ &\times f_X(x_n - h_n u) du \,, \\ &= \gamma(\tau) h_n^{l+1} f_X(\tau) \left(K_l(-1) - K_l(0) + o(1)\right). \end{split}$$

Un calcul similaire donne

$$\begin{split} & \mathbb{E}\left[Z_{1,3}^{2} 1\!\!1_{\{X_{1} \ge \tau_{n}\}}\right] \\ &= h_{n}^{2l+1} \int_{-1}^{\frac{x_{n} - \tau_{n}}{h_{n}}} u^{l} K^{2}(u) \left(m(x_{n} - h_{n}u) - m(x_{n}) + h_{n}um'(x_{n})\right)^{2} \\ &\times f_{X}(x_{n} - h_{n}u) du , \\ &= \mathcal{O}\left(h_{n}^{2l+1}\right) , \end{split}$$

d'où

$$\begin{aligned} R_{l}(x_{n},\tau_{n}) &= n \mathbb{E}\left[Z_{1,3} 1\!\!1_{\{X_{1} \geq \tau_{n}\}}\right] + \mathcal{O}_{P}\left(\sqrt{n \mathbb{E}\left[Z_{1,3}^{2} 1\!\!1_{\{X_{1} \geq \tau_{n}\}}\right]}\right), \\ &= \gamma(\tau) n h_{n}^{l+1} f_{X}(\tau) \left(K_{l}(-1) - K_{l}(0) + o_{P}(1)\right), \\ &= \gamma(\tau) n h_{n}^{l+1} f_{X}(\tau) \left((sign(r))^{l} K_{l}(0) + o_{P}(1)\right). \end{aligned}$$

1	]	

3) r < 0 et  $z \ge 0$   $(r \ge 0$  et z < 0).

**Lemme 4.5.** Soit  $\tau_n = \tau + r \frac{h_n}{a_n}$  et  $x_n = \tau + zh_n$ . telles que  $z \ge 0$ , r < 0 (z < 0,  $r \ge 0$ ,  $x_n \ge \tau_n$ ) et  $a_n \to \infty$ . Sous les hypothèses D1-D4 nous avons:

a) 
$$S_l(x_n, \tau_n) = nh_n^{l+1} f_X(\tau) \left( (-sign(z))^l K_l(-|z|) + o_P(1) \right),$$
  
b)  $T_l(x_n, \tau_n) = nh_n^{l+1} f_X(\tau) \sigma^2(\tau) \left( (-sign(z))^l L_l(-|z|) + o_P(1) \right).$   
c)  $Si a_n = o(nh_n),$   
 $R_l(x_n, \tau_n) = \frac{\tau_n - \tau}{h_n} \gamma(\tau) nh_n^{l+1} f_X(\tau) \left( (-sign(z))^l |z|^l K(-|z|) + o_P(1) \right).$ 

### Preuve du lemme 4.5.

Nous avons

$$\mathbb{E}\left[Z_{1,1}\mathbb{1}_{\{X_1 \ge \tau_n\}}\right] = h_n^{l+1} \int_{-1}^{\frac{x_n - \tau_n}{h_n}} u^l K_+(u) f_X(x_n - h_n u) du,$$
  
$$= h_n^{l+1} f_X(\tau) \left( (-sign(z))^l K_l(-z) + o(1) \right).$$

De même

$$\mathbb{E}\left[Z_{1,1}^{2} \prod_{\{X_{1} \geq \tau_{n}\}}\right] = h_{n}^{2l+1} \int_{-1}^{\frac{x_{n}-\tau_{n}}{h_{n}}} u^{2l} K^{2}(u) f_{X}(x_{n}-h_{n}u) du,$$
  
$$= \mathcal{O}\left(h_{n}^{2l+1}\right).$$

Ainsi

$$S_{l}(x_{n},\tau_{n}) = n \mathbb{E} \left[ Z_{1,1} \mathbb{1}_{\{X_{1} \geq \tau_{n}\}} \right] + \mathcal{O}_{P} \left( \sqrt{n \mathbb{E} \left[ Z_{1,1}^{2} \mathbb{1}_{\{X_{1} \geq \tau_{n}\}} \right]} \right),$$
  
$$= n h_{n}^{l+1} f_{X}(\tau) \left( (-sign(z))^{l} K_{l}(-z) + o_{P}(1) \right),$$

$$\mathbb{E}\left[Z_{1,2} \prod_{\{X_1 \ge \tau_n\}}\right] = h_n^{l+1} \int_{-1}^{\frac{x_n - \tau_n}{h_n}} u^l K^2(u) f_X(x_n - h_n u) \sigma^2(x_n - h_n u) du ,$$
  
$$= h_n^{l+1} f_X(\tau) \sigma^2(\tau) \left( (-sign(z))^l L_l(-z) + o(1) \right) ,$$

$$\mathbb{E}\left[Z_{1,2}^{2} 1\!\!1_{\{X_{1} \ge \tau_{n}\}}\right] = h_{n}^{2l+1} \int_{-1}^{\frac{x_{n}-\tau_{n}}{h_{n}}} u^{2l} K^{4}(u) f_{X}(x_{n}-h_{n}u) \sigma^{4}(x_{n}-h_{n}u) du$$
  
$$= \mathcal{O}\left(h_{n}^{2l+1}\right).$$

Nous en déduisons

$$T_{l}(x_{n},\tau_{n}) = n\mathbb{E}\left[Z_{1,2}1\!\!1_{\{X_{1}\geq\tau_{n}\}}\right] + \mathcal{O}_{P}\left(\sqrt{n\mathbb{E}\left[Z_{1,2}^{2}1\!\!1_{\{X_{1}\geq\tau_{n}\}}\right]}\right),$$
  
$$= nh_{n}^{l+1}f_{X}(\tau)\sigma^{2}(\tau)\left((-sign(z))^{l}L_{l}(-z) + o_{P}(1)\right).$$

D'autre part

$$\mathbb{E}\left[Z_{1,3}\Pi_{\{X_1 \ge \tau_n\}}\right] = h_n^{l+1} \int_{-1}^{\frac{x_n - \tau_n}{h_n}} u^l K(u) f_X(x_n - h_n u) \\ \times \left(m(x_n - h_n u) - m(x_n) + h_n um'(x_n)\right) du , \\ = h_n^{l+1} \int_{z}^{z - \frac{r}{a_n}} u^l K(u) f_X(x_n - h_n u) \\ \times \left(m(x_n - h_n u) - m(x_n) + h_n um'(x_n)\right) du + o(h_n^{l+1}), \\ = \frac{\tau_n - \tau}{h_n} \gamma(\tau) h_n^{l+1} f_X(\tau) \left(z^l K(z) + o(1)\right),$$

$$\mathbb{E}\left[Z_{1,3}^{2} \mathrm{I\!I}_{\{X_{1} \ge \tau_{n}\}}\right] = h_{n}^{2l+1} \int_{-1}^{\frac{x_{n}-\tau_{n}}{h_{n}}} u^{l} K^{2}(u) f_{X}(x_{n}-h_{n}u) \\ \times \left(m(x_{n}-h_{n}u)-m(x_{n})+h_{n}um'(x_{n})\right)^{2} du \\ = \mathcal{O}\left(\frac{h_{n}^{2l+1}}{a_{n}}\right),$$

lorsque  $a_n = o(nh_n)$  nous avons alors,

$$R_{l}(x_{n},\tau_{n}) = n\mathbb{E}\left[Z_{1,3}\mathbb{1}_{\{X_{1}\geq\tau_{n}\}}\right] + \mathcal{O}_{P}\left(\sqrt{n\mathbb{E}\left[Z_{1,3}^{2}\mathbb{1}_{\{X_{1}\geq\tau_{n}\}}\right]}\right),$$
  
$$= \frac{\tau_{n}-\tau}{h_{n}}\gamma(\tau)nh_{n}^{l+1}f_{X}(\tau)\left(z^{l}K(z)+o_{P}(1)\right).$$

Remplaçant dans les expressions de  $B_n$  et  $V_n$  ((4.7) et (4.8)), les termes  $S_l(\cdot, \cdot)$ ,  $T_l(\cdot, \cdot)$  et  $R_l(\cdot, \cdot)$ par leurs développements donnés par les lemmes 4.3, 4.4 et 4.5, nous obtenons les résultats énoncés dans le théorème 4.2.

Corollaire 4.2. Sous les hypothèses du théorème 4.2, nous avons

$$\int_{0}^{1} \mathbb{E} \left[ (\hat{m}(x,\tau_{n}) - m(x))^{2} \mid X_{1}, \dots, X_{n} \right] dx$$
  
=  $\frac{B_{K}^{2}h_{n}^{4}}{4} \left( \int_{0}^{1} \left[ m^{(2)}(x) \right]^{2} dx \right) + \frac{V_{K}^{2}}{nh} \left( \int_{0}^{1} \frac{\sigma^{2}(x)}{f(x)} dx \right)$   
+  $o_{P}(h_{n}^{4} + \frac{1}{nh_{n}}) + \mathcal{O}_{P}(|\tau_{n} - \tau|)$ 

**Remarque 4.6.** Notons que si nous avions utilisé, à la place de la régression linéaire locale, un estimateur à noyau de type Nadaraya-Watson ou Gasser-Müller, l'erreur quadratique moyenne intégrée <sup>1</sup> aurait été de l'ordre de  $\mathcal{O}_P\left(h_n^3 + \frac{1}{nh_n}\right)$  alors que le corollaire ci-dessus conduit à une vitesse de convergence meilleure dès que  $\tau_n$  converge suffisamment vite. Ce qui sera réalisé dans la pratique.

#### Preuve du corollaire 4.2.

Ce corollaire est une conséquence immédiate du théorème 4.2. En effet

$$\int_{0}^{1} \mathbb{E} \left[ (\hat{m}(x,\tau_{n}) - m(x))^{2} \mid X_{1}, \dots, X_{n} \right] dx$$
  
=  $\int_{|x-\tau| \ge h_{n}} \mathbb{E} \left[ (\hat{m}(x,\tau_{n}) - m(x))^{2} \mid X_{1}, \dots, X_{n} \right] dx$   
+  $\int_{\tau-h_{n}}^{\tau+h_{n}} \mathbb{E} \left[ (\hat{m}(x,\tau_{n}) - m(x))^{2} \mid X_{1}, \dots, X_{n} \right] dx$ .

 ${}^{1}\int_{0}^{1}\mathbb{E}\left[\left(\hat{m}(x,\tau_{n})-m(x)\right)^{2}\mid X_{1},\ldots,X_{n}\right]dx$ 

et

Tout d'abord, nous avons

$$\int_{|x-\tau| \ge h_n} \mathbb{E}\left[ \left( \hat{m}(x,\tau_n) - m(x) \right)^2 \mid X_1, \dots, X_n \right] dx$$
  
=  $\frac{B_K^2 h_n^4}{4} \left( \int_0^1 \left[ m^{(2)}(x) \right]^2 dx \right) + \frac{V_K^2}{nh} \left( \int_0^1 \frac{\sigma^2(x)}{f(x)} dx \right) + o_P(h_n^4 + \frac{1}{nh_n}).$ 

Si nous supposons, par exemple, que  $\tau_n \leq \tau$ , alors

$$\begin{split} &\int_{\tau-h_n}^{\tau+h_n} \mathbb{E} \left[ (\hat{m}(x,\tau_n) - m(x))^2 \mid X_1, \dots, X_n \right] dx \\ &= \int_{\tau-h_n}^{\tau_n} \mathbb{E} \left[ (\hat{m}(x,\tau_n) - m(x))^2 \mid X_1, \dots, X_n \right] dx \\ &+ \int_{\tau_n}^{\tau} \mathbb{E} \left[ (\hat{m}(x,\tau_n) - m(x))^2 \mid X_1, \dots, X_n \right] dx \\ &+ \int_{\tau}^{\tau+h_n} \mathbb{E} \left[ (\hat{m}(x,\tau_n) - m(x))^2 \mid X_1, \dots, X_n \right] dx , \\ &= o_P (h_n^4 + \frac{1}{nh_n}) + \mathcal{O}_P (|\tau_n - \tau|) + \mathcal{O}_P (\frac{(\tau_n - \tau)^2}{h_n}), \\ &= o_P (h_n^4 + \frac{1}{nh_n}) + \mathcal{O}_P (|\tau_n - \tau|). \end{split}$$

Nous en déduisons donc

$$\int_{0}^{1} \mathbb{E} \left[ (\hat{m}(x,\tau_{n}) - m(x))^{2} \mid X_{1}, \dots, X_{n} \right] dx$$
  
=  $\frac{B_{K}^{2}h_{n}^{4}}{4} \left( \int_{0}^{1} \left[ m^{(2)}(x) \right]^{2} dx \right) + \frac{V_{K}^{2}}{nh} \left( \int_{0}^{1} \frac{\sigma^{2}(x)}{f(x)} dx \right)$   
+  $o_{P}(h_{n}^{4} + \frac{1}{nh_{n}}) + \mathcal{O}_{P}(|\tau_{n} - \tau|).$ 

Ceci termine la démonstration du corollaire 4.2.

1		

# 4.3 Un seul point de rupture. Segmentation basée sur l'estimateur de la localisation de la rupture

Nous supposons dans ce paragraphe que la localisation de la rupture  $\tau$  n'est pas connue et montrons que si son estimateur  $\hat{\tau}$  converge de manière suffisamment rapide, la fonction  $m(\cdot)$  est estimée avec une erreur quadratique intégrée conditionnelle optimale.

Il est facile de voir que

$$\int Var\left[\hat{m}(x,\hat{\tau})|X_1,\ldots,X_n\right]dx = \mathcal{O}_P\left(\frac{1}{nh_n}\right).$$

Cherchons maintenant à contrôler le biais de  $\hat{m}(x, \hat{\tau})$ . Étant donné une suite  $(g_n)_n$  positive convergent vers zéro, nous pouvons écrire:

$$\begin{split} &|\mathbb{E}\left[\hat{m}(x,\hat{\tau}) - m(x)|X_{1},\dots,X_{n}\right]| \\ &\leq \left|\mathbb{E}\left[\left(\hat{m}(x,\hat{\tau}) - m(x)\right)\mathbf{1}\!\!1_{|\hat{\tau} - \tau| > g_{n}}|X_{1},\dots,X_{n}\right]\right| \\ &+ \left|\mathbb{E}\left[\left(\hat{m}(x,\hat{\tau}) - m(x)\right)\mathbf{1}\!\!1_{|\hat{\tau} - \tau| \le g_{n}}|X_{1},\dots,X_{n}\right]\right|, \\ &\leq \left(\mathbb{E}\left[\left(\hat{m}(x,\hat{\tau}) - m(x)\right)^{2}|X_{1},\dots,X_{n}\right]\right)^{1/2} \\ &\times \mathbb{P}^{1/2}\left[|\hat{\tau} - \tau| > g_{n}|X_{1},\dots,X_{n}\right] \\ &+ \left(\mathbb{E}\left[\left(\hat{m}(x,\hat{\tau}) - m(x)\right)^{2}\mathbf{1}\!\!1_{|\hat{\tau} - \tau| \le g_{n}}|X_{1},\dots,X_{n}\right]\right)^{1/2}. \end{split}$$

Il en résulte, en particulier, que

$$\int \left(\mathbb{E}\left[\hat{m}(x,\hat{\tau}) - m(x)|X_{1},\ldots,X_{n}\right]\right)^{2} dx$$

$$\leq \int \mathbb{E}\left[\left(\hat{m}(x,\hat{\tau}) - m(x)\right)^{2}|X_{1},\ldots,X_{n}\right]$$

$$\times \mathbb{P}\left[|\hat{\tau} - \tau| > g_{n}|X_{1},\ldots,X_{n}\right] dx$$

$$+ \int \mathbb{E}\left[\left(\hat{m}(x,\hat{\tau}) - m(x)\right)^{2} \mathbb{I}_{|\hat{\tau} - \tau| \le g_{n}|X_{1},\ldots,X_{n}\right] dx,$$

$$= \mathcal{O}_{P}\left(h_{n} + \frac{1}{nh_{n}}\right) \mathbb{P}\left[|\hat{\tau} - \tau| > g_{n}|X_{1},\ldots,X_{n}\right]$$

$$+ \mathcal{O}_{P}\left(h_{n}^{4} + \frac{1}{nh_{n}}\right) + \mathcal{O}_{P}\left(g_{n}\right),$$

où nous avons utilisé pour obtenir la dernière majoration, le résultat énoncé dans le corollaire 4.2.

Ainsi, nous en déduisons le théorème suivant

**Théorème 4.3.** Supposons que les hypothèses D1-D4 sont satisfaites et soit  $\hat{\tau}$  un estimateur de  $\tau$  tel qu'il existe une suite  $(g_n)$  vérifiant:

i)  $g_n = \mathcal{O}\left(h_n^4 + \frac{1}{nh_n}\right),$ ii)  $\operatorname{IP}\left[|\hat{\tau} - \tau| > g_n \ |X_1, \dots, X_n\right] = o_P(h_n^3).$ 

A lors:

$$\int \operatorname{I\!E} \left[ \left( \hat{m}(x, \hat{\tau}) - m(x) \right)^2 | X_1, \dots, X_n \right] dx = \mathcal{O}_P \left( h_n^4 + \frac{1}{nh_n} \right).$$

Naturellement les estimateurs de  $\tau$  étudiés dans les chapitres précédents satisfont sous des conditions convenables les conditions requises ci-dessus (cf paragraphe 4.5).

# 4.4 Plusieurs points de rupture

### 4.4.1 Notations et hypothèses

Dans ce paragraphe, nous supposons que la fonction de régression admet p points de rupture :  $0 < \tau_1 < \ldots < \tau_p < 1$ . Par convention, nous posons  $\tau_0 = 0$  et  $\tau_{p+1} = 1$ . Nous conservons les hypothèses D1, D3, D4, et nous remplaçons D2 par :

D2'. La fonction de régression  $m(\cdot) = \mathbb{E}(Y \mid X = \cdot)$  est deux fois continûment dérivable en tout point différent de  $\tau_1, \ldots, \tau_p$  et est *càdlàg* aux points  $\tau_1, \ldots, \tau_p$  ainsi que ses deux premières dérivées.

Pour estimer m(x), l'approche utilisée plus haut dans le cas d'un seul point de rupture, est généralisée de la façon suivante :

• Pour tout  $0 \le i \le p$ , si  $x \in [\tau_i, \tau_{i+1}]$ , nous ajustons localement une droite  $\alpha_x + \beta_x(x-u)$ par RLL pour obtenir  $\hat{m}(x) = \hat{\beta}_x$  en utilisant seulement les observations  $\tau_i \le X_j < \tau_{i+1}$ .

Soit  $\tau = (\tau_1, \ldots, \tau_p)$ , l'estimateur de m(x) est défini par :

$$\hat{m}(x,\tau) = \frac{\sum_{j=1}^{n} w_j(x,\tau) Y_j}{\sum_{j=1}^{n} w_j(x,\tau)},$$
(4.9)

avec lorsque  $x \in [\tau_i, \tau_{i+1}[,$ 

$$w_j(x,\tau) = \begin{cases} 0 & \text{si } X_j \notin [\tau_i, \tau_{i+1}[, \\ K(\frac{x-X_j}{h_n})(S_2(x,\tau) - (x-X_j)S_1(x,\tau)) & \text{sinon}, \end{cases}$$

où pour l = 0, 1, 2, ... et si  $x \in [\tau_i, \tau_{i+1}]$ 

$$S_l(x,\tau) = \sum_{\tau_i \le X_j < \tau_{i+1}} (x - X_j)^l K(\frac{x - X_j}{h_n}).$$
(4.10)

Lorsque p est inconnu, ainsi que  $\tau_1, \ldots, \tau_p$ , on leur substitue dans les expressions ci-dessous les estimateurs  $\hat{p}, \hat{\tau}_1, \ldots, \hat{\tau}_{\hat{p}}$ .

# 4.4.2 Résultats asymptotiques

Soit  $\hat{p}$  un estimateur du nombre de rupture de la fonction de régression. On note  $\hat{\tau}_1 < \ldots < \hat{\tau}_{\hat{p}}$  les estimateurs des points de ruptures de la fonction  $m(\cdot)$ . Pour simplifier les expressions, nous posons dans la suite :

i)  $\tau = (\tau_1, ..., \tau_p),$ 

ii) 
$$\hat{\tau} = (\hat{\tau}_1, \dots, \hat{\tau}_{\hat{p}}),$$
  
iii)  $\mathcal{E}_n = \left(\bigcap_{j=1}^{\hat{p}} \{ |\hat{\tau}_j - \tau_j| \le g_n| \} \right) \bigcap \{ \hat{p} = p \}.$ 

Nous avons d'une part

$$\int Var\left[\hat{m}(x,\hat{\tau})|X_1,\ldots,X_n\right]dx = \mathcal{O}_P\left(\frac{1}{nh_n}\right),$$

et d'autre part

$$\begin{split} |\mathbb{E}\left[\hat{m}(x,\hat{\tau}) - m(x)|X_{1},\ldots,X_{n}\right]| \\ &\leq |\mathbb{E}\left[\left(\hat{m}(x,\hat{\tau}) - m(x)\right)\mathbb{1}\left(\mathcal{E}_{n}^{c}\right)|X_{1},\ldots,X_{n}\right]| \\ &+ |\mathbb{E}\left[\left(\hat{m}(x,\hat{\tau}) - m(x)\right)\mathbb{1}\left(\mathcal{E}_{n}\right)|X_{1},\ldots,X_{n}\right]|, \\ &\leq \left(\mathbb{E}\left[\left(\hat{m}(x,\hat{\tau}) - m(x)\right)^{2}|X_{1},\ldots,X_{n}\right]\right)^{1/2} \\ &\times \mathbb{P}^{1/2}\left[\mathcal{E}_{n}^{c}|X_{1},\ldots,X_{n}\right] \\ &+ \left(\mathbb{E}\left[\left(\hat{m}(x,\hat{\tau}) - m(x)\right)^{2}\mathbb{1}\left(\mathcal{E}_{n}\right)|X_{1},\ldots,X_{n}\right]\right)^{1/2}. \end{split}$$

Ainsi

$$\int \left(\mathbb{E}\left[\hat{m}(x,\hat{\tau}) - m(x)|X_{1}, \dots, X_{n}\right]\right)^{2} dx$$

$$\leq \int \mathbb{E}\left[\left(\hat{m}(x,\hat{\tau}) - m(x)\right)^{2}|X_{1}, \dots, X_{n}\right]$$

$$\times \mathbb{P}\left[\mathcal{E}_{n}^{c}|X_{1}, \dots, X_{n}\right] dx$$

$$+ \int \mathbb{E}\left[\left(\hat{m}(x,\hat{\tau}) - m(x)\right)^{2} \mathbb{I}(\mathcal{E}_{n})|X_{1}, \dots, X_{n}\right] dx,$$

$$= \mathcal{O}_{P}\left(h_{n} + \frac{1}{nh_{n}}\right) \mathbb{P}\left[\mathcal{E}_{n}^{c}|X_{1}, \dots, X_{n}\right] + \mathcal{O}_{P}\left(h_{n}^{4} + \frac{1}{nh_{n}}\right) + \mathcal{O}_{P}\left(g_{n}\right).$$

D'où le résultat énoncé dans le théorème suivant

**Théorème 4.4.** Soit  $\hat{\tau}$  un estimateur de  $\tau$  tel que il existe une suite  $(g_n)$  vérifiant:

i)  $g_n = O\left(h_n^4 + \frac{1}{nh_n}\right),$ ii)  $\mathbb{P}\left[\mathcal{E}_n^c \mid X_1, \dots, X_n\right] = o_P(h_n^3).$ 

Supposons que les hypothèses D1, D2', D3, D4 sont satisfaites,

$$\int \mathbb{E}\left[ (\hat{m}(x,\hat{\tau}) - m(x))^2 | X_1, \dots, X_n \right] dx = \mathcal{O}_P\left( h_n^4 + \frac{1}{nh_n} \right).$$

Il n'est pas impossible que les coefficients, qui interviennent dans l'écriture asymptotique de l'erreur quadratique moyenne intégrée, ne soient pas modifiés lorsqu'on remplace le nombre et les localisations des points de rupture de la fonction de régression par leurs estimateurs. En d'autres termes, il est probable qu'on puisse montrer:

$$\int_0^1 \mathbb{E}\left[ \left( \hat{m}(x,\hat{\tau}) - m(x) \right)^2 \mid X_1, \dots, X_n \right] dx$$
  
=  $\frac{B_K^2 h_n^4}{4} \left( \int_0^1 \left[ m^{(2)}(x) \right]^2 dx \right) + \frac{V_K^2}{nh} \left( \int_0^1 \frac{\sigma^2(x)}{f(x)} dx \right) + o_P(h_n^4 + \frac{1}{nh_n}).$ 

En effet, nous montrons que la partie biais de l'expression asymptotique du MISE est inchangée lorsque p et  $\tau$  sont estimés.

$$\begin{split} &|\mathbb{E} \left[ \hat{m}(x,\hat{\tau}) - \hat{m}(x,\tau) | X_{1}, \dots, X_{n} \right] \\ &\leq \left| \mathbb{E} \left[ (\hat{m}(x,\hat{\tau}) - \hat{m}(x,\tau)) \, \mathrm{I\!I}(\mathcal{E}_{n}^{c}) | X_{1}, \dots, X_{n} \right] \right| \\ &+ \left| \mathbb{E} \left[ (\hat{m}(x,\hat{\tau}) - \hat{m}(x,\tau)) \, \mathrm{I\!I}(\mathcal{E}_{n}) | X_{1}, \dots, X_{n} \right] \right| , \\ &\leq \left( \mathbb{E} \left[ (\hat{m}(x,\hat{\tau}) - \hat{m}(x,\tau))^{2} | X_{1}, \dots, X_{n} \right] \right)^{1/2} \mathbb{P}^{1/2} \left[ \mathcal{E}_{n}^{c} | X_{1}, \dots, X_{n} \right] \\ &+ \left( \mathbb{E} \left[ (\hat{m}(x,\hat{\tau}) - \hat{m}(x,\tau))^{2} \, \mathrm{I\!I}(\mathcal{E}_{n}) | X_{1}, \dots, X_{n} \right] \right)^{1/2} . \end{split}$$

Ainsi

$$\begin{split} &\int \left(\mathbb{E}\left[\hat{m}(x,\hat{\tau}) - \hat{m}(x,\tau)|X_{1},\ldots,X_{n}\right]\right)^{2}dx \\ &\leq \int \mathbb{E}\left[\left(\hat{m}(x,\hat{\tau}) - \hat{m}(x,\tau)\right)^{2}|X_{1},\ldots,X_{n}\right]\mathbb{P}\left[\mathcal{E}_{n}^{c}|X_{1},\ldots,X_{n}\right]dx \\ &+ \int \mathbb{E}\left[\left(\hat{m}(x,\hat{\tau}) - \hat{m}(x,\tau)\right)^{2}\Pi(\mathcal{E}_{n})|X_{1},\ldots,X_{n}\right]dx , \\ &\leq \mathcal{O}_{P}\left(h_{n} + \frac{1}{nh_{n}}\right)\mathbb{P}\left[\mathcal{E}_{n}^{c}|X_{1},\ldots,X_{n}\right] \\ &+ \sum_{i=1}^{p}\int_{\tau_{i}-h_{n}-g_{n}}^{\tau_{i}+h_{n}+g_{n}}\mathbb{E}\left[\left(\hat{m}(x,\hat{\tau}) - \hat{m}(x,\tau)\right)^{2}\Pi(\mathcal{E}_{n})|X_{1},\ldots,X_{n}\right]dx , \\ &= \mathcal{O}_{P}\left(h_{n} + \frac{1}{nh_{n}}\right)\mathbb{P}\left[\mathcal{E}_{n}^{c}|X_{1},\ldots,X_{n}\right] + \sum_{i=1}^{p}\mathcal{O}_{P}\left(\frac{g_{n}^{2}}{h_{n}} + g_{n}\right). \end{split}$$

Il en résulte que sous les hypothèses :

i) 
$$g_n = o\left(h_n^4 + \frac{1}{nh_n}\right),$$
  
ii)  $\mathbb{P}\left[\mathcal{E}_n^c \mid X_1, \dots, X_n\right] = o_P(h_n^3),$ 

$$\int \left( \mathbb{E} \left[ \hat{m}(x,\hat{\tau}) - \hat{m}(x,\tau) | X_1, \dots, X_n \right] \right)^2 dx = o_p \left( h_n^4 + \frac{1}{nh_n} \right).$$

Nous en déduisons que le biais intégré peut s'écrire comme suit

$$\int \left( \mathbb{E} \left[ \hat{m}(x,\hat{\tau}) - m(x) | X_1, \dots, X_n \right] \right)^2 dx = \int \left( \mathbb{E} \left[ \hat{m}(x,\tau) - m(x) | X_1, \dots, X_n \right] \right)^2 dx + \int \left( \mathbb{E} \left[ \hat{m}(x,\hat{\tau}) - m(x,\tau) | X_1, \dots, X_n \right] \right)^2 dx + 2 \int \mathbb{E} \left[ \left( \hat{m}(x,\hat{\tau}) - \hat{m}(x,\tau) \right) \left( \hat{m}(x,\tau) - m(x) \right) | X_1, \dots, X_n \right] dx = \frac{B_K^2 h_n^4}{4} \left( \int_0^1 \left[ m^{(2)}(x) \right]^2 dx \right) + o_p \left( h_n^4 + \frac{1}{nh_n} \right).$$

Il reste maintenant à prouver que la variance conditionnelle n'est pas modifiée lorsqu'on remplace p et  $\tau = (\tau_1, \ldots, \tau_p)$  par leurs estimateurs respectifs  $\hat{p}$  et  $\hat{\tau} = (\hat{\tau}_1, \ldots, \hat{\tau}_p)$ .

# 4.5 Choix des estimateurs du nombre et des localisations des ruptures

Dans le chapitre précédent nous avons présenté un algorithme pour estimer le nombre de ruptures d'une fonction de régression. D'après le corollaire 3.2, il est possible de construire un estimateur de ce nombre de ruptures tel que

$$1 - \operatorname{I\!P}\left[\forall 1 \le i \le p, \ \hat{\tau}_i \in ]\tau_i - g_n, \tau_i + g_n\left[ \text{ et } \hat{p} = p \right] = o\left(g_n^4\right),$$
où  $g_n = n^{-\delta}$  avec  $0 < \delta < 1$ .

Or

$$\mathbb{E}\left(\mathbb{P}\left[\mathcal{E}_{n}^{c} \mid X_{1}, \dots, X_{n}\right]\right) = 1 - \mathbb{P}\left[\forall 1 \leq i \leq p, \ \hat{\tau}_{i} \in ]\tau_{i} - g_{n}, \tau_{i} + g_{n}\left[ \text{ et } \hat{p} = p\right], \\ = o\left(g_{n}^{4}\right).$$

Par conséquent

$$\mathbb{P}\left[\mathcal{E}_n^c \mid X_1, \ldots, X_n\right] = o_P(g_n^4).$$

Maintenant si on choisit  $\delta > 4/5$ , alors pour toute suite  $(h_n)_{n \ge 1}$  qui vérifie  $h_n \to 0$  et  $nh_n \to \infty$ , on a:
i)  $g_n = o\left(h_n^4 + \frac{1}{nh_n}\right),$ ii)  $\mathbb{P}\left[\mathcal{E}_n^c \mid X_1, \dots, X_n\right] = o_P(h_n^3).$ 

## 4.6 Simulations

Nous analysons les effets de la segmentation pour plusieurs modèles. Les facteurs déterminants sont le choix de la fenêtre de lissage et le rapport signal sur bruit. Pour nos comparaisons, nous avons retenu trois modèles M1, M2 et M3:

M1 
$$m(x) = 4x^2 + \prod_{[0.5,1]}, n = 250 \text{ et } \sigma^2 = 0.1.$$
  
M2  $m(x) = 4x^2 + \prod_{[0.5,1]}, n = 250 \text{ et } \sigma^2 = 0.5.$   
M3  $m(x) = 4\sin(5x) + 3x + \prod_{[0.7,1]}, n = 1000 \text{ et } \sigma^2 = 1.$ 

Tout au long de ces simulations, le plan d'expérience est supposé aléatoire de loi uniforme et le bruit est gaussien.

Notons:

• 
$$\underline{m}(x) = \mathbb{E}(\hat{m}(x,t)) - 2\sqrt{Var(\hat{m}(x,t))},$$

• 
$$\overline{m}(x) = \mathbb{E}(\hat{m}(x,t)) + 2\sqrt{Var(\hat{m}(x,t))}$$

#### Commentaires.

- Nous avons observé que l'augmentation du nombre d'observations s'accompagne d'un rétrécissement de la "bande de confiance" de la fonction de régression m(·) ([m(·), m(·)]). Pour ne pas trop rallonger ce paragraphe, nous ne présentons pas ici les simulations concernant le comportement en fonction du nombre d'observations.
- 2) Dans tous les modèles analysés, nous remarquons que l'écart entre m(x) et  $\hat{m}(x, \hat{\tau})$  s'accentue au voisinage de la rupture. Ceci s'explique par :
  - L'augmentation du biais au voisinage du point de discontinuité (voir les figures 4.12, 4.16 et 4.20). Ceci est cohérent avec les théorèmes 4.1 et 4.2 qui montrent qu'au voisinage de τ le biais se comporte en O(1).
  - L'augmentation de la variance autour de  $\tau$ . Ceci est aussi cohérent avec les résultats énoncés dans les théorèmes 4.1 et 4.2. Ces théorèmes montrent que même si l'ordre

de grandeur de la variance  $(\mathcal{O}(1/nh_n))$  reste le même au voisinage de  $\tau$ , les constantes  $V_K^2(-|z|)$  sont plus grandes autour de la rupture  $(z \to 0)$  que lorsqu'en s'éloigne de la rupture, |z| > 1 (cf la figure 4.4).

3) Lorsque la segmentation des observations est effectuée en utilisant la vraie localisation τ, nous avons aussi constaté une détérioration de la qualité de l'estimation de m(x) au voisinage de τ. Elle est dûe essentiellement à l'augmentation de la variance autour du point de discontinuité pour la raison évoquée dans 2). En ce qui concerne le biais au voisinage de la rupture, nous remarquons une diminution par rapport à celui créé loin de τ. En effet, (cf le théorème 4.2 pour τ<sub>n</sub> = τ), lorsqu'on utilise la vraie localisation de la rupture pour segmenter les observations, le biais au voisinage de τ reste de l'ordre de O(h<sub>n</sub><sup>2</sup>). Les constantes asymptotiques (B<sub>K</sub>(-|z|), cf la figure 4.3) sont cependant plus petites pour les points qui sont proches de τ(z → 0) que pour ceux qui en sont loins (|z| > 1). Notons que Fan & Gijbels [FG92] ont observé ce phénomène dans le cadre de l'estimation d'une fonction de régression lisse par polynômes locaux (augmentation de la variance et diminution du biais sur les bords).



**Figure 4.10 :** Effet de la segmentation en  $t = \tau$  : Estimation de  $\underline{m}(x)$  et  $\overline{m}(x)$  à partir de 200 échantillons simulés selon le modèle M1. La courbe en "-." représente une estimation de  $\underline{m}(x)$ , celle en "--" représente une estimation de  $\overline{m}(x)$ . La courbe en "-" représente la vraie fonction de régression m(x). Dans le graphe en haut à gauche on a aussi représenté avec des "." un échantillon des observations. Les trois figures en bas sont des agrandissements du voisinage de la discontinuité.



**Figure 4.11 :** Effet de la segmentation en  $t = \hat{\tau}$  : Estimation de  $\underline{m}(x)$  et  $\overline{m}(x)$  à partir de 200 échantillons simulés selon le modèle M1. La courbe en "-." représente une estimation de  $\underline{m}(x)$ , celle en "--" représente une estimation de  $\overline{m}(x)$ . La courbe en "-" représente la vraie fonction de régression m(x). Dans le graphe en haut à gauche on a aussi représenté avec des "." un échantillon des observations. Les trois figures en bas sont des agrandissements du voisinage de la discontinuité.



**Figure 4.12 :** Estimations du biais local et de l'écart-type local (de haut en bas) à partir de 200 échantillons simulés selon le modèle M1. On montre à gauche ces différentes estimations pour  $\hat{m}(\cdot, \tau)$  et à droite celles de  $\hat{m}(\cdot, \hat{\tau})$ .



**Figure 4.13 :** Estimations du biais au carré intégré, de la variance intégrée et de la MISE (de haut en bas) à partir de 200 échantillons simulés selon le modèle M1. On montre à gauche ces différentes estimations pour  $\hat{m}(\cdot, \tau)$  et à droite celles de  $\hat{m}(\cdot, \hat{\tau})$ .



**Figure 4.14 :** Effet de la segmentation en  $t = \tau$  : Estimation de  $\underline{m}(x)$  et  $\overline{m}(x)$  à partir de 200 échantillons simulés selon le modèle M2. La courbe en "-." représente une estimation de  $\underline{m}(x)$ , celle en "--" représente une estimation de  $\overline{m}(x)$ . La courbe en "-" représente la vraie fonction de régression m(x). Dans le graphe en haut à gauche on a aussi représenté avec des "." un échantillon des observations. Les trois figures en bas sont des agrandissements du voisinage de la discontinuité.



**Figure 4.15 :** Effet de la segmentation en  $t = \hat{\tau}$  : Estimation de  $\underline{m}(x)$  et  $\overline{m}(x)$  à partir de 200 échantillons simulés selon le modèle M2. La courbe en "-." représente une estimation de  $\underline{m}(x)$ , celle en "--" représente une estimation de  $\overline{m}(x)$ . La courbe en "-" représente la vraie fonction de régression m(x). Dans le graphe en haut à gauche on a aussi représenté avec de "." un échantillon des observations. Les trois figures en bas sont des agrandissements du voisinage de la discontinuité.



**Figure 4.16 :** Estimations du biais local et de l'écart-type local (de haut en bas) à partir de 200 échantillons simulés selon le modèle M2. On montre à gauche ces différentes estimations pour  $\hat{m}(\cdot, \tau)$  et à droite celles de  $\hat{m}(\cdot, \hat{\tau})$ .



**Figure 4.17 :** Estimations du biais au carré intégré, de la variance intégrée et de la MISE (de haut en bas) à partir de 200 échantillons simulés selon le modèle M2. On montre à gauche ces différentes estimations pour  $\hat{m}(\cdot, \tau)$  et à droite celles de  $\hat{m}(\cdot, \hat{\tau})$ .



**Figure 4.18 :** Effet de la segmentation en  $t = \tau$  : Estimation de  $\underline{m}(x)$  et  $\overline{m}(x)$  à partir de 200 échantillons simulés selon le modèle M3. La courbe en "-." représente une estimation de  $\underline{m}(x)$ , celle en "--" représente une estimation de  $\overline{m}(x)$ . La courbe en "-" représente la vraie fonction de régression m(x). Dans le graphe en haut à gauche on a aussi représenté avec des "." un échantillon des observations. Les trois figures en bas sont des agrandissements du voisinage de la discontinuité.



**Figure 4.19 :** Effet de la segmentation en  $t = \hat{\tau}$  : Estimation de  $\underline{m}(x)$  et  $\overline{m}(x)$  à partir de 200 échantillons simulés selon le modèle M3. La courbe en "-." représente une estimation de  $\underline{m}(x)$ , celle en "--" représente une estimation de  $\overline{m}(x)$ . La courbe en "-" représente la vraie fonction de régression m(x). Dans le graphe en haut à gauche on a aussi représenté avec des "." un échantillon des observations. Les figures trois en bas sont des agrandissements du voisinage de la discontinuité.



**Figure 4.20 :** Estimations du biais local et de l'écart-type local (de haut en bas) à partir de 200 échantillons simulés selon le modèle M3. On montre à gauche ces différentes estimations pour  $\hat{m}(\cdot, \tau)$  et à droite celles de  $\hat{m}(\cdot, \hat{\tau})$ .



**Figure 4.21 :** Estimations du biais au carré intégré, de la variance intégrée et de la MISE (de haut en bas) à partir de 200 échantillons simulés selon le modèle M3. On montre à gauche ces différentes estimations pour  $\hat{m}(\cdot, \tau)$  et à droite celles de  $\hat{m}(\cdot, \hat{\tau})$ .

## 4.7 Données réelles : Exemple du Nil.

Nous présentons dans cette section une application de l'analyse de ruptures par lissage linéaire local aux données du Nil (cf chapitre Introduction). Pour l'analyse, nous avons préféré supposer que le nombre de ruptures est au plus 1. Tout d'abord nous avons estimé, dans le cas où nous supposons qu'il y a une seule rupture, les paramètres de rupture ( $\tau$  et  $\gamma(\tau)$ ) pour différentes tailles de la fenêtre de lissage. Ensuite nous avons vérifié la pertinence de chacune des ruptures détectées à l'aide du test strictement local. Enfin et suite aux résultats des tests effectués, nous avons gardé comme paramètre de localisation de la rupture  $\hat{\tau} = 1898$  et nous avons estimé le débit annuel du Nil en utilisant comme point de segmentation  $t = \hat{\tau} = 1898$ .

Notons que le meilleur cadre pour analyser les données du Nil est celui des séries chronologiques. Néanmoins nous faisons comme si les conditions étaient remplies pour pouvoir appliquer notre approche d'analyse de ruptures. Les résultats obtenus sont cohérents avec ceux obtenus par Müller [MÜL92] à l'aide d'une méthode de noyau, ou ceux de Cobb [COB78] dans un cadre paramétrique.

	<i>K</i> 1			K2
$h_n$	$\hat{ au}$	$\hat{\gamma}(\hat{ au})$	$\hat{ au}$	$\hat{\gamma}(\hat{\tau})$
10	1915	269.1861	1897	-473.2106
12	1898	-283.3347	1898	-534.6239
14	1898	-316.9070	1898	-543.0585
16	1898	-307.8848	1898	-445.1601
18	1898	-296.1336	1899	-375.6468
20	1898	-293.0075	1899	-385.7173

#### 1) Estimation des paramètres de rupture.

**Table 4.1 :** Estimations de  $\hat{\tau}$  et  $\hat{\gamma}(\hat{\tau})$ . Les noyaux utilisés sont K1 et K2.

#### 2) Tests de rupture.

**Table 4.2 :** Test strictement local d'absence de rupture en  $\tau = 1898, 1902, 1915$  au seuil  $\alpha = 5\%$ . Dans chaque cellule figure la valeur de la statistique et la décision prise. Le noyau utilisé est K1.

	$\tau = 1898$	$\tau = 1902$	$\tau = 1915$
$h_n$	(TSL)	(TSL)	(TSL)
10	HO (1.3888)	HO (0.2806)	HO $(1.7250)$
12	H1 $(1.9890)$	HO $(0.3662)$	HO $(1.6145)$
14	H1 $(2.4029)$	HO $(0.2790.)$	HO $(1.5485)$
16	H1 $(2.4957)$	HO $(0.1807)$	HO $(1.3967)$
18	H1 $(2.5461)$	HO $(0.5442)$	HO $(1.3381)$
20	H1 $(2.6555)$	HO (0.7921)	HO (1.3868)

**Table 4.3 :** Test strictement local d'absence de rupture en  $\tau = 1898, 1902, 1915$  au seuil  $\alpha = 5\%$ . Dans chaque cellule figure la valeur de la statistique et la décision prise. Le noyau utilisé est K2.

	$\tau = 1898$	$\tau = 1902$	$\tau = 1915$
$h_n$	(TSL)	(TSL)	(TSL)
10	H1 (3.2319)	HO $(0.1869)$	H1 $(2.5969)$
12	H1 $(4.4730)$	HO $(0.4891)$	HO $(1.6405)$
14	H1 $(4.9076)$	HO (0.0914.)	HO $(1.1916)$
16	H1 $(4.3007)$	HO $(1.4642)$	HO $(0.6091)$
18	H1 (3.8276)	H1 (2.3802)	HO (0.4756)
20	H1 (3.7270)	H1 (2.7846)	HO (0.7208)

Rappelons que la statistique utilisée pour construire les différents tests strictement locaux

d'absence de ruptures dans le débit annuel du Nil sont définis par :

$$TSL = \sqrt{nh_n} \frac{\hat{\gamma}_n(\tau)}{\sqrt{2}V_+\hat{\sigma}}.$$

Sous l'hypothèse HO,  $TSL \xrightarrow{\mathcal{L}} \mathcal{N}(0, 1)$ . Pour  $\hat{\sigma}$  nous avons utilisé l'estimation fournie par Cobb [COB78] ( $\hat{\sigma} = 125$ ).

## 3) Estimation du débit annuel.



Figure 4.22 : Estimation du débit annuel du Nil entre 1870 et 1970.

# Bibliographie

- [AG93] A. Antoniadis and G. Grégoire. Nonparametric estimation in change-point hazard rate models for censored data: a counting process approach. Journal of Nonparametric Statistics, 3:135-154, 1993.
- [AG96] A. Antoniadis and I. Gijbels. Inference in change-point problems via wavelet decompositions. Technical report, Institut de Statistique, Université Catholique de Louvain, 1996.
- [BIL68] P. Billingsley. Convergence of probability measures. Wiley, New York 1968.
- [BN93] M. Basseville and I. V. Nikiforov. *Detection of abrupt changes: theory and application*. Prentice-Hall information and system sciences series. Prentice-Hall, 1993.
- [BR73] P. J. Bickel and M. Rosenblatt. On some global measures of the deviation of density function estimators. *Ann. Math. Statist*, 1:1071–1095, 1973.
- [CH91] D. B. H. Cline and J. D. Hart. Kernel estimation of densities with discontinuities or discontinuous derivatives. *Statistics*, 22:69–84, 1991.
- [CHE64] H. Chernoff. Estimation of the mode. Ann. Inst. Statist. Math, 16:31–41, 1964.
- [CIN 75] Erhan Cinlar. Introduction to stochastic processes. Prentice-Hall, Inc., 1975.
- [CLE79] W. S. Cleveland. Robust locally weighted regression and smoothing scatterplots.
   J. Amer. Statist. Assoc., 74(368):829-836, 1979.
- [COB78] G. W. Cobb. The problem of the nile: Conditional solution to a changepoint problem. *Biometrika*, 65:243-251, 1978.
- [DBE75] J. Durbin R. L. Brown and J. M. Evans. Techniques for testing the constancy of regression relationships over time. J. Royal Statist. Soc. B, 37:149–163, 1975.

- [DÜM91] L. Dümbgen. The asymptotic behavior of some nonparametric change-point estimators. *The Annals of Statistics*, 19(3):1471–1495, 1991.
- [ECS95] R. L. Eubank D. B. H. Cline and P. L. Speckman. Nonparametric estimation of regression curves with discontinuous derivatives. Technical report, Texas A&M University and University of Missouri-Columbia, 1995.
- [EDD82] W. F. Eddy. The asymptotic distributions of the mode. Ann. Statist., 59:279–290, 1982.
- [EED85] C. V. Eeden. Mean integrated squared error of kernel estimators when the density and its derivatives are not necessarily continuous. Ann. Inst. Statist. Math., 37(1):461-472, 1985.
- [ES94] L. Eubank and P. L. Speckman. Nonparametric estimation of functions with jump discontinuities. In H.-G. Müller E. Carlstein and D. Siegmund, editors, *Change-Point Problems*. Institute of Mathematical Statistics, Hayward, California, 1994.
- [FAN92] J. Fan. Design-adaptive nonparametric regression. J. Amer. Statist. Assoc., 87:998–1004, 1992.
- [FAN93] J. Fan. Local linear regression smoothers and their minimax efficiencies. Ann. Statist., 21:196-216, 1993.
- [FCM97] J. Fan M. Y. Cheng and J. S. Marron. On automatic boundary corrections. The Annals of Statistics, 25(4):1691–1708, 1997.
- [FEL66] William Feller. An introduction to probability theory and its applications, volume II.John Wily & Sons, Inc., 1966.
- [FG92] J. Fan and I. Gijbels. Variable bandwidth and local linear regression smoothers. Ann. Statist., 20:2008–2036, 1992.
- [FG95] J. Fan and I. Gijbels. Data-driven bandwidth selection in local polynomial fitting: variable bandwidth and spatial adaptation. J. R. Statist. Soc. B, 57(2):371-394, 1995.
- [FG96] J. Fan and I. Gijbels. Local Polynomial Modelling and Its Applications. Chapman and Hall, London, 1996.

- [FM94] J. Fan and J. S. Marron. Fast implementations of nonparametric curve estimators. Journal of Computational and Graphical Statistics, 3(1):35-56, 1994.
- [GIR 90] D. Girard. From template matching to optimal approximation by piecewise smooth curves. In SPIE Proceedings on "Curves and surfaces in computer vision and graphics, Santa Clara, CA, Feb. 11-16 1990.
- [GK68] B. V. Gnedenko and A. N. Kolmogorov. Limit distributions for sum of independent random variables. Addison-Wesley series in statistics. Addison-Wesley publishing company, 1968.
- [HÄ90] W. Härdle. Applied nonparametric regression. Cambridge University Press, 1990.
- [HGK95] P. Hall I. Gijbels and A. Kneip. Interval and band estimation for curves with jumps. Technical report, Institut de Statistique, Université Catholique de Louvain, 1995.
- [HGK96] P. Hall I. Gijbels and A. Kneip. On the estimation of jump points in smooth curves. Technical report, Institut de Statistique, Université Catholique de Louvain, 1996.
- [HL84] W. Härdle and S. Luckhaus. Uniform consistency of a class of regression function estimators. *The Annals of Statistics*, 12(2):612–623, 1984.
- [HL93] T. Hastie and C. Loader. Local regression: automatic kernel carpentry. Statistical Science, 8(2):120-143, 1993.
- [HY96] J. D. Hart and S. Yi. One-sided cross-validation. Technical report, Department of Statistics, Texas A & M University, 1996.
- [JOH82] G. J. Johnston. Probabilities of maximal deviations for nonparametric regression function estimates. *Journal of Multivariate Analysis*, 12:402–414, 1982.
- [KBP95] I. Koch M. Bunt and A. Pope. Kernel-based nonparametric regression for discontinuous functions. Technical report, Centre for Mathematics and its Applications, School of Mathematical Sciences, Australian National University, Canberra, ACT 0200, Australia, 1995.
- [KE96] A. Kneip and J. Engel. A remedy for kernel estimation under random design. Statistics, 28:201-225, 1996.
- [KM88] P. R. Krishnaiah and B. Q. Miao. Review about estimation of change points. Handbook of Statistics, 7:375-402, 1988.

- [KOR86] A. P. Korostelev. On minimax estimation of a discontinuous signal. *Probability* Theory and its Applications, pages 727–730, 1986.
- [LB87] J. P. Lecoutre D. Bosq. *Théorie de l'estimation fonctionnelle*. Economica, 1987.
- [LEE91] D. Lee. Detection, classification, and measurement of discontinuities. SIAM, J. Sc. Stat. Comput., 12(2):311-341, 1991.
- [LEJ85] M. Lejeune. Estimation non-paramétrique par noyaux: Régression polynomiale mobile. *Revue de Statistiques Appliquées*, 33:43-64, 1985.
- [LLR83] G. Lindgren M. R. Leadbetter and H. Rootzén. Extremes and related properties of random sequences and processes. Springer, Series in Statistics, New York, 1983.
- [LOA96] C. R. Loader. Change point estimation using nonparametric regression. *The Annals* of Statistics, 24:1667–1678, 1996.
- [LU86] P. J. Laurent and F. Utreras. Optimal smoothing of noisy broken data with spline functions. J. Approximation Theory and its Applications, 2:71–94, 1986.
- [MCS94] H. G. Müller E. Carlstein and D. Siegmund. Change-point problems, volume 23. Institute of Mathematical Statistics Lecture Note Series, 1994.
- [MO86] J. A. McDonald and A. B. Owen. Smoothing with split linear fits. *Technometrics*, 28(3):195-208, 1986.
- [MS82] Y. P. Mack and B. W. Silverman. Weak and strong uniform consistency of kernel regression estimates. Z. Wahrscheinlichkeitstheorie verw. Gebiete, 61:405-415, 1982.
- [MS87] H. G. Müller and U. Stadtmüller. Estimation of heteroscedasticity in regression analysis. *Ann. Statist.*, 15:610–625, 1987.
- [MS97] H. G. Müller and K. S. Song. Two-stage change-point estimators in smooth regression models. *Statistics & Probability Letters*, 34:323-335, 1997.
- [MÜL87] H. G. Müller. Weighted local regression and kernel methods for nonparametric curve fitting. Journal of the American Statistical Association, 82(397):231-238, 1987.
- [MÜL91] H. G Müller. Kernels estimators of zeros and locations and size of extrema of regression functions. *Scand. J. Statist.*, 12(3):1471-1495, 1991.

- [MÜL92] H. G. Müller. Change-points in nonparametric regression analysis. Ann. Statist., 20:737-761, 1992.
- [OUD95] C. G. M. Oudshoorn. Minimax estimation of a regression function with jumps: attaining the optimal constant. Technical report, Mathematical Institute, University of Utrecht, 1995.
- [PAG 57] E. S. Page. Estimating the point of change in a continuous process. *Biometrika*, 44:248-252, 1957.
- [PC72] M. B. Priestley and M. T. Chao. Non-parametric function fitting. J. Roy. Statist. Soc. B, 34:385–392, 1972.
- [PGC93] D. Picard and V. Genon-Catalot. Eléments de statistique asymptotique. SMAI. Springer-Verlag, 1993.
- [PK84] V. I. Piterbarg V. D. Konakov. On the convergence rate of maximal deviation distribution for kernel regression estimates. J. Multivariate Anal., 15:279-294, 1984.
- [POL84] D. Pollard. Convergence of stochastic processes. Springer-Verlag, series in statistics, New York Berlin Heidelberg Tokoyo, 1984.
- [QIU94] P. Qiu. Estimation of the number of jumps of the jump regression functions. *Statistics Theory and method*, 23:2141–2155, 1994.
- [QY98] P. Qiu and B. Yandell. A local polynomial jump detection algorithm in nonparametric regression. *To appear in Technometrics*, 1998.
- [RAI96] M. Raimondo. Modèles en rupture, situations non ergodique et utilisation de méthode d'ondelette. PhD thesis, Université de Paris VII, 1996.
- [RW94] D. Ruppert and M. P. Wand. Multivariate locally weighted least squares regression. Ann. Statist., 22:1346–1370, 1994.
- [SG96] B. Seifert and T. Gasser. Finite-sample variance of local polynomials: analysis and solutions. Journal of the American Statistical Association, 91(433):267–275, 1996.
- [SHA80] S. A. Shaban. Change point problem and two-phase regression: an annotated bibliography. *International Statistical Review*, 48:83–93, 1980.
- [SH185] J. H. Shiau. Smoothing spline estimation of functions with discontinuities. PhD thesis, University of Wisconsin, Madison, Wisconsin, 1985.

[SIL 78]	B. W. Silverman. Weak and strong uniform consistency of the kernel estimate of a density and its derivatives. <i>The Annals of Statistics</i> , 6(1):177–184, 1978.		
[Spe95]	P. L. Speckman. Detection of change-points in nonparametric regression. Technical report, Department of Statistics, University of Missouri-Columbia, 1995.		
[Spo97]	V. G. Spokoiny. Estimation of a function with discontinuities via local polynomial fit with an adaptative window choice. Technical report, Weierstrass Institute for Applied Analysis and Stochastics, Mohrenstr. 39, 10117 Berlin, 1997.		
[Sto77]	C. J. Stone. Consistent nonparametric regression. The Annals of Statistics, 5(4):595-645, 1977.		
[Sto82]	C. J. Stone. Optimal global rates of convergence for nonparametric regression. Annals of Statistics, 10:1040-1053, 1982.		
[TK93]	A. B. Tsybakov and A. P. Korostelev. <i>Minimax Theory of Image Reconstruction</i> . Springer-Verlag, Lecture Notes in Statistics, New York Berlin Heidelberg Tokoyo, 1993.		
[Tsy86]	A. B. Tsybakov. Robust reconstruction of function by local-approximation method. Problems of Information Transmission, 2:69–84, 1986.		
[Tsy90]	A. B. Tsybakov. Locally-polynomial algorithms of passive stochastic approxima- tion. <i>Problems of Control and Information Theory</i> , 3:181–195, 1990.		
[Wan95]	Y. Wang. Jump and sharp cusp detection by wavelets. <i>Biometrika</i> , 82:385–397, 1995.		
[WC93A]	J. S. Wu and C. K. Chu. Kernel-type estimators of jump points and values of a regression function. <i>The Annals of Statistics</i> , 21(3):1545-1566, 1993.		
[WC938]	J. S. Wu and C. K. Chu. Nonparametric function estimation and bandwidth selection for discontinuous regression functions. <i>Statistica Sinica</i> , 3:557–576, 1993.		
[WH170]	W. Whitt. Weak convergence of probability measures on the function space $C[0, \infty[$ . Ann. Math. Statist., 41:939–944, 1970.		
[WH180]	W. Whitt. Some useful functions for functional limit theorems. <i>Mathematics of Operations Research</i> , 5:67–84, 1980.		

- [WM90] J. L. Wang and H. G. Müller. Nonparametric analysis changes in hazard rates for censored survival data: An alternative to change-point models. *Biometrika*, 77:305-314, 1990.
- [YIN88] Y. Q. Yin. Detection of the number, locations and magnitudes of jumps. Commun. Statist. Stochastic Models, 4(3):445-455, 1988.
- [ZF96] J. Zhang and J. Fan. Minimax kernels for nonparametric estimation. Technical report, Department of Statistics, UNC-Chapel Hill, NC 27599-3260, 1996.

## Résumé

Nous nous intéressons dans cette thèse à l'estimation, dans un cadre non paramétrique, d'une fonction de régression présentant des discontinuités et, plus précisément aux problèmes de détection de ruptures, d'estimation des paramètres de rupture (nombre, localisations, amplitudes) et de segmentation de la fonction de régression (reconstitution de la fonction). La méthode utilisée est basée sur les propriétés du processus de saut estimé,  $\hat{\gamma}(t)$ , défini en tout t comme la différence entre un estimateur à droite et un estimateur à gauche, ces estimateurs étant obtenus régression linéaire locale.

Dans un premier temps, nous considérons la situation d'une seule discontinuité et étudions les propriétés de l'estimateur de l'amplitude de la discontinuité lorsque la localisation est connue. Nous donnons l'expression de l'erreur quadratique moyenne asymptotique et montrons la convergence et la normalité asymptotique de l'estimateur. Lorsque la localisation  $\tau$  n'est pas connue, nous construisons un estimateur de au à l'aide du processus de déviation locale associé à  $\hat{\gamma}(t)$  et montrons que cet estimateur converge avec une vitesse en  $n^{-1}$  ou arbitrairement proche de  $n^{-1}$  selon le noyau utilisé. Nous proposons ensuite trois tests d'existence d'une rupture : un test strictement local, un test local et un test global, tous trois définis en terme d'une statistique construite à l'aide du processus de saut estimé. Concernant le problème d'estimation du nombre de ruptures nous élaborons une procédure permettant à la fois d'estimer le nombre p de ruptures et les localisations  $\tau_1, \ldots, \tau_p$ . Nous montrons la convergence presque sûre de ces estimateurs et donnons aussi des résultats sur les vitesses de convergence. Enfin nous proposons une méthode de reconstitution d'une fonction de régression présentant des discontinuités basée sur la segmentation des observations. Nous montrons qu'en utilisant la procédure d'estimation du nombre de ruptures et des localisations développée auparavant, nous obtenons un estimateur de la fonction de régression qui a la même vitesse de convergence qu'en l'absence de ruptures. Des expérimentations numériques sont fournies pour chacun des problèmes étudiés de manière à mettre en évidence les propriétés des procédures étudiées et leur sensibilité aux divers paramètres.

Mots clés : régression, non-paramétrique, rupture, régression linéaire locale, estimation, détection, segmentation.

### Abstract

In this thesis, we consider the change-point problems in nonparametric regression model: detection of change-points, estimation of the change-point parameters and segmentation of the regression function are investigated. The proposed method is based on a comparison of left and right one-sided smoothers. These smoothers are obtained by local linear regression.

First, we propose an estimator of the jump size in the regression function when the location of the change-point is known. We derive the expression of its asymptotic MSE and its asymptotic normality. Next we propose an estimator of the location of the change-point and give its asymptotic distribution and corresponding rate of convergence. Afterwards, we are interested in testing for a change-point. The properties of the deviation process associated with the jump estimate process, allow to obtain the asymptotic distributions of some test statistics for a change-point. In the case where there are more than one change, we introduce a consistant algorithm to estimate the number of change-points together with their localizations. Finally, we propose a method for the estimation of a piecewise smooth regression fonction. We show that, provided discontinuities can be located with sufficient accuracy, our approach enjoys rates of convergence similar to those obtained in smooth case. For each of the proposed procedures, we provide numerical experiments that show evidence of the quality of the method and give some information about the dependence upon such parameters as the kernel and the bandwith.

**Keywords** : regression, nonparametric, change-point, local linear regression, estimation, detection, segmentation.